# Processing Flow-Sorted Cryopreserved Cells for C1 mRNA Sequencing Applications

## Introduction

Since Fluidigm introduced the C1™ system, single-cell omics and multi-omic studies have been growing in adoption. Due to the power of this technique in revealing cellular variation, researchers are sequencing a broad array of tissue types. Such samples might be sourced from rare tissues or collected from a site where it is logistically challenging to conduct an experiment. Therefore, cryopreservation and flow sorting have become necessary steps upstream of single-cell processing for numerous studies, including Human Cell Atlas projects. However, cryopreserved cells after recovery and flow sorting are often less robust than fresh tissue, with a large fraction of the cryopreserved cell population becoming inviable and RNA quality being compromised[1,2]. In this study, we evaluated the effect of serum on the recovery of cryopreserved cells. We find that a significantly higher proportion of cells sorted into serum are recovered compared to cells that are not sorted into serum. Following loading and single-cell amplification on the C1 mRNA-seq HT integrated fluidic circuit (IFC), sequencing metrics of cells are improved with serum, even when limited to viable cells confirmed by imaging. Here we discuss how the addition of serum during the recovery of cryopreserved cells helps ensure greater success in C1 single-cell studies.

## Sorting into 5% Serum Improves Cell Recovery

We first explored whether serum had an effect on the recovery of cryopreserved cells following sorting. To address this question, we evaluated cryopreserved samples consisting of mixed species cell types (50:50 human:mouse cells). Samples were thawed gradually in 13 mL of 37 °C Hibernate™-A Medium (Thermo Fisher Scientific PN A1247501) media followed by centrifugation at 4 °C for 5 minutes at 400 x *g* to pellet the cells. Supernatant was removed and cells were resuspended in 4 mL of PBS with 2 mM EDTA. For each sample, a pre-sorting cell number was obtained by counting cells (regardless of species) on a hemocytometer. Following the pre-sort count, cells were labeled with SYTOX™ Blue dead cell stain (Thermo Fisher Scientific PN S34857) and processed with a SH800 Cell Sorter (Sony Biotechnology). Selected cells were viable (negative for SYTOX stain) and collected in PBS with 2 mM EDTA in the absence (n=5 experiments) or presence (n=6 experiments) of 5% fetal bovine serum. Collected cells were centrifuged at 300 x *g* for 10 minutes at 4 °C. Supernatant was removed, with 100 µL remaining in the tube. Cells were resuspended in the

remaining buffer and a post-sorting total cell number was obtained by counting the cells on a hemocytometer. The percentage of recovered cells, which was determined by dividing the post-sort cell number by the pre-sort cell number, was significantly lower in the no-serum group compared to cells recovered with serum, as shown in Figure 1 ($p=0.003$, Welch's unequal variances t test), although one sample in the no-serum group showed 61.7% recovery (dataset 1912011224).

