Navigating Heterogeneity in Genome Editing Outcomes Using Single-cell Analysis

Brittany Enzmann[1] , Saurabh Gulati[1], Saurabh Parikh[1], Qawer Ayaz [1], Joanne Nguyen[1], Benjamin Miltz [1], Ben Geller1, Kate Thompson1, Jacqueline Marin[1], Daniel Mendoza[1], Edward Szekeres1, Benjamin Schroeder[1], Shu Wang1, Chieh-Yuan (Alex) Li[1]

1. Mission Bio. 400 E Jamie Ct, Suite 100, South San Francisco, CA 94080

mission bio

BOOTH 737

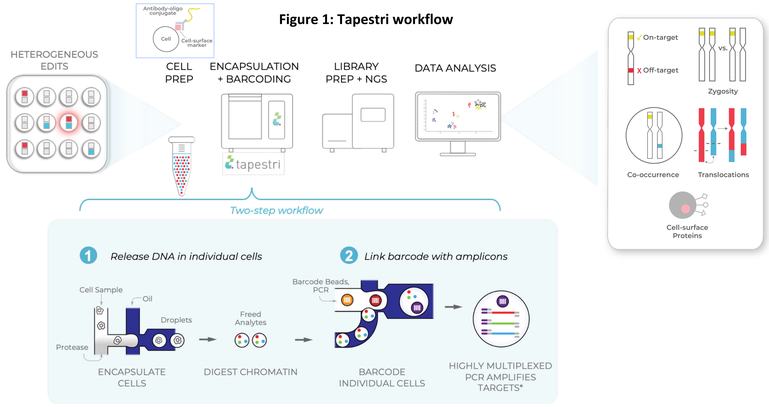
Introduction

Genome editing has emerged as a revolutionary force within the life sciences, wielding transformative potential in applications such as cell and gene therapy development, disease modeling, and functional genomics. Despite the precision of advanced genome editors, editing outcomes remain largely unpredictable. Different cells subjected to the same editing regimen can yield distinct combinations of edits, varying not only across multiple on-target sites but also between on-target and off-target locations. In particular, from the perspective of the fundamental biological unit—a single cell— the zygosity disparity (mono-allelic vs. bi-allelic), heterogeneity in variants (homozygous, heterozygous, compound heterozygous), and their phenotypic effects all contribute to the layer of complexity to the mosaicism of editing's outcomes. Current genome editing analyses primarily rely on bulk methods which provide only an average editing efficiency of a population. The nuanced cell-to-cell variation of edits remains elusive within these traditional approaches. Here, we present compelling evidence that the Tapestri Genome Editing Solution offers a breakthrough in the analysis of knockout (KO) and base editing (BE) experiments. We demonstrate the technology's unique single-cell multi-omics capability to furnish intricate details regarding zygosity and the co-occurrence of on- and off-target edits and immunophenotype, thereby affording researchers the granularity needed for precise experimental outcomes. We also illustrate the capability to multiplex samples via antibody hashing, which allows for economical and scalable analysis while maintaining performance.

Methods

The Tapestri® Platform utilizes droplet microfluidic technology to rapidly encapsulate, process, and profile up to 20,000 individual cells for multi-analyte detection. The platform is enabled by a novel two-step microfluidic workflow and a high multiplex PCR biochemistry scheme (Fig 1). In the first step, individual cells are encapsulated in oil droplets and the chromatin is released using a protease. In the second step, the cell lysate is encapsulated by another oil droplet and merged with a bead containing cell-specific barcodes, and PCR reagents. The targets are amplified along with the barcode via multiplex PCR, followed by library preparation. The final products are sequenced on an Illumina sequencing instrument. For multi-omics projects, the cell sample is initially stained with antibody-oligo conjugates specific for cell-surface proteins of interest. The oligo sequence on each AOC, which identifies the presence of a given protein in the sample, is read out through NGS.

Figure 1: Tapestri workflow



To evaluate the performance of the pipeline for indel calling (for KO applications), we edited Jurkat cells using **CRISPR-Cas9** (Synthego), targeting Programmed Cell Death (**PDCD1**) and T cell receptor α constant (**TRAC**). Both **edited cell pools** with and **isogenic clones** were obtained. The on- and off-targets' edits were verified through bulk NGS and ICE analysis. For based editing applications, where base substitutions are the primary edit types, we assess the pipeline's performance on SNV calls using well characterized cell lines, **GM12878** & **GM24385**, from the National Institute of Standards and Technology (NIST). Assays were designed targeting each genome coordinates of interests.

