BECKMAN COULTER
Life Sciences

Automation of Single-Cell RNA-Seq Library Preparation using the Biomek i7 Hybrid Automated Workstation

Michael Hayes[1], David Kind[2], Martin Giner[1], Coralie Viollet[2], Bhagya Wijayawardena[1]
[1]Beckman Coulter Life Sciences, [2]Boehringer Ingelheim Pharm GmbH & Co.

Abstract

The advent of next generation sequencing (NGS) has led to a wave of new methods in the field of genomics. One new method that has rapidly gained popularity is RNA sequencing (RNA-seq), which is a reliable way to analyze the transcriptome of biological samples. One twist on the traditional RNA-seq workflow that has gained an enormous amount of interest is single-cell RNA-seq (scRNA-seq), where individual cells are first separated into droplets. Their RNA is barcoded in a manner that each cell gives rise to a unique barcode that can be tracked during subsequent sequencing runs. This extends the ability of RNA-seq to look at changes in the transcriptomes of heterogenous biological samples, where the differences between cells can be due to multiple cell types and/or changes in key regulatory/signaling pathways. One bottleneck that may arise when performing scRNA-seq workflows at scale, is the generation of high-quality sequencing libraries. After processing of single-cell samples into cDNA, fragmentation, end repair, A-tailing, adaptor ligation and PCR amplification must be performed prior to library cleanup and sequencing. These time- and labor-intensive library preparation steps are an excellent opportunity for laboratory automation. Here the gene expression library preparation steps of the 10X Genomics Single Cell 3' kit were automated using a Biomek i7 Hybrid Automated workstation. The automated method had the key advantage of increased throughput, generating up to 48 scRNA-seq libraries per run. Comparing the automated method with samples prepared by hand showed that the Biomek libraries were comparable in terms of both quality and quantity. Sequencing of the libraries prepared by both methods also showed that the samples were nearly identical. Together the data presented here shows scRNA-seq library preparation is amenable to automation using a Biomek i7 Hybrid Automated workstation, providing a user-friendly, hands-free method.

Introduction

Next generation sequencing (NGS) has revolutionized the field of genomics. Researchers can now rapidly analyze changes in DNA or RNA samples in a cost-effective and high throughput manner. This has led to a plethora of new NGS techniques, each designed to assess different scientific inquiries. RNA sequencing using NGS (RNA-seq) is a reliable way to analyze the transcriptome of biological samples. Using this method researchers can analyze changes in mRNA, giving insights into gene expression levels and even mutational status via the observed sequences. This has led to novel findings in the genetic underpinnings of disease and how novel therapeutics might act to correct these abnormalities. In addition to analyzing diseased versus normal tissue, RNA-seq can be extended to analyzing differences between single cells, dubbed single-cell RNA-seq (scRNA-seq). The scRNA-seq method differs from bulk RNA-seq in that it allows the researcher to analyze minute differences in mixed cell populations, where perhaps only a small subset of cells is driving the observed phenotype. scRNA-seq methods also allow the clustering of single cells based on their transcriptome, so one can detect rare cell types and how the cell populations change over time or in response to a drug.

The basic workflow of scRNA-seq is to separate the cell population into single cells, isolate their mRNA, generate labeled cDNA via reverse transcription, prepare a sequencing library, and then perform standard short read sequencing. As interest in scRNA-seq has grown, several technologies have been developed to meet the needs of users. One of the most popular methods that has been developed is the 10X Genomics Chromium Single Cell 3' kit. The workflow for this kit is outlined in Figure 1. Briefly, a proprietary instrument (10X Chromium Controller) separates cells into Gel Bead-in-Emulsion (GEM) particles, where each GEM contains no more than one cell. Each GEM contains the reagents required for reverse transcription, incorporating primers that tag cDNA with a unique cell barcode, as well as Unique Molecular Identifiers (UMIs) for each mRNA molecule to determine gene expression levels. This ensures that each piece of cDNA that is generated from a single GEM has the same barcode that is unique to that cell. This is followed by disruption of the GEM, amplification of the pooled cDNA, and the generation of a sequencing library for gene expression analysis. While the entire workflow can take over 8 hours, the library preparation process alone can require 4 hours depending on sample count and is a laborious process with a number of laborious magnetic bead washing steps. Here we sought to use a Biomek i7 Hybrid Automated workstation to streamline the process of library generation following 10X Chromium generation of single-cell cDNA pools.

