BIOINFORMATICS

Finding a face in the crowd

An analytical algorithm for transcriptomic data reveals rare cell types that were previously hidden in larger populations.

Any given organ or tumor sample is likely to comprise a host of different cell types. Detailed profiles of mRNA content make it possible for scientists to broadly classify individual cells, but capturing the full diversity of heterogeneous cellular populations remains challenging.

An algorithmic approach devised by Alexander van Oudenaarden and colleagues at the Hubrecht Institute in the Netherlands now makes it possible to thoroughly classify such communities in a manner that allows for the ready identification of rare cell types. Starting from a collection of single-cell transcriptomic data, their RaceID algorithm uses a clustering approach to broadly classify similar cell populations. These populations are then analyzed more stringently to identify 'outliers' that are statistically likely to represent distinct subtypes within these groups, flagging rare cell types that might otherwise be overlooked.

As an initial test, van Oudenaarden and colleagues analyzed the composition of cultured 'organoids' derived from the mouse intestinal wall. In one set of experiments, they dissected the subpopulations of hormone-secreting enteroendocrine cells, which could be broadly identified on the basis of expression of the gene *Reg4*. In addition to distinguishing the subset of serotonin-releasing enterochromaffin cells, RaceID revealed the existence of three previously undiscovered enterochromaffin subtypes. Each of these populations expressed distinct genetic markers, and the researchers subsequently confirmed that they could be readily distinguished via fluorescent labeling in actual tissue samples.

This approach also made it possible to characterize the extent to which the intestinal stem cell pool represents a truly homogeneous population. By looking at the previously identified gene marker *Lgr5*, the researchers identified a pool of cells that was almost entirely composed of stem cells, but with a small subset of enteroendocrine cells and Paneth cells that could be readily distinguished on the basis of their mRNA profiles. These results are in keeping with a model in which Paneth cells retain the potential to revert back into stem cells to repair damage to the intestinal wall.

The combination of RaceID with other labeling and lineage-tracing methods could thus be a useful tool for understanding processes underlying both organ development and disease etiology at the single-cell scale.

Michael Eisenstein

RESEARCH PAPERS

Grün, D. *et al.* Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* doi:10.1038/nature14966 (19 August 2015).