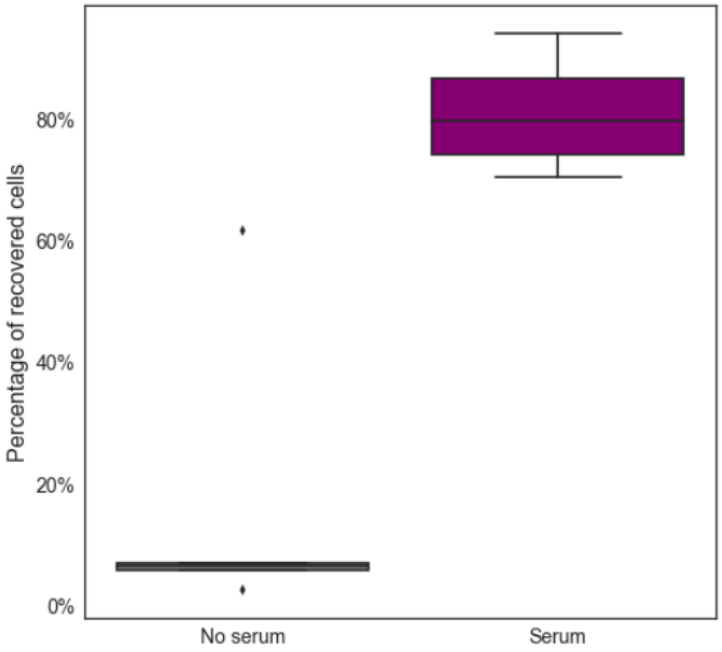


Figure 1. Adding 5% serum in collection media significantly improves the recovery of flow-sorted cells from cryopreserved samples.

## The Viability Effect of Serum Persists Through IFC Loading

While sorting without serum significantly impacted recovery, selection of viable cells was still possible regardless of whether serum was used during collection. We therefore asked whether group differences would be observed in the number of captured single viable cells following cell loading on the C1 Single-Cell mRNA Seq HT IFC, 10–17 µm (HT IFC). We evaluated four cell samples, two in the serum group and two in the no-serum group. For each sample, cells were stained with viability dye (Calcein-AM, Thermo Fisher Scientific PN C3100MP) prior to dilution to a concentration of 450 cells/µL. The stained, diluted cell preparation was then mixed with Fluidigm Cell Suspension reagent at a 7:3 ratio by volume. Following cell loading, each capture site of the HT (high-throughput) IFC was imaged and scored. A cell was called live when it stained positively for Calcein-AM. The number of viable single cells for each HT IFC is shown in Table 1. While the two groups had similar concentrations of cells loaded, fewer single, live cells were captured in the no-serum group.

This suggests that, even following selection of live cells during recovery, the viability effect of serum persists through IFC loading.

Table 1. Number of viable single cells on each HT IFC

| Dataset | Serum | Number of Viable Loaded Single Cells |
|---|---|---|
| 1912011224 | No serum | 497 |
| 1912011226 | No serum | 302 |
| 1912014277 | Serum | 654 |
| 1912014296 | Serum | 658 |

## 5% Serum Results in Significantly Higher Gene Detection Sensitivity in Single-Cell mRNA Sequencing

Following our observations in cell recovery and cell loading on the HT IFC, we asked whether the viability effect of serum was manifested in sequencing metrics such as number of genes detected, even when analysis was limited to capture sites scored by imaging as containing live cells. The four datasets described in Table 1 continued to be processed for 3′ mRNA-seq chemistry following manufacturer instructions[3]. After harvest and library preparation, samples were sequenced on an Illumina® NextSeq™ 500. Column FASTQ files were demultiplexed into individual cell FASTQ files with the C1 mRNA Seq HT Demultiplex script. Following trimming, FASTQ files were processed with a mixed species analysis pipeline. TopHat (2.0.13) was used to map reads to hg19 human reference genome and mm10 mouse reference genome. RSEM (2.1.14) was used for RefSeq transcriptome mapping of hg19 and mm10 to obtain gene expression measurements. Data analyses and visualizations were performed with Python and Python packages (Pandas for data munging, Matplotlib and Seaborn for data visualization, SciPy for statistical tests, and scikit-learn for linear regression).

As indicated above, analyses for sequencing metrics were confined to samples with the following filtering criteria: (1) Identify capture sites containing single, live cells via imaging, and then (2) confine analysis to single, live human cells by identifying a threshold for each HT dataset from the number of reads mapping to human transcriptome for each sample. After filtering, the number of single, live human cells was 224 in the no-serum group and 329 in the serum group. From this dataset, we visualized the relationship of mapped reads to gene number for each dataset in a scatter plot. To compare datasets, we determined the representative logarithmic function equation (example in Figure 2 inset). This was done by first log-transforming the number of reads mapping to RefSeq and then fitting a linear regression function to the log-transformed data with ordinary least squares (scikit-learn, linear_model.LinearRegression). Using the learned model coefficients, we predicted gene number over the range of mapped reads (x-axis on a linear scale). The resulting model curves for each dataset were then plotted together (Figure 2). The model curves for the serum datasets show a higher number of genes detected across the range of mapped reads

compared to the no-serum datasets. This suggests a viability effect of serum addition in gene detection.