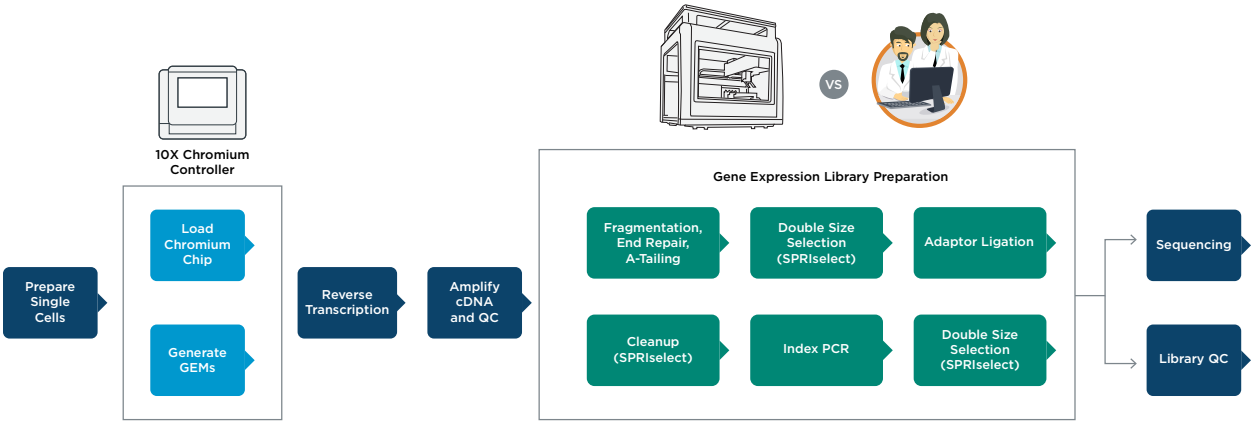

Figure 1. Workflow of the 10X Genomics Chromium Single Cell 3' Kit. Steps performed manually are indicated in dark blue, steps performed using the Chromium controller are indicated in light blue, and steps performed using the Biomek workstation via the automated method are highlighted in green.

The Biomek i7 Hybrid Automated workstation is an automated liquid handler that is capable of efficiently performing the complex liquid handling steps of scRNA-seq workflows (Figure 2). This minimizes the number of required user interactions and increases walkaway time, freeing the operator to attend to other laboratory tasks. The multichannel pod can be equipped with a 96-well head that can accurately pipette 1 to 1200 µL or a 384-well head that is accurate over the range of 0.5 to 60 µL. Additionally, the 8-channel Span-8 pod is accurate from 1 to 1000 µL. The Biomek i7 Hybrid Automated workstation supports 45 deck positions and can be directly fitted with orbital shakers, heating/cooling Peltiers, and tip-washers for plate and sample processing (Figure 2). Further, depending on user needs, the Biomek i7 Hybrid Automated workstation supports integration with other automated plate handling instruments, such as thermal cyclers, incubators, barcode readers, washers, multimode plate readers, centrifuges, and more. Here we show that automated 10X Chromium 3' gene expression library preparation using a Biomek i7 Hybrid Automated workstation provides excellent results that are equivalent to manually processed samples. The automated workflow can reduce the hands-on time and possibility of sample handling errors by the user.

Figure 2. Biomek i7 Hybrid Automated Workstation

Methods

Single-cell suspension generation

Murine bone marrow-derived cells were isolated from C57BL/6J animals, resuspended, expanded and differentiated in macrophage media with MCSF for 7 days in an incubator at 37°C with 5% CO_2. Bone marrow-derived macrophages (BMDM) cultures were then detached, and single cell suspensions were evaluated for viability (>90%), aggregation (<5%), and cell concentration using the Nucleocounter NC-3000 (ChemoMetec). Samples passing quality control steps were carried through for subsequent scRNA-seq analysis.

