10x Genomics Sequencing Handbook

Introduction

10x Genomics Chromium and Visium workflows generate sequencing-ready libraries that include assay-specific dsDNA insert and barcodes. These libraries are standard Illumina paired-end constructs. They begin with P5, end with P7, and have i5 and i7 sample indices. Additionally, they have Read 1 and Read 2 standard Illumina sequencing primer sites (TruSeq or Nextera). An example configuration of a generic dual index sequencing-ready 10x Genomics library is shown below.

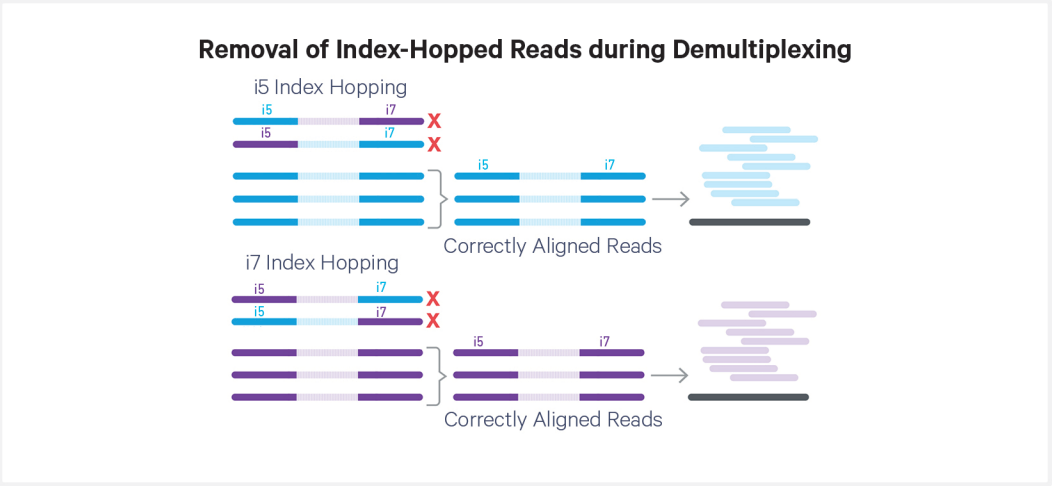
This document describes best practices and provides guidance for sequencing 10x Genomics libraries to maximize data output. General guidance for library storage, quality control (QC), pooling, and loading along with sequencing run recommendations are provided in this handbook. For additional product-specific information, consult the relevant User Guides on the 10x Genomics Support website.

Table of Contents

1. Library Storage & QC

1.0 Library Storage

Final 10x Genomics libraries can be stored at **4°C** for up to **72 h** or at **–20°C** for long-term storage. Libraries should not be pooled prior to storage as storing pooled libraries may result in elevated index hopping. Index hopping is a phenomenon that occurs during cluster generation of libraries on the sequencing flow cell and can result in incorrect assignment of insert reads to a sample. Dual indexing mitigates index hopping during demultiplexing by enabling the computational identification of reads that contain an expected pair of unique i5 and i7 index sequences.



If storing libraries for more than a month, requantifying the library before sequencing using the KAPA Library Quantification Kit is recommended.

1.1 Library Size QC

Bioanalyzer, TapeStation, LabChip, or Fragment Analyzer electrophoresis instruments are recommended to check fragment length distribution and to determine the average fragment size that will be used as the library size. Confirm that the fragment length distribution for a generated library is similar to the representative QC trace provided in the User Guide. If additional lower molecular weight or higher molecular weight fragments are observed, contact 10x Genomics Support for assistance for determining if additional cleanup steps are needed prior to sequencing. For example, lower molecular weight fragments in Gene Expression libraries may be from adaptors that were not removed via the cleanup step during library construction. If these are not

removed, adaptor dimers will generate sequencing data that will not map to the transcriptome, reducing overall usable yield.

1.2 Library Quantification

Accurate library quantification is important to ensure uniform sample pooling and optimal sequencing yield. The most common cause of under and overclustering is inaccurate library quantification. Ensure that the quantified library is diluted to the Quantitative Range (ng/µl) to ensure accurate concentration measurements.

The KAPA Library Quantification Kit for Illumina Platforms is recommended for quantifying 10x Genomics libraries. A protocol can be found in this Sequencing Handbook and the Appendix of each User Guide. The KAPA kit is recommended because it only quantifies sequenceable library molecules, such as fragments that contain both the Illumina P5 and P7 sequences. Other quantification methods (such as Bioanalyzer and Qubit) measure all nucleic acids, including incomplete library molecules. Compared to the Bioanalyzer measurements, the KAPA kit detects diverse DNA fragments with similar efficiency, making it ideal for accurate quantification for sequencing.

If using a quantification method other than KAPA, perform a comparison between the KAPA qPCR and the quantification method of choice before performing a sequencing run. When performing these comparisons, keep instruments and reagents consistent as results may vary.

Consider the following scenario:

- A library quantified by KAPA is measured to be 300 nM.
- The same library quantified by Qubit is measured at 200 nM.

Loading this library at the recommended concentration based on the Qubit quantification would result in loading 50% more library vs. using the measured concentration from KAPA. Overloaded sequencing runs may lead to overclustering on the flow cell. Overloading and overclustering can lead to poor run performance, lower Q30 scores, possible introduction of sequencing artifacts, and lower total data output (i.e., fewer usable reads) as compared to optimally loaded runs.

