# Assessment of cell type identification methods for single-cell and spatial transcriptomic analysis in mouse

**Yang Bai** (baiy4@rpi.edu), RPI, **Yang Bai**, Rensselaer Polytechnic Institute

Single-cell and spatial transcriptomic profiling is increasingly used in biomedical research and bioengineering development. Annotating known cell types with a reliable prediction of novel cell types is a crucial but challenging step in the bioinformatic analysis of single-cell and spatially resolved gene expression data. RPI Genomics Core provides end-to-end single-cell RNA-seq services using two 10x Chromium devices. Additionally, the core also actively develops low-cost spatial RNA-seq protocols for diverse freshly frozen types of tissue. Combining in-house single-cell and spatial RNA-seq profiles with publicly available datasets and references, we evaluated new and popular automated cell type annotation tools within the last five years. This assessment includes all three major annotation methods: marker gene database-based method, reference dataset correlation-based method, and AI/ML supervised classification-based method. From our study, scBERT, a pre-trained Transformer-based language representation deep learning model, showed outstanding performance in cell type identification, highlighting the powerful application of AI techniques such as gene embeddings, Transformer architecture, and BERT paradigm in single-cell and spatial transcriptomic analysis.