TECHNICAL NOTE

Sequencing Metrics & Base Composition of Single Cell 5' v2 Dual Index Libraries

Introduction

The Chromium Next GEM Single Cell 5' v2 (dual index), standard, and High Throughput (HT) workflow produces sequencing-ready Gene Expression, V(D)J, Cell Surface Protein, and/or CRISPR 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 5' v2 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 5' v2 Dual Index Libraries

The dual index library types that can be generated using the Chromium Next GEM Single Cell 5' v2 (dual index) or the Chromium Next GEM Single Cell 5' HT v2 (dual index) reagents and protocols are shown in the schematics below.

The libraries include cDNA/V(D)J 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 10 bp UMI and Read 2 is used for priming and sequencing the cDNA insert or the Feature Barcode. The two 10 bp sample indexes are sequenced in the i5 and i7 reads.

Chromium Single Cell 5' v2
Gene Expression
Dual Index Library

Chromium Single Cell 5' v2
V(D)J Amplified
Dual Index Library

Chromium Single Cell 5' v2
Cell Surface Protein
Dual Index Library

Chromium Single Cell 5' v2
CRISPR Screening
Dual Index Library

Methods Overview

Single Cell 5' v2 dual index Gene Expression libraries alone or in combination with V(D)J, Cell Surface Protein, and CRISPR screening libraries were generated from both standard and HT workflows as described in the respective user guides (see References). The libraries were sequenced in the following combinations:

- Gene expression, V(D)J, and Cell Surface Protein libraries
- Gene expression and V(D)J libraries
- V(D)J and Cell Surface Protein libraries
- Gene expression and CRISPR Screening libraries

Libraries were generated from a target of 1,000 human Peripheral Blood Mononuclear Cells (PBMC) using the standard Single Cell 5' v2 assay. Libraries were generated from a target of 20,000 human bone marrow cells using the Single cell 5' v2 HT assay.

These combinations were selected as the most representative library pooling strategies based on the most common experimental designs. 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 for the standard assay. Figure 1 shows representative "Data by Cycle" plots from the Illumina SAV software displaying the percentage of base calls and Q30 quality scores of Single Cell 5' v2 standard and the HT derived Gene Expression, V(D)J, and Cell Surface Protein libraries run on the NovaSeq. 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. Furthermore, sequencing performance between the library types generated using the Single Cell 5' v2 standard and the HT assay are comparable. These data serve as guidelines for assessing the quality of Single Cell 5' v2 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 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 or Agilent Tapestation QC.
- Sequencing platform loading concentration.

Gene Expression,V(D)J, & Cell Surface Protein Dual Index Libraries

Four Chromium Single Cell 5' Gene Expression, eight V(D)J (four TCR amplified, four BCR amplified) and four Cell Surface Protein (dual index) libraries were pooled and sequenced on indicated Illumina sequencers. Table 1 shows 'Data by Cycle' plots from the Illumina SAV software displaying the percentage of base calls and Q30 quality scores along with additional metrics.

Sequencing configuration & run parameters

Minimum sequencing depth:
Gene Expression 20,000 read pairs/cell; V(D)J amplified 5,000 read pairs/targeted cell*; Cell Surface Protein 5,000 read pairs/cell

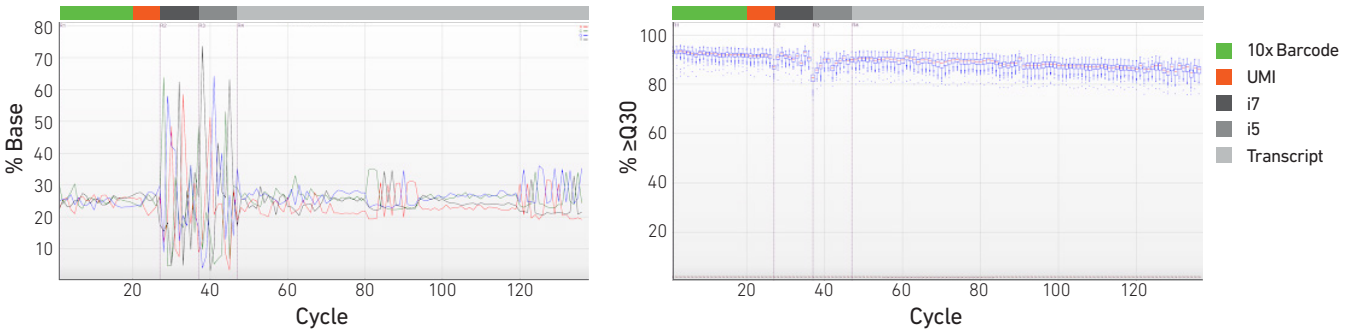
Paired-end, dual indexing

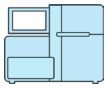
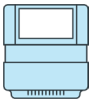
- Read 1: 26 cycles
- i7 Index: 10 cycles
- i5 Index: 10 cycles
- Read 2: 90 cycles

