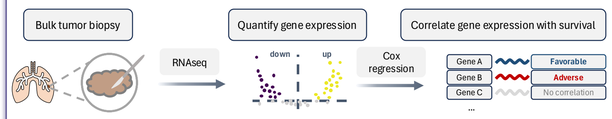
Pan-cancer identification of cell states associated with prognostic risk and therapy response at single-cell and spatial resolution

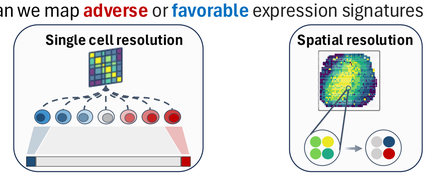
Brooks Benard[1], Chinmay Lalgudi[1], and Andrew Gentles[1,2,3]
Departments of [1]Pathology, [2]Medicine, and [3]Biomedical Data Science

Project motivation

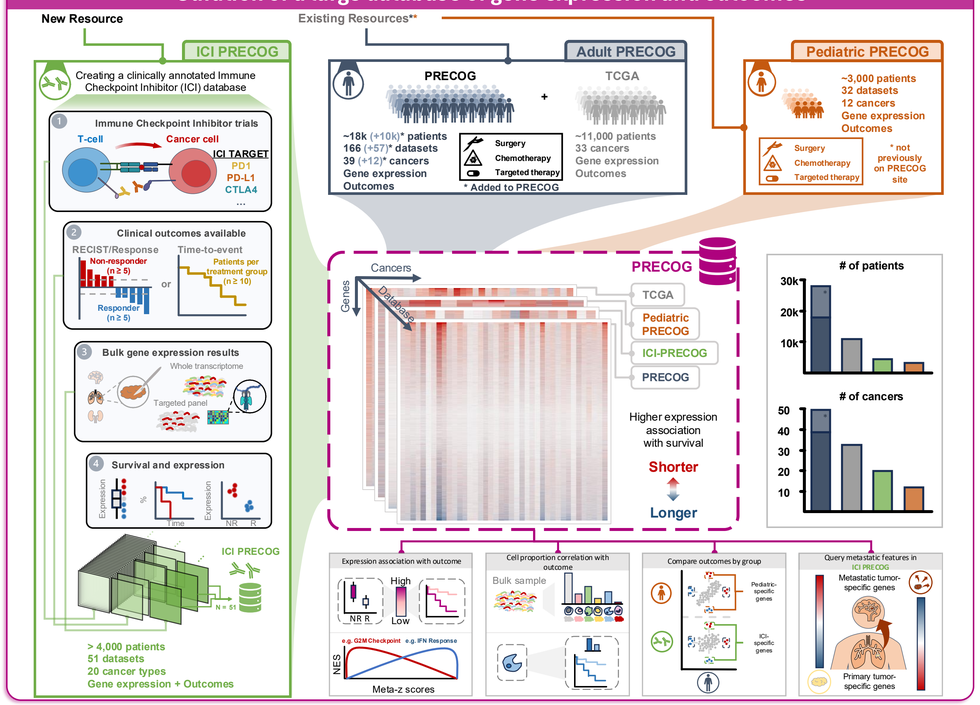
Robust discovery of therapeutically relevant single-cell states across cancers helps inform drug-target selection. To identify cells contributing to **adverse** or **favorable** outcomes in cancer, we first augmented our PRECOG[1] resource to generate a comprehensive prognostic map of gene expression across cancers, integrating data from over **46,000 adult and pediatric patients across 339 datasets and 60 histologies**, encompassing both standard and immunotherapy treatments. Next, we developed a semi-supervised method (scIMPEL) to map prognostic signal from these bulk cancer gene expression datasets onto individual cells in single cell and spatial transcriptomics data. We applied scIMPEL in a pan-cancer approach across **223 adult and pediatric single-cell RNAseq datasets representing 57 cancer subtypes and over 10 million cells**. We also applied scIMPEL to **~100 spatial transcriptomic samples from 4 cancers** to uncover spatial determinants of prognostic associations. Clinical annotations were available for most samples, allowing us to assess the association of scIMPEL results based on treatment interventions (e.g. chemo vs. immunotherapy), site of sample acquisition (e.g. primary vs. metastatic), and other features (e.g. diagnosis vs. relapse). Together, the pan-cancer single cell and spatial application of our scIMPEL method represents a rich resource from which to identify and test improved cancer therapies.



