



**Jörn Walter, Nina Gasparoni**

---

## **1. Einleitung**

In:

Walter, Jörn / Schickl, Hannah (Hrsg.): Einzelzellanalyse in Forschung und Medizin : eine Stellungnahme der interdisziplinären Arbeitsgruppe Gentechnologiebericht.

ISBN: 978-3-939818-84-7

Berlin-Brandenburgische Akademie der Wissenschaften, 2019. S. 10-22

Persistent Identifier: [urn:nbn:de:kobv:b4-opus4-32797](https://nbn-resolving.org/urn:nbn:de:kobv:b4-opus4-32797)

---

Die vorliegende Datei wird Ihnen von der Berlin-Brandenburgischen Akademie der Wissenschaften unter einer Creative Commons Namensnennung 4.0 International Lizenz zur Verfügung gestellt.



## 1. EINLEITUNG

### 1.1 VOM KOMPLEXEN GEWEBE ZUR EINZELZELLSIGNATUR – NEUE HORIZONTE FÜR DIE MODERNE ZELLBIOLOGIE

Seit dem Aufkommen der modernen Zellbiologie Anfang des 20. Jahrhunderts suchen Wissenschaftler nach Techniken zur Erfassung der molekularen Mechanismen, die biologische Programme in einzelnen Zellen des Organismus steuern. Mit der Einzelzellanalytik steht der Wissenschaft seit Kurzem ein Mittel zur Verfügung, das umfassende und präzise Daten über den molekularen Charakter und die Funktionsweise *einzelner* Zellen generiert. Die Interpretation dieser Daten eröffnet der Forschung tiefgreifende neue Möglichkeiten des Verständnisses komplexer biologischer Vorgänge in einer Zelle, von komplexen Entwicklungsprozessen und Alterung bis hin zur Anpassung an Umweltbedingungen, und von komplexen Vorgängen der Organentwicklung bis hin zu Ursachen und Folgen von Erkrankungen. Mithilfe der neuen Technologien lassen sich diese Prozesse zellgenau erfassen und zwar simultan für tausende bis zu Millionen einzelner Zellen. Neue Verfahren ermöglichen es, die so gewonnenen Daten für die Modellierung der räumlichen Zuordnung der Zelle im Gewebe und in ihrer entwicklungsbiologischen Dynamik zu nutzen. Damit bringt die Einzelzellanalytik Biologen ihrem ursprünglichen Ziel, Eigenschaften und Funktionsweisen von Zellen im Organismus präzise zu verorten, zu verstehen und zu beeinflussen, erheblich näher.

Bisherige funktionelle Konzepte zellulärer Programme gründen sich auf genetische, biochemische und molekulare Daten. Vor den Möglichkeiten der Einzelzellanalyse mussten umfassende Untersuchungen an Zellgemischen oder vermeintlich homogenen Zellpopulationen durchgeführt werden, die zuvor aus Geweben oder Körperflüssigkeiten isoliert wurden. Die molekularen Signaturen (wie bspw. Genexpressionsprofile) solcher Zellpopulationen spiegeln jedoch immer die Summe der Eigenschaften der einzelnen Zellen wider und erlauben es nicht, die individuelle Variations- und Funktionsbreite einzelner Zellen zu erfassen. Veränderungen, die während der Entwicklung, im Verlauf des Zellzyklus und bei Prozessen des individuellen Zellalters auftreten oder die als Auswirkungen ihrer räumlichen Anordnung entstehen, sind mit den herkömmlichen Methoden nur bedingt analysierbar. Derartige individuelle Eigenschaften lassen sich durch die Analyse von

Zellpopulationen nicht adäquat bestimmen. Hinzu kommt das Problem, dass die Analyse von Zellpopulationen einer Selektion oder Vorsortierung nach bestimmten Eigenschaften bedarf, sodass nicht alle Zellen eines Gewebes gleichzeitig erfasst werden können.

Mit der Entwicklung umfassender Einzelzell-Omics<sup>1</sup>-Technologien bieten sich nun tiefgreifende Lösungen für viele dieser Probleme. In Kombination mit neuen verzerrungsfreien Sortiertechniken, erweiterten Mikroskopie-Verfahren wie Multi-RNA-FISH<sup>2</sup> und neuartigen bioinformatischen Ansätzen werden Einzelzellanalysen Antworten auf bislang unlösbare Fragen liefern und neue systemische Einblicke in die Funktion einzelner Zellen in einer komplexen biologischen Umgebung eröffnen.