Table 1: Representative Plots and Sequencing Data

Plots shown are from a pool of four Gene Expression libraries, eight V(D)J libraries, and four Cell Surface protein libraries sequenced on a NovaSeq S2 flowcell.

NovaSeq S2



		% ≥Q30				Yield per Lane (Gb)		Reads Mapped to Reference (%)		
		R1	i7	i5	R2	R1	R2	Gene Expression	V(D)J	Cell Surface Protein
HiSeq 2500 RR										
	Loading Conc. (pM): 10 % PF**: 973 Phix (%): 1	97.5	96.7	95.3	95.8	4.1	14.6	98.3±0.2	78.1±11.0	96.7±0.2
NovaSeq 6000										
	Loading Conc. (pM): 300 % PF**: 79.5 Phix (%) 1	93.2	94.1	89.6	90.5	13.7	45.3	86.7±1.8	79.4±2.0	96.3±0.2
NextSeq 2000										
	Loading Conc. (pM): 650 % PF**: 84.0 Phix (%) 1	92.5	92.2	92.7	93.1	15.2	50.1	89.1±1.5	75.8±14.0	96.2±0.2

* Targeted cell refers to the number of V(D)J expressing cells in the sample. For TCR amplified libraries, targeted cell refers to the number of T cells. For BCR amplified libraries, targeted cell refers to the number of B cells

**Percent Pass Filter (% PF) is reported for NovaSeq and NextSeq 2000 instead of cluster density due to the patterned flow cell

Gene Expression & V(D)J Dual Index Libraries

Four Chromium Single Cell 5' Gene Expression and eight V(D)J (four TCR and four BCR amplified) 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 2.

Sequencing configuration & run parameters:

Minimum sequencing depth:
Gene Expression 20,000 read pairs/cell; V(D)J amplified 5,000 read pairs/targeted cell*

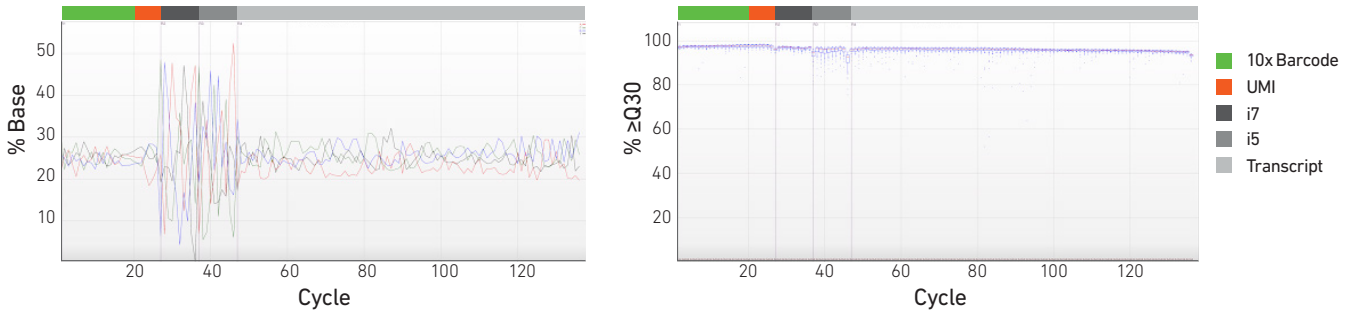
Paired-end, dual indexing

- Read 1: 26 cycles
- i7 Index: 10 cycles
- i5 Index: 10 cycles
- Read 2: 90 cycles

