High Throughput Single Cell Analysis Workflow for Accurate Measurement of Genotoxicity Arising from Gene Editing Experiments

Saurabh Gulati[1], Chieh-Yuan (Alex) Li[1], Saurabh Parikh[1], Dongjo Ban[1], Alex Majane[1], Qawer Ayaz [1], Benjamin Miltz [1], Madhumita Shrikhande[1], Lubna Nousheen[1], Indira Krishnan [1], Edward Szekeres[1],Nechama Kalter[2], Michael Rosenburg[2], Ayal Hendel[2], Benjamin Schroeder[1], Shu Wang[1]

1. Mission Bio. 400 E Jamie Ct, Suite 100, South San Francisco, CA 94080; 2. Bar-il Bar Ilan University, Ramat-Gan, 5290002 Israel

mission bio

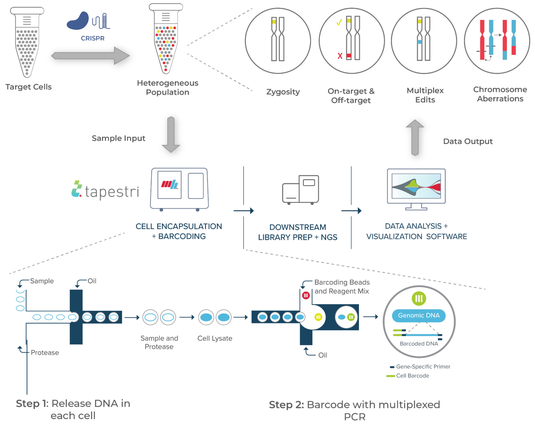
BOOTH 737

Introduction

Despite advancements in precise gene editing technologies such as CRISPR-Cas9, TALENs, and ZFNs, the fundamental editing process yields heterogeneous populations where some cells may have undesired outcomes that bear the risk of genome toxicity. Notably, these adverse outcomes include the introduction of structural variants, copy number alterations, or chromosomal translocations. Therefore, the development of efficacious gene therapies hinges on the ability to accurately measure and understand these events. Furthermore, since "cells" are the functional units of gene editing products, it is prudent to measure the co-occurrences of editing results and potential genotoxicity events in a single-cell context. Here, we demonstrate a microfluidics and multiplex PCR based single-cell technology that, in once assay, **simultaneously measures** the **co-occurrence** and zygosity of **on-target edit**, **off-target edits**, **translocations** between predicted edit sites, as well as the **genomic CNV landscape** in over thousands of cells in parallel. This single-cell technology offers a comprehensive view of the heterogeneous editing profile of gene edited products for a proper and fast evaluation of editing outcome and potential malignant events.

Methods

Figure 1: Tapestri Single cell Gene Editing workflow



The Tapestri® single cell Gene Editing workflow 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 (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 for NGS readout.

In the context of genome editing, the custom PCR panel targets both intended on-target sites and potential off-target sites. The single-cell data are evaluated using the **Tapestri Genome Editing (GE) Pipeline**. This pipeline provides insights into the co-occurrence and frequency of on-target and off-target editing, editing zygosity, precise editing activity per cell and per allele, and cell clonality (Figure 1).

Table1: Tapestri Genome Editing (GE) Pipeline Performance

	Sensitivity	Specificity	Accuracy	FPR	FNR	Limit of detection
Indels performance (Median)	99.77%	99.93%	99.92%	0.07%	0.23%	0.10%
CV	0.55%	0.06%	0.08%	62.26%	129.79%	-

Tapestri GE workflow performance metric Technical note

The performance of the Tapestri Genome Editing pipeline for detecting on-target and off-target edits was assessed using isogenic clonal Jurkat cell lines modified with CRISPR-Cas9 targeting programmed cell death (PDCD1) and/or T cell receptor α constant (TRAC). Each clonal cell line's editing profile was confirmed through bulk NGS, followed by individual processing through Tapestri, with each run including 100% of the cells from a single clone. We used a custom 49-plex amplicon panel targeting both on-target and predicted off-target sites. This setup allowed us to assume that all cells from the isogenic clone contain the identified co-occurring edits. The performance of the pipeline was assessed by comparing the editing status of each potential editing site (either on target or off target), as called by the pipeline, with the known truth for that site. Since the pipeline is capable of calling editing zygosity at each potential editing site, the editing status of each copy of the DNA (allele) was compared. The performance metric is show in table 1 [1].

