mission bio

tapestri
APPLICATION NOTE

Tapestri Platform Resolves Clonality of Heterogeneous Mouse Organoid Cancer Model through Single-Cell DNA Sequencing of Lentiviral Barcodes

Takeaways

- Tapestri Platform now supports the mouse genome
- Clonality of a heterogeneous mouse organoid bladder cancer model is resolved with the Tapestri Platform
- Custom panel of amplicons to lentiviral barcodes resulted in high sensitivity of subclone detection, below 1%, with co-occurring gene insertions
- A novel method using single-cell DNA sequencing was developed as a system for investigating the functional impact of higher-order genetic interactions in cancer

Abstract

The genomic heterogeneity of bladder cancer is expansive, with over 30 known oncogenic genetic alterations and each tumor harboring at least 5.[1] Studying these diverse higher-order genetic interactions that drive bladder cancer is difficult with current models of tumorigenesis and limited by bulk sequencing that fail to directly discern clonality and resolve mutational co-occurrence patterns. Here, we demonstrate an organoid model of bladder cancer using lentivirus to introduce gain or loss of function of multiple genes per cell in random combinations. Organoids are selected for in vivo tumorigenesis and profiled through single-cell DNA sequencing using the Tapestri Platform and a custom single-cell DNA panel targeting lentiviral barcodes. This method revealed combinations of altered genes that led to tumorigenesis within immunocompetent and immunodeficient mice.

Experiment & Methods

Over 30 genes were chosen for gain or loss of function based on recurrent frequencies of at least 5% in bladder cancer. Open reading frames (ORFs) or shRNAs were introduced to cells via lentivirus that also included a DNA barcode for each gene. Primary C57/BL6J mouse bladder urothelial cells were isolated and infected with lentivirus at a multiplicity of infection corresponding to an average of four gene aberrations per cell. In Matrigel, cells formed organoids after 72hrs. Organoids were then grafted into immunocompetent (C57BL/6J) or immunodeficient (NSG) mice. Organoids that formed tumors were sequenced at the single-cell level using the Tapestri Platform with a custom amplicon panel to lentiviral barcodes. Cell barcodes were extracted using umi-tools using pre-filtered sequencing reads. The reads were aligned to our custom reference by bwa (Burrows-Wheeler Aligner) and matrix of cell x amplicon was constructed using python. After additional filtering and read normalization the cell population was defined based on the reads of different amplicons.

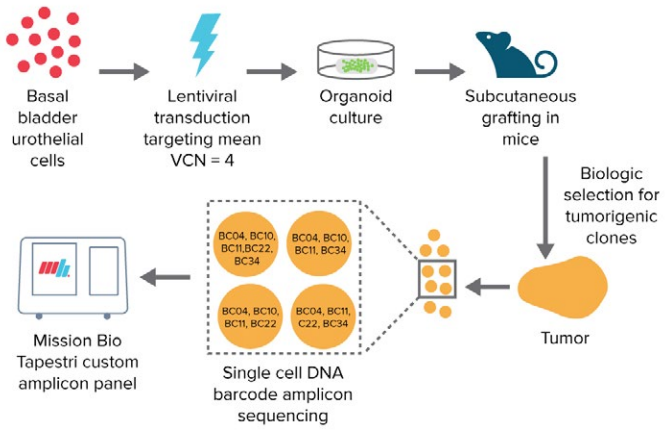


Figure 1 - Schematic of lentiviral transduction, organoid grafting, and single-cell DNA sequencing of a novel mouse model of bladder cancer to uncover higher-order genetic interactions that lead to tumorigenicity .

Results

Mouse organoid bladder cancer model enables investigation of higher-order genetic interactions through single-cell DNA sequencing

Traditional technologies to investigate the functional contribution of genetic alterations in cancer are inefficient, low-throughput, and do not address the complexities of higher-order compound genetic interactions. Available genetically defined model systems for cancer severely underrepresent the heterogeneity found in human cancers. To develop a robust, economical, and biologically relevant assay that enables the functional annotation of the compound and context-specific oncogenic drivers of bladder cancer, the Lee lab from the Fred Hutchinson Cancer Research Center combined genetically engineered organoids with mice grafting, and single-cell DNA sequencing (Figure 1).

In this assay, mouse urothelial cells were harvested and infected with a mixture of lentiviruses containing expression cassettes to mimic gain or loss of function of genes relevant to bladder cancer. Lentiviral cassettes also included a genetic barcode for downstream single-cell analysis. Infected cells averaged four gene aberrations per cell, and were grafted into immunodeficient mice after having formed in vitro organoids. The utilization of organoids in this assay allowed for easy manipulation of cells ex vivo, and tumor formation after grafting into mice.

Organoids that displayed tumorigenicity in mice were then resected and DNA sequenced at the single-cell level using the Tapestri Platform. A custom amplicon panel was designed to sequence lentiviral barcodes in order to distinguish clones containing gain or loss of function cassettes. Using this method, combinations of co-occurring gene aberrations were identified in single cells, establishing clonal populations and giving insight into higher-order genetic interactions that lead to disease.

Single-cell tumor analysis reveals four clonal populations with co-occurring genetic aberrations

As a proof-of-concept of the assay, and to develop a genetically defined model for squamous cell carcinoma of the bladder, genetically modified organoids were generated with combinations of lentiviral vectors expressing PPARG, PVRL4, YWHAZ, FGFR3 S243C, and PIK3CA E545K. These genes corresponded with lentiviral barcodes BC04, BC10, BC11, BC22, and BC34, respectively. The organoids were grafted into immunodeficient mice and allowed to grow into tumors. In one mouse that developed a tumor from the grafted organoid, the tumor was resected and then analyzed using the Tapestri Platform for single-cell DNA sequencing. 2,933 cells were analyzed from the tumor using custom amplicons for detection of the lentiviral barcodes.

Single-cell DNA sequencing revealed four distinct populations of cells within the heterogeneous tumor sample (Table 1). Each clonal population had several co-occurring genetic aberrations, with clones detectable under 2%. The most abundant clone had been modified with all five genes at over 90% of the tumor cell population. PPARG and YWHAZ expressing lentiviral vectors were found in all four clones, suggesting a critical role in bladder cancer formation.

LV Barcodes	PPARG	PVRL4	YWHAZ	FGFR3 S243C	PIK3CA E545K	Clonal Frequency
BC04, BC10, BC11, BC22, BC34	+	+	+	+	+	91.65%
BC04, BC10, BC11, BC34	+	+	+	-	+	3.92%
BC04, BC10, BC11, BC22	+	+	+	+	-	2.73%
BC04, BC11, BC22, BC34	+	-	+	+	+	1.70%

Table 1 - Single-cell tumor clonality analysis reveals four clonal populations, each with multiple co-occurring genetic aberrations, within a heterogeneous sample.

a

LV Barcodes	PPARG	PVRL4	YWHAZ	FGFR3 S243C	PIK3CA E545K	Clonal Frequency
BC04, BC10, BC11, BC22, BC34	+	+	+	+	+	89.70%
BC04, BC10, BC11, BC34	+	+	+	-	+	8.70%
BC04, BC10, BC11, BC22	+	+	+	+	-	0.76%
BC04, BC11, BC22, BC34	+	-	+	+	+	0.75%

b

LV Barcodes	PPARG	PVRL4	YWHAZ	FGFR3 S243C	PIK3CA E545K	Clonal Frequency
BC04, BC10, BC11, BC22, BC34	+	+	+	+	+	87.89%
BC04, BC10, BC11, BC34	+	+	+	-	+	6.38%
BC04, BC10, BC11, BC22	+	+	+	+	-	0.00%
BC04, BC11, BC22, BC34	+	-	+	+	+	5.73%

Table 2 - Single-cell tumor clonality analysis of tumors propagated in NSG mice, which are highly immunodeficient (a) and C57BL/6J mice, which are immunocompetent (b), showing loss of clone that lacks PIK3CA E545K expression. 10,282 and 3,666 cells were analyzed from NGS or C57BL/6J mice, respectively.

Changes in heterogeneity of tumors observed in immunocompetent versus immunodeficient mice through single-cell DNA sequencing

The tumor that developed from initial lentiviral transduction of vectors expressing PPARG, PVRL4, YWHAZ, FGFR3 S243C, and PIK3CA E545K was propagated in immunodeficient and immunocompetent mice in order to observe population changes based on immunity. Propagated tumors were then analyzed using the Tapestri Platform.

In immunodeficient mice, all four clones previously observed were still detectable (Table 2a). In contrast, only three clones were detectable in immunocompetent mice (Table 2b). The only clone that lacked expression of PIK3CA E545K was undetectable, suggesting a cell-intrinsic role of PIK3CA in activating mutations in tumor immune evasion.

Subclonal populations from these samples were detectable down to 0.75%, again illustrating extremely high sensitivity of single-cell DNA sequencing using the Tapestri Platform.

Conclusion

The introduction of systematic genetic aberrations to model heterogeneous tumors mixed with single-cell DNA sequencing results in a next-generation functional genomics assay to deconvolute complex genotype-to-phenotype relationships in cancer.

Multiplexed lentiviral transduction of primary mouse urothelial cells enabled interrogation of higher-order genetic interactions in cancer initiation and progression. Single-cell DNA sequencing was able to resolve multiple clones with co-occurring genetic aberrations from a heterogeneous mouse tumor sample at extremely high sensitivity. Sequencing based on lentiviral barcodes enabled robust screening for loss or gain of function of targeted genes. Tumor comparison between tumors propagated in immunodeficient or immunocompetent mice using single-cell sequencing revealed a potential role for PIK3CA in immune evasion.

References

1. https://www.ncbi.nlm.nih.gov/pubmed/28988769

mission bio

tapestri
APPLICATION NOTE

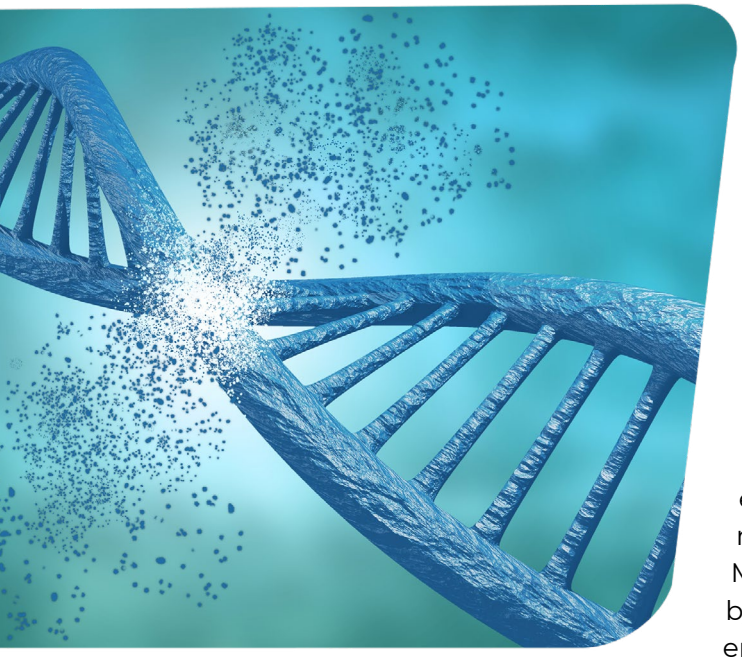
Single-cell DNA Analysis of a Novel Acute Erythroid Leukemia (AEL) CRISPR Induced Mouse Model

Abstract

Acute erythroid leukemia (AEL) is a rare, high-risk form of acute myeloid leukemia (AML) that has a distinct morphology and mutational spectrum. To better understand the biology of the mutations in the progression of AEL, mouse models of AEL were generated using a multiplexed genome editing approach of hematopoietic stem and progenitor cells (HSPCs). Single-cell DNA sequencing using the Tapestri Platform was used to analyze serially transplanted tumors for the precise genome edits induced by CRISPR/Cas9 and for the co-occurrence of secondary mutations. By combining this elegant CRISPR strategy with single-cell DNA sequencing, the investigators revealed the clonal architecture that led to the propagation and evolution of subsequent tumors. The Tapestri Platform is the only high-throughput system that can profile mouse DNA for SNVs and indels at the single-cell level. And now, with Tapestri Designer 2.0, a custom single-cell DNA mouse panel is easily designed to sequence any genomic DNA target across the entire mouse (mm10) or human (hg9) genome, a new feature that will aid researchers in identifying both on and off-target genome editing events.