GEM Generation and cDNA Preparation

Single-cell libraries were generated using the Single Cell 3′ Reagent Kit v3.1 chemistry (10X Genomics). Briefly, RT (reverse transcription) master mix was prepared on ice and combined with single-cell suspension generated above. ~8,700 cells per sample were loaded onto a G Chip in order to capture a final 5,000 cells, with an expected 4.6% doublet rate. Gel bead and partitioning oil were then added to the chip, and GEM generation was performed using the 10X Controller according to the manufacturer's protocol. Following GEM generation, the resulting mixtures were transferred to PCR tubes and isothermal RT reaction was performed at 53°C for 55 min, followed by 5 min at 85°C. The cDNA output from RT was then purified using Dynabeads and amplified using kit supplied reagents on a thermal cycler with the following program: 98°C for 3 minutes (initial denature), 12 cycles of 98°C for 15 sec (denature), 63°C for 20 sec (anneal), 72°C for 60 sec (extend), and 72°C for an additional 60 sec (final extend). These PCR amplified samples were then purified using SPRIselect (Beckman Coulter Life Sciences) and assessed for quantity and quality using a Qubit 4 Fluorometer (Thermo Fisher) and Fragment Analyzer (Agilent), respectively. cDNA was then normalized and 100 ng was used as starting material for the subsequent library preparation.

Gene Expression Library Construction

Following cDNA normalization, sequencing libraries for gene expression analysis were prepared manually and using an automated method on a Biomek i7 Hybrid Automated workstation from 100 ng cDNA input. The deck layout for the automated protocol is presented in Figure 3. For fragmentation and end repair, the cDNA was combined with assembled Fragmentation Mix and Buffer EB and incubated in a traditional thermal cycler (manual) or the on-deck ATC thermal cycler (Biomek). The end-treated cDNA was subjected to SPRIselect cleanup, diluted 1:1 with preassembled Adaptor Ligation Mix, incubated at 20°C, and cDNA was again purified via SPRIselect (Beckman Coulter Life Sciences).

For sample index PCR, the cDNA with adaptor sequences was diluted 1:2 in Sample Index PCR Mix, which was followed by addition of individual i7 indices. The gene expression libraries were amplified using the following thermal cycler conditions: 98°C for 45 sec (initial denature), 12 or 13 cycles of 98°C for 20 sec (denature), 54°C for 30 sec (anneal), 72°C for 20 sec (extend), and 72°C for an additional 60 sec (final extend). For manually prepared samples, 12 cycles in an off-deck thermal cycler were used, and the automated method called for 13 cycles in the on-deck ATC. Prior to the final library elution, both methods used an additional 1:1 SPRI cleanup to remove any residual adapter and primer dimer. Libraries were eluted in either 30 µL (manual) or 23 µL (Biomek). Libraries were then quantitatively and qualitatively assessed using the Qubit 4 Fluorometer (Thermo Fisher) and the Fragment Analyzer (Agilent), as above. The scRNA-seq gene expression libraries were then normalized, pooled, spiked with PhiX Control v3 (Illumina), and sequenced on an Illumina NextSeq instrument with single index, paired-end reads (Read parameters: Rd1: 28, Rd2: 8, Rd3: 91). Data was plotted and analyzed using Prism 9 software (GraphPad). Coefficient of variance (%CV) was calculated for each condition by dividing the standard deviation by the average value and multiplying the result by a factor of 100.

