АКТУАЛЬНЫЕ ТЕХНОЛОГИИ
Обзор / Review
УДК 577.215.3:575.112

Вавиловский журнал генетики и селекции. 2019;23(5):508-518
DOI 10.18699/VJ19.520

Новейшие технологии высокопроизводительного секвенирования транскриптома отдельных клеток

Е.А. Водясова ✉, Э.С. Челебиева, О.Н. Кулешова

Институт биологии южных морей имени А.О. Ковалевского Российской академии наук, Москва, Россия
✉ e-mail: eavodiasova@gmail.com

Огромное количество полногеномных и транскриптомных данных, полученных с помощью современных технологий секвенирования нового поколения для целых организмов, не смогло дать ответы на многие вопросы в онкологии, иммунологии, физиологии, нейробиологии, зоологии и других областях науки и медицины. Так как основой всех одноклеточных и многоклеточных организмов является клетка, то необходимо изучение биологических процессов на ее уровне. Это понимание дало толчок развитию нового направления и появлению технологий, позволяющих работать с единичными клетками (технологии single-cell). Быстрое развитие не только приборной базы, но и различных усовершенствованных протоколов для работы с единичными клетками обусловлено актуальностью этих исследований во многих областях науки и медицины. Изучение особенностей различных этапов онтогенеза, определение закономерностей дифференциации клеток и последующего развития тканей, проведение геномного и транскриптомного анализов в различных областях медицины (особенно востребовано в иммунологии, онкологии), классификация типов и состояний клеток, закономерностей биохимических и физиологических процессов с применением технологий single-cell позволяют проводить комплексные исследования на новом уровне. Разработанные первые платформы для осуществления секвенирования транскриптомов отдельных клеток (scRNA-seq) проводили изоляцию не более ста клеток единовременно, что оказалось недостаточным в связи с выявленной высокой гетерогенностью клеток, обнаруженными минорными типами клеток, которые не детектировались по морфологическим признакам, и сложными регуляторными путями в организме. В настоящее время появились методики изоляции, захвата и секвенирования транскриптомов (scRNA-seq) десятков тысяч клеток единовременно. Однако новые технологии имеют определенные отличия как на этапе пробоподготовки, так и во время проведения биоинформатического анализа. В работе рассмотрены наиболее эффективные методы множественного параллельного scRNA-seq на примере современной платформы для изоляции и баркодирования клеток 10XGenomics, а также особенности проведения такого эксперимента, дальнейший биоинформатический анализ полученных данных, перспективы использования и области применения новых высокопроизводительных технологий.
Ключевые слова: scRNA-seq; транскриптомика; Chromium 10XGenomics; секвенирование; единичные клетки.

Для цитирования: Водясова Е.А., Челебиева Э.С., Кулешова О.Н. Новейшие технологии высокопроизводительного секвенирования транскриптома отдельных клеток. Вавиловский журнал генетики и селекции. 2019;23(5):508-518. DOI 10.18699/VJ19.520

The new technologies of high-throughput single-cell RNA sequencing

E.A. Vodiasova ✉, E.S. Chelebieva, O.N. Kuleshova

A.O. Kovalevsky Institute of Biology of the Southern Seas, RAS, Moscow, Russia
✉ e-mail: eavodiasova@gmail.com

A wealth of genome and transcriptome data obtained using new generation sequencing (NGS) technologies for whole organisms could not answer many questions in oncology, immunology, physiology, neurobiology, zoology and other fields of science and medicine. Since the cell is the basis for the living of all unicellular and multicellular organisms, it is necessary to study the biological processes at its level. This understanding gave impetus to the development of a new direction – the creation of technologies that allow working with individual cells (single-cell technology). The rapid development of not only instruments, but also various advanced protocols for working with single cells is due to the relevance of these studies in many fields of science and medicine. Studying the features of various stages of ontogenesis, identifying patterns of cell differentiation and subsequent tissue development, conducting genomic and transcriptome analyses in various areas of medicine (especially in demand in immunology and oncology), identifying cell types and states, patterns of biochemical and physiological processes using single cell technologies, allows the comprehensive research to be conducted at a new level. The first RNA-sequencing technologies of individual cell transcriptomes (scRNA-seq) captured no more than one hundred cells at a time, which was insufficient due to the detection of high cell heterogeneity, existence of the minor cell types (which were not detected by morphology) and complex regulatory pathways. The unique techniques for isolating, capturing

© Водясова Е.А., Челебиева Э.С., Кулешова О.Н., 2019

and sequencing transcripts of tens of thousands of cells at a time are evolving now. However, new technologies have certain differences both at the sample preparation stage and during the bioinformatics analysis. In the paper we consider the most effective methods of multiple parallel scRNA-seq using the example of 10XGenomics, as well as the specifics of such an experiment, further bioinformatics analysis of the data, future outlook and applications of new high-performance technologies.
Key words: scRNA-seq; transcriptomics; Chromium 10XGenomics; sequencing; single cell.

For citation: Vodiasova E.A., Chelebieva E.S., Kuleshova O.N. The new technologies of high-throughput single-cell RNA sequencing. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(5):508-518. DOI 10.18699/VJ19.520 (in Russian)

Введение

Развитие методов секвенирования нового поколения в совокупности с технологиями множественного параллельного захвата и анализа единичных клеток подняло на более высокий уровень исследования во многих областях биологии, биотехнологии и медицины (Junker, Oudenaarden, 2014). Краткая история развития методов изучения в масштабе одной клетки показана на рис. 1. Основные направления, где внедряются инновационные методы, – нейробиология, иммунология, эмбрионология и онкология. Первые эксперименты по анализу транскриптома нейронов с использованием микрочипов проведены в 2003 г. Изоляцию нейронов у крыс осуществляли с помощью лазерной микродиссекции, данные выявили неожиданно высокую гетерогенность клеток (Kamme et al., 2003). С 2006 г. активно развивается секвенирование транскриптома нейронов (Moroz et al., 2006; Tang et al., 2009), и в то же время появляются методы, основанные на микрофлюидике, позволяющие анализировать до 100 клеток одновременно (Marcus et al., 2006). Параллельно внедряются методы qPCR, адаптированные для работы с единичными клетками (Subkhankulova et al., 2008). Впервые проведенное полногеномное секвенирование опухолевых клеток выявило многочисленную популяцию псевдодиплоидных клеток, не дающих метастазов (Navin et al., 2011). В последние пять лет развивается новое направление «эпигеномика», связанное с исследованием различных модификаций РНК и ДНК с помощью секвенирования. Современные технологии работы с единичными клетками позволяют характеризовать метилирование ДНК, доступность хроматина, пространственную укладку хроматина, сайты специфических модификаций гистонов, сайты связывания с белками (Goldberg et al., 2007; Nagano et al., 2013; Rotem et al., 2015).

Необходимость единовременного анализа огромного числа клеток привела к развитию технологий, направленных на повышение производительности секвенирования и разработки платформ для захвата множества клеток. С 2003 по 2016 г. количество одновременно обрабатываемых клеток возросло от нескольких клеток до сотен тысяч.

Анализ транскриптома, протеома и эпигенома на уровне клетки позволяет понять особенности формирования, развития, организации и взаимодействия различных клеток и тканей; исследовать проблемы стресса и адаптаций (Frieda et al., 2017); определить типы и состояния клеток; установить закономерности их дифференциации в процессе онтогенеза (Nowogrodzki, 2017; Mi et al., 2018); изучить различные заболевания в иммунологии, онкологии, нейро-

