TECHNICAL NOTE

Sequencing Metrics & Base Composition of Single Cell 3' v3.1 Dual Index Libraries

Introduction

The Chromium Next GEM Single Cell 3' v3.1 (dual index) workflow produces sequencing-ready Gene Expression libraries alone or in combination with either Cell Surface Protein or CRISPR Screening libraries from the same single cells. This enables simultaneous profiling of cellular features in combination with gene expression profiling. This Technical Note presents a comparison of sequencing metrics for various Single Cell 3' Dual Index library types across Illumina platforms. Individual results may vary depending on the specific sequencing instrument and/or particular sample and loading characteristics.

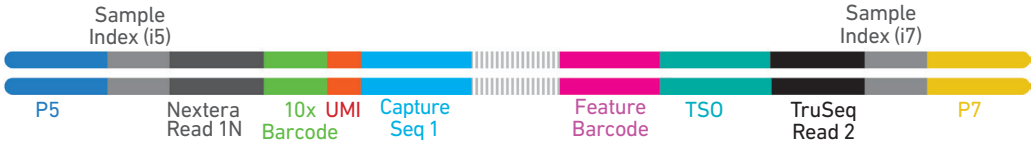
Single Cell 3' Dual Index Libraries

The three dual index library types that can be generated using the Chromium Next GEM Single Cell 3' v3.1 (dual index) reagents and protocols are shown in the schematics below.

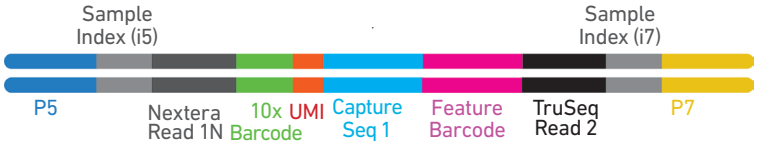
The libraries include cDNA insert or Feature Barcode constructs which begin with P5 and end with P7, sequences necessary for binding to the Illumina flow cell. Read 1 is used to sequence 16 bp 10x Barcodes and 12 bp UMI and Read 2 is used for priming and sequencing the cDNA insert or the Feature Barcode as illustrated above. The two 10 bp sample indexes are sequenced in the i5 and i7 reads.

Chromium Single Cell 3' Gene Expression Dual Index Library

Chromium Single Cell 3' CRISPR Screening Dual Index Library



Chromium Single Cell 3' Cell Surface Protein Dual Index Library



Methods Overview

Single Cell 3' Gene Expression libraries alone or in combination with Cell Surface Protein or CRISPR Screening libraries were generated (from PBMCs or A375 CRISPR Screening cells repectively) as described in the respective user guides (see References).

1,000-5,000 cells were targeted using the Chromium Next GEM Chip G to generate the libraries. The libraries were quantified and sequenced as indicated in the results (Tables 1-4).

Results Overview

Tables 1-4 show representative sequencing metrics and base composition data derived from the indicated libraries. The Q30 quality scores, representative Data by Cycle plots, and other metrics for each sequencer/workflow is shown. Individual results may vary depending on the specific sequencing instrument and/or particular sample and loading characteristics.

Conclusions

In summary, % Bases by cycle and % ≥Q30 Quality Score distribution showed highly consistent profiles for all sequencing platforms and workflows tested. The data serve as guidelines for assessing the quality of Single Cell 3' Dual Index library sequencing. Additional factors that may contribute to overall success of a sequencing run and impact downstream application performance metrics include:

- Sample preparation to obtain a high quality single cell/nuclei suspension.
- Final libraries with fragment lengths within the expected size range for each library type, for optimal cluster formation on Illumina flow cells.
- Reliable and accurate library quantification using the KAPA DNA Quantification Kit using the average insert size determined by Agilent Bioanalyzer QC.
- Sequencing platform loading concentration.

Gene Expression Dual Index Library

Three Chromium Single Cell 3' Gene Expression (dual index) libraries were pooled and sequenced on indicated Illumina sequencers. 'Data by Cycle' plots from the Illumina SAV software displaying the percentage of base calls and Q30 quality scores along with additional metrics are shown in Table 1.

