Multiple Platforms for Single Cell Genomics to Enable Biomarker Discovery in Immunotherapy

Ankita Das[1], Jing Wang[1], Kayla Lee[1], Derek Vargas[1], Gavin Washburn[1], Niyati Thosani[1], Vasumathi Kode[1], Nitin Mandloi[2], Angelica Lavallee[1], Amit Chaudhuri[1,2] and Papia Chakraborty[1]

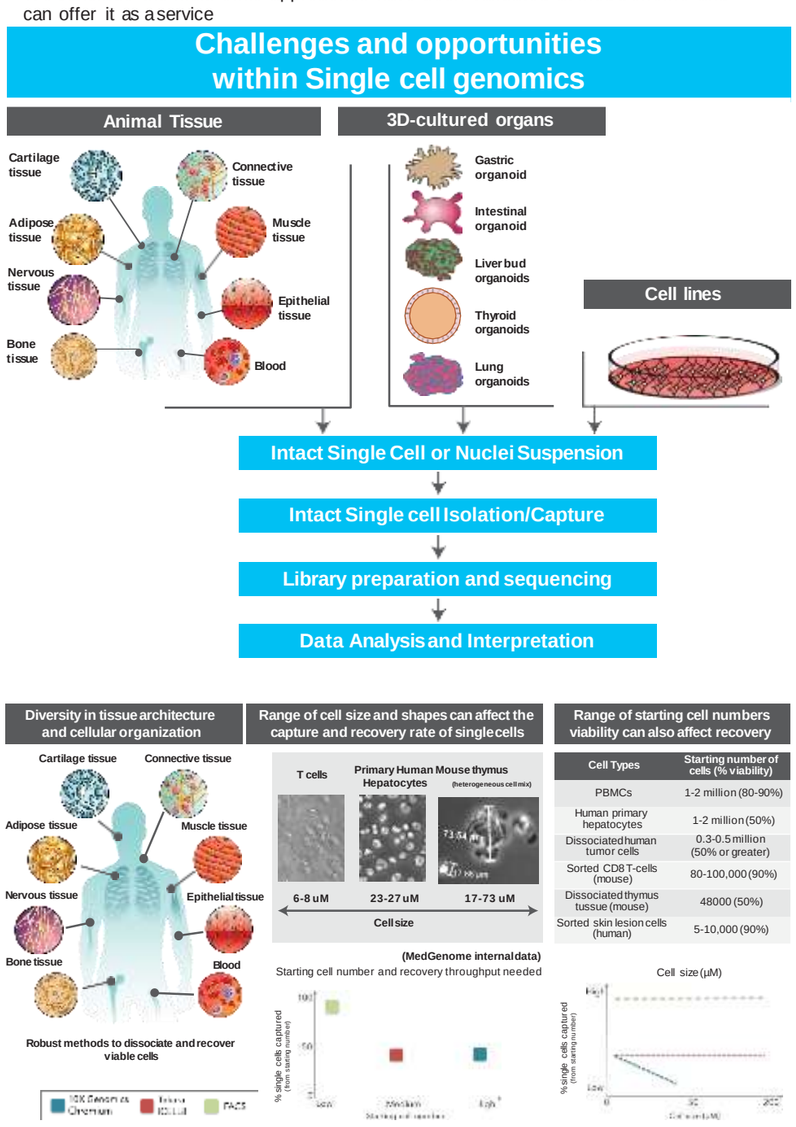
1) MedGenome Inc, Foster City, CA, 2) MedGenome Labs, Bangalore, India; corresponding author : ankita.d@medgenome.com

Abstract

Single cell genomic approaches can provide valuable insights into the complexity and heterogeneity of the cell types in the context of a tissue or tumor. However, challenges with the preparation of single cell suspensions, good cell viability and efficiently capturing diverse cell types in a mix via appropriate cell capture methods can override the utility of the approaches. We provide validation and application data generated from a range of single cell input types (number of starting cells, viability, and research question) and diverse commercially available platforms (the 10X Genomics Chromium and the FACS based SMART-Seq (Takara Bio). We demonstrate the utility of using the appropriate single cell genomics approach to get relevant information. Recent studies have shown that capturing additional information on cellular phenotypes or features can provide valuable information on cell identities otherwise missed and additional heterogeneity, which in turn facilitates discovery of meaningful biomarkers and understanding of molecular mechanisms of development and disease. To enable such discovery, we have validated the sC-ATAC seq, CITE-Seq and CUT&Tag approaches and present data on the utility of those approaches and highlight the capabilities at MedGenome.

