NEON Metagenomics Standard Operating Procedure, v.5

Prepared for:
Battelle Ecology, Inc.

Prepared By:
Battelle Memorial Institute
505 King Avenue
Columbus, Ohio 43201

Effective Date: 02/28/2024

BATTELLE

I. Version History

Version 5	• Updated KAPA HyperPrep Manual from KR0961– v9.20 to newest manual version KR0961– v10.23. • Replaced Covaris microTUBE-15 with microTUBE-50 to allowed up to 55 µL sample volume input for fragmentation. • Removed Qiagen MinElute Cleanup and replaced with 0.8X bead-based cleanup recommended in KAPA HyperPrep Manual. • Updated reagents on Materials list to include BluePippin cassette with control DNA for BluePippin run, NovaSeq Xp 4-lane kit with NovaSeq™ Xp 4-Lane Manifold manifold, 1M Tris-HCl, pH 8.0 for neutralization and NovaSeq 6000 S4 sequencing reagent kit. • Increased the number of samples prepared and sequenced from 60 to 192. • Increased the upper limit of the double-sided size selection from 350 bp to 450 bp. • Included a Sage Science Blue-Pippin size selection after pooling normalized libraries to completely remove adapter dimers. • Sequencing instrument changed from Illumina NextSeq 550 to NovaSeq 6000 sequencing on an S4 flow cell using the Illumina Xp workflow.
Version 4	Updated references to manufacturer manuals and catalog numbers, clarified statements regarding use of negative controls, clarified NextSeq Sample Sheet Preparation, combined DNA denaturation reagent tables for simplicity, added information for confirming presence of DNA libraries prior to sequencing using an automated electrophoresis system, added calculations used for determining sequencing library size.
Version 3	The number of samples prepared and sequenced on each run was increased to 60 from 40.

II. Objective and Overview

Using the KAPA HyperPrep kit from Roche Kapa Biosystems, two 96-sample plates (a total of 192 samples) are prepared for metagenomic shotgun sequencing (**Figure *1***). Each 96-well plate contains a negative control well (an extraction blank well not containing sample processed with the extraction reagents kit). The 96-well sample plates are indexed using the entire KAPA Unique Dual Indexed-Adapter kit (Roche KAPA Biosystems).

The final libraries are bead-based cleaned, quantified by qPCR, normalized, and combined into four pools (4) of forty-eight (48) samples per pool. If free adapters or adapter dimers are present after pooling, the affected library pool(s) will undergo a BluePippin size selection to remove the adapters or adapter dimers before qPCR quantification and normalization.

Using the Illumina NovaSeq Xp 4-Lane Kit v1.5, each lane in a 4-lane Illumina NovaSeq S4 flow cell is primed with one of the 4 size-selected library pools, resulting in 48 adapter-indexed libraries per lane. The S4 flow cell containing 192 libraries is sequenced on an Illumina NovaSeq 6000 (Manual: Document # 1000000019358 v17).

