3- Illumina Library Prep & Sequencing

Special focus on SARS-CoV-2 genome sequencing protocols
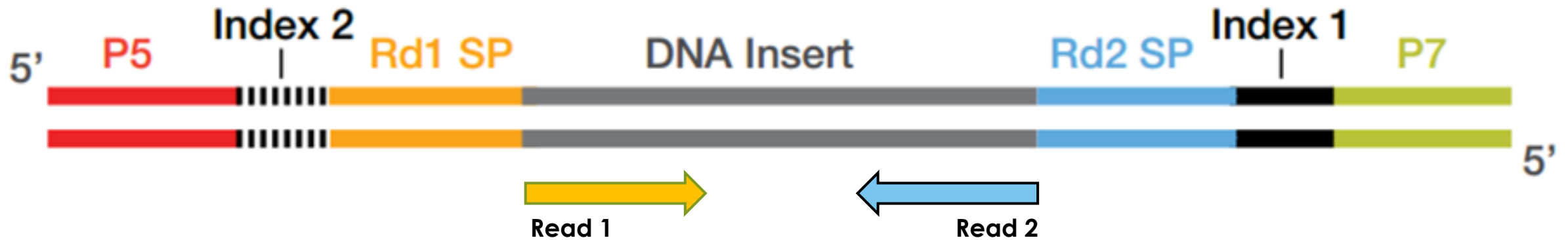
DARRELL L. DINWIDDIE, PHD

DARYL B. DOMMAN, PHD

Illumina Sequencing

Illumina Library Structure

Rd1 SP- Read 1 Sequencing Primer
Rd2 SP- Read 2 Sequencing Primer

https://www.illumina.com/techniques/sequencing/ngs-library-prep.html

Choosing a Tagmentation Kit for Your Experiment

Applications	Product	Benefits
16s rRNA Sequencing, Amplicon Sequencing, *De Novo* Sequencing, Shotgun Sequencing, Whole-Genome Sequencing	**Nextera XT**	•Optimized for research on small genomes, PCR amplicons, and plasmids
Amplicon Sequencing, *De Novo* Sequencing, Shotgun Sequencing, Whole-Genome Sequencing	**Illumina DNA Prep**	•No library quantification needed
Human Whole-Genome Sequencing	**Illumina DNA PCR-Free**	•Bead-linked transposome technology •Avoid PCR duplicates •No library quantification needed

Nextera XT

Assay Time	~5.5 hours from DNA extraction to normalized library. (Library prep time: ~90 minutes).
Hands-On Time	15 minutes
Mechanism of Action	Enzymatic fragmentation
Multiplexing	Up to 384 uniquely indexed samples may be pooled and sequenced together.
Input Quantity	1 ng DNA
Species Category	Any Species, Nematode, Plant, Zebrafish, Fungal, Mouse, Mammalian, Virus, Bacteria, Drosophila, Rat, Human, Yeast
Species Details	Compatible with any species
Target Insert Size	300 bp–1.5 kb
Description	Fast library prep optimized for research on small genomes, PCR amplicons, and plasmids.
Specialized Sample Types	Low-Input Samples, Not FFPE-Compatible, Single Cells
Technology	Sequencing
Method	16s rRNA Sequencing, Amplicon Sequencing, De Novo Sequencing, Shotgun Sequencing, Whole-Genome Sequencing
System Compatibility	iSeq 100, MiniSeq, MiSeq, NextSeq 1000, NextSeq 2000, NextSeq 500, NextSeq 550
Automation Capability	Liquid Handling Robots
Variant Class	Single Nucleotide Polymorphisms (SNPs), Structural Variants
Nucleic Acid Type	DNA

DNA Prep

Assay Time	~3-4 hours (from DNA extraction to normalized library)
Hands-On Time	1-1.5 hours
Input Quantity	<mark>Small genomes (e.g. microbial): 1-500 ng DNA</mark>. Large genomes (e.g. human): 100-500 ng DNA.
Mechanism of Action	<mark>Bead-linked transposome</mark>
Multiplexing	Up to 384 unique dual (UD) combinations and 96 combinatorial dual (CD) combinations
Species Details	Compatible with any species
System Compatibility	HiSeq 2500, HiSeq 3000, HiSeq 4000, HiSeq X Five, HiSeq X Ten, iSeq 100, MiniSeq, MiSeq, MiSeqDx in Research Mode, NextSeq 1000, NextSeq 2000, NextSeq 550, NextSeq 550Dx in Research Mode, NovaSeq 6000
Specialized Sample Types	Blood, Not FFPE-Validated, Saliva
Sample Type Details	Supports multiple sample types, including genomic DNA, blood (and dried blood spots), saliva, PCR amplicons, plasmids, and bacterial colonies
Species Category	Any Species, Human, Mouse, Rat, Plant, Drosophila, Virus, Yeast, Zebrafish, Bacteria, Mammalian, Nematode
Description	A fast, flexible workflow for a wide range of research applications and sample types, from human to microbial whole-genome sequencing and more.
Target Insert Size	<mark>~350bp</mark>
Technology	Sequencing
Method	Amplicon Sequencing, De Novo Sequencing, Shotgun Sequencing, Whole-Genome Sequencing
Variant Class	Chromosomal Abnormalities, Copy Number Variants (CNVs), Gene Fusions, Germline Variants, Insertions-Deletions (indels), Loss of Heterozygosity (LOH), Single Nucleotide Polymorphisms (SNPs), Somatic Variants, Structural Variants
Automation Capability	Liquid Handling Robots
Nucleic Acid Type	DNA

Choosing Sequencing Reagents

Illumina MiSeq V2 vs V3 Reagents

- ▶ **Cluster Density**
  - ▶ **V2- 467-583 k/mm2**
  - ▶ **V3- 727-827 k/mm2**

- ▶ **Read Length**
  - ▶ **V2- 2x25bp, 2x150bp, 2x250bp**
  - ▶ **V3- 2x75bp or 2x300bp**

- ▶ **Total Output**
  - ▶ **V2- 24-30 Million Paired Reads**
  - ▶ **V3- 44-50 Million Paired Reads**

- ▶ **Sequencing Time**
  - ▶ **V2- 5.5-39 hours**
  - ▶ **V3- 21-56 hours**

Illumina MiSeq Specifications

Table 1: MiSeq System performance parameters

Read length	Total time[a]	Output	Quality scores[b]	Single reads[c]	Paired-end reads[c]
MiSeq Reagent Kit v2					
2 × 25 bp	~5.5 hours	750-850 Mb	> 90% bases higher than Q30		
2 × 150 bp	~24 hours	4.5-5.1 Gb	> 80% bases higher than Q30	12-15M	24-30M
2 × 250 bp	~39 hours	7.5-8.5 Gb	> 75% bases higher than Q30		
MiSeq Reagent Kit v3					
2 × 75 bp	~21 hours	3.3-3.8 Gb	> 85% bases higher than Q30	22-25M	44-50M
2 × 300 bp	~56 hours	13.2-15 Gb	> 70% bases higher than Q30		
MiSeq Reagent Kit v2 Micro					
2 × 150 bp	~19 hours	1.2 Gb		4M	8M
MiSeq Reagent Kit v2 Nano					
2 × 150 bp	~17 hours	300 Mb		1M	2M
2 × 250 bp	~28 hours	500 Mb			

Illumina Product Numbers

Ordering information

Product	Catalog no.
MiSeq System	SY-410-1003
MiSeq Reagent Kit v2 (50-cycles)[a]	MS-102-2001
MiSeq Reagent Kit v2 (300-cycles)[a,b]	MS-102-2002
MiSeq Reagent Kit v2 (500-cycles)[a,b]	MS-102-2003
MiSeq Reagent Kit v3 (150-cycle)[b]	MS-102-3001
MiSeq Reagent Kit v3 (600-cycle)[b]	MS-102-3003
MiSeq Reagent Micro Kit v2 (300-cycles)[b]	MS-103-1002
MiSeq Reagent Nano Kit v2 (300-cycles)[b]	MS-103-1001
MiSeq Reagent Nano Kit v2 (500-cycles)	MS-103-1003

