mission bio

Decoding the Mosaicism of Genome Editing with Single-cell Multi-omics Analysis

Saurabh Gulaiti[1], Chieh-Yuan (Alex) Li[1], Benjamin Miltz[1], Qawer Ayaz[1], Joanne Nguyen[1], Yilong Yang, Khushali Patel, Brittany Enzmann[1], Vanee Pho[1], Matt Cato[1], Kathy Tran[1], Edward Szekeres[1], Shu Wang[1], Benjamin Schroeder[1]
1. Mission Bio. 400 E Jamie Ct, Suite 100, South San Francisco, CA 94080

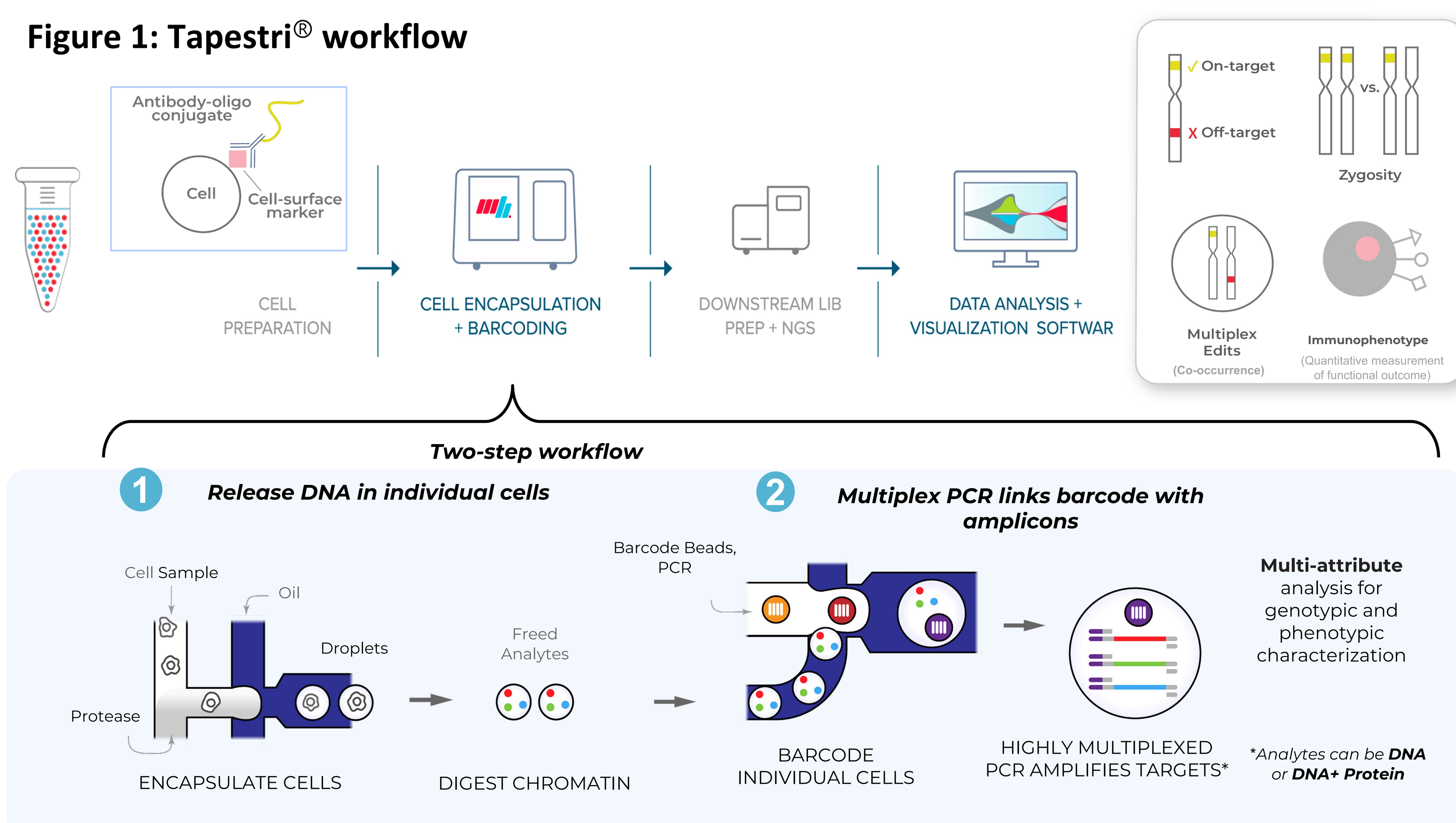
Introduction

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 promise of 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. From the perspective of the fundamental biological unit—a single cell— the zygosity disparity (Mono-allelic, Bi-allelic), heterogeneity in variants (homozygous, heterozygous, compound heterozygous), and their functional impact all contribute to the layer of complexity in the mosaicism of editing outcomes. Current genome editing analyses primarily rely on bulk methods, which, though valuable, provide only an average editing efficiency (at the allelic level) 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 (GE) 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, thereby affording researchers the granularity needed for precise experimental outcomes.

Methods

The Tapestri® single cell DNA 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. The two-step microfluidics allows for efficient access to DNA for downstream genomic reactions and provides flexibility to adapt for additional applications and multi-omics analysis (with oligo conjugated antibody during cell preparation). The multiplex PCR chemistry is developed and co-optimized with an AI-powered panel design pipeline and enables direct and efficient amplification of targeted genomic regions within barcoded individual cells. The final products are sequenced on an Illumina sequencing instrument (Figure 1).

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** and **isogenic clones** were obtained. The on- and off-target edits were verified through bulk NGS and Inference of CRISPR Edits (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 interest.

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) followed by Tapestri® workflow. The gene editing results were analyzed using the Tapestri® Genome Editing Solution Protein + DNA Pipeline.

Results

Tapestri® GE Pipeline has high level of sensitivity, specificity & reproducibility

A total of 13 and 10 different indels and SNVs were assessed between 24 Tapestri single-cell runs. The performance metric was calculated by comparing the target's expected editing status (from expected truth) to the target's editing status called by the pipeline. The pipeline performance metrics at a sample level uses aggregate counts of events across all cells and targets. High **sensitivity** (SNV = 98.71% with %CV of 0.51%; Indel = 99.87% with %CV of 1.7%), **specificity** (SNV = 95.27% with %CV of 0.51%; Indel = 99. 93% with %CV of 0.03%) and **accuracy** (SNV = 95.34% with %CV of 0.51%; Indel = 99. 93% with %CV of 0.03%); as well as low false positive rate (SNV=4.73%, Indel =0.07%) and false negative rates (SNV=1.29%, Indel =0.13%) were observed for all samples with high reproducibility and a limit of detection of 0.1% (0.06%-0.09%).