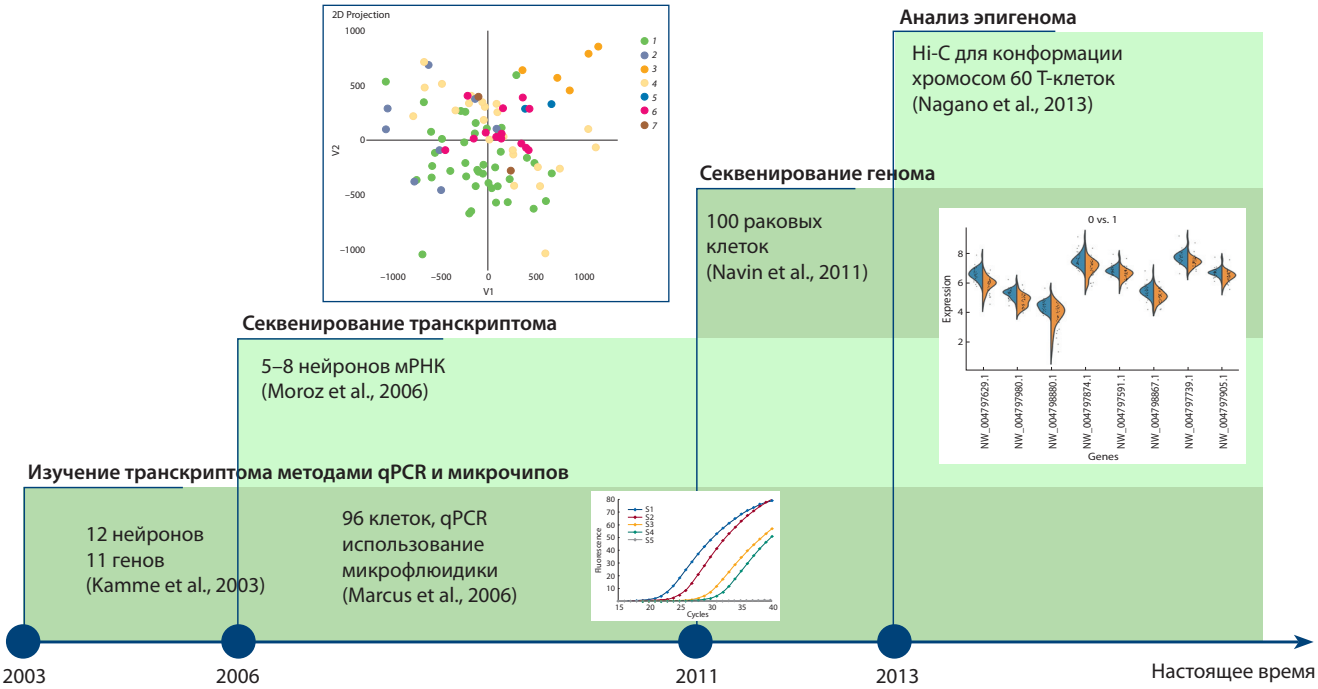

Рис. 1. История развития технологий для изучения единичных клеток.

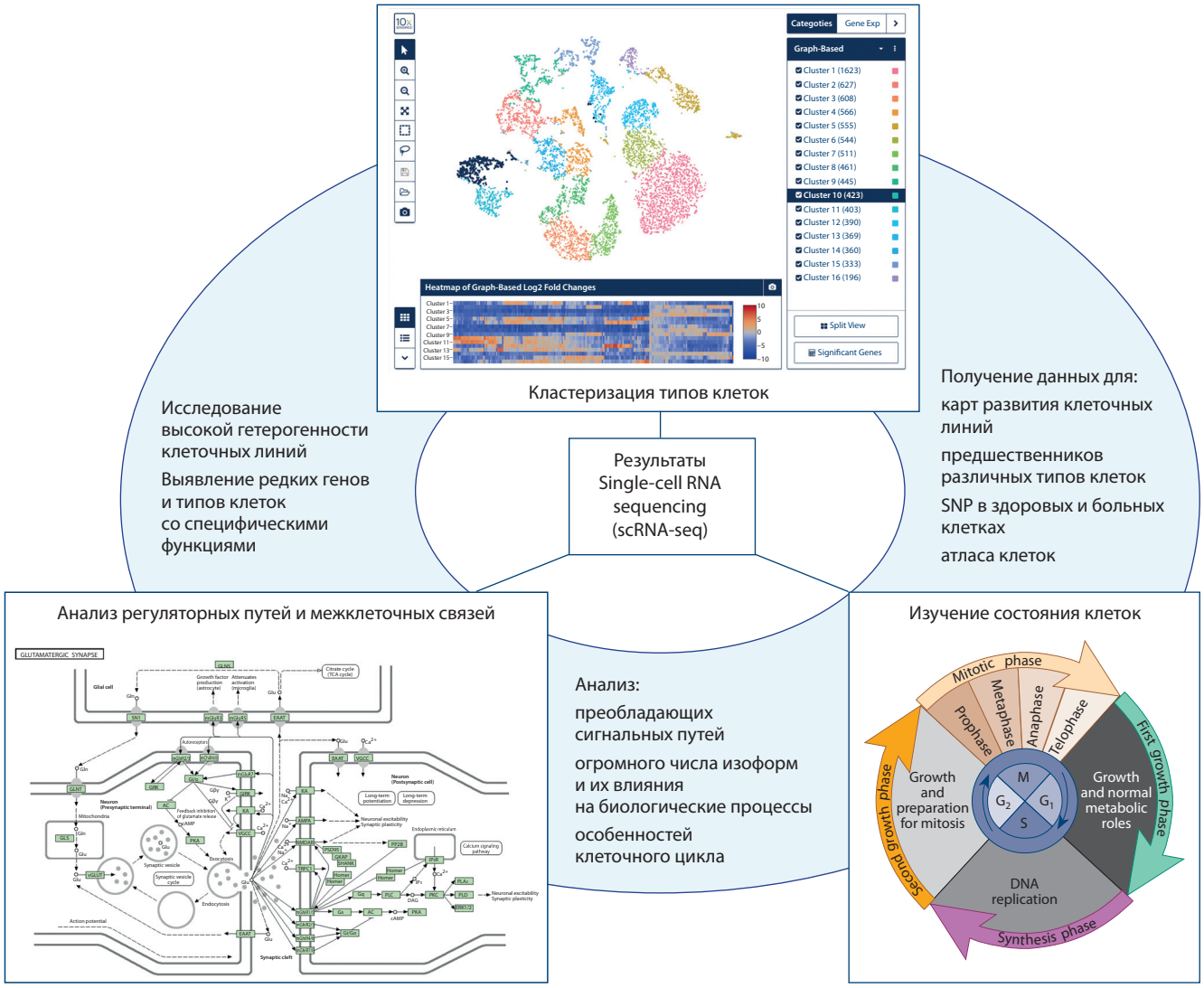

Рис. 2. Основные направления анализа данных scRNA-seq.

биологии и других областях медицины (Leung et al., 2017); ответить на вопросы эволюции, видообразования и формирования глобального биоразнообразия (Moroz, 2018).

Настоящая работа посвящена обзору наиболее эффективных и востребованных платформ для проведения высокопроизводительного секвенирования транскриптомов единичных клеток (single-cell RNA sequencing, scRNA-seq). Описаны этапы получения и анализа транскриптома единичных клеток, рассмотрены основные протоколы для таких исследований и перспективы использования технологий scRNA-seq.

Применение технологий single-cell RNA-seq

Основные области и методы применения анализа данных scRNA-seq в различных биологических и медицинских исследованиях показаны на рис. 2. Биоинформатический анализ данных scRNA-seq позволяет исследоовать различные биохимические и регуляторные пути за счет изучения дифференциальной экспрессии различных генов для каждой отдельной клетки, функция которых может быть определена через поиск гомологии и исследования онтологий (Janes et al., 2010; Shalek et al., 2013; Trapnell et al., 2014; Treutlein et al., 2014). Новая технология дала возможность кластеризовать клетки по типу или состоянию (Jaitin et al., 2014); регистрировать редкие гены, которые при секвенировании общего транскриптома отбрасываются как минорные фракции (Gerber et al., 2017); изучать точечные мутации (Gawad et al., 2016; Ludwig et al., 2019).

Благодаря исследованию и анализу профилей экспрессии тысяч генов в сотнях тысяч отдельных клеток установлены их новые типы, состояния, обнаружены не исследованные ранее регуляторные пути (Saliba et al., 2014; Grun et al., 2015; Okaty et al., 2015; Zeisel et al., 2015; Poulin et al., 2016; Tirosh et al., 2016; Callaway, 2017; Lavin et al., 2017). Созданы карты развития различных клеточных линий (Segal et al., 2004; Pijuan-Sala et al., 2019; Taylor et al., 2019), показана их высокая гетерогенность (Mahata et al., 2014; Wang, Song, 2017).

Сочетание различных молекулярных методов с применением технологий scRNA-seq позволяет проводить эксперименты на принципиально новом уровне. Так, для

изучения состояний клеток и их дифференцировки во время онтогенеза возможно использование системы CRISPR-Cas9, с помощью которой происходит баркодирование клеток за счет внедрения специфических мутаций в геном или применения различных флуоресцентных меток (Adamson et al., 2016; Jaitin et al., 2016; McKenna et al., 2016; Kalhor et al., 2017). Например, показано, что длинные некодирующие РНК, представленные всего в нескольких копиях в клетке, могут иметь важные регуляторные функции (Derrien et al., 2012). При использовании технологий захвата и секвенирования единичных клеток в сочетании с определением конформации хромосом (метод Hi-C) (Belton et al., 2012) и трехмерным (3D) моделированием структуры хроматина установлено, что структуры отдельных топологически связанных доменов и петель существенно различаются от клетки к клетке (Stevens et al., 2017), а хромосомные перестройки, контактная изоляция, топологически связанные домены (TADs) или устойчивые петли хроматина находятся под управлением определенной динамики клеточного цикла (Nagano et al., 2017).