Loading Illumina NGS Library
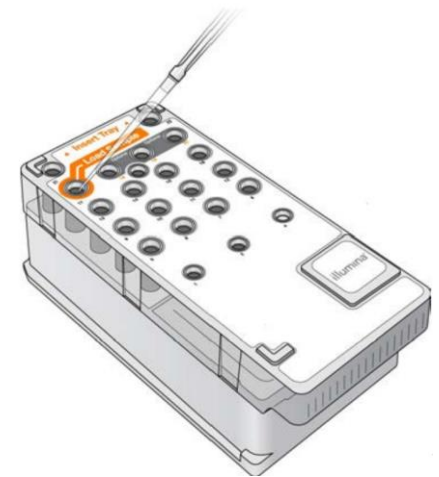
Determine Molarity

Standard Normalization Calculations

1. After running the libraries on the Qubit and Fragment Analyzer, the following formula was used to calculate the nM concentration of each library:

$$nM = ((ng/ul)/ (avg\ bp\ size \times 660\ g/mol))* 1{,}000{,}000$$

Qubit= 55ng/ul

**Ave bp size
475bp**

$$nM= ((55ng/ul)/(475bp*660g/mol)) * 1{,}000{,}000$$

$$nM= 175.4$$
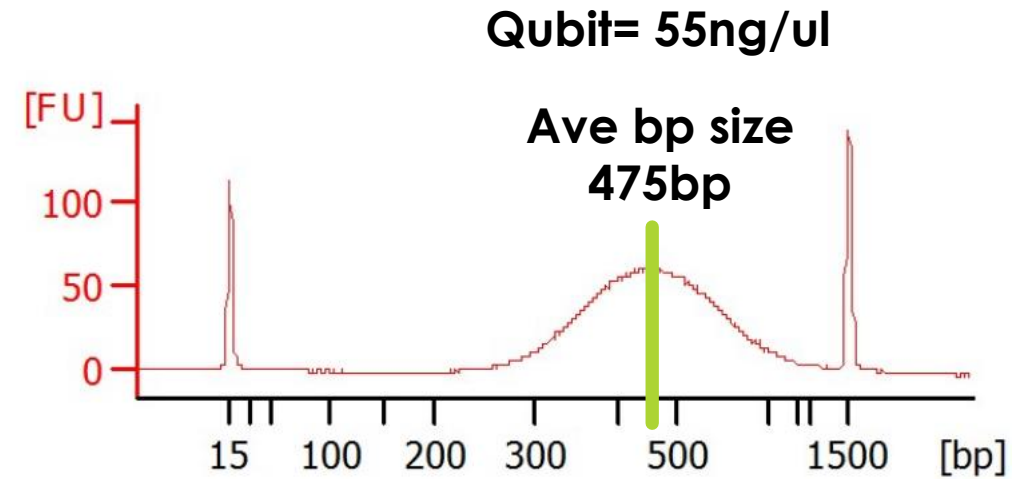
Determining Molarity

| Number	Sample	Concentration (ng/ul)	Fragment Size	nM
1	Virus_RNAseq_Pool_2	55	475	175.4
2	Virus_RNAseq_Pool_3	10	475	31.9
3	Virus_RNAseq_Pool_4	10	400	37.9
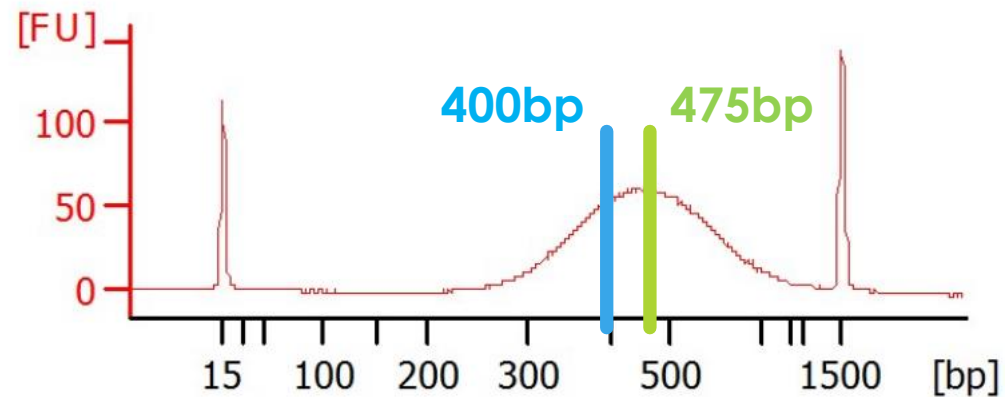
Generate 4nM Library

Number	Sample	Concentration (ng/ul)	Fragment Size	nM	Total ul of stock	Library to make 4 nM stock (ul)	TE Buffer
1	Virus_RNAseq_Pool_2	55	475	175.4	100	2.28	97.72
2	Virus_RNAseq_Pool_3	10	475	31.9	50	6.27	43.73
3	Virus_RNAseq_Pool_4	10	400	37.9	50	5.28	44.72