1.3 Post Library Construction Quantification using KAPA qPCR

a. Thaw KAPA Library Quantification Kit for Illumina Platforms.

b. Dilute **2 µl** of sample with deionized water to appropriate dilutions that fall within the linear detection range of the KAPA Library Quantification Kit for

Illumina Platforms. (For more accurate quantification, make dilutions in duplicate).

c. Make enough Quantification Master Mix for the DNA dilutions per sample and the DNA Standards (plus 10% excess) using the guidance for 1 reaction volume below.

Quantification Master Mix	1X (µl)
SYBR Fast Master Mix + Primer	12
Water	4
Total	**16**

d. Dispense **16 µl** Quantification Master Mix for sample dilutions and DNA Standards into a 96 well PCR plate.

e. Add **4 µl** sample dilutions and **4 µl** DNA Standards to appropriate wells. Centrifuge briefly.

f. Incubate in a thermal cycler with the following protocol.

Step	Temperature	Run Time
1	95°C	00:03:00
2	95°C	00:00:05
3	67°C Read Signal	00:00:30
4	Go to Step 2, 29X (Total 30 cycles)	

g. Follow the manufacturer's recommendations for qPCR-based quantification, as well as additional best practices from Illumina. For library quantification for sequencer clustering, determine the concentration based on library size derived from the Bioanalyzer, TapeStation, LabChip, or Fragment Analyzer trace.

2. Library Pooling

2.0 Library Pooling Overview

Pooling 10x Genomics libraries for sequencing will depend on factors such as library type, target cell number (Chromium libraries), tissue section size (Visium libraries), and desired read depth. Consult the appropriate 10x Genomics User Guides for recommended pooling ratios. Additional information can also be found in the “Sequencing” section of the product-specific support page on the 10x Genomics support website (e.g. Single Cell Gene Expression - Sequencing). Accurate quantification of final libraries is critical for obtaining the desired pool representation. See 1.2 Library Quantification on page 4, the KAPA Kit User Guide, and the 10x Genomics User Guides for additional information. The KAPA kit will yield a concentration in nM for each library.

2.1 Unique and Diverse Indices

To ensure that each library is correctly assigned to an input sample, select a unique index combination for each library type when pooling them. Samples using the same sample index should not be pooled together, or run on the same flow cell lane, as this would not enable correct sample demultiplexing. Consider Illumina guidelines for sample index selection. Illumina does not support pooling single and dual index libraries in the same lane. See this page on the 10x Genomics support website for sample index sequences for Universal 3' Gene Expression and this page for Universal 5' Gene Expression.

Color balance in the index read is important for optimal quality and demultiplexing results. For optimal run performance and optimal demultiplexing results, Illumina recommends ensuring there is sufficient nucleotide diversity in the index reads. For sequencing platforms with two channel chemistry, Illumina recommends selecting index sequence combinations such that signal will be present in both channels for each index cycle. When pooling samples on the NovaSeq X Series instruments, Illumina recommends avoiding index combinations which only have signals from A bases, G bases, or G+A bases in any given cycle. See Illumina’s guidance on index color balancing.

For more information, consult the following Illumina resources:

- What is nucleotide diversity and why is it important?
- Index color balancing for XLEAP SBS reagents on the NextSeq 1000/2000 and NovaSeq X/X Plus
- Index color balancing for the NovaSeq 6000 system
- Index color balancing for the NextSeq 1000/2000 system using standard SBS reagents

The tables below contain suggested index combinations when sequencing low-plex pools on 2-channel systems. 10x Genomics has not generated a list of suggested index combinations for other Illumina instruments. See Illumina resources linked above for general guidance on color balancing recommendations for different Illumina instruments.

The following criteria were used for selecting these index combinations:

- For 2-channel systems with XLEAP-SBS chemistry (e.g. NovaSeq X Series, NextSeq 1000/2000 XLEAP-SBS), each cycle of the index read should contain at least one C or T. This avoids index combinations that have signal in the blue channel only (from A or A+G) or no signal (only G)
- For 2-channel systems with non-XLEAP SBS chemistry (e.g. NovaSeq 6000, NextSeq 2000 standard SBS), each cycle of the index read should contain at least one non-G base. This provides signal in at least one channel, preferably both channels (blue channel - A or C; green channel - A or T)

These index combinations are suggested theoretical combinations to optimize color balance. Not all index combinations have been tested. 10x Genomics has not performed extensive testing of low plex pools. The tables below are not exhaustive and may not cover all possible index combinations.

The guidance below is for 10x Genomics libraries prepared using Dual Index Plates TT or TS. Consult this article for a list of library types that are prepared using Dual Index Plates TT or TS. 10x Genomics has not generated a list of suggested low plex index combinations for 10x Genomics Dual Index Plates NT, NN, and TN. These Index Plates are used for Feature Barcode libraries (e.g. Cell Surface Protein, CRISPR, etc.), which 10x Genomics does not recommend sequencing alone. Feature Barcode libraries should be pooled with their associated Gene Expression libraries for sequencing.

Illumina instruments with 2-channel sequencing require at least one base other than G for the first two cycles of each index read. To accommodate this requirement, the 10x Genomics Dual Index Plates were designed such that none of the i5 or i7 index sequences begins with “GG”.

For optimal demultiplexing results, accurate library quantification and appropriate loading concentrations are also critical, as discussed in this article.