Существующие платформы для изоляции отдельных клеток

Основная сложность при изучении транскриптомов индивидуальных клеток – диссоциация, захват каждой клетки и ее подготовка для дальнейшего секвенирования (мечение всех транскриптов каждой клетки). Кроме этого, необходимо ввести в каждую молекулу мРНК определенный баркод, который позволит на этапе биоинформатического анализа сортировать данные по отдельным клеткам. Современное оборудование для scRNA-seq представлено коммерческими моделями, отличающимися производительностью: Puncher Platform Vycap, CellRaft AIR System, PEPArray System, Fluidigm C1, Wafergen ICELL8, BioRad Illumina ddSEQ, Dolomite Bio Nadia и RNA-Seq System, Tapestri Platform MissionBio, 1CellBio InDrop, BD Rhapsody, Chromium 10XGenomics (Kolodziejczyk et al., 2015; Valihrach et al., 2018).

Все методы, используемые для изоляции (захвата) одиночных клеток, можно разделить на низкопроизводительные (медленное разделение, захват от десяти до сотен клеток) и высокопроизводительные (быстрое разделение, захват от сотен до нескольких тысяч клеток) (Wang, Navin, 2015; Poulin et al., 2016). К первым относятся последовательное разведение (Ham, 1965), механические микроманипуляции (Brehm-Stecher et al., 2004), микропипетирование, лазерная захватывающая микродиссекция (Laser-Capture Microdissection, LCM); ко вторым – сортировка по степени флуоресценции (Fluorescence Activated Cell Sorting, FACS), микрофлюидика (microfluidics), работа с микрокаплями (microdroplets) (Navin et al., 2011; Landry et al., 2013; Mazutis et al., 2013). Так как в настоящее время наиболее востребованы высокоэффективные платформы захвата клеток, рассмотрим их более подробно. FACS – один из наиболее эффективных и экономически выгодных методов изоляции сотен тысяч клеток в минуту, основанный на размерности, гранулярности и флуоресцентных свойствах клеток. Однако применение флуоресцентных красителей может отрицательно влиять на жизнеспособность клеток (Lindström et al., 2010).

Активно развиваются и широко применяются при анализе не только транскриптома, но и полного генома отдельных клеток, а также эпигеномных модификаций методы изоляции клеток, основанные на микрофлюидике и использовании микрокапель (microdroplets) (Zheng et al., 2017). Эти способы позволяют работать с нанолитрами жидкостей, выделять ДНК отдельных органелл и при этом получать точные результаты (Whitesides, 2006; Salafi et al., 2016). В настоящее время существуют высокопроизводительные коммерческие платформы для изоляции клеток, основанных на микрофлюидике, которые не только разделяют клетки, но и обеспечивают дальнейшее проведение биохимических реакций для молекулярно-генетических исследований: процессы баркодирования, проведение обратной транскрипции, синтез кДНК и амплификация (Poulin et al., 2016; Valihrach et al., 2018).

«Капельные» методы стали прорывом в РНК-секвенировании и сделали возможными параллельную обработку большого числа клеток и дифференциацию клеточного происхождения каждой мРНК. Этот подход реализован в таких системах, как Drop-seq (Macosko et al., 2015), InDrop (Klein et al., 2015) и коммерческая платформа Chromium 10XGenomics (Kolodziejczyk et al., 2015). Он основан на технологии, которая отделяет высокомолекулярные фрагменты ДНК или целые одиночные клетки в мицеллы, содержащие адаптеры и уникальные баркоды (Coombe et al., 2016). В результате получается суспензия, содержащая нанокапли, в каждой из которых находятся одна клетка и все необходимые реактивы в наноколичествах для лизиса, баркодирования, обратной транскрипции и синтеза кДНК. Основное преимущество этих платформ – высокая производительность захвата и подготовки единичных клеток для дальнейшего секвенирования. Один из лидеров в этой области – Chromium 10XGenomics, позволяющий одновременно проводить захват до 80 тыс. клеток, в то время как число анализируемых клеток у других высокопроизводительных платформ колеблется от 10 до 48 тыс. клеток (Valihrach et al., 2018).

Анализ транскриптома единичных клеток с использованием высокопроизводительных платформ

Изучение транскриптома единичных клеток состоит из экспериментальной и биоинформатической частей (Kumar et al., 2017; Li et al., 2017). Прежде чем начинать эксперимент по секвенированию транскриптомов единичных клеток, необходимо иметь референсный геном и транскриптом, на который будет происходить картирование данных scRNA-seq (Gawad et al., 2014). Наиболее производительная платформа – Chromium™ 10XGenomics. Ниже рассмотрим особенности эксперимента scRNA-seq на ее основе в сравнении с другими методами для изучения отдельных клеток.

Изоляция клеток и методы scRNA-seq

Основная задача пробоподготовки – получение суспензии жизнеспособных неагрегированных клеток (диссоциация клеток). Необходимо определиться с концентрацией клеток, так как для удачного захвата требуется определенное их число в зависимости от выбранного метода изоляции

клеток: от десяти клеток (при использовании микропипетирования, цитоплазматической аспирации, лазерной микродиссекции) до тысяч клеток (при использовании приборов на основе технологий FACS, микрофлюидики и микрокапель) (Der et al., 2017). В случае работы с высокоэффективными приборами, например Chromium 10XGenomics, концентрация клеток должна быть примерно 10^6 клеток в мл (Der et al., 2017). При работе с животными или тканями, содержащими число клеток меньше необходимого, нужно увеличивать число образцов на пробу. Все этапы диссоциации клеток проводят в минимальном объеме раствора (от 50 до 1000 мкл) для повышения концентрации и уменьшения возможных потерь клеток.

Каждый протокол секвенирования транскриптомов одиночных клеток состоит из трех этапов: обратная транскрипция, амплификация кДНК (WTA – полнотранскриптомная амплификация) и подготовка библиотеки. Несмотря на нежелательность амплификации кДНК (из-за возможности возникновения ошибок полимеразы или потери редких транскриптов), данный этап необходим для создания библиотеки, так как принято, что количество общей РНК в клетке составляет около 10 пг, что недостаточно для успешного секвенирования (Wang, Song, 2017). В зависимости от задачи и используемой платформы для изоляции клеток будут отличаться протоколы для обратной транскрипции и получения кДНК (Haque et al., 2017; Kumar et al., 2017; Ziegenhain et al., 2017).

В настоящее время можно выделить три основных подхода. Первым был предложен метод с использованием олиго-dT-праймеров, конъюгированных с адаптерами, для обратной транскрипции и избирательной амплификации полиаденилированной мРНК с помощью ПЦР (Tang et al., 2009). Этот протокол имеет существенный недостаток: из-за смещения в область сгенерированных 3′-концов во время обратной транскрипции происходит потеря информации для анализа альтернативного сплайсинга.

Позже был разработан подход, позволяющий конструировать полноразмерную кДНК, – так называемый синтез кДНК со сменой матрицы (template switching cDNA synthesis) (Ramskölld et al., 2012). Преимущество данного метода заключается в получении и амплификации полноразмерной кДНК, что позволяет определять варианты альтернативного сплайсинга и аллель-специфическую экспрессию (ASE) (Kolodziejczyk et al., 2015). Такой подход используется в протоколах STRT (Islam et al., 2014), SMART-seq и SMART-seq2 (Ramsköld et al., 2012; Deng et al., 2014). Особенность перечисленных выше протоколов – амплификация, в ходе которой происходит экспоненциальный рост числа транскриптов, что будет приводить к смещению в ходе анализа и потере минорных экспрессированых генов. В качестве альтернативы был разработан подход транскрипции *in vitro* (IVT) для линейной амплификации кДНК, который представлен в таких протоколах для анализа единичных клеток, как CEL-Seq (Hashimshony et al., 2012) и MARS-Seq (Jaitin et al., 2014).

Третий подход заключается в дополнительном использовании уникальных молекулярных идентификаторов (UMI), представляющих собой случайные короткие последовательности от 6 до 10 п. н., встраиваемые в олиго-dT-праймер и помогающие различить отдельные молекулы. Эта технология показана в таких протоколах для scRNA-seq, как CEL-Seq (Hashimshony et al., 2012) и CEL-Seq2 (Hashimshony et al., 2016), Drop-seq (Macosko et al., 2015), MARS-Seq (Jaitin et al., 2014), SCRB-seq (Soumillon et al., 2014), STRT (Islam et al., 2014), In-Drop (Klein et al., 2015). Один из последних протоколов с использованием молекулярных идентификаторов – Quartz-Seq2 (Sassagawa et al., 2018), позволяет анализировать до 1536 клеток из одной пробы и повышает эффективность преобразования UMI с 22 % (для других протоколов scRNA-seq) до 35 %. Это дает возможность получить информацию о большом числе генов.