Determine Average bp Size
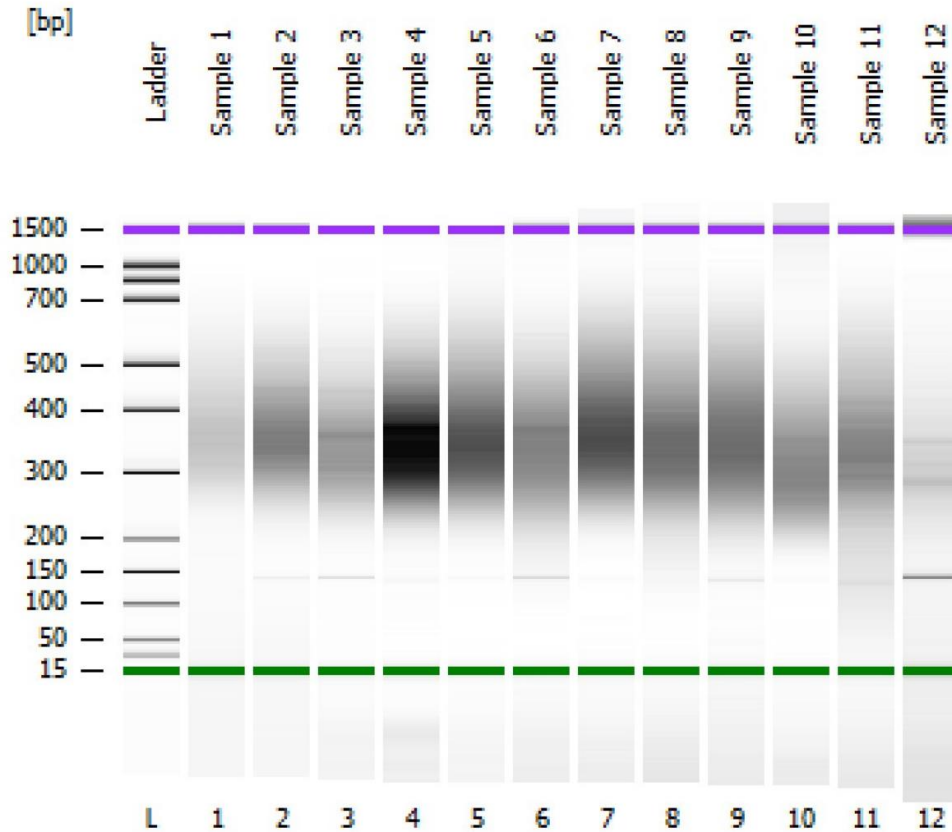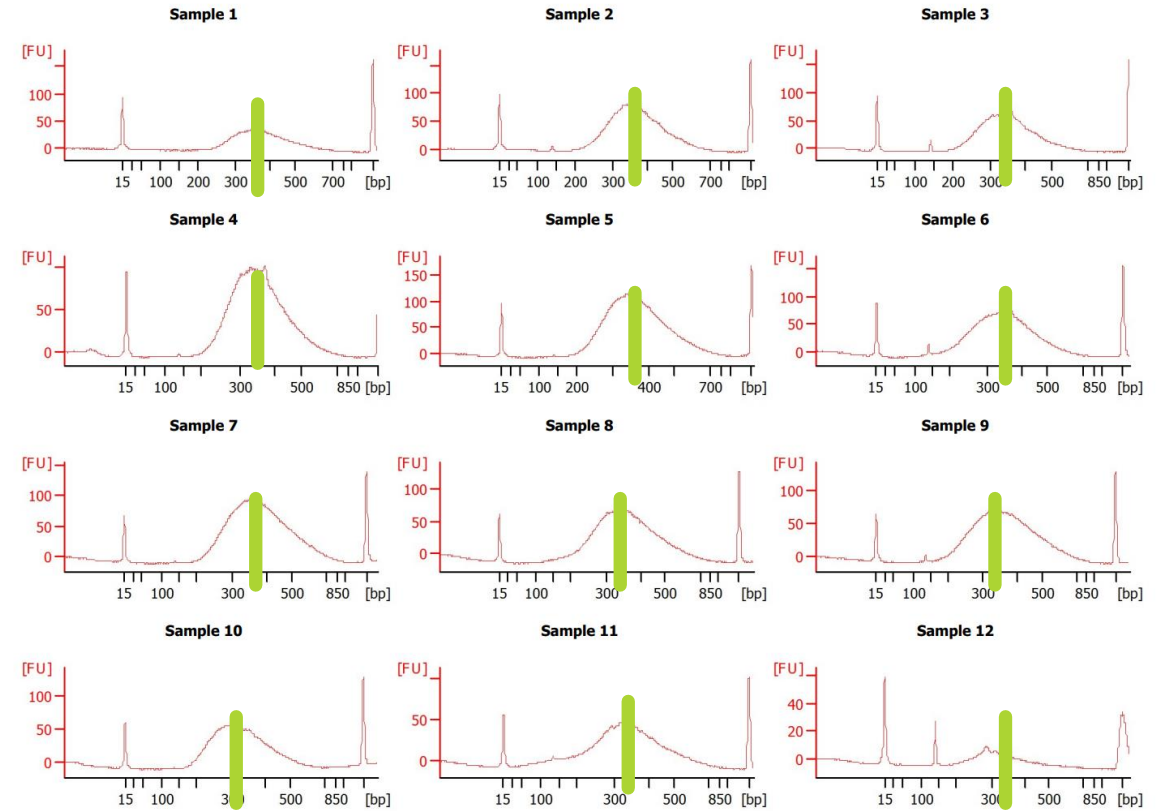
Denaturation

Denature a 4 nM Library

1   Combine the following volumes in a microcentrifuge tube.
   ▶   4 nM library (5 µl)
   ▶   0.2 N NaOH (5 µl)

2   Vortex briefly and then centrifuge at 280 × g for 1 minute.

3   Incubate at room temperature for 5 minutes.

4   Add 990 µl prechilled HT1 to the tube containing denatured library.
    The result is 1 ml of a 20 pM denatured library.

Dilute Denatured 20 pM Library

1   Dilute to the desired concentration using the following volumes.

Concentration	6 pM	8 pM	10 pM	12 pM	15 pM	20 pM
20 pM library	180 µl	240 µl	300 µl	360 µl	450 µl	600 µl
Prechilled HT1	420 µl	360 µl	300 µl	240 µl	150 µl	0 µl

# Loading Concentration

Loading Volume and Concentration

This procedure denatures and dilutes libraries to a final volume of 600 μl. The recommended loading concentration varies depending on the version of MiSeq Reagent Kit used for the sequencing run. In practice, loading concentration can vary depending on library preparation and quantification methods.

Chemistry	Recommended Final Loading Concentration
MiSeq Reagent Kit v3	Supports 6–20 pM loading concentration. Requires at least a 4 nM library before diluting and denaturing.
MiSeq Reagent Kit v2	Supports 6–10 pM loading concentration.

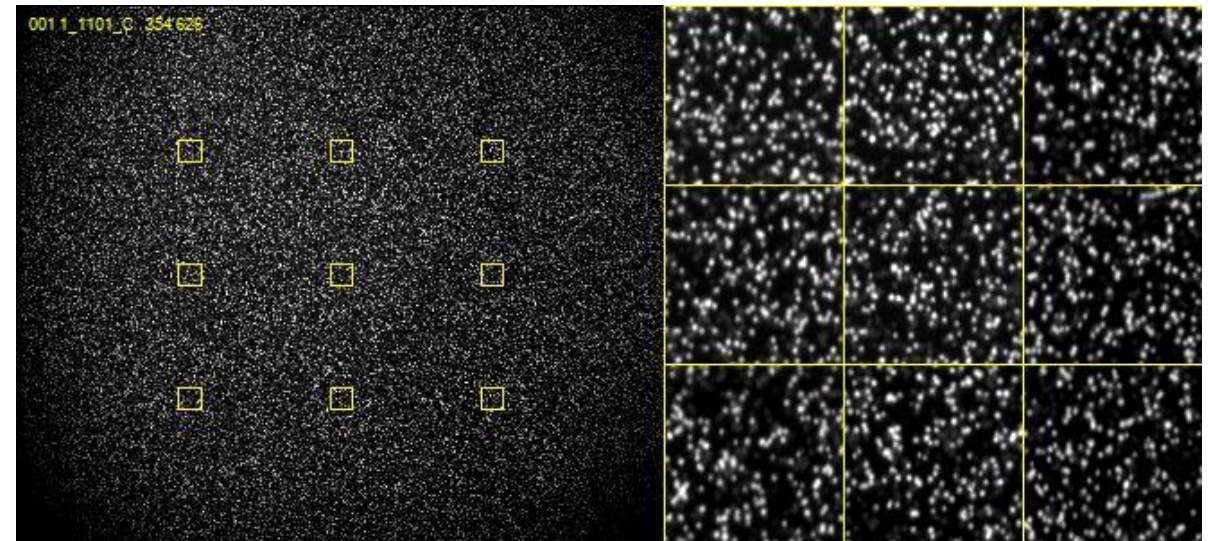MiSeq System- Denature and Dilute Libraries Guide Document # 15039740 v10