## 1.2 TECHNOLOGISCHE ENTWICKLUNGEN UND DIE ENTSTEHUNG DER MODERNEN EINZELZELLBIOLOGIE

Die Grundlage der modernen „Einzelzell-Omics“ wurde durch die funktionelle Annotation des menschlichen Genoms und der Genome aller wichtigen Modellorganismen gelegt. Die Lokalisierung von Genen und anderen funktionellen bzw. regulatorischen Teilen des Genoms hat zahlreiche Studien zur funktionellen Genomik beflügelt, die zum Ziel haben, verschiedenen Zelltypen ihre spezifischen molekularen Programme zuzuordnen. Diese „funktionelle Genomrevolution“ wurde durch die schnelle Entwicklung neuartiger technologischer Fortschritte im Bereich der Hochdurchsatz-Sequenzierung ermöglicht, den sogenannten Next-Generation-Sequencing-Technologien (NGS-Technologien). NGS-Methoden, die ursprünglich für die Genomsequenzierung entwickelt worden sind, wurden schnell für Funktionsanalysen von Zellen nutzbar gemacht, zum Beispiel für die umfassende Erstellung von Genexpressionsprofilen (zur Darstellung der jeweils in einer Zelle exprimierten Gene) mit NGS-basierten RNA-Sequenzierungsmethoden (RNA-seq)<sup>3</sup>.

- 1 „Omics“ ist ein Neologismus, der verschiedene Forschungsfelder in den Lebenswissenschaften beschreibt, die die Wortendung „omics“ enthalten, z. B. Genomik, Transkriptomik, Metabolomik und Proteomik. Die Wortendung zeigt an, dass der Fokus der Studien auf dem gesamten zellulären Gehalt an untersuchten Molekülen liegt (d. h. auf der Gesamtheit der Gene, Gentranskripte, Stoffwechselprodukte oder Proteine der Zelle).
- 2 FISH (Fluoreszenz-in-situ-Hybridisierung) ist eine Methode, bei der bestimmte Moleküle einer Probe mit fluoreszierenden Markern versehen werden, um sie nachweisen zu können. „In situ“ bringt zum Ausdruck, dass die Moleküle in ihrer natürlichen Position (also in der Zelle, im Gewebe) untersucht werden können.
- 3 Die RNA-Sequenzierung nutzt NGS, um die Anzahl und Anwesenheit von RNA-Transkripten in einer Probe zum Zeitpunkt der Untersuchung nachzuweisen (auch „Gesamttranskriptom-Shotgun-Sequenzierung“ genannt). Da diese Methode genutzt wird, um das zelluläre Transkriptom zu analysieren, gehört sie zur Transkriptomik.

Vor etwa 10 Jahren wurden die ersten manuellen Versuche unternommen, umfassende mRNA<sup>4</sup>-seq-Signaturen aus einigen wenigen sortierten einzelnen Säugetierzellen zu erhalten (Tang et al., 2009). Diese ersten erfolgreichen Anwendungen haben das neue Forschungsgebiet beflügelt und sehr schnell wurden neue Hochdurchsatzmethoden für die Einzelzellisolierung und die NGS-Prozessierung entwickelt, um umfassende RNA-seq-basierte Genexpressionsprofile für *vieler* Einzelzellen auf einmal zu erhalten. Alle diese neuartigen Techniken kombinieren Mikroprozessierungs-Techniken mit anspruchsvollen Protokollen zur Erstellung komplexer Sequenzier-Bibliotheken einzelner Zellen.<sup>5</sup> Die neuen Technologien erweitern die Forschungsmöglichkeiten in zwei wichtige Richtungen: i) eine immer umfassendere Darstellung molekularer Signaturen in einzelnen Zellen und ii) die Möglichkeit, eine große Anzahl von Einzelzellen in kostengünstigeren Hochdurchsatzsystemen gleichzeitig zu analysieren.<sup>6</sup>

Das Forschungsfeld begann mit der Erstellung einer Reihe von hochauflösenden Einzelzellkarten für Blutzellen, die neue tiefe Einblicke in die Entwicklung sowie die Anpassung der verschiedenen Immunzelltypen im Verlauf von Erkrankungen gewähren (Kowalczyk et al., 2015; Wilson et al., 2015). Bald darauf wurden die ersten umfassenden Einzelzellkarten von Modellorganismen erstellt, gefolgt von jüngsten Veröffentlichungen zu Geweben und Organen bei Menschen und Mäusen (Han et al., 2018), wie Leber (Halpern et al., 2017; MacParland et al., 2018, Aizarani et al., 2019), Gehirn (Darmanis et al., 2015; Lake et al., 2018; Rosenberg et al., 2018), Niere (Magella et al., 2018), Lunge (Treutlein et al., 2014; Xu et al.,