Sequencing configuration & run parameters:

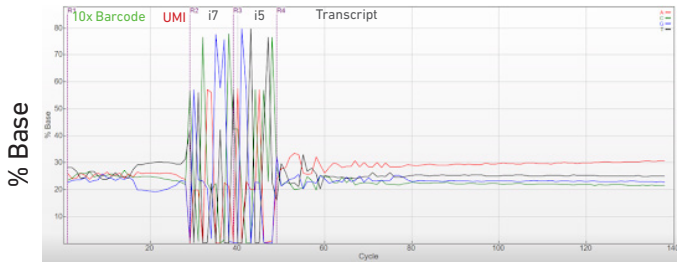
Minimum sequencing depth 20,000 read pairs per cell

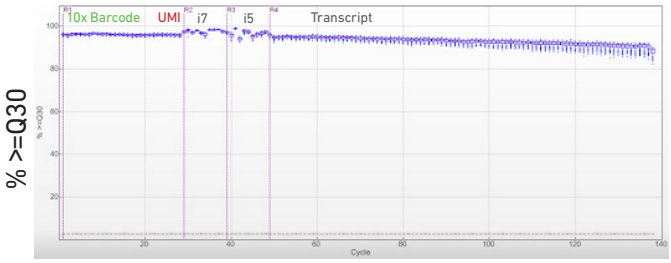
Paired-end, dual indexing

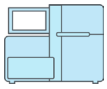
- Read 1: 28 cycles
- i7 Index: 10 cycles
- i5 Index: 10 cycles
- Read 2: 90 cycles

Table 1: Representative Plots and Sequencing Data

Plots shown are from a pool of three Gene Expression libraries sequenced on a NovaSeq SP flowcell.





		% ≥Q30				Yield per Lane (Gb)		Reads Mapped to Genome (%)
		R1	i7	i5	R2	R 1	R 2	
NextSeq 550								
	Loading Conc. (pM): 1.8 Cluster Density: 235 K/mm^2 Phix (%): 1	96.0	91.6	96.2	89.7	3.7	12.1	95.7
HiSeq 2500 RR								
	Loading Conc. (pM): 10 Cluster Density: 1047 K/mm^2 Phix (%): 1	97.6	96.4	97.7	89.0	4.9	16.2	95.4
NovaSeq (SP flow cell)								
	Loading Conc. (pM): 300 % PF*: 80.9 Phix (%) 1	96.2	96.7	97.7	93.1	14.0	46.1	96.5

*Percent Pass Filter (% PF) is reported for NovaSeq instead of cluster density due to the patterned flow cell

Cell Surface Protein Dual Index Library

Three Chromium Single Cell 3' Cell Surface Protein (dual index) libraries were pooled and sequenced on indicated Illumina sequencers. 'Data by Cycle' plots from the Illumina SAV software displaying the percentage of base calls and Q30 quality scores along with additional metrics are shown below.

Sequencing configuration & run parameters:

Minimum sequencing depth 5,000 read pairs per cell

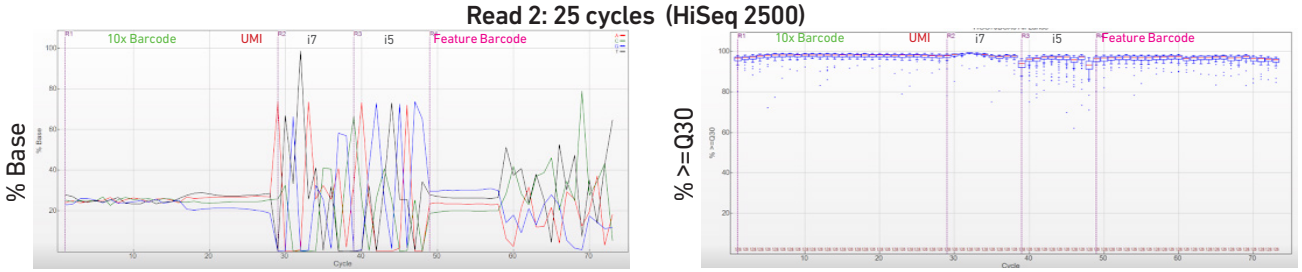
Paired-end, dual indexing

- Read 1: 28 cycles
- i7 Index: 10 cycles
- i5 Index: 10 cycles
- Read 2: 25 or 90 cycles