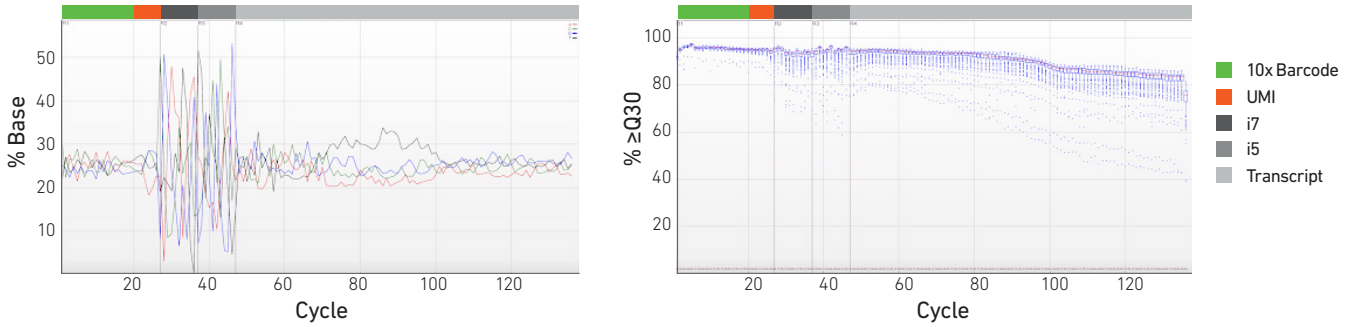
Table 2: Representative Plots and Sequencing Data

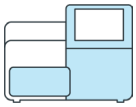
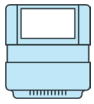
Plots shown are from a pool of four Gene Expression libraries and eight V(D)J libraries, sequenced on a MiSeq or NextSeq 500. Percentage of bases >Q30 were slightly higher in MiSeq, resulting in slightly higher usable reads for both library types.

MiSeq



NextSeq 500



		% ≥Q30				Yield per Lane (Gb)		Reads Mapped to Reference (%)	
		R1	i7	i5	R2	R1	R2	Gene Expression	V(D)J
MiSeq									
	Loading Conc. (pM): 10 % PF**: 1,040 Phix (%): 1	98.2	97.9	97.7	95.6	0.6	2.0	96.7±0.2	80.3±9.2
NextSeq 500									
	Loading Conc. (pM): 1.5 % PF**: 236 Phix (%): 1	95.3	93.5	93.8	89.1	4.4	15.8	88.0±1.6	77.0±11.6
NovaSeq 6000									
	Loading Conc. (pM): 300 % PF**: 80 Phix (%): 1	93.8	94.5	90.8	90.5	13.9	45.5	86.7±2.0	79.5±10.6
NextSeq 2000									
	Loading Conc. (pM): 650 % PF**: 83.3 Phix (%): 1	92.0	91.5	92.2	91.8	15.0	49.4	88.6±1.4	76.5±13.0

* Targeted cell refers to the number of V(D)J expressing cells in the sample. For TCR amplified libraries, targeted cell refers to the number of T cells. For BCR amplified libraries, targeted cell refers to the number of B cells

**Percent Pass Filter (% PF) is reported for NovaSeq and NextSeq 2000 instead of cluster density due to the patterned flow cell

V(D)J & Cell Surface Protein Dual Index Libraries

Eight Chromium Single Cell V(D)J (four TCR amplified and four BCR amplified), and four Cell Surface Protein 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 3.

Sequencing configuration & run parameters:

Minimum sequencing depth: V(D)J amplified 5,000 read pairs/targeted cell*; Cell Surface Protein 5,000 read pairs/cell