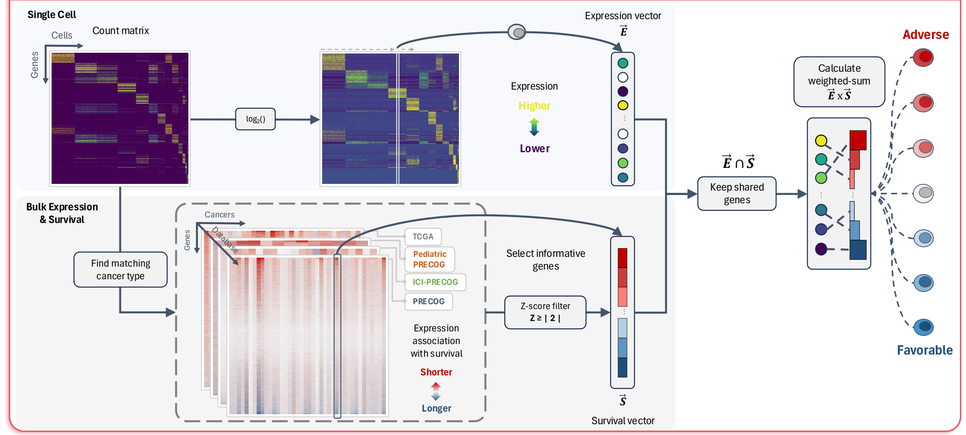
Can we map **adverse** or **favorable** expression signatures at:



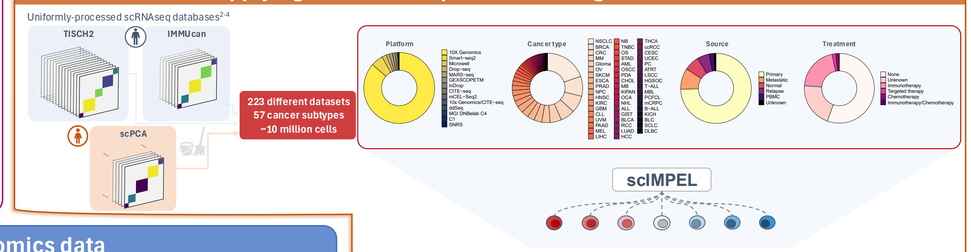
Curation of a large database of gene expression and outcomes



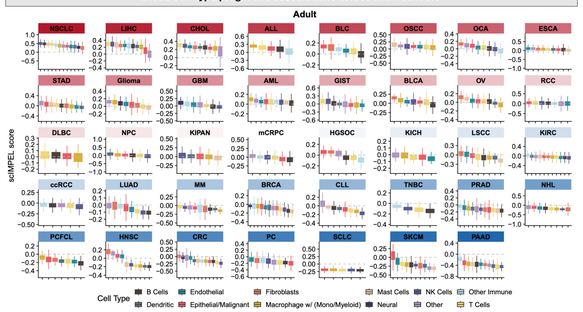
scIMPEL approach for assigning bulk prognostic signal to single cells

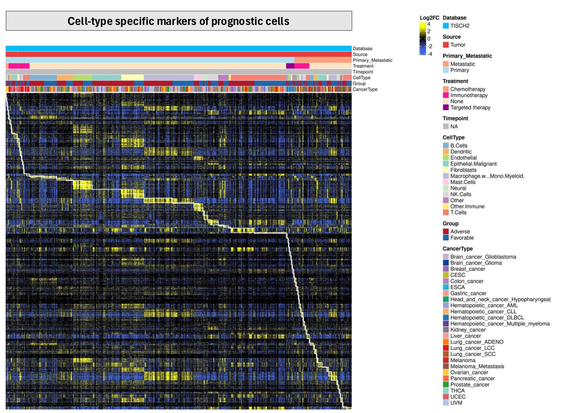


Applying scIMPEL to a pan-cancer single cell database



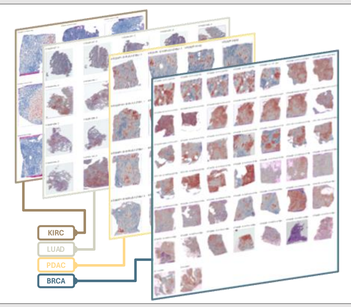
Prognostic cells and marker genes across cancers