- 4 mRNA ist die Abkürzung für messenger- (deutsch „Boten-“) RNA und bezeichnet das Molekül, das Produkt der Genexpression ist. Im Zuge eines Prozesses, der „Transkription“ genannt wird, dient die DNA als Vorlage für den Aufbau der RNA, die wiederum zu mRNA prozessiert wird. Die mRNA verlässt den Zellkern und wird in eine Aminosäuresequenz übersetzt und bildet so ein Protein. mRNA ist daher das Transkript (Abschrift) der korrespondierenden DNA, und das Studium des RNA-Gehaltes einer Zelle wird Transkriptomik genannt. Das Transkriptom einer Zelle enthält ihre gesamte RNA und ermöglicht Rückschlüsse darauf, welche Gene in einer bestimmten Zelle (zu einem bestimmten Zeitpunkt) exprimiert sind.
- 5 Vom Transkriptom werden DNA-Kopien erstellt (mRNA wird in DNA umgeschrieben, sogenannte cDNA für komplementäre DNA), d. h. es wird eine „Bibliothek“ der in einer Zelle vorhandenen einzelnen mRNA Moleküle erstellt, die dann mit Hilfe von Hochdurchsatz-Sequenzierung (NGS) ausgelesen werden kann. Man bestimmt so das Vorhandensein und die Anzahl der mRNA-Kopien eines Gens.
- 6 Aktuelle RNA-seq-Methoden nutzen entweder Mikrofluidik-Systeme, wie die gängigsten „Drop-seq“-Methoden, oder sie verwenden „zellcontainerähnliche“ Lösungen in Mikrotiterplatten. In beiden Systemen werden die RNA-Bibliotheken einzelner Zellen mittels ausgefeilter individueller (Adapter-) Kodierungstechniken markiert, um die Expressionsprofile einzelner Zellen unterscheiden zu können. Einige der neuen Technologien wie „MARS-seq“ (Jaitin et al., 2014; Keren-Shaul et al., 2019), „Drop-seq“ (Macosko et al., 2015), „Seq-Well“ (Gierahn et al., 2017), „SPLiT-seq“ (Rosenberg et al., 2018) – um nur einige der aktuellen „Spitzenreiter“ zu nennen – haben einen hohen Durchsatz erreicht, der die Generierung von RNA-seq-Signaturen für Millionen von Zellen zu angemessenen Sequenzierungskosten ermöglicht.

2016), oder sogar zu ganzen Tieren (Drosophila, Embryonalstadien der Maus: Karaiskos et al., 2017; Mohammed et al., 2017). Diese ersten umfassenden Analysen von Geweben und sich entwickelnden Organismen zeigen, dass es neben der Identifizierung neuer (bisher nicht spezifizierter) Zelltypen oder Zellzustände möglich ist, Ähnlichkeiten und Veränderungen von Zellfunktionen über Gewebe hinweg zu identifizieren. So kann die Dynamik von Zellpopulationen, d. h. deren Auftreten bzw. Verschwinden, während der Entwicklung ebenso erfasst werden wie die Heterogenität von Zelltypen in erkrankten Geweben (z. B. Krebs) oder die Variationsbreite der Zellzusammensetzung in alternden Geweben – um nur einige der faszinierendsten Facetten zu nennen.

### 1.3 INITIATIVEN ZUR ENTWICKLUNG DES BEREICHS DER EINZELZELL-OMICS

Mit dem Aufkommen von Einzelzellmessungen wurde schnell klar, dass für vergleichende Analysen eine Art Standardisierung sowohl auf experimenteller als auch auf der Ebene der Dateninterpretation erforderlich sein wird. Internationale Forschungskonsortien wie der Human Cell Atlas (HCA)<sup>7</sup> und auch LifeTime<sup>8</sup> wurden gegründet, um das sich schnell entwickelnde Forschungsfeld in diesen Aufgaben voranzubringen durch die Bereitstellung von qualitativ hochwertigen Einzelzelldaten mit definierten Standards und Kontrollen in Referenzdatenbanken. Erste Datenbanken wurden für Maus, Mensch und Drosophila eingerichtet, von denen solche zellspezifischen Einzelzelldaten abgerufen werden können.<sup>9</sup> Der Human Cell Atlas war das erste internationale Konsortium, das 2017 mit dem Ziel gegründet wurde, einen umfassenden Einzelzellatlas aller menschlichen Zellen zu erstellen und neue „cloud“-basierte Informatiklösungen zur Datenspeicherung und -analyse zu entwickeln. Die im Jahr 2018 gestartete europäische Initiative LifeTime ergänzt diese Bemühungen, indem sie sich auf medizinische Anwendungen in verschiedenen krankheitsrelevanten Bereichen konzentriert. Ein Hauptziel ist die Entwicklung und Analyse krankheitsrelevanter Modelle sowie die Entwicklung neuartiger Ansätze, die auf die klinische Verwendung von Einzelzelldaten übertragen werden können. In Deutschland wurde darüber hinaus 2018 das vom BMBF geförderte Netzwerk Single Cell Omics Germany (SCOG)<sup>10</sup> mit dem Ziel gegründet, ein erstes Netzwerk von Laboren aufzubauen,

7 Siehe: <https://www.humancellatlas.org/> [13.08.2019].

8 Siehe: <https://lifetime-fetflagship.eu/> [13.08.2019]. Siehe auch: Junker, Popp, Rajewsky, Kapitel 2.

9 Eine umfassende Datenbankliste ist abrufbar unter: <https://www.singlecell.de/index.php/resources/databases/> [13.08.2019].