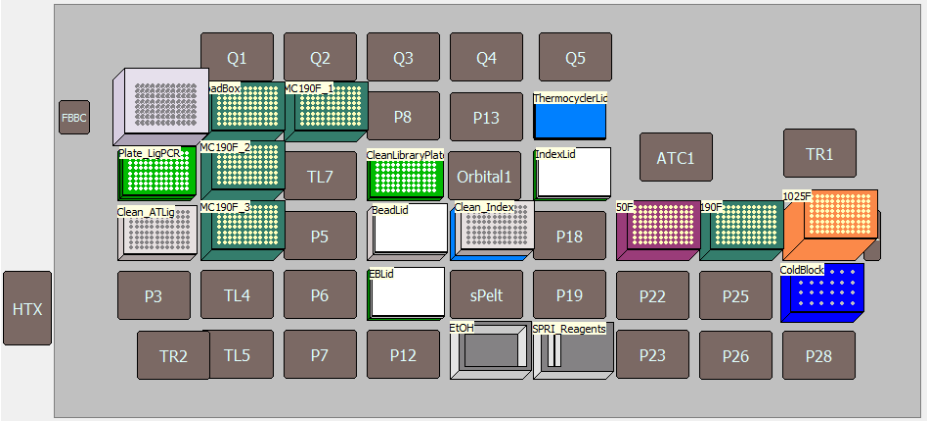

Figure 3. Deck layout for Biomek i7 Hybrid Automated Workstation protocol. The automated scRNA-seq library preparation workflow used the following components: 4 tip-loading ALPs, 13 1X1 ALPs, 1 orbital shaking ALP, 1 heating/cooling Peltier, and an on-deck thermal cycler.

Results and Discussion

The generation of gene expression libraries for scRNA-seq analysis is a time- and labor-intensive process that is an attractive target for laboratory automation. Additionally, being able to generate multiple libraries at once can have significant time and cost savings over manual preparations. Here we developed a method capable of generating up to 48 scRNA-seq libraries in parallel. In order to validate that the newly developed Biomek i7 Hybrid Automated Workstation method performed as expected, it was used directly alongside scRNA-seq libraries that were generated by hand. Prior to running the automated method, six UMI-barcoded cDNA samples from single-cell suspensions of murine macrophage cells were generated using the conventional 10X Genomics GEM workflow. Gene expression libraries were then generated for each of these six samples using both the automated and manual methods for each.

The six samples of cDNA were fragmented, end-repair treated, A-tailed, ligated with adaptors, and indexed both manually and via Biomek. One of the first steps toward validation of the automated method was comparison of the quantity and quality of the sequencing libraries following completion of the protocol (Figure 4). Using a Qubit Fluorometer (Thermo Fisher), library concentrations for each sample were measured (Figure 4A). For each of the six samples, a higher molarity was observed in the libraries prepared using the Biomek. The automated samples had an average concentration of 135 ± 17 nM, and the manual samples averaged 80 ± 5 nM (Avg ± SD). This concentration was then adjusted for the elution volume of each method to calculate the total yield for each sample (Figure 4B). The manual method used an elution volume of 30 µL, and the automated workflow eluted the libraries in 23 µL. Again, a higher average yield was observed with the automated method.

During the writing of the automated method, there was concern that the ATC located on the Biomek's deck may provide lower library yields than a traditional, off-deck thermal cycler. To assuage these concerns, 13 PCR cycles were used in the indexing PCR step of the automated method, whereas the manual method used only 12 cycles. This accounts for differences in the amount of library present in the samples presented here. As PCR amplification is 2n, where n is the number of cycles, and there is not twice as much library present in the automated sample, we can extrapolate that indeed the ATC may have been slightly less efficient than the off-deck cycler. Regardless, both methods produced well above the minimum amount of library required for sequencing, so the use of 12 versus 13 cycles in the indexing PCR step is not critical to obtaining quality results.

Prior to sequencing, the quality of the sequencing libraries made by hand or automation were compared using a Fragment analyzer (Figure 4C). The libraries that were generated were nearly indistinguishable. The manual workflow generated sequencing libraries with an average size of 457 ± 7 bp, whereas the automated method libraries averaged 466 ± 7 bp. Moreover, no short fragments in the 50 to 75 bp range due to adaptor or primer dimer were observed for any of the samples tested. Taken together this data indicated that scRNA-seq libraries of high quality and quantity could be prepared using a Biomek i7 Hybrid Automated workstation.

