MAPPING A HUMAN STEM CELL STATE SPACE USING A MICROSCOPY BASED PIPELINE

Winfried Wiegraebe and the Allen Institute for Cell Science Team,
Allen Institute for Cell Science, Seattle, WA, USA
winfriedw@alleninstitute.org

KEY WORDS: Cell state, human induced pluripotent stem cells (hiPSC), light microscopy, automation, pipeline, concentration, FCS

Cells occupy many types, residing in multiple states. They can differentiate into highly structured cardiomyocytes or transition in any of a number of states, e.g., interphase, cycling, migratory or apoptotic, to name just a few. All of these changes are characterized by changes in cell shape and organization. We are conjoining single cell imaging, single cell RNAseq, and phenotypic readouts to identify and understand different cell types and states. The goal is to create a high dimensional state space, elicit the principles of cellular morphogenesis, and determine how cells transition between states.

Towards this goal, we created a library of more than 25 hiPSC lines (human induced pluripotent stem cell) with each expressing a genome-edited protein that labels each of the major cellular structures (nucleus, golgi, mitochondria, etc.) These structures are tagged by eGFP monoallelically at the endogenous locus using the CRISPR/Cas9 system and available to the public for research purposes (https://www.allencell.org/cell-catalog.html).

To examine the morphology of these cells and structures, we created a pipeline consisting of robotics for cell handling; microscopy for live-cell, 3D imaging; cluster computing for image analysis; and a LIMS system for data organization. The pipeline is optimized to generate high quality images in high replicates (thousands of images). Automation, with human operators with biological knowledge involved, ensures reproducibility and biological fidelity.

We have also created intensity/concentration calibration standards using purified eGFP in solution to measure the concentrations of proteins under study. We calibrated the concentration standards with FCS (Fluorescence Correlation Spectroscopy). The concentration standards allow quantitative comparison of data collected on different systems. We validated this approach with quantitative Western blots, ELISA, and UV/VIS spectroscopy. We will discuss the advantages and limitations of the different standards, as well as the limitations, corrections, and controls necessary because of mono-allelic expression of eGFP-tagged proteins, inactive eGFP, and environmental changes of eGFP brightness.

Finally, we developed deep-learning based methods to combine thousands of images into an integrated cell derived from independent experiments. With similar algorithms we train a network to predict the location of multiple structures, both discretely and statistically, in transmitted light images (https://www.allencell.org/allen-integrated-cell.html).

All data and methods are available at https://www.allencell.org/3d-cell-viewer.html.