1149-B

Scaling up plate-based single-cell profiling: introducing automated STORM-seq

Cedric Uytingco[1*], Chris Waddling[3], Kamila Koprowska[3], Arnab Mandal[2], Smriti Chawla[2], Mike Covington[1], Matthieu Pesant[1*], Benjamin K. Johnson[4], Mary F. Rhodes[4], Saloni Pasta[1], Anitha Jayaprakash[1], Michael Phelan[1], Yue Yun[1], Hui Shen[4] & Andrew Farmer[1]

1. Takara Bio USA, Inc., 2560 Orchard Pkwy, San Jose, CA, 95131, USA
2. DSS Takara Bio India, Pvt. Ltd., Jayanagar 4th Block, Bengaluru, KA, 560011, India
3. SPT Labtech, Melbourn Science Park, Melbourn Hertfordshire SG8 6HB, United Kingdom
4. Department of Epigenetics, Van Andel Institute, Grand Rapids, MI, 49503, USA

*Corresponding Authors

Background

Full-length, single-cell RNA-seq (scRNA-seq) allows for more complete transcriptomic profiling compared to 5' or 3' end-counting methods. However, labor-intensive, non-modular workflows with high reagent costs limit scalability. Advancements in laboratory automation, protocol design, and reaction miniaturization have overcome some of these constraints. In collaboration with the Van Andel Institute, we present Single-cell TOtal RNA Miniaturized sequencing (STORM-seq), which pairs Takara Bio's SMART® technology with SPT Labtech's mosquito® HV genomics liquid handling platform. This miniaturized, plate-based scRNA-seq approach is scalable, automated, and offers a solution for single-cell transcriptomic profiling (i.e., coding/noncoding transcripts, isoforms, gene fusions, etc.) at reduced cost and increased cell throughput.

Methods

Single viable K562 cells, peripheral blood mononuclear cells (PMBCs), or renal cell carcinoma cells (RCCs) were sorted into 384-well plates using fluorescence-activated cell sorting (FACS). Following lysis and RNA fragmentation, RNA was reverse transcribed into cDNA. Unique molecular identifiers (UMIs) and unique dual indices (UDIs) were incorporated into the cDNA library following a modified protocol from Johnson and Rhodes et al. (2022, bioRxiv). Miniaturized (1/6) reactions were facilitated using mosquito HV genomics (SPT Labtech). After indexing, pooling, and cleanup, the cDNA libraries were treated with the ZapR™ ribo-depletion technology to remove rRNA-derived cDNAs. The libraries were amplified and sequenced using an Illumina® NextSeq® 500. Sequencing results were analyzed using Cogent™ NGS Software (Takara Bio).

1 STORM-seq with UMIs: a flexible and modular workflow

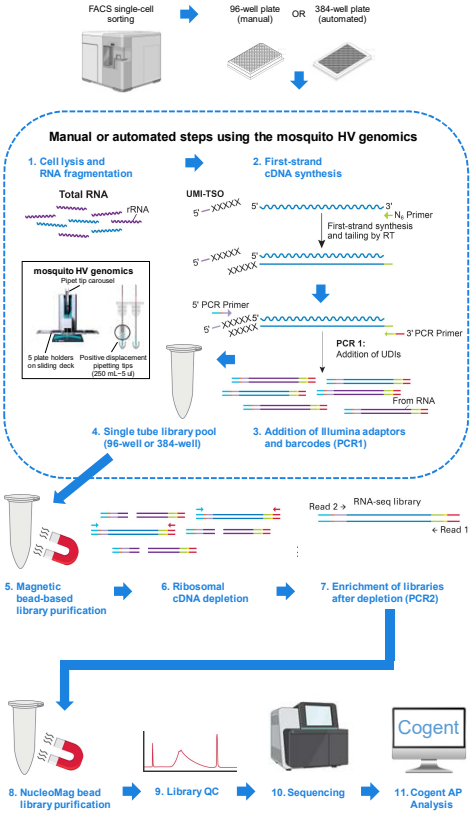

Figure 1. STORM-Seq with UMIs features an easy workflow that generates Illumina-compatible scRNA-seq libraries. Actual processing time is around 8 hours, and it may vary depending on the number of samples, third-party instrumentation, and cycling conditions. The workflow includes multiple safe stopping points: post-cell sorting, after pooling (Step 4), and after PCR2 amplification (Step 7).

2 Comparison of single-cell library recovery using a manual or automated approach

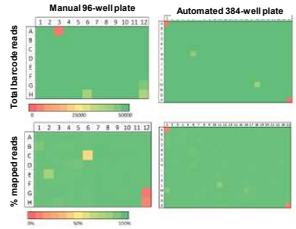

Figure 2. 96-well manual and 384-well automated STORM-Seq reaction plates exhibit >95% recovery of libraries and single cells. Total barcode reads (top) and % mapped reads (bottom) were determined for both the 96-well manual and the 384-well automated STORM-seq protocol. Total barcode reads, and % mapped reads were largely uniform across the reaction plate for both the manual and automated protocol. Negative controls are shown in wells G12 and H12 for the 96-well manual protocol and A1 and P24 for the 384-well automated protocol.