10 Siehe: <https://www.singlecell.de/> [16.08.2019].

die Einzelzellanalysen durchführen, und Weiterbildungen in Einzelzelltechnologien anzubieten, insbesondere in aufstrebenden Bereichen wie der Einzelzell-Multi-Omics, der umfassenden Datenanalyse und -interpretation.

Diese gemeinsamen Bemühungen sind darauf ausgerichtet, eine wissenschaftliche Gemeinschaft aufzubauen und zu fördern, die zusammen einen vollständigen Atlas aller Zelltypen des menschlichen Körpers in Einzelzellauflösung ausarbeitet, welcher auch komplexe Szenarien wie Gewebe, Organalterung und Krankheiten umfasst. Diese Daten sollen der Forschungsgemeinschaft frei zugänglich gemacht und neue informatische Ansätze für ihre tiefgehende biologische Interpretation entwickelt werden. Die Komplexität der Einzelzelldaten und die vielen neuen Fragestellungen, die mit dieser Art von hochauflösenden Daten beantwortet werden können, erfordern die Entwicklung neuer bioinformatischer Ansätze, die weit über die für NGS-Daten entwickelten Anwendungen hinausgehen. Dieser Aspekt wird in Kapitel 4 von Aliee, Sacher und Theis näher erläutert.

#### 1.4 AKTUELLE ENTWICKLUNGEN IM BEREICH DER EINZELZELL-OMICS

Gegenwärtig konzentriert sich die Mehrzahl der Einzelzell-Omics-Messungen auf RNA-seq, wobei hauptsächlich Expressionssignaturen des letzten Exons<sup>11</sup> eines Gens erfasst werden (z. B. Chromium (Zheng et al., 2017), Drop-seq). Solche Ansätze sind recht robust und eignen sich gut für die Erzeugung von groben Zellsignaturen, die eine Unterscheidung von Hauptzelltypen ermöglicht. Allerdings erlauben sie nicht die Untersuchung komplexerer Veränderungen in genregulatorischen Programmen, wie alternative Transkript- oder Spleiß-Varianten von Genen, die häufig eine andere und wichtige funktionelle Rolle spielen. Um das Spektrum der Gentranskription in größerer Tiefe zu erfassen, werden umfassendere RNA-seq-Methoden zur Herstellung von Bibliotheken entwickelt, wie zum Beispiel Smart-seq2 (Picelli et al., 2013) und andere (Chen et al., 2019). Solche Ansätze sind zwar noch recht kostspielig, werden aber durch sinkende Sequenzierungskosten immer mehr zur Routine, da die Fülle solcher Daten eine bessere und tiefergehende Interpretation ermöglicht.

Die Erzeugung hochauflösender Signaturen einzelner Zellen geht zu Lasten der räumlichen Orientierung der einzelnen Zellen. Das bedeutet, dass die Forscher

<sup>11</sup> Exons sind Teile des ursprünglichen RNA-Transkripts, die die mRNA infolge eines als „Spleißen“ (Splicing) bezeichneten Prozesses bilden. Die Detektion des letzten Exons dient als Nachweis dafür, dass das Gen transkribiert worden ist.

nicht wissen, welche Umgebung und Position die Zelle innerhalb des Gewebes hatte. Dieses Wissen ist jedoch wichtig, um die Einzelzelldaten richtig interpretieren und in das vorhandene Wissen über Gewebe und den gesamten Organismus integrieren zu können. Zur Lösung dieses Problems werden derzeit Methoden entwickelt, die eine räumliche Anordnung der analysierten Einzelzellen in „virtuelle Gewebe“ erlauben. Die wichtigsten gängigen Verfahren verwenden hochauflösende Bildgebungsverfahren, um Expressionssignaturen in Geweben für eine ausreichende Anzahl von Genen und Zellen zu erstellen, zum Beispiel durch mehrfarbigen RNA-FISH. Derartige *in situ* quantifizierte Expressionsbilder von Genen in einzelnen Zellen von Gewebeschnitten nutzt man dann als „Ankerpunkte“ für eine räumliche Rekonstruktion der Einzelzell-RNA-seq-Signaturen, um anschließend aus den Einzelzelldaten eine Art „virtuell“ rekonstruiertes Gewebe zu erzeugen. Damit wissen die Forscher genau, wann und wo ein Gen in einer Zelle des Organismus angeschaltet ist. Auf diese Verfahren gehen Junker, Popp und Rajewsky in Kapitel 2 genauer ein. Ein alternativer Ansatz, um Ankerpunkte zu generieren, ist das Sammeln weniger oder einzelner Zellen aus definierten Regionen innerhalb von Gewebeschnitten mittels Laser-Capture-Mikroskopie, gefolgt von einer tiefen (Einzel-)Zellsequenzierung (Nichterwitz, 2016; Chen et al., 2017).