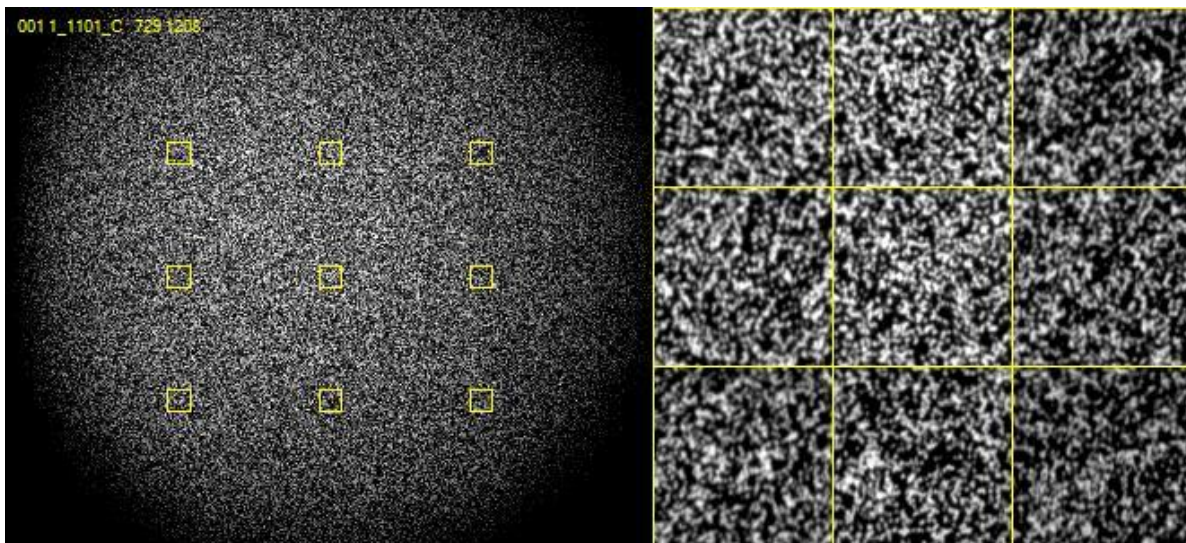# Optimal Clustering

Under Clustered
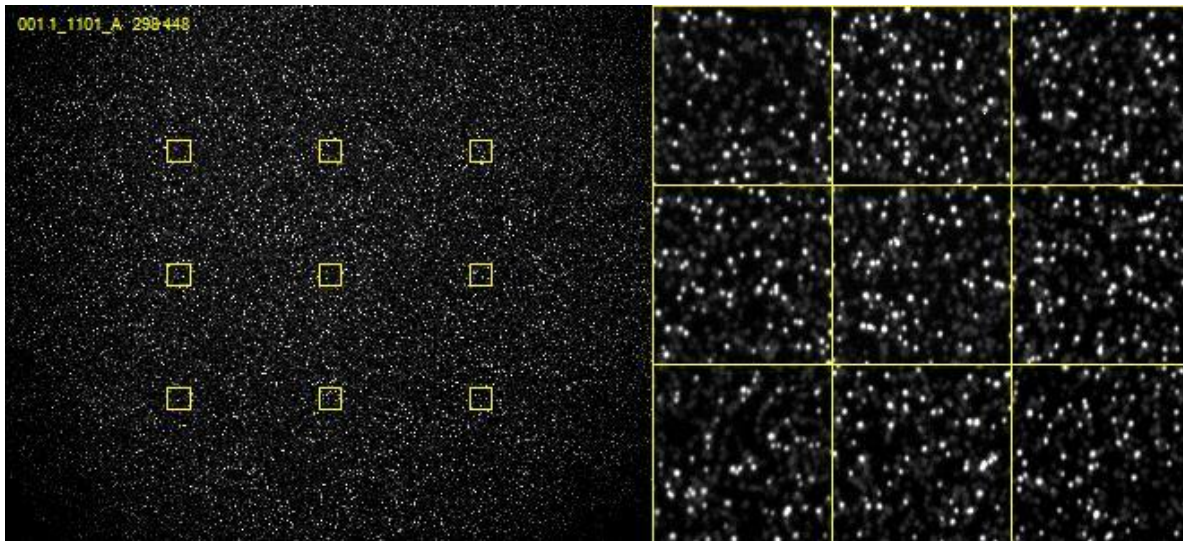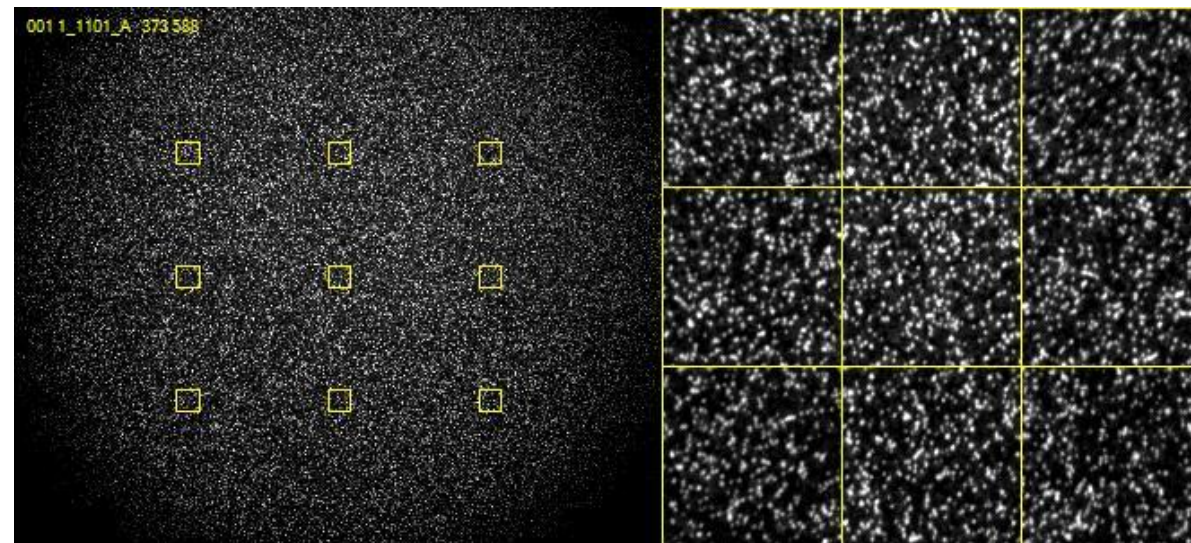

6 pM

Optimal Clustered

10 pM

Optimal Clustering

Over Clustered

Optimal Clustered

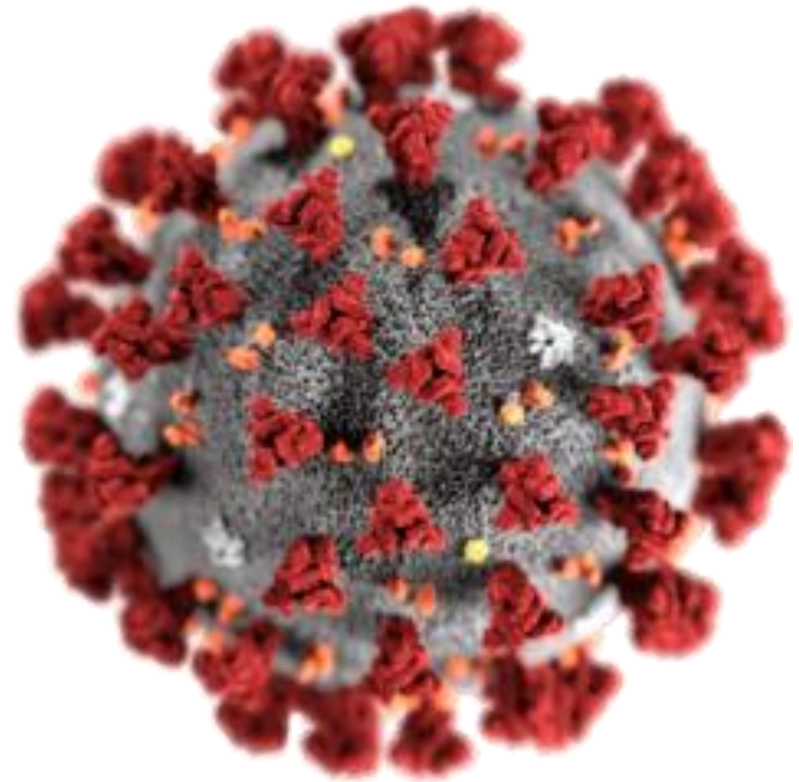**14 pM**

12 pM

SARS-CoV-2 Genome Sequencing

Clinical Diagnostics vs Genome Sequencing

Clinical Testing
Detects presence of 2-3 sites typically by qPCR

SARS Coronavirus

attatctagagtattaggtttgaaaacccttgctactcatggtttagctgctgttaatagattatctagagtattaggtttgaaaaaaaa

