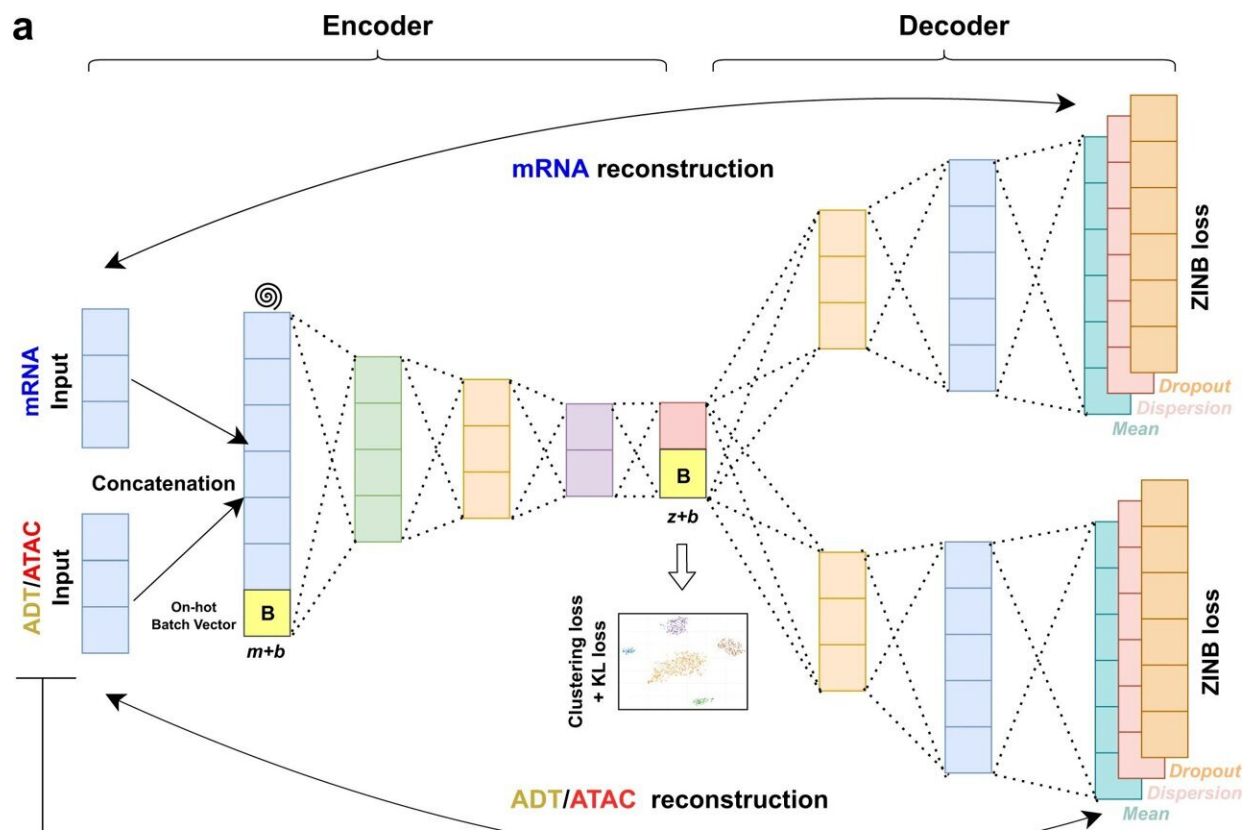


A new tool for studying multiple characteristics of a single cell

December 21 2022



The architecture of scMDC. scMDC has one encoder for the concatenated data and two decoders for each modal in the multimodal data (a). It can be used for clustering CITE-seq data and 10x Single-Cell Multiome ATAC + Gene Expression (SMAGE-seq) data. The spiral symbols indicate the artificial noises added to the data. For multi-batch datasets, scMDC will work in a conditional autoencoder manner. A one-hot batch vector **B** (in dimension b) will be concatenated to the input feature of the encoder (with raw feature dimension, m) and the decoders (with latent feature dimension, z). This is designed for batch

effect correction. scMDC learns a latent representation \mathbf{Z} (in dimension z) of data on which different modalities are integrated. A deep K-means algorithm and a KLD loss are implemented on \mathbf{Z} . Based on the clustering results, scMDC employs an ACE model to detect markers in different clusters (**b**). Then, pathway analyses can be conducted based on the gene ranks learned by ACE (**c**).
Credit: *Nature Communications* (2022). DOI: 10.1038/s41467-022-35031-9

Researchers from Children's Hospital of Philadelphia (CHOP) and New Jersey Institute of Technology (NJIT) developed new software that integrates a variety of information from a single cell, allowing researchers to see how one change in a cell can lead to several others and providing important clues for pinpointing the exact causes of genetic-based diseases.

The findings were published by *Nature Communications*.

Single-cell sequencing allows researchers to look at specific aspects of a cell to determine how it interacts with its microenvironment. This is particularly relevant in [cancer research](#) since it can be used to determine the effects of a mutation that may only affect a small portion of cells. At the single-cell level, researchers can study gene expression as well as messenger RNA, proteins and even organelles within the cells in much greater detail and resolution than before.

However, because each of the characteristics of a single cell has been studied individually, their connections with one another—for example, how a genetic variant might directly impact messenger RNA, protein synthesis or epigenetics—may not be apparent, even when comparing data generated from the same cell.

To address this statistical and computational dilemma, the researchers

developed an automated single-cell multimodal sequencing clustering [software tool](#) to profile what is happening within the cell across multiple [biological processes](#) simultaneously and better characterize relationships between changes in a cell.

"With this tool, we can better understand a single cell as an entity and not just as a fragmented unit," said Hakon Hakonarson, MD, Ph.D., director of the Center for Applied Genomics at CHOP and a senior author of the study. "This is a significant advancement and allows us to integrate and put all of this information into biological perspective, which is particularly important when considering information on different diseases."

The software, referred to as single-cell multimodal deep clustering (scMDC), uses machine learning to analyze data about different characteristics of a single cell. The researchers conducted extensive simulation and real-data experiments and found that scMDC outperformed existing single cell single-modal and multimodal clustering methods on [single-cell](#) multimodal data sets. It also utilizes linear scalability, meaning that more data sources provided to the scMDC yield better results.

More information: Xiang Lin et al, Clustering of single-cell multi-omics data with a multimodal deep learning method, *Nature Communications* (2022). [DOI: 10.1038/s41467-022-35031-9](https://doi.org/10.1038/s41467-022-35031-9)

Provided by Children's Hospital of Philadelphia

Citation: A new tool for studying multiple characteristics of a single cell (2022, December 21) retrieved 27 April 2024 from <https://phys.org/news/2022-12-tool-multiple-characteristics-cell.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.