Die umfassende und mechanistische Interpretation von Einzelzell-Omics-Daten ist zudem angewiesen auf den Vergleich mit anderen (idealerweise ebenfalls Einzelzell-)Omics-Daten wie zum Beispiel genomischen und funktionellen Epigenomik-Daten (Chromatin und DNA-Modifikationen). Bisher wurden Referenzdaten zur Genomik und funktionellen Genomik nur für Zellgemische von Konsortien wie dem Internationalen Krebs-Genom-Konsortium (ICGC),<sup>12</sup> 4DNucleome,<sup>13</sup> der Enzyklopädie der DNA-Elemente (ENCODE),<sup>14</sup> dem Internationalen Human-Epigenom-Konsortium (IHEC)<sup>15</sup> und anderen generiert. Für einige Methoden, zum Beispiel ATAC-seq, wurden einzelzellbasierte Protokolle entwickelt (Buenrostro et al., 2015), die auch bereits kommerzialisiert sind. Die meisten der anderen genomischen und epigenomischen NGS-basierten Methoden sind jedoch technisch anspruchsvoll und auf Einzelzellebene nur bedingt anwendbar. Bis vor Kurzem erschien es unmöglich, die Genexpression einzelner Zellen direkt mit deren epigenetischen Profilen zu verknüpfen. Pilotversuche von Clark et al. (2018) haben nun gezeigt, dass auch Einzelzelldaten zur Genexpression,

12 Siehe: <https://icgc.org/> [13.08.2019].

13 Siehe: <https://www.4dnucleome.org/> [13.08.2019].

14 Siehe: <https://www.encodeproject.org/> [13.08.2019].

15 Siehe: <http://ihec-epigenomes.org/> [13.08.2019].

DNA-Methylierung und Chromatinzugänglichkeit<sup>16</sup> gleichzeitig von derselben Zelle erhalten werden können. Darüber hinaus wurden die ersten hochtechnischen Ansätze entwickelt, um die dreidimensionale Konfiguration von Chromosomen in Einzelzellen zu bestimmen. Man gewinnt daraus Erkenntnisse über die räumliche Anordnung von Genen im Zellkern und die Bedeutung für die Regulierung der Genaktivität (Nagano et al., 2017).

Die integrierte Interpretation solcher Multi-Omics-Einzelzelldaten stellt ein wichtiges Zukunftsfeld in der Einzelzellbiologie dar. Sie schlägt Brücken zwischen (deskriptiven) transkriptionellen Signaturen einzelner Zellen und den Mechanismen, wie diese Genprogramme festgelegt und ausgeführt werden. Funktionelle Multi-Omics-Daten ermöglichen die Betrachtung biomedizinischer Fragen mit einer nie zuvor erreichten Auflösung und ebnen damit den Weg für ein präzises Verständnis der Mechanismen der Genregulation. Die Komplexität der generierten Daten stellt die Einzelzell-Bioinformatik jedoch vor große Herausforderungen, da unterschiedliche Datentypen (mit unterschiedlichen Dynamikbereichen) integriert und analysiert werden müssen.

## 1.5 EINZELZELLANALYSEN UND ENTWICKLUNGSBIOLOGIE

Studien im Bereich der Entwicklungsbiologie werden in hohem Maße von der Verwendung der Einzelzell-(Multi-)Omics profitieren. Die Auflösung auf Einzelzellebene bietet einen neuartigen umfassenden Überblick über Änderungen des Zellprogramms und deren dynamische Anpassung während der Entwicklung. Es ist wahrscheinlich, dass Einzelzelldaten unsere derzeitige Sicht auf (stochastische und gerichtete) Mechanismen, die Differenzierungsprozesse vorantreiben, sowie möglicherweise auch unsere (eher statische und auf Vorwissen basierende) Sicht auf die Definition von Zelltypen verändern. Sie werden sicherlich auch unser Verständnis dafür verbessern, wie Zellen sich an veränderte Umweltbedingungen anpassen.