Genome Sequencing

Detects and generates sequence data for entire genome

Whole genomes are the highest resolution data
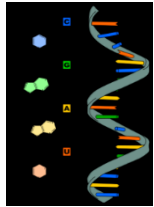
SARS-CoV-2 Sequencing Considerations

▶ **If Ct values are available prioritize values less than 30**

▶ **Use fresh properly stored VTM and/or RNA**

    ▶ **Room temp <24 hrs**

    ▶ **4 degrees <1 week**

    ▶ **-20 degrees upto 1 year**

▶ **Avoid repeated freeze-thaw of VTM and/or isolated RNA**

Next-Generation Sequencing of SARS-CoV-2 Genomes
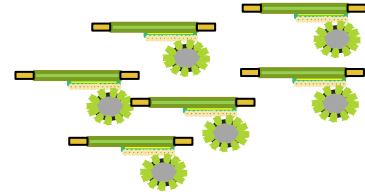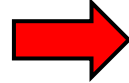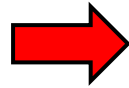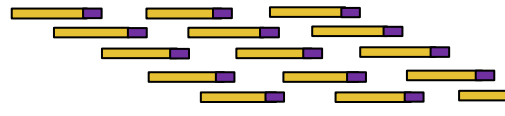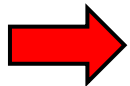
Illumina Provided SARS-CoV-2 Protocols

- **COVIDSeq**
  - **PCR & Tagmentation**
  - **Research Use Only (RUO) & USA FDA Emergency Use Authorization**

- **AmpliSEQ**
  - **PCR & Restriction Digestion & Adapter Ligation**

- **Respiratory Virus Oligo Panel v2 (RVOP)**
  - **Target hybridization capture**

Other Manufacturer Protocols for Illumina

- **Qiagen**
  - **QIAseq DIRECT SARS-CoV-2 Kit (PCR)**

- **IDT/SWIFT**
  - **xGen Amplification Panel (PCR) & xGen Hybridization Probes (Target Capture)**

- **Roche/Kapa**
  - **KAPA SARS-CoV-2 Target Enrichment Panel by Roche (Target Capture)**

PCR Primer Sets

- **ARTIC**
  - **V3- 400bp amplicons (98 primers pairs)**
  - **V4- Updated for optimal coverage of Delta- Illumina COVIDSeq**
  - **V4.1- Includes primers optimized for Omicron**
  - **Midnight- 1,200bp**
    - **Omicron dropout of amplicon 28, possibly amplicon 2**

- **Manufacturer Specific**
  - **Qiagen, IDT**

Illumina COVIDSeq

8.5 μl RNA

RNA

Reverse Transcribe Random Hexamers

25°C for 5 minutes
50°C for 10 minutes
80°C for 5 minutes

Single Stranded cDNA

PCR

98°C for 3 minutes
35 cycles of:
98°C for 15 seconds
63°C for 5 minutes
Hold at 4°C

Primer Pool A

Primer Pool B

Illumina COVIDSeq

Primer Pool A

Primer Pool B

10μl PCR DNA

Mix

10μl PCR DNA

Tagmentation

55°C for 5 minutes
Hold at 10°C

Bead Cleanup

Tagged SARS-CoV-2 Amplicons

PCR Index Primers (IDT UDI)

72°C for 3 minutes
98°C for 3 minutes 7 cycles of:
98°C for 20 seconds
60°C for 30 seconds
72°C for 1 minute
72°C for 3 minutes
Hold at 10°C

Pool & Cleanup Libraries

Amplicon Dropouts

Amplicon Dropouts

- Due to mismatches in primer sequence and genome sequence of virus

- More problematic in low viral load samples

- If amplicon contains lineage defining mutations, may be incorrectly classified
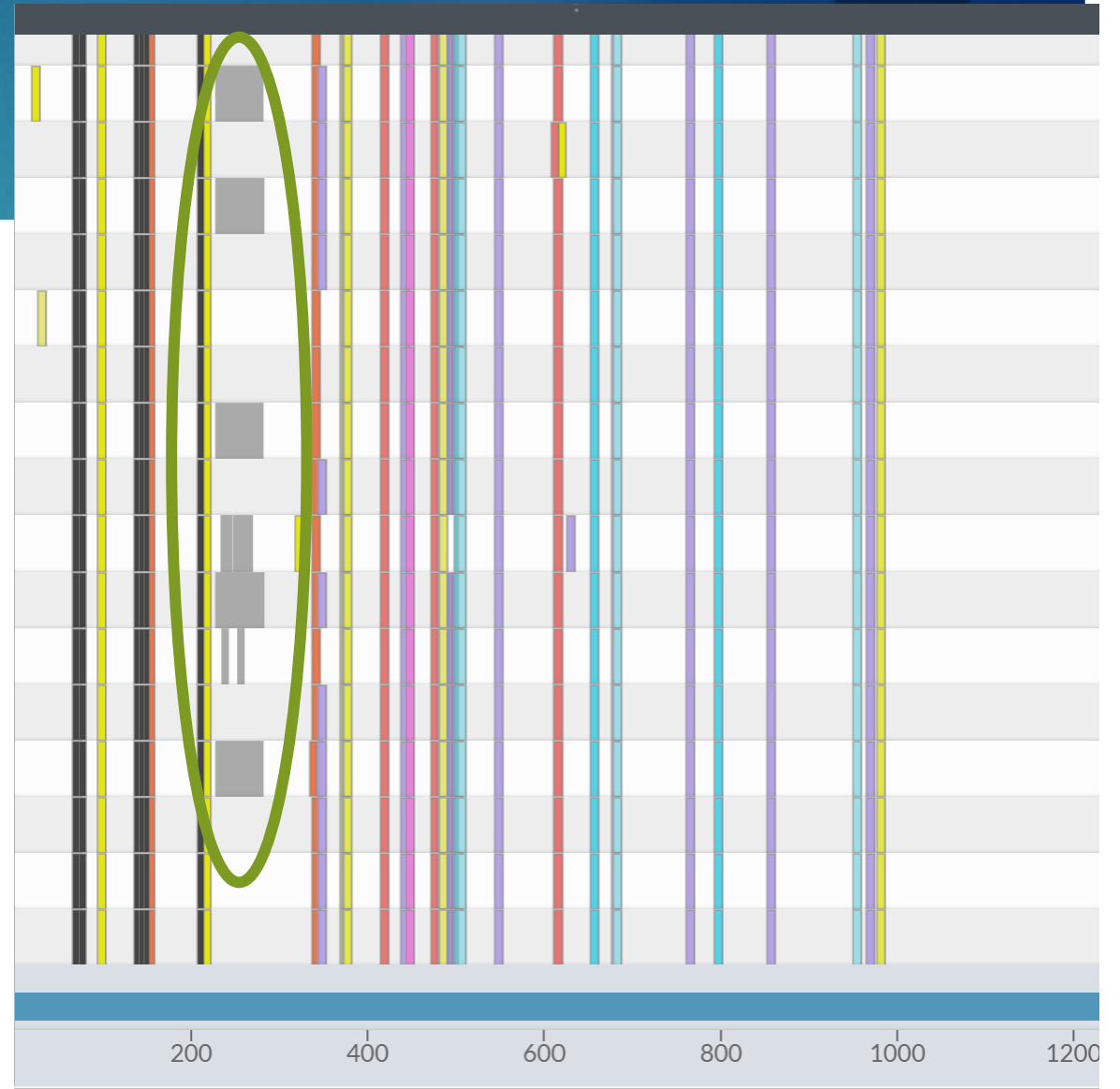
Illumina COVIDSeq Considerations

- Illumina recommends ~1 million reads/samples
  - 48 samples/MiSeq
  - 384 samples/NextSeq 1000/2000

- Primer sets can be changed
  - Ships with ARTIC v4
  - ARTIC v4.1 for better Omicron coverage
  - Midnight?, may depend on read length used for sequencing

- Automated solutions available

Illumina AmpliSeq

▶ **2-pool design, containing 247 amplicons/primer pairs**

    ▶ **(242 unique amplicons (Pool 1: 125 amplicon, Pool 2: 122 amplicons):**

    ▶ **237 viral specific SARS-CoV-2 targets and 5 human gene expression controls) ranging from 125-275 bps in length that covers >99% of the viral genome and all potential serotypes of the virus.**

▶ **Higher Cost/Sample than COVIDSeq**

ThermoFisher
SCIENTIFIC   **illumina®**

Introducing
AmpliSeq™ Solutions for Illumina®

For Research Use Only. Not for use in diagnostic procedures.