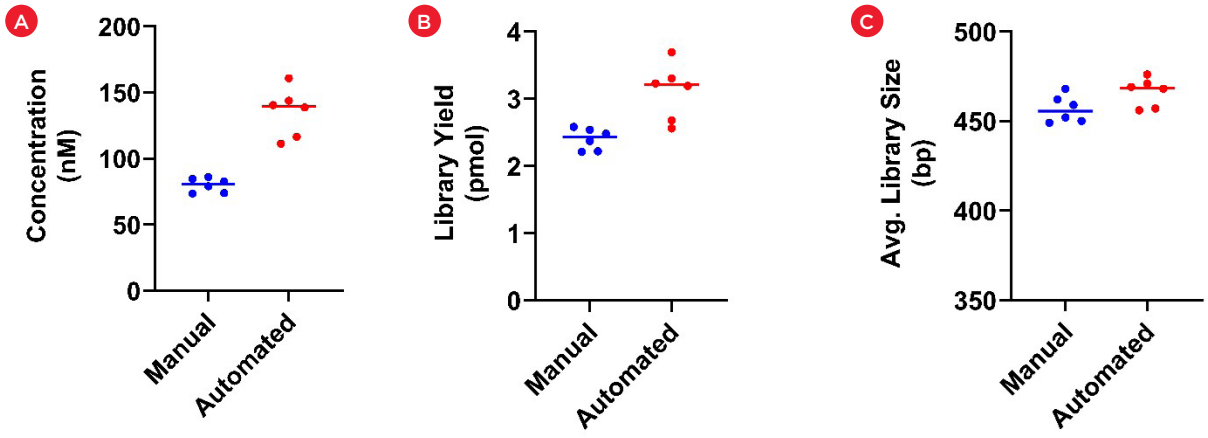

Figure 4. Library Preparation Quantity and Quality. Gene expression libraries were generated from six cDNA samples using the Biomek i7 Hybrid and manually. A) DNA concentration was quantified using a Qubit Fluorometer (Thermo Fisher) and the resulting yield (B) was calculated using the elution volume for each method. C) The average library size was measured using a Fragment Analyzer (Agilent).

After determining that the scRNA-seq libraries were of adequate quantity and quality, the gene expression profile of each sample was analyzed using NGS. After the sequencing runs were performed, analysis was performed to evaluate single-cell QC parameters, including cell count, the number of UMI, the number of genes, and the sequencing saturation. For each of the six cDNA samples tested, libraries were prepared twice, once by hand and once by the Biomek. The manual libraries were then sequenced twice to get an estimation of the sequencing variance/noise in the system, while the Biomek libraries were sequenced once (Figure 5).

The first parameter analyzed was the cell count, which should have been approximately 5,000, as that many cells were introduced in the GEM generation step for each of the six samples. While cell count did vary somewhat between samples (Sample 1 vs Sample 4), the differences observed between the manual and Biomek libraries were relatively minor, with an average variance (%CV) of approximately 10% (Figure 5A). The next parameter evaluated was the median number of UMI counts per cell (Figure 5B). In this parameter, there was, again, some variability between samples, with values ranging from approximately 2700 (Sample 5, manual) to 4500 (Sample 1, automated), but only minor differences are observed between the manual and automated results, with an average variance of 11%. A third parameter that was evaluated was the median number of genes detected per cell (Figure 5C).

As the human genome contains ~30,000 genes, approximately 5% of the human genome was detected in each sample, and the differences between sample-matched results was negligible, as only 7% average variance was observed. The final sequencing metric measured was sequencing saturation, which is dependent on the complexity of the library that was used as input and how deeply it was sequenced. For each sample tested here, the Biomek produced a slight increase in the sequencing saturation observed, though the relevance of this increase remains to be determined and could be reflective of the library pool balance. Together this data further illustrated how the Biomek i7 Hybrid Automated workstation provided increased sample throughput while maintaining sequencing results nearly indistinguishable from libraries made by hand.