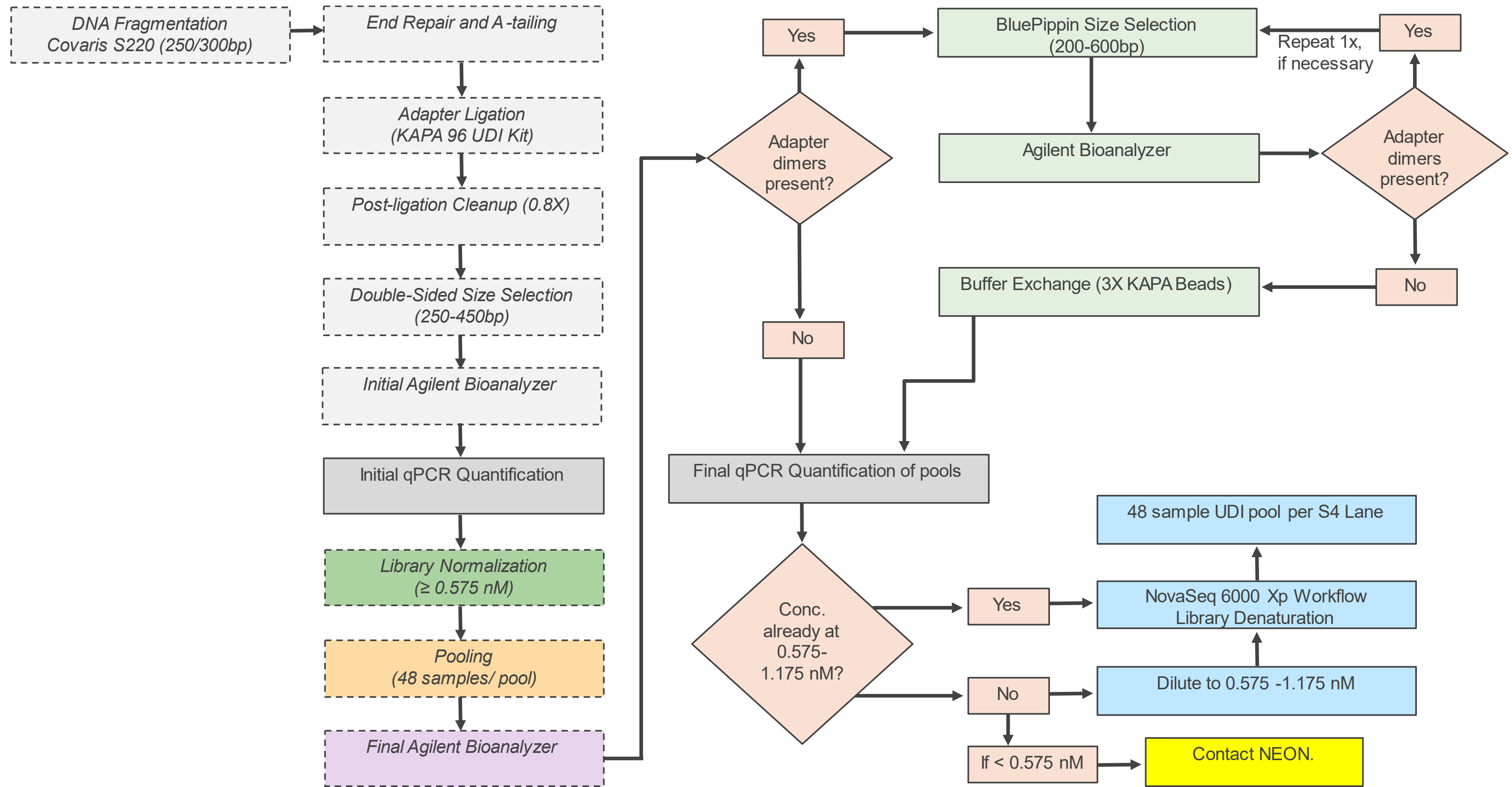


Figure 1. NEON Metagenomics Library Preparation Process Workflow.

