SARS-CoV-2 NGS Training

NGS Academy for the Africa Pathogen Genomics Initiative

Module 4 - Illumina workflow

SARS-CoV-2 NGS Training

Module 4: Illumina workflow

- Session 1: Concepts, library prep, starting a sequencing run
- Session 2: Library prep QC and sequencing run QC
- Session 3: IDSeq/Data Processing

CHAN ZUCKERBERG BIOHUB

Cristina Tato,
Director
Rapid Response (RR) group

Manu Vanaerschot,
Scientist
RR group

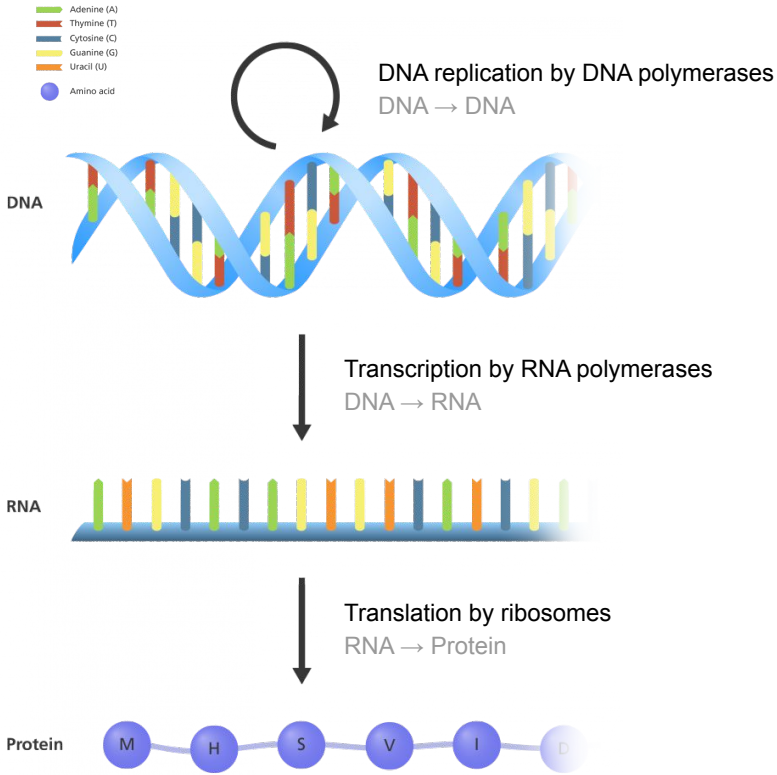
Vida Ahyong,
Scientist
RR group

CHAN ZUCKERBERG INITIATIVE

Katrina Kalantar,
Computational Biologist

Liz Fahsbender,
IDseq Application Scientist

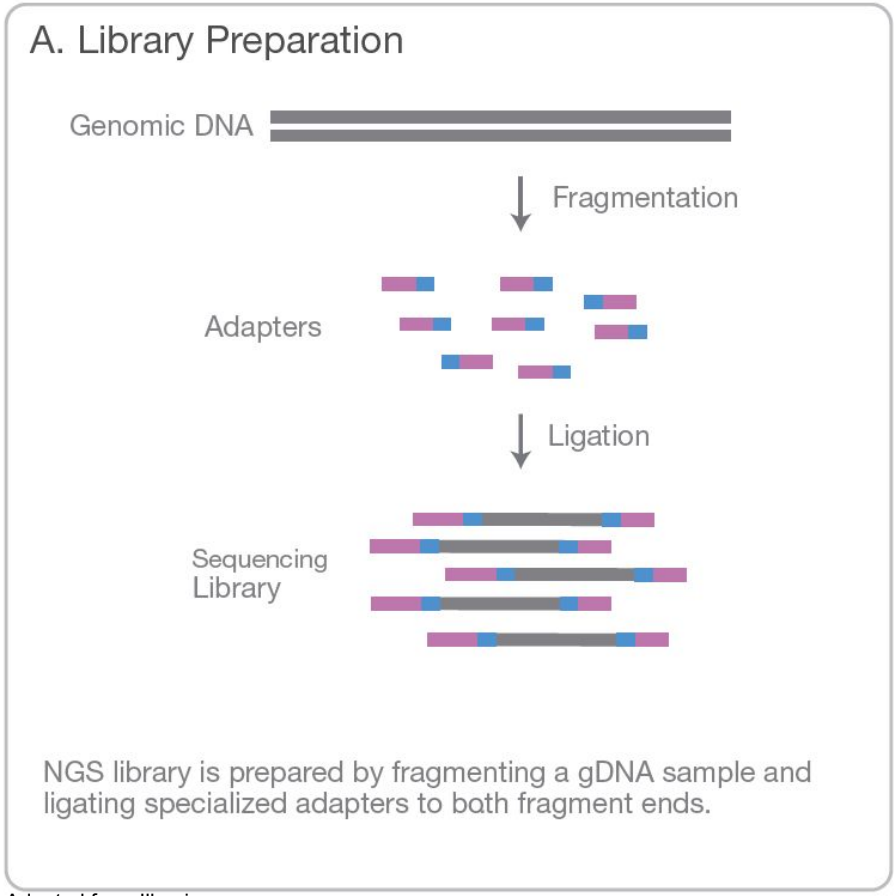
NGS determines the sequence of DNA or RNA in samples

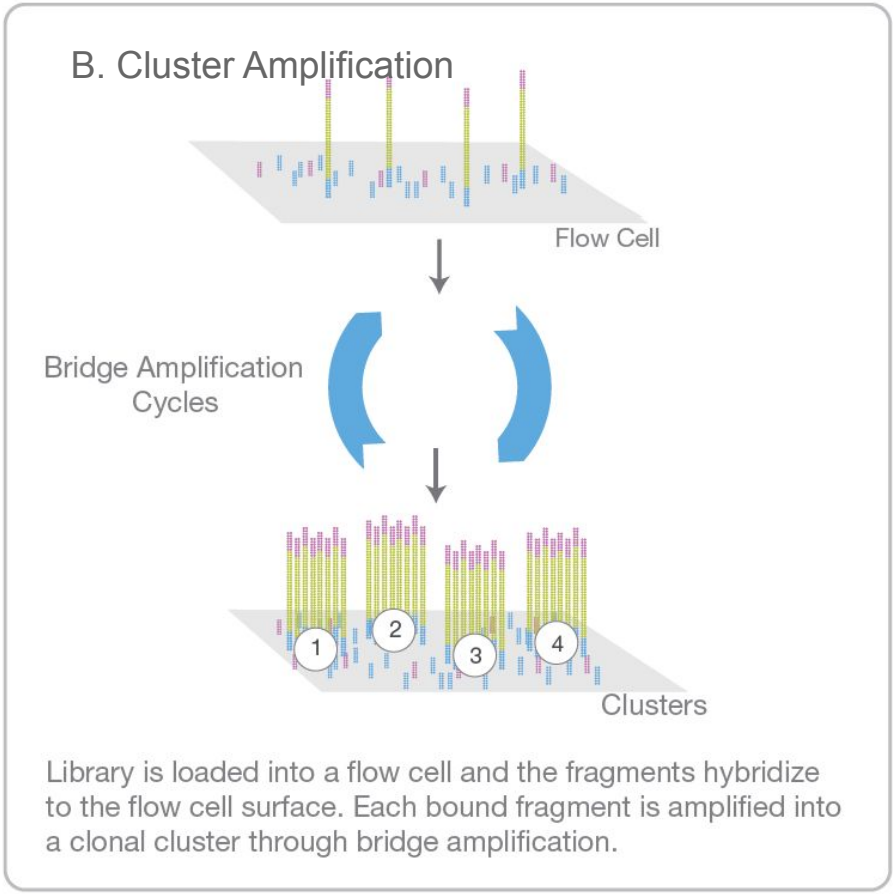


Adapted from yourgenome.org

	Viruses	Bacteria	Eukaryotic microbes (yeast, amoebas, worms, etc)
Genome composition	ssRNA, ssDNA, dsDNA, or dsRNA	dsDNA	dsDNA
Ribosomal RNA	None, uses host ribosomes	16S rRNA	18S rRNA
Genome size	2 Kb - 1 Mb	1 Mb- 5 Mb	200 Kb - 100,000 Mb

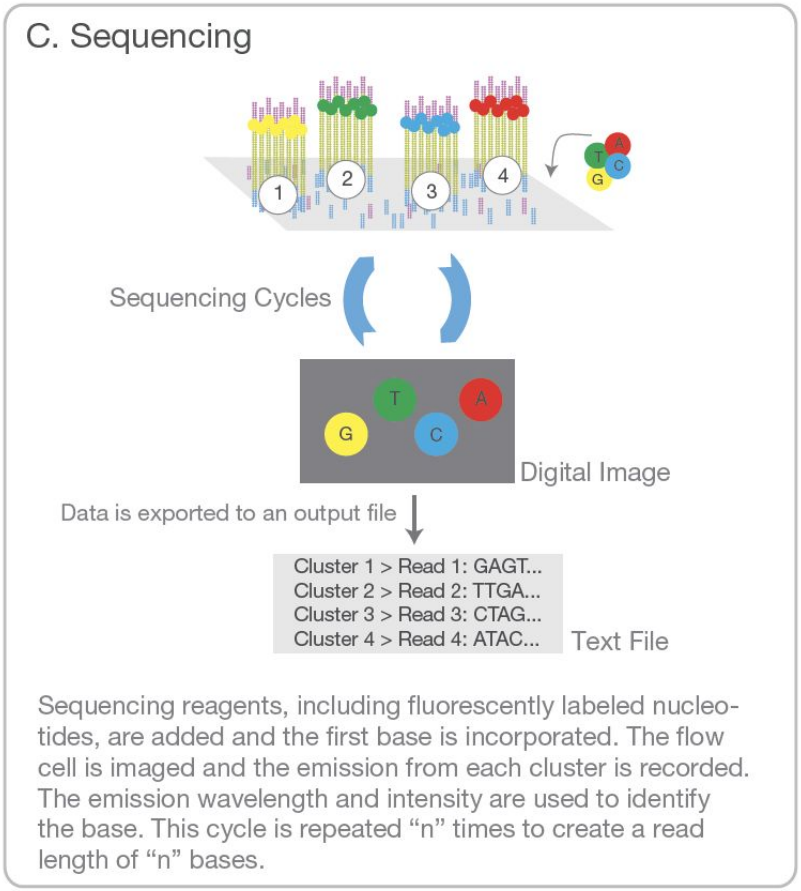
Illumina sequencing

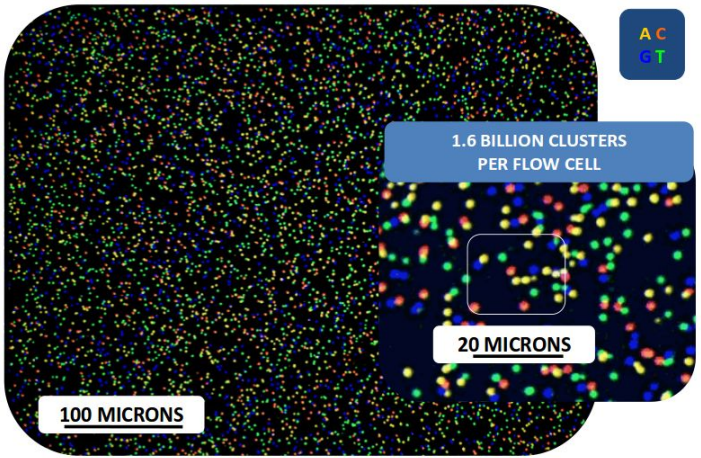




Adapted from Illumina.com

Illumina sequencing





Adapted from Illumina.com

NGS terminology

What is…..