Takeaways

- A novel mouse model of AEL was generated using a multiplex CRISPR gene editing strategy
- Tapestri Designer 2.0 enables the design of a large panel of mouse genes to determine CRISPR induced and spontaneous secondary mutations in single cells
- Single-cell data revealed the clonal architecture of the fittest clones that led to propagation in subsequent tumors



Materials & Methods

Six lentiviral pools of RFP-gRNAs targeting combinations of genes found to be mutated in AEL (*Trp53, Tet2, Dnmt3a, Asxl1, Rb1, Stag2, Ezh2, Nfix, Bcor, Ppm1d*) were used to infect HSPCs from Cas9-GFP mice, and then transplanted into recipient mice. Analyzing one of the resultant mouse models of AEL, primary tumors with edits in *Trp53* and *Dnmt3a* were serially transplanted into secondary, tertiary, and quaternary mice. Leukemia cells from each transplanted mouse were analyzed with the Tapestri Platform using a custom DNA panel of 30 mouse genes with 75 amplicons directed at CRISPR targeted DNA edit sites and secondary AEL genes of interest. The CRISPR/mouse single-cell DNA custom panel was designed using Mission Bio's White Glove Service, which is now replaced by the newly released Tapestri Designer 2.0 software that enables whole genome coverage for both mouse and human.

Results

Generation of leukemia mouse models using multiplex CRISPR/Cas9

Based on previously published genomic data in AEL[1] (Figure 1a) and to examine the role and cooperativity of the different alterations in leukemogenesis, Dr. Ilaria Iacobucci from Dr. Mullighan's lab at St. Jude Children's Research Hospital used multiplex CRISPR/Cas9-based genome editing to generate tumors in mouse models (Figure 1b). Six different lentiviral pools of guide RNAs (gRNAs) targeting key mutated genes in human AEL were generated and used to transduce Cas9-eGFP-mouse HSPCs. Guide RNAs were designed to generate loss-of-function mutations and edited cells were transplanted into mice and analyzed for tumor development.

Looking more closely at one of the six pools of gRNAs that targeted the epigenetic and tumor suppression genes, *Trp53, Bcor* and *Dnmt3a*, two models of AEL emerged. These results illustrate that mouse models of leukemias were reliably generated using an informed strategy of CRISPR-Cas9 induced co-occurring mutations in HSPCs. Subsequent single-cell DNA sequencing experiments in one of the AEL mouse models (Figure 2 shown in the red box) with mutations in *Trp53* and *Dnmt3a* in the primary tumor, enabled the dissection of clonal heterogeneity and tumor evolution in serial passages in mice.

Single-cell DNA sequencing of AEL mouse model reveals genetic basis for tumor evolution

The evolution of tumors often involves the accumulation of mutations that lead to better clonal survival. To investigate whether this evolution is driven by specific genetic alterations or combinations thereof, the researchers turned to single-cell DNA sequencing of serially transplanted tumors. To determine which mutations were present in the primary tumor, and tumors that evolved through serial transplantation, a single-cell sequencing amplicon panel of 30 mouse genes was designed. Genes were selected based on the genomic locations targeted by CRISPR/Cas9 and recurrent sites of acquired somatic mutations (Figure 3a). Single-cell DNA sequencing not only revealed spontaneously occurring mutations at the clonal level, but also validated edits generated through CRISPR-Cas9 directly from the DNA.

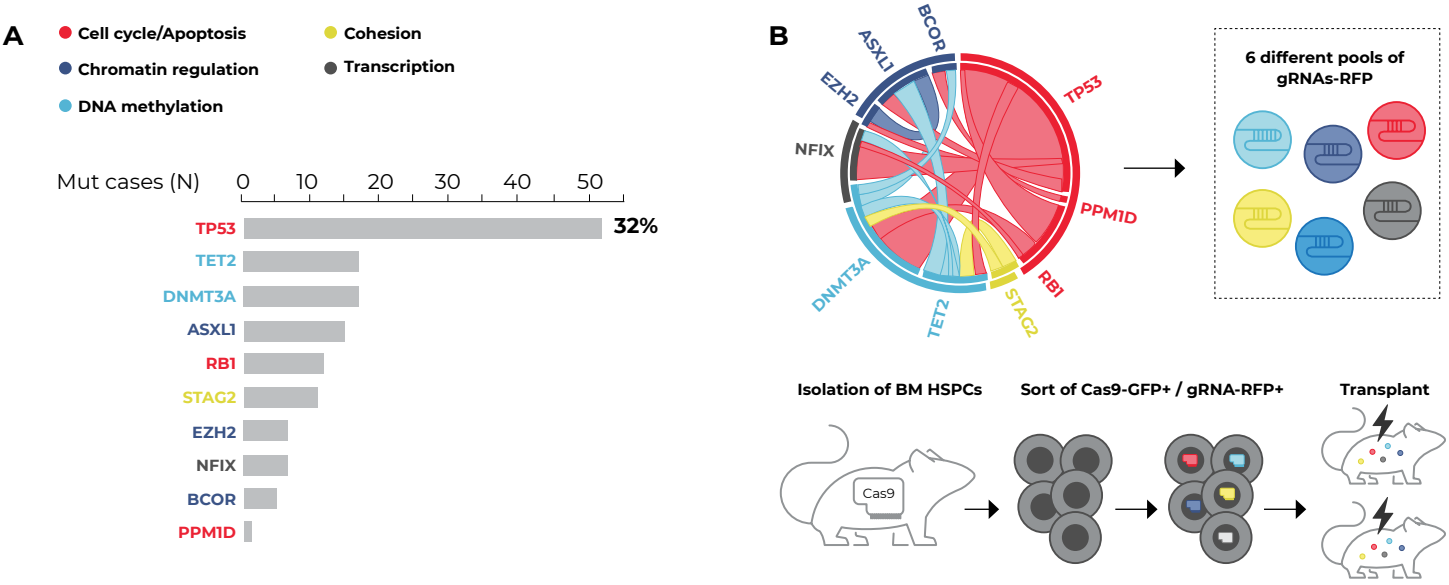


Figure 1 - Frequency of commonly mutated genes in human AEL (A) and pattern of co-occurrence that was used to inform various combinations of RFP-gRNAs for multiplex gene editing of Cas9-GFP mouse HSPCs using CRISPR-Cas-9 (B-upper). Edited HSPCs were sorted for RFP and GFP and then transplanted into mice for subsequent analysis of tumors (B-lower).

Typically, cells containing reporters attached to gRNA and Cas9, in this case RFP and GFP, are assumed to contain both the gRNA and Cas9. However, this alone does not verify that the cells are edited for the desired mutation. In addition, there were many different mutations that were induced through genome editing with the gRNA targeting *Trp53* (Figure 3b). These results highlight the power of using targeted single-cell DNA sequencing to directly confirm the edits that are made to the DNA using genome editing strategies for thousands of single cells.

The primary tumor from the AEL model was serially transplanted into secondary, tertiary, and quaternary mice. Tumor samples were collected at each timepoint and analyzed by single-cell DNA sequencing on the Tapestri Platform (Figure 4a). The primary tumor was found to have two clonal populations, one with mutations in *Trp53, Dnmt3a,* and *Kif1a* (Figure 4b, lime clone) and the other that acquired an additional mutation in a fourth gene, *Ptpn11* (Figure 4b, blue clone). In the secondary tumor, the later clone gained mutations in *Kit, Fancd2,* and *Nudt18*, which persisted as the dominant clone in the tertiary and quaternary tumors. Single-cell DNA-Seq revealed the precise genetic basis for the phenotypic evolution of this mouse model of AEL, confirm the genome edits that were made directly from the DNA, and show that primary and secondary muations in multiple genes co-occurred within the same single cells.

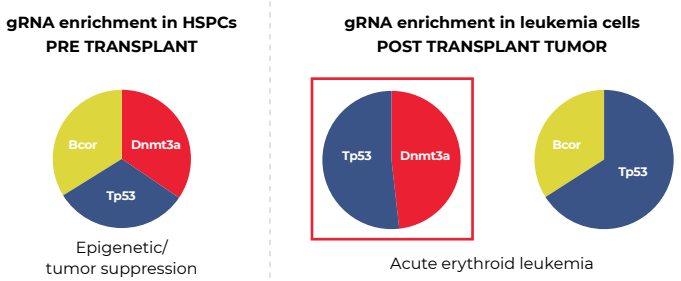


Figure 2 - One of the six pools of gRNAs targeting *Trp53, Bcor* and *Dnmta3* was used to edit HSPCs that were subsequently transplanted into mice (left). Different AEL tumors were induced with different combinations of initiating mutations. The red box indicates the tumor used for serial transplantation expanded upon in Figure 4.

A

30 - Gene Mouse Panel				
Asxl1	Dnah8	Jak3	Notch1	Rb1
Bcor	Dnm2	Kif1a	Nr3c1	Setd2
Cenpe	Dnmt3a	Kit	Nras	Sf3b3
Ddx3x	Ezh2	Kras	Nudt18	Stag2
Dhx57	Fancd2	Lmo7	Ppm1d	Tet2
Dnah7b	Jak1	Nfix	Ptpn11	Trp53

B

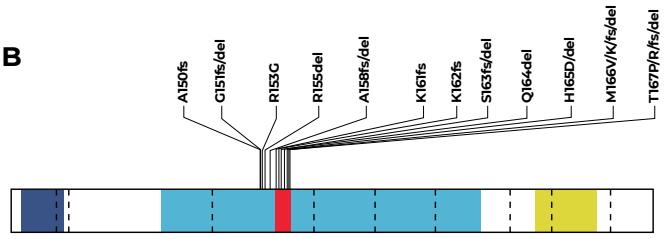


Figure 3 - Panel of amplicons designed to sequence 30 genes related to AEL. Red genes were mutated through CRISPR gene editing in the original mouse models and black genes are of interest for secondary mutations (A). *Trp53* mutations induced by CRISPR-Cas9 illustrates abundance of different mutations generated from the gRNA design shown by the red box (B).