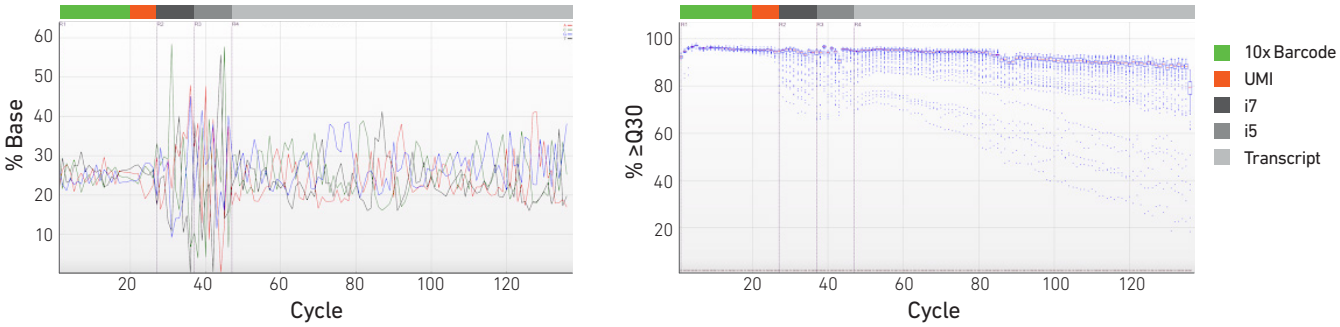
Paired-end, dual indexing

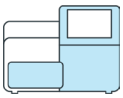
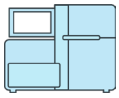
- Read 1: 26 cycles
- i7 Index: 10 cycles
- i5 Index: 10 cycles
- Read 2: 90 cycles

Table 3: Representative Plots and Sequencing Data

Plots shown are from a pool of eight V(D)J libraries and four Cell Surface Protein libraries, sequenced on a HiSeq 2500 in Rapid Run mode

HiSeq 2500



		% ≥Q30				Yield per Lane (Gb)		Reads Mapped to Reference (%)	
		R1	i7	i5	R2	R1	R2	V(D)J	Cell Surface Protein
MiSeq									
	Loading Conc. (pM): 10 % PF**: 987 K/mm^2 Phix (%): 1	98.3	97.7	97.5	95.9	0.6	2.0	80.1±9.2	96.7±0.1
NextSeq 550									
	Loading Conc. (pM): 1.5 % PF**: 211 K/mm^2 Phix (%): 1	95.9	94.2	94.0	91.4	3.9	14.0	75.8±13.1	96.2±0.1
HiSeq 2500 RR									
	Loading Conc. (pM): 10 % PF**: 969 K/mm^2 Phix (%): 1	98.0	96.7	95.7	96.0	4.3	15.2	76.6±12.6	96.7±0.2
NovaSeq 6000									
	Loading Conc. (pM): 300 % PF**: 75 Phix (%): 1	92.0	90.8	85.6	87.9	77.4	255.0	78.7±10.6	96.3±0.2
NextSeq 2000									
	Loading Conc. (pM): 650 % PF**: 83 Phix (%): 1	92.8	91.5	92.1	93.7	15.0	49.3	81.3±8.9	96.1±0.2

*Targeted cell refers to the number of V(D)J expressing cells in the sample. For TCR amplified libraries, targeted cell refers to the number of T cells. For BCR amplified libraries, targeted cell refers to the number of B cells

**Percent Pass Filter (% PF) is reported for NovaSeq and NextSeq 2000 instead of cluster density due to the patterned flow cell

Gene Expression & CRISPR Screening Dual Index Libraries

Twelve Chromium Single Cell Gene Expression, and ten CRISPR Screening 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 4.

Sequencing configuration & run parameters:

Minimum sequencing depth:
Gene Expression 20,000 read pairs/targeted cell;
CRISPR Screening 5,000 read pairs/cell

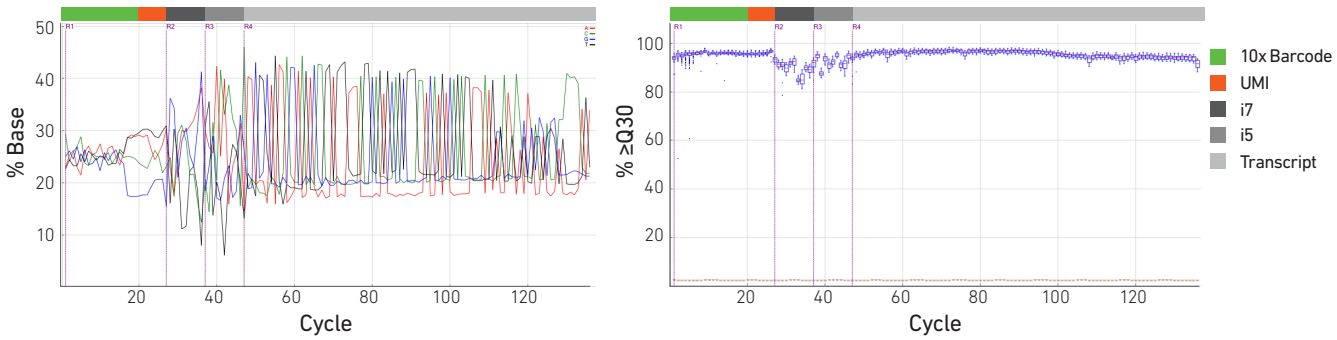
Paired-end, dual indexing

- Read 1: 26 cycles
- i7 Index: 10 cycles
- i5 Index: 10 cycles
- Read 2: 90 cycles

Table 4: Representative Plots and Sequencing Data

Plots shown are from a pool of twelve Gene Expression and ten CRISPR screening libraries, sequenced on a NextSeq 2000.