Recommended pooling for 2-channel systems with XLEAP-SBS chemistry *(e.g. NovaSeq X Series, NextSeq 2000 XLEAP-SBS)*		
	SI-TT Plate	**SI-TS Plate**
2-plex	A5, B9	A8, B7
	B2, B10	A8, C5
	C1, F8	B2, B7
	C2, H3	D1, F7
	C10, F8	E4, F7
	D5, F11	E11, F7
	D5, H3	F7, H9
	D5, H6	G3, H2
	F6, H3	G3, H9
	SI-TT Plate	**SI-TS Plate**
3-plex	D6, E6, F6	A8, B7, B8
	C7, D7, E7	A8, B8, C8
	D7, E7, F7	C1, C2, D2
	F8, G8, H8	C1, D1, D2
	C1, C2, C3	C8, C9, D9
	C2, C3, C4	D10, E9, E10
	C9, C10, C11	F5, F6, G5
	D4, D5, D6	G1, H1, H2
	F5, F6, F7	G2, G3, H2
		G3, H2, H3
		G11, H10, H11
	SI-TT Plate	**SI-TS Plate**
4-plex	D1, E1, F1, G1	B4, C4, D4, E4
	A2, B2, C2, D2	A8, B8, C8, D8
	B2, C2, D2, E2	E8, F8, G8, H8
	A6, B6, C6, D6	D9, E9, F9, G9
	C6, D6, E6, F6	B10, C10, D10, E10
	D6, E6, F6, G6	D10, E10, F10, G10
	B7, C7, D7, E7	E11, F11, G11, H11
	C7, D7, E7, F7	
	D7, E7, F7, G7	
	E8, F8, G8, H8	
	A11, B11, C11, D11	
	B11, C11, D11, E11	

Recommended pooling for 2-channel systems with non-XLEAP SBS chemistry *(e.g. NovaSeq 6000, NextSeq 2000 standard SBS)*		
	SI-TT Plate	**SI-TS Plate**
2-plex	A1, B1	G2, H2
	A2, B2	G4, H4
	G3, H3	A5, B5
	A4, B4	C5, D5
	E4, F4	A7, B7
	E7, F7	E7, F7
	A8, B8	A8, B8
	G8, H8	G8, H8
	E9, F9	C9, D9
	G9, H9	E10, F10
	C10, D10	G10, H10
	E10, F10	A11, B11
	G11, H11	E11, F11
	A12, B12	
	G12, H12	
	SI-TT Plate	**SI-TS Plate**
3-plex	A1, B1, C1	A1, B1, C1
	F1, G1, H1	F1, G1, H1
	A2, B2, C2	F2, G2, H2
	A3, B3, C3	F3, G3, H3
	F3, G3, H3	A4, B4, C4
	A4, B4, C4	F4, G4, H4
	A6, B6, C6	A5, B5, C5
	F6, G6, H6	F5, G5, H5
	A7, B7, C7	A6, B6, C6
	A8, B8, C8	A7, B7, C7
	F8, G8, H8	F7, G7, H7
	F9, G9, H9	A8, B8, C8
	A10, B10, C10	F8, G8, H8
	A11, B11, C11	F9, G9, H9
	F11, G11, H11	F10, G10, H10
	A12, B12, C12	A11, B11, C11
	F12, G12, H12	F11, G11, H11

Recommended pooling for 2-channel systems with non-XLEAP SBS chemistry *(e.g. NovaSeq 6000, NextSeq 2000 standard SBS)*		
	SI-TT Plate	**SI-TS Plate**
4-plex	A1, B1, C1, D1	A1, B1, C1, D1
	E1, F1, G1, H1	E1, F1, G1, H1
	A2, B2, C2, D2	E2, F2, G2, H2
	E3, F3, G3, H3	E3, F3, G3, H3
	A4, B4, C4, D4	A4, B4, C4, D4
	E4, F4, G4, H4	E4, F4, G4, H4
	A5, B5, C5, D5	A5, B5, C5, D5
	E5, F5, G5, H5	E5, F5, G5, H5
	A6, B6, C6, D6	A6, B6, C6, D6
	E6, F6, G6, H6	A7, B7, C7, D7
	A7, B7, C7, D7	E7, F7, G7, H7
	E7, F7, G7, H7	A8, B8, C8, D8
	A8, B8, C8, D8	E8, F8, G8, H8
	E8, F8, G8, H8	A9, B9, C9, D9
	E9, F9, G9, H9	E9, F9, G9, H9
	A10, B10, C10, D10	E10, F10, G10, H10
	E10, F10, G10, H10	A11, B11, C11, D11
	A11, B11, C11, D11	E11, F11, G11, H11
	E11, F11, G11, H11	E12, F12, G12, H12
	A12, B12, C12, D12	
	E12, F12, G12, H12	

2.2 Library Pooling Calculations

To calculate the necessary volumes for library pooling, gather the following info:

- Average Library Size (bp): Determined by Bioanalyzer (or TapeStation / LabChip / Fragment Analyzer)
- Library Concentration (nM): Determined by qPCR
- Target Volume: can be adjusted up or down to make sure the 'final transfer volume' is not too high (e.g. use up too much of the final library); or too low (pipette volume too low / inaccurate).
- Target Pool Concentration: This is the expected concentration of the final pool. 10x Genomics highly recommends re-quantifying the pool prior to sequencing. If the expected pool concentration is significantly different from the observed concentration (by qPCR) this could indicate a problem

with pooling, individual library quantification, or final library re-quantification.

Information necessary for Chromium libraries:

- Cell Number Expected: How many cells expected for each library (e.g. targeted cell recovery)
- Reads per Cell: Outlined in each User Guide and in 3.4 Sequencing Parameters on page 19. Different sequencing depths may be necessary depending on the experimental design and cell types of interest

Information necessary for Visium libraries:

- Percentage of Capture Area covered by tissue (Visium HD) or number of tissue-covered spots (Visium v1, Visium v2)
- Read Depth Requirements: Outlined in each User Guide and in 3.4 Sequencing Parameters on page 19.

