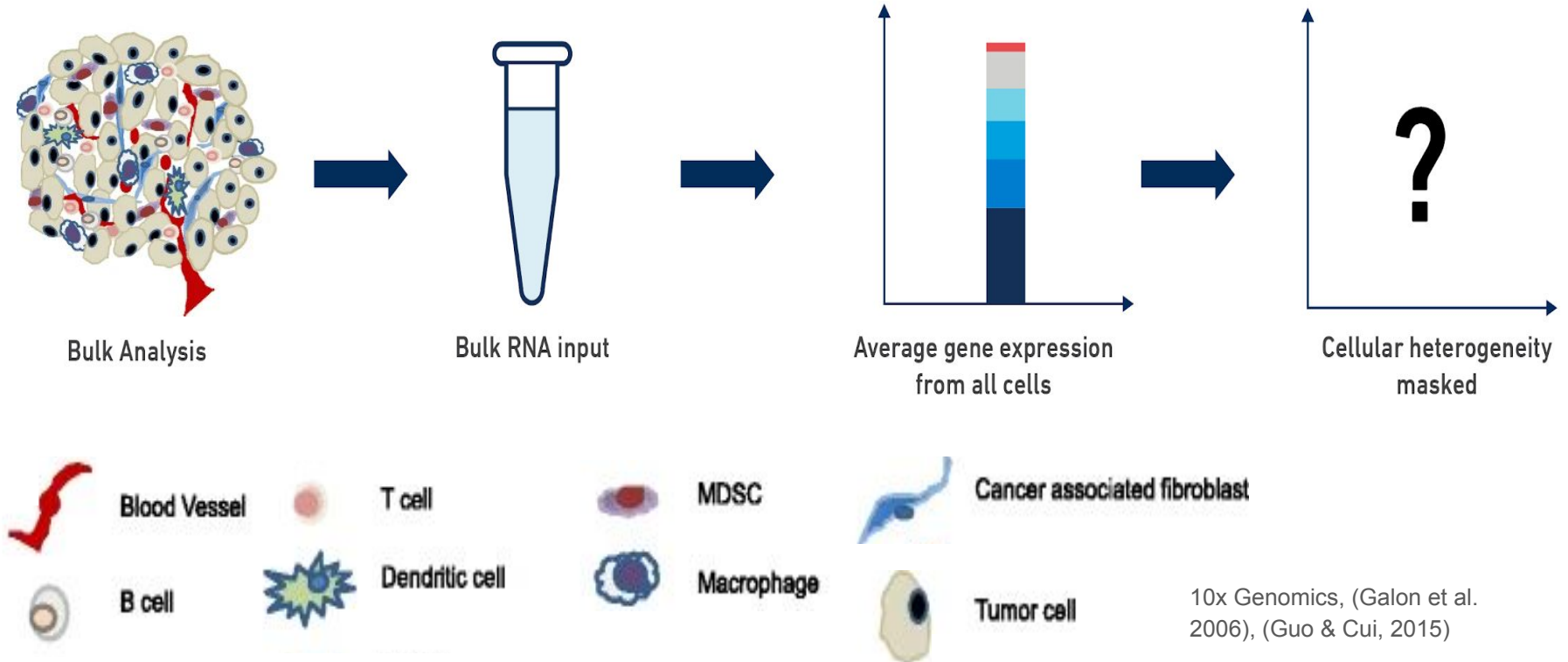# Analysis of Solid Tumor Tissue Archetypes through Bulk RNA-seq Deconvolution and Differential Gene Expression Analysis

Stephen Hwang, Tiana Pereira

Bulk RNA-seq of the tumor microenvironment masks immune cell type composition predictive of patient outcome

Bulk Analysis → Bulk RNA input → Average gene expression from all cells → Cellular heterogeneity masked

Blood Vessel   T cell   MDSC   Cancer associated fibroblast
B cell   Dendritic cell   Macrophage   Tumor cell

10x Genomics, (Galon et al. 2006), (Guo & Cui, 2015)

