A Complete Workflow for Single Cell Gene Expression Profiling

Lucas Dennis*, Janelle Crane*, Xiaoyang Wang*

*NanoString Technologies, Seattle WA 98109 | †Becton Dickinson Biosciences, San Jose CA 95131

Introduction

Certain populations of cells, once thought homogeneous, have recently revealed a startling amount of gene expression variation. New understanding has been driven largely by recent advances in the sensitivity of gene expression detection technologies, which are now capable of analysis at the individual cell level. These new tools enable researchers to characterize cell heterogeneity by elucidating previously hidden relationships between individual cells within a population.

To capitalize on the wealth of expression information contained within a single cell, a robust process must be employed to identify, isolate and profile an individual cell. Such a process would ideally ensure cell viability during sorting and be fast, scalable, precise, and flexible, as well as allow for expression information to be accurately obtained from a large number of genes or gene pathways. Here, we describe a complete workflow for single cell gene expression using indexed fluorescence-activated cell sorting (FACS) on the BD FACSJazz™ cell sorter (BD Biosciences) in combination with nCounter® Single Cell Gene Expression Assays (NanoString Technologies, Inc.) that meets all of these criteria (FIGURE 1).

FACS is a powerful technology capable of delivering individual cells for downstream molecular analyses quickly and accurately. It can identify, characterize, and isolate individual cells based on cell light scattering information and specific cell surface or intracellular marker fluorescence staining intensity information. The BD FAC SJazz cell sorter is a benchtop instrument designed for individual laboratories that are new to flow cytometry. With up to six-color/eight-parameter detection, and index sorting, the BD FAC SJazz offers robust performance for genomics researchers. The BD FAC SJazz also offers the flexibility to deposit cells quickly and precisely into a variety of collection devices, including 96- and 384-well plates, tubes, or custom-made devices. By tracking and documenting sorted cell information as positions in a collection device, index sorting enables researchers to correlate cellular phenotypes with subsequent molecular analyses, such as gene expression profiles.

The nCounter Single Cell Gene Expression Assay utilizes NanoString’s optical-barcode technology to enable gene expression analysis of up to 800 genes with a simple and highly automated workflow. The ability to analyze up to 800 genes in a single tube eliminates the need for sample partitioning, minimizes the number of amplification cycles required, and allows researchers to assay entire molecular pathways or groups of pathways. The digital read-out provides highly reproducible results and allows for easy comparison with other digital technologies such as next-generation sequencing.

To demonstrate the capability and complementarity of the BD FAC SJazz and nCounter Analysis System, we performed a blind experiment by sorting single cells from a mixed cell population followed by nCounter gene expression analysis. Cellular identity was determined independently by FACS and gene expression analysis using the BD FAC SJazz and nCounter, respectively. Our results showed a 100% correlation in cell identity based on gene expression data and on flow cytometry phenotyping. These results described below validate the combined BD FAC SJazz and nCounter Analysis system workflow as a powerful method to obtain gene expression profiles from rare cells within mixed cell populations.

Complete Workflow

BD FAC SJazz™ Cell Sorter

**TaqMan® Preamp Master Mix + Pooled MTE Primers
**Reporter CodeSet, Buffer, Capture ProbeSet

FIGURE 1: Complete single cell workflow using the BD FAC SJazz and nCounter Analysis System. A population of cells is analyzed and sorted based on size and cell surface markers to obtain specific individual cells. Each selected cell is placed into a well of a 96-well plate and lysed. After cDNA conversion and target-specific enrichment, each sample is hybridized with a CodeSet containing target-specific Capture and Reporter probes. After a simple overnight hybridization, cDNA/target probe complexes are purified and counted by the nCounter Analysis System.