Input the information above into these example 10x Genomics worksheets for 8 reactions or 32 reactions to determine necessary input volumes.

Alternatively, each library can be diluted to the same nM (by qPCR) and pooled in equal molar ratios.

2.3 Pooling Verification

To verify correct library pooling, sequencing QC may be performed on a small output sequencer or as a spike-in. Refer to Illumina's recommendations for assessing library representation in a pool using iSeq for pooling QC and for final pool quantification. Select a sequencer with comparable clustering chemistry when possible. Sequencing QC will help verify the library quality, the sample representation in the pool, and confirm that the sample indices are represented in the expected ratios before proceeding with the final sequencing run.

2.4 Library Pooling Compatibility

10x Genomics recommends pooling Feature Barcode libraries with their respective Gene Expression libraries to increase nucleotide diversity. For example, pool Single Cell 5' v3 CRISPR Screening libraries with Single Cell 5' v3 Gene Expression libraries. Refer to the relevant User Guide for recommended pooling ratios based on the sequencing read depth requirements for each library type.

It may also be possible to pool libraries together from different 10x Genomics assays. This is low risk if the libraries share the same sequencing

configurations and have similar recommended loading concentrations. 10x Genomics has not extensively tested all possible pooling configurations across all sequencing platforms. When pooling different library types, consider that short library molecules cluster more efficiently. This different clustering efficiency is also dependent on instrument type; therefore, the optimal pooling ratio will have to be determined empirically. Optimization of loading concentrations and % PhiX input may also be required depending on the pooled library types and sequencing platform.

The following table provides guidance for pooling libraries from different 10x Genomics assays.

Product	
Universal 3' Gene Expression	
Single Cell 3' v4 Gene Expression	
Single Cell 3' v4 Cell Surface Protein	
Single Cell 3' v3.1/HT Gene Expression	
Single Cell 3' v3.1/HT Cell Surface Protein	
Single Cell 3' v3.1/HT CRISPR	
Single Cell 3' v3.1/HT CellPlex	
Universal 5' Gene Expression	
Single Cell 5' v3 Gene Expression	It is low risk to pool these library types together for sequencing as they share the same sequencing configuration (R1: 28, i7: 10, i5:10, R2:90) * 5' v2/HT libraries require 26 cycles for Read 1 if sequenced alone. If pooled with other 10x library types, Read 1 should be increased to 28 cycles. ** Flex Multiplex libraries require higher PhiX spike-in (5% or 10% PhiX depending on sequencing platform).
Single Cell 5' v3 V(D)J	
Single Cell 5' v3 Cell Surface Protein	
Single Cell 5' v3 CRISPR	
Single Cell 5' v2/HT Gene Expression*	
Single Cell 5' v2/HT V(D)J*	
Single Cell 5' v2/HT Cell Surface Protein*	
Single Cell 5' v2/HT CRISPR*	
Epi Multiome	
Single Cell Multiome - Gene Expression	
Flex	
Flex Singleplex Gene Expression	
Flex Singleplex Protein Expression	
Flex Multiplex Gene Expression**	
Flex Multiplex Protein Expression**	
Visium	
Fresh Frozen Direct Placement (v1) Gene Expression	

Product	
Visium	
CytAssist Spatial Gene Expression (v2)	CytAssist v2 libraries have different sequencing configurations than the library types above (R1: 28, i7:10, i5:10, R2: 50). It may be possible to pool these libraries with the library types above as long as (1) Read 2 is increased to 90 cycles, (2) 10% PhiX is used if pooling with Multiplex Flex libraries, and (3) the CytAssist v2 libraries do not comprise >40% of the pool. When pooling, optimization may be required depending on the sequencing platform. For example, for NovaSeq X Series instruments, the optimal loading concentrations differ between CytAssist v2 libraries and other 10x Genomics library types; thus optimization may be required to determine the optimal pooling ratios and the optimal loading concentration for the library pool.
CytAssist Spatial Gene & Protein Expression	
HD Spatial Gene Expression	Pooling Visium HD libraries with other 10x Genomics libraries is not recommended due to their unique sequencing configuration: R1: 43, i7:10, i5:10, R2: 50. Additionally, on NovaSeq X Plus, Visium libraries cluster optimally at a higher loading concentration (300-400 pM) than most Single Cell Gene Expression libraries (150-200 pM). Limited testing of Visium HD and Flex libraries pooled at a 1:1 molar ratio at range of concentrations on NovaSeq X Plus did not yield any equal sequencing representation of Visium HD and Flex libraries, likely due to the differences in optimal loading concentrations for each library type.
Epi ATAC	
Single Cell ATAC v2	ATAC v2 libraries have a different sequencing configuration (R1: 50, i7: 8, i5: 16, R2: 50). 10x Genomics has not tested and does not recommend pooling these libraries with other 10x Genomics library types (exception: see notes below for Multiome ATAC).
Epi Multiome ATAC	
Single Cell Multiome - ATAC	Multiome ATAC libraries have a different sequencing configuration (R1: 50, i7: 8, i5: 24, R2: 49). 10x Genomics has not tested and does not recommend pooling these libraries with other 10x Genomics library types, other than standalone ATAC libraries (see below). Pooling Multiome ATAC libraries with standalone ATAC libraries may be possible when using Illumina sequencers that run the forward strand workflow (eg. MiSeq), with the following adjustment: R1: 50, i7 :8 , i5: 16, R2: 49. 10x Genomics does not recommend pooling these libraries on reverse complement sequencing workflows.