… a library ?

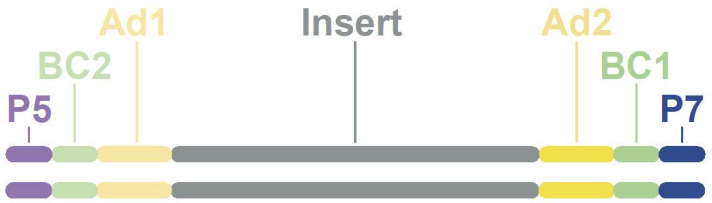
… (de-)multiplexing ?

… a read ?

… coverage depth ?

… paired-end sequencing ?

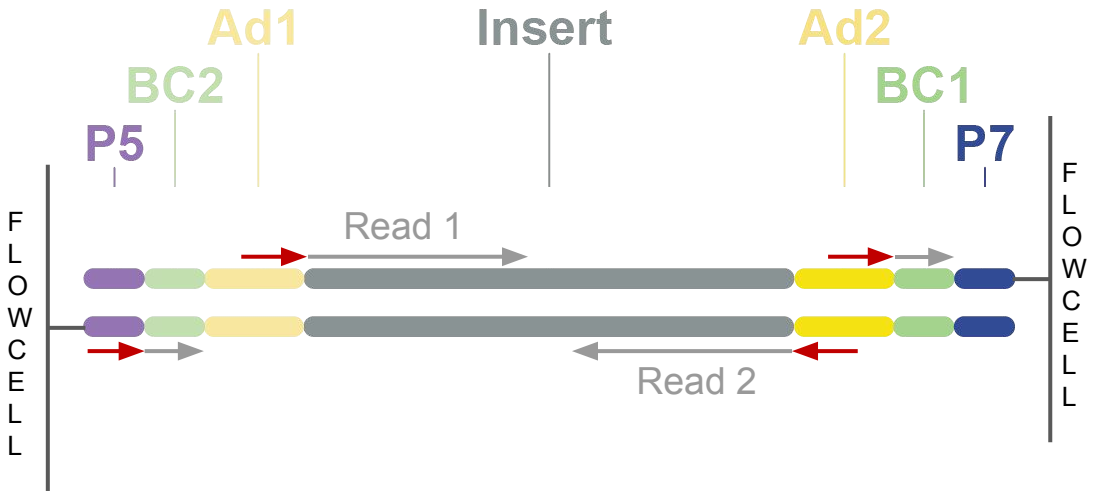
Sequencing library = collection of fragments ready for sequencing



Insert	= DNA fragment to be sequenced
Ad1/Ad2	= Adapters that allow initiation of sequencing
BC1/BC2	= Barcodes specific for each sample (one for each strand)
P5/P7	= Flow cell binding sites

Sequencing library = collection of fragments ready for sequencing

- Sequencing process:



Multiplexing = sequencing multiple samples together in one reaction

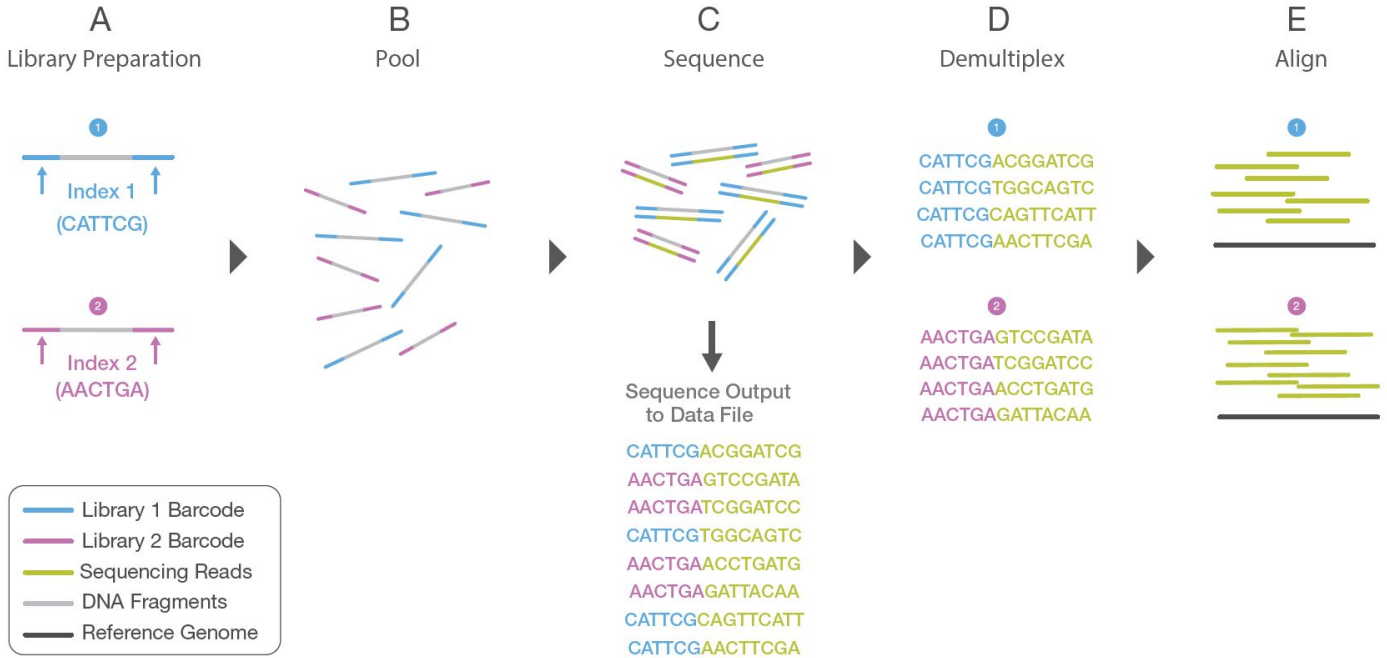


Figure 5: Library Multiplexing Overview.

a. Two distinct libraries are attached to unique index sequences. Index sequences are attached during library preparation.

b. Libraries are pooled together and loaded into the same flow cell lane.

c. Libraries are sequenced together during a single instrument run. All sequences are exported to a single output file.

d. A demultiplexing algorithm sorts the reads into different files according to their indexes.

e. Each set of reads is aligned to the appropriate reference sequence.

Adapted from Illumina.com

Illumina paired-end sequencing

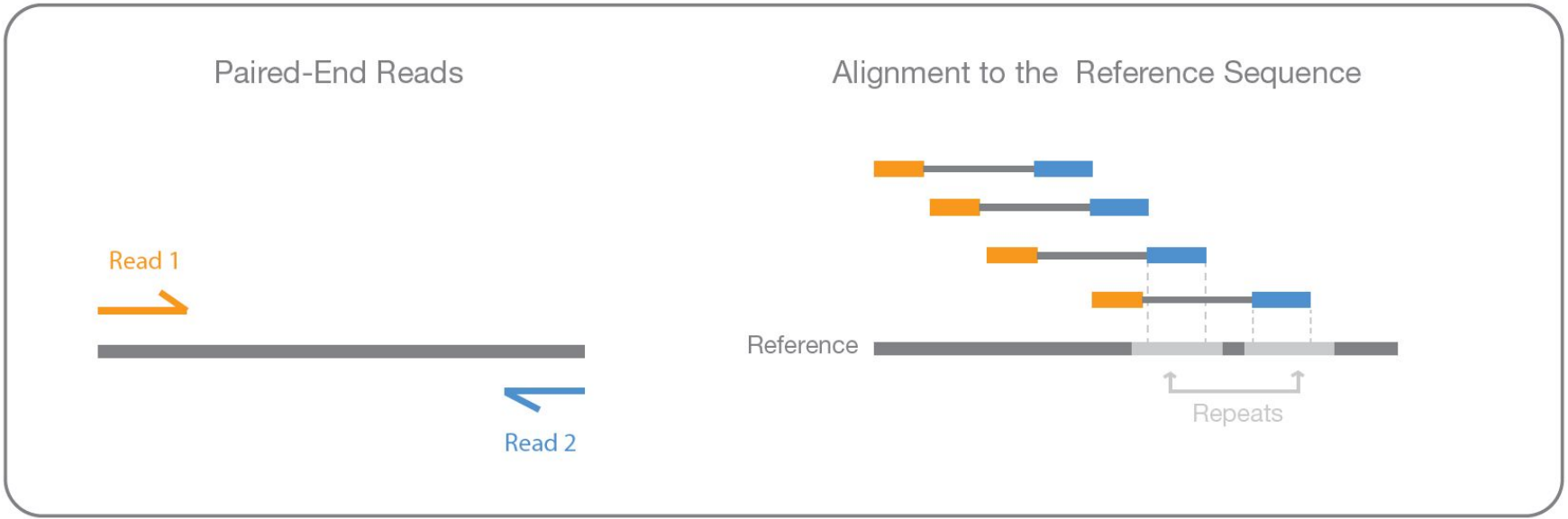
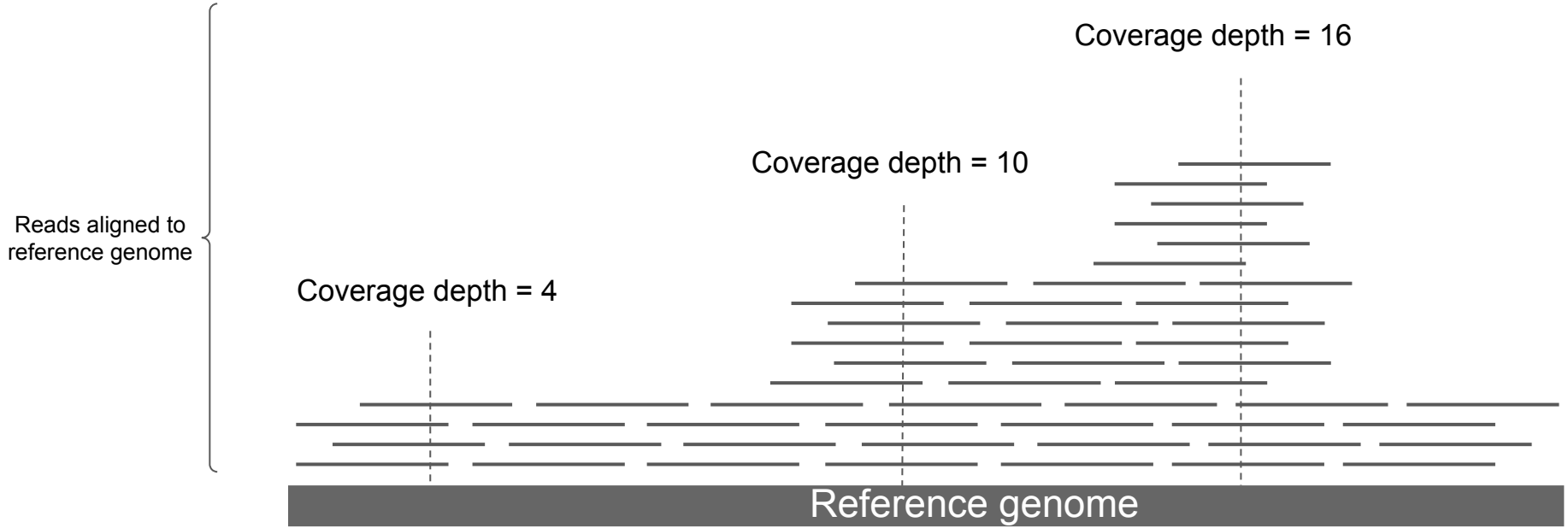


Figure 4: Paired-End Sequencing and Alignment—Paired-end sequencing enables both ends of the DNA fragment to be sequenced. Because the distance between each paired read is known, alignment algorithms can use this information to map the reads over repetitive regions more precisely. This results in much better alignment of the reads, especially across difficult-to-sequence, repetitive regions of the genome.