- We present a platform agnostic and flexible approach to utilize single cell profiling of gene expression depending on the research question and the number of cells available
- We present epigenomic profiling data using the CUT&TAG workflow and show that this approach can be utilized to study epigenetic modifications and examine chromatin accessibility can identify cellular identities and understand epigenetic mechanisms for lineage commitment and disease progression
- We conclude the multi-omic approaches have been validated at MedGenome and can offer it as a service

Challenges and opportunities within Single cell genomics



Cell Types	Starting number of cells (% viability)
PBMCs	1-2 million (80-90%)
Human primary hepatocytes	1-2 million (50%)
Dissociated human tumor cells	0.3-0.5 million (50% or greater)
Sorted CD8 T-cells (mouse)	80-100,000 (90%)
Dissociated thymus tussue (mouse)	48000 (50%)
Sorted skin lesion cells (human)	5-10,000 (90%)

Figure 1: Schematic representation of the different single cell workflows at MedGenome: Starting with single cell suspensions from fresh or frozen dissociated tissues, primary cells, cell lines and FACS sorted cells, cell viability and counting assessments are performed. Depending on the downstream application, cell processing steps are performed. For single cell gene expression, depending on the input cell type or the research question, single cells can be loaded onto the 10X Genomics Chromium controller and library preparation and sequencing can be performed. Single cells can also be sorted into plates with Lysis buffer for the SMARTSeq approach. To perform single cell ATAC seq, nuclei isolation is performed from single cell suspensions and after transposition of the nuclei in bulk, the 10X Chromium controller can be used to generate single nuclei libraries and be sequenced on the Illumina platform. To process cells for characterization of cell surface markers to identify subtypes of cells with CITE-Seq, single cell suspensions can be labeled with antibodies against the cell surface markers. Downstream labeling with feature barcode libraries along with gene expression libraries or single cell TCR repertoire libraries can be generated using the 10X Immune profiling chemistries. A combination multi-omic approach with single cell genomics can enable discovery of additional cellular heterogeneity and characterize biomarkers.

Single cell genomics solutions at MedGenome

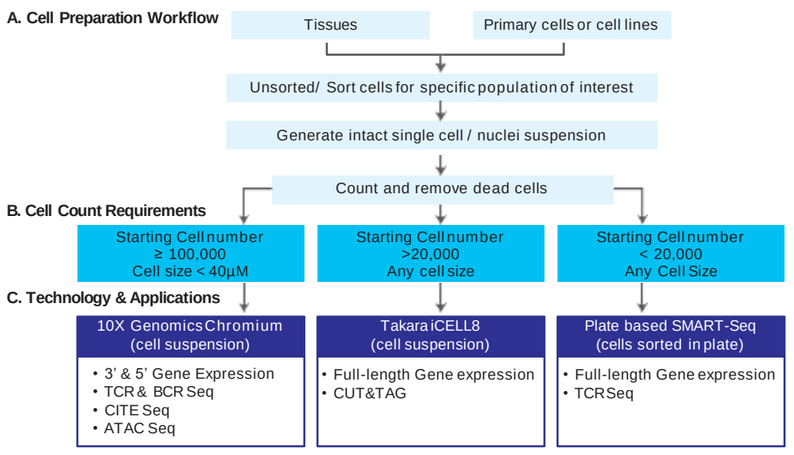
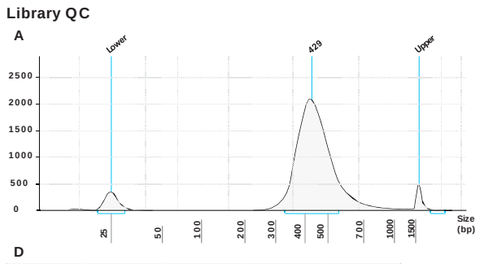
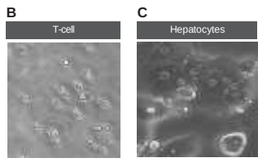


Figure 2: A) Starting with intact single cell / nuclei suspensions obtained from dissociated tissues, primary cells, cell lines or sorted cells, cell viability checks (trypan blue staining and microscopic examination) and dead cell removal steps are performed via washes. B) The rationale for selecting the platform for single cell isolation/encapsulation depends on the starting number of viable cells available, the recovery rate needed and the cell size and shape of the cell. C) At MedGenome, we have integrated and validated three of the well validated approaches. C.1) If there are greater than 100,000 cells available with good viability, and cells are less than 40µM, the 10X Genomics Chromium is recommended. If cells are larger than 40µM, the following two options are available for single cell genomics. C.2) If the application calls for sequencing over a thousand cells, and imaging the cells might be of value, the Takara ICELL8 platform is an ideal platform to select and image single cells. C.3) Lastly if there are a few hundred to thousands of cells available for sequencing, the plate based SMART-Seq technology is an ideal approach to capture the full-length transcriptome diversity.

