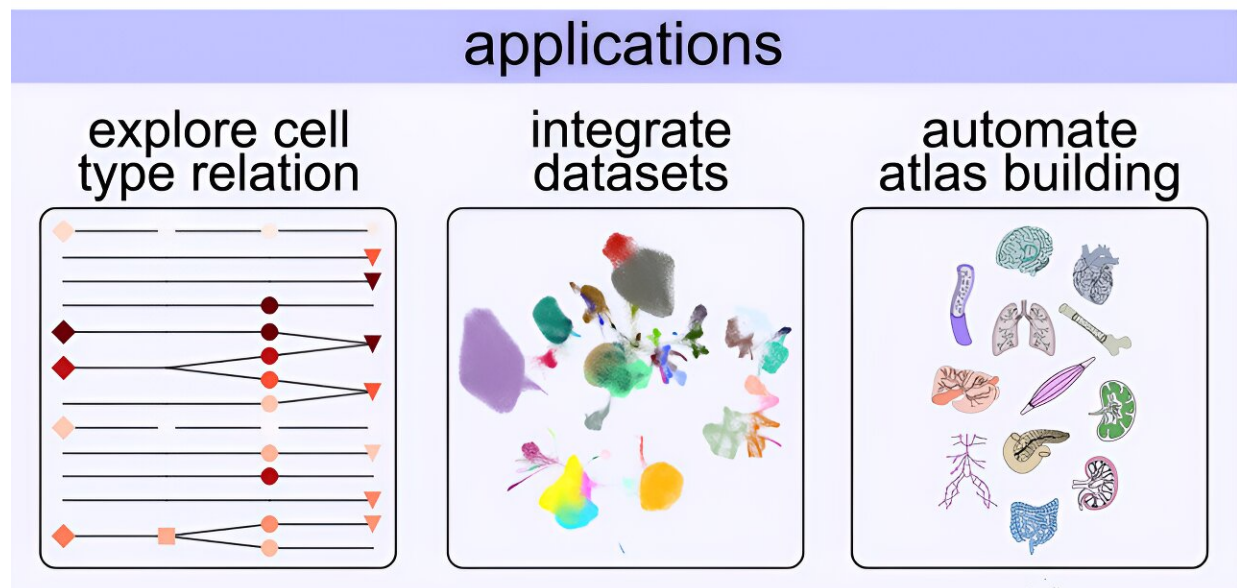
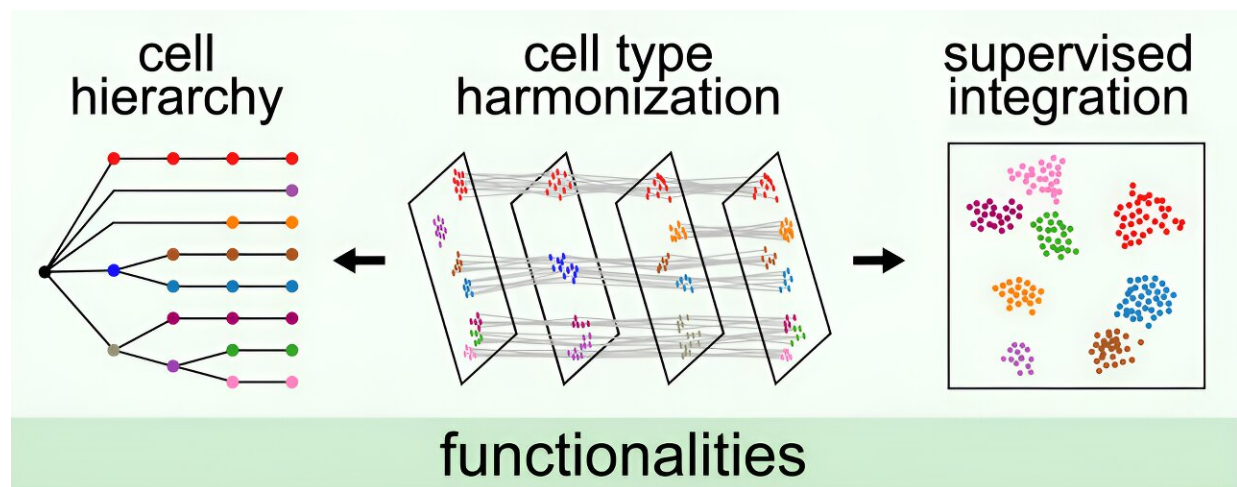