Adapted from Illumina.com

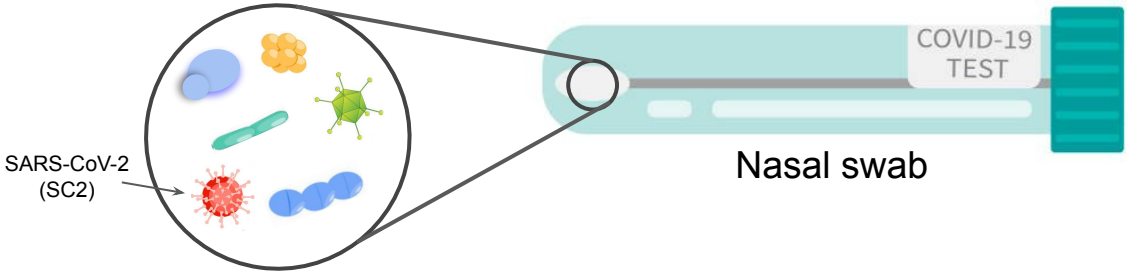
Coverage or sequence depth

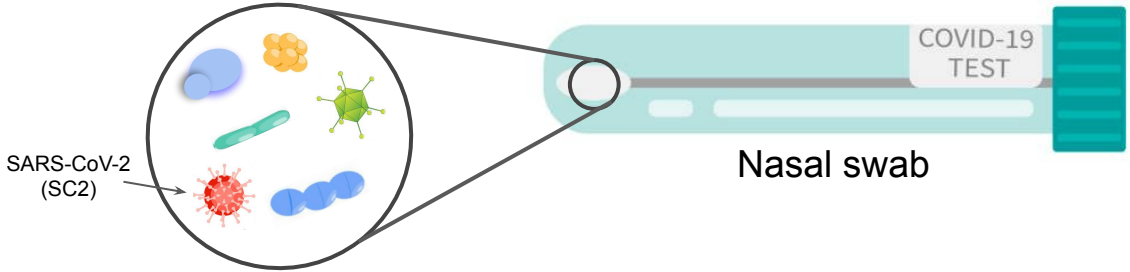
Coverage = How many times each base has been sequenced on average throughout the genome



SARS-CoV-2

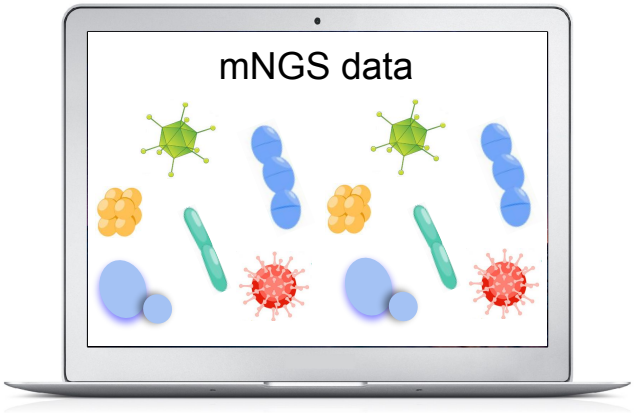
Sequencing options





Option 1: Metagenomics (mNGS)

Sequence EVERYTHING
using random primers



- Detect all (unknown) pathogens & co-infections
- Hard to get full SC2 genomes since sequencing space is shared with everything else present in the sample

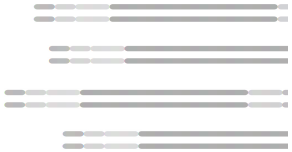
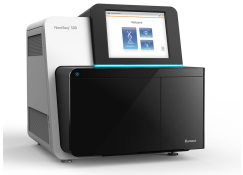
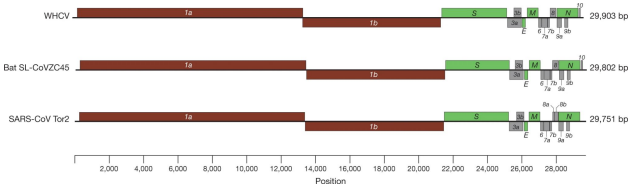
Article | Open Access | Published: 03 February 2020

A new coronavirus associated with human respiratory disease in China

Fan Wu, Su Zhao, Bin Yu, Yan-Mei Chen, Wen Wang, Zhi-Gang Song, Yi Hu, Zhao-Wu Tao, Jun-Hua Tian, Yuan-Yuan Pei, Ming-Li Yuan, Yu-Ling Zhang, Fa-Hui Dai, Yi Liu, Qi-Min Wang, Jiao-Jiao Zheng, Lin Xu, Edward C. Holmes & Yong-Zhen Zhang

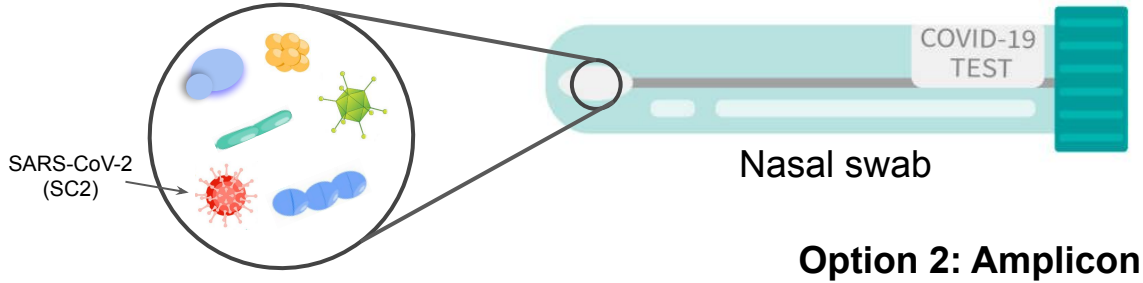
Nature **579**, 265–269 (2020) | Cite this article

mNGS identified SARS-CoV-2 and provided the first genome

Sampling	xNA extraction	Library prep	Sequencing	Data analysis
				
Bronchoalveolar lavage fluid from symptomatic patient	Zymo Viral Magbead kit	Total RNA-seq kit (TaKaRa)	Illumina Miniseq Paired-end 150bp reads 56.6M reads generated	De novo assembly of reads: longest and most abundant contig aligned to bat SARS-like coronavirus

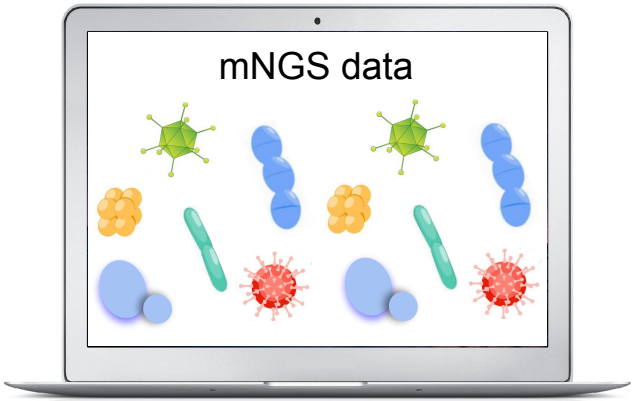
Primer design based on contig for reference genome verification

Final SC2 reference genome MN908947



Option 1: Metagenomics (mNGS)

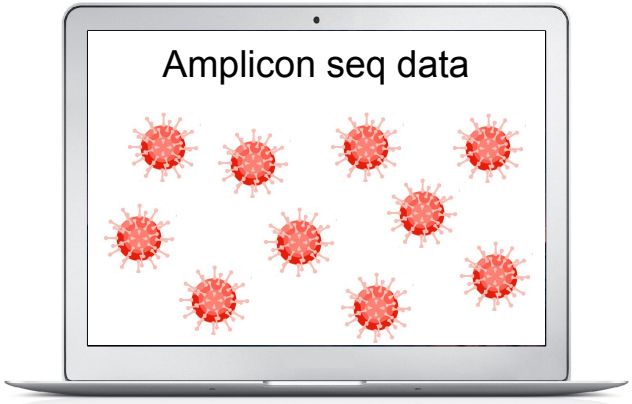
Sequence EVERYTHING
using random primers



- Detect all (unknown) pathogens & co-infections
- Hard to get full SC2 genomes since sequencing space is shared with everything else present in the sample

Option 2: Amplicon sequencing

Sequence ONLY SC2 using
multiplex PCR with SC2-specific primers



- Reference genome of interest required for primer design
- Co-infections NOT detected
- Best option to obtain full SC2 genomes -> great for genomic epidemiology

ACCEPTED MANUSCRIPT

Estimation of secondary household attack rates for emergent spike L452R SARS-CoV-2 variants detected by genomic surveillance at a community-based testing site in San Francisco

James Peng, Jamin Liu, Sabrina A Mann, Anthea M Mitchell, Matthew T Laurie, Sara Sunshine, Genay Pilarowski, Patrick Ayscue, Amy Kistler, Manu Vanaerschot, Lucy M Li, Aaron McGeever, Eric D Chow, Carina Marquez, Robert Nakamura, Luis Rubio, Gabriel Chamie, Diane Jones, Jon Jacobo, Susana Rojas, Susy Rojas, Valerie Tulier-Laiwa, Douglas Black, Jackie Martinez, Jamie Naso, Joshua Schwab, Maya Petersen, Diane Havlir, Joseph DeRisi ✉, IDseq Team

Clinical Infectious Diseases, ciab283, https://doi.org/10.1093/cid/ciab283

Amplicon sequencing allows to obtain full genomes of many samples at once

Sampling	xNA extraction	Library prep	Sequencing	Analysis +	Epidemiological data
					
Anterior nasal swabs collected in DNA/RNA Shield (Nov '20 - Jan'21)	Qiagen RNAeasy Plus mini kit	ARTIC V3 PCR, SC2-specific primers (=amplicons!) + Nextera DNA library reagents	Illumina NovaSeq Paired-end 150bp reads	Consensus genome generation using IDSEQ n>900	Household infection rate

ACCEPTED MANUSCRIPT

Estimation of secondary household attack rates for emergent spike L452R SARS-CoV-2 variants detected by genomic surveillance at a community-based testing site in San Francisco

James Peng, Jamin Liu, Sabrina A Mann, Anthea M Mitchell, Matthew T Laurie, Sara Sunshine, Genay Pilarowski, Patrick Ayscue, Amy Kistler, Manu Vanaerschot, Lucy M Li, Aaron McGeever, Eric D Chow, Carina Marquez, Robert Nakamura, Luis Rubio, Gabriel Chamie, Diane Jones, Jon Jacobo, Susana Rojas, Susy Rojas, Valerie Tulier-Laiwa, Douglas Black, Jackie Martinez, Jamie Naso, Joshua Schwab, Maya Petersen, Diane Havlir, Joseph DeRisi ✉, IDseq Team

