

Image-based multi-dimensional analysis pipeline for high-throughput cell screening data

Yang Shen¹, Nard Kubben², Julián Candia³, Alexandre V. Morozov⁴, Tom Misteli², Wolfgang Losert¹

1. Chemical Physics Graduate Program, University of Maryland, College Park, MD 20742, USA
2. National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA
3. Center for Human Immunology, Autoimmunity and Inflammation, National Institutes of Health, 10 Center Drive, 10/7N115, Bethesda, MD 20892, USA
4. Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey, USA

Motivation: Large scale RNA-interference screens combined with multichannel imaging yield a powerful technique for unbiased systematic analysis of biological processes. However, the enormous amount of single cell information available in the resulting imaging datasets, and the heterogeneity at the single cell level are emerging as a significant obstacle toward high throughput image-based cell biology. The majority of analyses to-date ignore the richness of the information – using only one or two features for each cell – and the heterogeneity of the cell population.

Results: We have developed a computational analysis pipeline suitable for image based high-throughput screening experiments that involve two sets of controls (e.g., healthy and diseased cell populations) and multiple staining channels. For each staining channel, we used multi-dimensional single-cell analysis to learn what combination of parameters in each image channel best distinguishes the two controls. This allows us to quantify how effectively each RNAi changes cell properties, and to select RNAi “hits” that significantly shift the cell population from the negative to the positive control. We can determine “hits” both based on each image channel separately, or combinatorially in multiple channels. We develop and demonstrate our approach using data from an experiment of model progeria cells. Our pipeline identifies more potential RNAi hits than standard analysis methods, and provides more insights into both average and single cell properties. One insight from our first application is that some RNAi are double “hits”, i.e. shift cell population behavior for two features on average but not at the single-cell level.