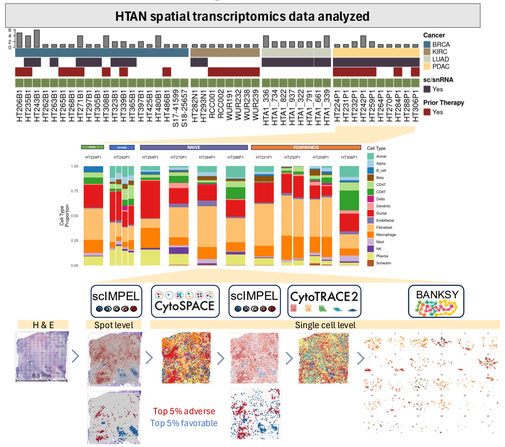


Localizing prognostic signal at single cell resolution in spatial transcriptomics data

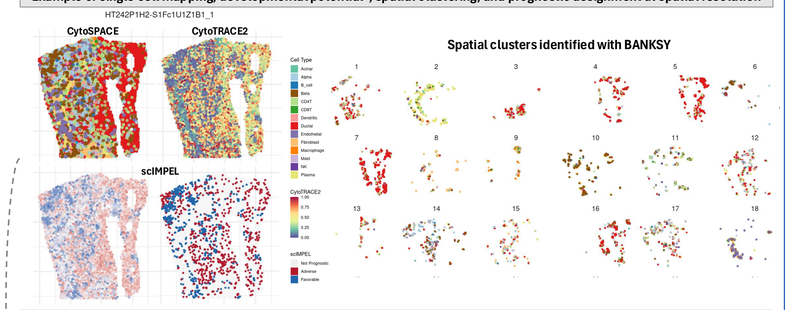
Spot-level scIMPEL scores for all available HTAN Visium samples



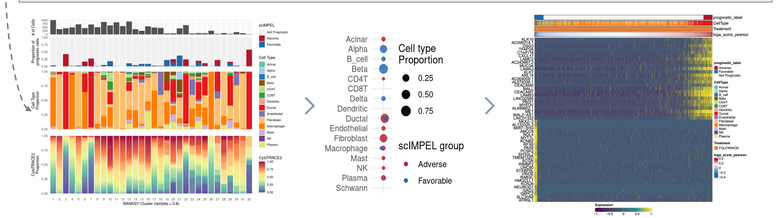
HTAN spatial transcriptomics data analyzed



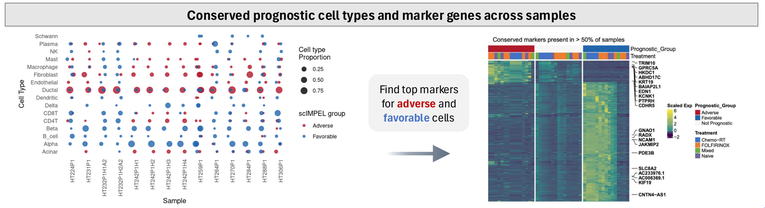
Example of single cell mapping, developmental potential[5], spatial clustering, and prognostic assignment at spatial resolution



Example level summary of cluster enrichment, cell type association, and marker genes for prognostic cells



Conserved prognostic cell types and marker genes across samples



Conclusions

- PRECOG now contains gene expression and outcomes annotations for **>46,000 adult and pediatric patients across 339 datasets and 60 histologies**, encompassing both standard and immunotherapy treatments.
- We developed a semi-supervised method to map pan-cancer prognostic expression signatures onto single cell and spatial transcriptomics datasets.
- We aggregated over **10 million cells from 223 single cell datasets** representing 56 subtypes across adult and pediatric cancers.
- We identified pan-cancer, cancer specific, and cell type specific marker genes for adverse and favorably prognostic cell types.
- Application of our method to spatial transcriptomics datasets identified strong association of our prognostic signal with spatial clusters and cell types.

Works cited

1. Gentles, Andrew J., et al. "The prognostic landscape of genes and infiltrating immune cells across human cancers." *Nature medicine* 21.8 (2015)
2. Han, Ya, et al. "TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment." *Nucleic acids research* 51.D1 (2023)
3. Camps, Jordi, et al. "Meta-analysis of human cancer single-cell RNA-seq datasets using the IMMUcan database." *Cancer Research* 83.3 (2023)
4. Hawkins, Allegra, et al. "The Single-cell Pediatric Cancer Atlas: Open-source data and tools for single-cell transcriptomics of pediatric tumors" *Cancer Research* (2024)
5. Kang , Armenteros1 "Mapping single-cell developmental potential in health and disease with interpretable deep learning" *BioRxiv* (2024)
6. Jin, Suoqin, et al. "CellChat for systematic analysis of cell-cell communication from single-cell and spatially resolved transcriptomics" BioRxiv (2023)

Stanford MEDICINE | School of Medicine
bbenard@stanford.edu