NextSeq 2000



	% ≥Q30				Yield per Lane (Gb)		Reads Mapped to Reference (%)	
	R1	i7	i5	R2	R1	R2	GEX	CRISPR Screening
MiSeq								
Loading Conc. (pM): 10 % PF**: 94 K/mm^2 Phix (%): 1	98.6	98.5	98.3	95.7	0.4	1.5	98.6±0.2	88.6±6.0
NextSeq 500								
Loading Conc. (pM): 1.2 % PF**: 95 K/mm^2 Phix (%): 1	98.1	97.0	96.7	95.1	1.7	5.9	95.1±0.4	90.3±6.0
NextSeq 2000								
Loading Conc. (pM): 650 % PF**: 72 K/mm^2 Phix (%): 1	95.8	90.0	92.1	95.7	12.0	42.7	95.0±0.5	93.8±5.8

**Percent Pass Filter (% PF) is reported for NovaSeq and NextSeq 2000 instead of cluster density due to the patterned flow cell

Gene Expression, V(D)J, and Cell Surface Protein Dual Index Libraries from Single Cell 5' v2 standard & HT assays

A pool (1:1) of two Chromium Single Cell 5' v2 Gene Expression, two BCR, and two Cell Surface Protein (standard assay) libraries were compared to a pool (1:1) of two Chromium Single Cell 5' v2 Gene Expression, two BCR, and two Cell Surface Protein (HT assay) libraries. "Data by Cycle" plot from the Illumina SAV software displaying the percentage of base calls and Q30 quality scores are shown in Figure 1.

Sequencing configuration & run parameters:

Minimum sequencing depth:
Gene Expression 20,000 read pairs/targeted cell,

V(D)J amplified 5,000 read pairs/targeted cell*, Cell Surface Protein 5,000 read pairs/cell*

Paired-end, dual indexing

- Read 1: 26 cycles
- i7 Index: 10 cycles
- i5 Index: 10 cycles
- Read 2: 90 cycles

Figure 1: Representative Plots

Plots shown are from a pool (1:1) of two Gene Expression libraries, two BCR, and two Cell Surface protein libraries, sequenced together on a NovaSeq flowcell.