III. Recommended Materials

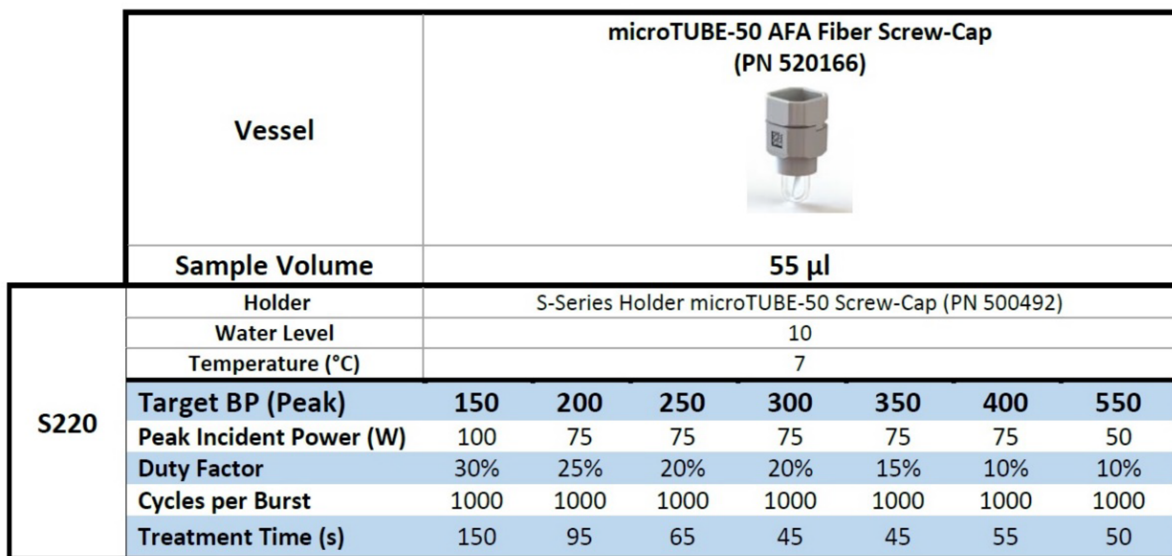
Material	Manufacturer	Catalog #
microTUBE-50 AFA Beads Screw-Cap	Covaris	520166/ 520167
KAPA HyperPrep Kit	KAPA Biosystems	KK8505
KAPA Pure Beads	KAPA Biosystems	KK8002
KAPA Unique Dual-Indexed Adapter	KAPA Biosystems	KK8727
KAPA Library Quant (Illumina) Universal qPCR Mix Kit	KAPA Biosystems	KK4824
2X KAPA SYBR Fast qPCR Master Mix Universal Kit	KAPA Biosystems	KK4600
BluePippin; 2% Agarose, dye-free, with internal/external markers; 100-600 bp	Sage Science	BDF2010
Control DNA for BluePippin 2% Agarose Cassettes with External Marker	Sage Science	CON2004
MinElute PCR Purification Kit	Qiagen	28006
1M Tris-HCl, pH 8.5	Teknova	T5080
1M Tris-HCl, pH 8.0	Thermo Scientific	AAJ22638AE
2N NaOH	Fisher	SS264-1
NovaSeq Xp 4-Lane Kit v1.5	Illumina	20043131
NovaSeq™ Xp 4-Lane Manifold Pack	Illumina	20021667
NextSeq PhiX Control Kit	Illumina	FC-110-3002
NovaSeq 6000 SP Reagent Kit v1.5 (300 cycles): • **S4 Flow Cell** • **Library Tube** • **SBS Cartridge** • **Cluster Cartridge** • **Buffer Cartridge**	Illumina	20028400 20025962 20014012 20031051 20031056 20003187

IV. Procedure

All work takes place in a decontaminated biosafety cabinet (BSC). Prior to work, the BSC is wiped down with 10% bleach, followed by 70% isopropyl alcohol. The germicidal light is turned on for a minimum of 10 minutes.

A. DNA Fragmentation

The initial step in the NEON metagenomics procedure is to mechanically fragment EDTA-free double-stranded DNA (dsDNA) in a Covaris microTUBE-50 to a target base pair (bp) peak of 250 bp using the Covaris S220 focused-ultrasonicator (S220 User Manual; PN 010122, Revision I, May 2017) settings in **Figure *2*** (Covaris DNA Shearing Quick Guide: 010368 Rev D, January 2017). The Covaris shears DNA via an acoustic transducer that emits pulses of energy at user-specified frequencies.



