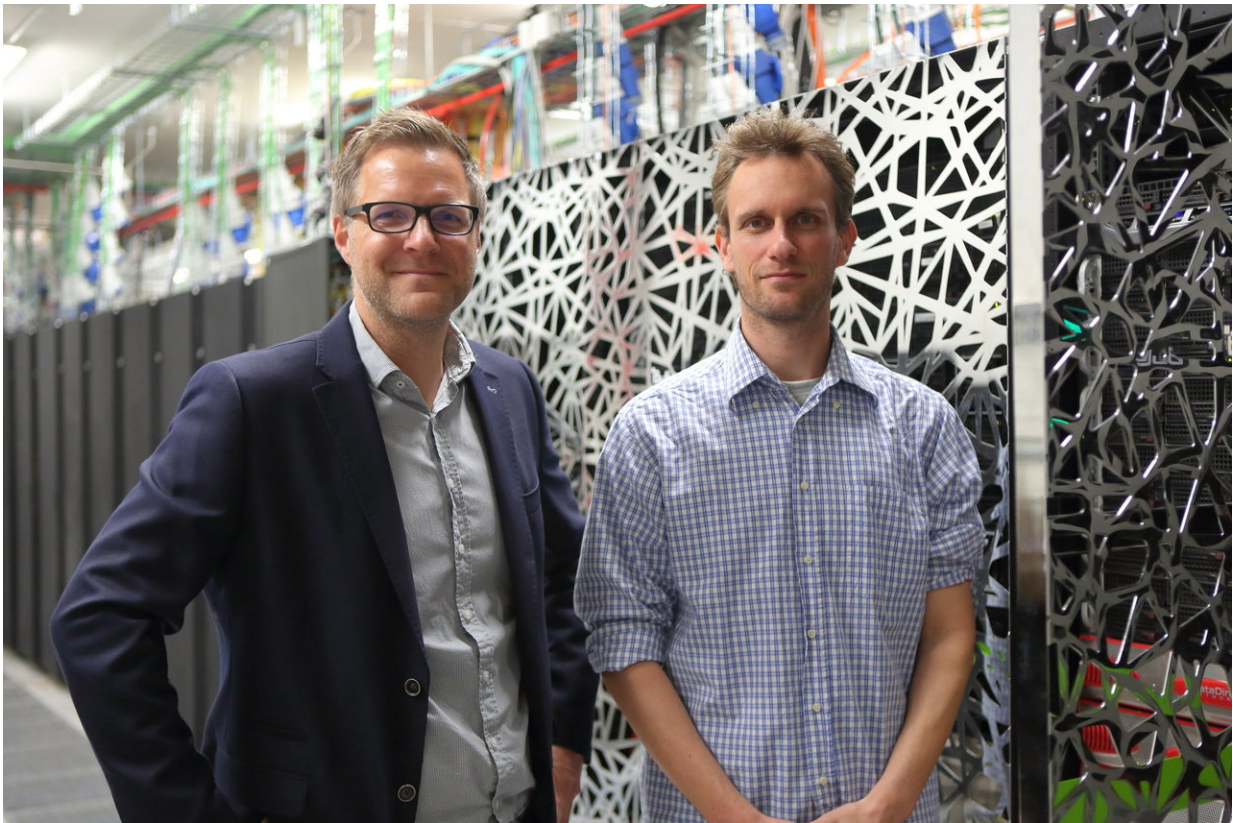


New tool enables large-scale analysis of single cells

June 6 2018



Holger Heyn and Giovanni Iacono developed BigScale, a new computational tool that enables big-scale analysis of single cells. Credit: CNAG-CRG

New research led by Holger Heyn at the Centro Nacional de Análisis Genómico of the Centre for Genomic Regulation (CNAG-CRG),

presents a sophisticated computational framework to analyze single-cell gene expression levels, scalable to process millions of individual cells. The work, published in the current issue of the scientific journal *Genome Research*, reports the first tool capable of analyzing such large single-cell RNA-seq dataset. This dramatically extends the limits of single-cell genome research.

All the cells of a single organism share the same genome, but each cell has the potential to differentiate into any tissue or organ via [gene expression](#). Scientists worldwide are looking at what differentiates one cell from another. One of the current challenges of [genome research](#) is to analyze many [individual cells](#) in order to find and to identify those differences. Looking at gene expression of one cell at a time, scientists are now able to describe a sample's heterogeneity at unprecedented resolution and without prior knowledge of its composition.

Accordingly, large-scale single-cell projects led to the identification of previously unknown cell types and to the drawing of comprehensive cellular maps of organisms. Within the framework of the Human Cell Atlas project, researchers aim at creating an atlas of all cell types that make up a human body. However, such studies create massive amounts of sequencing data and analyzing large datasets is a major challenge.

A group of scientists of the CNAG-CRG, in collaboration with researchers from the University Pompeu Fabra (UPF) and the Spanish Biomedical Research Consortium on Rare Diseases (CIBERER), has now developed an efficient computational framework that enables processing, analysis and interpretation of such large-scale single-cell experiments. The group illustrated the power of their strategy by analyzing one of the largest single-cell studies encompassing the 1.3 million individual cells of the developing mouse brain.

"BigScale is extremely powerful in identifying cell type-specific genes,

which greatly helps in the downstream interpretation of experiments," says Holger Heyn, CNAG-CRG team leader and senior author of the study. The novelty of the analytic tool BigScale lies in a numerical model that sensitively determines differences between single cells. Having charted how individual cells differ from each other, they can be grouped together into populations of cells to describe the cellular complexity of a given tissue. As virtually all tissues are composed of different cell types and subtypes, such an analysis can guide an unbiased, in-depth characterization without initial hypotheses. Differentially expressed marker genes between subpopulation help the researcher to link cells to prior knowledge about the tissue anatomy or to describe the functions of newly discovered cell types.

In addition, the tool was designed to tackle future challenges of large datasets. "The costs to derive single-cell profiles are decreasing and we are seeing studies of increasing cell numbers," adds Dr. Heyn. In this regard, a module in the BigScale workflow enables the analysis of millions of cells through a directed convolution strategy. Here, single-cell transcriptomes from similar cells are merged into index cells, which greatly reduces the amount of data to be processed.

With the new tool in hand, the group analyzed a dataset of 1.3 million cells, a resource publicly available from 10x Genomics. "BigScale allowed us to look deep into the developmental processes of the mouse brain and to characterize even rare neuronal [cell types](#)," said Giovanni Iacono, the first author of the work. Specifically, the high number of cells enabled the group to zoom into a small transient cell population of Cajal-Retzius [cells](#) and to describe major substructures related to distinct differentiation stages, spatial organization and cellular function. "The BigScale framework provides a powerful solution for virtually any species, and is even applicable outside the RNA sequencing context," explains Dr. Heyn. "We expect it to contribute to the interpretation of large-scale studies, such as the Human Cell Atlas project."

More information: Giovanni Iacono et al, bigSCale: an analytical framework for big-scale single-cell data, *Genome Research* (2018). [DOI: 10.1101/gr.230771.117](https://doi.org/10.1101/gr.230771.117)

Provided by Center for Genomic Regulation

Citation: New tool enables large-scale analysis of single cells (2018, June 6) retrieved 24 April 2024 from <https://phys.org/news/2018-06-tool-enables-large-scale-analysis-cells.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.