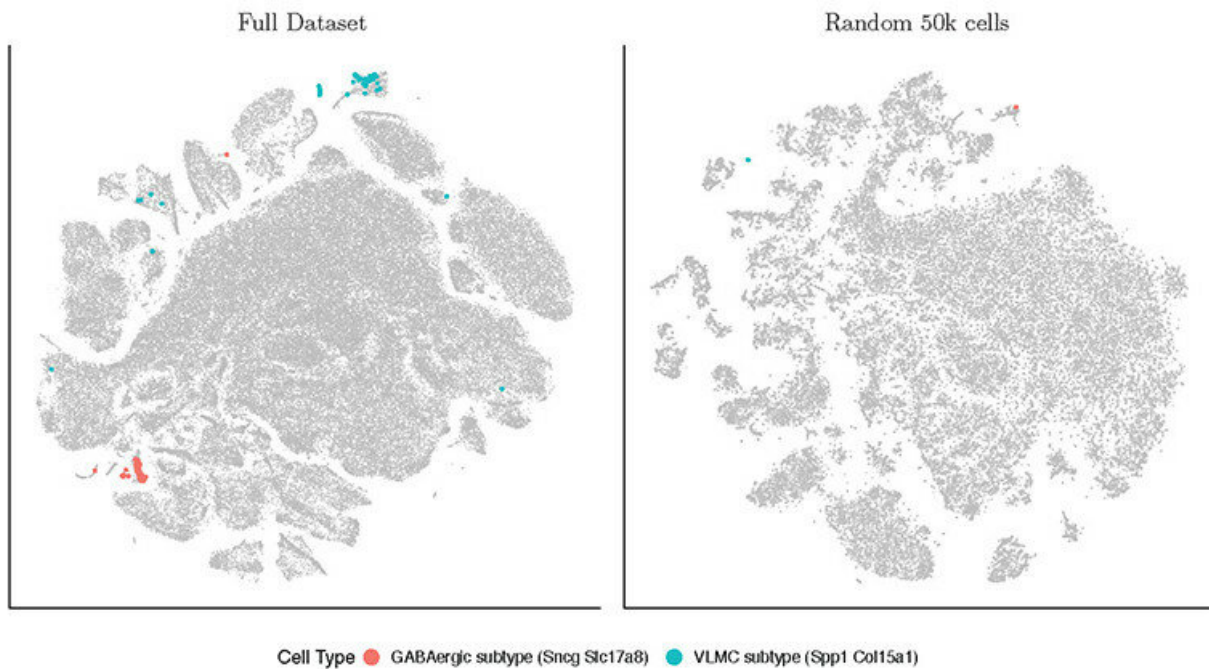


Improved RNA data visualization method gets to the bigger picture faster

February 14 2019, by Kendall Teare



A full 1.3 million mouse brain cell dataset using the new FI-t-SNE (left) visualization method versus a downsampling to a random 50,000 cells with the old t-SNE technique (right). Credit: Yale University

Like going from a pinhole camera to a Polaroid, a significant mathematical update to the formula for a popular bioinformatics data visualization method will allow researchers to develop snapshots of

single-cell gene expression not only several times faster but also at much higher-resolution. Published in *Nature Methods*, this innovation by Yale mathematicians will reduce the rendering time of a million-point single-cell RNA-sequencing (scRNA-seq) data set from over three hours down to just fifteen minutes.

Scientists say the existing decade-old method, t-distributed Stochastic Neighborhood Embedding (t-SNE), is great for representing patterns in RNA sequencing data gathered at the single cell level, scRNA-seq data, in two dimensions. "In this setting, t-SNE 'organizes' the cells by the genes they express and has been used to discover new cell types and cell states," said George Linderman, lead author and a Yale M.D.-Ph.D. student specializing in applied mathematics.

By computational standards, t-SNE is quite slow. Thus, researchers often "downsample" their scRNA-seq dataset—take a smaller sample from the initial sample—before applying t-SNE. However, downsampling is a poor compromise, as it makes it unlikely for t-SNE to capture rare cell populations, which are often what researchers most want to identify.

More than 30 years ago, another team of Yale mathematicians developed the fast multipole method (FMM), a revolutionary numerical technique that sped up the calculation of long-ranged forces in the n-body problem. The researchers on this study recognized that the principles behind the FMM could also be applied to nonlinear dimensional reduction problems, such as t-SNE, and accelerated t-SNE until it earned its new name: FIt-SNE, or fast interpolation-based t-SNE.

"Using our approach, researchers can not only analyze single cell RNA-sequencing data faster, but it also can be used to characterize rare cell subpopulations that cannot be detected if the data is subsampled prior to t-SNE," said Yuval Kluger, senior author and Yale professor of pathology. Additionally, the team used a heatmap-style visualization for

its FIt-SNE results, which makes it easy for researchers to see the expression patterns of thousands of genes at the level of single [cells](#) simultaneously.

The researchers said 2019 couldn't be a better new year for t-SNE to get "FIt." In December 2018, Science Magazine named tracking development of embryos cell by cell—impossible to accomplish without visualizations based on scRNA-seq data—the Breakthrough of the Year. FIt-SNE will speed up further work in this field of developmental biology as well as in fields such as neuroscience and [cancer research](#), where single-cell sequencing has become an invaluable tool for mapping the brain and understanding tumors, said the researchers.

More information: George C. Linderman et al, Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data, *Nature Methods* (2019). [DOI: 10.1038/s41592-018-0308-4](https://doi.org/10.1038/s41592-018-0308-4)

Software for FIt-SNE and the heatmap-style visualization is available at the following links:

[Fast Fourier Transform-accelerated Interpolation-based t-SNE \(FIt-SNE\)](#)

[Beta version of 1D t-SNE heatmaps to visualize expression patterns of hundreds of genes simultaneously in scRNA-seq](#)

Provided by Yale University

Citation: Improved RNA data visualization method gets to the bigger picture faster (2019, February 14) retrieved 18 April 2024 from <https://phys.org/news/2019-02-rna-visualization-method-bigger-picture.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.