

New technique improves efficiency and accuracy of single cell RNA sequencing

November 21 2018



A hairpin loop from a pre-mRNA. Highlighted are the nucleobases (green) and the ribose-phosphate backbone (blue). Note that this is a single strand of RNA that folds back upon itself. Credit: Vossman/ Wikipedia

In the era of personalized medicine, scientists are using new genetic and genomic insights to help them determine the best treatment for a given patient. In the case of cancer, the first step toward these treatments is an

investigation into how tumor cells behave in an effort to figure out the best drugs to use to attack them.

Researchers then use DNA- and RNA- sequencing to look at populations of [cells](#), examining which genes are expressed within a sample of cancerous tissue. However, traditional sequencing methods can hide that fact that not all [tumor cells](#) necessarily behave in the same way. Not recognizing this means that if you target a tumor with a specific type of drug, some cells may be just different enough to survive and thrive.

In a major advance for genomics, it is now possible to look at what one single cell is doing at any given time with a technique called single-cell RNA sequencing (scRNA-seq). This method looks at the amount of messenger RNAs (mRNAs) in a cell and compares those to other cells to look for differences in gene expression.

However, what information you find can depend on how you run your experiment and how the data are analyzed. Lana Garmire, Ph.D., associate professor of the department of computational medicine & bioinformatics at Michigan Medicine and her team is studying ways to eliminate some of the biases that can make interpreting scRNA-seq data difficult.

"A lot of the noise in this type of sequencing comes from the fact that you have to measure samples in extreme low quantities and in different batches," she explains. For example, the tissue sample a [researcher](#) is analyzing may not fit on one plate, a piece of equipment used to house cell samples, and therefore have to be split onto two plates. Differences that arise due to this split are called batch effects. Genomics researchers must correct for these batch effects, but this process can raise a conundrum: how do you know if a difference is a batch effect or a true difference between cells?

New uses for data

Bioinformatics is the term for collecting and analyzing complex biological data using computer programs. It is a relatively new field born out of the ability to gather enormous amounts of biological data, such as DNA and [protein sequences](#).

Researchers rely on bioinformatics techniques to determine which genes are expressed in single cells. But they've had to work around the noise introduced through different research protocols and batch effects.

Garmire, who recently joined U-M from the University of Hawaii and is the new faculty director University of Michigan Medical School Bioinformatics Core, has discovered a more efficient way of identifying differences between cells using the same set of data produced during sequencing experiments. Instead of relying on gene expression, she found that looking at what are known as single nucleotide variants (SNVs) can eliminate some of this uncertainty. "With SNVs, you are dealing with numbers that are binary, 0 and 1. Either the mutation is there or not."

Recall that genes are made up of nucleotides represented by the letters A, T, G and C that make up a code that is translated into a protein. Garmire's method looks for differences in single nucleotides, knowing that an A can only be replaced by a T and a G by a C. This new work, described in *Nature Communications*, developed a new set of procedures to process scRNA-seq data and retrieve this variant information. Further, using a computer program called SSrGE, they can link this variant information to more traditional gene expression information.

"This gives us information on different subpopulations of tumor cells and becomes sort of like a fingerprint that can be marked to identify cell-to-cell differences," says Garmire.

What it all means

Ultimately, drug makers and clinicians use these targets to guide pharmaceutical treatments. "When you want to attack the issue, you go at it by attacking the fundamental features of that issue: the mutations. Clinicians may be able to use this information later on to guide their therapeutics." Garmire looks forward to bringing bioinformatics out of the lab, helping researchers who amass large amounts of data to use them and develop downstream clinical applications. "We divide the body up and specialize but at the end of the day, you need to look holistically and ask, what am I doing and who is this helping? We are developing computational tools to bring [bioinformatics](#) researchers and bench scientists and clinicians together to connect the dots and ultimately make change.

More information: Olivier Poirion et al, Using single nucleotide variations in single-cell RNA-seq to identify subpopulations and genotype-phenotype linkage, *Nature Communications* (2018). [DOI: 10.1038/s41467-018-07170-5](#)

Provided by University of Michigan

Citation: New technique improves efficiency and accuracy of single cell RNA sequencing (2018, November 21) retrieved 25 April 2024 from <https://phys.org/news/2018-11-technique-efficiency-accuracy-cell-rna.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>