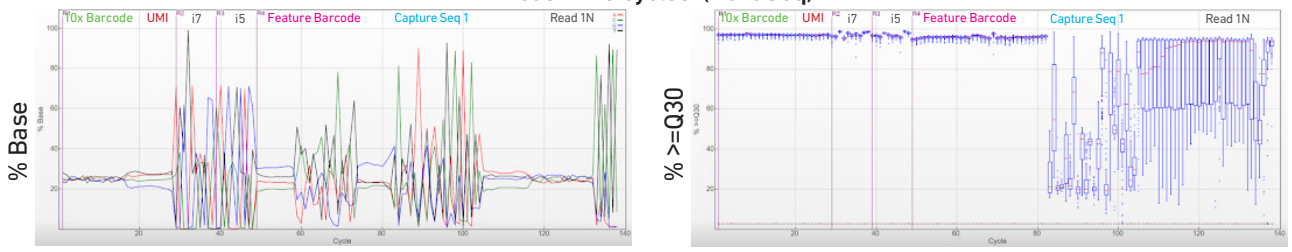
Read 2 can be sequenced at either 90 or 25 cycles.
Sequencing to 90 bp leads to a drop in quality scores after the first 25 Read 2 cycles due to sequencing through the Capture Sequence. However, this does not impact assay performance. This skewed base distribution can be normalized by pooling the Cell Surface Protein library with Gene Expression libraries (see Table 3).

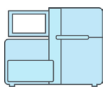
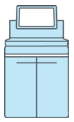
Table 2: Representative Plots and Sequencing Data

Read 2: 25 cycles (HiSeq 2500)



Read 2: 90 cycles (NovaSeq)



		% ≥Q30				Yield per Lane (Gb)		
		R1	i7	i5	R2	R1	R2	CSP Fraction Antibody Reads (%)
NextSeq 550								
	Loading Conc. (pM): 1.8 Cluster Density: 172 K/mm^2 Phix (%): 1	97.4	97.4	96.8	96.1**	2.8	2.5	93.6
HiSeq 2500 RR								
	Loading Conc. (pM): 10 Cluster Density: 1,094 K/mm^2 Phix (%): 1	97.7	95.7	98.1	96.5**	5.2	4.6	94.0
NovaSeq (SP flow cell)								
	Loading Conc. (pM): 300 % PF*: 81.3 Phix (%): 1	97.1	96.9	97.1	75.6***	14.0	46.3	93.8

*Percent Pass Filter (% PF) is reported for NovaSeq instead of cluster density due to the patterned flow cell
**Read 2: 25 cycles; ***Read2: 90 cycles

Gene Expression + Cell Surface Protein Dual Index Libraries

Three Chromium Single Cell 3' Gene Expression and three Cell Surface Protein (dual index) libraries were pooled (3:1 ratio) and sequenced on indicated Illumina sequencers. 'Data by Cycle' plot from the Illumina SAV software displaying the percentage of base calls and Q30 quality scores along with additional metrics are shown below.

Sequencing configuration & run parameters:

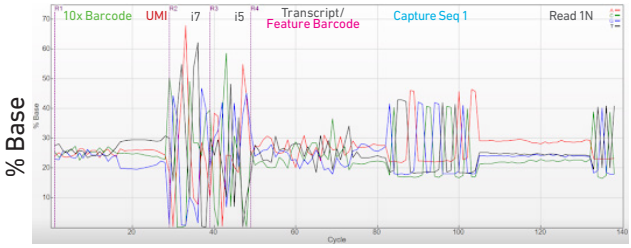
Minimum sequencing depth 20,000 read pairs per cell for Gene Expression and 5,000 read pairs per cell for Cell Surface Protein library

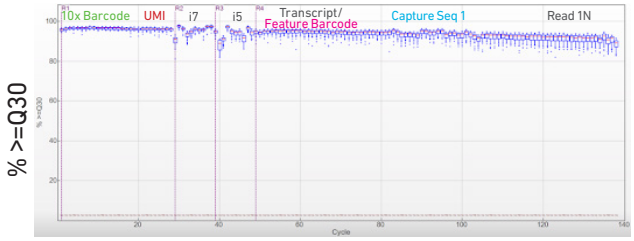
Paired-end, dual indexing

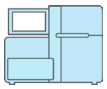
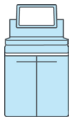
- Read 1: 28 cycles
- i7 Index: 10 cycles
- i5 Index: 10 cycles
- Read 2: 90 cycles

Table 3: Representative Plots and Sequencing Data

Plots shown are from a pool of three Gene Expression libraries and three Cell Surface Protein libraries sequenced together on a NovaSeq SP flowcell.