55 µl sample volume - from 150 to 550 bp

	Vessel	microTUBE-50 AFA Fiber Screw-Cap (PN 520166)						
	Sample Volume	55 µl						
S220	Holder	S-Series Holder microTUBE-50 Screw-Cap (PN 500492)						
	Water Level	10						
	Temperature (°C)	7						
	Target BP (Peak)	150	200	250	300	350	400	550
	Peak Incident Power (W)	100	75	75	75	75	75	50
	Duty Factor	30%	25%	20%	20%	15%	10%	10%
	Cycles per Burst	1000	1000	1000	1000	1000	1000	1000
	Treatment Time (s)	150	95	65	45	45	55	50

Figure 2. DNA Fragmentation Settings When Using MicroTUBE-50 on a Covaris S220 Focused-Ultrasonicator.

B. KAPA PCR-Free HyperPrep Library Preparation

The fragmented DNA samples are converted to Metagenomic DNA sequencing libraries using the KAPA HyperPrep Kit (Manual: KR0961 – v10.23) for sequencing on an Illumina sequencing platform as described below.

1. End Repair and A-tailing

In the same plate containing the fragmentation reactions, the end repair and A-tailing reactions are assembled according to the manufacturer's instructions (KAPA HyperPrep Kit Manual: KR0961 – v10.23). This allows unique dual-index adapters to be ligated to fragmented DNA samples in the next step.

2. Indexed-Adapter Ligation

Prior to ligating the KAPA unique dual-indexed (UDI) adapters (KAPA Unique Dual-Indexed Adapter Manual KR1736 – v2.19) to an end-repaired DNA fragment, each KAPA UDI Adapter stock is diluted with a KAPA Adapter Dilution buffer to the appropriate concentration corresponding to the starting input DNA as outlined in Table 1 below.

Table 1. Adapter stock concentrations for Sequencing Libraries Constructed from 1 ng – 1 µg.

Input DNA	Adapter Stock Concentration	Adapter : Insert molar ratio
1 µg	15 µM	10:1
500 ng	15 µM	20:1
250 ng	15 µM	40:1
100 ng	15 µM	100:1
50 ng	15 µM	200:1
25 ng	7.5 µM	200:1
10 ng	3 µM	200:1
5 ng	1.5 µM	200:1
2.5 ng	750 nM	200:1
1 ng	300 nM	200:1

In the same plate used to perform the end repair and A-tailing, the indexed-adapter ligation reactions are assembled according to the manufacturer's instructions (KAPA HyperPrep Kit Manual: KR0961 – v10.23).

3. Post-Ligation Cleanup

After the indexed-adapter ligation step, a 0.8X KAPA Pure Beads cleanup is used to clean the post-ligation samples to eliminate any remaining reagents from the library preparation that may interfere with sequencing, according to the manufacturer's instructions (KAPA HyperPrep Kit Manual: KR0961 – v10.23; page 12).

4. Double-Sided Size Selection

Next, a double-sided size selection using KAPA Pure beads is performed to select DNA library molecules (inclusive of adapters) in the range of 250-450 bp (KAPA HyperPrep Kit Manual: KR0961 – v10.23).

5. Initial Agilent Bioanalyzer

Prior to library quantification, a negative extraction control and ten randomly selected PCR-free libraries are loaded onto a High Sensitivity DNA Chip (Agilent) and analyzed with an automated electrophoresis system, such as the Agilent Bioanalyzer, to confirm the presence of DNA libraries. The negative control should be loaded "neat" on the Bioanalyzer chip with no peaks detected on the electropherogram in the range of 250-450 bp. If a peak is detected in the negative control within the 250-450 bp range, contact the principal investigator.

C. Initial qPCR Library Quantification

Prior to normalization and pooling, the PCR-free double-sided size selected DNA libraries are quantified using a KAPA Biosystems Library Quantification Universal qPCR (Quantitative Polymerase Chain Reaction) kit for Illumina Platforms, according to the manufacturer's instructions (KAPA Library Quantification Kit Manual: KR0405 – v12.23). The qPCR assay is run on a 7500 Real-Time system (Thermo Fisher) according to the conditions described Table 2.