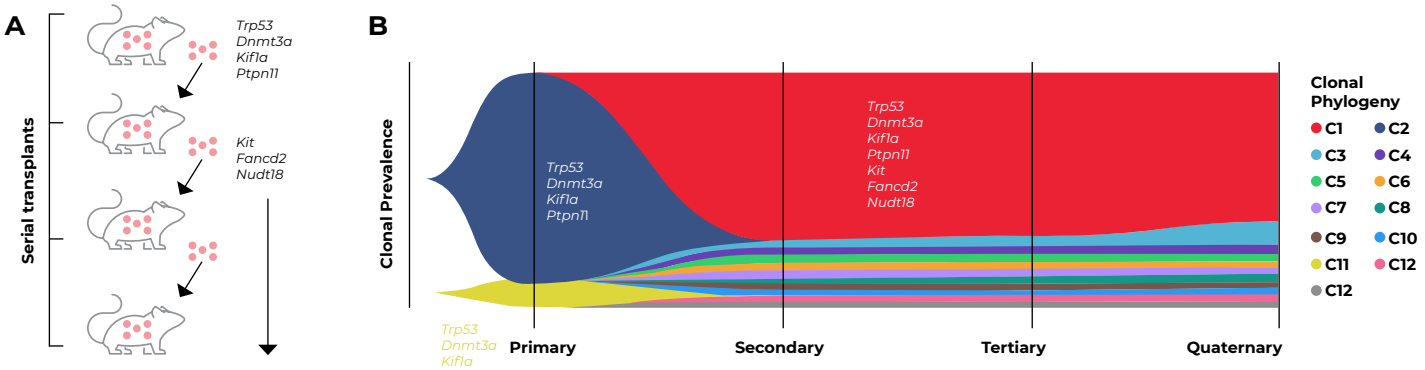


Figure 4 - The primary tumor induced through gene editing of *TP53* and *DNMT3A* was serially transplanted into secondary, tertiary, and quaternary mice (A). SSingle-cell DNA sequencing revealed that the primary tumor also had mutations in *KIF1A* and *PTPN11* giving rise to two clonal populations (B). Evolution within the secondary tumor gave rise to a dominant population with the addition of mutations in *FANCD2* and *NUDT18*, which persisted throughout the tertiary and quaternary tumors.

Conclusion

Mouse models of AEL were successfully generated through multiplex CRISPR/Cas9 editing of genes found to be mutated in human AEL patients. Tumor evolution was precisely visualized through single-cell DNA sequencing, which enabled detection and quantification of many loss-of-function mutations in a large panel of genes, and also determined which mutations co-occur in the same cells to unravel the clonal architecture within each tumor. In addition, edits induced by multiplex CRISPR in the primary tumor were fully resolved at the DNA level and with single cell-resolution and not inferred based on reporters or bulk sequencing. The Tapestri Platform is available for high-throughput single-cell DNA sequencing experiments across the entire mouse and human genomes, and targeted panels are designed in minutes using Tapestri Designer 2.0 software. These fast custom panel designs enable researchers to interrogate gene edits directly from the DNA for any genome editing system, and enable generation of animal models of disease as shown here and in a recent publication[2], as well as support for the development and production of cell and gene therapies.

References

1. Iacobucci, et. al., Modeling and targeting of erythroleukemia by hematopoietic genome editing, Blood, 2021.
2. Ten Hacken, el. al., High throughput single-cell detection of multiplex CRISPR-edited gene modifications, Genome Biology, 2020.

mission bio

tapestri
APPLICATION NOTE

Deep Understanding of Cell and Gene Therapy Genome Editing Protocols Enabled With Single-cell Sequencing

Takeaways

- Replace multiple traditional bulk assays with one single-cell sequencing assay
- Simultaneously identify edits, zygosity, co-occurrence and translocations in up to 10,000 single cells per sample
- Optimize genome editing protocols during development and get superior control during production manufacturing and release testing

Abstract

Cell and gene therapies are yielding new treatments, which were inconceivable only a few decades ago, for genetic disorders including cancers and inherited diseases. Single-cell DNA sequencing on the Tapestri Platform provides the most sensitive and nuanced quantification of genetic edits made to single cells with a simple one-day high-throughput workflow. Quantification of on- and predicted off- target edits for multiple targets, zygosity, and translocations are all measured simultaneously from a single assay for up to 10,000 single cells per sample. After making genomic edits, single-cell sequencing garners richer information than bulk assays, resulting in accurate quantification of all genome editing cell outcomes without the need for multiple cumbersome assays and cell culture. The Tapestri Platform integrates with cell and gene therapy workflows during development as well as during final product testing and release in manufacturing. Having faster turnaround and more complete data at these steps will aid in reducing the time and cost for cell and gene therapies to go to market.

Introduction

Cell and gene therapy is revolutionizing the treatment of many intractable genetic disorders, including hemophilia, sickle-cell anemia, Huntington's disease, and cancer. With these therapies, patients get a reprieve from life-threatening diseases or finally find relief from life-long, debilitating symptoms. Cell and gene therapy repairs genetic errors in a population of cells. After the cells are edited, they are placed in the patient where they either begin replacing the dysfunctional cells or provide an immune response against cancerous cells (Figure 1).

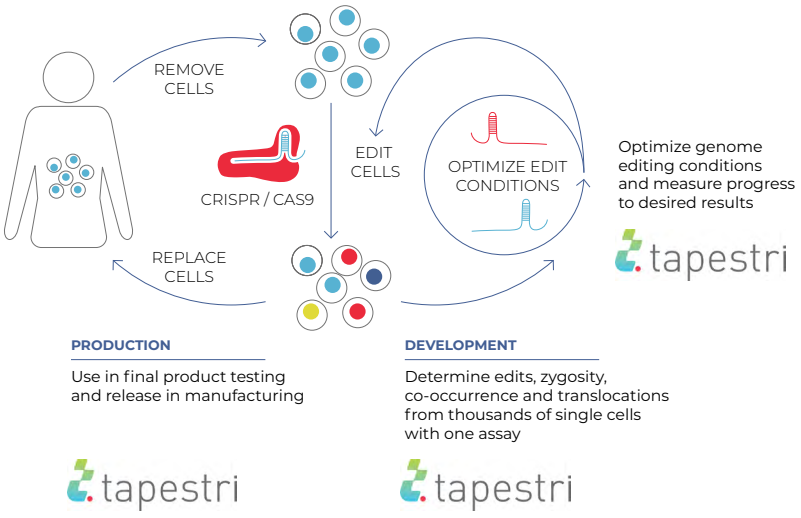


Figure 1 - A schematic of cell and gene therapy during development and production.

Two types of systems are commonly used to edit genomes for cell and gene therapies. The first method uses genome editing tools, such as CRISPR/Cas9 or other similar technologies, to target the gene or DNA sequence of interest and disable the malfunctioning gene. The second uses integration systems, such as viral vectors or homology directed repair (HDR) to incorporate large pieces of DNA into the genome. Some of these edits result in the restoration of normal gene function.

Before genetically modified cells can be placed in a patient, they must undergo rigorous testing to identify both desirable and undesirable edits. For the genome editing approach, careful selection and optimization of genome editing protocols, including the selection and design of guide RNAs (gRNAs), is an important initial step when creating cell and gene therapies. After execution of these genome editing protocols, the Tapestri Platform is used to thoroughly assess the results and quality of the protocol and can be used to help understand the ideal gRNA with the highest editing efficiency. Specifically, single-cell DNA sequencing with the Tapestri Platform identifies expected on-target edits, discerns whether the locus or loci is homozygous or heterozygous, and finds the proportion of cells that have co-occurrence of edits in multiplexing experiments (>1 target per cell). All of these characteristics are important to measure as they affect the efficacy of the therapy. In addition, undesirable events, including off-target edits and translocations, are quantified to determine the efficiency of the gene editing protocol and the possibility of dangerous changes to gene function. It is the selection of the right gRNAs and the characterization of the overall genome editing protocol that is critical for getting cell and gene therapies to market faster.

Following gene editing protocols, cell populations contain a mixed assemblage of different edits. Current methods used to identify genomic edits are both time-consuming and do not provide a detailed understanding of the individual edited cells. qPCR, ddPCR, TIDE (Tracking of Indels by DEcomposition), and ICE (Inference of CRISPR Edits) are the most common techniques used to analyze the genetic edits made to cells. These techniques return bulk data for the entire mixed population of cells or require clonal outgrowth to provide data on individual cells. Alternatively, researchers use karyotyping to identify large chromosomal translocations, but can only obtain information from a handful of cells.

Mission Bio's Tapestri Platform sequences DNA and protein expression in up to 10,000 single cells per sample after cells are edited in cell and gene therapy protocols. Single-cell sequencing provides the most detailed analysis of on- and predicted off-target edits, combination of edits, zygosity and co-occurrence within single cells, and the detection of very low-frequency events like translocations, which can have important effects on the safety of the therapy (Figure 2). This information can then be used to optimize edit conditions during development and then used for final product testing and release in manufacturing.

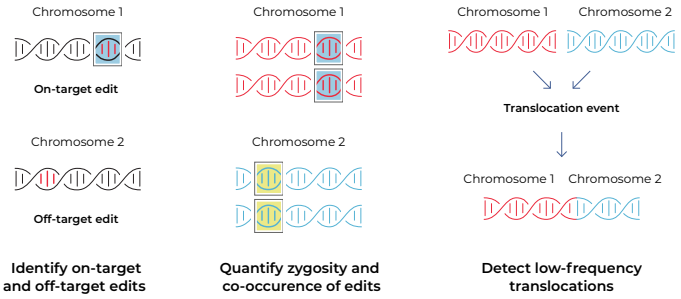


Figure 2 - Single-cell DNA sequencing detects desirable genetic edits, such as on-target events and the zygosity and co-occurrence of edits in a single cell and aberrant events including low-frequency translocations.

Experimental Design and Results

In partnership with Agilent Technologies, three genes, *HBB*, *CLTA* and *RAB11A* were edited with CRISPR/Cas9 in a cancer cell line. Agilent Technologies' CRISPR SureGuide Chemically

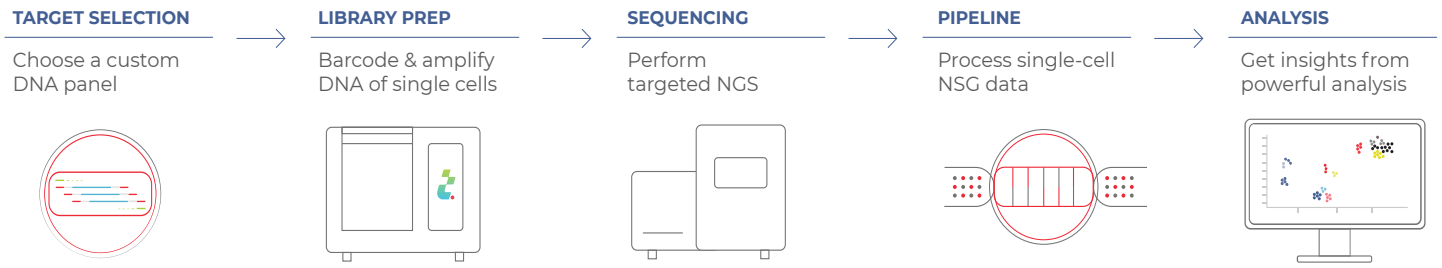


Figure 3 - Tapestri workflow begins with the design of a custom panel, followed by single-cell encapsulation, barcoding, and library preparation. Tapestri software is used for data analysis.

