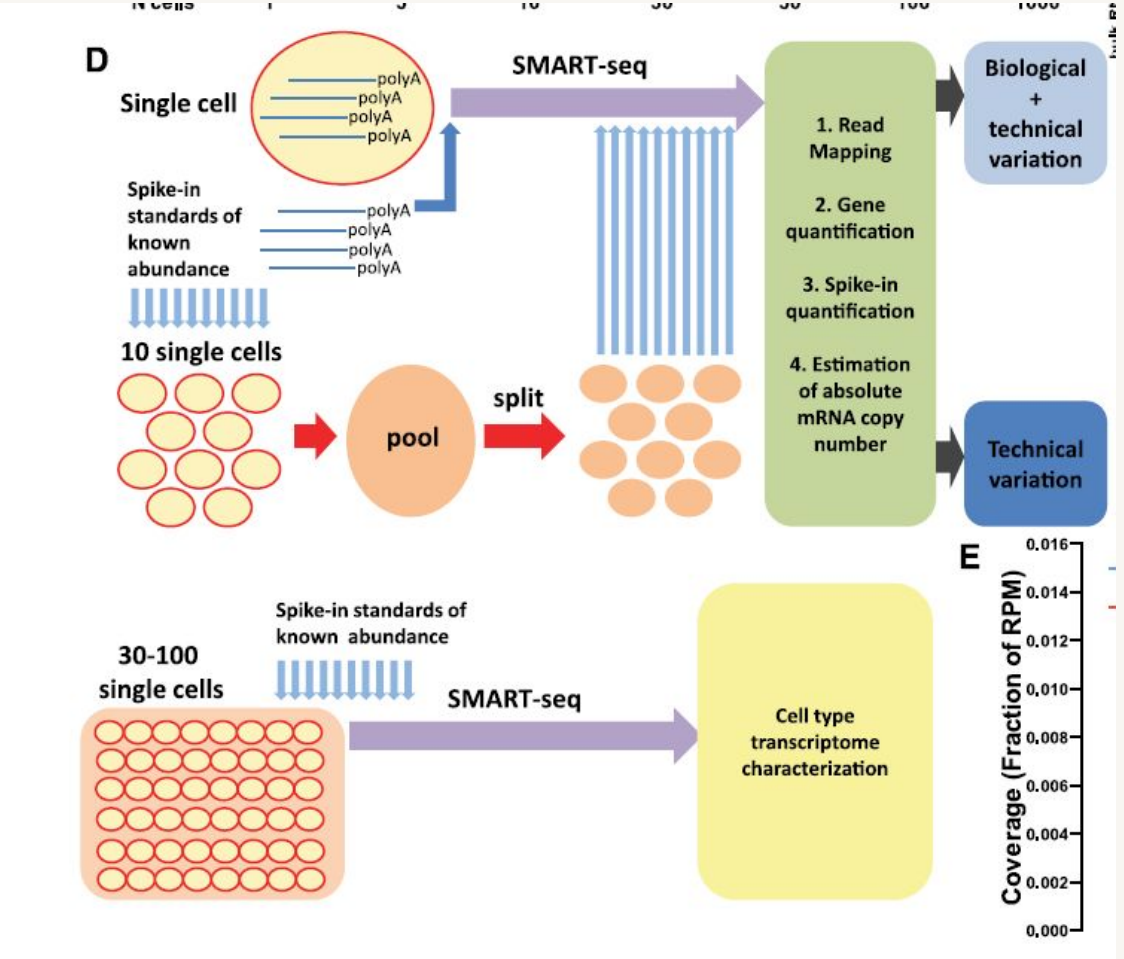
“From single-cell to cell-pool transcriptomes” (Marinov et al., 2017)

Presented by Seth D Temple
UW Statistical Genetics Seminar
February 15, 2021

Two major challenges

- “[E]fficiency and uniformity with which each mRNA is copied into cDNA and ultimately represented in the library”
  - Single-molecule capture efficiency
  - scRNA-seq protocols biased towards 3’ ends
- Transcriptomic characterization of rare cells
- ***Conduct lab experiments to inform an experimental design (cell pools) to address these***