Последние достижения в параллельной работе с тысячами клеток потребовали усовершенствования баркодирования транскриптов. Наиболее современным и инновационным подходом, используемым в платформах, основанных на микрофлюидике и технологии микрокапель, служит применение дополнительно клеточного баркода (олигонуклеотид длиной ~14 п. н.) одновременно с праймерами, несущими на себе UMI, которые потом помещаются в каждую каплю с отдельными клетками. Клеточный баркод служит идентификатором всех последовательностей нуклеотидов из различных клеток. Преимущество такого двойного баркодирования – высокая точность и возможность определения клетки, из которой получена каждая отдельная РНК (Islam et al., 2014). Для секвенирования на платформе Drop-Seq разработан протокол STAMPs (Single-cell Transcriptomes Attached to Microparticles) и протокол Cell-Seq – для платформы InDrop (Wang, Song, 2017).

Наиболее высокопроизводительная коммерческая платформа Chromium 10XGenomics интегрировала технологию Gemcode, которая разделяет в каплях длинные молекулы ДНК и баркодирует их для создания библиотек под секвенирование (Eisenstein, 2015; Coombe et al., 2016). Использование двух баркодов на Chromium 10XGenomics при работе с единичными клетками позволяет уменьшить технический шум и проанализировать одновременно тысячи различных клеток, идентифицируя принадлежность каждого транскрипта, что особенно актуально при работе со сложными тканями. Это дает возможность определять профили экспрессии генов в масштабе одной клетки (Zheng et al., 2017).

Существуют различные адаптеры, позволяющие подготовить библиотеки микроРНК, например 3′-концевой адаптер, содержащий 5′,5′-аденил пирофосфорилированный участок (Hafner et al., 2008; Chen et al., 2012).

После получения кДНК происходит процесс секвенирования. Выбор технологии секвенирования должен непосредственно зависеть от поставленных целей. Различная длина получаемых последовательностей (от пятидесяти до нескольких тысяч нуклеотидов), точность, производительность – все это необходимо учитывать при выборе метода секвенирования (Liu et al., 2012).

Биоинформатический анализ данных

С развитием высокопроизводительных платформ для одновременного захвата десятков тысяч клеток потребовались и новые биоинформатические подходы для рабо-

ты с такими массивами данных. Следует отметить, что при анализе scRNA-seq не всегда возможно использование традиционных методов анализа суммарного транскриптома. Ниже рассмотрим сложности, возникающие при работе с данными транскриптомов единичных клеток, и возможные пути их решения.

Один эксперимент scRNA-seq способен дать информацию о каждом транскриптоме для сотен тысяч клеток. Это приводит к получению огромного массива данных и техническим сложностям при их анализе. Выравнивание и подсчет прочтений осуществляются независимо для каждой клетки и требуют «параллелизации процесса», а, соответственно, большей вычислительной мощности (Tokunaga et al., 2014; Yu, Lin, 2016).

scRNA-seq позволяет анализировать профили экспрессии для каждой отдельной клетки. Благодаря этому существуют три основных направления анализа (см. рис. 2): идентификация клеточных популяций (типы клеток и их состояния), реконструкция клеточной иерархии (дифференциация клеток при эмбриогенезе или клеточные ответы на различные стимулы) и поиск регуляторных сетей (на основе экспрессии генов) (Stegle et al., 2015). Каждая из этих задач сводится к следующим этапам биоинформатического анализа: контроль качества прочтений; фильтрация; выравнивание; картирование на референсный геном/транскриптом; формирование матрицы прочтений; сокращение размерности; нормализация; поиск внешних факторов, вносящих погрешность в анализ; кластеризация клеток; поиск маркерных и высоковариабельных генов, анализ дифференциальной экспрессии и изоформ; поиск корреляций в экспрессии различных генов (Stegle et al., 2015; Hwang et al., 2018). Для некоторых этапов возможно использование стандартных программ и подходов, разработанных для общего транскриптома. Ниже кратко рассмотрены особенности анализа только данных scRNA-seq.

При анализе профилей экспрессии клеток необходимо учитывать факторы, приводящие к ошибкам на этапе анализа и интерпретации данных. Их можно разделить на две категории: технический шум (например, различная эффективность захвата клеток, амплификация дуплетов, неамплифицированные гены – “dropout”, влияние пробоподготовки – так называемый batch effect) и биологические факторы (стохастическая экспрессия генов, наличие клеточного цикла, влияние окружающей среды) (Stegle et al., 2015; Andrews, Hemberg, 2018; Hwang et al., 2018).

Особенность протоколов при захвате клеток – вероятность попадания двух клеток в одну каплю (1–10 % в зависимости от типа используемой платформы), что приводит к появлению дуплетов на этапе секвенирования и возникновению ошибки при идентификации типов клеток (Segerstolpe et al., 2016). С другой стороны, на этапе диссоциации при пробоподготовке в части клеток могли произойти деградация РНК или некачественный лизис клетки, что опять же приведет к неправильной классификации (Brennecke et al., 2013; Hwang et al., 2018). Такие клетки должны быть исключены из дальнейшего анализа на этапе контроля и оценки качества данных. Наиболее эффективно применение протоколов, включающих использование уникальных молекулярных идентификаторов (UMI) (Hashimshony et al., 2012; Macosko et al., 2015) и дополнительной чужеродной РНК (spike-in) (Jiang et al., 2011) известной последовательности и концентрации, которые добавляются к каждой клетке на этапе захвата и подготовке к секвенированию. Это позволяет рассчитать критерии, определяющие качество полученных транскриптомов для каждой клетки: долю картированных ридов чужеродной РНК и долю амплифицированной мРНК в каждой клетке (Stegle et al., 2015). Можно также выявить технические погрешности на этапе постановки эксперимента и уменьшить количество ошибок в ходе биоинформатического анализа с помощью повторности scRNA-seq.

Кроме того, внедрение чужеродной РНК позволяет более точно оценить различия в количестве РНК у разных клеток на этапе нормализации данных (Stegle et al., 2015). Наиболее часто используемые параметры для нормализации данных – RPKM (Mortazavi et al., 2008), FPKM и TPM (Li et al., 2010), альтернативные подходы – TMM и DESeq (Robinson, Oshlack, 2010; Li et al., 2012). В то же время были разработаны методы нормализации специально для данных scRNA-seq (Lun et al., 2016; Bacher et al., 2017).

Так как при секвенировании транскриптомов для тысяч клеток мы получаем огромное количество данных по всем генам, то это создает трудности при статистическом анализе. Поэтому важный этап при работе с такими данными – выбор стратегии, уменьшающей размерность массива данных (Andrews, Hemberg, 2018). Существует два подхода: сокращение размерности и удаление из массива неинформативных генов. К первому относятся: метод главных компонент – PCA (Pierson, Yau, 2015); стохастические методы – tSNE (Maaten, Hinton, 2008); диффузионные карты – DM (Moon et al., 2018). Для определения неинформативных генов можно проводить поиск высоковариабельных генов (HGV), оценку неамплифицированных генов в различных клетках (M3Drop), анализировать положительную или отрицательную корреляцию экспрессии генов между различными клетками, использовать методы, основанные на чужеродной ДНК (spike-in) (Andrews, Hemberg, 2018). Некоторые программы позволяют применять обе стратегии, например PAGODA (Fan et al., 2016), что наиболее эффективно.

При кластеризации клеток и выделении отдельных типов следует помнить, что огромную ошибку может вносить и различное биологическое состояние, в котором находится та или иная клетка. Так, уровень экспрессии различных генов будет отличаться в зависимости от фазы клеточного цикла – G1 или G2 (Stegle et al., 2015).

Для построения регуляторных сетей и реконструкции клеточной иерархии был разработан алгоритм, который невозможно осуществить для данных RNA-seq. Он основан на упорядочивании транскрипционных состояний различных клеток, которые размещаются на траектории, характеризующей какой-либо развивающийся биологический процесс в организме, например апоптоз (Haghverdi et al., 2016).