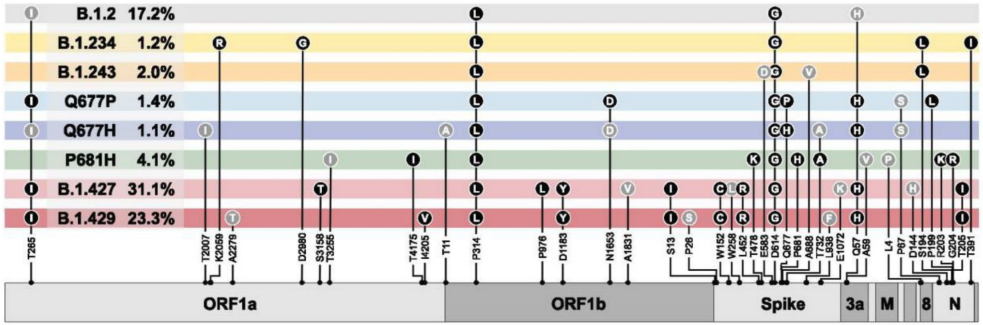
Clinical Infectious Diseases, ciab283, https://doi.org/10.1093/cid/ciab283

Amplicon sequencing allows to obtain full genomes of many samples at once

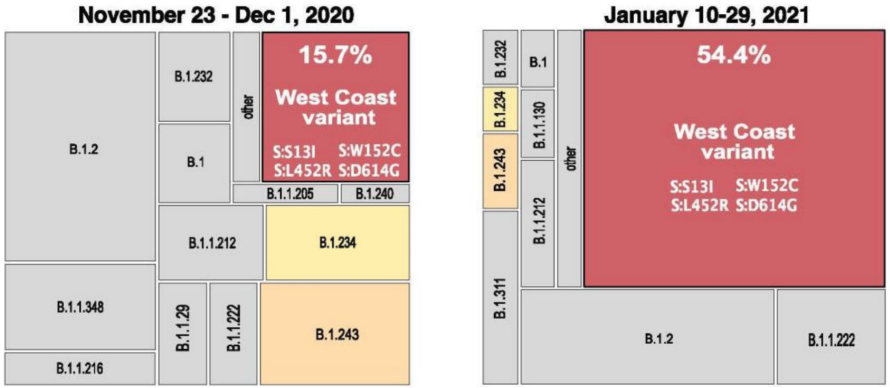
Household contacts exposed to West Coast variant had estimated 28% higher risk of secondary infection

Combine with household infection data

Variants detected in the study (Nov '20 - Jan '21)

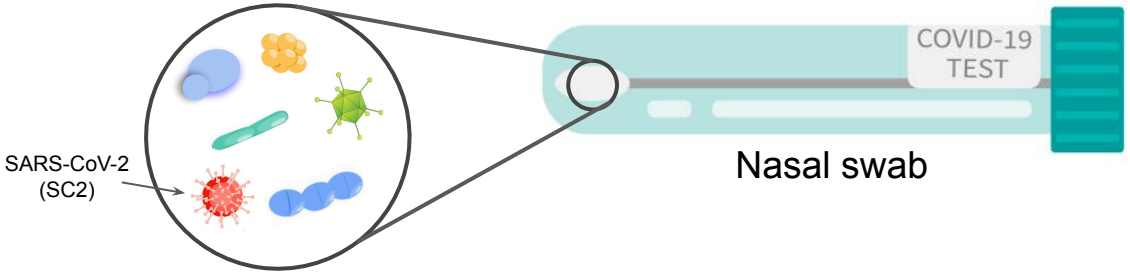


West Coast variant increased in prevalence over time



SARS-CoV-2

Amplicon sequencing



SARS-CoV-2 sequencing concepts

1. Illumina sequencing requires DNA as input material, but SARS-CoV-2 is an RNA virus
    - Reverse transcribe viral RNA to cDNA
2. A clinical sample contains many host cells and many different organisms
    - Amplify SARS-CoV-2 by PCR
3. Library prep requires fragmentation, adapter ligation and barcoding
    - Use a suitable DNA library prep kit
4. QC individual sample libraries
5. Pool samples together for maximum throughput
    - # depends on sequencer and viral content of original sample
6. Final QC of sequencing library & loading the sequencer
7. Run the Sequencer
8. QC of sequencing run

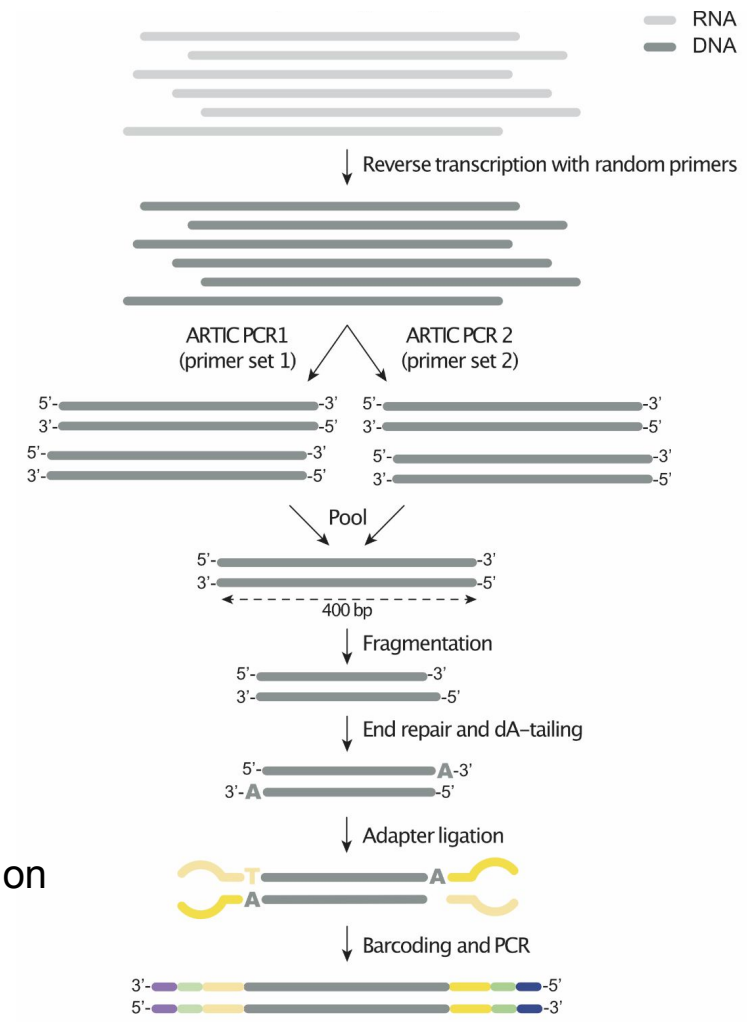
SARS-CoV-2 amplicon sequencing

1. Reverse transcription:
   a. Random hexamer/oligodT priming
   b. Reverse transcription of RNA into cDNA

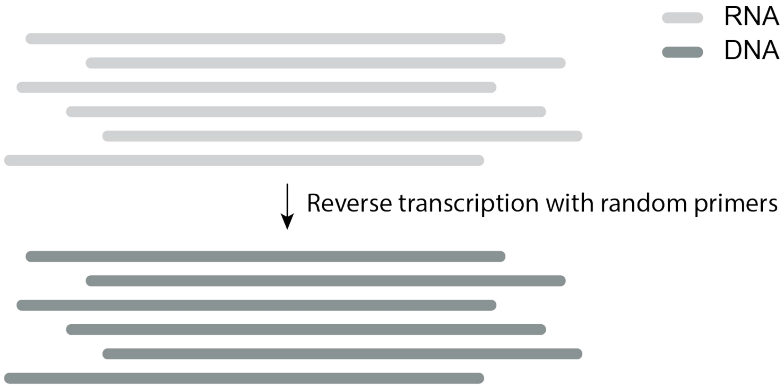
2. SARS-CoV-2 amplification:
   a. ARTIC PCR 1
   b. ARTIC PCR 2

   | pool PCR products

3. Library prep:
   a. Fragmentation, end-repair & dA tailing, adapter ligation
   b. Barcoding and PCR enrichment



Step 1: Reverse transcription



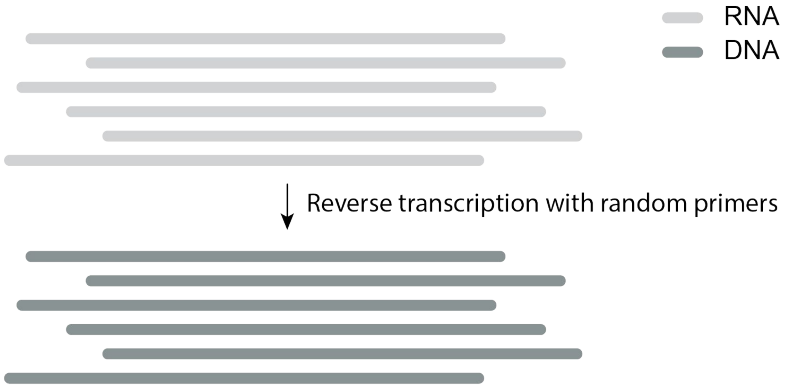
1. Dilute sample if needed:
    a. Ct <15: 1/10 dilution
    b. Ct 15-35: no dilution

LunaScript RT SuperMix (NEB)

2. Priming and reverse transcription (1 single rxn)
    a. Add mastermix (incl. primers) to RNA
    b. Incubate at
        i. 25 °C for 2 minutes: initiate RT reaction with random primers
        ii. 55 °C for 10 minutes: RT reaction
        iii. 95 °C for 1 minutes: inactivate RT enzyme

Keep cDNA on ice until next step

Step 1: Reverse transcription

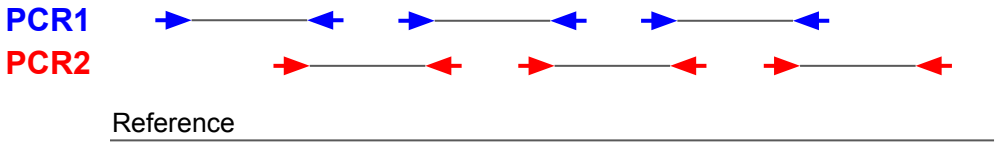


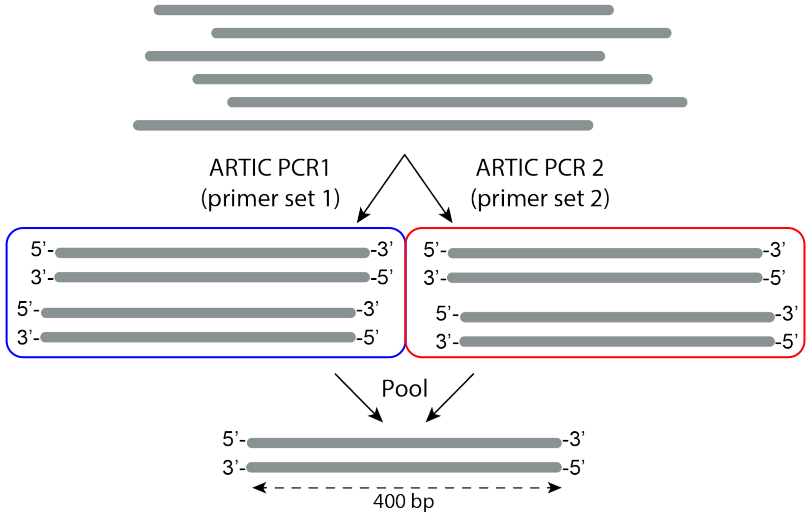
Superscript IV (ThermoFisher)

