Estimating the Effect of Cell Cycle on Transcriptomes and Biophysical Properties of Single Cancer Cells

Ryland Mortlock, Brian Joughin, Peter Winter, Scott Manalis, and Douglas Lauffenburger

Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA

Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA

Introduction

Despite advances in the treatment of cancer, relapse remains a significant barrier to patient survival. Relapse can be driven by therapeutic resistance in a small number of cells, pointing to the need to study cancer at the single cell level. This project aims to estimate how each cell's cell cycle state contributes to its gene expression and buoyant mass. An accurate, high-resolution estimate of cell cycle state will strengthen models to predict response to treatment as a function of single cell gene expression and biophysical properties.

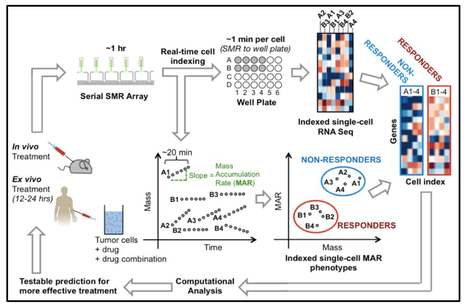


Figure 1 | The MIT Cancer Systems Biology Consortium (CSBC) aims to functionally characterize therapeutic resistance at the single cell level. [1]

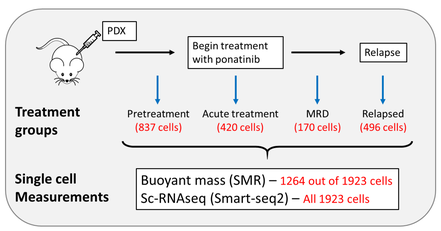
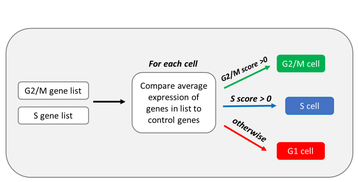


Figure 2 | Resistance is studied within a patient-derived xenograft (PDX) model of BCR-ABL+ B-cell acute lymphoblastic leukemia (B-ALL).

Motivation



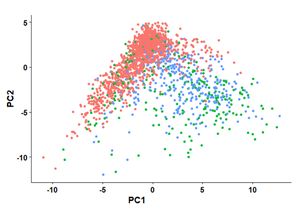


Figure 3 | Cell cycle effects constitute a significant signal in PDX expression data. (A) Workflow for Seurat's built-in function for cell-cycle phase assignment [2]. **(B)** The first principal component of the PDX expression data correlates highly with cell cycle phase.

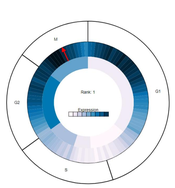


Figure 4 | Cyclebase [3] contains time course data for bulk expression of cell-cycle related genes in synchronized HeLa cells.
Averaged time course data for expression of PLK1, a characteristic G2/M gene, is shown by the outer colored ring. Red arrow indicates peak expression.

Methods

Figure 4 | A dot-product based calculation can be used to estimate the correlation between expression of cell-cycle related genes in single cells and time course expression data from synchronized HeLa cells.

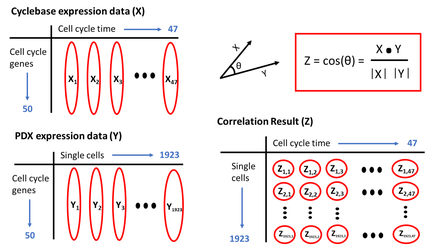
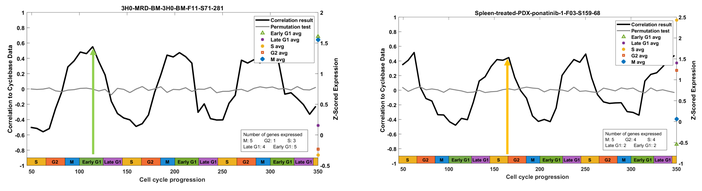


Figure 5 | Cells correlate with Cyclebase expression at characteristic timepoints. (A) A single PDX cell from MRD treatment group correlates with Early G1 / Late M phase expression and **(B)** a single cell from acute treatment correlates with S phase expression.