Многие этапы проводят с использованием целых программных пакетов, предусмотренных для работы с данными транскриптомов единичных клеток (Valihrach et al., 2018). Так, для Chromium 10XGenomics разработчик предоставляет программный конвейер CellRanger (Zheng

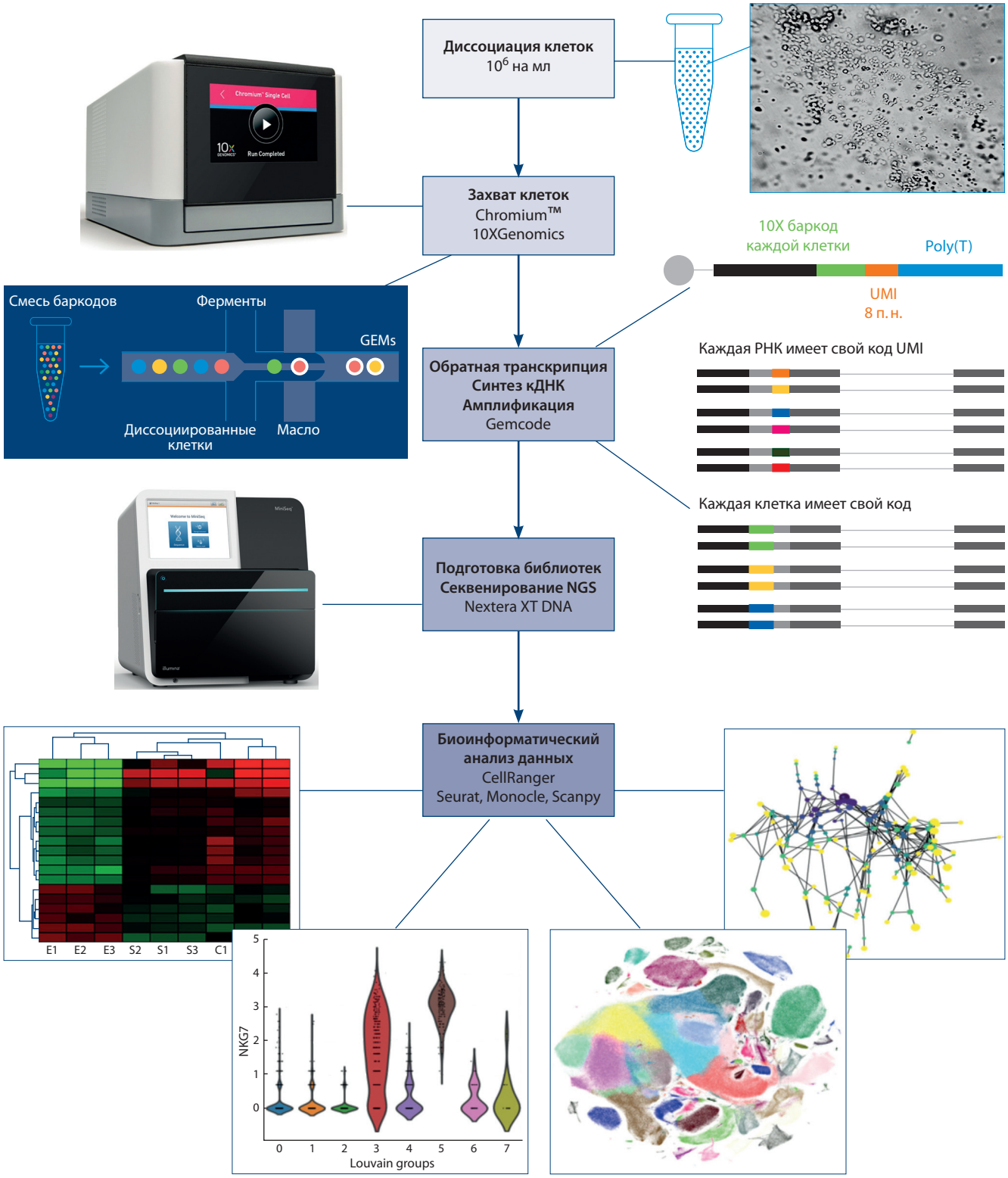

Рис. 3. Процесс секвенирования транскриптомов единичных клеток на высокопроизводительной платформе Chromium 10XGenomics.

et al., 2017); для Drop-Seq нет программного обеспечения для первичного анализа от разработчика, но есть несколько сторонних разработок, например zUMIs (Parekh et al., 2018), scPipe (Tian et al., 2018), Dr.Seq2 (Zhao et al., 2017).

Наиболее популярные программные пакеты для дальнейшего анализа данных scRNA-seq – Seurat (Butler et al., 2018) и Monocle (Qiu et al., 2017). Seurat – это программный пакет, обеспечивающий контроль качества, анализ и исследование данных единичных клеток RNA-seq. Программное обеспечение включает в себя три компонента: неконтролируемую кластеризацию и обнаружение типов и состояний клеток, пространственную реконструкцию и

интегрированный анализ единичных клеток RNA-seq по условиям, технологиям и видам. Monocle – комплексный программный пакет, который обеспечивает инструменты для анализа экспериментов с экспрессией единичных клеток. Однако эти программные пакеты не масштабируются для все более доступных больших наборов с количеством, превышающим миллион клеток, или при объединении данных нескольких экспериментов и сравнительном анализе. Альтернативой может стать Scanpy, который преодолевает это ограничение и предоставляет аналогичные возможности анализа (Wolf et al., 2018).

Поэтапная реализация эксперимента по секвенированию транскриптомов единичных клеток с использованием наиболее высокопроизводительной платформы Chromium 10XGenomics показана на рис. 3.

Несмотря на быстрое развитие этих технологий, возможно возникновение ошибок на различных этапах эксперимента scRNA-seq (Fustin et al., 2013; Nikolenko et al., 2013; Schwartz et al., 2013; Gawad et al., 2014, 2016; Poulin et al., 2016). Это необходимо учитывать при планировании и проведении РНК-секвенирования отдельных клеток и на этапе биоинформатического анализа.

Заключение

Секвенирование РНК единичных клеток показало, что изменение профилей экспрессии генов наблюдается не только в клетках различных тканей, но и в процессе онтогенеза, а также в результате воздействия различных внешних факторов. Благодаря своей разрешающей способности метод секвенирования транскриптомов отдельных клеток используется в различных областях медицины и биологии и является инновационным для изучения механизмов конвергентной эволюции, независимого возникновения систем в различных таксономических группах, эволюционных изменений в клеточных линиях, определения новых типов клеток и определения их функций. Рассмотренные в этой статье работы демонстрируют необходимость выполнения полных исследований, например целых организмов, опухолей или тканей, что требует развития технологий параллельного секвенирования транскриптомов огромного числа клеток. Показано, что в настоящее время основные критерии выбора технологии проведения scRNA-seq – производительность приборов для захвата и метод баркодирования. Наиболее инновационной и эффективной по совокупности параметров на сегодняшний день представляется коммерческая платформа Chromium 10XGenomics с интегрированной технологией Gemcode, позволяющая проводить такой анализ параллельно до 80 тыс. клеток.

Список литературы / References

Adamson B., Norman T.M., Jost M., Cho M.Y., Nunez J.K., Chen Y., Villalta J.E., Gilbert L.A., Horlbeck M.A., Hein M.Y., Pak R.A., Gray A.N., Gross C.A., Dixit A., Parnas O., Regev A., Weissman J.S. A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. Cell. 2016;167(7):1867-1882.e21. DOI 10.1016/j.cell.2016.11.048.

Andrews T.S., Hemberg M. Identifying cell populations with scRNASeq. Mol. Aspects Med. 2018;59:114-122. DOI 10.1016/j.mam.2017. 07.002.

Bacher R., Chu L.-F., Leng N., Gasch A.P., Thomson J.A., Stewart R.M., Newton M., Kendziorski C. SCnorm: robust normalization of single-cell RNA-seq data. Nat. Methods. 2017;14(6):584-586. DOI 10.1038/nmeth.4263.

Belton J.M., McCord R.P., Gibcus J.H., Naumova N., Zhan Y., Dekker J. Hi-C: a comprehensive technique to capture the conformation of genomes. Methods. 2012;58(3):268-276. DOI 10.1016/j.ymeth.2012.05.001.

Brehm-Stecher B.F., Johnson E.A. Single-cell microbiology: tools, technologies, and applications. Microbiol. Mol. Biol. Rev. 2004; 68(3):538-559.

Brennecke P., Anders S., Kim J.K., Kolodziejczyk A.A., Zhang X., Proserpio V., Baying B., Benes V., Teichmann S.A., Marioni J.C., Heisler M.G. Accounting for technical noise in single-cell RNA-seq experiments. Nat. Methods. 2013;10:1093-1095.

Butler A., Hoffman P., Smibert P., Papalexi E., Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 2018;36(5):411-420. DOI 10.1038/nbt.4096.