1. Dilute sample if needed:
    a. Ct <15: 1/10 dilution
    b. Ct 15-35: no dilution
2. Random hexamer priming (1 rxn)
    a. Add primers to RNA
    b. Incubate at 65°C for 5 minutes
3. Reverse transcription (1 rxn)
    a. Add RT enzyme and buffers to primed RNA
    b. Incubate at
        i. 25 °C for ~5 minutes: initiate RT reaction with random primers
        ii. 42 °C for 50 minutes: RT reaction
        iii. 70 °C for 10 minutes: inactivate RT enzyme

Step 2: Multiplex PCR to generate amplicons

1. Perform multiplex PCRs
    a. Standard PCR mastermix
    b. Two sets of primers, resulting in overlapping fragments





ARTIC V3 primer set:

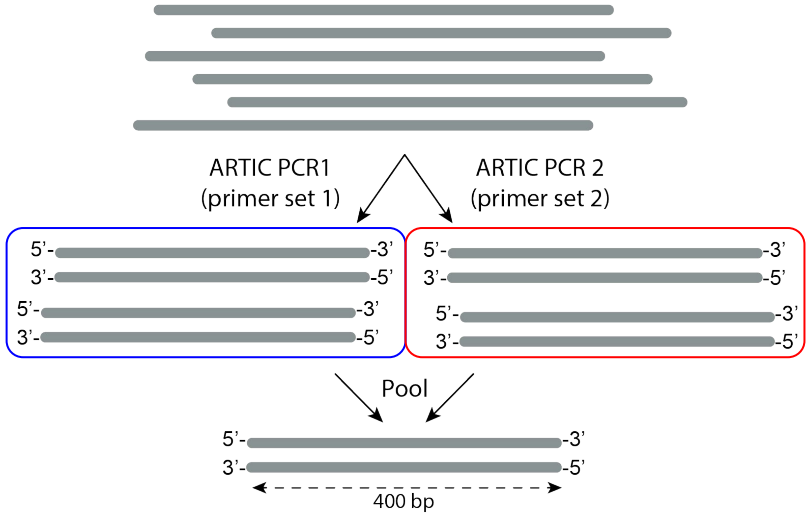
- ~98 primer pairs distributed over 2 primer sets
- Generates ~400bp amplicons
- Designed using primalscheme.com (Quick *et al.*)

Other available primer sets:

- ARTIC V4 (June 2021): to account for primer erosion due to emerging variants
- 'Midnight' set: ~1.2 kb amplicons
- AmpliSeq research panel: ~125-275 bp amplicons
- …

Step 2: Multiplex PCR to generate amplicons

1. Perform multiplex PCRs
    a. Standard PCR mastermix
    b. Two sets of primers, resulting in overlapping fragments
2. Pool PCR products per sample
3. Clean up or dilute pooled PCR products



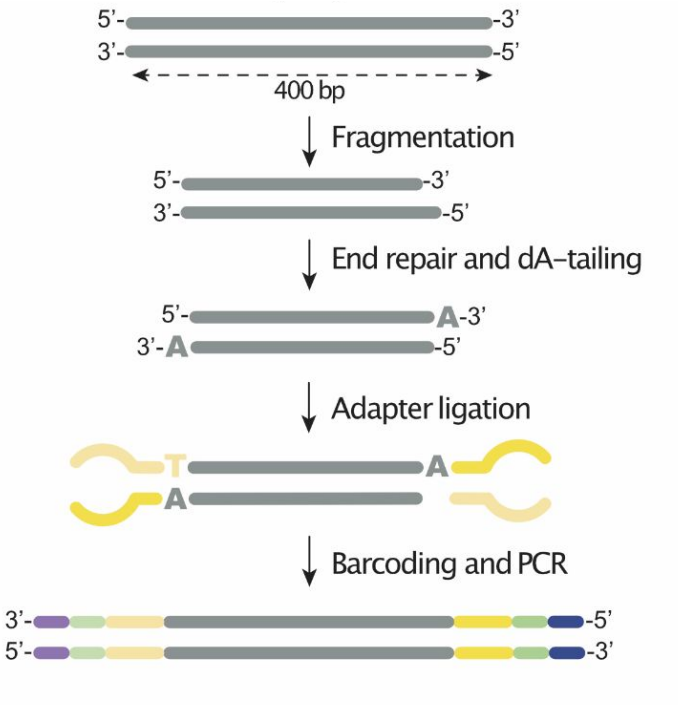
AVOID CROSS-CONTAMINATION AT ALL COST!!

- Amplicons contaminate pipets, surfaces, …
- Cross-contamination may lead to false positive/negative SNPs
- Pipet slowly, avoid bubbles
- Use water controls to monitor for cross-contamination

Step 3: Library prep

Ligation protocol

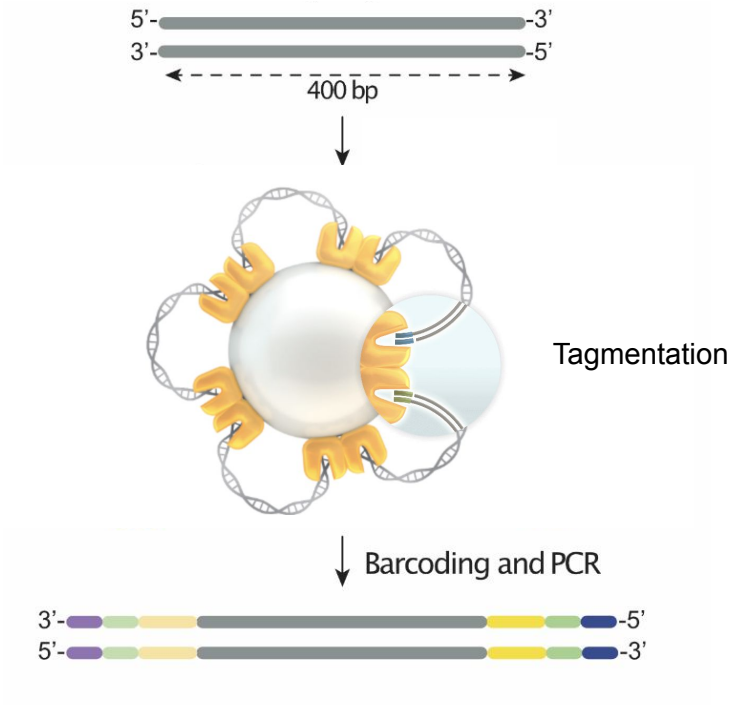
Illumina TruSeq DNA / NEBNext Ultra II DNA



OR

Tagmentation protocol

Illumina DNA prep (= Nextera Flex)



Step 3: Library prep

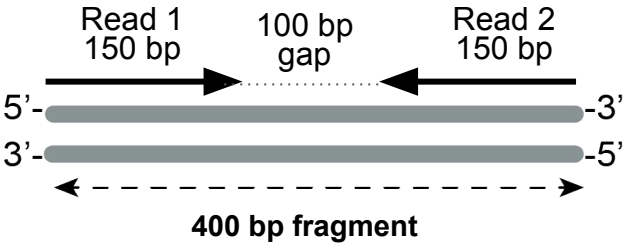
Why is fragmentation required for ARTIC sequencing?

- Sequencing kits specify the max number of cycles that can be performed

- # cycles = total # bp that can be sequenced, e.g.:
  - 150 cycles -> 75 bp per read for paired-end sequencing
  - 300 cycles -> 150 bp per read for paired-end sequencing
  - ...
- Kits available:
  - iSeq: 300 cycles
  - MiniSeq: 75, 100 or 150 cycles
  - MiSeq V3: 150 or 600 cycles
  - NextSeq1000/2000: 50, 100, 200 or 300 cycles
  - NovaSeq: 35 to 500 cycles

- Running more cycles requires a longer runtime!

E.g. 300 cycle kit



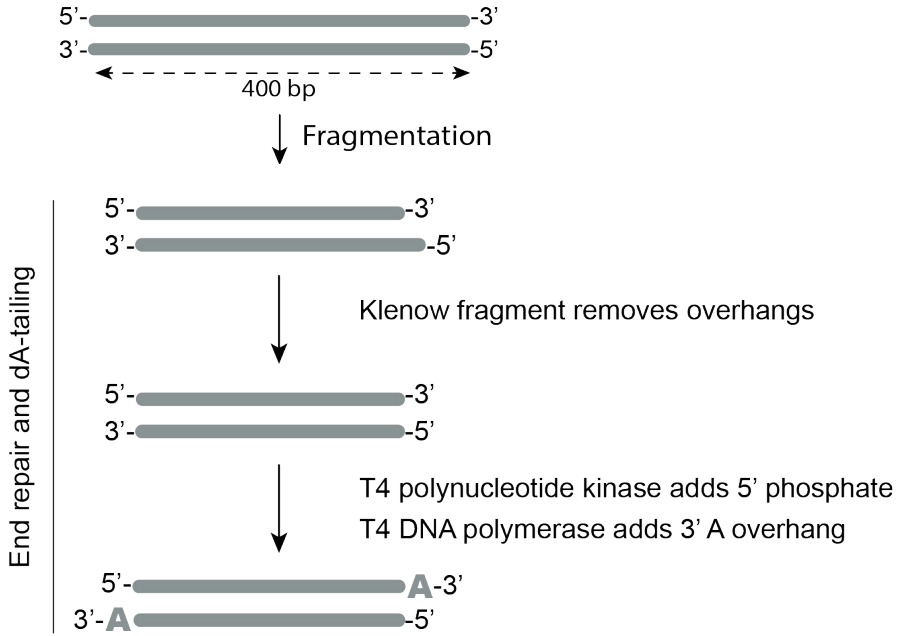
Only MiSeq 600 cycle or NovaSeq 500 cycle kits would be able to fully sequence a 400 bp amplicon without fragmentation

Step 3: Library prep - TruSeq/NEBNext Ultra II DNA protocol

A. Fragmentation and end-repair (1 rxn total)

1. Fragmentation: enzymatic or mechanical
2. Remove overhangs to get blunt ends
3. Prep fragment to clone on adapters
   a. Add 5’ phosphate (T4 kinase)
   b. Add single 3’ A overhang to clone on adapters (T4 DNA polymerase)

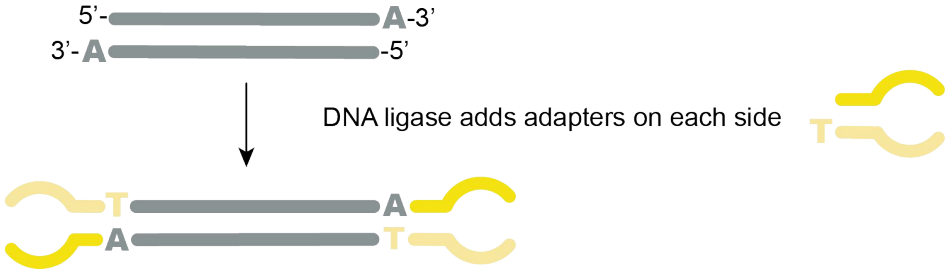
Note: After end-repair, move on to adapter ligation step ASAP (5’ phosphate is fragile)



Step 3: Library prep - TruSeq/NEBNext Ultra II DNA protocol

B. Adapter ligation

1. Adapters are added to allow sequencing (1 rxn)



2. Remove free adapter dimers by bead clean-up (0.9x ratio)

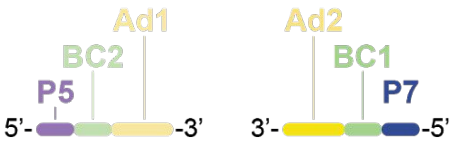
   Adapter dimers are our enemy #1 after this step!!

