

## KI findet Fehler bei RNA-Analyse

### Neue Algorithmen ermitteln Messfehler bei der Einzelzell-RNA-Sequenzierung

**Warum verhalten sich manche Zellen im menschlichen Körper nicht so, wie sie sollten, und bilden beispielsweise Krebsgeschwüre? Von der sogenannten Einzelzellanalyse versprechen Forscherinnen und Forscher sich Antworten. Bisher war diese Methode allerdings anfällig für Fehler. Ein Team der Technischen Universität München (TUM), des Helmholtz Zentrum München und des englischen Wellcome Sanger Institute hat neue Algorithmen entwickelt, die die Fehlerquellen mithilfe Künstlicher Intelligenz (KI) ausfindig machen.**

Alle Zellen des menschlichen Körpers zu kartieren um damit die Diagnose, Überwachung und Behandlung von Krankheiten verbessern zu können - das ist die Vision hinter dem internationalen Human-Cell-Atlas-Projekt. Mit einer solchen Referenzdatenbank zur Entwicklung personalisierter Medizin sollen sich gesunde mit kranken Zellen vergleichen lassen. Möglich macht das die sogenannte Einzelzell-RNA-Sequenzierung. Mit dieser Methode lässt sich nachvollziehen, welche Gene für die Produktion einer Zelle eine Rolle spielen. Bei der Erzeugung eines Proteins für den Zellaufbau werden nur bestimmte Abschnitte der DNA eines Menschen abgelesen und in RNA übersetzt. Diese dient als Grundlage für die Proteinbiosynthese.

Für die Einzelzell-RNA-Sequenzierung sind enorm feine Messungen nötig, die häufig durch die genutzten Geräte, die Umwelt oder auch die Zellbiologie selbst gestört werden. Abweichungen in den Messungen entstehen beispielsweise, wenn die Temperatur des Messgeräts sich verändert hat oder die Verarbeitungszeit der Zellen schwankt. Dieser sogenannte Batch-Effekt lässt sich zwar durch verschiedene mathematische Modelle herausrechnen, dafür müssen Forscherinnen und Forscher allerdings bemessen können, wie groß der Effekt ist. [Fabian Theis](#) ist Professor für die [Mathematische Modellierung biologischer Systeme](#) an der TUM und Direktor des Institute of Computational Biology am Helmholtz Zentrum. Sein Team hat ein neues Maß namens kBET entwickelt, mit dem sich die Unterschiede zwischen den Experimenten bemessen lassen. Die Ergebnisse wurden in Nature Methods veröffentlicht.

### Ein Algorithmus, der Null-Messungen aufspürt

Auch sogenannte Null-Messungen stellen eine Herausforderung für die Einzelzellsequenzierung dar. „Wir sequenzieren hin und wieder eine Zelle und stellen fest, dass ein bestimmtes Gen in dieser Zelle überhaupt kein Signal von sich gibt. Dahinter kann sich ein biologischer oder ein technischer Grund verbergen: Entweder wird das Gen nicht abgelesen, weil es keine Rolle spielt, oder aber die Sequenz konnte aus technischen Gründen nicht erfasst werden“, sagt Fabian Theis.

Ob eine Null-Messung biologisch oder technisch begründet ist, kann nun ein Algorithmus ermitteln, den Theis' Gruppe entwickelt hat. Die in Nature Communications vorgestellte Software beruht auf einem neuen Wahrscheinlichkeitsmodell und vergleicht die ursprünglichen mit den rekonstruierten Daten. „Wir bauen aber keine Software, um Ergebnisse künstlich zu glätten“, sagt Theis. „Unser Ziel ist vor allem, Fehler ausfindig zu machen und zu korrigieren. Mit diesen möglichst korrekten Daten können wir dann in den Austausch mit unseren Kolleginnen und Kollegen weltweit gehen und unsere Ergebnisse mit ihnen vergleichen.“ Die Verlässlichkeit und Vergleichbarkeit der Daten ist von Bedeutung, wenn diese in den Human Cell Atlas eingespeist werden sollen. „Unser neuer

Algorithmus basiert als einer der ersten im Bereich der Einzelzell-Genomik auf Deep Learning und ist bisher der schnellste auf diesem Gebiet“, sagt Theis.

**Publikationen:**

Büttner, M.; Theis F. et al. (2019): A test metric for assessing single-cell RNA-seq batch correction. Nature Methods, [DOI: 10.1038/s41592-018-0254-1](https://doi.org/10.1038/s41592-018-0254-1)

Eraslan, G.; Simon, L.M.; Theis F. et al. (2019): Single cell RNA-seq denoising using a deep count autoencoder. Nature Communications, [DOI: 10.1038/s41467-018-07931-2](https://doi.org/10.1038/s41467-018-07931-2)