3. Library Sequencing

3.0 Loading Concentration

The following table provides recommended final library loading concentrations (pM) based on internal testing. Recommended loading concentrations are starting points and may need adjustments.

Product	Kit	MiSeq (pM)	NextSeq 500/550 (pM)	NextSeq 1000/2000 (pM)	NovaSeq 6000 (pM)	NovaSeq X Series (pM)
Universal 3' Gene Expression	**Single Cell 3' v4**					
	Gene Expression	12	1.6	650	150*/150	150-200
	Cell Surface Protein	Recommended Pooling - Pool with 3' Gene Expression libraries				
	Single Cell 3' v3.1/HT					
	Gene Expression	11	1.8	650	150*/150	150-200
	CRISPR	Required Pooling - Pool with 3' Gene Expression libraries				
	CellPlex	Required Pooling - Pool with 3' Gene Expression libraries				
	Cell Surface Protein	Recommended Pooling - Pool with 3' Gene Expression libraries				
Universal 5' Gene Expression	**Single Cell 5' v3**					
	Gene Expression	10	1.6	650	150*/150	150-200
	V(D)J					
	CRISPR	Required Pooling - Pool with 5' Gene Expression libraries				
	Cell Surface Protein	Recommended Pooling - Pool with 5' Gene Expression libraries				
	Single Cell 5' v2, 5' HT					
	Gene Expression	10	1.5	650	150*/150	150-200
	V(D)J					
	CRISPR	Required Pooling - Pool with 5' Gene Expression libraries				
	Cell Surface Protein	Recommended Pooling - Pool with 5' Gene Expression libraries				
Epi ATAC	**Single Cell ATAC v2**					
	ATAC v2	11	1.7	650	150*/150	100-150

Product	Kit	MiSeq (pM)	NextSeq 500/550 (pM)	NextSeq 1000/2000 (pM)	NovaSeq 6000 (pM)	NovaSeq X Series (pM)
Epi Multiome	**Single Cell ATAC Multiome ATAC + GEX**					
	Multiome - ATAC	10	1.5	650	150*/150	100-150
	Multiome - Gene Expression	11	1.8	650	150*/150	150-200
Flex	**Flex**					
	Singleplex Gene Expression	12	2.5	650	150-200*/ 100-150	150-200
	Multiplex Gene Expression					
	Protein Expression	Recommended - Pool with Gene Expression libraries				
	Multiplex Protein Expression	Recommended - Pool with Gene Expression libraries				
Spatial	**Visium**					
	Fresh Frozen Direct Placement (v1)	11	1.8	650	150*/300	150-200
	CytAssist Spatial Gene Expression (v2)	11	1.8	650	150-200*/ 100-150	300-400
	CytAssist Spatial Gene & Protein Expression	11	1.8	650	150-200*/ 100-150	300-400
	HD Spatial Gene Expression	-	-	650	150-200*/ 100-150	300-400

*Refers to guidance for the XP (lane splitting) workflow

10x Genomics recommends a PhiX spike-in for all library types. The exact recommendation can be found in the respective User Guides. Generally, a 1% PhiX spike-in is recommended for QC purposes, unless sequencing a low diversity library or a low diversity stretch (e.g. Flex libraries when reading the Probe Barcode information). In these cases, a 5% PhiX spike-in is recommended (10% PhiX for Flex Multiplex libraries on the NovaSeq 6000 and NovaSeq X Series).

3.1 Denaturing and Diluting

Once quantified and normalized, libraries should be denatured and diluted as recommended for Illumina sequencing platforms. Consult Illumina documentation for denaturing and diluting libraries.

3.2 Sequencer Compatibility

Choosing a sequencing platform will depend on many factors including, but not limited to:

- Number of reads desired
- Price
- Access

Compatibility of the sequencers listed below has been verified by 10x Genomics for many, but not all, 10x Genomics libraries. Consult the relevant User Guides to confirm sequencer compatibility for a specific library type. Some variation in assay performance is expected based on sequencer choice.

- MiSeq
- NextSeq 500/550
- NextSeq 1000/2000
- NovaSeq 6000
- NovaSeq X Series

The 10x Genomics Flowcell Capacity Calculator (CG000604) may be used to calculate the total reads required to sequence a pool of 10x Genomics Single Cell libraries, and to select an Illumina platform and flow cell with sufficient read output. Not all library types can be pooled and sequenced together. This calculator will NOT call out these incompatibilities.

10x Genomics libraries can also be modified to enable sequencing on various long and short-read sequencing platforms, with some platforms requiring third-party analysis tools. The 10x Genomics Compatible Products page contains a list of sequencing platforms compatible with 10x Genomics libraries.

3.3 Sequencing by Synthesis Kit Selection

In addition to selecting a sequencer and flow cell, an appropriate SBS cycle kit with sufficient reagents to sequence a specific library type is required. Illumina provides SBS reagents for all sequencing platforms. Having sufficient volume of sequencing reagents is critical for proper run performance. Each cycle of sequencing generates one base of sequencing information for each cluster on the flow cell. Consult this article for more information on how many cycles of sequencing reagents are contained in various Illumina kits.

For example, the NextSeq 1000/2000 P3 flow cell v3 100 cycle kit can generate ~1.2 billion clusters in the flow cell and contains enough sequencing reagents for 138 cycles of sequencing. This means a maximum of 138 bp could be sequenced with this SBS kit, including the index cycles.

