Poster #148

Powerful Insights with Single-Cell Multi-Omics: Co-Detecting both Genotype and Phenotype from the Same Cell

Jim Flynn[1], Adam Sciambi[1], Aik Ooi[1], Dalia Dingha[1], Pedro Mendez[1], David Ruff[1], Saurabh Gulati[1], Shu Wang[1], Nigel Beard[1]

[1] *Mission Bio, South San Francisco, CA, USA. All are employees and shareholders.*

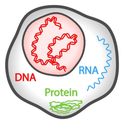
Abstract

Modern medicine has given rise to an array of treatment options for diseases such as acute myeloid leukemia (AML). Cancer is a heterogeneous mixture of cells with varied states, and the best treatment options to an individual patient requires an understanding of the disease state at the cellular level. A single-cell multi-omics approach is the only way to achieve full resolution of the disease at the cellular level, revealing the interplay between genotype and phenotype. Here, we show that the Tapestri Platform enables comprehensive identification of cell subpopulations through single-cell multi-omics profiling. The platform discerned single nucleotide variants (SNVs), copy number variations (CNVs), and cell surface protein expression in single cells from a mixture of AML cell lines and an AML research sample. Our results demonstrate that subpopulations are not consistently defined by one genetic or phenotypic factor alone, but by multiple parameters in conjunction and irretrievable by bulk sequencing methods. We also show that multi-omics data generated by the Tapestri Platform are consistent with the gold-standard techniques currently used for traditional omics analysis of AML samples. Taken together, our results demonstrate the power of the Tapestri Platform to uncover and define various cell states of AML using single-cell multi-omics.

The Multi-Omic Workflow

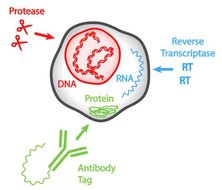
The Tapestri Platform is able to simultaneously analyze both genotypic (SNV, Indels, and CNV) and phenotypic (RNA and Protein) factors from thousands of cells individually. This comprehensive, single-cell analysis is performed by a series of analyte-specific processing steps to generate associated DNA sequences: proteins are tagged with DNA-conjugated antibodies, RNA is reverse transcribed into cDNA, and gDNA in the nucleus is freed from chromatin via protease. The resulting DNA sequences can then be barcoded with a cell-specific sequence for later identification.

Mission Bio enables single-cell DNA, RNA, and protein detection



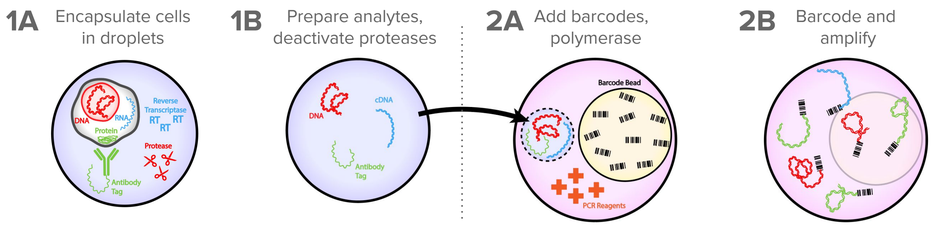
Multi-Omic Strategy:

- Tag proteins with antibodies.
- Reverse transcribe RNA.
- Free DNA from nucleus.
- **Barcode all.**



The workflow is enabled by its two droplet steps (below). Cells are individually encapsulated in a first droplet (1A) where analyte preparation can occur, including aggressive digestion and subsequent protease heat inactivation (1B). Afterwards, sensitive enzymes that are incompatible with the earlier preparation can be added to make a new drop (2A) where barcoding and amplification proceed (2B).