NovaSeq

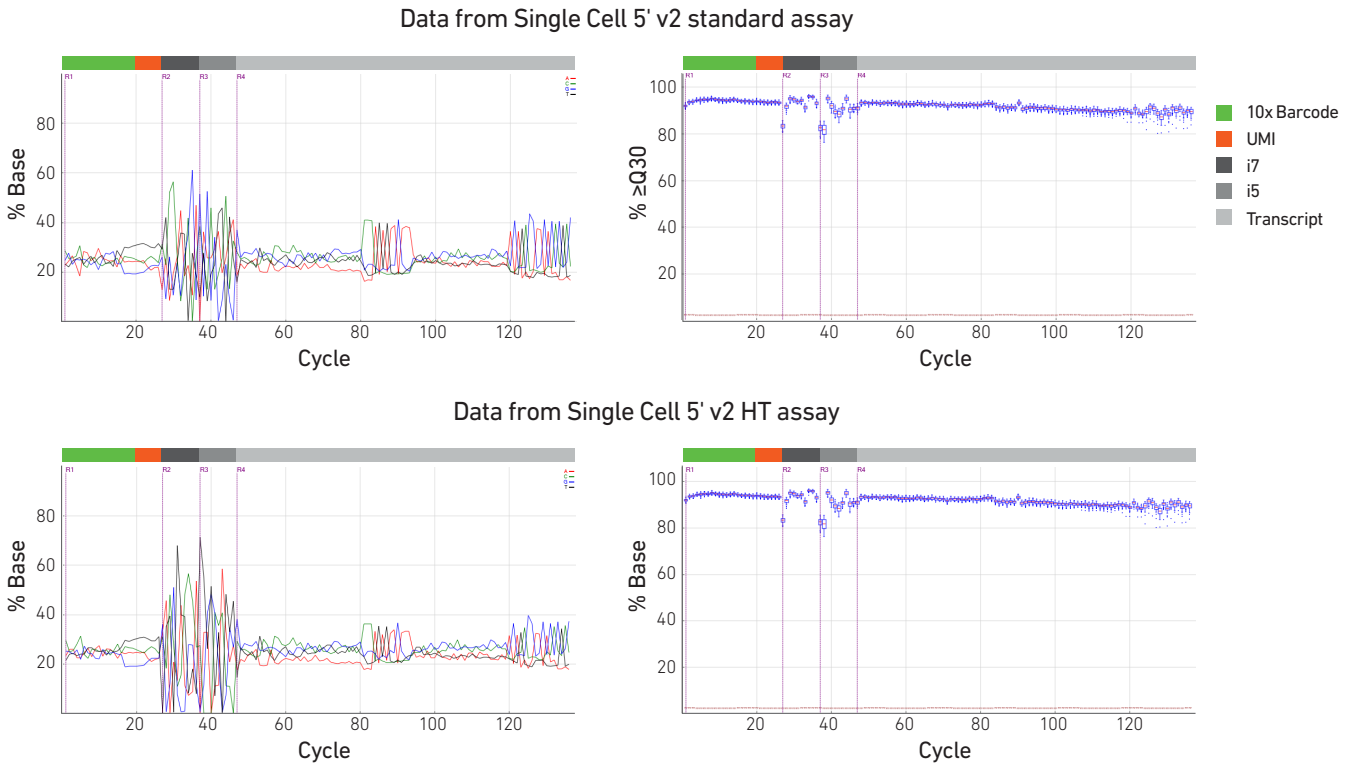


Figure 1. Representative plots derived from Single Cell 5' v2 standard assay (top panel) and HT assay (bottom panel) libraries sequenced on a Novaseq flowcell. Metrics were comparable across all other sequencers tested (data not shown).

* Targeted cell refers to the number of V(D)J expressing cells in the sample. For TCR amplified libraries, targeted cell refers to the number of T cells. For BCR amplified libraries, targeted cell refers to the number of B cells

Note

Depending on the size of the antibody panel, Cell Surface Protein libraries may vary in library complexity. Follow Illumina best practices for working with low complexity libraries, such as optimizing the percentage of Phi-X spike-in.

References

- Chromium Next GEM Single Cell 5' Reagent Kits v2 (Dual Index) User Guide (CG000331)
- Chromium Next GEM Single Cell 5' Reagent Kits v2 (Dual Index) with Feature Barcode technology for Cell Surface Protein & Immune Receptor Mapping User Guide (CG000330)
- Chromium Next GEM Single Cell 5' Reagent Kits v2 (Dual Index) with Feature Barcode technology for CRISPR Screening User Guide (CG000510)
- Chromium Next GEM Single Cell 5' Reagent Kits v2 (Dual Index) with Feature Barcode technology for CRISPR Screening and Cell Surface Protein (CG000511)
- Chromium Next GEM Single Cell 5' HT Reagent Kits v2 (Dual Index) (CG000423)
- Chromium Next GEM Single Cell 5' HT Reagent Kits v2 (Dual Index) with Feature Barcode technology for Cell Surface Protein & Immune Receptor Mapping (CG000424)
- Chromium Next GEM Single Cell 5' HT Reagent Kits v2 (Dual Index) with Feature Barcode technology for CRISPR Screening (CG000512)
- Chromium Next GEM Single Cell 5' HT Reagent Kits v2 (Dual Index) with Feature Barcode technology for CRISPR Screening and Cell Surface Protein (CG000513)

Document Revision Summary

Document Number	CG000401
Title	Sequencing Metrics & Base Composition of Single Cell 5' v2 Dual Index Libraries
Revision	Rev B to Rev C
Revision Date	June 2022
Specific Changes	1. Updated sequencing data for CRISPR Screening sequencing libraries 2. Updated sequencing data for Single Cell 5' v2 standard and HT Gene Expression, BRC and CRISPR Screening libraries.
General Changes	Updated for general minor consistency of language and terms throughout

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