Results

On-target and off-target co-occurrence, zygosity and protein expression

Figure 2: On-target indel length

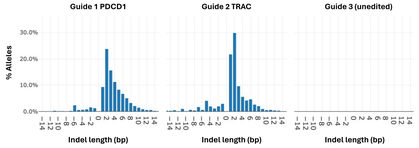
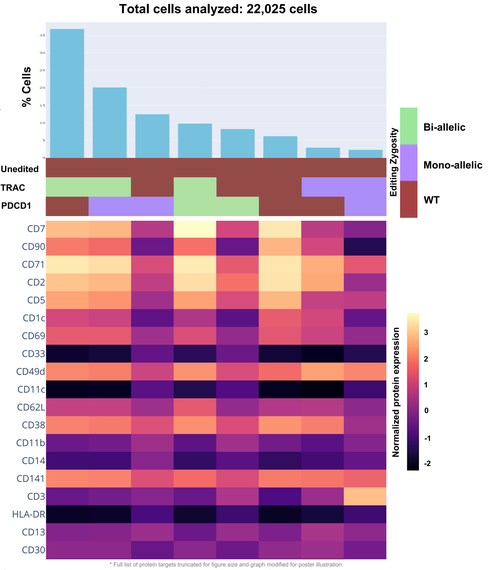


Figure 3: Editing co-coorrence, zygosity and protein expression



Tapestri's single-cell Genome Editing (GE) workflow analysis of edited genomes alongside cell-surface protein expression, confirming protein-level knockouts in edited cells. In this study, we evaluated a mixture of PDCD1 and TRAC -edited Jurkat cells and **unedited** peripheral blood mononuclear cells (PBMCs) using the Tapestri DNA + Protein workflow. For the on-target analysis, by looking at the allele level analysis, we see both PDCD1 (guide 1)and TRAC (guide 2) produces heterogeneous edits with various indel lengths (Figure 2). A separate loci that does not have a guide was included in the analysis as a WT for unedited control (guide 3), where we see no indels. From single-cell level analysis, in figure 3, each cell's SNV, indel and zygosity were determined individually with the sample wide information reconstructed. From the report, we can appreciate the full extent of the mosaicism of editing outcome.
Furthermore, the Tapestri GE DNA + Protein pipeline reports a sample's edit co-occurrence and zygosity, along with their corresponding quantitative surface protein expressions.

This comprehensive approach allows for direct assessment of gene editing outcomes, immunotyping of samples, and quantitative functional validation of protein expression.

Scan for single-cell multi-omics workflow

Single-cell Translocation detection

Figure 4: Translocation detection schema

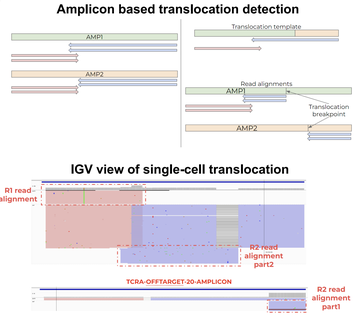


Table 2: Tapestri (single cell) vs. rhAMPseq (bulk) performance

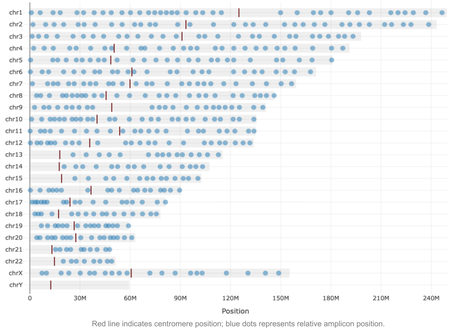
Translocation pairs		Single cell translocation call cell number (% representation in sample)			
Target1	Target2	Tapestri replicate 1	Tapestri replicate 2	Tapestri replicate 3	Mock sample
Guide 1_ON target1	Guide 2_ON target	4 (0.09%)	6 (0.14%)	2 (0.05%)	0
Guide 1_ON target1	Guide 3_OFF target-40	0	2 (0.05%)	0	0
Guide 1_ON target1	Guide 3_ON target-1	23 (0.54%)	22 (0.53%)	27 (0.64%)	0
Guide 2_OFF target-38	Guide 3_ON target-1	0	0	1 (0.02%)	0
Guide 2_ON target	Guide 3_OFF target-51	2 (0.05%)	3 (0.07%)	2 (0.05%)	0
Guide 3_OFF target-127	Guide 2_ON target	0	0	0	0
Guide 3_ON target-1	Guide 2_ON target	17 (0.4%)	9 (0.22%)	10 (0.24%)	0
Guide 3_ON target-1	Guide 3_OFF target-40	0	0	1 (0.02%)	0