Illumina AmpliSeq

1 Quantify and Dilute RNA
Hands-on: 10 minutes
Total: 10 minutes
Reagents: Nuclease-Free Water

2 Reverse Transcribe RNA
Hands-on: 10 minutes
Total: 50 minutes
Reagents: 5X AmpliSeq cDNA Reaction Mix, 10X AmpliSeq RT Enzyme Mix, Nuclease-Free Water

Safe Stopping Point

3 Amplify Targets
Hands-on: 15 minutes
Total: 1.5–4 hours
Reagents: AmpliSeq Custom RNA Panel, 5X AmpliSeq HiFi Mix, Nuclease-Free Water

Safe Stopping Point

4 Partially Digest Amplicons
Hands-on: 10 minutes
Total: 50 minutes
Reagents: FuPa Reagent

Safe Stopping Point

5 Ligate Indexes
Hands-on: 15 minutes
Total: 55 minutes
Reagents: DNA Ligase, AmpliSeq CD Indexes, Switch Solution

Safe Stopping Point

6 Clean Up Library
Hands-on: 15 minutes
Total: 25 minutes
Reagents: 70% EtOH, AMPure XP Beads

Illumina AmpliSeq- Cont.

Standard Workflow

7 **Amplify Library**
Hands-on: 10 minutes
Total: 45 minutes
Reagents: 1X Library Amp Mix, 10X Library Amp Primers

Safe Stopping Point

8 **Perform Second Cleanup**
Hands-on: 15 minutes
Total: 35 minutes
Reagents: 70% EtOH, AMPure XP Beads, Low TE

Safe Stopping Point

9 **Check Libraries**
Total: 1–1.5 hours

10 **Dilute to Starting Concentration**
Hands-on: 20 minutes
Total: 20 minutes
Reagents: Low TE

Equalizer Workflow

7 **Amplify Library**
Hands-on: 10 minutes
Total: 45 minutes
Reagents: 1X Library Amp Mix, 10X Library Amp Primers

8 **Wash Equalizer Beads**
Hands-on: 5 minutes
Total: 5 minutes
Reagents: Equalizer Wash Buffer, Equalizer Beads

9 **Perform Capture and Cleanup**
Hands-on: 10 minutes
Total: 10 minutes
Reagents: Equalizer Capture, Equalizer Wash Buffer, Equalizer Beads

10 **Elute Library**
Hands-on: 15 minutes
Total: 15 minutes
Reagents: Equalizer Elution Buffer

AmpliSeq Considerations

- **Uses AmpliSeq CD Indexes Set A, B, C, or D plate (96 indexes, 96 samples) or AmpliSeq UD Indexes for Illumina (24 Indexes, 24 Samples),**
  - **Limited capacity to multiplex for higher output sequencers, NextSeq 2000 or NovaSeq**

- **Illumina recommends 2x151bp sequencing**

- **MiSeq v3 (600 cycle kit) sequencing reagents loaded at 7-9pM**

Target Capture for Virus Genome Sequencing

Comprehensive viral enrichment enables sensitive respiratory virus genomic identification and analysis by next generation sequencing

Brigid M. O'Flaherty,[1,2,6] Yan Li,[1,6] Ying Tao,[1] Clinton R. Paden,[1,2] Krista Queen,[1,2] Jing Zhang,[1,3] Darrell L. Dinwiddie,[4] Stephen M. Gross,[5] Gary P. Schroth,[5] and Suxiang Tong[1]
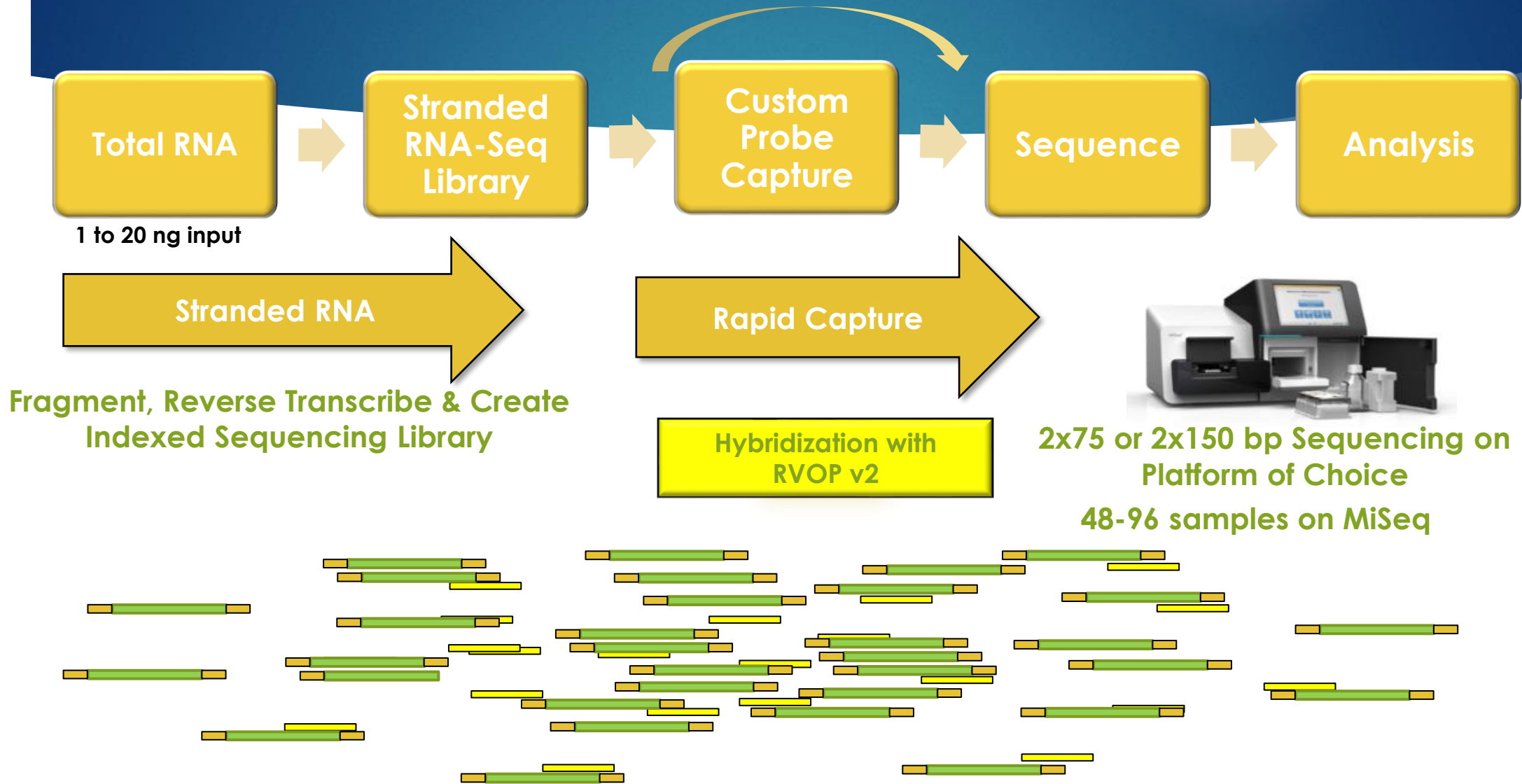
[1] Centers for Disease Control and Prevention, NCIRD, DVD, Atlanta, Georgia 30329, USA; [2] Oak Ridge Institute for Science Education, Oak Ridge, Tennessee 37830, USA; [3] IHRC Incorporated, Atlanta, Georgia 30346, USA; [4] Department of Pediatrics, Clinical Translational Science Center, University of New Mexico, Albuquerque, New Mexico 87131, USA; [5] Illumina, Incorporated, San Diego, California 92122, USA