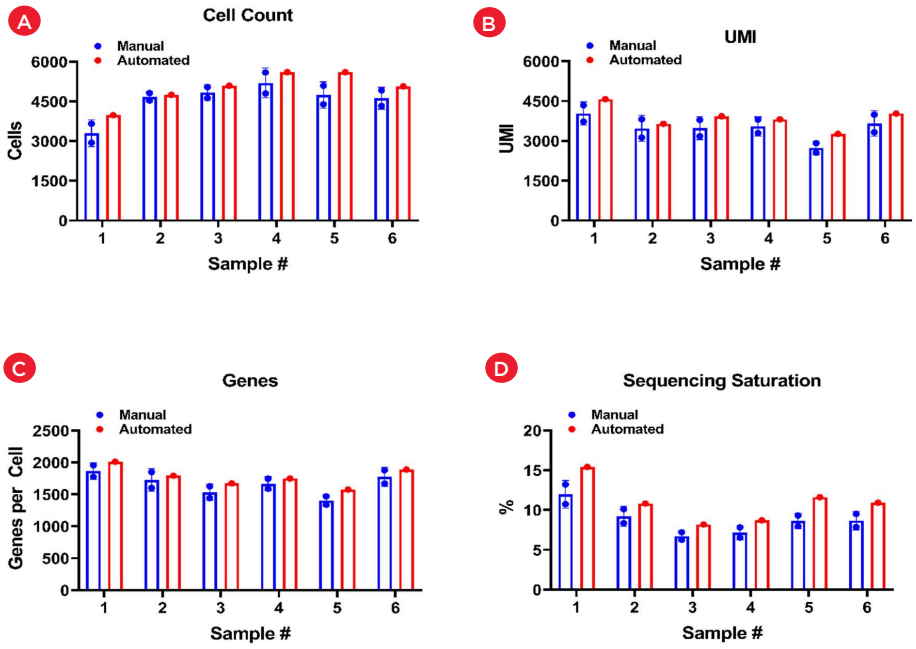

Figure 5. Gene Expression Library Sequencing Metrics. Sequencing was performed in parallel on scRNA-seq libraries prepared manually and via Biomek. The following metrics were compared for both manually and Biomek prepared samples: cell count (A), number of unique molecular identifiers (UMI) (B), number of genes observed in each sample (C), and the percent sequencing saturation (D).

Summary

Next generation sequencing (NGS) has revolutionized the field of genomics, leading to an influx of new methods and variations on existing ones. One new workflow that is becoming extremely relevant within the scientific community is scRNA-seq. This allows researchers to analyze the transcriptional changes occurring in heterogenous biological samples, which has impacts ranging from basic science to drug discovery. The preparation of NGS libraries is a critical step in scRNA-seq workflows. This involves fragmentation/end-repair/A-tailing preparation steps, adaptor ligation reactions, index PCR, and several bead purification steps. Altogether this can be both time- and labor-intensive. Here the library preparation steps of the 10X Genomics Single Cell 3' kit were automated using a Biomek i7 Hybrid Automated workstation. This automated method had the key advantage of reduced hands-on time and increased throughput, generating up to 48 scRNA-seq libraries per run, while retaining high quality and quantity. Subsequent sequencing runs showed that key parameters, such as the number of cells and genes/UMI per cell were consistent between automated and manual methods. Together the data presented here shows scRNA-seq library preparation is amenable to automation using a Biomek i7 Hybrid Automated workstation, providing a user-friendly, hands-free method with increased throughput.

Materials

Equipment	Manufacturer
Biomek i7 Hybrid Automated Workstation	Beckman Coulter Life Sciences
Static Heating/Cooling Peltier ALP	
Fragment Analyzer	Agilent
NucleoCounter NC-3000	ChemoMetec
Qubit Fluorometer	Thermo Fisher
Automated Thermal Cycler (ATC)	Applied Biosystems
NextSeq 500	Illumina
Chromium Controller	10X Genomics

Table 1. Instruments used

Reagents	Manufacturer	Part #
Chromium Next 3' GEM Chip G Single Cell Kit	10X Genomics	1000120
Qubit dsDNA HS Assay Kit	Thermo Fisher	Q32851
HS NGS Fragment Kit	Agilent	DNF-474-0500
NextSeq 500/550 High Output Kit v2.5 (150 Cycles)	Illumina	20024907
PhiX Sequencing Control v3	Illumina	FC-110-3001

Table 2. Reagents used

Consumables	#	Manufacturer	Part #
Biomek i-Series, 50 µL pipette tip, sterile, filter	1	Beckman Coulter Life Sciences	B85888
Biomek i-Series, 190 µL pipette tip, sterile, filter	4		B85911
Biomek i-Series, 1025 µL pipette tip, sterile, filter	1		B85955
Cold block	1		A83054
96-Well Skirted PCR Plate	1	BioRad	HSP9641
twin.tec PCR Plates 96, skirted	3	Eppendorf	EP951020401

Table 4. Consumables used per run

For Beckman Coulter's worldwide office locations and phone numbers, please visit Contact Us at beckman.com

21.06.1892.AUTO