A comparison of ambient RNA and gene detection between split pool combinatorial barcoding and droplet-based technologies

Tammy Lyons (tammy@parsebiosciences.com), Parse Biosciences, **Ajay Sapre**, Parse Biosciences, **Simone Marrujo**, Parse Biosciences, **Alec Salvino**, Parse Biosciences, **Charlie Roco**, Parse Biosciences, **Alex Rosenberg**, Parse Biosciences, **Anastasia Potts**, Parse Biosciences

Cross-contamination with ambient RNA is a common issue in single cell RNA sequencing platforms when floating RNA is incorporated in the droplets and barcoded with the cell's own RNA, leading to poor clustering resolution and data misinterpretation.

Parse Evercode is a hardware-free technology that uses the cell as a reaction vessel, making it less susceptible to ambient RNA.

In a head-to-head evaluation, we compared the Chromium droplet-based microfluidics approach and Parse combinatorial barcoding. We chose the mouse kidney, an organ containing multiple cell types, for an in-depth evaluation of cell type resolution and ambient RNA contamination. Both libraries were sequenced on Illumina Novaseq™ 6000 by the same 3rd party provider. Both datasets were downsampled to 9,256 cells, 18,898 mean reads/cells, and independently analyzed.

Evercode detected 79% more genes than Chromium. Annotation of the two datasets showed a large cluster of seemingly RBCs in the Chromium data, and hemoglobin RNA was found across all clusters, absent in the Evercode data.

To closely examine subcellular resolution, minor tubules cells were subclustered. Evercode data showed a clear separation between the clusters and canonical cell type markers - including resolving two types of collecting duct cells, while there is a single cluster in the Chromium data, further illustrating the negative impact of ambient RNA.

Overall, Evercode WT detected more genes per cell, had substantially less ambient RNA contamination, and demonstrated higher cluster resolution than Chromium Next GEM 3' v3.1.