3 Sequencing metrics, UMI and gene detection, and gene body coverage

A

Sequencing metrics		
RNA source	Single K562 cells	
Plate type	384-well	
Final concentration, total yield	137 ng/µl, 2,189 ng	
Replicates	N = 78	
	Mean	**CV (%)**
Mapped reads (%)	98	2
Uniquely mapped reads (%)	90	2
Exonic reads (%)	76	11
Intronic reads (%)	19	41.4
Intergenic reads (%)	5	25
Number of UMIs	53,444	9
Number of genes	10,248	12
Number of transcripts	17,126	16

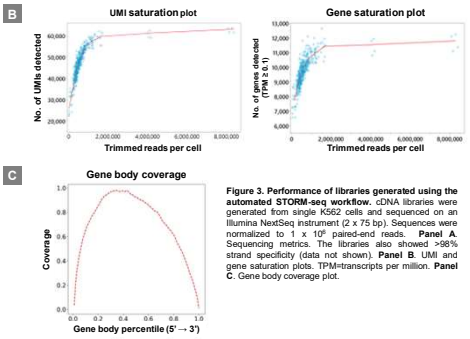

Figure 3. Performance of libraries generated using the automated STORM-seq workflow. cDNA libraries were generated from single K562 cells and sequenced on an Illumina NextSeq instrument (2 x 75 bp). Sequences were normalized to 1 x 10^6 paired-end reads. **Panel A.** Sequencing metrics. The libraries also showed >98% strand specificity (data not shown). **Panel B.** UMI and gene saturation plots. TPM=transcripts per million. **Panel C.** Gene body coverage plot.

4 Gene expression is reproducible between single K562 cells

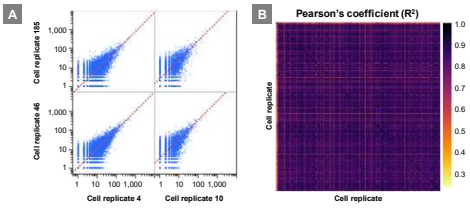

Figure 4. Gene expression data is reproducible between cells in the automated STORM-seq protocol. Panel A. Correlation plots from four randomly selected K562 cell replicates demonstrate >0.9 Pearson correlation between single-cell samples. **Panel B.** Heatmap showing Pearson's Correlation of gene expression derived from 381 cells.

5 Detection of multiple RNA biotypes

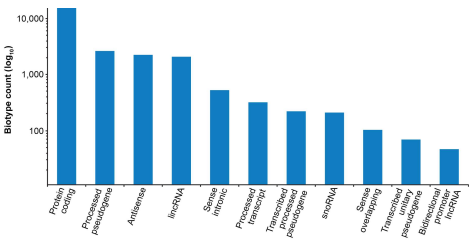

Figure 5. STORM-seq identifies multiple RNA biotypes. Counts of selected biotypes for the automated STORM-seq protocol identified from 381 K562 cells.

6 Detection of gene fusions in K562 cells

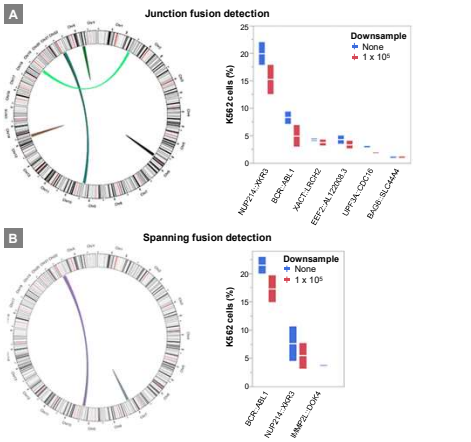

Figure 6. STORM-seq identifies gene fusion events in K562 cells. Panel A. Circos plot visualization and quantification of junction fusion events. **Panel B.** Circos plot visualization and quantification of spanning fusion events.

7 Detection of distinct single-cell populations

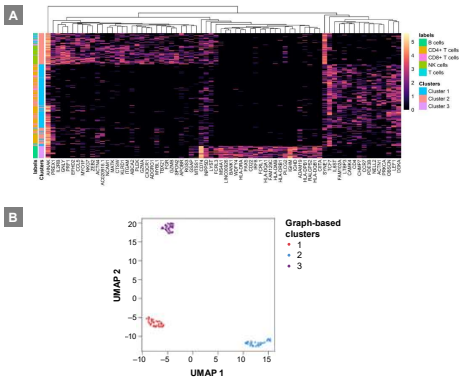

Figure 7. STORM-seq identifies distinct cell populations. Panel A. Heatmap with dendrogram illustrating differentially expressed genes in PBMCs. **Panel B.** uMAP from RCCs, K562 cells, and human brain control RNA. **Panel C.** Heatmap with dendrogram illustrating differentially expressed genes in RCCs, K562 cells, and human brain control RNA.

Conclusions

- STORM-seq is a flexible (automated or manual) and cost-efficient method for transcriptome-wide gene expression analysis of single cells with relatively even gene body coverage.
- Scalable from several to thousands of cells, STORM-seq has an adaptable experimental design to accommodate various inputs, such as cryopreserved or FACS-isolated cells from cell lines, primary cells, and tumor cells.
- Primer design and uniform gene body coverage in STORM-seq enable the detection of coding and noncoding RNA biotypes as well as different transcript isoforms and gene fusions.

References

Johnson, B. and Rhodes, M. et. al. Single-cell Total RNA Miniaturized sequencing (STORM-seq) reveals differentiation trajectories of primary human fallopian tube epithelium. *bioRxiv* (2022).

Figure 1 was created with BioRender.com

For further inquiries, please email
Saloni Pasta at Saloni_Pasta@takarabio.com

800.662.2566
takarabio.com