Callaway A. The trickiest family tree in biology. Nature. 2017; 547(7661):20-22. DOI 10.1038/547020a.

Chen Y.R., Zheng Y., Liu B., Zhong S., Giovannoni J., Fei Z. A cost-effective method for Illumina small RNA-Seq library preparation using T4 RNA ligase 1 adenylated adapters. Plant Methods. 2012;8(1):41. DOI 10.1186/1746-4811-8-41.

Coombe L., Warren R.L., Jackman S.D., Yang C., Vandervalk B.P., Moore R.A., Pleasance S., Coope R.J., Bohlmann J., Holt R.A., Jones S.J.M., Birol I. Assembly of the complete Sitka spruce chloroplast genome using 10× Genomics' GemCode sequencing data. PLoS One. 2016;11(9):e0163059. DOI 10.1371/journal.pone.0163059.

Deng Q., Ramsköld D., Reinius B., Sandberg R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science. 2014;343(6167):193-196. DOI 10.1126/science.1245316.

Der E., Ranabothu S., Suryawanshi H., Akat K.M., Clancy R., Morozov P., Kustagi M., Czuppa M., Izmirly P., Belmont H.M., Wang T., Jordan N., Bornkamp N., Nwaukoni J., Martinez J., Goilav B., Buyon J.P., Tuschl T., Putterman C. Single cell RNA sequencing to dissect the molecular heterogeneity in lupus nephritis. JCI Insight. 2017;2(9). pii: 93009. DOI 10.1172/jci.insight.93009.

Derrien T., Johnson R., Bussotti G., Tanzer A., Djebali S., Tilgner H., Guernec G., Martin D., Merkel A., Knowles D.G., Lagarde J., Veeravalli L., Ruan X., Ruan Y., Lassmann T., Carninci P., Brown J.B., Lipovich L., Gonzalez J.M., Thomas M., Davis C.A., Shiekhattar R., Gingeras T.R., Hubbard T.J., Notredame C., Harrow J., Guigo R. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012;22(9):1775-1789. DOI 10.1101/gr.132159.111.

Eisenstein M. Startups use short-read data to expand long-read sequencing market. Nat. Biotechnol. 2015;33(5):433-435. DOI 10.1038/nbt0515-433.

Fan J., Salathia N., Liu R., Kaeser G.E., Yung Y.C., Herman J.L., Kaper F., Fan J.-B., Zhang K., Chun J., Kharchenko P.V. Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. Nat. Methods. 2016;13:241-244. DOI 10.1038/nmeth.3734.

Frieda K.L., Linton J.M., Hormoz S., Choi J., Chow K.-H.K., Singer Z.S., Budde M.W., Elowitz M.B., Cai L. Synthetic recording and *in situ* readout of lineage information in single cells. Nature. 2017;541(7635):107-111. DOI 10.1038/nature20777.

Fustin J.M., Doi M., Yamaguchi Y., Hida H., Nishimura S., Yoshida M., Isagawa T., Morioka M.S., Kakeya H., Manabe I., Okamura H. RNA-methylation-dependent RNA processing controls the speed of the circadian clock. Cell. 2013;155(4):793-806. DOI 10.1016/j.cell.2013.10.026.

Gawad C., Koh W., Quake S.R. Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics. Proc.

Natl. Acad. Sci. USA. 2014;111(50):17947-17952. DOI 10.1073/pnas.1420822111.

Gawad C., Koh W., Quake S.R. Single-cell genome sequencing: current state of the science. Nat. Rev. Genet. 2016;17(3):175-188. DOI 10.1038/nrg.2015.16.

Gerber T., Willscher E., Loeffler-Wirth H., Hopp L., Schadendorf D., Schartl M., Anderegg U., Camp G., Treutlein B., Binder H., Kunz M. Mapping heteroge-neity in patient-derived melanoma cultures by single-cell RNA-seq. Oncotarget. 2017;8(1):846-862. DOI 10.18632/oncotarget.13666.

Goldberg A.D., Allis C.D., Bernstein E. Epigenetics: a landscape takes shape. Cell. 2007;128(4):635-638.

Grun D., Lyubimova A., Kester L., Wiebrands K., Basak O., Sasaki N., Clevers H., Oudenaarden A. Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature. 2015;525(7568):251-255. DOI 10.1038/nature14966.

Hafner M., Landgraf P., Ludwig J., Rice A., Ojo T., Lin C., Holoch D., Lim C., Tuschl T. Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. Methods. 2008;44(1):3-12.

Haghverdi L., Büttner M., Wolf F.A., Buettner F., Theis F.J. Diffusion pseudotime robustly reconstructs lineage branching. Nat. Methods. 2016;13:845-848.

Ham R.G. Clonal growth of mammalian cells in a chemically defined, synthetic medium. Proc. Natl. Acad. Sci. USA. 1965;53:288-293.

Haque A., Engel J., Teichmann S.A., Lonnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome Med. 2017;9:75.

Hashimshony T., Senderovich N., Avital G., Klochendler A., Leeuw Y., Anavy L., Gennert D., Li S., Livak K.J., Rozenblatt-Rosen O., Dor Y., Regev A., Yanai I. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol. 2016;17:77. DOI 10.1186/s13059-016-0938-8.

Hashimshony T., Wagner F., Sher N., Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep. 2012;2(3):666-673. DOI 10.1016/j.celrep.2012.08.003.

Hwang B., Lee J.H., Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp. Mol. Med. 2018;50:96.

Islam S., Zeisel A., Joost S., Manno G.L., Zajac P., Kasper M., Lonnerberg P., Linnarsson S. Quantitative single-cell RNA-seq with unique molecular identifiers. Nat. Methods. 2014;11(2):163-166. DOI 10.1038/nmeth.2772.

Jaitin D.A., Kenigsberg E., Keren-Shaul H., Elefant N., Paul F., Zaretsky I., Mildner A., Cohen N., Jung S., Tanay A., Amit I. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science. 2014;343(6172):776-779. DOI 10.1126/science.1247651.

Jaitin D.A., Weiner A., Yofe I., Lara-Astiaso D., Kern-Shaul H., David E., Salame T.M., Tanay A., Oudenaarden A., Amit I. Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. Cell. 2016;167(7):1883-1896. DOI 10.1016/j.cell.2016.11.039.

Janes K.A., Wang C.C., Holmberg K.J., Cabral K., Brugge J.S. Identifying single-cell molecular programs by stochastic profiling. Nat. Methods. 2010;7(4):311-317. DOI 10.1038/nmeth.1442.

Jiang L., Schlesinger F., Davis C.A., Zhang Y., Li R., Salit M., Gingeras T.R., Oliver B. Synthetic spike-in standards for RNA-seq experiments. Genome Res. 2011;21:1543-1551.

Junker J.P., Oudenaarden A. Every cell is special: genome-wide studies add a new dimension to single-cell biology. Cell. 2014;157(1):8-11. DOI 10.1016/j.cell.2014.02.010.

Kalhor R., Mali P., Church G.M. Rapidly evolving homing CRISPR barcodes. Nat. Methods. 2017;14(2):195-200. DOI 10.1038/nmeth.4108.

Kamme F., Salunga R., Yu J., Tran D.T., Zhu J., Luo L., Bittner A., Guo H.Q., Miller N., Wan J., Erlander M. Single-cell microarray analysis in hippocampus CA1: demonstration and validation of cellular heterogeneity. J. Neurosci. 2003;23(9):3607-3615.

Klein A.M., Mazutis L., Akartuna I., Tallapragada N., Veres A., Li V., Peshkin L., Weitz D.A., Kirschner M.W. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell. 2015;161(5):1187-1201. DOI 10.1016/j.cell.2015.04.044.

Kolodziejczyk A.A., Kim J.K., Svensson V., Marioni J.C., Teichmann S.A. The technology and biology of single-cell RNA sequencing. Mol. Cell. 2015;58(4):610-620. DOI 10.1016/j.molcel.2015.04.005.

Kumar P., Tan Y., Cahan P. Understanding development and stem cells using single cell-based analyses of gene expression. Development. 2017;144:17-32. DOI 10.1242/dev.133058.

Landry Z.C., Giovanonni S.J., Quake S.R., Blainey P.C. Optofluidic cell selection from complex microbial communities for single-genome analysis. Methods Enzymol. 2013;531:61-90. DOI 10.1016/B978-0-12-407863-5.00004-6.