Library QC





D

Well	Conc. [pg/µl]	Sample Description	Alert	Observations
B1	5660	1 GEX		

F

Size [bp]	Calibrated Conc. [pg/µl]	Assigned Conc. [pg/µl]	Peak Molarity [pmol/l]	% Integrated Area	Peak Comment	Peak Comment
25	409	-	25200	-		Lower Marker
429	5600	-	20300	100.00		-
1500	250	250	256	-		Upper Marker

E

Cell types	Fresh frozen/ methanol fixed	Cell viability	Number of genes detected
PBMCs-10XData	Not reported	Not reported	1,297
PBMCs-Source1	Fresh/frozen	99 %	1,539
PBMCs-Source2	Fresh	87 %	889
PBMCs-Source3	Fresh/frozen	89 %	733
Primary hepatocytes	Fresh/frozen	80 %	6,380
Mouse thymus dissociated tissue	Fresh/frozen	60 %	851
Primary coloncells	Fresh/frozen	92 %	4,127
Cancer cell lines treated with	Methanol fixed	Tumor cell: 92% T cells: 57%	6,485
T Cells-Sources1	Fresh	95 %	2,850
T Cells-Sources2	Methanol fixed	95 %	2,841

Figure 3: Validation of single cell gene expression workflow using the 10X platform: A) Example of Tape Station plot of a final gene expression library. B&C) Shows quantification of the final library profile. D) Bright-field image taken of a sample PBMC where we see an ideal sample with cell size 8-10 µM. F) Bright-field image of hepatocytes (40-50 µM) which is not ideal for processing with 10X Genomics platform and might lead to lower yield of final cells being processed into libraries. As part of our workflow, we evaluate the samples and process, and provide our recommendations on how to analyze the downstream data obtained from the sequencing. F) Example table with cell viability measurements after thawing the fresh frozen samples and post wash and clean up viability. The viability is determined using Trypan blue staining. Before proceeding with processing cells through the 10X Genomics chromium controller, we remove the dead cells using the EasySep™ (Stem Cell Technologies) dead cell removal kit.

Identification of T cell functional response and novel biomarkers of T cell subpopulations using the 10X platform