		% ≥Q30				Yield per Lane (Gb)		Reads Mapped to Genome (%)	CSP Fraction Antibody Reads (%)
		R1	i7	i5	R2	R1	R2		
NextSeq 550									
	Loading Conc. (pM): 1.8 Cluster Density: 222 K/mm^2 Phix (%): 1	96.3	93.6	94.4	91.1	3.5	11.4	95.6	93.4
HiSeq 2500 RR									
	Loading Conc. (pM): 10 Cluster Density: 1,030 K/mm^2 Phix (%): 1	97.9	96.2	97.2	90.1	4.9	16.0	95.2	94.1
NovaSeq (SP flow cell)									
	Loading Conc. (pM): 300 % PF*: 81.8 Phix (%): 1	96.4	93.4	95.3	93.3	14.1	46.5	96.2	93.9

*Percent Pass Filter (% PF) is reported for NovaSeq instead of cluster density due to the patterned flow cell

Gene Expression + CRISPR Screening Dual Index Libraries

Two Chromium Single Cell 3' Gene Expression and two CRISPR Screening (dual index)libraries were pooled (3:1 ratio) and sequenced on indicated Illumina sequencers. 'Data by Cycle' plots from the Illumina SAV software displaying the percentage of base calls and Q30 quality scores along with additional metrics are shown below.

Sequencing configuration & run parameters:

Minimum sequencing depth 20,000 read pairs per cell for Gene Expression and 5,000 read pairs per cell for CRISPR Screening library

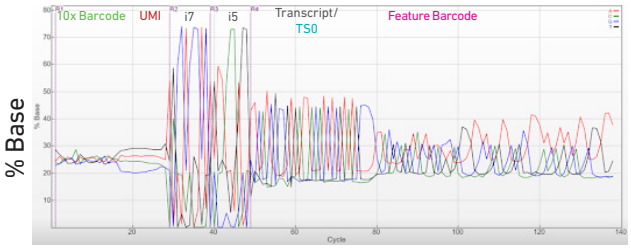
Paired-end, dual indexing

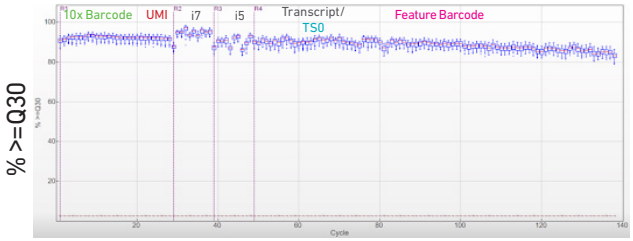
- Read 1: 28 cycles
- i7 Index: 10 cycles
- i5 Index: 10 cycles
- Read 2: 90 cycles

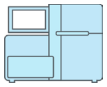
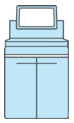
Pooling Single Cell 3' Gene Expression & CRISPR Screening dual index libraries is recommended for sequencing to maintain nucleotide diversity.

Table 4: Representative Plots and Sequencing Data

Plots shown are from a pool of two Gene Expression libraries and two CRISPR Screening libraries sequenced together on a NovaSeq SP flow cell.





		% ≥Q30				Yield per Lane (Gb)		GEX Reads Mapped to Genome (%)	CRISPR Fraction Guide Reads (%)
		R1	i7	i5	R2	R1	R2		
NextSeq 550									
	Loading Conc. (pM): 1.8 Cluster Density: 223 K/mm^2 Phix (%): 1	96.3	94.0	95.9	91.2	3.5	11.6	96.0	67.0
HiSeq 2500 RR									
	Loading Conc. (pM): 10 Cluster Density: 951 K/mm^2 Phix (%): 1	98.0	96.3	97.6	88.1	4.5	14.9	95.0	65.3
NovaSeq (SP flow cell)									
	Loading Conc. (pM): 300 % PF*: 76.9 Phix (%): 1	92.3	89.7	94.2	88.6	13.3	43.7	96.3	66.1

*Percent Pass Filter (% PF) is reported for NovaSeq instead of cluster density due to the patterned flow cell

References

- Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (Dual Index) User Guide (CG000315)
- Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (Dual Index) with Feature Barcode technology for CRISPR Screening User Guide (CG000316)
- Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (Dual Index) with Feature Barcode technology for Cell Surface Protein User Guide (CG000317)

Contact:
support@10xgenomics.com
10x Genomics
6230 Stoneridge Mall Road
Pleasanton, CA 94588 USA