Per-cell translocations between multiple edited sites can be identified using the PCR amplicon schema shown in Figure 4, which generates chimeric (fusion) reads for translocations. In this analysis, we evaluate translocation detection performance using a sample edited with three guides (1, 2, and 3) in primary T cells. The translocation detection was compared to rhAMP-seq on the same samples analyzed by CRISPECTOR[2]. The Tapestri single-cell assay proves to be more sensitive than bulk detection methods due to mutually exclusive droplet amplification of fusion amplicons (table 2: green highlights shows translocations detected by both Tapestri and rhAMP seq, whereas blue highlighted indicates translocation detected by only Tapestri; IGV view demonstrated in Figure 4) and offers a detailed assessment of the frequency of translocation events in the sample.

Genome intergrity: Genome wide copy number analysis identifies both SNV and CNV sub-clonal populatio⁻

Beyond translocations, gene editing can also cause unwarranted copy number changes, leading to large structural alterations at the chromosomal level. To evaluate the genome integrity of genome-edited cells, we developed a genome-wide CNV panel to monitor copy number variations (CNV) and ensure the safety of edited cell products. The base panel consists of 500 amplicons distributed across the genome, with an average spacing of 5.8Mb and a uniformity of 96.7% (Fig 5). The panel also includes a probe on the Chr20q locus, which is commonly altered during stem cell differentiation. This 500-plex CNV panel can be combined with custom genome editing, DNA, or protein panels to assess editing co-occurrence, zygosity, translocations, protein expression validation, and gene editing-related focal and/or large CNV events.

Figure 5: Genome wide CNV panel amplicon location



Red line indicates centromere position; blue dots represents relative amplicon position.

In Figure 6, a mixture of GM12878 (diploid reference), two batches of Raji cell lines with different CNV profiles (determined by bulk whole genome sequencing), and a myeloma cell line NCI-H929 were analyzed using Tapestri with the Genome Integrity panel. The Tapestri CNV pipeline identifies subclones based on distinct CNV profiles, revealing five clusters. Verification with cell line-specific SNVs (Raji and H929) confirms the high accuracy of CNV sub-clonal identification. Notably, in the Raji cell lines, single-cell CNV analysis identifies three distinct subclones (red, blue, and yellow), which are indistinguishable by SNVs. These subclones are distributed differently across the two batches: batch 1 consists of a mix of red and blue clones, while batch 2 contains a mix of red and yellow clones. This contrasts with bulk CNV data, which only reports an average CNV profile across the sample (Figure 6, top). Similarly, the analysis shows that the myeloma cell line harbors distinct subclones defined by their unique CNV profiles.

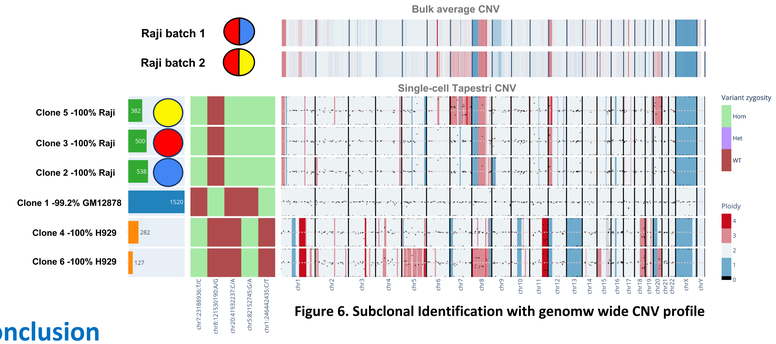


Figure 6. Subclonal Identification with genomw wide CNV profile

Conclusion

The Tapestri single cell Gene Editing DNA + Protein workflow **simultaneously measures** over thousands of cells in parallel the **co-occurrence** and **zygosity** of **on-target edit**, **off-target edits**, **translocations, protein expressions** between predicted edit sites in individual cells, as well as the **genomic CNV landscape**. This single-cell technology offers a powerful and comprehensive view of the heterogeneous editing profile of gene edited products for a proper and fast evaluation of editing outcome and potential malignant events.

Reference
1. Performance of the Tapestri Platform for single-cell targeted DNA sequencing
2. Amit, I., et al. Nat Commun 12, 3042 (2021). https://doi.org/10.1038/s41467-021-22417-4

Scan QR code for information on the Tapestri Gene Editing Solutions.