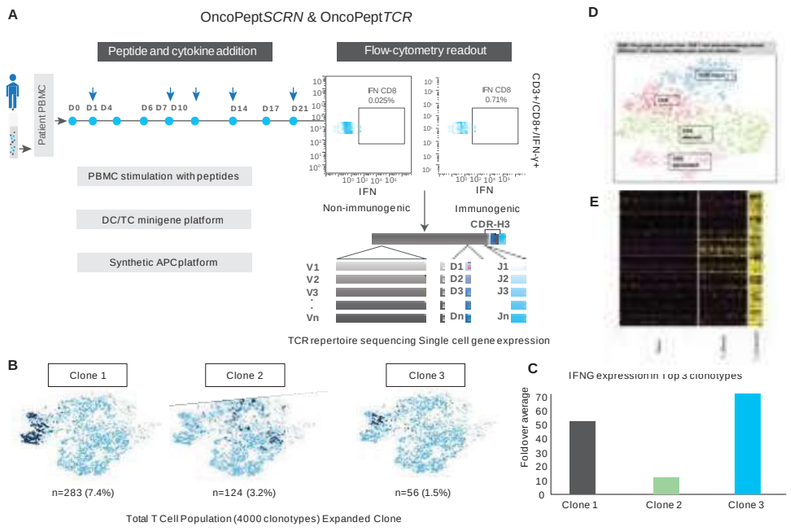


Figure 3: Application of single cell gene expression coupled to TCR repertoire : A) Graphic shows the workflow of an in-vitro neoantigen vaccine screening assay, with a read-out of CD8 T cell activation from PBMCs treated with the peptides. Functional CD8 T cell response is quantitated by IFN-γ production on a flow cytometer. TCR repertoire sequencing and single cell gene expression is then analyzed. B) Single cell TCR repertoire analysis of PBMCs (with positive responses to the peptide, assayed by IFNγ production) identified over 4000 TCR clonotypes, with t-SNE plots showing the top 3 clonotypes and the frequencies are listed below. C) Single cell gene expression of the amplified T cells reveal that there is varying levels of IFNγ in T-cells that are clonally expanded and this information can provide insights on the functional state of the expanded T-cell clones. The single cell TCR and gene expression experiment was performed using the 10X Genomics Immune profiling solution and sequencing was performed using the Illumina Novaseq 6000. 10X Cell Ranger analysis was performed downstream. D) Shows single cell plots from CD8 T cell activation assays. There are different T cell transition states post vaccine stimulation including CD8 native, CD4, CD8 effector and CD8 exhausted. E) shows clustering of the marker genes in the subpopulations identified from the gene expression data.

Full Length Gene Expression Analysis of Single Cells Using SMART-Seq

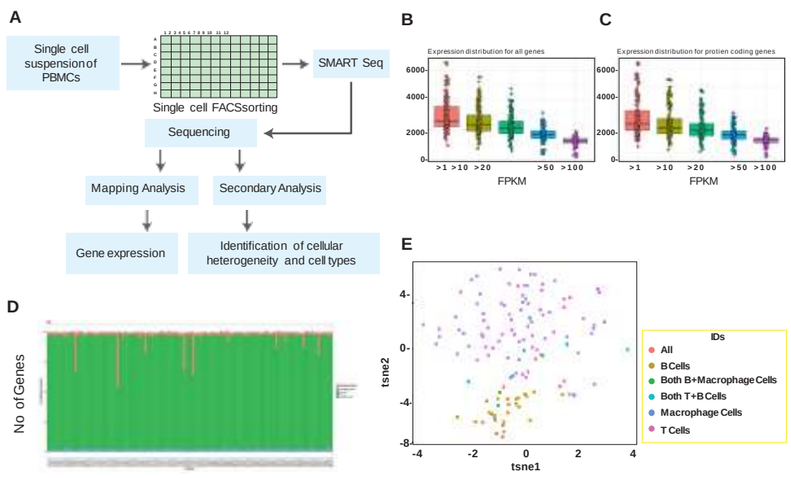
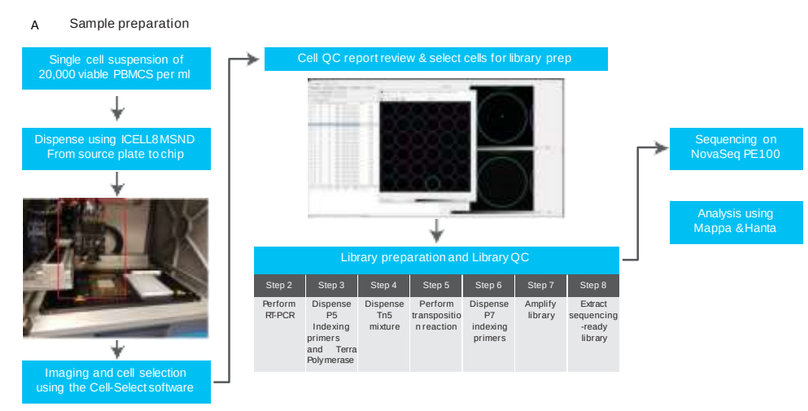


Figure 4: A) Shows workflow for full-length single cell gene expression using PLATE-SEQ. Briefly, cells are sorted into single cell plates and full-length RNA seq libraries are generated using the Takara SMART-Seq v4 Kit, and sequenced using the NovaSeq PE100 sequencing. Analysis is performed using the standard workflows, and cell types are identified by secondary analysis. B &C) Mapping statistics shows good quality mapping to exonic, and low ribosomal. In addition we show that using this approach we show that we identify over 3000 genes with FPKM >1, with several cells with more genes, making this a sensitive method for single cell gene expression analysis. E) shows results from splicing analysis where novel junctions identified using this approach.