Molecules That Count®
Translational Research • Gene Expression • Single Cell • CNV
Cell Staining and Individual Cell Sorting on the BD FACSJazz

Before sorting, the BD FACSJazz was cleaned with ethanol and DEPC treated phosphate buffered saline. The fluidics system settings were optimized for sorting one cell into the center of each well of a 96-well plate (FIGURE 2). To verify the precise positioning in the center of well, one fluorescent bead (Rainbow Fluorescent Particles, Spherotech) was sorted into each well of a 96 well plate and imaged on an imaging system (BD Pathway™ 855) prior to sorting the cells[2].

Three cell lines, Raji, Jurkat, and HEL were separately stained with their specific cell surface markers: CD19 APC for Raji, CD3 FITC for Jurkat, and CD235a PE for HEL, respectively [all reagents were from BD Biosciences] (FIGURES 3A). After staining, cells were mixed in a ratio of 80:15:5 (Raji:Jurkat:HEL). The mixed cells were analyzed and sorted on a BD FACSJazz sorter (FIGURES 3B).

Single cells were selected (“gated”) based on parameters that describe their relative size and structure, forward scatter (FSC) and side scatter (SSC) (FIGURES 3C–E). Cells from each cell line were subsequently gated based on their unique markers (FIGURES 3F–H) so that identity could be tracked by the index sorting feature of BD FACS™ Sortware software[1].

![FIGURE 2: Fluidics system settings optimized for sorting one cell into the center of each well of a 96-well plate. Prior to sorting beads (red arrow) are sorted in order to verify positioning of subsequently sorted cells. A representative plate for well position accuracy is shown. In the experiment described here, a 96-well PCR plate was used.]

![FIGURE 3: FACS analysis of a mixed population. A) Raji, Jurkat, and HEL cells are mixed at 80:15:5 ratios. B) Gating hierarchy showing how each cell population is derived and proportion of each population after sorting. C) Live cells were gated via forward and side scatter. D–E) Individual FSC and SSC gates were used to eliminate doublets and ensure that only single cells were collected. F) Raji cells were gated based on positive CD19 staining. G) Jurkat cells were gated based on positive CD3 staining. H) HEL cells were gated based on positive CD235a staining.]

<table>
<thead>
<tr>
<th>Populations</th>
<th>Events</th>
<th>% Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSC Singlets</td>
<td>96</td>
<td>100.00%</td>
</tr>
<tr>
<td>Raji</td>
<td>78</td>
<td>81.25%</td>
</tr>
<tr>
<td>Jurkat</td>
<td>8</td>
<td>8.33%</td>
</tr>
<tr>
<td>HEL</td>
<td>5</td>
<td>5.21%</td>
</tr>
</tbody>
</table>
Prior to sorting, each well of a 96-well polypropylene PCR plate (GeneMate, BioExpress) was pre-loaded with 3 μL of a chemical lysis solution (Cells-to-Ct™ lysis buffer, Life Technologies). A BD FACSJazz sorter was used to deposit single cells from the mixed population (1 cell per well from the SSC Singlets gate, FIGURE 3E) directly into the 96-well plate. Cells were transported at low pressure at the speed of 500 events per second through a large nozzle (100 micron diameter), minimizing damage to live cells and helping to ensure cell viability. Dispensing 96 single cells took less than two minutes. FIGURE 4 shows each sorted cell within the population of analyzed cells and the location of these cells on the 96-well plate. The 96-well plates containing the single cells were immediately frozen at -80°C, and then shipped to NanoString on dry ice.

**Reverse Transcription and Multiplexed Target Enrichment**

The sample identities were blinded and cells were prepared for nCounter gene expression analysis via Multiplexed Target Enrichment (MTE) following steps outlined in the nCounter Single Cell Expression Assay Protocol. MTE is a two-step process in which input RNA is converted to cDNA and then amplified with target-specific primers. MTE primers were pooled at a final concentration of 500 nM per oligo in Tris-EDTA (pH 7.5) to ensure a working concentration of 50 nM per oligo in the MTE. The entire cell lysate (3 μL) was used as input for reverse transcription with the SuperScript® VILO™ Kit. After cDNA conversion, a highly multiplexed enrichment of target sequences was performed using TaqMan® PreAmp Master Mix and the MTE primer pool using 18 cycles of amplification, resulting in 11 μL of amplified cDNA per cell.

