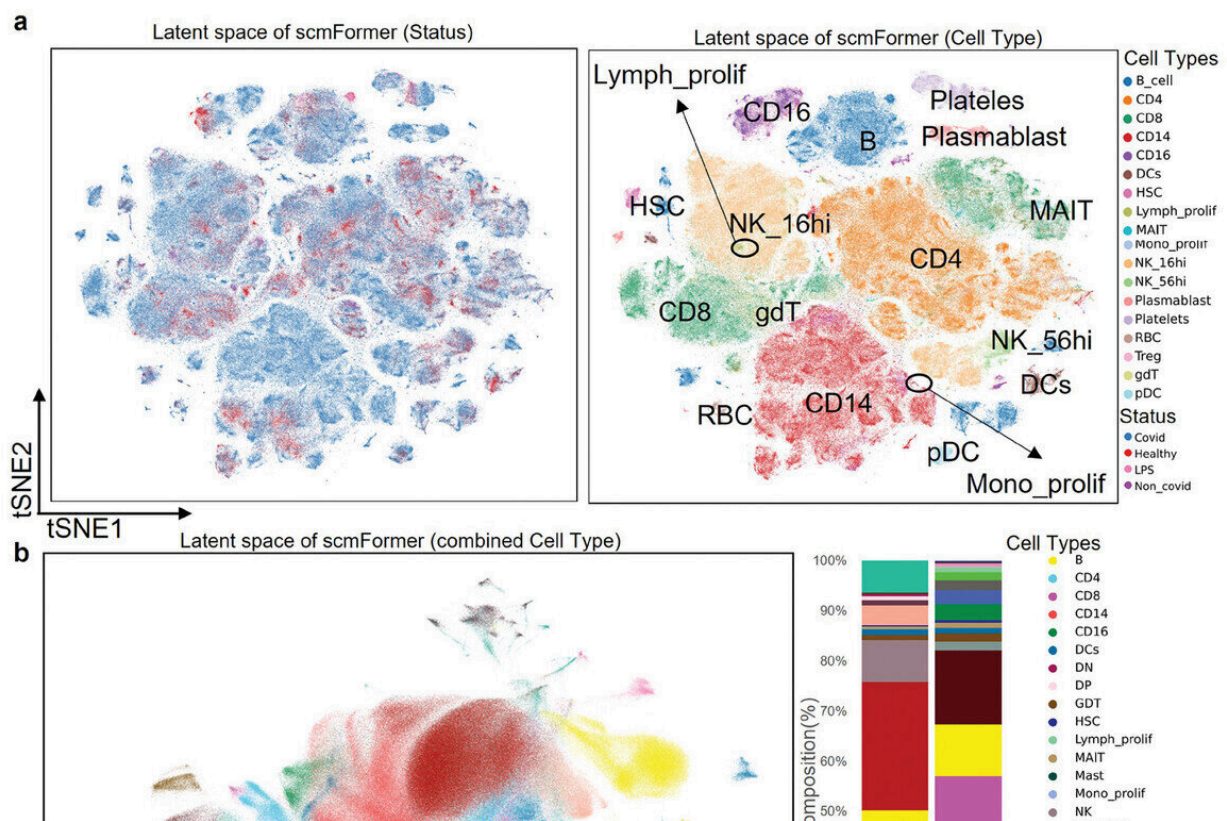


New multi-task deep learning framework integrates large-scale single-cell proteomics and transcriptomics data

April 26 2024, by Zhang Nannan



Integration of COVID-19 cell atlas. Credit: *Advanced Science* (2024). DOI: 10.1002/advs.202307835

The exponential progress in single-cell multi-omics technologies has led

to the accumulation of large and diverse multi-omics datasets. However, the integration of single-cell proteomics and transcriptomics (or epigenomics) data poses a significant challenge to existing methods. Several transformer-based models, such as Geneformer, have significantly changed the paradigm of single-cell transcriptome analysis. However, these methods place significant demands on computational resources.

To address these [challenges](#), researchers at the Wuhan Botanical Garden of the Chinese Academy of Sciences have developed a Transformer-based method, called scmFormer, to integrate large-scale single-cell proteomics and transcriptomics data using a multi-task transformer. The study titled "scmFormer Integrates Large-Scale Single-Cell Proteomics and Transcriptomics Data by Multi-Task Transformer" was [published](#) in *Advanced Science*.

The researchers presented a comprehensive evaluation and made case studies of this method, the results showed that scmFormer exhibited remarkable proficiency in harmonizing large-scale single-cell omics plus proteomics datasets at both the cell type and finer-scale cell level with limited computer resources.

In addition, scmFormer possesses the ability to integrate multiple single-cell paired multimodal datasets, leading to the dual benefit of reduced high cost and improved biological insights.

Moreover, scmFormer shows an outstanding ability to eliminate technical differences between different omics modalities while preserving the underlying biological information inherent in the data, spanning both [cell types](#) and experimental conditions.

The application of scmFormer for the integration of two COVID-19 datasets with 1.48 million cells further demonstrated the distinct

advantage of scmFormer for handling large datasets on regular laptops.

More information: Jing Xu et al, scmFormer Integrates Large-Scale Single-Cell Proteomics and Transcriptomics Data by Multi-Task Transformer, *Advanced Science* (2024). [DOI: 10.1002/adv.202307835](https://doi.org/10.1002/adv.202307835)

Provided by Chinese Academy of Sciences

Citation: New multi-task deep learning framework integrates large-scale single-cell proteomics and transcriptomics data (2024, April 26) retrieved 20 June 2024 from <https://phys.org/news/2024-04-multi-task-deep-framework-large.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.