Für Planarien (Strudelwürmer; Cao et al., 2017), frühe Mausembryonen (Peng et al., 2016; Mohammed et al., 2017) und Drosophila-Larven (Taufliegenlarven;

<sup>16</sup> Chromatin ist ein Komplex aus DNA und begleitenden Proteinen. Dabei ist die DNA um die sogenannten Histon-Proteine gewickelt und in sich verdrillt. Der Grad der Verdrillung (Kondensation) beeinflusst die Zugänglichkeit des Chromatins für weitere Proteine, die etwa Gene aktivieren oder inaktivieren können. Die DNA-Methylierung ist eine biochemische Modifizierung der DNA, die sich auf das Bindeverhalten regulatorischer Proteine und auf die Chromatin-Konformation auswirkt und so die Genexpression beeinflusst.



Karaiskos et al., 2017) wurden erste eindrucksvolle Beispiele für „Karten des Entwicklungsschicksals“ erstellt. Diese Daten liefern Einblicke in dynamische Veränderungen zellulärer Programme, die in kurzen Phasen selbst-reorganisierender Prozesse auftreten (z. B. während der Gastrulation). Sie ermöglichen es, die Bildung und Organisation von Zellen während der Entwicklung zum Beispiel des Gehirns, der Niere, des Herzens usw. räumlich und zeitlich zu verfolgen und die vorprogrammierten und stochastischen Mechanismen zu identifizieren, die die Diversifikation und Differenzierung der Zellen (z. B. bei der frühen Säugetierentwicklung) antreiben. Es wurden komplexe Rechenmodelle etabliert, um die Dynamik der Entwicklungstrajektorien von Zelllinien abzuleiten und zelluläre Veränderungsabläufe, die die Zelltypentwicklung festlegen, zu verfolgen. In Kombination mit genetischer Markierung oder einer Überlagerung mit mikroskopischen Referenzdaten ermöglichen solche hochauflösenden Omics-Daten ein tiefgehendes Verständnis der Mechanismen, die die räumliche und zeitliche Organisation von Entwicklungsübergängen in verschiedenen Organismen regulieren. Diese Aspekte werden im Kapitel 2 von Junker, Popp und Rajewsky sowie im Kapitel 4 von Aliee, Sacher und Theis näher erläutert.

## 1.6 EINZELZELLANALYSEN IN DER KRANKHEITSFORSCHUNG

Die Zusammensetzung und relative Lokalisierung von Zellen in einem Gewebe bzw. Organ sind wichtige Parameter, um deren physiologische Funktionen in Zusammenhang mit der natürlichen Organfunktion, der Homöostase, dem Altern, der Regeneration, aber auch mit Krankheiten zu verstehen. Einzelzell-Omics bieten einen unvoreingenommenen<sup>17</sup> Ansatz, um den genauen Zusammenhang zwischen der Zellzusammensetzung und der Organbiologie zu untersuchen. Darüber hinaus lassen sich die Folgen einer lokalen Funktionsstörung von Zellen im Organ verfolgen, zum Beispiel bei Prozessen, die zu Verletzungen, Narbenbildung, Fibrose (Vermehrung von Bindegewebe), Steatose (Verfettung) usw. führen. Die Einzelzellanalyse wird es ermöglichen, Veränderungen in der Zellzusammensetzung, die in pathologischen Situationen wie abnormaler Organentwicklung, Autoimmunerkrankungen, chronischen Krankheiten oder Krebs auftreten, direkt zu untersuchen. Die Bestimmung der zellulären Heterogenität in Tumoren oder in einzelnen Leukämiezellen eröffnet eine neue diagnostische Ebene zur Bestimmung der Herkunft, Progression und Heterogenität des Tumors für eine genauere,

17 „Unvoreingenommen“ bedeutet in diesem Zusammenhang, dass Forschende Ergebnisse erlangen können, ohne die Bandbreite möglicher Ergebnisse durch eigene, der Durchführung des Experiments vorangehende, Hypothesen einschränken zu müssen.

tumorspezifische Diagnose und bessere Vorhersagen für den erfolgreichen Einsatz therapeutischer Maßnahmen. Aschenbrenner, Mass und Schultze gehen in Kapitel 3 näher hierauf ein. Die umfangreichen medizinischen Anwendungsmöglichkeiten der Einzelzellanalyse lassen dabei auch die Frage aufkommen, inwiefern dieses Arbeitsfeld ethische Fragen aufwirft. Dieser Aspekt wird in Kapitel 6 von Fangerau, Marx-Stöltzing und Osterheider untersucht.