Deconvolution to immune cell type proportions coupled with DGE and unsupervised clustering may reveal tumor archetypes
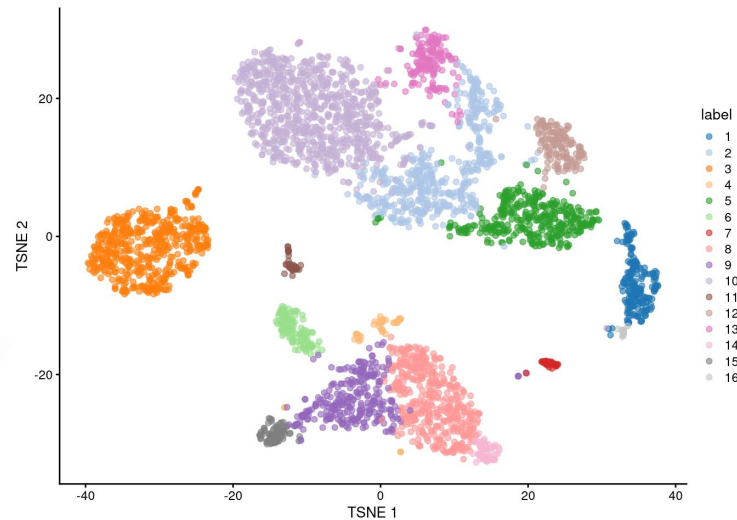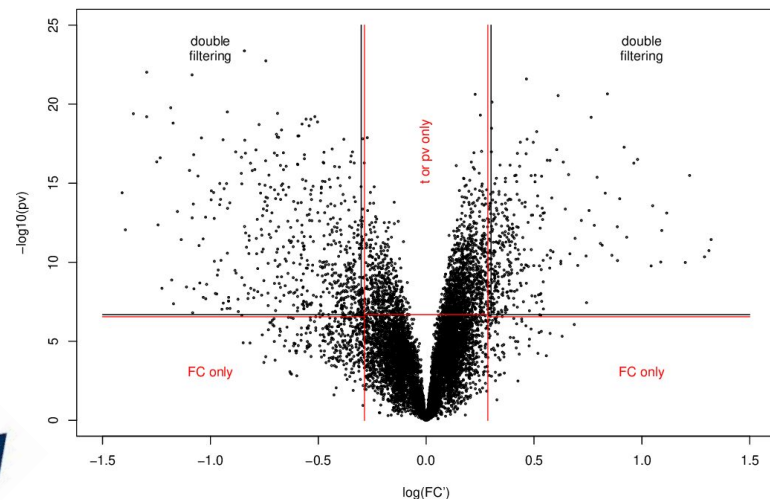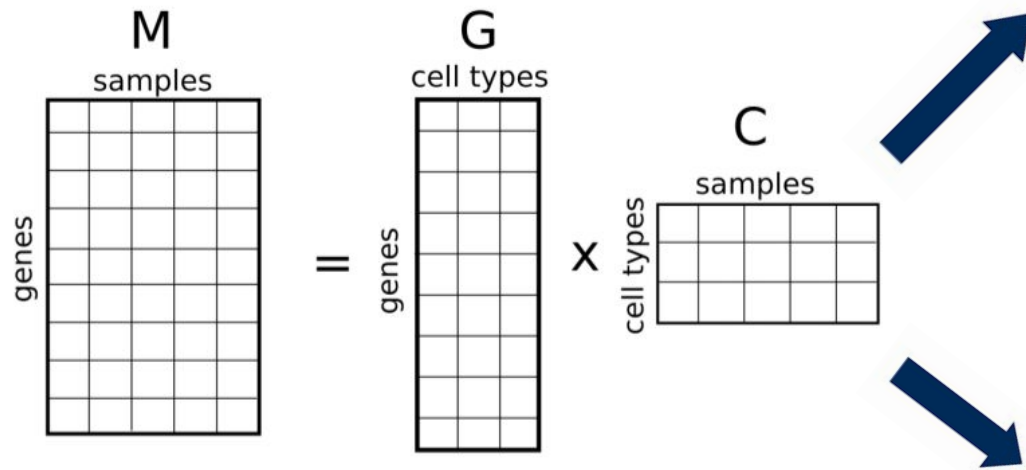
Objective

Goal: To identify solid tumors archetypes by differential gene expression and deconvolved immune cell type composition to better understand and predict clinical outcomes.

Hypothesis: Solid tumor immune cell type composition and the tumor transcriptome are predictors of patient outcome.

Significance: Median patient outcome for some solid tumors (i.e. GBM) can be low as 15-16 months. Gaining insight to clinical outcome and tumor heterogeneity may lead to better, personalized treatments.

Proposed methods

- Select solid tumor bulk RNA-seq count data from The Cancer Genome Atlas (TCGA)
  - Selecting pre-calculated counts quantified by **STAR-HTSeq-counts** and **Kallisto** (Pseudo-Bake-off)
  - Quality Control: Plot PCA and gene expression count boxplots
- Deconvolution
  - CIBERSORTx using LM22 signature matrix
- Clustering (Bake-off)
  - **K-means** vs **Louvain Clustering**
  - Gene expression
  - Deconvolved immune cell types proportion
  - Compare clusters using Adjusted Rand Index and Silhouette Score (validation)
- Differential gene expression across clusters (Bake-off)
  - **DEseq2** vs **Sleuth**
  - Compare up/down-regulated genes between outputs of tools (validation)
  - Identify GO terms that match immune cell type proportion enrichment per cluster (validation)

Proposed methods (cont.)

- ● Survival analysis
  - ○ Determine whether there are any cell types associated with survival time
  - ○ Kaplan-Meier Curve
- ● Cox Proportional-Hazards Model
  - ○ Analyze associations between survival time and predictor variables
    - ■ Immune cell type proportion
    - ■ Gene ontology scores
    - ■ Up/down-regulation