Synthesized sgRNAs were used to ensure a high on-target to predicted off-target ratio of edits.[1] High-throughput single-cell sequencing was performed using the Tapestri Platform and analyzed with Tapestri Pipeline and Tapestri Insights software (Figure 3).

Single-cell sequencing quantified all on- and predicted off-target events for the three genes in the multiplexing CRISPR experiment (Figure 4a). While *CLTA* had only 0.08% predicted off-target edits, the *HBB* gene had 26%, showing the value of quantifying multiple on- and predicted off-target events for several genes in the same assay. This allows researchers to select gRNAs and optimize conditions that increase the proportion of on- to predicted off-target edits for all of the genes simultaneously.

Measuring the zygosity and co-occurrence of edits at multiple loci in single cells helps determine the efficacy of the cell and gene therapy. If the therapy is most effective when all of the targeted sequences have homozygous edits, then understanding the distribution of edits within the population of single cells is essential. Single-cell sequencing

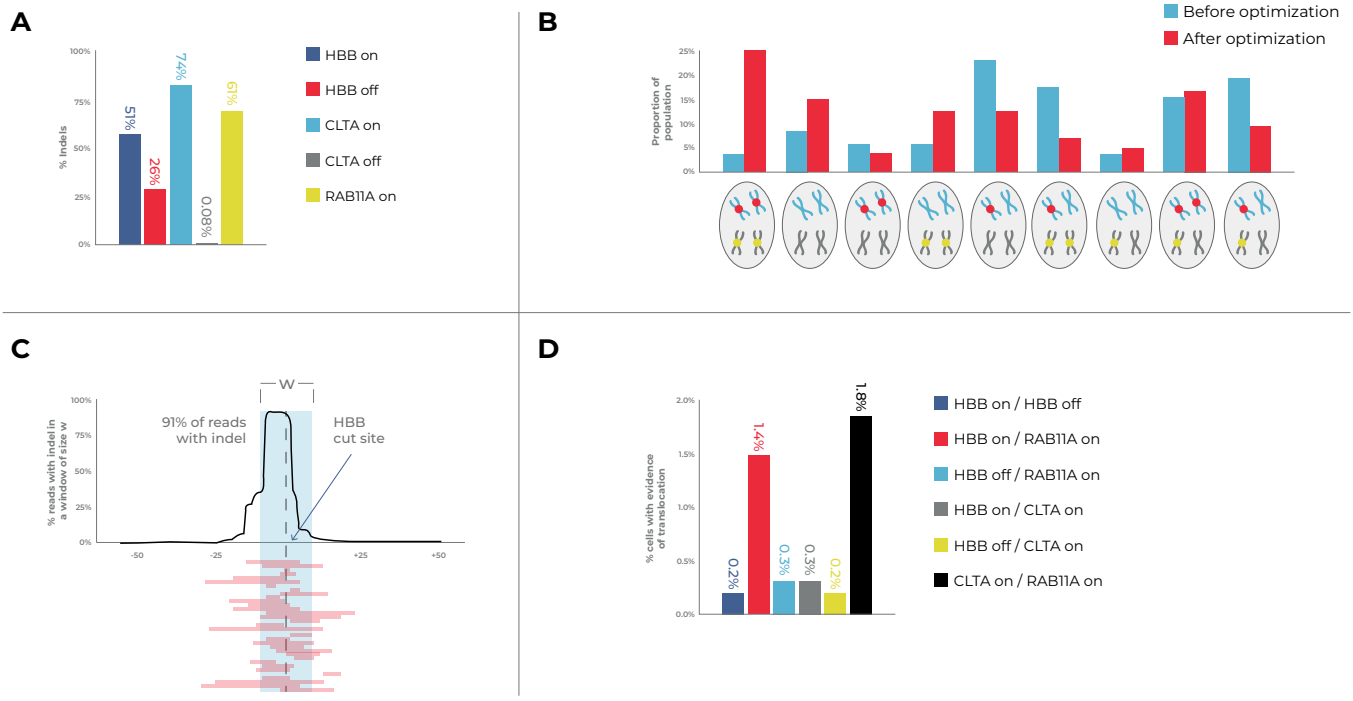


Figure 4 - Simultaenous measurement of on- and predicted off-target edits (a), co-occurrence, zygosity (b), indels (c) and translocations (d) from thousands of single cells.

provides an in-depth quantification of all possible combinations of zygosity and edit co-occurrence, and this information is used to improve CRISPR editing protocols to increase the occurrence of homozygous edits at all target sequences (Figure 4b).

The characterization of CRISPR-induced indels helps define the potential effects of frameshift and in-frame mutations on gene function. The benefits of using a sequencing technology are that the exact indels are measured directly from the DNA, and not extrapolated from reporter molecules. Data showed indel length and location varied considerably around the cut site in the *HBB* gene, providing detailed analysis of exactly what changes and at what quantities were being made to the DNA at the cut site of interest (Figure 4c).

Finally, a translocation in an edited cell can alter gene and cell function in unexpected, deleterious ways, which is a known-side effect of genome-editing but hard to quantify in a high-throughput fashion using traditional karyotyping methods. The ability to measure amplicons that arise from mis-matched primer pairs allows the direct quantification of translocation events, while simultaneously quantifying on- and predicted off-target events, co-occurrence, and zygosity of edits in the same single cells. More complete information is obtained from a single assay concerning the changes to each cell after cell and gene therapy protocols have been employed. Here we show rare translocations between 0.2% and 1.8% that occur in this multi-gene CRISPR experiment.

Conclusion

Single-cell sequencing technology offers exciting new capabilities for the development of cell and gene therapies. By simultaneous assessing desired and undesired editing events at the single-cell level, researchers can better optimize their protocols to build both in vivo and ex vivo cell and gene therapies. When edits were made to cancer cells using CRISPR technology, single-cell DNA sequencing characterized the frequency of on-target and predicted off-target events, the zygosity and co-occurrence of edits, indel length, location, and frequency, and translocations within a populations of single cells. Rare events, such as translocations were identified in as few as 0.2% of the sequenced cells. The high sensitivity and ability to provide comprehensive information from up to 10,000 cells per sample make single-cell sequencing an ideal approach to integrate into cell and gene therapy workflows from development to production.

References

1. Hendel A, Bak RO, Clark JT, et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol.* 2015;33(9):985-989. doi:10.1038/nbt.3290

mission bio

tapestri
APPLICATION NOTE

In-depth Quantification of Cell and Gene Therapy Transduction Efficiency Enabled with Single-cell Sequencing

Takeaways

- Significantly reduce the time between manufacturing and testing by eliminating clonal outgrowth protocols
- Identify transduced versus non-transduced cells with exceptional precision within populations of up to 10,000 single cells
- Rapidly optimize viral transduction protocols during development and measure the quality of transduced cells during production

Abstract

Cell and gene therapies are altering the treatment landscape of intractable genetic disorders, including cancer and inherited diseases. Single-cell DNA sequencing on the Tapestri Platform provides several advantages over other currently used techniques. PCR-based methods (qPCR, ddPCR) measure the average transduction efficiency of heterogeneous populations. Alternatively, colony-forming cells can first be isolated and outgrown over weeks in order to assess the transduction of clonal populations. Single-cell DNA sequencing quantifies the presence or absence of exogenous DNA without the need for lengthy cell culture protocols. The Tapestri Platform provides an unprecedented level of quantification of transduction (or transfection) efficiency in significantly less time compared to bulk approaches and does so at single-cell resolution. These advantages allow for the rapid characterization of cells and optimization of protocols during the preclinical development of cell-based therapies. Tapestri analysis can also can be adapted for lot release testing of therapeutic products.

Introduction

Cell and gene therapies are transforming the treatment of genetic diseases and are rapidly being seen to cure cancers and inherited diseases that were once thought incurable. These therapies provide hope to patients who live with chronic, life-threatening conditions, such as lymphoma, sickle cell anemia, and Fabry disease. Gene-modified cell therapies involve altering the genetic material of cells ex vivo. The modified cells are then placed in the patient and either provide an immune response to diseased cells or proliferate and replace cells responsible for the genetic disease (Figure 1).

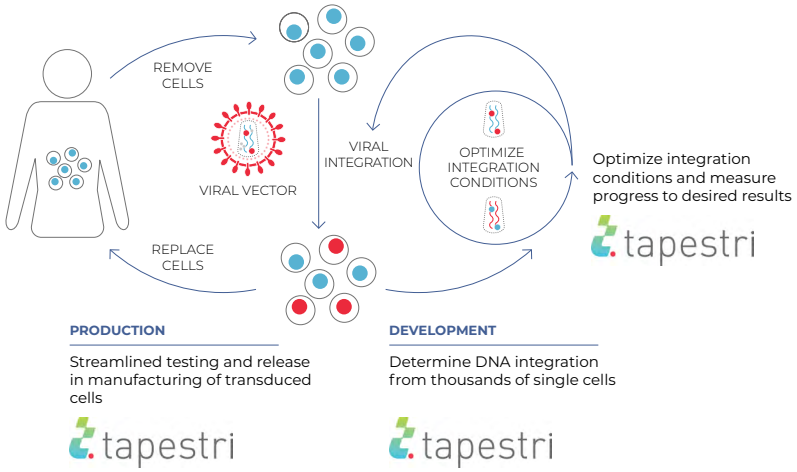


Figure 1 - Single-cell analysis on Tapestri can enable and gene therapy development and production.

Gene transfer is often achieved using viral vectors, such as lentiviruses, to transport genetic material into cells. Some cell therapies, such as CAR-T therapies, rely on the integration of DNA into the host cell genome[1]. Several such therapies have already been approved by the FDA[2], while others are rapidly advancing in clinical trials.

Rigorous testing is required during all stages of development and manufacturing of gene-modified cell therapies. Analytical characterization is conducted during preclinical research to thoroughly describe the alterations that are made to cells. Conventional quantification of transduction efficiency is achieved using PCR-based methods like qPCR and ddPCR. The assessment of heterogeneous populations of infected cells reports a population average of transduction efficiency. Measuring transduction in clonal populations will likely provide a more accurate measurement, but it requires isolating single cells and 4-14 days of clonal outgrowth. During this cell culture, the genetics of the cells in the sample could change via selection. Additionally, the sample size for measuring transduction efficiency may be limited depending on how many plates one has the capacity can monitor during the culture period.

In contrast, single-cell DNA sequencing eliminates the need for lengthy clonal outgrowth protocols, reducing sample processing time from a month or more down to a single week. Moreover, single-cell sequencing provides detailed quantification of viral transduction in 1000s of individual cells (Figure 2).

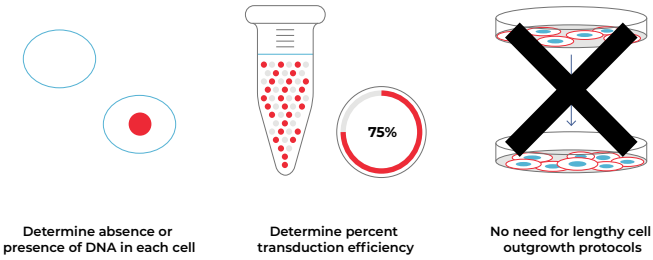


Figure 2 - Single-cell DNA sequencing detects transduced and non-transduced cells.