At 10x Genomics, libraries are sequenced as specified in the Sequencing section of the appropriate User Guides. It may be possible to modify the sequencing configuration to utilize a SBS kit with fewer cycles; however, this may compromise assay performance and is not recommended.

To determine the appropriate flow cell and SBS kit:

- Based on the total number of read pairs required for a library or library pool, determine the appropriate flow cell. The desired read pairs per cell correspond to Illumina's single-end reads (i.e., the number of clusters) when assessing yield specifications.
- Calculate the total number of required sequencing cycles to identify the appropriate SBS kit by summing all the read lengths from R1, R2, and the index reads (see examples below). Illumina sequencers with forward strand workflows will require an additional 7 cycles of sequencing reagents when performing a dual indexing run. This will also need to be accounted for in the total number of cycles. Consult this article to ensure that the bp required is less than the cycles included for each kit.

Example 1

10x Genomics 3' v3.1 Dual Index libraries have the following recommended sequencing confiugation, which can be found in the relevant User Guides.

	Read 1	i7 Index	i5 Index	Read 2
Purpose	Cell Barcode & UMI	Index	Index	Insert
Length	28	10	10	90

Given this sequencing configuration, 138 bp (28 + 10 + 10 + 90) are needed. This would be appropriate for any Illumina 100 cycle kit, as the cycle number

contained in these kits (138) contains the required number of base pairs (138 bp).

Example 2

10x Genomics 5' v2 Dual Index libraries have the following recommended sequencing confiugation, which can be found in the relevant User Guides.

	Read 1	i7 Index	i5 Index	Read 2
Purpose	Cell Barcode & UMI	Index	Index	Insert
Length	26	10	10	90

Given this sequencing configuration, 136 bp (26 + 10 + 10 + 90) are needed. This would be appropriate for any Illumina 100 cycle kit, as the cycle numbercontained in these kits (138) exceeds the required number of base pairs (136 bp).

Some sequencers may not have an available 100 cycle kit. Exceptions include:

- NovaSeq S4: 200 cycle kit available
- iSeq: 300 cycle kit available
- MiSeq, MiniSeq, and NextSeq: 150 cycle kit is available

3.4 Sequencing Parameters

Number of Cycles Required for Each Read

The following table summarizes the number of cycles required for each read for 10x Genomics libraries. The recommended configuration can also be found in the relevant User Guide. Read lengths described below have been optimized by 10x Genomics. In general, modification of these lengths is not recommended due to negative impacts on assay performance.

Product	Read 1	i7	i5	Read 2	Total
Universal 3' Gene Expression					
Single Cell 3' v4 Gene Expression	28	10	10	90	138
Single Cell 3' v4 Cell Surface Protein					
Single Cell 3' v3.1/HT Gene Expression					
Single Cell 3' v3.1/HT Cell Surface Protein					
Single Cell 3' v3.1/HT CRISPR					
Single Cell 3' v3.1/HT CellPlex					
Universal 5' Gene Expression					
Single Cell 5' v3 Gene Expression	28	10	10	90	138
Single Cell 5' v3 V(D)J					
Single Cell 5' v3 Cell Surface Protein					
Single Cell 5' v3 CRISPR					
Single Cell 5' v2, 5' HT Gene Expression	26	10	10	90	136
Single Cell 5' v2, 5' HT V(D)J					
Single Cell 5' v2, 5' HT Cell Surface Protein					
Single Cell 5' v2, 5' HT CRISPR					
Epi ATAC					
Single Cell ATAC v2	50	8	16	50	124
Epi Multiome					
Single Cell Multiome - ATAC	50	8	24	49	131
Single Cell Multiome - Gene Expression	28	10	10	90	138
Flex					
Flex Singleplex Gene Expression	28	10	10	90	138
Flex Multiplex Gene Expression					
Flex Singleplex Protein Expression					
Flex Multiplex Protein Expression					

Product	Read 1	i7	i5	Read 2	Total
Visium					
Fresh Frozen Direct Placement (v1)	28	10	10	50	90
CytAssist Spatial Gene Expression (v2)					98
CytAssist Spatial Gene & Protein Expression					
HD Spatial Gene Expression	43	10	10	50	113

Sequencing Depth

The following table summarizes the sequencing depth for 10x Genomics libraries. In the sequencing section of the User Guide, the sequencing depth required for data analysis is listed as read pairs per cell. Read pair per cell (reads per cell) correspond to the number of clusters passing filter needed for single cell analysis. One single end read = 1 cluster = 1 read pair = 2 paired end reads.For information on library pooling, see 2.4 Library Pooling Compatibility on page 11.

Product	Minimum Sequencing Depth	Additional Notes
Universal 3' Gene Expression	**Read Pairs per Cell**	
Single Cell 3' v4 Gene Expression	20,000	
Single Cell 3' v3.1 Gene Expression	20,000	
Single Cell 3' v4 Cell Surface Protein	5,000	Pool libraries with Single Cell 3' v4 Gene Expression libraries to maintain nucleotide diversity
Single Cell 3' v3.1/HT Cell Surface Protein/CRISPR/CellPlex	5,000	Pool libraries with Single Cell 3' v3.1 Gene Expression libraries to maintain nucleotide diversity
Universal 5' Gene Expression	**Read Pairs per Cell**	
Single Cell 5' v3 Gene Expression	20,000	
Single Cell 5' v2, 5' HT Gene Expression	20,000	
Single Cell 5' v3 V(D)J	5,000	Pool libraries with Single Cell 5' Gene Expression libraries to maintain nucleotide diversity
Single Cell 5' v3 Cell Surface Protein		
Single Cell 5' v3 CRISPR		
Single Cell 5' v2, 5' HT Gene Expression	20,000	
Single Cell 5' v2, 5' HT V(D)J	5,000	
Single Cell 5' v2, 5' HT Cell Surface Protein		Pool libraries with Single Cell 5' HT Gene Expression libraries to maintain nucleotide diversity
Single Cell 5' v2, 5' HT CRISPR		Pool libraries with Single Cell 5' HT Gene Expression libraries to maintain nucleotide diversity
Epi ATAC	**Read Pairs per Nucleus**	
Single Cell ATAC v2	25,000	