Figure 2: Tapestri GE pipeline performance metrics for SNV and Indel

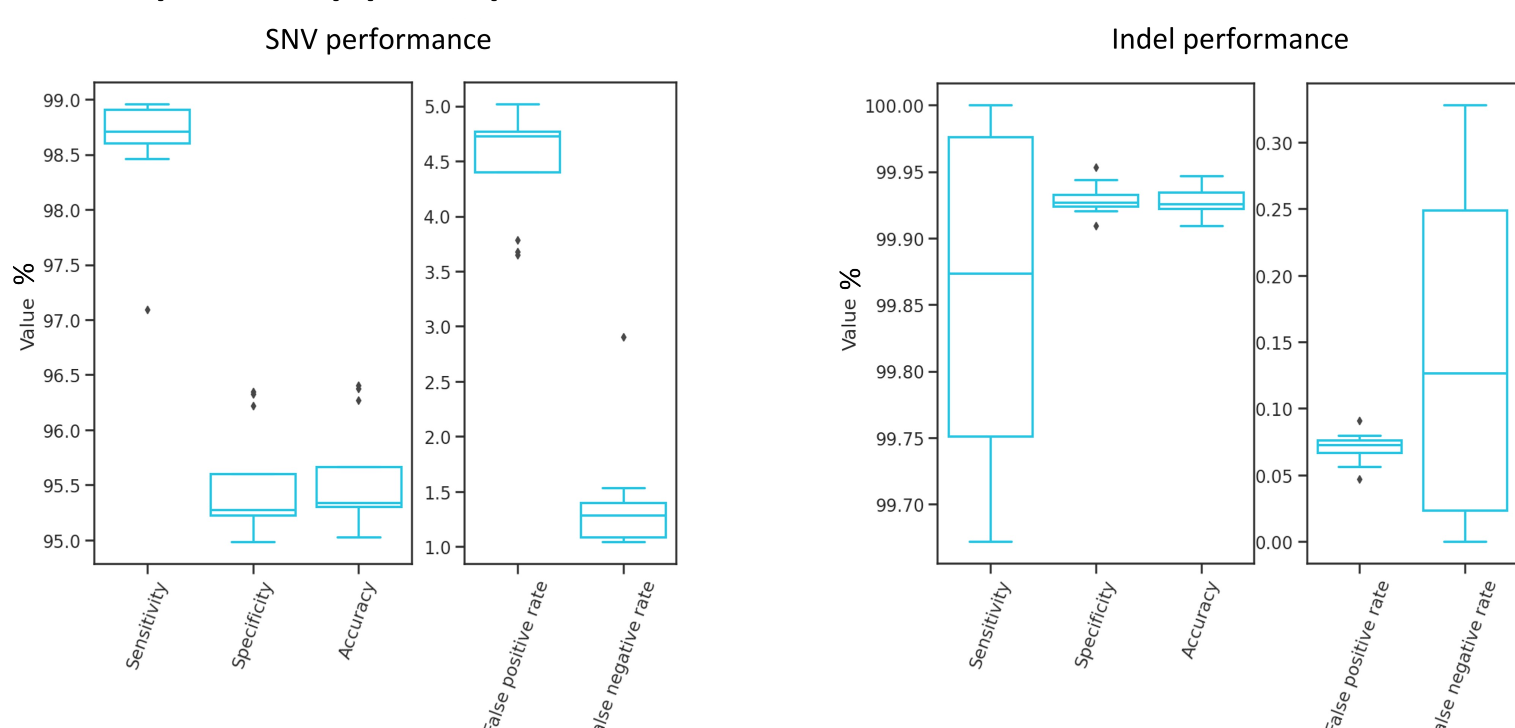


Table 1: performance metrics

Tapestri Performance	Sensitivity	Specificity	False positive rate	False negative rate	Accuracy	Limit of detection
SNVs	98.71%	95.27%	4.73%	1.29%	95.34%	0.10%
Indels	99.87%	99.93%	0.07%	0.13%	99.93%	0.10%

