Single-cell sequencing

A brief overview of how to derive a genome or transcriptome from a single cell.

Genome and transcriptome sequencing require orders of magnitude more starting material than what is found in an individual cell, pushing the limits of amplification technology. Handling such small quantities means that degradation, sample loss and contamination can have a pronounced effect on sequence quality and robustness. Heavy amplification also propagates errors and biases, which can lead to uneven coverage, noise and inaccurate quantification.

Recent technical advances have helped mitigate these challenges, making single-cell sequencing an appealing way to address an expanding set of problems. Rare cell types, heterogeneous samples, phenotypes associated with mosaicism or variability, and microbes that cannot be cultured are good candidates for single-cell approaches. Single-cell sequencing can enable the discovery of clonal mutations, cryptic cell types or transcriptional features that would be diluted or averaged out in bulk tissue.

Picking the right cell

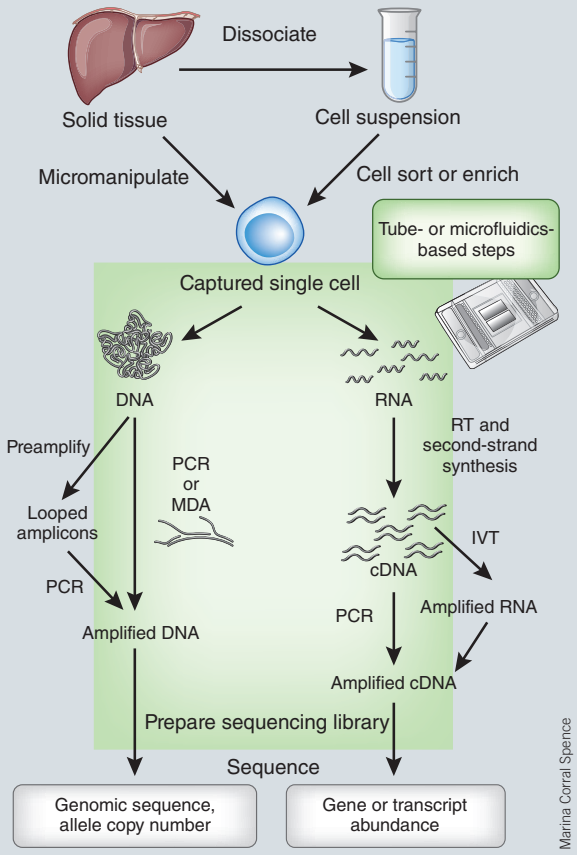
Micromanipulation is a precise but laborious way to target a single cell, and microcapillaries can be used to extract a cell's contents directly. Many tissues can be dissociated to produce cell suspensions, which are easier to handle and allow cells expressing specific markers to be enriched with a cell sorter. Instruments that trap very rare cells on the basis of their surface markers are also being used to isolate tumor cells that circulate in blood.

Single transcriptomes

Many single-cell RNA sequencing protocols are now available, though every variation begins with the conversion of RNA to the first strand of complementary DNA (cDNA) by reverse transcriptase. Some methods sequence full transcripts and others sequence tags only at the 5′ or 3′ end.

The common goal of these methods is to capture the original RNA population and amplify it evenly and accurately. Capture efficiency is influenced by how comprehensively reverse transcriptase samples RNA from the cell—a stochastic process that can be improved using small reaction volumes and, potentially, better enzymes. In addition, a technique known as template switching may ensure that a greater proportion of captured transcripts are full length.

Amplification can also be improved by minimizing the number of cycles, inhibiting primer byproduct amplification by 'suppression PCR', and pooling barcoded cDNA from different samples to provide enough starting material for linear amplification by *in vitro* transcription. Unique molecular identifier sequences can also be used to label individual RNA molecules, allowing the absolute number of original molecules to be counted directly even after subsequent (uneven) amplification.



Marina Corral Spence

Workflows for amplifying and sequencing the RNA or genomic DNA of a cell. MDA, multiple displacement amplification; RT, reverse transcription; IVT, *in vitro* transcription.

Single genomes

Whole-genome amplification starts with a tiny amount of material: just a single molecule of DNA. This makes uneven coverage and preferential amplification of one allele—or loss, known as allelic dropout—a widespread problem. The most common approach is multiple displacement amplification, which uses random primers that bind throughout the genome and a polymerase that can displace other fragments on the same template that it copies, forming iterative branching structures that massively amplify DNA. Early cycles have a strong effect on the uniformity of amplification. One variation uses special primers that cause amplicons to form closed loops and prevent further copying, allowing a few cycles of linear amplification prior to PCR. Scaling up and monitoring reactions in real time can help to overcome low success rates in genome amplification, and lower-input sequencing library preparations that rely on less amplification are another promising direction.

One cell for all

Scaling up is important to ensure that biological variability is well sampled in single-cell studies. Microfluidics or microwell technologies provide higher throughput and standardized handling, and are often efficient because reactions are concentrated in small volumes; however, microfluidics can be restricted to certain cell size ranges. Barcoding and pooling will also help to increase throughput.

Technologies for single-cell amplification and sequencing are maturing. As the cost and ease of examining individual cells improves, the approach will enter the hands of more researchers as a standard tool for understanding biology at high resolution.

Tal Nawy

ACKNOWLEDGMENTS

We thank J. Levin and K. Zhang for help preparing this Primer.