Product	Minimum Sequencing Depth	Additional Notes
Epi Multiome	**Read Pairs per Nucleus**	
Single Cell Multiome - ATAC	25,000	
Single Cell Multiome - Gene Expression	20,000	
Flex	**Read Pairs per Cell**	
Flex Singleplex Gene Expression	10,000	
Flex Multiplex Gene Expression		
Flex Singleplex Protein Expression	5,000	Pool Protein Expression libraries with Gene Expression libraries to maintain nucleotide diversity
Flex Multiplex Protein Expression		
Visium	**Read Pairs per tissue-covered spot (v1/v2) or fully-covered Capture Area (HD)**	
Fresh Frozen Direct Placement (v1)	50,000	Calculating sequencing depth requires estimating the approximate Capture Area (%) covered by tissue. This may be performed visually or by using the Visium Manual Alignment Wizard in Loupe Browser for a more accurate measurement. For more information, consult the Manual Alignment for Visium page on the 10x Genomics Support website.
CytAssist Spatial Gene Expression (v2)	Gene Expression: 25,000 Protein Expression: 5,000	
CytAssist Spatial Gene & Protein Expression		
HD Spatial Gene Expression	275,000,000	Calculating sequencing depth requires estimating the approximate Capture Area (%) covered by tissue. This may be performed visually or by using the Visium Manual Alignment Wizard in Loupe Browser for a more accurate measurement. For more information, consult the Manual Alignment for Visium page on the 10x Genomics Support website. 275M minimum recommendation achieved >50% sequencing saturation for >50% formalin fixed, paraffin embedded (FFPE) tissues tested at 10x Genomics. To achieve similar saturation with fixed and fresh frozen tissues deeper sequencing will be required due to higher sensitivity expected with these tissue preservation methods. For more information, consult the Visium HD Spatial Gene Expression User Guide (CG000685).

4. Library Demultiplexing

4.0 Library Demultiplexing Overview

Multiplexing is the technique of combining multiple samples in a single sequencing run. Each sample is tagged with a unique identifier called an index barcode, allowing them to be processed together. During the sequencing process, the machine reads all the combined samples as a single dataset, producing a mixture of sequences from all the indexed samples. Therefore, demultiplexing is a crucial step in bioinformatics analysis, involving the identification of indices, the sorting of sequences into separate files (one for each original sample), and concatenating the bases from the individual sequencing cycles to a continuous sequence. If demultiplexing is done incorrectly, sequences could be assigned to the wrong samples or no samples, leading to potential errors in downstream analysis. Once the sequences have been demultiplexed from the binary base calling files (BCL files), they are typically stored in the FASTQ format—a text-based file format that contains both nucleotide sequence information and quality scores for each nucleotide. These FASTQ files, usually provided by a sequencing core, are essential inputs for most Cell Ranger and Space Ranger pipelines. However, if the starting with BCL files, they must be converted to FASTQ format through the demultiplexing process. 10x Genomics recommends using Illumina's BCL Convert software to generate compatible FASTQ files.

4.1 Sample Sheets

A sample sheet instructs the sequencer how many cycles each read should be and how to assign generated FASTQ files to the appropriate sample. Though a sample sheet is not required for sequencing 10x Genomics libraries when running the sequencer in 'standalone' mode, a sample sheet is required for demultiplexing after sequencing. If a sample sheet will not be used, the run setup must specify the appropriate number of cycles for Read 1, the i7 and i5 index reads, and Read 2. Information on product-specific sequencing configurations can be found in 3.4 Sequencing Parameters on page 19.

MiSeq instruments running newer versions of the MiSeq Control Software (MCS) v3.1 and above can be run without a sample sheet. MCS v3.1 added manual mode (similar to standalone mode), which allows users to manually enter run parameters without a sample sheet.

A basic sample sheet for BCL Convert has three sections. Each section is described here and example sample sheets are provided for both single and dual indexed samples.

`[Header]`: Can be used to specify the BCL sample sheet version.

`[BCLConvert_Settings]`: In a V2 sample sheet, this section is used to specify several FASTQ conversion settings including whether or not to create FASTQ files for indices. Use `[Settings]` in a V1 sample sheet.

`[BCLConvert_Data]`: In a V2 sample sheet, this section is used to sort samples and index adaptors based on the following column headers. The `[BCLConvert_Data]` section must be renamed `[Data]` or `[data]` for a V1 sample sheet:

Column Name	Description
`Lane`	Optional. Generates FASTQ files only for the samples with the specified lane number. Allows only one valid integer. If the same sample has been run on multiple lanes of the flow cell, add a new row for each lane. If the lane is not specified, indices are searched in all lanes.
`Sample_ID`	The sample ID.
`index`	i7 index sequence
`index2`	i5 index sequence
`Sample_ Project`	Optional Used when --bcl-sampleproject-subdirectories is specified in BCL Convert run. Only alphanumeric characters, dashes, and underscores are allowed. Logs or Reports should not be used as directory names for this flag, as they are already default output directories. Additional information may be found here.