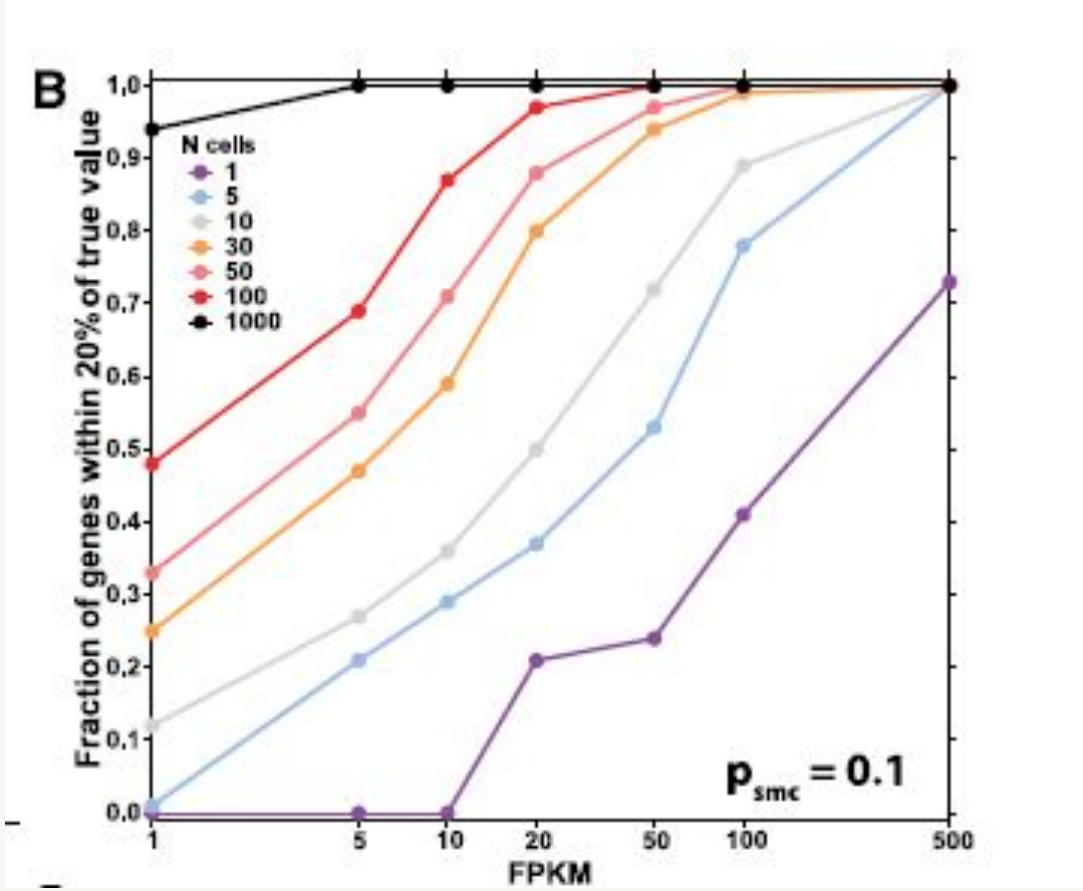


Major factors informing sc-RNAseq

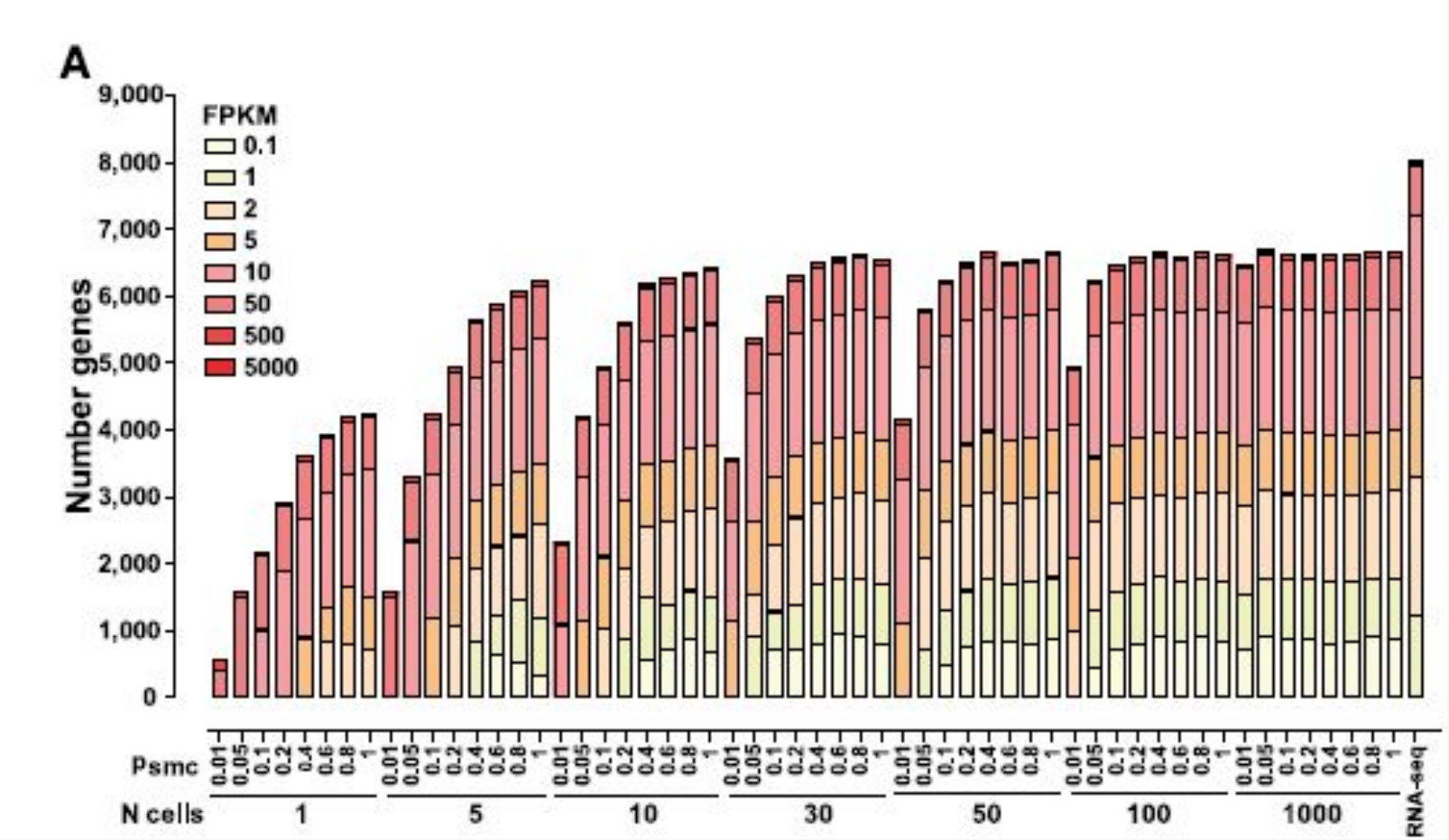
1. Single-molecule capture efficiency
   a. This is technical variation!
   b. In practice, about 10% efficient
   c. What impacts this? (Taylor)
   d. Are there other sources of technical variation? (Nick)