New tool unifies single-cell data for the global research community

December 21 2023, by Rachael Smith



Graphical Abstract. Credit: *Cell* (2023). DOI: 10.1016/j.cell.2023.11.026

A new methodology that allows for the categorization and organization of single-cell data has been launched. It can be used to create a harmonized dataset for the study of human health and disease.

Researchers at the Wellcome Sanger Institute, the University of Cambridge, EMBL's European Bioinformatics Institute (EMBL-EBI), and collaborators developed the tool, known as CellHint. CellHint uses machine learning to unify data produced across the world, allowing it to be accessed by the wider research community, potentially driving [new discoveries](#).

In a new study, [published](#) today in *Cell*, researchers applied CellHint to reveal underexplored connections between healthy and diseased lung cell states. They looked at eight diseases, such as [interstitial lung disease](#) and chronic obstructive pulmonary lung disease, and showed the possible benefits of this tool. They also applied CellHint to 12 tissues from 38 datasets, providing a deeply curated cross-tissue database with around 3.7 million [cells](#).

Cellhint is freely available worldwide and was created as part of the Human Cell Atlas initiative, which aims to map every cell type in the [human body](#) to transform understanding of health and disease.

Single-cell genomics enables the understanding of every cell in the context of the human body at high resolution. Currently, a challenge in assembling the diverse datasets produced by single-cell research is that there is no unified system for naming and organizing data.

To address this, researchers from the Wellcome Sanger Institute, and

collaborators developed CellHint, which can unify [cell types](#) produced by independent laboratories. CellHint then places the data into a defined graph that shows the relationships between cell subtypes, giving a full picture of all the cells identified across different datasets.

The team applied CellHint to current data and revealed underexplored relationships between healthy and diseased lung cell states in eight diseases. It also identified cell types in adult human hippocampus that could be of potential interest for future research.

The researchers also applied CellHint to 12 tissues from 38 datasets, providing a deeply curated cross-tissue database with around 3.7 million cells. Each cell was annotated, which is the process of labeling cells with particular information. They also showed how it can create various models for automatic cell annotation across human tissues.

Dr. Chuan Xu, first author from the Wellcome Sanger Institute, said, "CellHint stands out from other tools because it makes full use of the often inconsistent but valuable cell annotation information from individual studies, to achieve biologically-driven data integration."

"We are excited that with CellHint, cells from independent laboratories can be re-annotated and researchers can utilize the resulting information to put each cell into different contexts beyond the original study. We hope that this tool will greatly facilitate the reuse of molecular and cellular data and information across laboratories, potentially driving new discoveries in biology."

Dr. Sarah Teichmann, senior author from the Wellcome Sanger Institute and co-founder of the Human Cell Atlas, said, "The Human Cell Atlas is creating detailed reference maps of all cells in the human body to transform our understanding of biology, health and disease, and single-cell technologies underpin this hugely ambitious project. Global

collaboration and open data sharing are vital to achieve the aim of a representative Human Cell Atlas that will benefit humanity worldwide."

"CellHint enables the unification and sharing of single-cell data, which allows the global research community to contribute to and benefit from the ongoing research that is happening around the world, and help drive advances in health and [health](#) care."

More information: Chuan Xu et al, Automatic cell-type harmonization and integration across Human Cell Atlas datasets, *Cell* (2023). [DOI: 10.1016/j.cell.2023.11.026](https://doi.org/10.1016/j.cell.2023.11.026)

CellHint can be found at <https://github.com/Teichlab/cellhint>

Provided by Wellcome Trust Sanger Institute

Citation: New tool unifies single-cell data for the global research community (2023, December 21) retrieved 9 May 2024 from <https://medicalxpress.com/news/2023-12-tool-single-cell-global-community.html>

<p>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.</p>
--