   (Short fragments get preferentially sequenced)

Step 3: Library prep - TruSeq/NEBNext Ultra II DNA protocol

C. Barcoding/Indexing PCR

1. Add specific i5 and i7 oligos to each sample (includes barcodes)
2. PCR to add barcodes and flow cell binding sequences to fragments
3. Bead clean-up (0.8x)

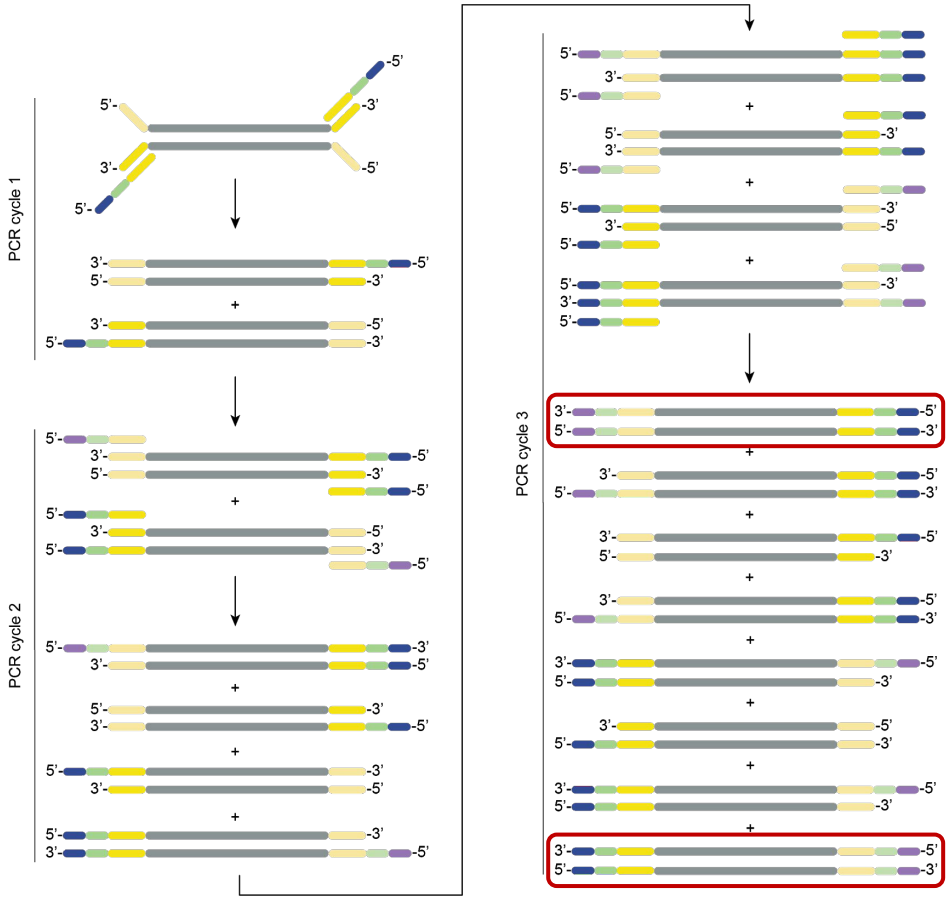


i5 barcoded primer i7 barcoded primer

Ad1/Ad2 = Adapters that allow initiation of sequencing

BC1/BC2 = Barcodes specific for each sample

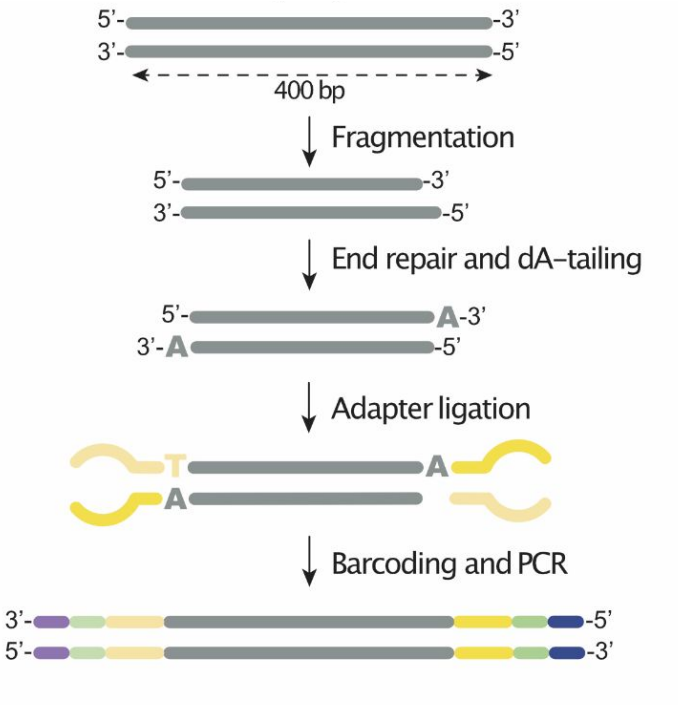
P5/P7 = Flow cell binding sites



Step 3: Library prep

Ligation protocol

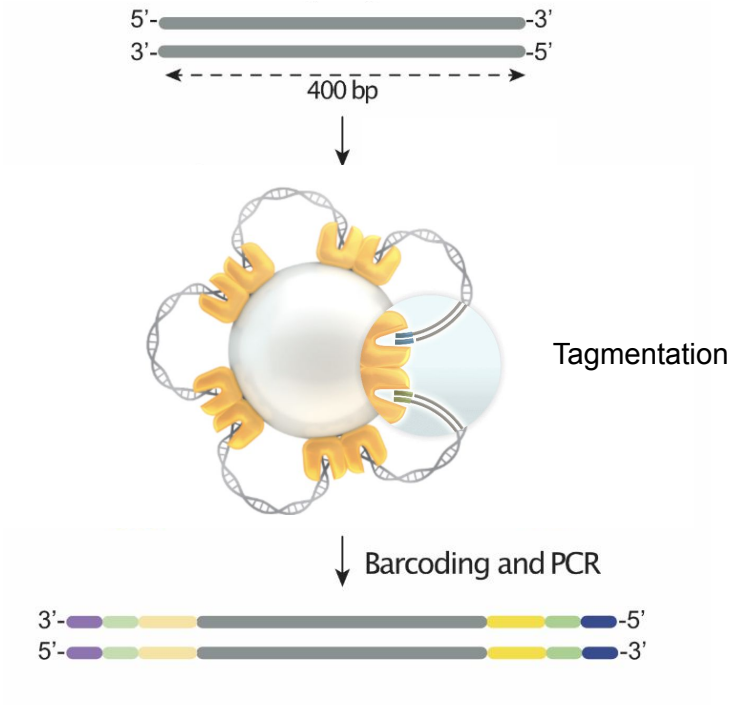
Illumina TruSeq DNA / NEBNext Ultra II DNA



OR

Tagmentation protocol

Illumina DNA prep (= Nextera Flex)



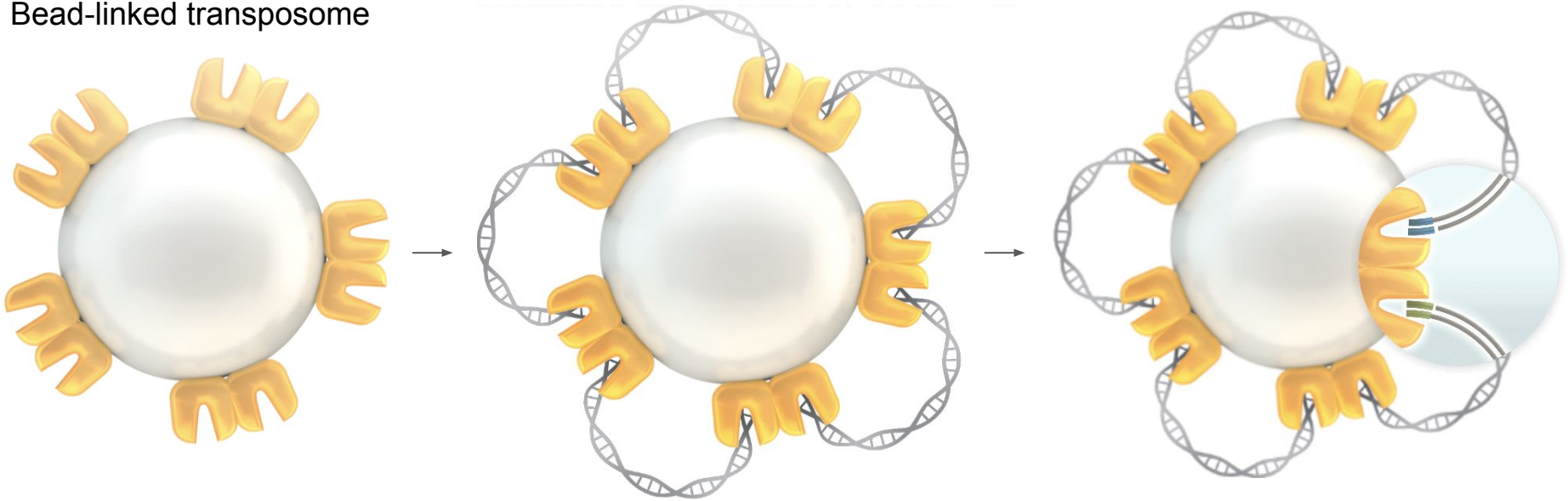
Step 3: Library prep - Illumina DNA prep (=Nextera) protocol

A. Tagmentation

Bead-linked transposome

DNA wraps around bead, forming a DNA-BLT complex

Transposome tagments DNA

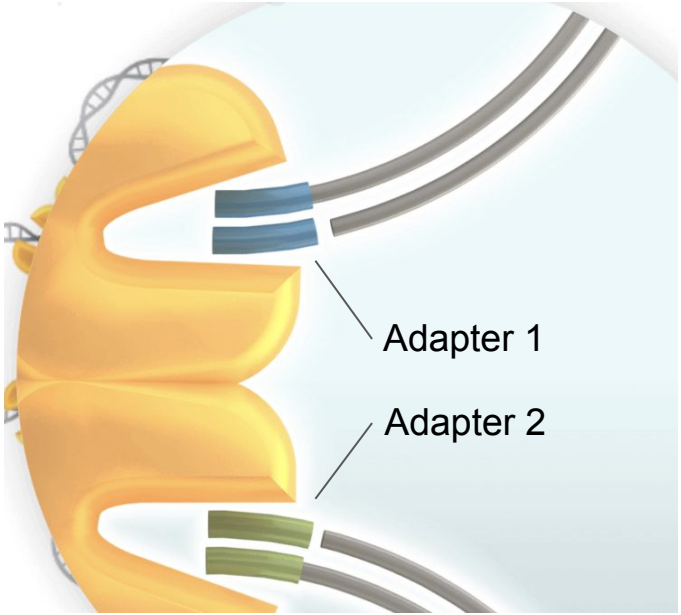


Tagmentation = cleaving and tagging of DNA with adapters (target insert size = 350 bp)

Adapted from illumina.com

Step 3: Library prep - Illumina DNA prep (=Nextera) protocol

A. Tagmentation



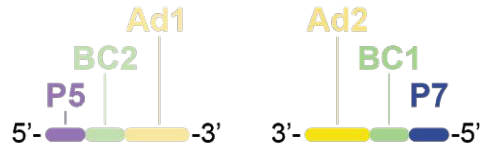
One single reaction fragments DNA and adds adapters