Lavin Y., Kobayashi S., Leader A., Amir E.D., Elefant N., Bigenwald C., Remark R., Sweeney R., Becker C.D., Levine J.H., Meinhof K., Chow A., Kim-Shulze S., Wolf A., Medaglia C., Li H., Rytlewski J.A., Emerson R.O., Merad M. Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses. Cell. 2017;169(4):750-765.e17. DOI 10.1016/j.cell.2017.04.014.

Leung M.L., Davis A., Gao R., Casasent A., Wang Y., Sei E., Vilar E., Maru D., Kopetz S., Navin N.E. Single-cell DNA sequencing reveals a late-dissemination model in metastatic colorectal cancer. Genome Res. 2017;27(8):1287-1299. DOI 10.1101/gr.209973.116.

Li B., Ruotti V., Stewart R.M., Thomson J.A., Dewey C.N. RNA-seq gene expression estimation with read mapping uncertainty. Bioinformatics. 2010;26:493-500.

Li H., Courtois E.T., Sengupta D., Tan Y., Chen K.H., Goh J.J.L., Kong S.L., Chua C., Hon L.K., Tan W.S., Wong M., Choi P.J., Wee L.J.K., Hillmer A.M., Tan I.B., Robson P., Prabhakar S. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. Nat. Genet. 2017;49:708-718.

Li J., Witten D.M., Johnstone I.M., Tibshirani R. Normalization, testing, and false discovery rate estimation for RNA-sequencing data. Biostatistics. 2012;13:523-538.

Lindström S., Andersson-Svahn H. Overview of single-cell analyses: microdevices and applications. Lab. Chip. 2010;10(24):3363-3372. DOI 10.1039/c0lc00150c.

Liu L., Li Y., Li S., Hu N., He Y., Pong R., Lin D., Lu L., Law M. Comparison of next-generation sequencing systems. J. Biomed. Biotechnol. 2012;2012:251364. DOI 10.1155/2012/251364.

Ludwig L.S., Lareau C.A., Ulirsch J.C., Christian E., Muus C., Li L.H., Pelka K., Ge W., Oren Y., Brack A., Law T., Rodman C., Chen J.H., Boland G.M., Hacohen N., Rozenblatt-Rosen O., Aryee M.J., Buenrostro J.D., Regev A., Sankaran V.G. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. Cell. 2019;176(6);1325-1339. DOI 10.1016/j.cell.2019.01.022.

Lun A.T.L., Bach K., Marioni J.C. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. Genome Biol. 2016;17:75. DOI 10.1186/s13059-016-0947-7.

Maaten L. van der, Hinton G. Visualizing data using t-SNE. J. Mach. Learn. Res. 2008;9: 2579-2605.

Macosko E.Z., Basu A., Satija R., Nemesh J., Shekhar K., Goldman M., Tirosh I., Bialas A.R., Kamitaki N., Martersteck E.M., Trombetta J.J., Weitz D.A., Sanes J.R., Shalek A.K., Regev A., McCarroll S.A. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell. 2015;161(5):1202-1214. DOI 10.1016/j.cell.2015.05.002.

Mahata B., Zhang X., Kolodziejczyk A.A., Proserpio V., Haim-Vilmovsky L., Taylor A.E., Hebenstreit D., Dingler F.A., Moignard V., Gottgens B., Arlt W., McKenzie A.N.J., Teichmann S.A. Single-cell RNA sequencing reveals T helper cells synthesizing steroids *de novo* to contribute to immune homeostasis. Cell Rep. 2014;7(4):1130-1142. DOI 10.1016/j.celrep.2014.04.011.

Marcus J.S., Anderson W.F., Quake S.R. Microfluidic single-cell mRNA isolation and analysis. Anal. Chem. 2006;78(9):3084-3089.

Mazutis L., Gilbert J., Ung W.L., Weitz D.A., Griffiths A.D., Heyman J.A. Single-cell analysis and sorting using droplet-based microfluidics. Nat. Protoc. 2013;8(5):870-891. DOI 10.1038/nprot.2013.046.

McKenna A., Findlay G.M., Gagnon J.A., Horwitz M.S., Schier A.F., Shendure J. Whole-organism lineage tracing by combinatorial and cumulative genome editing. Science. 2016;353(6298):aaf7907. DOI 10.1126/science.aaf7907.

Mi D., Li Z., Lim L., Li M., Moissidis M., Yang Y., Gao T., Hu T.X., Pratt T., Price D.J., Sestan N., Marin O. Early emergence of cortical interneuron diversity in the mouse embryo. Science. 2018;360(6384):81-85. DOI 10.1126/science.aar6821.

Moon K.R., Dijk D., Wang Z., Burkhardt D., Chen W.S., Yim K., Elen A., Hirn M.J., Coifman R.R., Ivanova N.B., Wolf G., Krishnaswamy S. Visualizing structure and transitions for biological data exploration. Cell. 2018;65. http://dx.doi.org/10.2139/ssrn.3155891.

Moroz L.L. NeuroSystematics and periodic system of neurons: model vs reference species at single-cell resolution. ACS Chem. Neurosci. 2018;9:1884-1903. DOI 10.1021/acschemneuro.8b00100.

Moroz L.L., Edwards J.R., Puthanveettil S.V., Kohn A.B., Ha T., Heyland A., Knudsen B., Sahni A., Yu F., Liu L., Jezzini S., Lovell P., Iannucculli W., Chen M., Nguyen T., Sheng H., Shaw R., Kalachikov S., Panchin Y.V., Farmerie W., Russo J.J., Ju J., Kandel E.R. Neuronal transcriptome of Aplysia: neuronal compartments and circuitry. Cell. 2006;127(7):1453-1467.

Mortazavi A., Williams B.A., McCue K., Schaeffer L., Wold B. Mapping and quantifying mammalian transcriptomes by RNA-seq. Nat. Methods. 2008;5:621-628.

Nagano T., Lubling Y., Stevens T.J., Schoenfelder S., Yaffe E., Dean W., Laue E.D., Tanay A., Fraser P. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature. 2013;502(7469):59-64. DOI 10.1038/nature12593.

Nagano T., Lubling Y., Varnai C., Dudley C., Leung W., Baran Y., Cohen N.M., Wingett S., Fraser P., Tanay A. Cell-cycle dynamics of chromosomal organization at single-cell resolution. Nature. 2017;547(7661):61-67. DOI 10.1038/nature23001.

Navin N., Kendall J., Troge J., Andrews P., Rodgers L., McIndoo J., Cook K., Stepansky A., Levy D., Esposito D., Muthuswamy L., Krasnitz A., McCombie W.R., Hicks J., Wingler M. Tumour evolution inferred by single-cell sequencing. Nature. 2011;472(7341):90-94. DOI 10.1038/nature09807.

Nikolenko S.I., Korobeynikov A.I., Alekseyev M.A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. BMC Genomics. 2013;14(Suppl. 1):S7. DOI 10.1186/1471-2164-14-S1-S7.

Nowogrodzki A. The cell seeker. Nature. 2017;547:24-26.

Okaty B.W., Freret M.E., Rood B.D., Brust R.D., Hennessy M.L., Bairos D., Kim J.C., Cook M.N., Dymecki S.M. Multi-scale molecular deconstruction of the serotonin neuron system. Neuron. 2015;88(4):774-791. DOI 10.1016/j.neuron.2015.10.007.

Parekh S., Ziegenhain C., Vieth B., Enard W., Hellmann I. zUMIs – A fast and flexible pipeline to process RNA sequencing data with UMIs. GigaScience. 2018;7(6). DOI 10.1093/gigascience/giy059.

Pierson E., Yau C., ZIFA: Dimensionality reduction for zero-inflated single-cell gene expression analysis. Genome Biol. 2015;16:241. DOI 10.1186/s13059-015-0805-z.

Pijuan-Sala B., Griffiths J.A., Guibentif C., Hiscock T.W., Jawaid W., Calero-Nieto F.J., Mulas C., Ibarra-Soria X., Tyser R.C.V., Ho D.L.L., Reik W., Srinivas S., Simons B.D., Nihols J., Marioni J.C., Gottgens B. A single-cell molecular map of mouse gastrulation and early organogenesis. Nature. 2019;566;490-495. DOI 10.1038/s41586-019-0933-9.

Poulin J.F., Tasic B., Hjerling-Leffler J., Trimarchi J.M., Awatramani R. Disentangling neural cell diversity using single-cell transcriptomics. Nat. Neurosci. 2016;19(9):1131-1141. DOI 10.1038/nn.4366.

Qiu X., Mao Q., Tang Y., Wang L., Chawla R., Pliner H., Trapnell C. Reversed graph embedding resolves complex single-cell developmental trajectories. Nat. Methods. 2017;14(10):979-982. DOI 10.1038/nmeth.4402.