Experimental Methods and Results

Cells were transduced with lentiviral particles containing a unique genetic sequence and clonal cell lines were produced. The transduced clonal cell lines were serially diluted with non-transduced cells to achieve the following approximate percentages of transduced cells: 0, 25, 50, 75, and 100%. (Figure 3).

Single-cell DNA sequencing on the Tapestri Platform was performed on 5 replicates from each concentration of transduced cells and analyzed with Tapestri Pipeline and Tapestri Insights software. Analysis across the 25 samples reported the percentage of transduced versus non-transduced cells at expected ratios (Figure 4a). For the 5 samples that comprised only non-transduced cells, the false positive rate was below 0.03%. The dilution series thereafter showed excellent linearity and precision among replicates between the

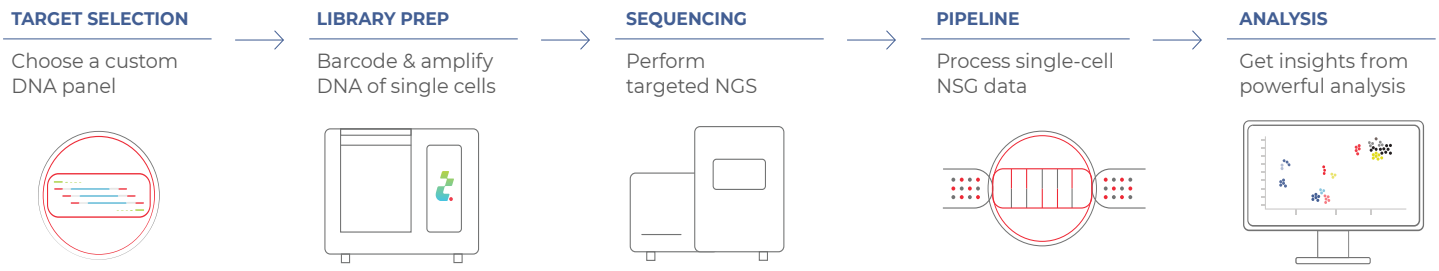


Figure 3 - Tapestri workflow begins with the design of a custom panel, followed by single-cell encapsulation, barcoding, and library preparation. Tapestri software is used for data analysis.

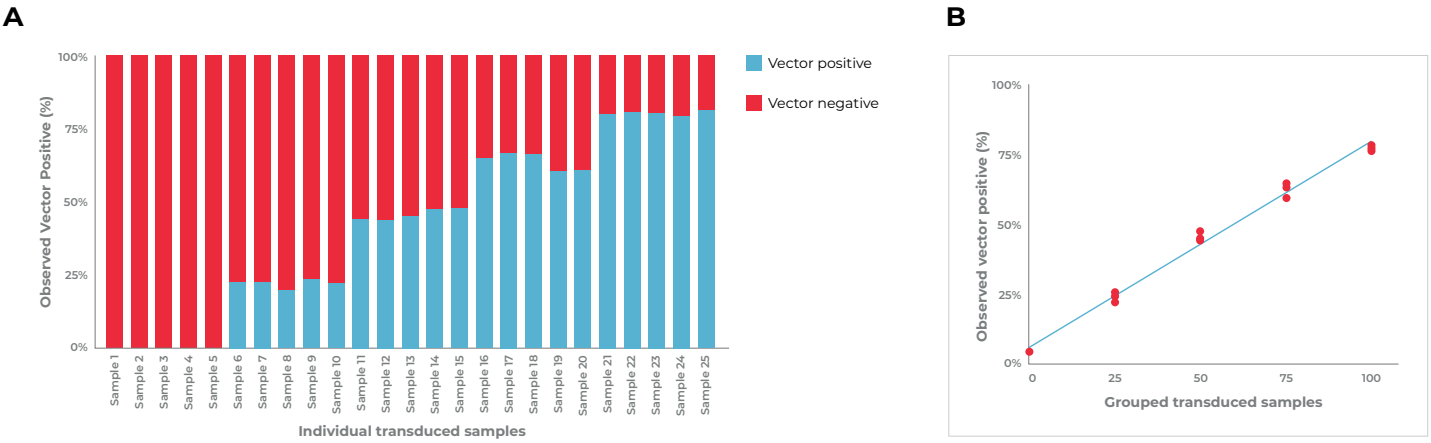


Figure 4 - Assessment of linearity with admixture of transduced cells and non-transduced cells.

expected and observed transduction

percentages (Figure 4b). Overall, these data showcase the ability to use the Mission Bio Tapestri Platform to interrogate 1000s of single cells per sample and garner the percentage of transduced and non-transduced cells in the sample. Importantly, this single-cell analysis was achieved without any lengthy clonal outgrowth or cell culture upfront of the analysis.

Conclusion

Single-cell DNA sequencing technology offers exciting new capabilities for the development of in vivo and ex vivo cell and gene therapies. By precisely measuring the presence or absence of transduction or transfection from thousands of individual cells, researchers can better optimize their protocols and reduce the time to go to market. Tapestri analysis showed a high correlation between the expected and observed percentages of transduced cells and exceptional precision among sample replicates. In addition, single-cell DNA sequencing took significantly shorter time than conventional methods for assessing "single cells" by generating clonal populations.

Single-cell DNA sequencing can provide a higher-resolution picture of the efficiency of gene transfer during the analytical characterization of transduced or transfected cells. In addition, these assays can be adapted to be GMP compliant and used for lot release testing of manufactured products.

References

1. Maude SL, Laetsch TW, Buechner J, et al. Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. *N Engl J Med*. 2018;378(5):439-448. doi:10.1056/NEJMoa1709866
2. Braendstrup P, Levine BL, Ruella M. The long road to the first FDA-approved gene therapy: chimeric antigen receptor T cells targeting CD19. *Cytotherapy*. 2020;22(2):57-69. doi:10.1016/j.jcyt.2019.12.004

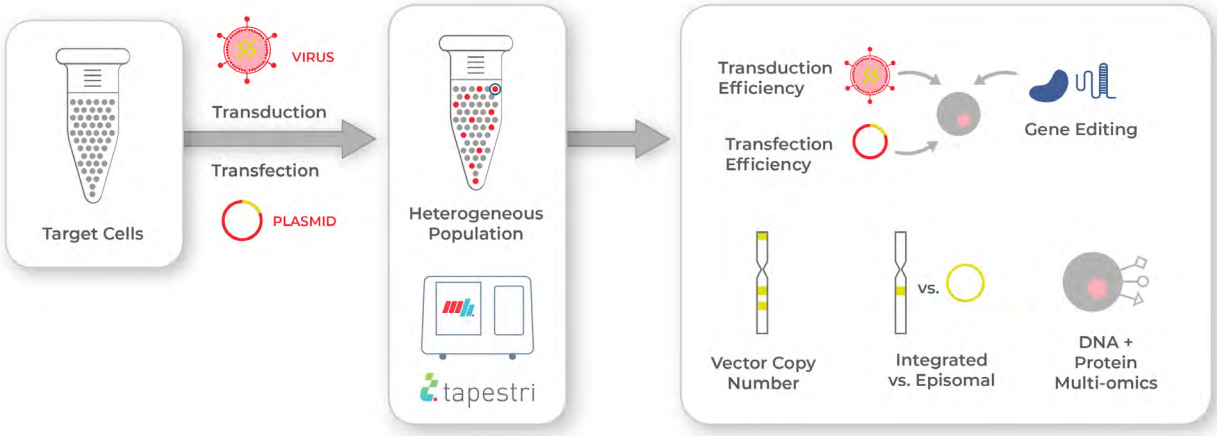
mission bio

Gene Transfer Analysis at Single-cell Resolution

Measure Multiple Attributes in 1000s of Single Cells

Many cell and gene therapies introduce transgenes into target cells utilizing viral and nonviral vectors. The accurate assessment of transduction/ transfection efficiency and the number of vector copies are critical aspects of therapy characterization.

The Tapestri Platform enables multiple genotypic and phenotypic attributes to be co-measured across 1000s of individual cells. This high-throughput platform collapses multiple assays into one workflow, accelerating CMC programs from preclinical analysis to product characterization, and can be adapted for lot release testing.



Introducing genes into cells using viral or nonviral vectors yields heterogeneous populations. Tapestri can measure cell-to-cell variation in transduction/ transfection, gene editing, and vector copy number. Additionally, Tapestri assays can evaluate integrated vs. episomal transgenes and even co-measure cell-surface proteins along with DNA.

Increase the Resolution of Your Characterization Assays

Conventional PCR assays like qPCR and ddPCR measure population averages of transduction/ transfection, and thus fail to report cell-to-cell variability in these attributes. Tapestri can assess transduction/ transfection and vector copy number in a single assay. Moreover, Tapestri's multi-omics assays can simultaneously evaluate genotype and immunophenotype.

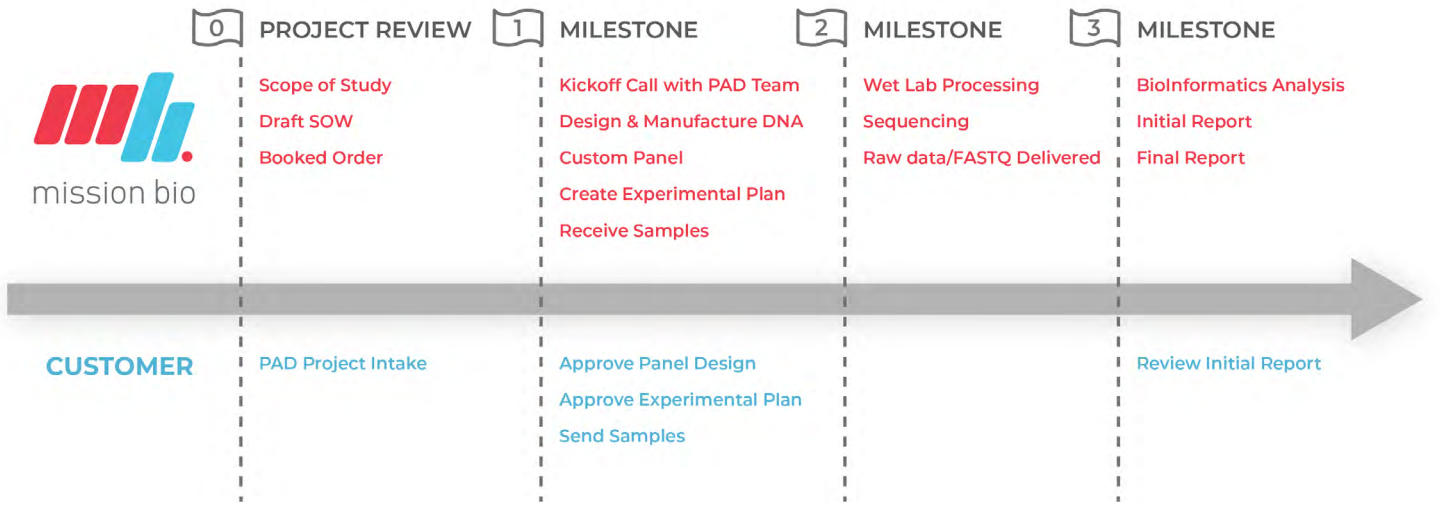
Single-cell Analysis Provides Information that PCR Cannot

	PCR Assays	Tapestri Analysis
Level of measurement	Population	Individual cells
Vector copy number distribution	×	✓
>1 vector in the individual cells	×	✓
Multi-omics in 1 assay	×	✓

End-to-End Service with Pharma Assay Development

Mission Bio offers **custom single-cell assay development** for cell and gene therapies. Our **Pharma Assay Development (PAD)** team provides a full suite of services — from experimental design and sample processing to data analysis and interpretation. For each project, we will provide a comprehensive report with your assay results. We can work with you to transfer your assay to your institution or a qualified CRO/CDMO and support IND filing. Enquire today!