Accurate measurement of co-occurrence and zygosity of edits at cellular level

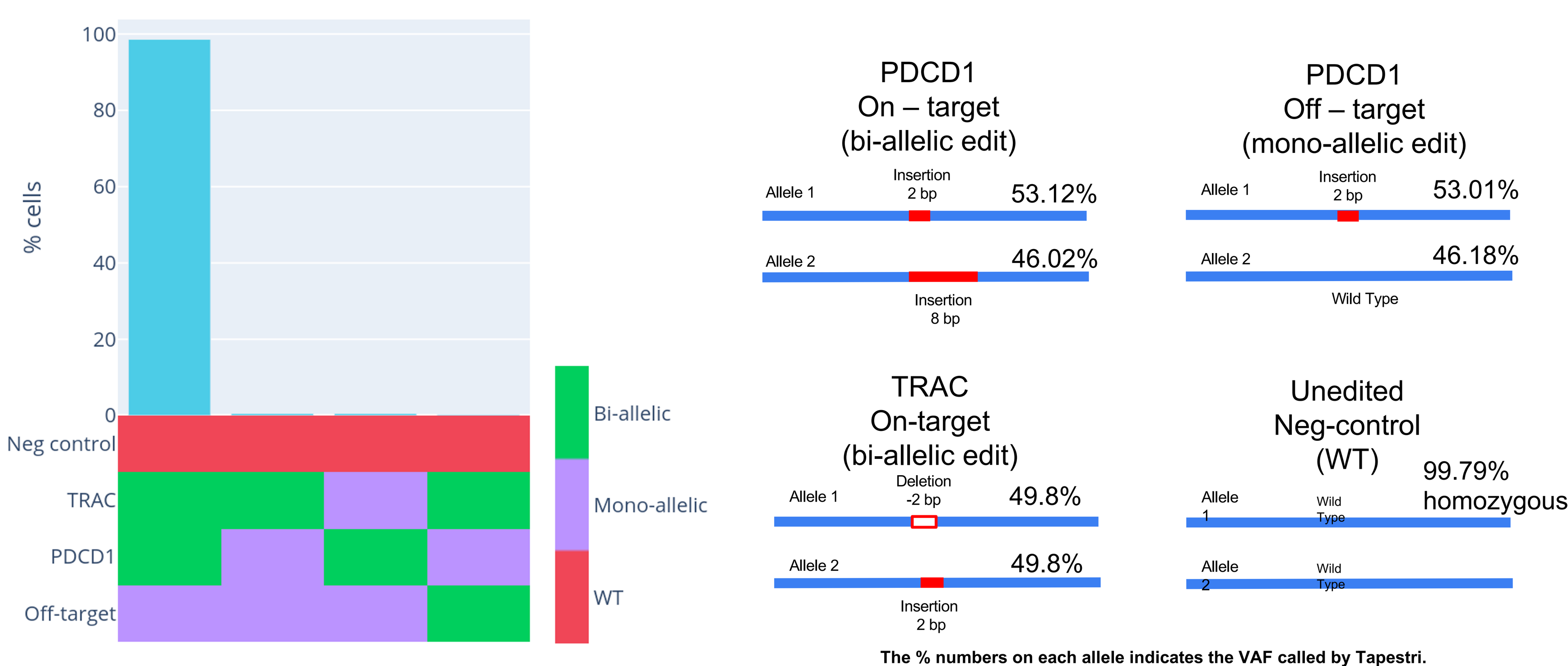
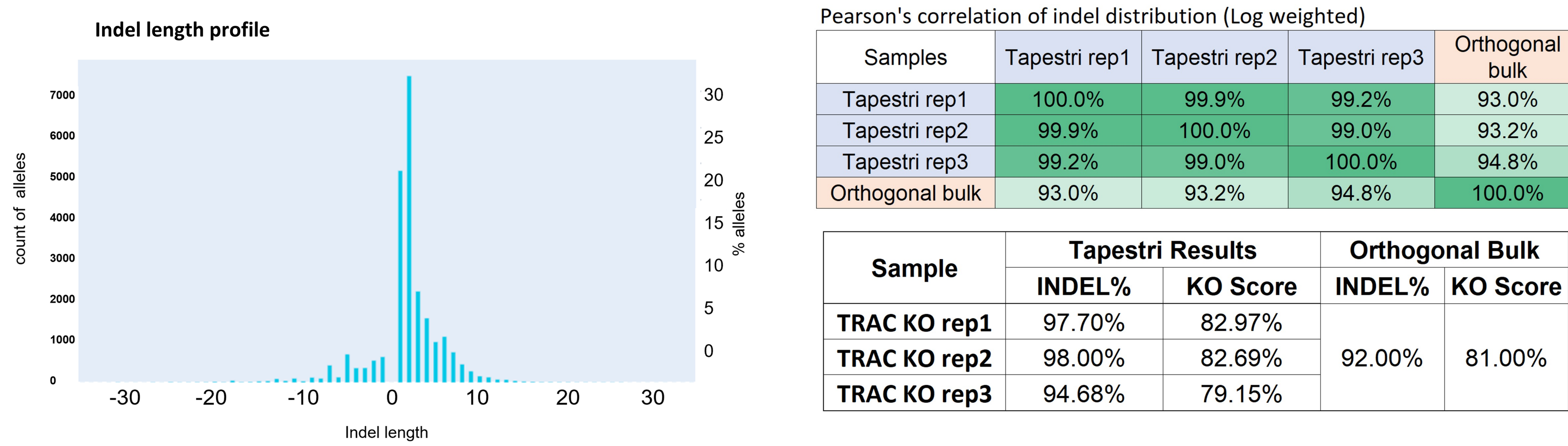