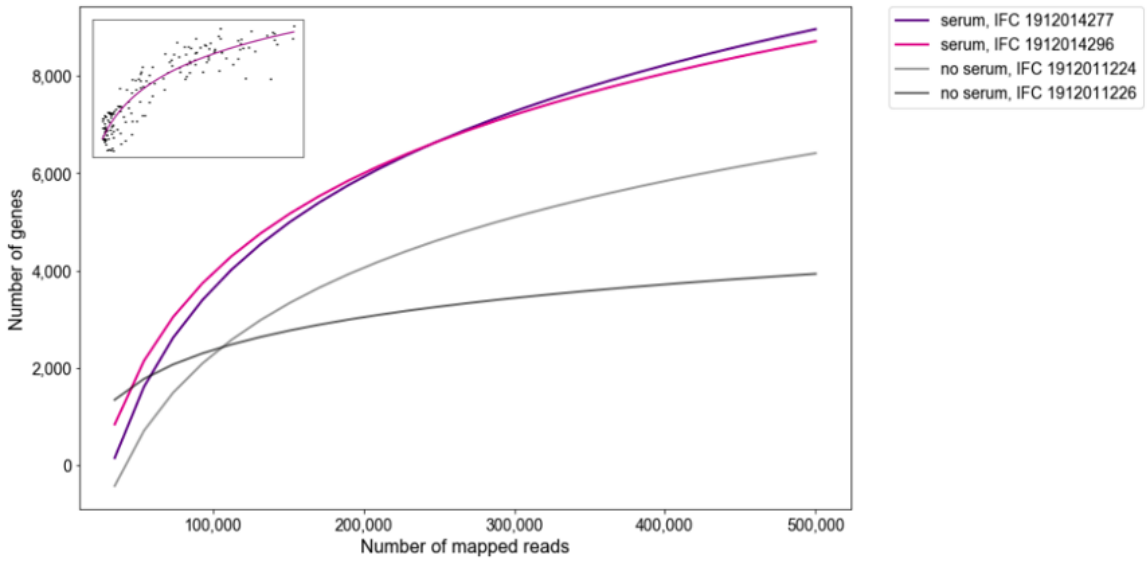


Figure 2. Cells sorted in serum show higher gene expression than cells not in serum. Inset: Depiction of learned logarithmic function model with raw cell data of serum, IFC 1912014296.

## Significantly Different Sequencing Metrics Observed for Single Viable Cells Sorted into 5% Serum

Given the observed difference in gene count, we hypothesized that other sequencing metrics would also be affected by serum state. To minimize the possibility that the metrics are different solely due to read-depth variability, we focused on metrics that can be evaluated as a percentage of total reads. We observed that the percentage of reads mapping to the transcriptome (RefSeq) was significantly higher with single-cell samples sorted in serum (Figure 3, $p < 0.05$, unpaired Student t test; scipy, stats.ttest_ind). Accordingly, the percent mapping to ribosomal RNA was higher in non-serum samples. The percentage of mitochondrial reads was not significantly higher in non-serum cells ($p=0.13$, data not shown). Collectively, this suggests that a viability effect also was reflected in transcriptome quality for cells sorted in serum, even when analysis was limited to single, live cells from IFC imaging.