For single-cell multi-omics analysis, a mixture of PBMC and CRISPR-Cas9 edited (TRAC) samples were stained with a 45-plex oligo conjugated antibodies (Total Seq-D, Heme Oncology panel, BioLegend) folowed by Tapestri workflow. The Gene editing results were analyzed using the Tapestri® Genome Editing Solution Protein + DNA Pipeline.

Results

Tapestri GE Pipeline has high sensitivity, specificity & reproducibility

A total of 13 and 10 different indels and SNVs were assessed in 24 Tapestri runs. Each performance metric was calculated by comparing the target's expected editing status (from known truth) to the target's editing status called by the pipeline. The pipeline performance metrics at a sample level used aggregate counts of events across all cells and targets. The analyses revealed that the pipeline had high sensitivity, specificity, and accuracy for calling both SNVs and indels, as well as low LOD and error rates (Table 1)

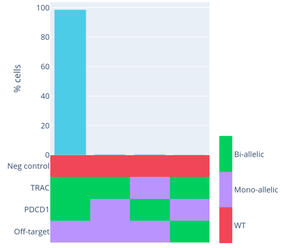
Table 1. Tapestri performance metrics for SNV & indel calling

Variant type	Sensitivity	Specificity	False positive rate	False negative rate	Accuracy	Limit of detection (LOD)
SNVs	98.71% (0.51% CV)	95.27% (0.51% CV)	4.73%	1.29%	95.34% (0.51% CV)	0.10%
Indels	99.87% (1.7% CV)	99.93% (0.03% CV)	0.07%	0.13%	99.93% (0.03% CV)	0.10%

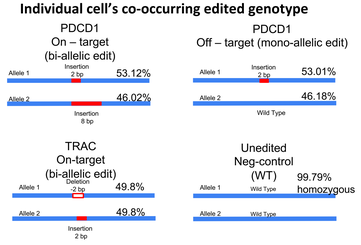
Accurate measurement of edit co-occurrence & zygosity

The accuracy of the Tapestri GE Pipeline was assessed using a **double-edited PDCD1 and TRAC clonal cell line** that also contained 1 off-target edit. The cell line has bi-allelic compound heterozygous edits at both intended targets (PDCD1: +2/+8bp; TRAC: -2/+2bp), as well as a mono-allelic heterozygous off-target edit (+2/0bp) (Fig 2, right). The Tapestri GE pipeline correctly measured the **co-occurrence** of all three targets in the cells (n=2 reps), reporting an avg of **98.6% of cells containing all three edits** (Fig 2, left). An additional locus that was not targeted for genome editing was included as a negative control (WT). The pipeline also reports the frequency and sequence of each variant (not shown).

Figure 2. Analysis of a co-edited clonal line

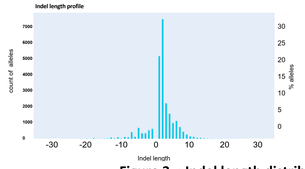


Individual cell's co-occurring edited genotype



The % numbers on each allele indicates the VAF called by Tapestri.

Sensitive report of indel length distribution and predicted editing activity



Pearson's correlation of indel distribution (Log weighted)

Samples	Tapestri rep1	Tapestri rep2	Tapestri rep3	Orthogonal bulk
Tapestri rep1	100.0%	99.9%	99.2%	93.0%
Tapestri rep2	99.9%	100.0%	99.0%	93.2%
Tapestri rep3	99.2%	99.0%	100.0%	94.8%
Orthogonal bulk	93.0%	93.2%	94.8%	100.0%

Sample	Tapestri Results		Orthogonal Bulk	
	INDEL%	KO Score	INDEL%	KO Score
TRAC KO rep1	97.70%	82.97%	92.00%	81.00%
TRAC KO rep2	98.00%	82.69%		
TRAC KO rep3	94.68%	79.15%		

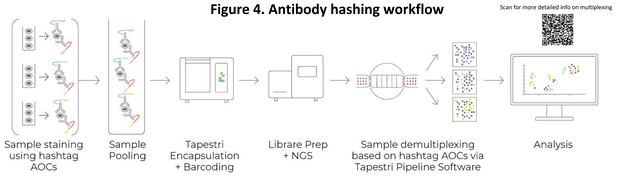
Figure 3. : Indel length distribution, indel % and predicted KO score

The Tapestri GE pipeline reports sample's indel % and indel length distribution with high sensitivity. In Figure 5, left bar plot, illustrates the allelic level indel length profiles of a TRAC edited pool measured by Tapestri. The measurements of indel length profile is highly reproducible with Pearson's r >0.99 (log weighted) amongst Tapestri replicates (Figure 3, top right table). The INDEL% (% alleles with indels) and KO score (this can be defined by user, i.e. frame shift, non-frame shift) can also be calculated. In this particular edited sample, between replicates, the average INDEL% and inferred KO score are 96.79% (std 1.84%) and 81.60% (std 2.13%). The Indel length, indel %, and KO score calculated by the Tapestri GE pipeline corroborated those calculated by the orthogonal bulk analysis (Synthego).