**nCounter CodeSet Hybridization and Analysis**

Amplified cDNA was hybridized with an nCounter CodeSet containing a panel of probes for 526 genes. After an overnight incubation at 65°C, cDNA/probe complexes were purified and deposited on an imaging surface by the automated nCounter Prep Station. Individual cDNA/probe complexes were counted and tabulated by the nCounter Digital Analyzer, resulting in a digital read-out of expression quantity for each target gene. Raw data was imported into nSolver™ analysis software, normalized for hybridization efficiency using internal positive controls, and exported for downstream analysis. Unsupervised clustering of gene expression counts was used to identify subsets of cells based on gene expression patterns. To achieve this, normalized data was filtered by excluding genes and samples with fewer than 50 counts for greater than 95% of probes; resulting in the exclusion of genes that were unexpressed and samples with insufficient amplification. This resulted in a final data set of 272 probes across 92 queried samples. Count data was then log converted and clustered, identifying three distinct cell populations (groups A, B, and C) with 76 cells comprising group A, 9 cells comprising group B, and 7 cells comprising group C (FIGURE 5). Interestingly, within the largest group of cells, A, two distinct subpopulations were observed, highlighting the value of examining expression at the single cell level.

**FIGURE 4: Sorting of individual cells into a 96-well plate on the BD FACSJazz.** Indexed sorting enables tracking of cellular identity and position on a collection device after a sort. A) Sorted Raji cells [orange]. B) Sorted HEL cells [green]. C) Sorted Jurkat cells [blue]. D) Location of index-sorted single cells on a 96-well plate.

**FIGURE 5: Cluster analysis of the 272 gene expression profiles (y-axis) for each cell (x-axis) identified three distinct clusters.** Expression is shown on a blue-yellow axis with blue representing low expression (log transformed counts) and yellow high expression (log transformed counts).
FACS and Gene Expression Data Comparison

When well positions associated with cell identity of the three index sorted cell lines were compared with well positions associated with the three cell populations defined by gene expression, there was complete agreement between the FACS analysis and the gene expression profile (see TABLE 1). The number of cells from each cell line observed after sorting was within the expected range (see TABLE 2) based on sampling statistics. In this experiment, both FACS and gene expression failed to identify a small percentage of cells. However, when information from both types of analysis was considered, 100% of cells could be identified, demonstrating the complementarity of examining expression at both the protein/cell surface marker and RNA levels using the BD FACSJazz and nCounter Analysis System.

A Complete Single Cell Workflow

This complete end-to-end workflow allows for the rapid isolation of individual cells from a heterogeneous cell population and permits quantification of expression for up to 800 transcripts from each cell, making it ideally suited to support the study of pathway-based biology at the single cell level. FACS captures cells quickly and gently, so that highly accurate gene expression profiles of healthy living cells are obtained. The speed of single cell sorting by FACS also means that a large population of cells can be processed easily and efficiently. In cases where cell surface or internal markers can be used to discriminate cells of interest, rapid segmentation of cell types using FACS permits experimental design to be more targeted, reducing the cost and complexity of experiments and helping to simplify interpretation of the data. We show that individual cells that are 5% of the overall cellular population can be readily distinguished by the methods described here, providing a valuable tool for researchers interested in characterizing rare cell populations.

The results also demonstrate the ability of gene expression data to be used to identify markers of individual cell subpopulations, which can be used to further refine sorting parameters for subsequent experiments. Finally, the data indicate the potential for identifying subpopulations with high resolution from within seemingly homogenous populations of cells using highly multiplexed gene expression analysis. With the ability to measure up to 800 genes in an individual cell, this complete workflow permits researchers to assay expression of entire molecular pathways or groups of pathways within a single cell allowing for an unprecedented view of biology.

References