Illumina RNA Enrichment

Total RNA → **Stranded RNA-Seq Library** → **Custom Probe Capture** → **Sequence** → **Analysis**

1 to 20 ng input

Stranded RNA →

Fragment, Reverse Transcribe & Create Indexed Sequencing Library

Rapid Capture →

Hybridization with RVOP v2

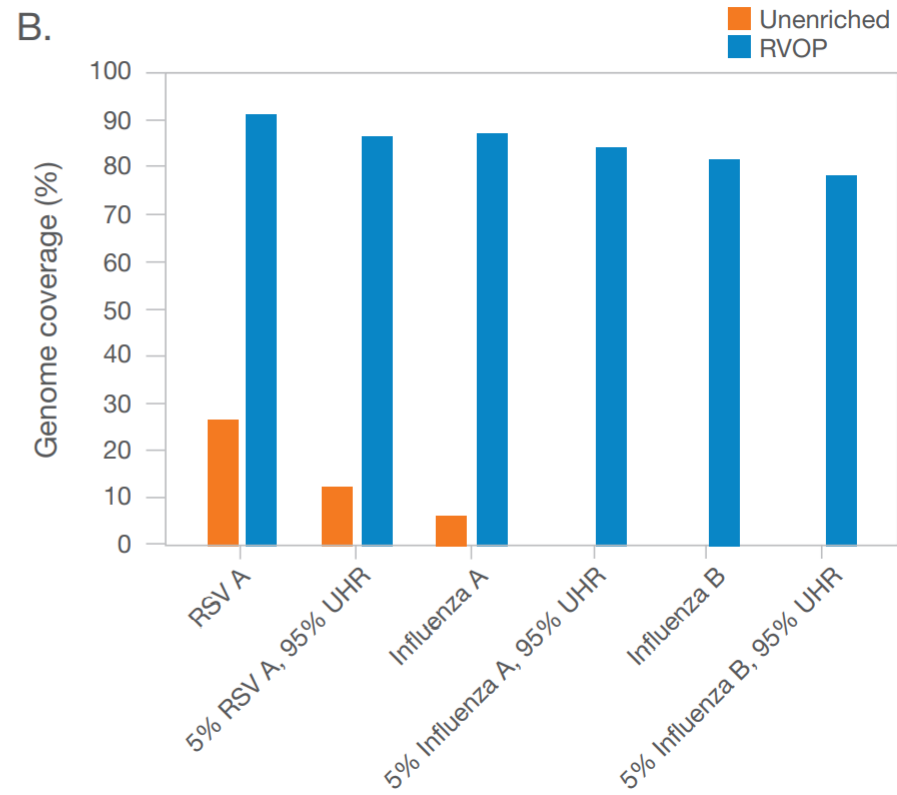2x75 or 2x150 bp Sequincing on Platform of Choice

48-96 samples on MiSeq

Illumina RNA Prep with Enrichment

Table 3: Viruses targeted by the Respiratory Virus Oligos Panel

|			
Human coronavirus 229E	Human parainfluenza virus 1	KI polyomavirus Stockholm 60	Influenza A virus (A/Hong Kong/1073/99(H9N2))
Human coronavirus NL63	Human parainfluenza virus 2	WU Polyomavirus	Influenza A virus (A/Texas/50/2012(H3N2))
Human coronavirus OC43	Human parainfluenza virus 3	Human parechovirus type 1 PicoBank/HPeV1/a	Influenza A virus (A/Michigan/45/2015(H1N1))
Human coronavirus HKU1	Human parainfluenza virus 4a	Human parechovirus 6	Influenza B virus (B/Lee/1940)
SARS-CoV-2	Human metapneumovirus (CAN97-83)	Human rhinovirus A89	Influenza B virus (B/Wisconsin/01/2010)
Human adenovirus B1	Respiratory syncytial virus (type A)	Human rhinovirus C (strain 024)	Influenza B virus (B/Brisbane/60/2008)
Human adenovirus C2	Human Respiratory syncytial virus 9320 (type B)	Human rhinovirus B14	Influenza B virus (B/Colorado/06/2017)
Human adenovirus E4	Influenza A virus (A/Puerto Rico/8/1934(H1N1))	Human enterovirus C104 strain: AK11	Influenza B virus (B/Washington/02/2019)
Human bocavirus 1 (Primate bocaparvovirus 1 isolate st2)	Influenza A virus (A/Korea/426/1968(H2N2))	Human enterovirus C109 isolate NICA08-4327	Human control genes
Human bocavirus 2c PK isolate PK-5510	Influenza A virus (A/New York/392/2004(H3N2))		
Human bocavirus 3	Influenza A virus (A/goose/Guangdong/1/1996(H5N1))		
Human bocavirus 4 NI strain HBoV4-NI-385	Influenza A virus (A/Zhejiang/DTID-ZJU01/2013(H7N9))		