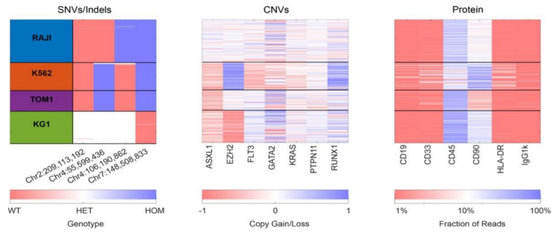
The Two-Step Workflow



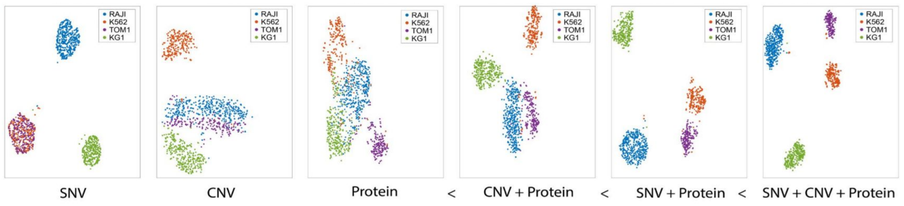
The Power of Single-cell Multi-Omics

Multi-omics tell a complete biological story. Any particular analyte might contain a critical disease signature that would otherwise be missed. Taken together, multiple analytes can show how a mutation or copy change directly alters RNA and protein expression. Yet another benefit is the extra power to resolve subpopulations. As an example of the latter, the plots below show four cell lines that have been combined in one run. Single-cell SNV, CNV, and protein clustering do not fully reconstruct the four subtypes, but together they can.