Table 2. **qPCR conditions on 7500 Real-Time System.**

Step	Temperature	Time	Cycles
Initial Denaturation	95°C	5 minutes	1
Denaturation	95°C	30 seconds	35
Annealing/Extension/Data acquisition	60°C	45 seconds	

The average Quantification Cycle (Cq) score obtained from the qPCR data for each DNA Standard is plotted against log10 (concentration in pM) to generate a standard curve. The concentrations of diluted library samples are then calculated against the standard curve using absolute quantification. Any library with a concentration >0.0002 pM can be quantified with the kit.

Due to the "forked" adapters in PCR-free workflows, the DNA libraries will appear to have a longer than expected size and/or may display a broad or bimodal size distribution on an Agilent Bioanalyzer electropherogram. Due to this configuration, a base pair size of 373 bp representing a complete full-length library (dual-indexed adapter ligated DNA fragment) will be used instead of the size provided by the Bioanalyzer traces.

The sample libraries concentrations are calculated using the results of the qPCR and the expected library size. Sample concentrations are determined based on the following formula:

$$\frac{(Average\ Conc.[pM])(452\ bp\ standards)}{373\ bp\ average\ amplicon} = W$$

W X D = Library Conc in pM

W = Library concentration in pM, D = dilution of samples input into Kapa Quant, i.e., if samples are diluted 1:10,000, then D=10,000

The library concentration is converted from picomolar (pM) to nanomolar (nM) using the conversion factor below:

$$\frac{pM\ Library\ conc}{1{,}000} = nM\ library\ conc.$$

D. Library Normalization

Each DNA library is diluted with sample dilution buffer (10 mM Tris-HCl pH 8.5) to the same concentration (0.575 nM or higher) for multiplexing. If the library concentration is less than 0.575 nM, contact the principal investigator.

Any size-selection proceeding this normalization step could potentially lead to sample loss; therefore, it is recommended to normalize to a value higher than the NovaSeq 6000 Xp Workflow loading concentration of 0.575 – 1.175 nM.

E. Pooling

Each of the two 96-well library plates is split into two pools resulting in a total of four library pools (see **Figure *3*** below) .