## 1.7 HERAUSFORDERUNGEN UND GRENZEN DER EINZELZELLTECHNOLOGIEN

Wie bei allen neuen technologischen Entwicklungen sind auch Einzelzellanalysen mit technischen und konzeptionellen Herausforderungen verbunden. Herausforderungen für die breite und routinemäßige Anwendung stellen insbesondere Schwierigkeiten der Vorbereitung von Einzelzellsuspensionen aus komplexen Geweben dar sowie Einschränkungen, RNA oder DNA (aber auch Lipide und Proteine) in ausreichender Menge und qualitativ hochwertig aus Einzelzellen gewinnen zu können (siehe Müller-Röber, Kapitel 5, zur weiteren Erörterung dieses Problems in Bezug auf Pflanzenzellen). Physikalische Einschränkungen wie Schwierigkeiten, bestimmte Zelltypen zum Beispiel in Gehirngeweben zu vereinzelnen, können Experimente stark beeinträchtigen. Neue Ansätze zur Verwendung isolierter Kerne und zur Analyse neugebildeter RNA können einige dieser Probleme überwinden (Krishnaswami et al., 2016) und sogar die Untersuchung von Krankheiten wie Alzheimer in konservierten Post-mortem-Geweben ermöglichen (siehe auch Aschenbrenner, Mass, Schultze, Kapitel 3).

Für die Interpretation von Multi-Omics-Einzelzelldaten und deren Modellierung wird es wesentlich sein, Parameter auf Einzelzellebene zu erschließen, die nicht ausschließlich auf Nukleinsäuren basieren, sondern auch das Vorhandensein von Proteinen, Lipiden und Stoffwechselprodukten in einer Zelle im Blick zu haben. Erste experimentelle Erfolge, die zeigen, dass dies leistbar ist, wurden bereits veröffentlicht, jedoch bleiben die technischen Möglichkeiten einer umfassenden Darstellung von Proteom- und Metabolom- sowie Lipidomspektren aus Einzelzellen noch immer sehr eingeschränkt (Duncan et al., 2019; Marx, 2019; Pasarelli et al., 2019).

Da sich Einzelzellanalysen auf Methoden stützen, die Zellen nach ihren Ähnlichkeiten in Expressionsprofilen gruppieren, ist es wichtig zu kontrollieren, ob die erhaltenen Gruppen und Muster das erwartete und vollständige Zellspektrum in

der Ausgangs-Einzelzellsuspension repräsentieren. Bei vielen Anwendungen, insbesondere beim Nachweis seltener Zellen, muss auch die Effizienz bei der Generierung einer Einzelzellbibliothek und bei der Sequenzierungstiefe berücksichtigt werden, da diese Anwendungen üblicherweise mit hohen Vorbereitungs- und Sequenzierungskosten verbunden sind. Darüber hinaus müssen experimentelle und bioinformatische Standards festgelegt werden, um eine Überinterpretation der NGS-Daten einzelner Zellen zum Beispiel aufgrund von technisch bedingten Ausfällen<sup>18</sup> einzelner Gensignaturen zu vermeiden (Van den Berge et al., 2018). Schließlich erfordert die Identifizierung und biologische Interpretation gruppierter Zellen bestimmte Vorkenntnisse, für die Schätzung der Zellzusammensetzung, zum Beispiel eine ungefähre Kenntnis der Anzahl der Zelltypen, und für die räumliche Rekonstruktion eine Orientierung durch zellspezifische „Marker- bzw. Anker-Gene“ (siehe Aliee, Sacher, Theis, Kapitel 4). Neue und alte Ansätze müssen (weiter-) entwickelt werden, um diesem Bedarf gerecht zu werden.

## 1.8 SCHLUSSBEMERKUNG

Einzelzell-Omics sind ein schnell wachsender und äußerst wichtiger Bereich der funktionellen Genomik. Ihr breites Anwendungs- und Datennutzungsspektrum wird die moderne Biologie und Medizin revolutionieren, in vielen Aspekten bereichern und in eine neue, tiefgreifend molekulare Richtung führen. Konzepte der Zell- und Systembiologie werden tiefer ergründet und teilweise in neuem Licht betrachtet werden. NGS-basierte Einzelzelldaten werden nahezu jedes biologische Forschungsfeld beeinflussen: von der grundlegenden Zellbiologie bis zur Entwicklungsbiologie, von der Physiologie bis zur Pathologie, von der Taxonomie bis zur Ökologie. Für die Medizin ist die Einzelzelldiagnostik eines der heißesten aufstrebenden Gebiete in der personalisierten Medizin mit einem hohen Potenzial, die Präzisionsdiagnostik auf eine neue Ebene zu heben. Der Erfolg der Einzelzellanalyse hängt dabei in hohem Maße von der Entwicklung neuartiger experimenteller und bioinformatischer Lösungen ab. Die Kernstrukturen für eine solche Entwicklung sind gegeben, erfordern aber eine ständige Investition und Anpassung.