Streamlined PAD Workflow



Project Inputs	Project Outputs
• Vector sequences • Vector maps » Integrated vs non-integrated sequence • Attribute(s) to measure: » Transduction efficiency (1+ vectors) » Transfection efficiency (1+ vectors) » Vector copy number » Integrated vs episomal genome » DNA + protein multi-omics » Gene editing*	• Raw Data Files » FASTQ files from ILMN » h5 files from Tapestri Pipeline • Final Project Report » QC analytics: sequencing & amplicon panel performance metrics for experimental and sample QC » Project data

*See Gene Editing flyer for more information

CONTACT US TO LEARN MORE
info@missionbio.com | www.missionbio.com | +1 (415) 854-0058

mission bio

mission bio

Gene Editing Analysis at Single-cell Resolution

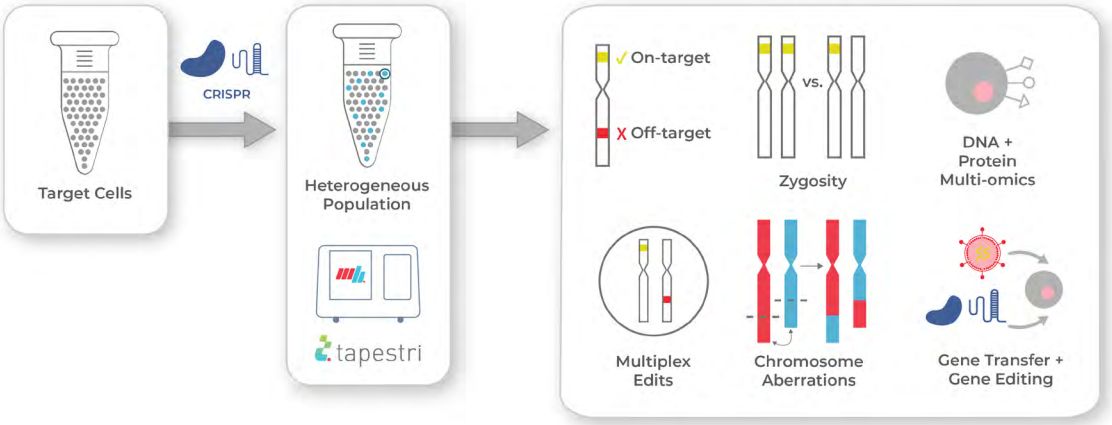
Product Heterogeneity Complicates Characterization

Gene-editing technologies, like CRISPR, are enabling the development of advanced cell and gene therapies. Genome engineering, however, yields heterogeneous cell populations in which cells vary in the edits they contain Moreover, some cells may contain unwanted or even deleterious alterations suchas off-target editing and chromosomal aberrations like translocations.

Tapestri Reveals Gene Edits in 1000s of Single Cells

Single-cell DNA sequencing on the Tapestri Platform provides a high-resolution analysis of gene-edited cells, including on-/off-target editing, multiplex edits, zygosity, and chromosomal aberrations.

For projects that involve gene transfer and gene editing, Tapestri can co-measure these attributes simultaneously in each cell. The platform's powerful multi-omics capability enables the co-analysis of DNA and cell-surface proteins so you can evaluate cell type/ state along with your genomic changes.



Genome engineering yields cell populations with genetic heterogeneity. Single-cell DNA sequencing on Tapestri enables co-measurement of genotypic attributes, including on-/off-target editing, multiplex editing, zygosity, and chromosome aberrations.

Get Higher Resolution Data Than Bulk Sequencing

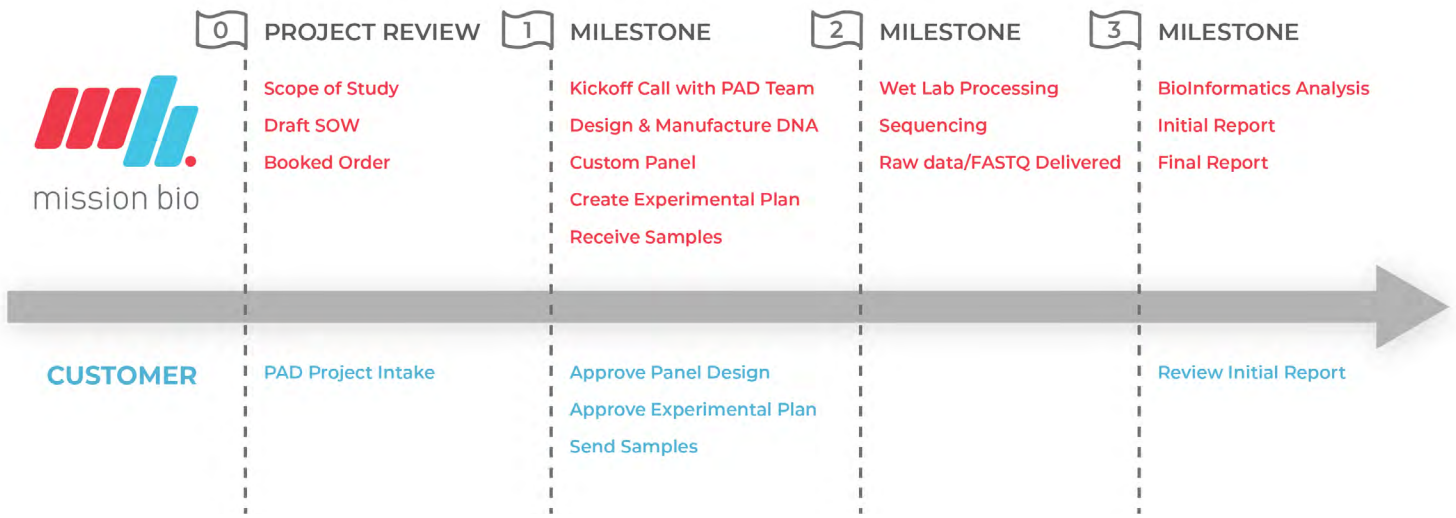
Unlike bulk technologies that report average editing efficiencies, Tapestri provides information from individual cells — enabling attributes like zygosity and editing co-occurrence to be directly measured in 1000s of individual cells.

	Bulk DNA Sequencing	Tapestri Analysis
Co-occurrence of edits	×	✓
Zygosity of edits	×	✓
Gene editing + gene transfer in individual cells	×	✓
DNA + protein in individual cells	×	✓

Easily Incorporate Single-cell into Your Workflows

Mission Bio offers **custom single-cell assay development** for gene editing. Our **Pharma Assay Development (PAD)** team will work closely with you to design an assay that fits your needs. We handle every step of the process — from experimental design to data analysis — and provide a final report upon completion. If desired, we will work with you to transfer the assay to your institution or a qualified CRO/ CDMO. Contact us to learn more about how you can incorporate single-cell analysis of gene editing into your workflow today.

Streamlined PAD Workflow



Project Inputs	Project Outputs
• Chromosomal coordinates for CRISPR target sites • Expected edit(s) » SNV, short indels, long indels, etc • Reference genome • Gene-editing system (e.g., CRISPR) • Amount of padding sufficient for the analysis • Attribute(s) to measure: » On-targets site(s) » Off-target site(s) » Translocations (breakpoint coordinates) » Other chromosome aberrations » DNA + protein multi-omics » Gene transfer*	• Raw Data Files » FASTQ files from ILMN » h5 files from Tapestri Pipeline • Final Project Report » QC analytics: sequencing & amplicon panel performance metrics for experimental and sample QC » Project data

*See Gene Transfer flyer for more information

CONTACT US TO LEARN MORE
info@missionbio.com | www.missionbio.com | +1 (415) 854-0058

mission bio

mission bio BioLegend®

tapestri
APPLICATION NOTE

Single-cell Multi-omics Reveals Novel Correlations Between Genomic Variants and Protein Expression in AML Patient Samples

Takeaways

- The Tapestri Platform is the only single-cell multi-omics platform that measures genotypes and phenotypes simultaneously from thousands of cells
- Simple visualization and data analysis software integrates multi-faceted DNA and protein single-cell data
- The optimized and fully validated 45-plex TotalSeq™-D Heme Oncology Cocktail v1.0 from BioLegend complements Mission Bio's pre-designed and custom targeted DNA panels

Abstract

Cancer is a heterogeneous disease, both genotypically and phenotypically, however, the ability to combine genotype-phenotype relationships at the single-cell level has remained a challenge for researchers. Instead, researchers have relied on bulk measurements from multiple different assays to piece together information for clinical samples. For the first time, we demonstrate, in partnership with BioLegend, a commercially available wet-lab validated protein panel that works specifically with the Tapestri Platform to simultaneously measure DNA variations for up to 1,000 targets and protein expression for 45 targets in thousands of single cells. This novel capability has the power to reveal cell identities and subtle cell states and link genomic variation to protein expression, leading to more informed research on disease and therapeutic development.

Introduction

Acute myeloid leukemia (AML) is a cancer that develops in the myeloid tissue of the bone marrow and has several subtypes with variable outcomes. Various states of leukemic cells, such as leukemic stem cells or progenitors, myeloid blasts, and normal cells, can be defined by cell surface protein markers or morphology. In addition, DNA mutational data and genomic analysis of leukemic cells can define varied clonal populations and lineages. Here, we show for the first time, the ability to combine genotype and phenotype information from the same single cells in order to obtain proteogenomic information with a fast, cost-effective, high-throughput, and wet-lab validated method.

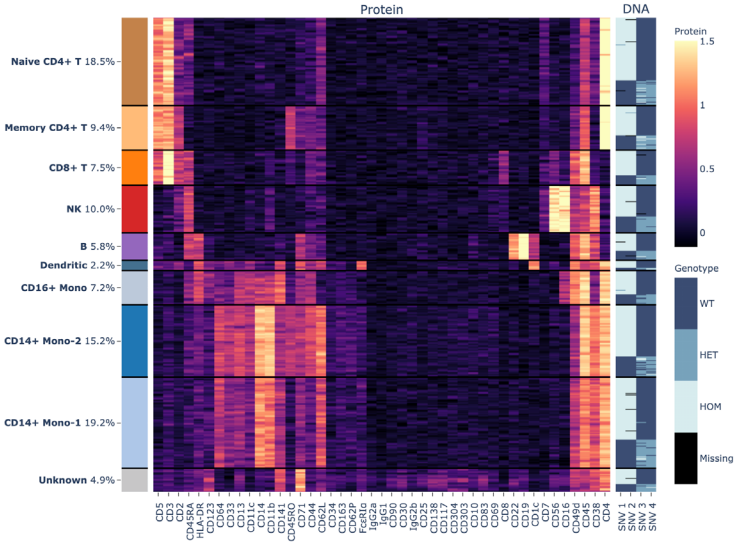


Figure 1 - PBMCs from two donor samples were mixed together. Heatmap visualization reveals the power to obtain genotype and phenotype from the same cells across thousands of cells.