Adapted from illumina.com

Step 3: Library prep - Illumina DNA prep (=Nextera) protocol

B. Barcoding PCR

1. Add specific i5 and i7 oligos to each sample (includes barcodes)
2. PCR to add indices (=barcodes) and flow cell binding sequences to fragments
3. Bead clean-up (0.8x)

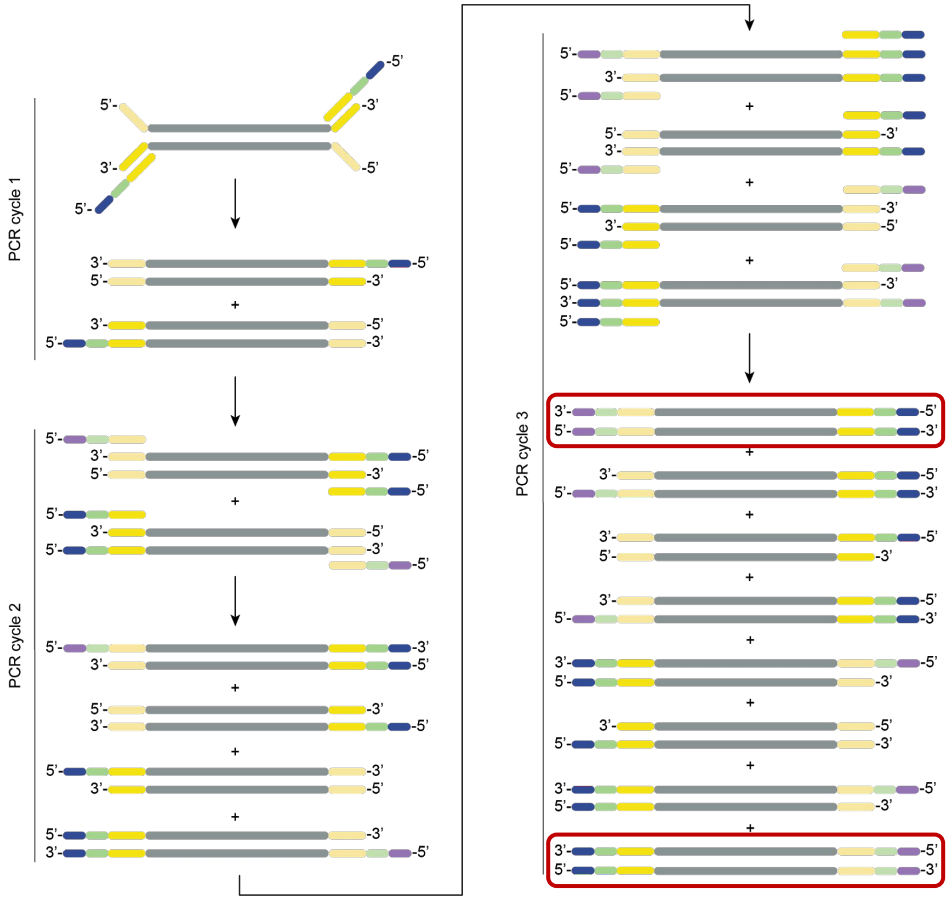


i5 barcoded primer i7 barcoded primer

Ad1/Ad2 = Adapters that allow initiation of sequencing

BC1/BC2 = Barcodes specific for each sample

P5/P7 = Flow cell binding sites

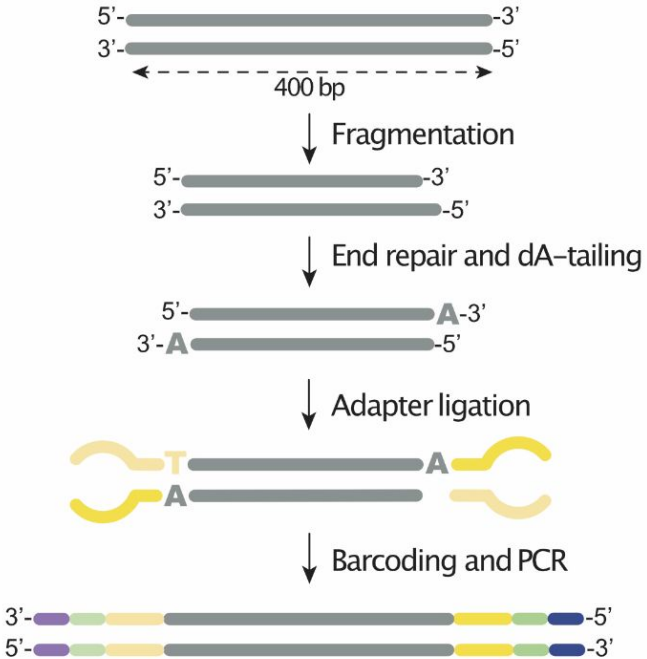


Step 3: Library prep

+ *Faster, no adapter dimer issues*
- *Little control over insert size*

Ligation protocol

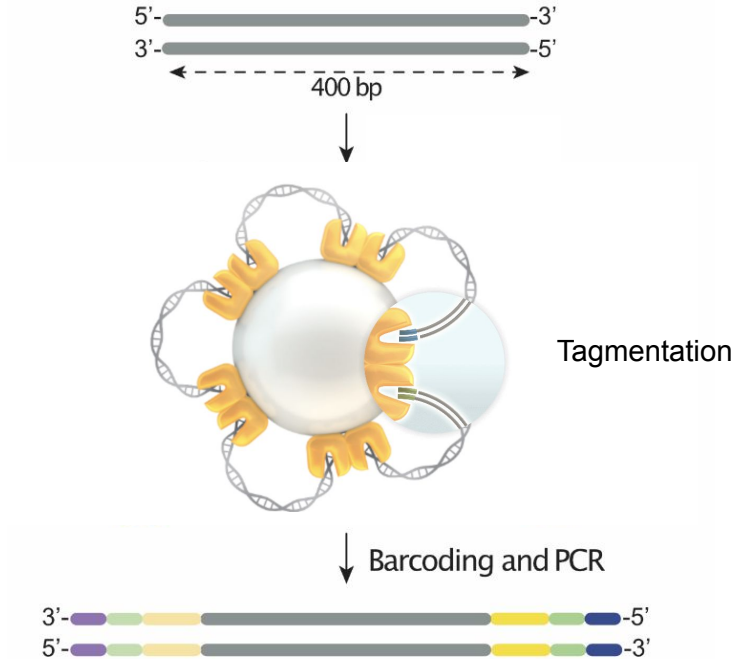
Illumina TruSeq DNA / NEBNext Ultra II DNA



OR

Tagmentation protocol

Illumina DNA prep (= Nextera Flex)



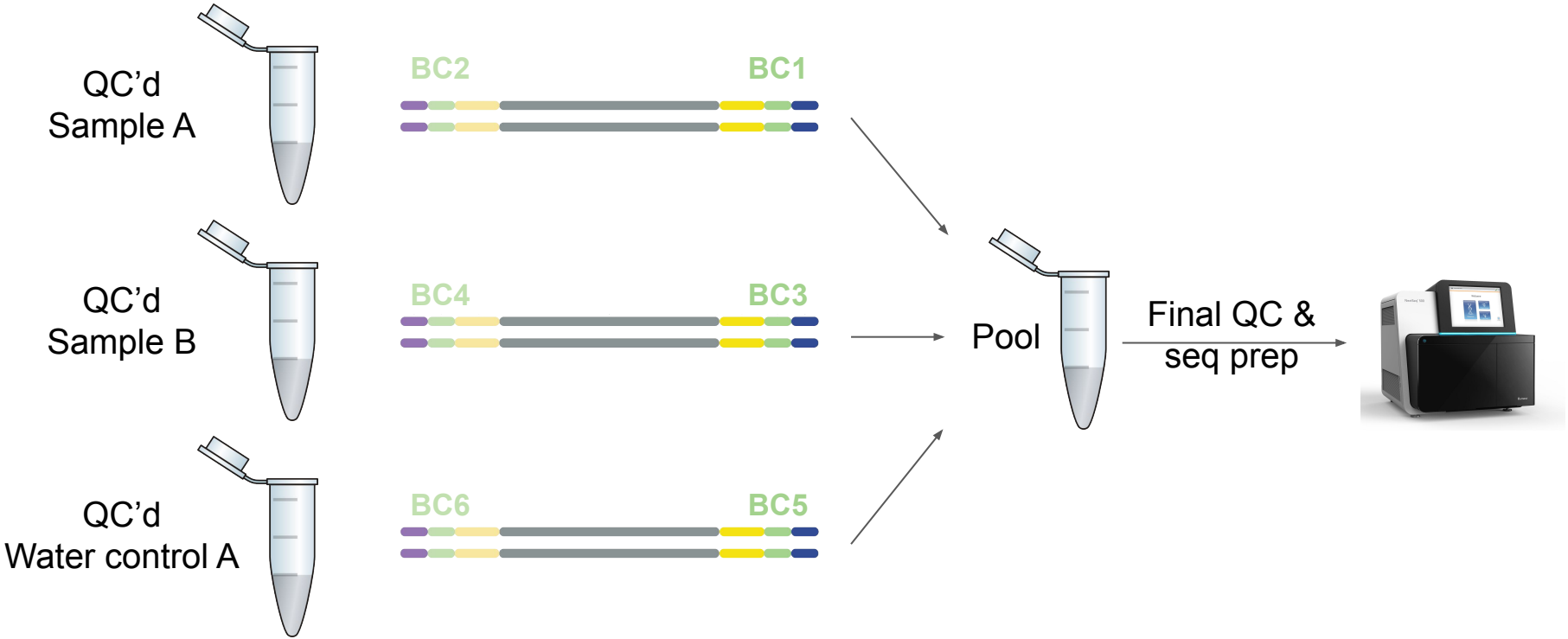
Step 4: QC individual libraries

- Important QC metrics:
  - Concentration ideal concentration 0.5-10 ng/µl
  - Fragment length ideal insert size ~300-350 bp
  - Absence of adapter dimers

More on library QC in Session 2

Step 5: Pooling samples for sequencing

- Each sample has a unique set of barcodes -> multiple samples can be pooled in one sequencing reaction



Step 5: Pooling samples for sequencing

How many samples can I pool together?

Depends on:

- Genome length: 29,903 kb for SARS-CoV-2
- Sequencing kit (sequencer-dependent): number & length of reads
- # Unique barcodes/indices available (max 384 from Illumina)

Examples: when aiming for average 1000x coverage of the SC2 genome:

- 1000 (coverage) x 29,903 (genome length) = 29,903,000 bp
  -> need to sequence ~ 30 M bp per sample

- iSeq 100 i1 Reagent Kit v2 (300 cycle): yields ~4 M reads
  - 4 M (# reads) x 300 (max #bp/read) = 1,200 M or 1.2 Gb
  - -> Can pool 1,200M/30M or 40 samples together

- MiSeq Reagent Kit v3 (150 cycle): yields 25 M reads
  - 25 M (# reads) x 150 (max #bp/read) = 3,750 M
  - -> Can pool 125 samples together

Illumina instruments (output)


iSeq (1.2 Gb)


MiniSeq (7.5 Gb)


MiSeq (15 Gb)


NextSeq (330 Gb)


NovaSeq (6 Tb)

SARS-CoV-2 amplicon sequencing

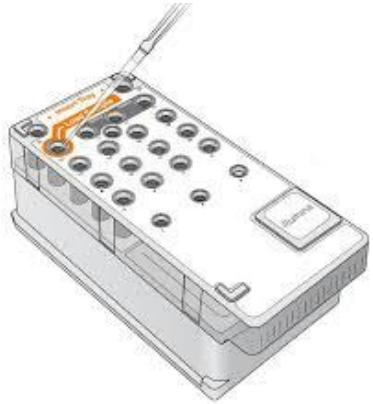
- ‘All-in-one’ SARS-CoV-2 sequencing kits available from Illumina, but are generally more expensive than standard ARTIC protocols:

  - AmpliSeq: smaller amplicons, no fragmentation required, adapters ligated on, includes human targets as control, online integrated analysis pipeline available (2,223 for + adapters)

  - CovidSeq : ARTIC-based, tagmentation protocol, includes human targets as control, online integrated analysis pipeline available, marketed for NovaSeq and NextSeq.