Figure 3: Co-occurrence and zygosity of edits

The accuracy of the co-occurrence of edits was assessed using a 100% double-edited PDCD1 and TRAC with one off-target **clonal cell line**. The clonal cell line (Figure 3, right) has on-target PDCD1 and TRAC edits that are compound heterozygous, bi-allelic (PDCD1: +2/+8bp; TRAC: -2/+2bp), as well as harbors a PDCD1, heterozygous, mono-allelic off-target edits (+2/0bp). The assumption is that all cells would contain the aforementioned co-occurring edits. Indeed, in Figure 3 (left graph), the Tapestri® GE pipeline measured the **co-occurrence** across replicates (n=2) and reported an average of **98.6%** of cells containing both TRAC and PDCD1 on-target edits and the predicted off-target edit. Furthermore, it correctly assessed that the **zygosity** of on-target edits were bi-allelically edited, and the off-target was mono-allelically edited, with the edit indel frequency close to the expected 50:50 range for a diploid cell (46.01%-53.12%, Figure 3, right side, above each allele represents the measured allele frequency). An additional locus that was not targeted for genome editing was included as a negative control (WT). The pipeline also reports the sequence of each indels (not shown).

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%		

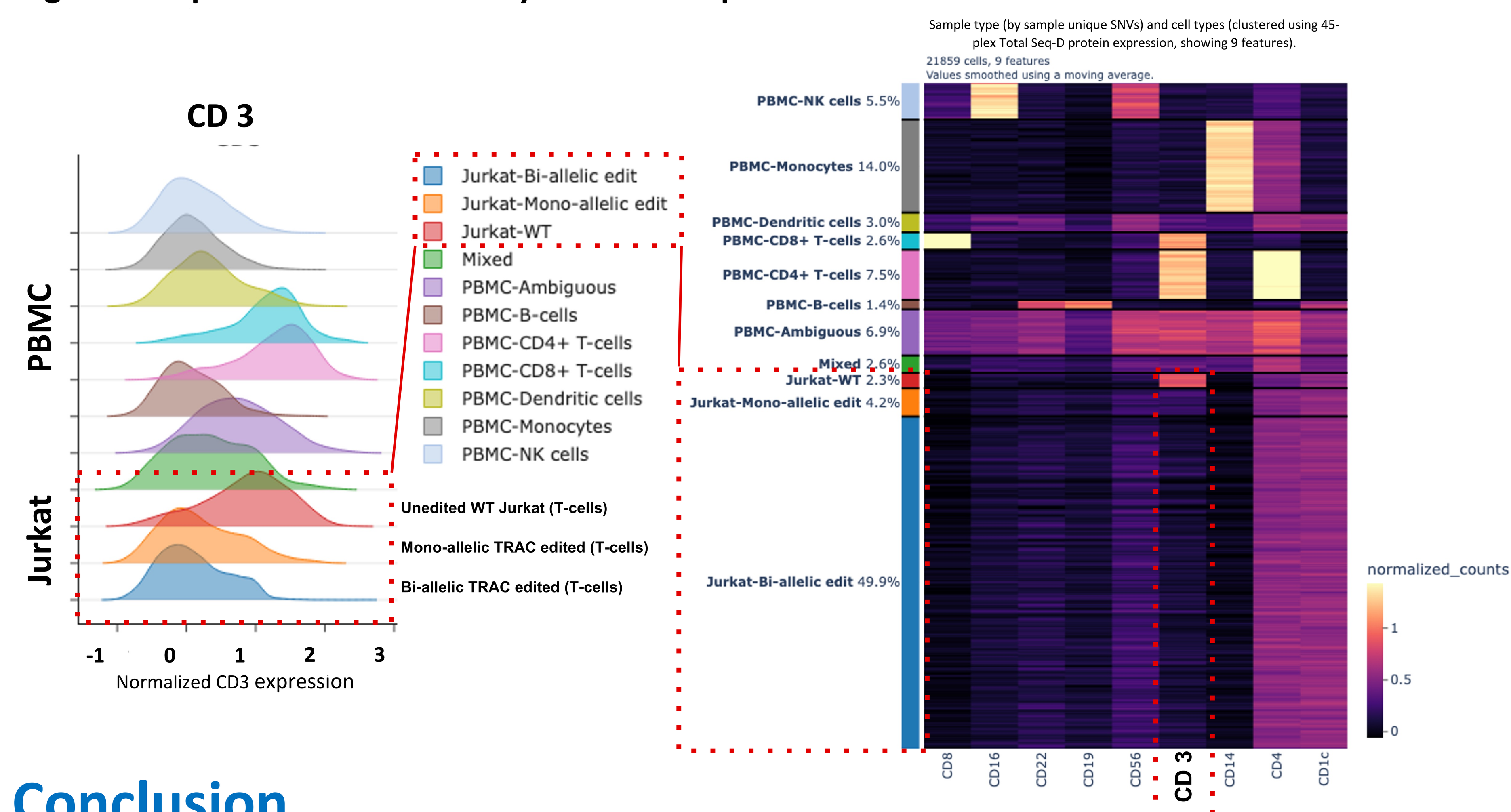
Figure 4: Indel length distribution, indel % and predicted KO score