Sample multiplexing by antibody hashing

The Tapestri Platform is compatible with sample multiplexing by antibody (Ab) hashing, enabling up to three samples to be processed in a single run (even if they have the same genetic background). The workflow for antibody hashing can be found in figure 4.

Figure 4. Antibody hashing workflow



- Two antibody-oligo conjugates (AOCs) that bind to common cell-surface proteins.
- The AOCs for each sample have a unique oligo tag that is incorporated into each target during Tapestri processing.
- These oligos are read out during NGS, allowing samples to be demultiplexed during analysis (Fig 4).

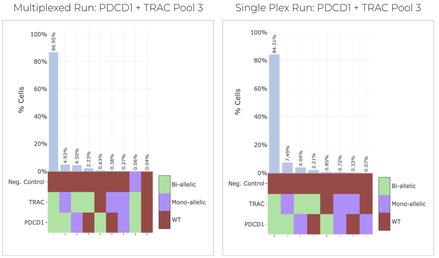
To assess the performance of the GE pipeline with hashed samples, allele calling was evaluated for a multiplexed run (3 samples) and compared to a single-sample run (Table 2). The sensitivity, specificity, and accuracy were comparable to those of a single-sample run, ranging between 98 and 99%.

Table 2 Performance metrics of hashed samples vs single-sample runs

Number of samples	Variant sensitivity	Variant specificity	Variant accuracy
1	≥ 98%	≥ 99%	≥ 99%
3	98.50%	99.83%	99.79%
	98.28%	99.84%	99.78%
	98.65%	99.83%	99.79%
Median	**98.50%**	**99.83%**	**99.79%**

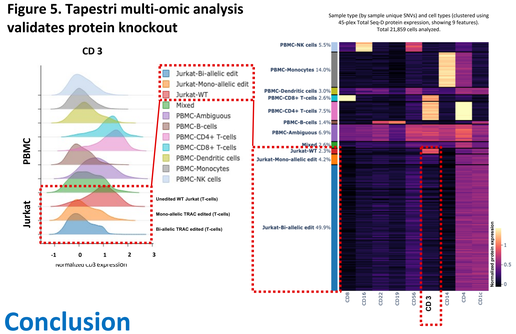
In the right figure, the sample multiplexed analysis has retained its gene editing co-occurrence, zygosity and performance compared to single-sample run.

Multiplexed Run: PDCD1 + TRAC Pool 3 | Single Plex Run: PDCD1 + TRAC Pool 3



Direct measurement of editing genotype and functional KO outcome enabled by single-cell DNA + Protein analysis

Figure 5. Tapestri multi-omic analysis validates protein knockout



The Tapestri GE Protein + DNA pipeline reports sample edit co-occurrence and zygosity along with quantitative surface protein expressions. Analysis of edited genome and cell-surface protein expression enabled the confirmation of a protein-level knockout in edited cells.

For instance, figure illustrates a mix of CRISPR-Cas9 TRAC-edited (KO) Jurkat cells (CD3+) with PBMCs analyzed by the pipeline. Corroborating with prior studies, Tapestri shows that TCRα knockout disrupts the CD3-TCR complex, impacting CD3 surface expression.

Conclusion

We have demonstrated that the Tapestri Genome Editing Solution is a robust and comprehensive tool for addressing challenges in genome engineering in uncovering the inherent heterogeneity of editing outcomes, both genotypically and phenotypically. It empowers researchers to understand the complex landscape of editing variance across thousands of individual cells, with high sensitivity, specificity, accuracy, and reproducibility. Notably, it can detect rare variants down to a remarkable 0.1% of cells, providing unparalleled insight into more sophisticated genome editing endeavors and their safety and therapeutic efficacy. Sample multiplexing using Ab hashing offers both time and cost savings while maintaining high sensitivity and specificity.

Reference 1. DOI https://doi.org/10.1038/mtm.2014.21