  - Respiratory Pathogen ID/AMR Enrichment Kit (enrichment by hybridization, NOT amplicon sequencing): Detects >280 respiratory pathogens, including SARS-CoV-2, designed to yield full genomes for SARS-CoV-2 and Influenza A/B. (VERY expensive)

Step 6: Final QC library and loading sequencer

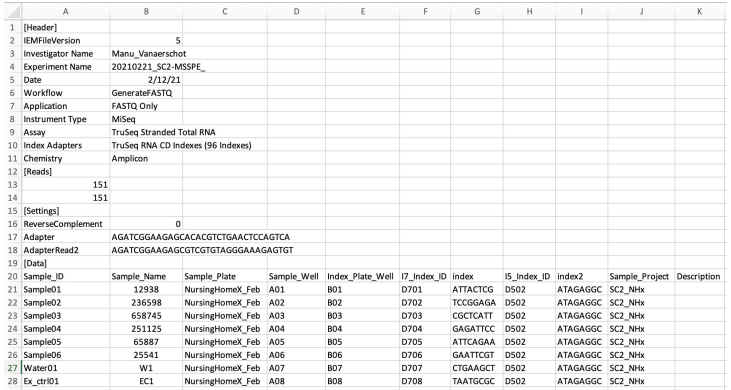


- Final QC: concentration - fragment length - absence of dimers
- Loading the sequencer: procedure depends on sequencer!
  - Properly thaw and mix sequencing kits
    - Thaw Reagent cartridge overnight at 4°C
    - Let flow cell and incorporation buffer adjust to room temp at least 30 min before loading
    - Invert cartridge 5 times to mix, avoid making bubbles
  - Make sure library is denatured if needed (e.g. for MiSeq runs)
    - Denaturing is part of library dilution process for MiSeq
  - Use proper loading concentration
    - May vary between individual instruments of the same type, keep a log.
    - Add 1-5% PhiX to library

More details on library QC in Session 2

Step 7: Start sequencing run

- Prepare Sample Sheet
  - Records experiment identifiers
  - List data analysis workflows that sequencer needs to perform
  - Provides all data required to demultiplex sequencing data
    - Sample names plus associated index/barcode sequences
  - Can be prepared manually or using Illumina Experiment Manager



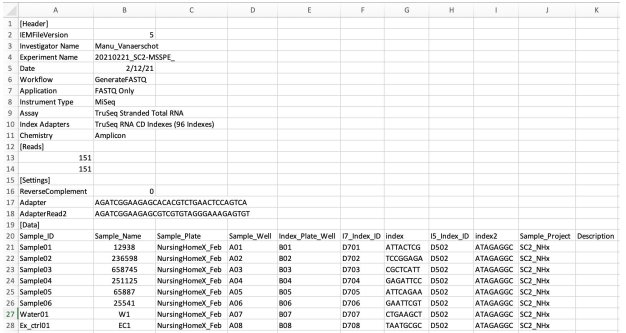
	A	B	C	D	E	F	G	H	I	J	K
1	[Header]										
2	IEMFileVersion	5									
3	Investigator Name	Manu_Vanaerschot									
4	Experiment Name	20210221_SC2-MSSPE_									
5	Date	2/12/21									
6	Workflow	GenerateFASTQ									
7	Application	FASTQ Only									
8	Instrument Type	MiSeq									
9	Assay	TruSeq Stranded Total RNA									
10	Index Adapters	TruSeq RNA CD Indexes (96 Indexes)									
11	Chemistry	Amplicon									
12	[Reads]										
13	151										
14	151										
15	[Settings]										
16	ReverseComplement	0									
17	Adapter	AGATCGGAAGAGCACACGTCTGAACTCCAGTCA									
18	AdapterRead2	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT									
19	[Data]										
20	Sample_ID	Sample_Name	Sample_Plate	Sample_Well	Index_Plate_Well	I7_Index_ID	index	I5_Index_ID	index2	Sample_Project	Description
21	Sample01	12938	NursingHomeX_Feb	A01	B01	D701	ATTACTCG	D502	ATAGAGGC	SC2_NHx	
22	Sample02	236598	NursingHomeX_Feb	A02	B02	D702	TCCGGAGA	D502	ATAGAGGC	SC2_NHx	
23	Sample03	658745	NursingHomeX_Feb	A03	B03	D703	CGCTCATT	D502	ATAGAGGC	SC2_NHx	
24	Sample04	251125	NursingHomeX_Feb	A04	B04	D704	GAGATTCC	D502	ATAGAGGC	SC2_NHx	
25	Sample05	65887	NursingHomeX_Feb	A05	B05	D705	ATTCAGAA	D502	ATAGAGGC	SC2_NHx	
26	Sample06	25541	NursingHomeX_Feb	A06	B06	D706	GAATTCGT	D502	ATAGAGGC	SC2_NHx	
27	Water01	W1	NursingHomeX_Feb	A07	B07	D707	CTGAAGCT	D502	ATAGAGGC	SC2_NHx	
28	Ex_ctrl01	EC1	NursingHomeX_Feb	A08	B08	D708	TAATGCGC	D502	ATAGAGGC	SC2_NHx	

Example sample sheet

[Header]								
Date	2020-05-01	General tracking info about the run						
Experiment Name	20200501_NH2							
Workflow	GenerateFASTQ	This tells the machine which type of analysis to perform on the data						
Application	FASTQ Only							
Chemistry	Amplicon							
[Reads]								
150	# of sequencing cycles to perform							
150								
						Index/barcode info		
[Data]								
Study_ID	Study_Description	Sample_ID	Sample_Name	Sample_Owner	Index_ID	Index	Index2_ID	Index2
1243	NursingHome1	Sample001	Sample001	John_Doe	Index1	AGTGCATTCGAG	Index20	TGCTAGTCCACT
7382	NursingHome1	Sample002	Sample002	John_Doe	Index2	TCACATTCGTCC	Index21	GGATTGGTGAAC
5927	NursingHome1	Sample003	Sample003	John_Doe	Index3	ACGAATCGACCA	Index22	AAGTGCCTCTTG
9283	NursingHome1	Sample004	Sample004	John_Doe	Index4	CTCATGCATCTC	Index23	CATGCAGATCGT
7493	NursingHome1	Sample005	Sample005	John_Doe	Index5	ACGATACGGCAA	Index24	ATGTCGGAGAAC
2932	NursingHome1	Sample006	Sample006	John_Doe	Index6	GTTCAGATGCTG	Index25	AGACTCCTGAAC
4543	NursingHome1	Sample007	Sample007	John_Doe	Index7	GAGTTCTTCGGT	Index26	TCTCTTGCTCGT
2131	NursingHome1	Sample008	Sample008	John_Doe	Index8	CAAGGCACTGTT	Index27	TAAGGCACGAAC
4432	NursingHome1	covid_neg	covid_neg	John_Doe	Index9	CTCATGTACACC	Index28	AGAGACCATGAG
5293	NursingHome1	water_control	water_control	John_Doe	Index10	GAGTCTCGATTC	Index29	AGCAGAGTACTC

Step 7: Start sequencing run

- Prepare Sample Sheet
  - Records experiment identifiers
  - List data analysis workflows that sequencer needs to perform
  - Provides all data required to demultiplex sequencing data
    - Sample names plus associated index/barcode sequences
  - Can be prepared manually or using Illumina Experiment Manager



	A	B	C	D	E	F	G	H	I	J	K
1	[Header]										
2	IEMFileVersion	5									
3	Investigator Name	Manu_Vanaerschot									
4	Experiment Name	20210221_SC2-MISSPE_									
5	Date	2/12/21									
6	Workflow	GenerateFASTQ									
7	Application	FASTQ Only									
8	Instrument Type	MiSeq									
9	Assay	TruSeq Stranded Total RNA									
10	Index Adapters	TruSeq RNA CD Indexes (96 Indexes)									
11	Chemistry	Amplicon									
12	[Reads]										
13	151										
14	151										
15	[Settings]										
16	ReverseComplement	0									
17	Adapter	AGATCGGAAGAGCACACGTCTGAACTCCAGTCA									
18	AdapterRead2	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT									
19	[Data]										
20	Sample_ID	Sample_Name	Sample_Plate	Sample_Well	Index_Plate_Well	I7_Index_ID	index	I5_Index_ID	index2	Sample_Project	Description
21	Sample01	12938	NursingHomeX_Feb	A01	B01	D701	ATTACTCG	D502	ATAGAGGC	SC2_NHx	
22	Sample02	236598	NursingHomeX_Feb	A02	B02	D702	TCCGGAGA	D502	ATAGAGGC	SC2_NHx	
23	Sample03	658745	NursingHomeX_Feb	A03	B03	D703	CGCTCATT	D502	ATAGAGGC	SC2_NHx	
24	Sample04	251125	NursingHomeX_Feb	A04	B04	D704	GAGATTCC	D502	ATAGAGGC	SC2_NHx	
25	Sample05	65887	NursingHomeX_Feb	A05	B05	D705	ATTCAGAA	D502	ATAGAGGC	SC2_NHx	
26	Sample06	25541	NursingHomeX_Feb	A06	B06	D706	GAATTCGT	D502	ATAGAGGC	SC2_NHx	
27	Water01	W1	NursingHomeX_Feb	A07	B07	D707	CTGAAGCT	D502	ATAGAGGC	SC2_NHx	
28	Ex_ctrl01	EC1	NursingHomeX_Feb	A08	B08	D708	TAATGCGC	D502	ATAGAGGC	SC2_NHx	

- How to start a run:
  - Follow on-screen instructions
  - Learn more with this Illumina video on MiSeq runs:

    https://support.illumina.com/content/dam/illumina-support/courses/MiSeq_How_to_Start_a_Run/story.html
  - Explore illumina.com/training.html for more videos (incl other instruments)

Summary SARS-CoV-2 amplicon sequencing

DNA ← RNA

1. Reverse transcribe RNA into DNA

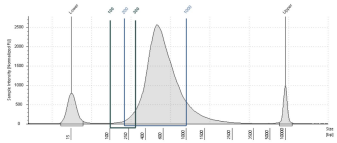
2. Amplify SARS-CoV-2 by multiplex PCR
   - Different primer sets available: ARTIC V3/V4, Midnight, ….

3. Library prep requires fragmentation, adapter ligation and barcoding
   - Use a suitable DNA library prep kit: ligation or tagmentation

4. QC individual sample libraries
   - Concentration, length, no adapter dimers

Session 2

5. Pool samples together for maximum throughput
   - # depends on sequencer and viral content of original sample



6. Final QC of sequencing library & loading the sequencer

Session 2

7. Run the Sequencer

8. QC of sequencing run

Session 2