Abstract

Despite advances in breast cancer diagnoses and treatment, it continues to be the second leading cause of cancer death in women. Recent interest in managing and prognosticating breast cancer has been centered around examining single-cell and spatial expression patterns on molecular imaging, such as immunofluorescence and immunohistochemistry ((Wu, F. *et al.* Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat. Commun.* **12**, 2540 (2021).)). However, current techniques for this single-cell quantification are disjointed and steps are not formatted to be compatible.

The purpose of this project is to develop a simple, easy to use, end-to-end pipeline to quantify how much a patient's cancer changes at different timepoints in their treatment. To accomplish this, serial cyclic immunofluorescence images acquired over the course of a patient's breast cancer treatment were gathered from the Human Tumor Atlas Network (HTAN). Two images of different samples were chosen from time point 0 and one image from time point 1 was chosen. Each image consisted of 25 unique channels, including DAPI, HER2, and ER. These images were used to train a patient-specific Variational Auto-Encoder (VAE), which was then used to quantify how much the cancer changed. In order to train the model, individual cells were segmented using an open-source tool called MCMICRO, resulting in around 14,000 cells in each image.

By running these analyses, the tumor phenotype of the patient (e.g., HER2 negative, estrogen receptor [ER] positive) was able to be automatically determined. In addition to this the MCMICRO quantification outputs show that the marker expression changed at a later time point as the patient underwent hormone therapy for their breast cancer. For example, the ER expression decreased 2-fold and the Ki67 expression increased 12-fold. Additionally, the VAE latent space plots show significantly different distributions at different time points, implying that the network was able to encode the baseline single-cell expression patterns.

This research shows that automated single-cell quantification of marker expression in breast cancer can be important to track tumor evolution over time, which can be quantified, visualized, and predicted using a VAE.