Multimodal analysis of DNA, surface proteins, and intracellular proteins in single cells

Prithvi Singh, Dalia Dhingra, Saurabh Parikh, Adam Sciambi, and Aik Ooi.
Mission Bio, Inc. South San Francisco, CA, USA.

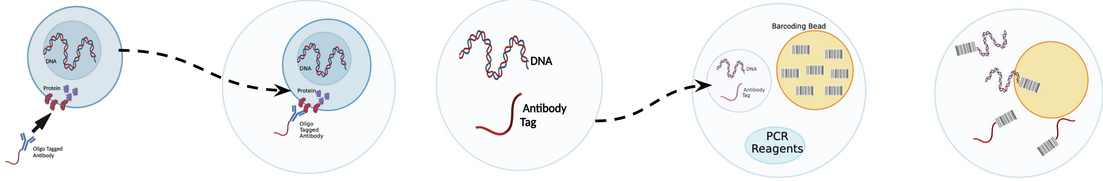
Introduction

Cancer is a disease driven by genomic mutation and selection. While the advent of next generation sequencing has revolutionized genomic analysis, the data obtained from bulk sequencing do not represent the heterogeneity in cancer. Single-cell RNA sequencing has advanced the knowledge of cellular heterogeneity, but this method is limited to the analysis of transcriptomes. As proteins are the key functional machinery of cells, capturing the differential protein expression from cell to cell would benefit the study of cancer biology. Multimodal approaches such as CITE-seq only partially address this problem by allowing simultaneous analysis of transcriptome and surface proteins, but not intracellular proteins that play vital roles in cellular functions. Additionally, these methods do not directly analyze DNA, hence do not provide a readout of genotypic information such as single-nucleotide variants (SNVs) and copy number variations (CNVs).

Here we describe a technology to overcome these hurdles. Using a novel workflow, cells treated with barcoded antibodies are encapsulated in droplets using the Mission Bio Tapestri® Platform and subsequently processed to obtain DNA and protein information from the same single cells.

The workflow

To perform single-cell multiomic DNA and protein analysis, cells are first treated with oligonucleotide-barcoded antibodies targeting surface proteins, then fixed and permeabilized, followed by incubation with barcoded antibodies for intracellular proteins (**1**). The resulting cells are processed on the Mission Bio Tapestri Platform through a two-droplet workflow. In the first droplet, cells are individually encapsulated and lysed (**2**), and gDNA is released from the chromatin by protease digestion (**3**). The first droplet is then captured in a second droplet where PCR reagents and cell barcodes are added for barcoding and amplification of the DNA targets (**4a-b**). DNA sequencing libraries are generated from the single cells using a multiplex panel of primers targeting regions of interest. Protein sequencing libraries are separately generated from the oligonucleotides off the antibodies. Both libraries are later paired through the cell barcodes for joint analysis.



1. Incubate cells with oligonucleotide-barcoded antibodies

2. Encapsulate single cells

3. Lyse and treat cells with protease.

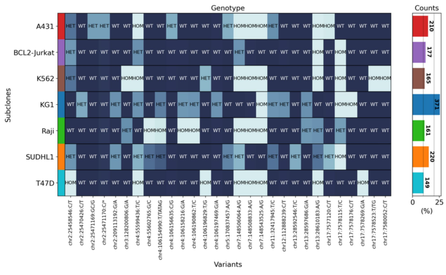
4a. Add barcoding bead and PCR reagents.

4b. Barcode and amplify analytes.

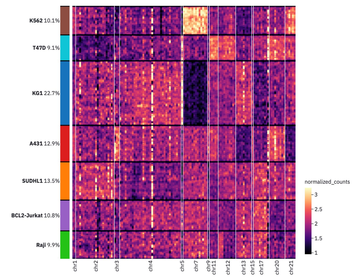
Concurrent analysis of DNA, surface proteins, and intracellular proteins

In the following experiment, seven cell lines were mixed and analyzed through the Tapestri Single Cell DNA AML panel, which examines 127 amplicons across 20 genes. Cell surface protein detection was enabled through the Total Seq™-D Heme Oncology Cocktail from Biolegend, which contains 45 antibodies including 3 isotype controls. In addition to the Total Seq-D surface protein panel, 5 intracellular protein targets were included as a proof of principle.

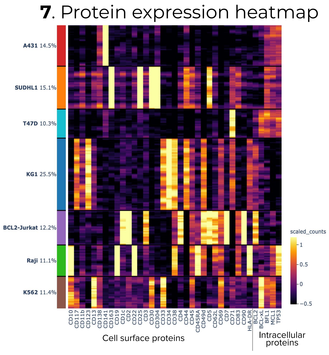
5. SNV-based cell identification



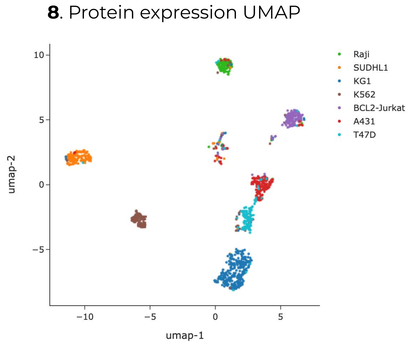
6. Normalized DNA reads



7. Protein expression heatmap



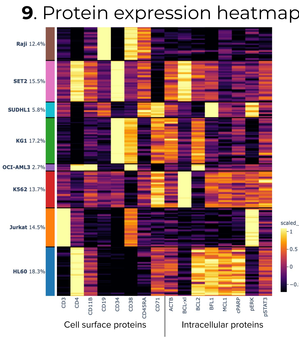
8. Protein expression UMAP



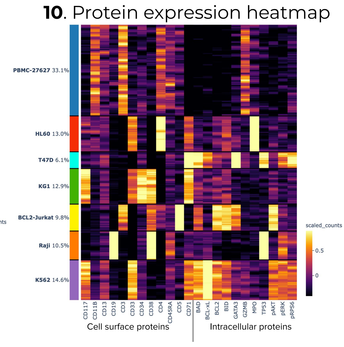
Cells were identified and grouped based on their genotypic signature through SNV analysis from the single-cell DNA data (**5**). To infer CNV information, normalized read counts from the DNA data were analyzed, with BCL2-Jurkat serving as a diploid reference (**6**). Through the read count analysis, it was determined that there was a copy number gain in chromosome 7 for K652 cells and a loss in the same chromosome for KG1 cells. Expression profiles of 37 proteins out of 50 were graphed on a heatmap to visualize differential expression (**7**). UMAP analysis was performed on normalized protein reads with the resulting plot showing cell types color coded based on SNV analysis (**8**).

The capability of this method was further validated with the detection of two types of post-translationally modified proteins: phosphorylated (phospho-ERK, phospho-STAT3), and cleaved proteins (cleaved-PARP) (**9**). Lastly, the workflow was demonstrated on frozen PBMCs mixed with cell lines to detect nuclear proteins (GATA3, TP53, MPO) along with additional phosphorylated proteins (phospho-AKT, phospho-RPS6) (**10**).

9. Protein expression heatmap



10. Protein expression heatmap



Conclusion

This is the first method to effectively link surface and intracellular protein measurements with targeted DNA analysis. This approach enables the correlation of genotypic and phenotypic readouts from single cells, allowing for concurrent multimodal analysis of cancer evolution and the protein expression that drives it.

singh@missionbio.com