Results

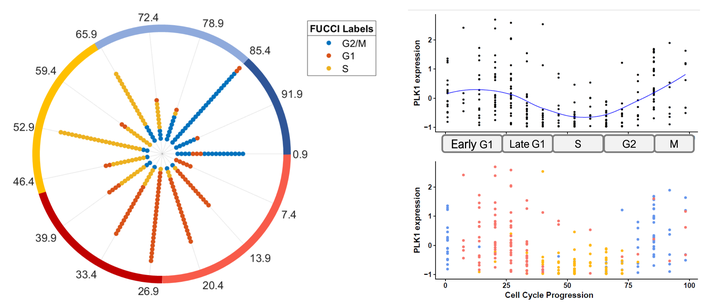


Figure 6 | New method is validated with expression data from FUCCI-labeled human embryonic stem cells [4]. (**A**) Cyclic histogram, points represent single cells (n = 247) binned in the theta direction by computational assignment of cell cycle progression 0-100. Outer color ring marks computational phase cutoffs, cells are colored by their FUCCI label. **(B)** PLK1 expression versus computationally assigned cell cycle progression, local regression line shown in blue. **(C)** Same data colored by FUCCI labels.

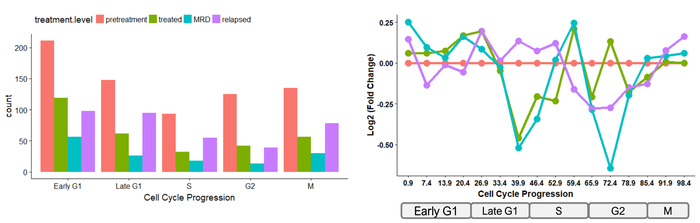


Figure 7 | New method reveals change in cell cycle state across treatment level in PDX model of B-ALL. (A) 1531 PDX cells binned by computationally assigned cell cycle progression. **(B)** Log fold change of each treatment level's proportion of cells at each cell cycle state, as compared to pretreatment. Cells treated acutely or at MRD are depleted near the G1/S transition relative to untreated cells.

Results cont.

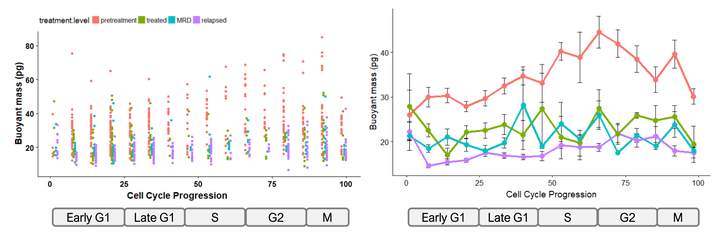


Figure 8 | Buoyant mass of pretreatment cells changes across cell cycle progression. (A) 1052 PDX cells with paired mass and gene expression data. **(B)** Average mass (±SE) of pretreatment cells increases with computationally assigned cell cycle progression until the S/G2 transition. Acute treatment, MRD, and relapsed cell show lower buoyant mass and smaller changes in buoyant mass across cell cycle progression, compared to untreated.

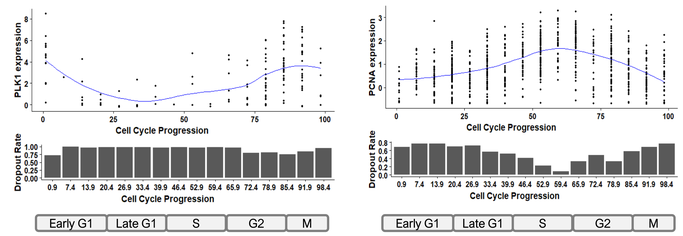


Figure 9 | Dropout rate for individual genes is higher for cells at cell cycle states when expression is lower. (A) Expression of PLK1 and **(B)** PCNA with dropouts removed. Dropout rate is the proportion of cells assigned to that cell cycle progression which had a transcript count of zero. Imputing dropout values would increase the number of low expression values, possibly strengthening the model.

Conclusion & Future Directions

A new method was developed to estimate **cell cycle progression** in cancerous B cells by mapping single cell expression data onto time course expression data from synchronized HeLa cells. The method performed well on **validation data**.

The method revealed differences in **cell-cycle state composition** across treatment levels in a PDX model of B-ALL. The method indicates a treatment-level specific relationship between **buoyant mass** and cell cycle progression.

Next steps include refining the model, investigating transcriptomic signatures for quiescence, and identifying differences in gene expression between sub-groups of B-ALL cells based on treatment level, mass, and cell cycle progression.

References

[1] MIT CSBC U54 Research Center website. [2] Butler, et al. *Nature biotechnology* (2018). [3] Santos, et al. *Nucleic acids research* (2014). [4] Leng, et al. *Nature methods* (2015).

Acknowledgements

This work was supported by the National Cancer Institute (NCI) Cancer Systems Biology Consortium (CSBC) and the Lauffenburger Lab (MIT Biological Engineering).