First, we demonstrate in Peripheral Blood Mononuclear Cells (PBMCs) from healthy donors, the ability to cluster thousands of cells by phenotype and genotype using the newly available 45-plex TotalSeq-D Heme Oncology Cocktail v1.0 by BioLegend with the Tapestri Single-cell DNA Myeloid Panel from Mission Bio (Figure 1). Next, to test the system on AML cancer samples and obtain for the first time, proteogenomic information from these samples, we collaborated with Dr. Ross Levine's laboratory at Memorial Sloan Kettering Cancer Center. Six AML patient samples that had previously been analyzed with conventional bulk sequencing were analyzed on the Tapestri Platform for DNA and protein[1]. Each sample was analyzed with a custom DNA panel and a custom oligo-conjugated antibody panel. Over 20,000 cells were sequenced in total, and the data were analyzed with Tapestri Pipeline, Tapestri Insights, and Tapestri Mosaic analysis tools. Overall, the results demonstrate a new single-cell multi-omics approach that integrates genotype and phenotype data for the same cell across thousands of cells.

Materials & Methods

Experiments for DNA and protein were performed using the complete end-to-end workflow that includes a cell staining step with oligo-conjugated antibodies upstream of the Tapestri Platform (Figure 2). PBMCs from two healthy donors were mixed together and analyzed with the Tapestri Single-cell DNA Myeloid Panel and the BioLegend 45-plex TotalSeq-D Heme Oncology Cocktail v1.0. In collaboration with Dr. Ross Levine's laboratory, six AML patient samples were analyzed with a custom DNA panel targeting 32 genes across 109 amplicons together with a custom protein antibody panel for six proteins targeting CD3, CD11b, CD34, CD38, CD45RA, and CD90. Secondary data analysis was performed using Tapestri Pipeline software. Tertiary data analysis was done with Tapestri Insights software and the Tapestri Mosaic analysis package.

Results

45-plex TotalSeq-D Heme Oncology Cocktail v1.0 by BioLegend

Cancer researchers rely heavily on protein markers to identify different cell types and cell states, but they have not been able to combine this phenotypic information with genotypic information on a large scale, particularly in single cells. In conjunction with BioLegend, a 45-plex TotalSeq-D Heme Oncology Cocktail v1.0 was developed that was curated by world leaders in heme oncology across academia and biotech and then extensively wet-lab optimized and validated. Data from a mixture of PBMCs from two healthy donors were clustered on a UMAP plot using the protein expression data from the 45-plex panel with Tapestri Insights software. Results showed distinct clustering of monocytes, CD8+ T cells, CD4+ T cells, NK cells and B cells (Figure 3a) that was verified with cell-specific markers (Figure 3b). Further analysis using the Tapestri Mosaic analysis package to generate a hierarchical clustering heatmap showed even greater resolution of PBMCs into nine cellular subpopulations, including naive versus memory CD4+ T cells, CD14+ versus CD16+ monocytes and

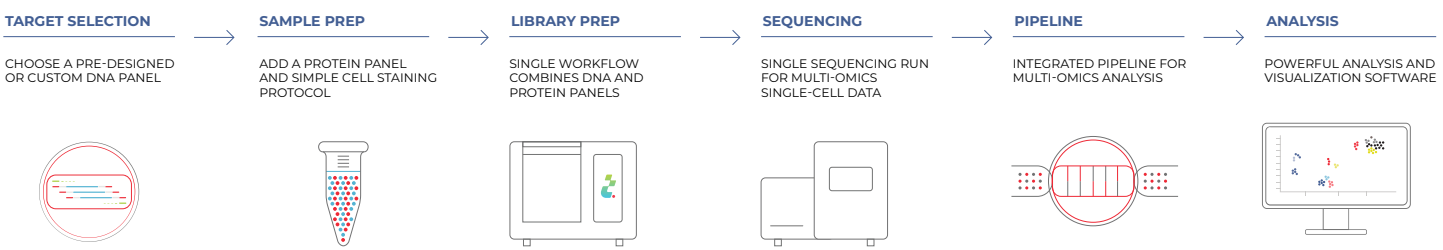


Figure 2 - Multi-omics DNA and protein Tapestri workflow.

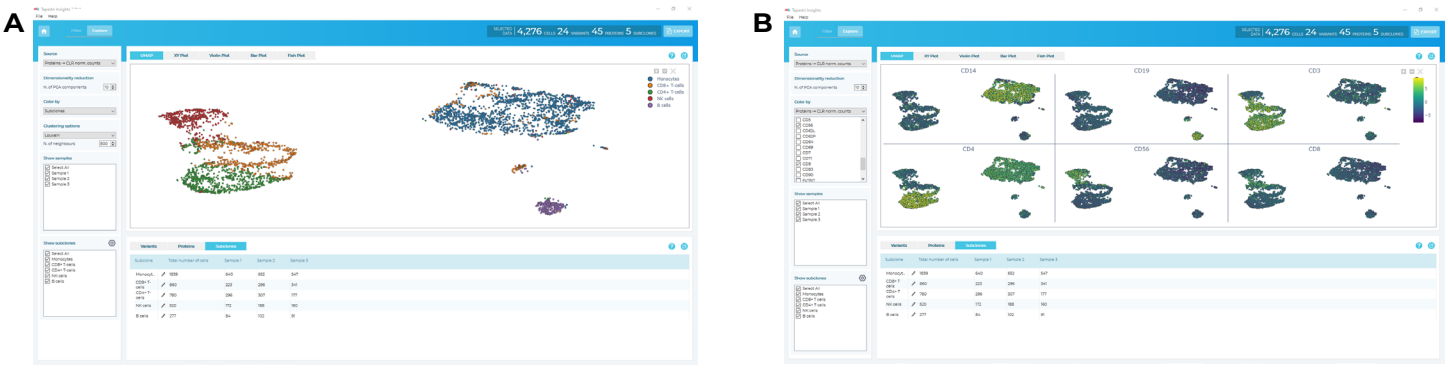


Figure 3 - UMAP plot of PBMCs using the 45-plex TotalSeq-D Heme Oncology Cocktail 1.0 from BioLegend and analyzed with Tapestri Insights software (a-b).

a minor population of dendritic cells (Figure 1). Moreover, protein information was paired with genotype information to identify the PBMCs that came from donor 1 versus donor 2. For the first time, researchers have a large-scale validated protein panel that pairs with large-scale targeted DNA panels to obtain high-throughput single-cell multi-omics insights.

Analysis of an AML patient sample using Tapestri Insights software

While SNVs and indels within clonal populations can be indicative of AML disease status, progression, and relapse, high and low levels of cellular protein expression can be indicative of AML prognosis. An AML patient sample was analyzed using Tapestri Insights software, and the variants *DNMT3A R882H,* and *IDH1 R132C* were identified and selected based on previous bulk sequencing data, their prevalence in the sample, and their likely pathogenicity using ClinVar annotations. UMAP plots were generated using the DNA variant allele frequencies (VAFs) and two clear clusters emerged, one that was wildtype and one that was mutated, containing both pathogenic variants (Figure 4a). Zygosity and co-occurrence status showed that the majority of the mutated cells harbored both mutations in the same single cells in a heterozygous state, while rare populations of homozygous cells were also observed for each variant. From this UMAP plot, protein expression was overlaid, and CD3

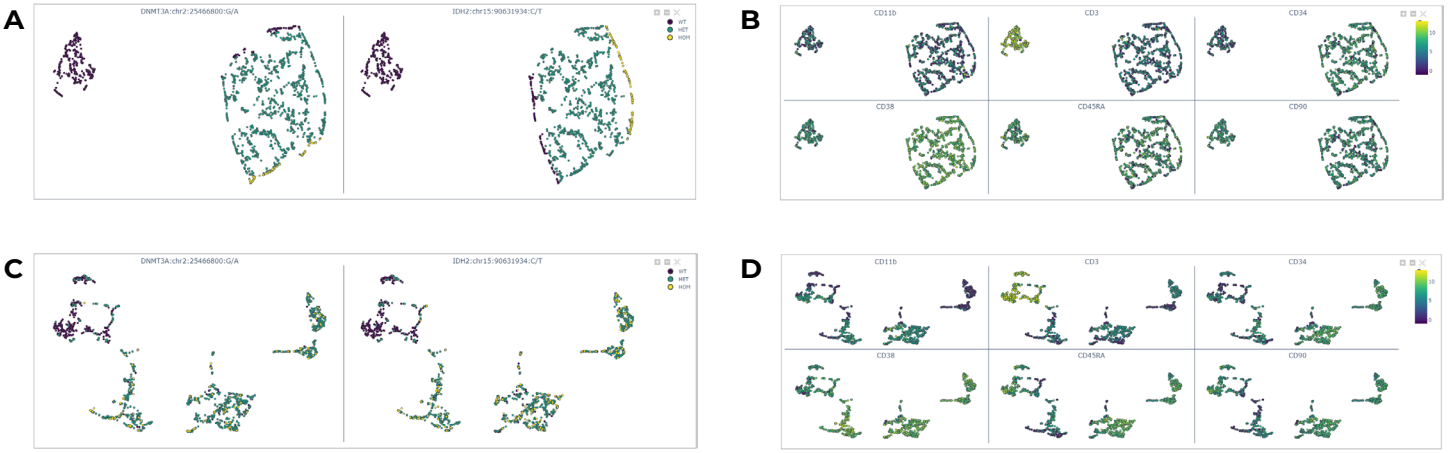


Figure 4 - An AML patient sample analyzed using Tapestri Insights shows UMAP plots clustered by DNA (a-b) or protein (c-d) and identifies distinct cell populations.

expression was observed in the genotypic wildtype population while CD34 and CD38 expression was observed in the genotypic mutated cell population (Figure 4b).

Alternatively, for researchers who want to start with protein data first, UMAP plots can be generated in Tapestri Insights based on clustering from the protein expression. Cells were clustered by the expression of the six protein markers, and the two DNA pathogenic variants were overlaid, showing a clear wildtype population and several mutated populations (Figure 4c). When the protein expression of the six markers was overlaid on the UMAP plot, data showed high, medium, and low expression in the different clusters of cells and revealed how the various combinations of expression profiles for the six markers led to a more refined grouping of the single cells into distinct populations (Figure 4d). Whether researchers want to start first with DNA variants or first with protein expression, Tapestri Insights software supports both analysis pathways and allows researchers to go between the two analyses interchangeably.