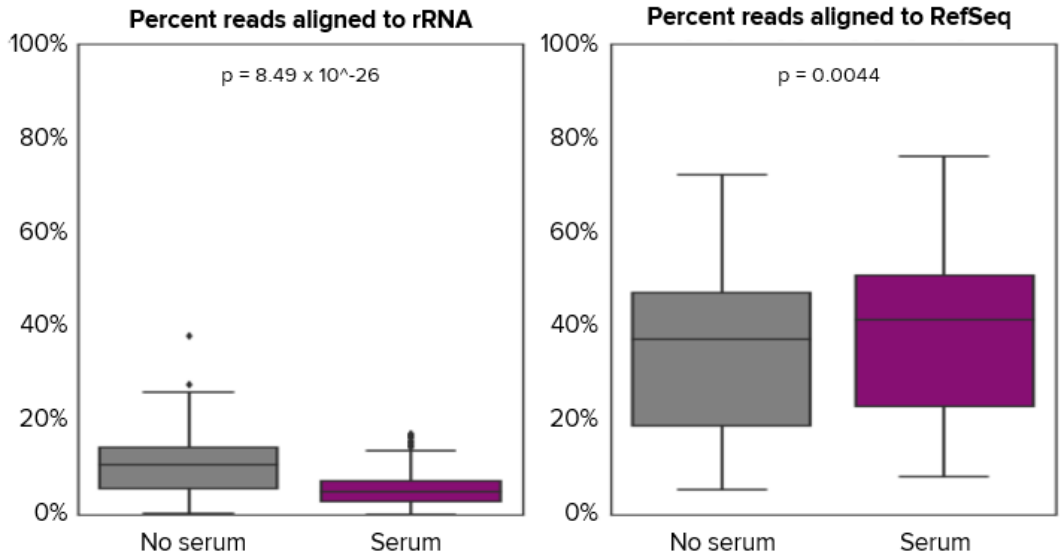


Figure 3. Significantly different alignment statistics were observed for cells collected in the absence and presence of serum.

# Conclusion

We demonstrate a workflow using serum to improve the recovery of cells processed by flow sorting from Human Cell Atlas sample sets archived by cryopreservation. The presence of serum in collection media improves the recovery of cells during cell sorting, allowing a greater number of cells to be available for processing on C1 Single-Cell mRNA Seq HT IFCs. During the on-IFC cell isolation process we observed that cryopreserved cells collected in 5% serum demonstrate a higher viability rate, enabling the capture of more high-quality cells per IFC. Likewise we experience higher reads mapping to transcriptome and lower reads mapping to ribosomal RNA, showing superior data quality when serum is added in cell collection media. Taken together, the results support use of this method as a solution for using our C1 Single-Cell mRNA Seq HT IFCs for cell characterization studies from flow-sorted cells archived by cryopreservation.

# References


1 Vieira Braga, F.A., Teichmann, S.A., and Stubbington, M.J. "Are cells from a snowman realistic? Cryopreserved tissues as a source for single-cell RNA-sequencing experiments." *Genome Biology* 18 (2017) 54.

2 Guillaumet-Adkins A., Rodríguez-Esteban G., Mereu E., Mendez-Lago M., Jaitin D.A., Villanueva A., Vidal A., Martinez-Marti A., Felip E., Vivancos, A., Keren-Shaul, H., Heath, S, Gut, M, Amit, I., Gut, I, and Heyn, H. "Single-cell transcriptome conservation in cryopreserved cells and tissues." *Genome Biology* 18 (2017) 45.

3 Generate cDNA Libraries with the C1 Single-Cell mRNA Seq HT IFC and Reagent Kit v2 Protocol (2017) (Fluidigm PN 101-4964).



**CORPORATE HEADQUARTERS**
7000 Shoreline Court, Suite 100
South San Francisco, CA 94080 USA
Toll-free: 866 359 4354 in the US and Canada
Fax: 650 871 7152
fluidigm.com

**SALES**
North America | +1 650 266 6170 | info-us@fluidigm.com
Europe/EMEA | +33 1 60 92 42 40 | info-europe@fluidigm.com
China (excluding Hong Kong) | +86 21 3255 8368 | info-china@fluidigm.com
Japan | +81 3 3662 2150 | info-japan@fluidigm.com
All other Asian countries | +1 650 266 6170 | info-asia@fluidigm.com
Latin America | +1 650 266 6170 | info-latinamerica@fluidigm.com