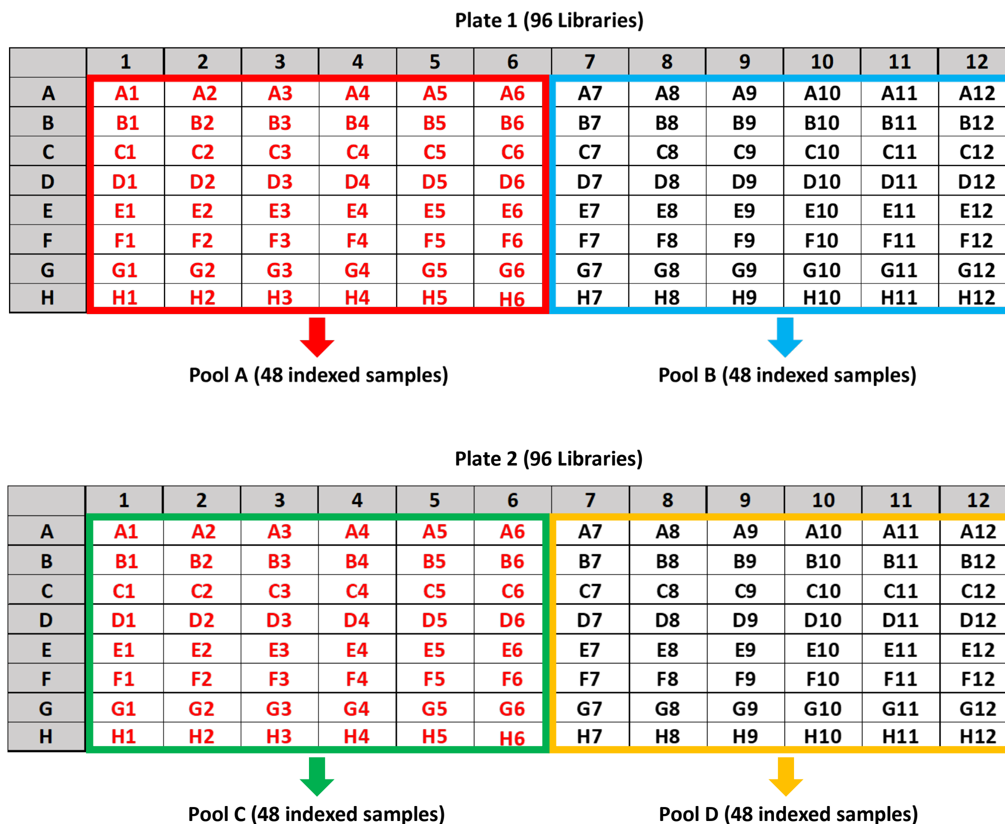


Plate 1 (96 Libraries)

	1	2	3	4	5	6	7	8	9	10	11	12
A	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
B	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
C	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
D	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
E	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
F	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
G	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
H	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12

Plate 2 (96 Libraries)

	1	2	3	4	5	6	7	8	9	10	11	12
A	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
B	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
C	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
D	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
E	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
F	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
G	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
H	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12

Figure 3. Individual Library Pools (A, B, C and D) Prepared for Each Lane of NovaSeq S4 Flow Cell.

F. Final Agilent Bioanalyzer

All the library pools (A, B, C and D) are loaded onto a High Sensitivity DNA Chip (Agilent) and analyzed with an automated electrophoresis system, such as the Agilent Bioanalyzer, to confirm the absence of adapters and adapter dimers. If adapters and/or adapter dimers are present, the library pools are size selected using a Sage Science BluePippin size selection instrument in Section IV. G. BluePippin Size Selection and Buffer Exchange. However, if no adapters and/or adapter dimers are present, the library pools will skip the BluePippin Size Selection and Buffer Exchange step and proceed to Section IV. H. Final qPCR Library Quantification.

G. BluePippin Size Selection and Buffer Exchange

Library pools with adapters or adapter dimers present are loaded onto a Sage Science BluePippin 2% (100-600 bp) Agarose Gel Cassette (Manual; Quick Guide BDF2010 Marker V2; 2021) and size selected by running on a Sage Science BluePippin gel electrophoresis system (BluePippin Manual; 460013 Rev K) to remove all free adapters and adapter dimers prior to loading on the NovaSeq 6000 sequencer.

Next, the size selected library pools are loaded onto a High Sensitivity DNA Chip and analyzed with an automated electrophoresis system, such as the Agilent Bioanalyzer, to confirm the removal of the free adapters and adapter dimers. If adapters and/or adaptor-dimers are still present, one additional size selection using the BluePippin gel electrophoresis system may be performed. If the additional size selection is not successful in removing the remaining adapter dimers, contact NEON to determine path forward.

Each library pool that undergoes BluePippin size selection will be buffer exchanged to 10 mM Tris-HCl pH 8.5 using a 3X KAPA bead-based cleanup (KAPA Pure Beads Manual; KR1245 – v4.17) with an elution volume of at least 35 µL.

H. Final qPCR Library Quantification

Prior to sequencing, all four library pools (A, B, C, and D) will be quantified by qPCR as described in Section IV. C Initial qPCR Library Quantification.

I. Metagenomics Shotgun Sequencing

1. NovaSeq 6000 Sample Sheet Preparation

Initiating the sequencing run on the NovaSeq 6000 does not require the use of a sample sheet. However, a sample sheet is required to demultiplex the run data. The sample sheet required to demultiplex the NovaSeq 6000 run data is created by using a ".csv" file containing both the sample ID associated with its corresponding unique dual-indexed adapter sequence. The adapter

sequences are provided in the KAPA Unique Dual-Indexed Adapter kit manual for Illumina Platforms (KAPA Unique Dual-Indexed Adapter Manual; KR1736 – v3.20).