Virus Enrichment
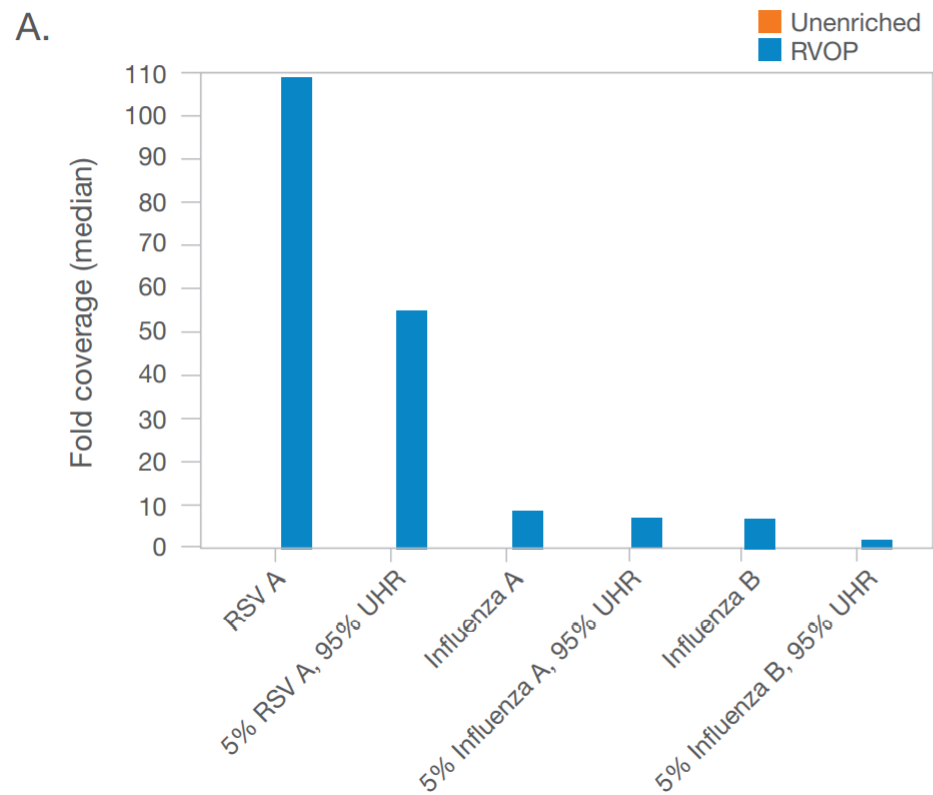
SARS-CoV-2 Enrichment
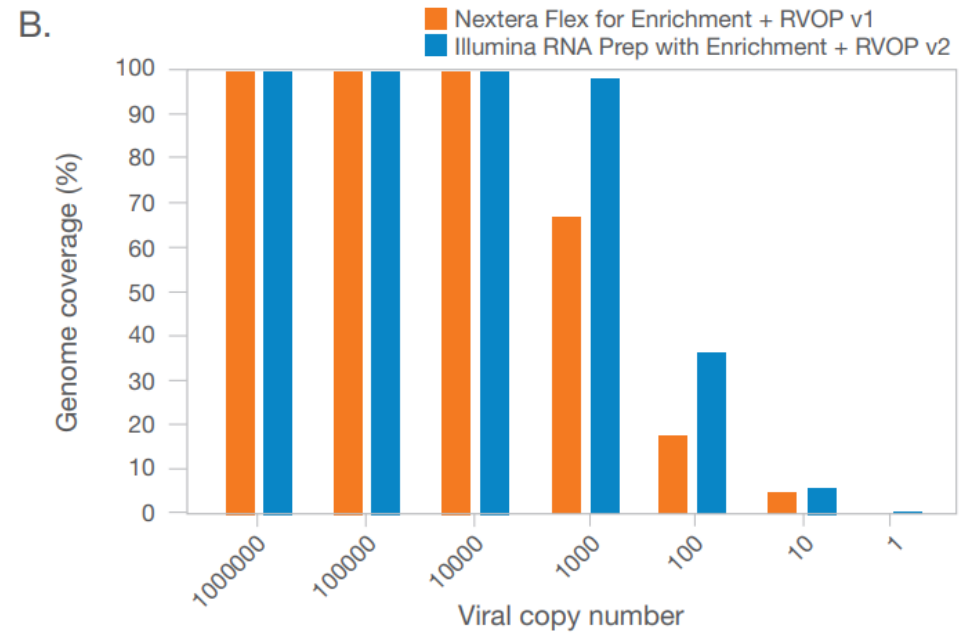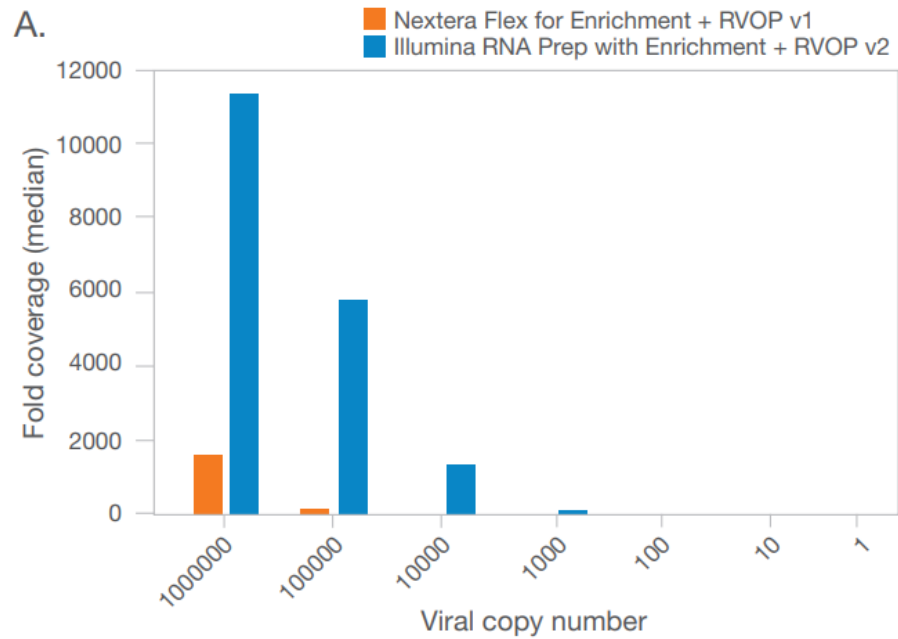
SARS-CoV-2 Enrichment

C. SARS-CoV-2 10^5 copies, Nextera Flex for Enrichment + RVOP v1
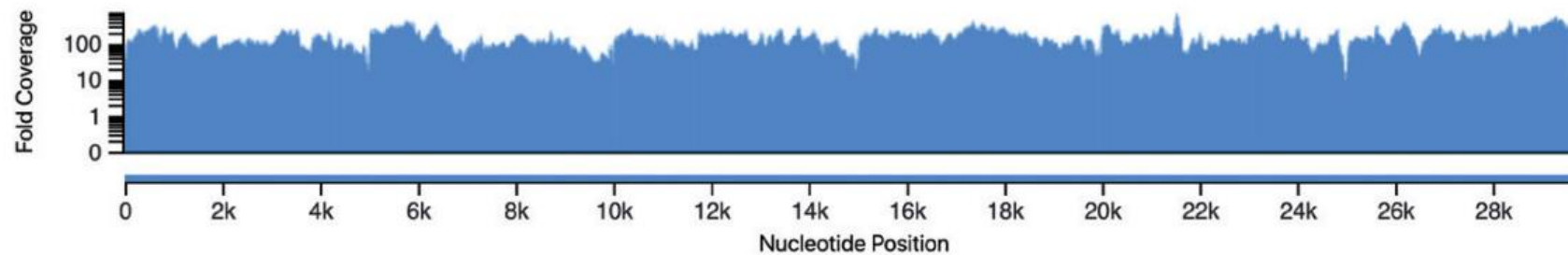
SARS-CoV-2 10^5 copies, Illumina RNA Prep with Enrichment + RVOP v2
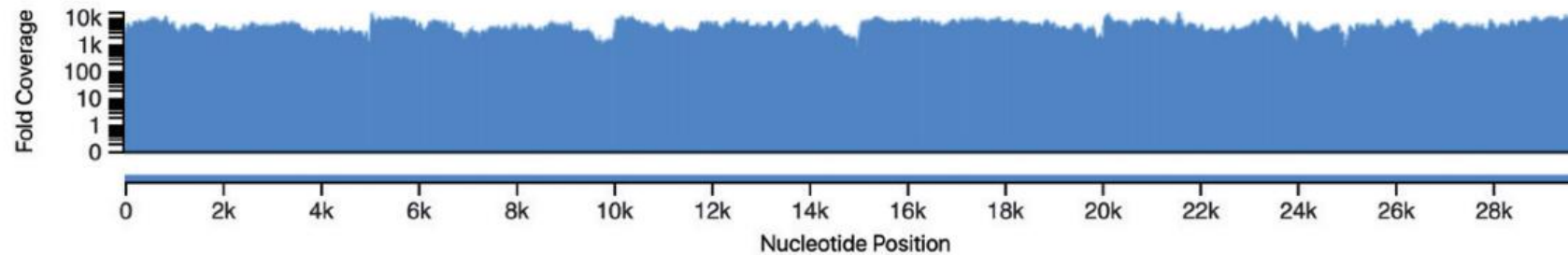
Considerations for Enrichment

▶ **Can be used for SARS-CoV-2 and additional respiratory viruses**

▶ **Less susceptible to virus mutations**

▶ **Higher costs & lower throughput**

Questions?