In addition to unsupervised clustering, groups or subclones can additionally be defined in Tapestri Insights using k-means or Louvain algorithms. Here we defined two subclones that corresponded with the wildtype and mutated clusters on the UMAP plot colored in by k-means grouping (Figure 5a). In addition to multiple and comprehensive UMAP tools for both DNA and protein analysis, Tapestri Insights also offers additional features for visualizations and analysis. XY scatter plots show two protein markers in relation to each other, giving investigators a readout more akin to fluorescence-activated cell sorting (FACS) analysis (Figure 5b). Multiple violin plots are easily selected that show DNA variant depth of coverage (DP), genomic quality (GQ), and allele frequency (AF), while violin plots based on protein data display inverse

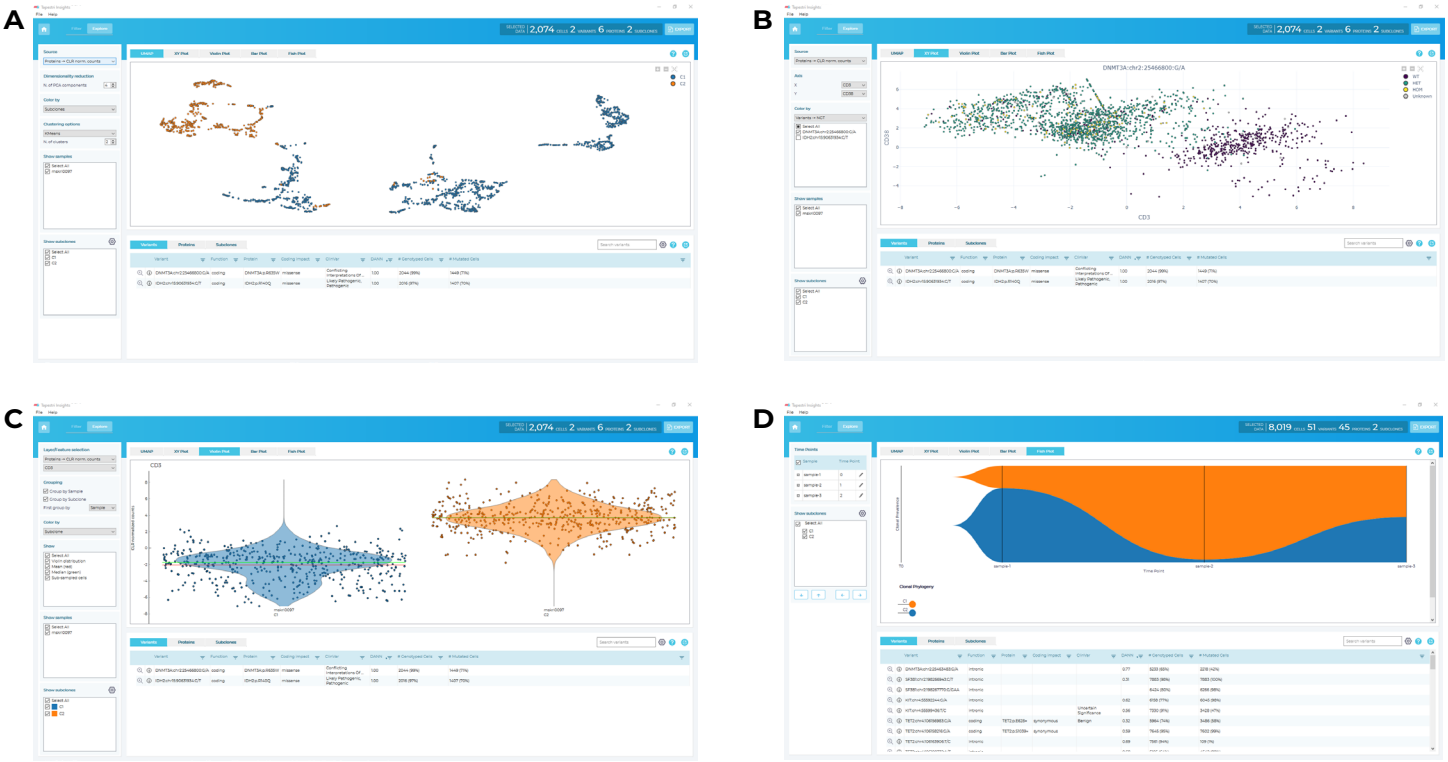


Figure 5 - Multiple analysis and visualization tools in Tapestri Insights software. UMAP plot (a), XY scatter plot (b), violin plot (c), and fish plot (d).

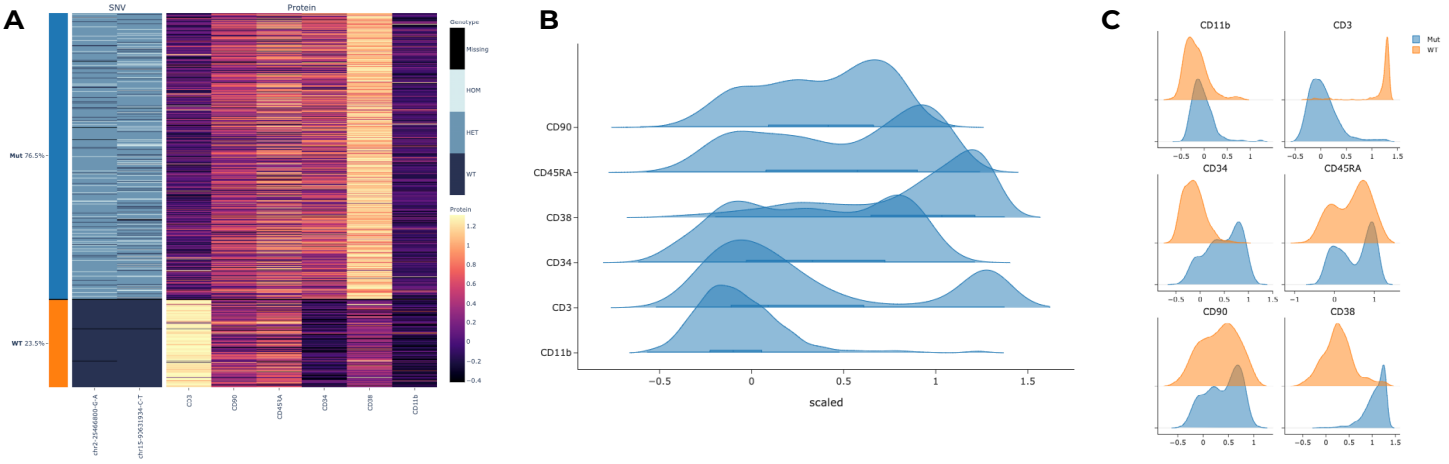


Figure 6 - Heatmap showing unsupervised clustering of an AML patient sample based on SNV (left) and cell surface proteins (right) (a). Ridge plots for each protein marker across the entire population of cells (b). Ridge plots by cell type based on genotype (c).

hyperbolic sine function (Asinh) and centered log ratio (CLR) normalized counts (Figure 5c). Bar plots quantify the number of cells in each subclone while fish plots visualize clonal evolution from time course or multi-sample experiments (Figure 5d). Tapestri Insights software gives researchers a quick and easy way to visualize complex single-cell multi-omics data in multiple formats.

Analysis of AML patient samples using the Tapestri Mosaic analysis package

In addition to the easy and intuitive Tapestri Insights software, more advanced analysis tools are available with the Tapestri Mosaic analysis package of Jupyter Notebooks available on GitHub (https://github.com/MissionBio/mosaic). A heatmap with both DNA and protein data side-by-side visualizes the high CD3 expression in the wildtype cells and the high CD38 and CD34 expression in the mutated cells (Figure 6a). Ridge plots for each protein marker show the distribution of protein expression across all cells (Figure 6b), while ridge plots broken down by genotype reveal the complexity of protein expression within each subclone (Figure 6c). These data exhibit the complex and interesting expression patterns of bimodal and continuum profiles and support a recent report from Dr. Adam Abate

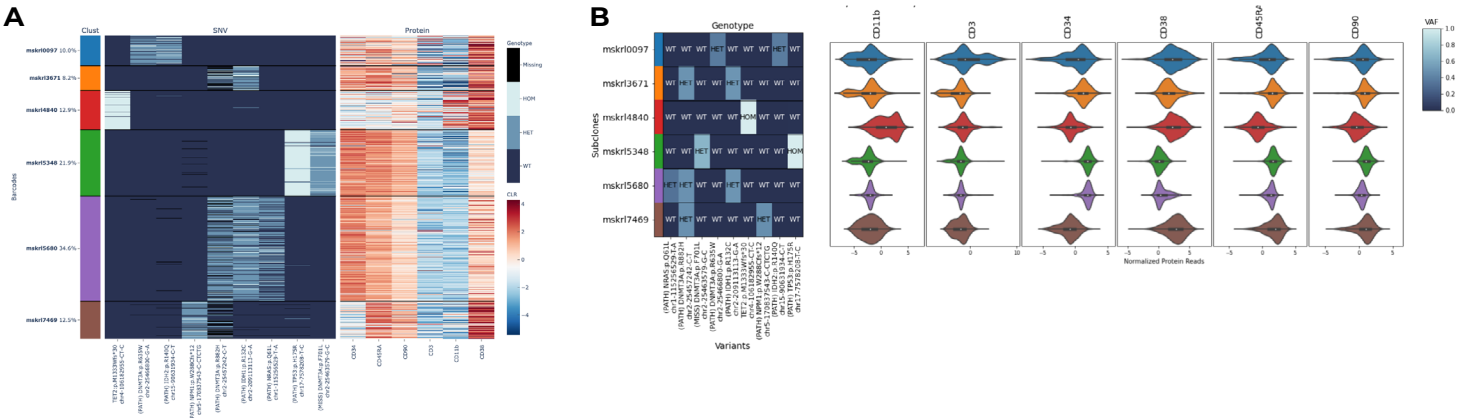


Figure 7 - Heat maps showing six AML patient samples and their SNV (left) and cell surface proteins (right) status (a). A simplified plot of genotype (left) and violin plots of phenotype (right) for each of the six samples (b).

and Dr. Cathy Smith of UCSF that found similar findings in their cohort of AML patient samples across multiple treatment time points and clinical recurrences[2].

Finally, data for the six AML patient samples were merged and simultaneously analyzed for DNA and protein using the Tapestri Mosaic analysis package. Two visualizations tools, heatmaps (Figure 7a) and violin plots (Figure 7b), showed where genotype and phenotype correlated and where they differed for each patient sample as well as between the six different patient samples. These data showed the heterogeneity and complexity within and between each patient sample, and directly linked genetic variation with protein expression in clonal populations using the single-cell multi-omics Tapestri Platform.

Conclusion

The Tapestri Platform has multi-omic capabilities that can determine cell surface protein expression with genotypes simultaneously in thousands of single cells[1-3]. Using a simple cell staining technique, cell identification and subtle cell states are determined, and genomic variants are correlated to protein expression. Here, we validated this new technique on multiple healthy and cancer blood samples. And we present the first commercially available platform from Mission Bio and the first commercially available reagent by BioLegend that detects proteins and DNA simultaneously from the same single cells to unravel complex genotype-phenotype relationships in thousands of cancer cells.

References

1. Miles L, et al. Single cell mutation analysis of clonal evolution in myeloid malignancies. Nature (2020).
2. Demaree B, et al. Joint profiling of proteins and DNA in single cells reveals extensive proteogenomic decoupling in leukemia. BioRxiv (2020).
3. Morita K, et al. Clonal evolution of acute myeloid leukemia revealed by high-throughput single-cell genomics. Nature Communications (2020).