Background While immunotherapy has revolutionized the treatment of many solid tumors, the efficacy of immunotherapy regimens is comparatively lower in breast cancer. Immunotherapy efficacy is often negatively correlated with intratumor heterogeneity. Novel breast cancer immunotherapy approaches should leverage how intratumor heterogeneity affects immune cells in the tumor microenvironment. However, current definitions of heterogeneity in breast cancer have limited resolution. Single-cell RNA-seq (scRNA-seq) provides an unprecedented opportunity to further define cancer epithelial cell heterogeneity and identify how it influences immune interactions.

Methods We generated a novel dataset of 236,363 cells from 119 primary breast tumors biopsied from 88 patients taken from 8 publicly available datasets, currently the largest published scRNA-seq dataset in breast cancer. To define cancer epithelial cell heterogeneity, we performed unsupervised and supervised clustering based on molecular subtype and clinical target expression of all cancer epithelial cells. This identified 10 gene elements (GEs), which reflect molecular features that vary between cancer epithelial cells. Receptor-ligand pairing analysis determined how cells that highly express each GE interact with immune cells. We developed InteractPrint, a score which predicts the predominant tumor-interacting immune cells based on GE composition of a patient’s tumor.

Results In our dataset, 20% of samples were HER2+, 46% HR+, and 32% TNBC. This dataset was statistically powered to characterize cancer epithelial cell heterogeneity. For the 10 GEs, we predicted interactions with immune cells. GEs with predicted NK cell interactions were resistant to NK cell cytotoxicity. In a spatial transcriptomics dataset, GEs with predicted T cell interactions demonstrated colocalization with CD8+ T cells. To infer GE-immune interactions at the patient level, we developed InteractPrint and assessed its accuracy in predicting response to anti-PD-1 therapy. Across two trials and clinical subtypes, T cell InteractPrint demonstrated significant improvement over PD-L1 in predicting response to anti-PD-1 therapy. In an scRNA-seq dataset of samples from patients treated with pembrolizumab, we observed AUC of 81% for T cell InteractPrint versus 54% for PD-L1. In patients treated with paclitaxel + pembrolizumab in the I-SPY2 trial, we observed AUC of 84% for T cell InteractPrint versus 73% for PD-L1.

Conclusions We defined intratumor heterogeneity and leveraged it to predict immune cell interactions within a patient’s tumor. We developed T cell InteractPrint which captures heterogeneity in cancer epithelial and CD8+ T cell interactions and is predictive of anti-PD-1 therapy response at higher AUC than PD-L1. This provides a path forward for the interpretation of intratumor heterogeneity in a clinically meaningful way.