Full Length Gene Expression Analysis of Single Cells Using ICELL8



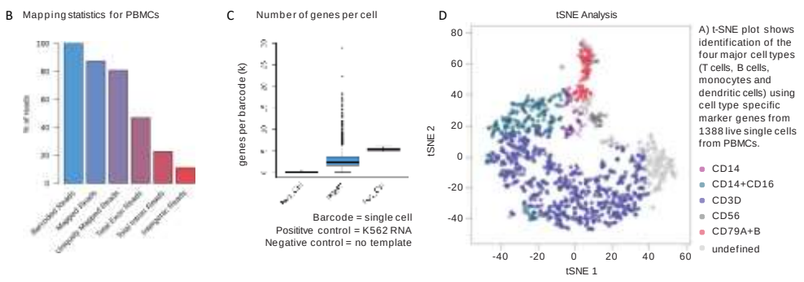
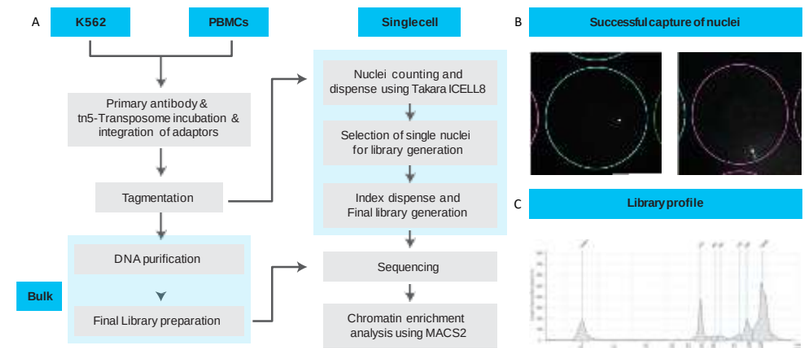


Figure 5: Validation of single cell gene expression workflow using the Takara iCell8 platform: A) Workflow for sample preparation and library preparation for full-length single cell gene expression. Starting with 20,000 PBMCs in a single cell suspension, cells are dispensed into the chip using the MSND (multi-sample nanodispenser). Assessment of live vs dead cells and single cells vs multiplets is completed by fluorescence microscopy based detection. Intact single cells are selected for downstream processing and library preparation. The reagents for library preparation are dispensed and library preparation is completed on the platform with a final step off the platform. The libraries are sequenced at a depth of 300,000 reads per cell using the Illumina NovaSeq and analysis is performed using the Takara Mappa and Hanta analysis software. B) The mapping statistics shows good quality of mapping with 80% or greater reads aligning to the genome, over 40% exonic and low percentage of intergenic reads. C) The median number of genes per cell is 3000, with some cells with over 9000 genes expressed. D) SNE plot shows identification of the four major cell types (T cells, B cells, monocytes and dendritic cells) using cell type specific marker genes from 1388 live single cells from PBMCs.

Experimental workflow for proof of concept CUT & Tag analysis



(adopted from Kaya-Okur et al, Nature communications, vol 10, 1930)

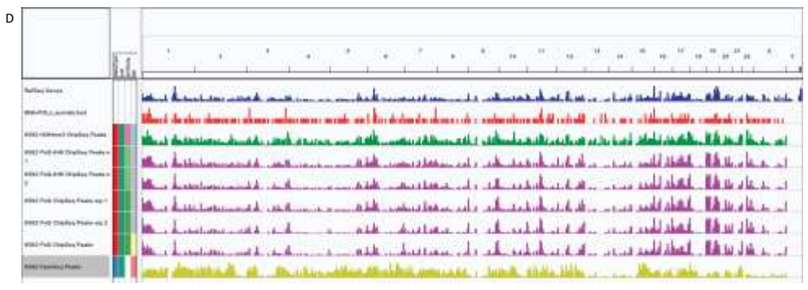


Figure 6: CUT&TAG workflow to identify chromatin accessibility and chromatin protein occupancy : A) Workflow of bulk and single cell CUT & Tag approach . B) Representative Tapestation profile of CUT&Tag libraries. C) Shows peaks called from bulk CUT&Tag libraries and compared with ChIP-Seq data obtained from public databases.

Conclusions and future directions

1) Single cell genomics assays at MedGenome can identify cell types and biomarkers from immune cells in a sensitive manner
2) Recent feature barcoding (CITE-Seq) and epigenomic single cell profiling technologies are currently being validated and will be offered as an end to end solution at MedGenome

Acknowledgements

1) Stanford FACS core
2) Takara NGS Field applications teamscientis for the single cell ICELL8 platform validation
3) Steve Henikoff lab for Cut&Tag reagents