The Tapestri GE pipeline reports the sample's indel % and **indel length distribution** with high sensitivity. In Figure 4, the left bar plot illustrates the allelic level indel length profiles of a TRAC edited pool measured by Tapestri. The measurements of indel length profile are highly reproducible with Pearson's r >0.99 (log weighted) among Tapestri replicates (Figure 4, top right table). The **INDEL%** (% alleles with indels) and **KO score** (defined as the percentage of INDELs that are either longer than 21bp in length or, shorter than 21bp in length but not a multiple of 3bp that results in a coding frameshift) 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%), respectively. The indel length, indel %, and KO score calculated by the Tapestri GE pipeline corroborated those calculated by the orthogonal analysis (Synthego).

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

The analysis of edited genome and cell-surface protein expression enabled the confirmation of a protein-level knockout in edited cells. Here, we assessed a mixture of TRAC edited Jurkat cells and PBMC with Tapestri DNA + Protein workflow. The Tapestri GE Protein + DNA pipeline reports the sample's edit co-occurrence and zygosity (Figure 3, left), as well as their quantitative surface protein expressions simultaneously. This enables the direct measurement of gene editing outcomes, immunotyping of the samples, and quantitative functional validation of protein expression. To illustrate this, in Figure 5, a mixture of CRISPR-Cas 9, TRAC edited (edit results in Figure 4) Jurkat cells (immortalized human T-cell leukemia cell line, CD3+) are mixed with PBMCs. Indeed, corroborating with previous studies, Tapestri reports TCRa KO disrupts CD3-TCR complex formation and hence impacts surface expression of CD3 [1]. In addition, from the same data set, more sophisticated single-cell resolution studies, such as per cell allele editing sequence bias (frameshift / indel length) vs. functional protein outcome can be performed.

Figure 5: Tapestri multi-omic analysis validates protein knockout



Conclusion

We have demonstrated that the Tapestri Genome Editing Solution emerges as a robust and comprehensive tool for addressing challenges in genome engineering in decoding the inherent heterogeneity of editing outcomes, both genotypically and functionally. It empowers researchers to navigate 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 to more sophisticated genome editing endeavors, and their safety and therapeutic efficacy.

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