For additional information on using BCL Convert to generate FASTQ files, please refer to the Direct Demultiplexing pages on the 10x Genomics Support site for each product. These links are collected in this article.

If using the Illumina Experiment Manager sample sheet, the sheet may ask for adaptor sequences for trimming. Leave this line blank, as 10x Genomics does not recommend adaptor trimming or preprocessing FASTQ reads before input into 10x Genomics software pipelines. Remove the following lines from the Settings section of the Illumina Experiment Manager sample sheet: Adaptor or AdaptorRead1 or AdaptorRead2.

BCL Convert can also run in Illumina's BaseSpace Sequencing Hub or on-board the instrument. Consult this article for more information on setting up a sample sheet in BaseSpace. When configuring a sample sheet in BaseSpace, the *OverrideCycles* setting may be changed to adjust the read length configuration of the FASTQ if it is different from the RunInfo.xml file. For more information on using *OverrideCycles*, consult this article.

The following is an example sample sheet for libraries created with the Dual Index Plate TT, Set A or Dual Index Plate TN, Set A. The parameter `CreateFastqForIndexReads,0` under `[BCLConvert_Settings]` tells BCL Convert not to generate FASTQ files for indices. Cell Ranger does not require FASTQ files for indices.

```
[Header]
FileFormatVersion,2
[BCLConvert_Settings]
CreateFastqForIndexReads,0
[BCLConvert_Data]
Lane,Sample_ID,index,index2
1,test_sample,GTAACATGCG,AGGTAACACT
```

For general support on using BCL Convert, contact Illumina technical support.

4.2 Undetermined Reads

A high percentage of reads in the 'Undetermined' category is most likely an issue with the sample sheet or a problem during demultiplexing. For example, incorrect indices may be used, causing the reads to not be recognized. This can be checked in the 'Top Unknown Barcodes' section of the `Top_Unknown_Barcodes.csv` file from `bcl-convert`. Additionally, poor quality of the index reads can also lead to a large number of 'Undetermined' reads.

4.3 Specifying Input FASTQ Files for 10x Genomics Pipelines

To serve as inputs for Cell or Space Ranger, FASTQ files should conform to the naming conventions of `bcl-convert`:

[Sample Name]_S1_L00[Lane Number]_[Read Type]_001.fastq.gz

Where Read Type is one of:

- I1: Sample index read (optional)
- I2: Sample index read (optional)
- R1: Read 1
- R2: Read 2

The FASTQ files are specified by providing the path to the folder containing them (via the `--fastqs` argument) and then optionally restricting the selection by specifying the samples and or lanes of interest.

Additionally, Space Ranger accepts file names without the lane number [Lane Number], e.g., sample1_S1_R1_001.fastq.gz.

For more information on specifying FASTQ files for 10x Genomics Pipelines, consult the following:

- Specifying input FASTQs for Cell Ranger
  - Universal 3' Gene Expression
  - Universal 5' Gene Expression
  - Flex
- Specifying input FASTQs for Space Ranger
  - Visium
- Specifying input FASTQs for Cell Ranger ARC
  - Epi Multiome
- Specifying input FASTQs for Cell Ranger ATAC
  - Epi ATAC

4.4 Running BCL Convert

The command to run BCL Convert:

```
/path/to/binary/bcl-convert --bcl-input-directory <folder-with-bcls> \
--output-directory <name-of-output-dir-for-FASTQs> \
--sample-sheet <samplesheet-filename.csv>
```

Required arguments:

- `--bcl-input-directory`: path to the input directory containing BCL files
- `--output-directory`: path to an output directory for newly created FASTQ files. This directory must not exist before command execution.
- `--sample-sheet`: path to a CSV file containing sample information as described in 4.1 Sample Sheets on page 24. Providing a path to the directory instead of the specific CSV file can cause the software to hang.

4.5 FastQC

FastQC is a quality control tool designed to assess the quality of high-throughput sequencing data. Run FastQC if poor sequencing quality is expected or a large number of N's are present in the 10x Barcode read (Read 1). A low number of valid 10x Barcodes can lead to a loss of certain reads during data processing and result in a low median gene count per cell. In such cases, generating a quality control report can be helpful for assessing data quality. For instance, generating a FastQC report on the raw sequencing data (i.e. FASTQ files) is a useful routine check after receiving data from a high-throughput sequencing run. For instructions on downloading FastQC, see this page. For installation instructions, see this page.

Example syntax to run FastQC:

```
fastqc -t 8 -f fastq -o ./output_directory_name/ /path_to_fastq_files_
dir/*.gz
```

The example above uses the downloaded FastQC executable file (fastqc), the number of threads for the program to use (`-t`), the input file type (`-f`), an output directory name to save the FastQC reports to a location other than the current directory (`-o`), and the full path to the input FASTQ files.

Notices

Document Number

CG000809 | Rev A

Legal Notices

Instrument & Licensed Software Updates Warranties

Updates to existing Instruments and Licensed Software may be required to enable customers to use new or existing products.

Support

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

Document Revision Summary

Document Number

CG000809

Title

10X Genomics Sequencing Handbook Handbook

Revision

Rev A

Revision Date

February 14, 2025

Description of Changes

- N/A - Initial release

Take 1 minute to evaluate this document. Scan this code or click here.

Updated for general minor consistency of language and terms throughout.