2. NovaSeq 6000 Xp Workflow for Library Denaturation for S4 Flow Cell

Prior to sequencing, each library pool is diluted to a final volume of 30 µL with a Pool Loading Concentration in the range of 0.575-1.175 nM (Table 3) using 10 mM Tris-HCl pH 8.5.

Table 3. Pool Loading Concentration Range Required Per Library Type on NovaSeq 6000.

Library Type	Pooled Loading Concentration (nM)	Final Loading Concentration on NovaSeq 6000 (pM)
PCR-free library	0.575 - 1.175	115 - 235

A volume of 1.1 µL of nondenatured PhiX (0.25 nM) is added into each pooled library (A, B, C and D) for a 1% spike. Freshly diluted 0.2 N NaOH is combined with each of the library/PhiX mixture in a microcentrifuge tube according to Table 4. Then, 400 mM Tris-HCl (pH 8) is added to each sample tube according to Table 4 to neutralize.

Table 4. *Illumina Xp Workflow for Denaturing Libraries for NovaSeq 6000 S4 Flow Cell.*

Mode	Total Volume of Pool Per Lane	Nondenatured 0.25 nM PhiX for 1% PhiX Spike-in Volume	0.2 N NaOH Volume	400 mM Tris-HCl pH 8.0 Volume
S4	30 µL	1.1 µL	7 µL	8 µL

3. Loading and Running the NovaSeq 6000

Prior to loading the S4 flow cell with denatured Library/PhiX mixture, the NovaSeq 6000 S4 flow cell, SBS cartridge, Cluster cartridge (with an empty library tube inserted into position #8) and buffer are prepared according to Illumina's Denature and Dilute Libraries Guide (NovaSeq 6000 System Denature and Dilute Libraries Guide Document # 1000000106351 v04; June 2022) and loaded onto the NovaSeq 6000. The sequencing run is set up by following the appropriate prompts on the NovaSeq 6000. During run set up, the "paired-end" option is selected, as well as "151bp" for both Cycle Read 1 and Cycle Read 2 and "8 bp" for both Index 1 and Index 2.

Illumina ExAmp Master Mix (DPX1/JPX1, DPX2/JPX2, and DPX3) is prepared according to manufacturer's instructions (NovaSeq 6000 Sequencing System Guide Document # 1000000106351 v04; June 2022) and added to each of the 4 denatured pools. The S4 flow cell is removed from the NovaSeq 6000 instrument and loaded onto an Illumina NovaSeq Xp Flow Cell Dock. Each Library/ExAmp mixture is added to a single lane on the Illumina Xp 4-lane manifold well and allowed to sit for at least 2 minutes for the mixed to reach the opposite end of each lane on the S4 flow cell. The sequencing run is started within 30 minutes of loading libraries onto the S4 flow cell.

V. Quality Review

Individual samples with fewer than 5,000,000 reads are reanalyzed by the lab. Samples should have greater than 5 million reads, a minimum Q-score of 20, and a maximum of 1 ambiguous base per read as shown in Table 5.

Table 5. Quality Control and Acceptance Criteria.

QA/QC measurement	Frequency	Requirement	Corrective Action
Presence of adapter dimers after size selection using a Sage Science BluePippin gel electrophoresis system	Per library pool	None present	If adapter dimers are still present following the size selection, one additional size selection step may be performed. If adapter dimers are present following a second size selection, halt processing the batch and contact NEON immediately to evaluate resolution. NEON must approve the resolution. Potential outcomes of evaluation: continue with analysis as normal; continue with analysis with additional quality flags; restart library preparation from Covaris fragmentation
Q-Score	Per batch	> 20	Contact NEON for resolution.
Filtered Reads	Per sample	> 5 million reads	Contact NEON for resolution.
Ambiguous bases per read	Per sample	≤ 1	Contact NEON for resolution.