<sup>18</sup> Ausfall bedeutet in diesem Zusammenhang, dass Transkripte aus technischen Gründen nicht nachgewiesen werden, z. B. unter anderem durch ineffizientes Umschreiben von RNA in cDNA.

## 1.9 LITERATUR

- Aizarani, N. et al. (2019): A human liver cell atlas reveals heterogeneity and epithelial progenitors. In: *Nature* 572: 199–204.
- Buenrostro, J. D. et al. (2015): Single-cell chromatin accessibility reveals principles of regulatory variation. In: *Nature* 523(7561): 486.
- Cao, J. et al. (2017): Comprehensive single-cell transcriptional profiling of a multi-cellular organism. *Science* 357(6352): 661–667.
- Chen, J. et al. (2017): Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. In: *Nature protocols* 12(3): 566.
- Chen, G./Shi, T. (2019): Single-cell RNA-seq technologies and related computational data analysis. In: *Frontiers in genetics* 10: 317.
- Clark, S. J. et al. (2018): scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. In: *Nature communications* 9(1): 781.
- Darmanis, S. et al. (2015): A survey of human brain transcriptome diversity at the single cell level. In: *Proceedings of the National Academy of Sciences* 112(23): 7285–7290.
- Duncan, K. D. et al. (2019): Advances in mass spectrometry based single-cell metabolomics. In: *Analyst* 144(23): 7285–7290.
- Gierahn, T. M. et al. (2017): Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. In: *Nature methods* 14(4): 395.
- Halpern, K. B. et al. (2017): Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. In: *Nature* 542(7641): 352.
- Han, X. et al. (2018): Mapping the mouse cell atlas by microwell-seq. In: *Cell* 172(5): 1091–1107.
- Jaitin, D. A. et al. (2014): Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. In: *Science* 343(6172): 776–779.
- Karaiskos, N. et al. (2017): The *Drosophila* embryo at single-cell transcriptome resolution. In: *Science* 358(6360): 194–199.
- Keren-Shaul, H. et al. (2019): MARS-seq2. 0: an experimental and analytical pipeline for indexed sorting combined with single-cell RNA sequencing. In: *Nature protocols* 14(6): 1841.

Kowalczyk, M. S. et al. (2015): Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. In: *Genome research* 25(12): 1860–1872.

Krishnaswami, S. R. et al. (2016): Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. In: *Nature protocols* 11(3): 499.

Lake, B. B. et al. (2018): Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. In: *Nature biotechnology* 36(1): 70.

Macosko, E. Z. et al. (2015): Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. In: *Cell* 161(5): 1202–1214.

MacParland, S. A. et al. (2018): Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. In: *Nature communications* 9(1): 4383.

Magella, B. et al. (2018): Cross-platform single cell analysis of kidney development shows stromal cells express Gdnf. In: *Developmental biology* 434(1): 36–47.

Marx, V. (2019): A dream of single-cell proteomics. In: *Nature methods*, online publication 12.08.2019. doi: 10.1038/s41592-019-0540-6.

Mohammed, H. et al. (2017): Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. In: *Cell reports* 20(5): 1215–1228.

Nagano, T. et al. (2017): Cell-cycle dynamics of chromosomal organization at single-cell resolution. In: *Nature* 547(7661): 61.

Nichterwitz, S. et al. (2016): Laser capture microscopy coupled with Smart-seq2 for precise spatial transcriptomic profiling. In: *Nature communications* 7: 12139.

Passarelli, et al. (2013): Single-cell lipidomics: characterizing and imaging lipids on the surface of individual *Aplysia californica* neurons with cluster secondary ion mass spectrometry. In: *Analytical chemistry* 85(4): 2231–2238.

Peng, G. et al. (2016): Spatial transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. In: *Developmental cell* 36(6): 681–697.

Picelli, S. et al. (2013): Smart-seq2 for sensitive full-length transcriptome profiling in single cells. In: *Nature methods* 10(11): 1096.

Rosenberg, A. B. et al. (2018): Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. In: *Science* 360(6385): 176–182.

Tang, F. et al. (2009): mRNA-Seq whole-transcriptome analysis of a single cell. In: *Nature methods* 6(5): 377.

Treutlein, B. et al. (2014): Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. In: *Nature* 509(7500): 371.

Van den Berge, K. et al. (2018): Observation weights unlock bulk RNA-seq tools for zero inflation and single-cell applications. In: *Genome biology* 19(1): 24.

Wilson, N. K. et al. (2015): Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. In: *Cell stem cell* 16(6): 712–724.

Xu, Y. et al. (2016): Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. In: *JCI insight* 1(20).

Zheng, G. X. Y. et al. (2017): Massively parallel digital transcriptional profiling of single cells. In: *Nat. Commun.* 8(14049), Online-Publikation 16.01.2017. DOI: 10.1038/ncomms14049.