Ramsköld D., Luo S., Wang Y.C., Li R., Deng Q., Faridani O.R., Daniels G.A., Khrebtukova I., Loring J.F., Laurent L.C., Schroth G.P., Sandberg R. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat. Biotechnol. 2012;30(8):777-782.

Robinson M.D., Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11:R25. DOI 10.1186/gb-2010-11-3-r25.

Rotem A., Ram O., Shoresh N., Sperling R.A., Goren A., Weitz D.A., Bernstein B.E. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat. Biotechnol. 2015;33(11):1165-1172. DOI 10.1038/nbt.3383.

Salafi T., Zeming K.K., Zhang Y. Advancements in microfluidics for nanoparticle separation. Lab. Chip. 2016;17(1):11-33.

Saliba A.E., Westermann A.J., Gorski S.A., Vogel J. Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res. 2014;42(14):8845-8860. DOI 10.1093/nar/gku555.

Sassagawa Y., Danno H., Takada H., Ebisawa M., Tanaka K., Hayashi T., Kurisaki A., Nikaido I. Quartz-Seq2: a high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads. Genome Biol. 2018;19:29. DOI 10.1186/s13059-018-1407-3.

Schwartz S., Agarwala S.D., Mumbach M.R., Jovanovic M., Mertins P., Shishkin A., Tabach Y., Mikkelsen T.S., Satija R., Ruvkun G., Carr S.A., Lander E.S., Fink G.R., Regev A. High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. Cell. 2013;155(6):1409-1421. DOI 10.1016/j.cell.2013.10.047.

Segal E., Friedman N., Koller D., Regev A. A module map showing conditional activity of expression modules in cancer. Nat. Genet. 2004;36(10):1090-1098.

Segerstolpe A., Palasantza A., Eliasson P., Andersson E.M., Andreasson A.C., Sun X., Picelli S., Sabirsh A., Clausen M., Bjursell M.K., Smith D.M., Kasper M., Ammala C., Sandberg R. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. Cell Metab. 2016;24:593-607. DOI 10.1016/j.cmet.2016.08.020.

Shalek A.K., Satija R., Adiconis X., Gertner R.S., Gaublomme J.T., Raychowdhury R., Schwartz S., Yosef N., Malboeuf C., Lu D., Trombetta J.J., Gennert D., Gnirke A., Goren A., Hacohen N., Levin J.Z., Park H., Regev A. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature. 2013;498(7453):236-240. DOI 10.1038/nature12172.

Soumillon M., Cacchiarelli D., Semrau S., van Oudenaarden A., Mikkelsen T.S. Characterization of directed differentiation by high-throughput single-cell RNA-seq. bioRxiv. 2014. DOI 10.1101/003236.

Stegle O., Teichmann S.A., Marioni J.C. Computational and analytical challenges in single-cell transcriptomics. Nat. Rev. Genet. 2015;16(3):133-145. DOI 10.1038/nrg3833.

Stevens T.J., Lando D., Basu S., Atkinson L.P., Cao Y., Lee S.F., Leeb M., Wohlfahrt K.J., Boucher W., O'Shaughnessy-Kirwan A., Cramard J., Faure A.J., Ralser M., Blanco E., Morey L., Sanso M., Palayret M.G.S., Lehner B., Croce L.D., Wutz A., Hendrich B., Klenerman D., Laue E.D. 3D structures of individual mammalian genomes studied by single-cell Hi-C. Nature. 2017;544(7648):59-64. DOI 10.1038/nature21429.

Subkhankulova T., Gilchrist M.J., Livesey F.J. Modelling and measuring single cell RNA expression levels find considerable transcriptional differences among phenotypically identical cells. BMC Genomics. 2008;9:268. DOI 10.1186/1471-2164-9-268.

Tang F., Barbacioru C., Wang Y., Nordman E., Lee C., Xu N., Wang X., Bodeau J., Tuch B.B., Siddiqui A., Lao K., Surani M.A. mRNA-Seq whole-transcriptome analysis of a single cell. Nat. Methods. 2009;6(5):377-382. DOI 10.1038/nmeth.1315. Epub 2009 Apr 6.

Taylor D.M., Aronow B.J. The pediatric cell atlas: defining the growth phase of human development at single-cell resolution. Dev. Cell. 2019;49. DOI 10.1016/j.devcel.2019.03.001.

Tian L., Su S., Dong X., Amann-Zalcenstein D., Biben C., Seidi A., Hilton D.J., Naik S.H., Ritchie M.E. scPipe: a flexible R/Bioconductor preprocessing pipeline for single-cell RNA-sequencing data. PLoS Comput. Biol. 2018;14(8):e100636. DOI 10.1371/journal.pcbi.1006361.

Tirosh I., Izar B., Prakadan S.M., Wadsworth M.H., Treacy D., Trombetta J.J., Rotem A., Rodman C., Lian C., Murphy G., Fallahi-Sichani M., Dutton-Regester K., Lin J.-R., Cohen O., Shah P., Lu D., Genshaft A.S., Hughes T.K., Ziegler C.G.K., Kazer S.W., Gaillard A., Kolb K.E., Villani A.C., Johannessen C.M., Andreev A.Y., Allen E.M.V., Bertagnolli M., Sorger P.K., Sullivan R.J., Flaherty K.T., Frederick D.T., Jane-Valbuena J., Yoon C.H., Rozenblatt-Rosen O., Shalek A.K., Regev A., Garraway L.A. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016;352(6282):189-196. DOI 10.1126/science.aad0501.

Tokunaga T., Hirose O., Kawaguchi S., Toyoshima Y., Teramoto T., Ikebata H., Kuge S., Ishihara T., Iino Y., Yoshida R. Automated detection and tracking of many cells by using 4D live-cell imaging data. Bioinformatics. 2014;30:43-51.

Trapnell C., Cacchiarelli D., Grimsby J., Pokharel P., Li S., Morse M., Lennon N.J., Livak K.J., Mikkelsen T.S., Rinn J.L. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. 2014;32(4):381-386. DOI 10.1038/nbt.2859.

Treutlein B., Brownfield D.G., Wu A.R., Neff N., Mantalas G.L., Espinoza F.H., Desai T.J., Krasnow M.A., Quake S.R. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature. 2014;509(7500):371-375. DOI 10.1038/nature 13173.

Valihrach L., Androvic P., Kubista M. Platforms for single-cell collection and analysis. Int. J. Mol. Sci. 2018;19:807.

Wang J., Song Y. Single cell sequencing: a distinct new field. Clin. Transl. Med. 2017;6(1):10. DOI 10.1186/s40169-017-0139-4.

Wang Y., Navin N.E. Advances and applications of single-cell sequencing technologies. Mol. Cell. 2015;58(4):598-609. DOI 10.1016/j.molcel.2015.05.005.

Whitesides G.M. The origins and the future of microfluidics. Nature. 2006;442(7101):368-373.

Wolf F.A., Angerer P., Theis F.J. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018;19(1):15. DOI 10.1186/s13059-017-1382-0.

Yu P., Lin W. Single-cell transcriptome study as big data. Genom. Proteom. Bioinf. 2016;14:21-30.

Zeisel A., Munoz-Manchado A.B., Codeluppi S., Lonnerberg P., Manno G., Jureus A., Margues S., Munguba H., He L., Betsholtz C., Rolny C., Castelo-Branco G., Hjerling-Leffler J., Linnarsson S. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science. 2015;347(6226):1138-1142. DOI 10.1126/science.aaa1934.

Zhao C., Hu S., Huo X., Zhang Y. Dr.seq2: a quality control and analysis pipeline for parallel single cell transcriptome and epigenome data. PLoS One. 2017;12(7):e0180583. DOI 10.1371/journal.pone.0180583.

Zheng G.X., 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., Bielas J.H. Massively parallel digital transcriptional profiling of single cells. Nat. Commun. 2017;8:14049. DOI 10.1038/ncomms14049.

Ziegenhain C., Vieth B., Parekh S., Reinius B., Guillaumet-Adkins A., Smets M., Leonhardt H., Heyn H., Hellmann I., Enard W. Comparative analysis of single-cell RNA sequencing methods. Mol. Cell. 2017;65:631-643. DOI 10.1015/j.molcel.2017.01.023.

ORCID ID
E.A. Vodiasova orcid.org/0000-0003-3886-2880
E.S. Chelebieva orcid.org/0000-0002-7662-2573
O.N. Kuleshova orcid.org/0000-0003-3745-7066

Благодарности. Работа выполнена при финансовой поддержке Правительства РФ, грант № 14.WO3.31.0015.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 14.09.2018. После доработки 30.05.2019. Принята к публикации 30.05.2019.