Raw Single-Cell Data: SNV, CNV, Protein

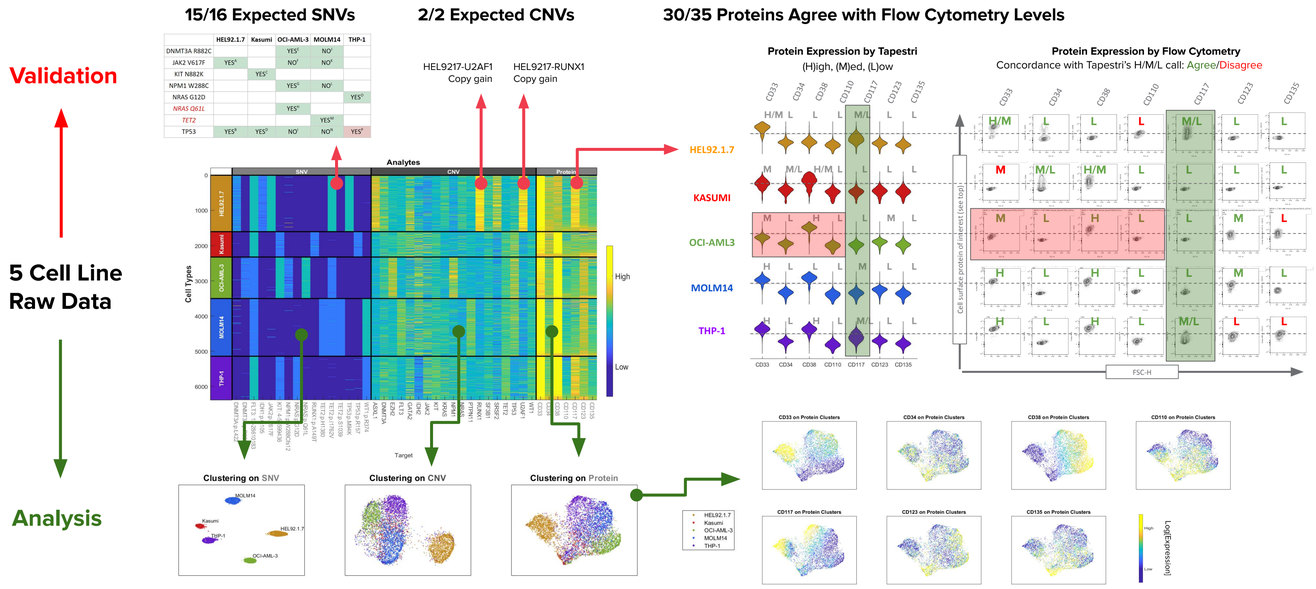


Clustering by Analyte



Validating the Platform

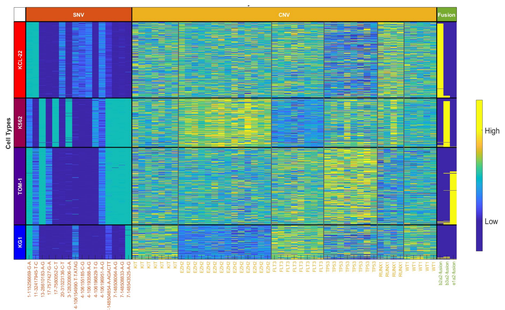
To demonstrate the accuracy of the system, many well known samples have been run and found to have concordance with literature. Given below is a 6000-cell customer Tapestri run where five cell lines were combined and evaluated against a targeted SNV, CNV, and Protein panel. From the raw data (center heatmap), the customer validated 15/16 expected SNVs, 2/2 expected CNVs, and 30/35 protein levels measured on those same cell lines with flow cytometry (above, right). At the same time, each set of analytes could be investigated de novo (below in green) to see how they organize. A product of that analysis is a mapping of individual protein expression on top of panel-wide protein clustering to resolve cell lines (bottom right).



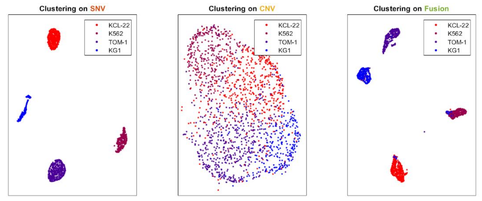
Detecting RNA Fusions

To demonstrate RNA fusion detection, a run was performed with four cell lines where each exclusively expressed one of three BCR-ABL fusions or none of them. SNV and RNA fusions were each able to distinguish the populations, and CNV had a robust signature. Sensitivities and specificities were above 90% for two of the fusion transcripts with a third at 70% sensitivity and 99% specificity, as denoted by the yellow bars representing expression >20 copies.

Heatmap: Cell Types vs SNV, CNV, RNA Fusion



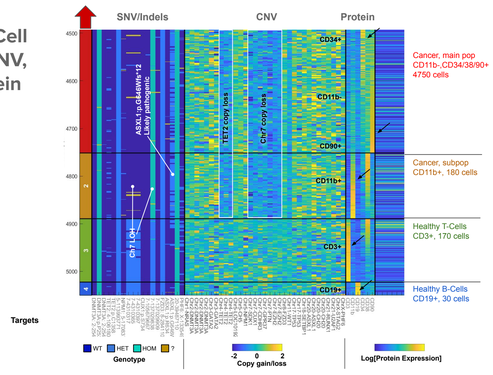
Clustering on SNV, CNV, and RNA Fusion



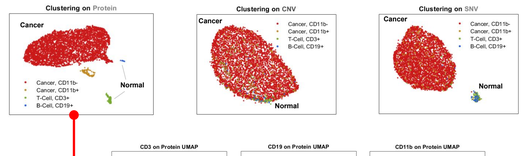
An AML Patient Sample: SNVs, Copy Loss/Loss of Heterozygosity, Protein Expression

In this experiment, an AML patient sample was analyzed by a Tapestri customer and 5000+ cells were recovered. The sample was found to be 95% cancerous as identified by any of the three analytes. Most striking was the Chromosome 7 copy loss, a known pathologic mechanism of AML, as detected in CNV and orthogonally validated by loss of heterozygosity in the SNV data (see heatmap, lower left). In addition, the cancerous population was identified by a pathogenic mutation (ASXL1 p.G646V) and higher expression of CD11b, CD34, CD38, and CD90 in two sub populations (lower right.)

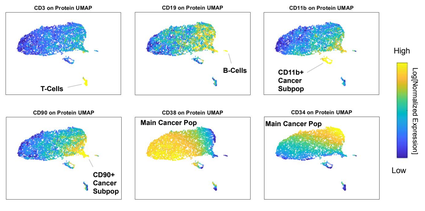
Heatmap: Cell Types vs SNV, CNV, Protein



Clustering on SNV, CNV, and Protein



Overlaying Individual Markers on Protein Clusters



Conclusions

Within a heterogeneous sample various populations are defined by inherent patterns, which may be defined by both genotype and phenotype at single-cell resolution revealing additional biological insight not visible with conventional bulk analysis. Taking a single-cell multi-omics approach ensures coverage of both parameters to capture the characteristic with the strongest signal to define critical subpopulations. Our data illustrates the power of integrated single-cell, CNV, fusion, and protein expression data to differentiate cancerous subpopulations and strengthen the conclusions made from individual analytes alone.