2. Total mRNA molecules per cell
   a. Depends on cell type, size, state
3. Expression of genes in cells

*** FPKM is a measure of relative mRNA abundance

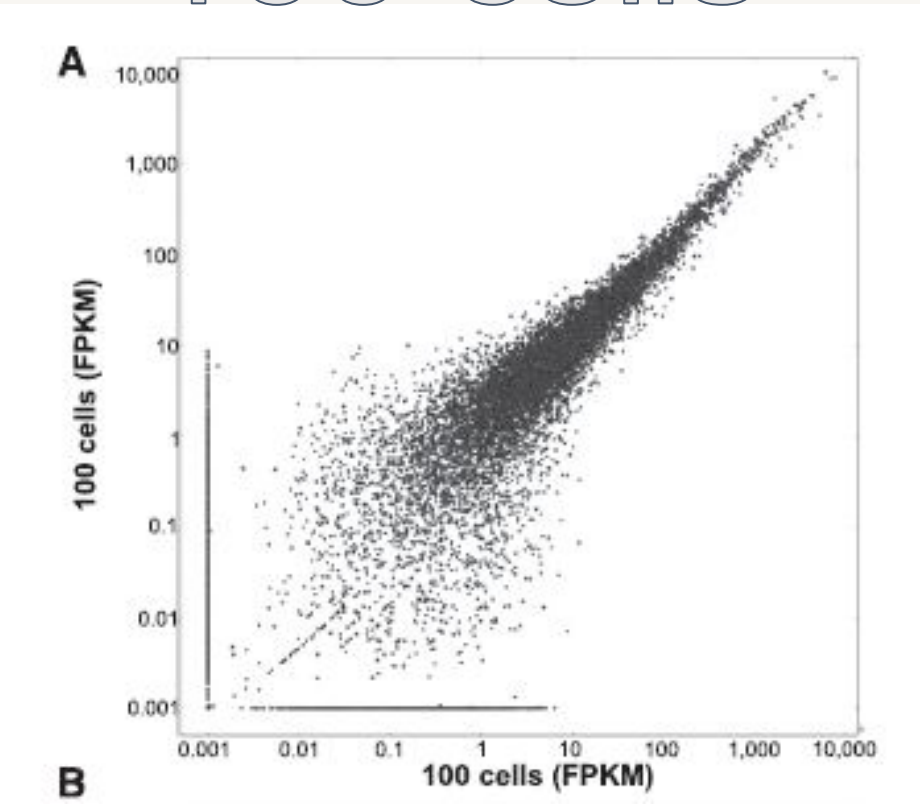


Why 30 cells is enough

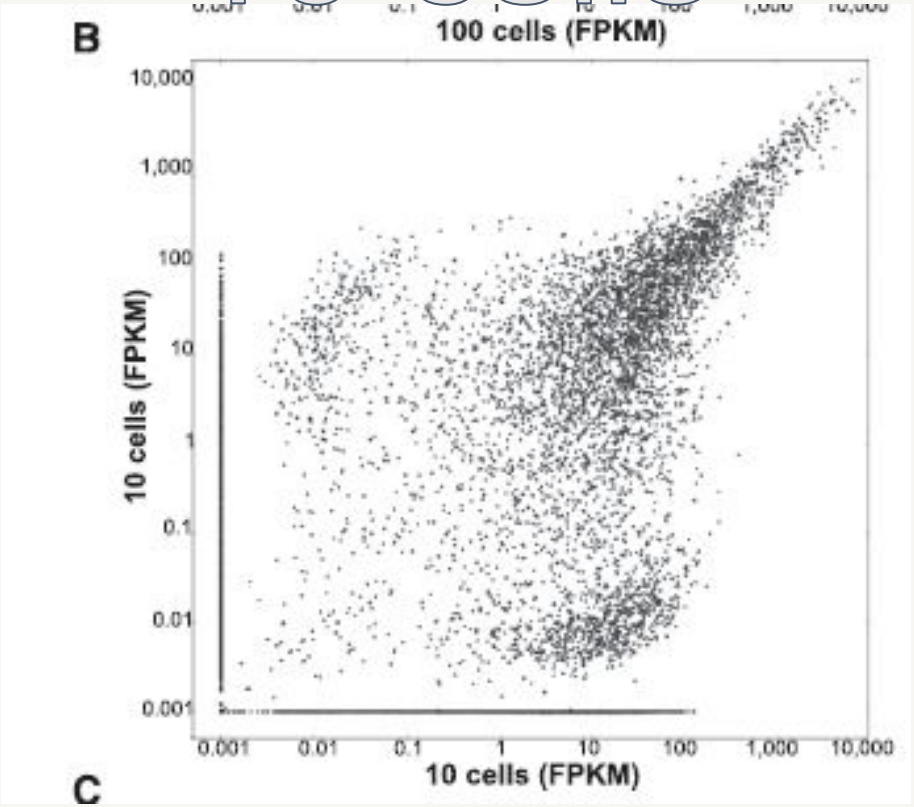


More cells, more correlated expression

100 cells



10 cells



Questions (experiments)

- How do the cell pool methods differ from the pool/split methods?
  - This is fundamental to understanding this paper!

- Why did they define “within 20% of the true value” for accuracy?
  - The true value ought to have standard deviation that depends on genes and cells.

- Why is variation in pool/split methods due mostly to technical variation?

Questions (statistical methodology)

- Hierarchical clustering
  - What are the units they are clustering?
  - What is the distance? What is the linkage?
  - What do these dendrograms demonstrate?

- Gene Ontology Network Analysis
  - Why do they perform gene ontology network analysis?
  - How does gene ontology network analysis work?

Questions (grab-bag)

- How does the SMART-seq protocol work? (Nick)
- Is transcriptional bursting a cause of alternative splicing, or is alternative splicing a subtype of transcriptional bursting? (Diane)
- To what extent do these concerns around technical variation apply to bulk RNA-seq experiments? (Taylor)
- What are the advantages of scRNA-seq compared to bulk RNA-seq? (Hanley)