

# Illumina: systems & strategies



<http://www.illumina.com/systems/sequencing.html>

# Overview

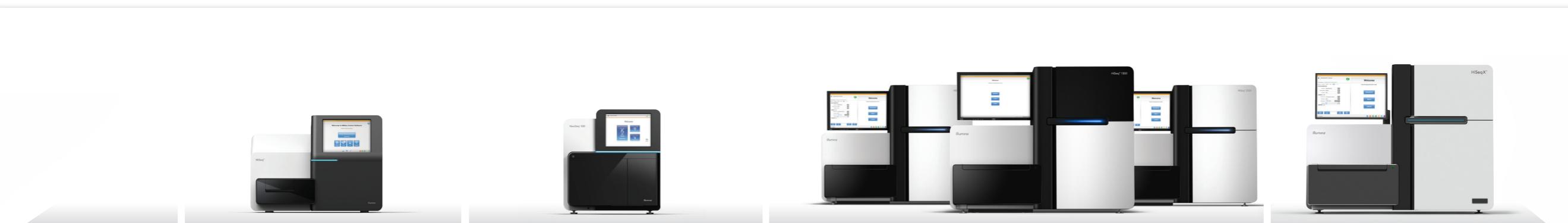
- ✓ Illumina's sequencing systems
- ✓ Standard library prep (Tru-Seq)
- ✓ Tagmentation-based approach (Nextera)
  - Long-insert library prep (mate pairs)
- ✓ Modifications/Strategies for application-specific libraries
- ✓ Multiplexing

# Overview

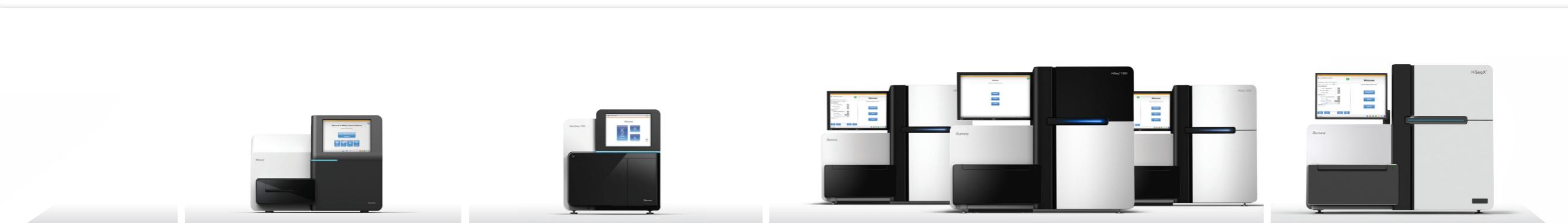
- ✓ Illumina's sequencing systems
- ✓ Standard library prep (Tru-Seq)
- ✓ Tagmentation-based approach (Nextera)
  - Long-insert library prep (mate pairs)
- ✓ Modifications/Strategies for application-specific libraries
- ✓ Multiplexing

				
<b>MiniSeq System</b>	<b>MiSeq Series</b>	<b>NextSeq Series</b>	<b>HiSeq Series</b>	<b>HiSeq X Series*</b>
Amplicon, targeted RNA, small RNA, and targeted gene panel sequencing.	Small genome, amplicon, and targeted gene panel sequencing.	Everyday exome, transcriptome, and targeted resequencing.	Production-scale genome, exome, transcriptome sequencing, and more.	Population- and production-scale whole-genome sequencing.

# Illumina's sequencing systems

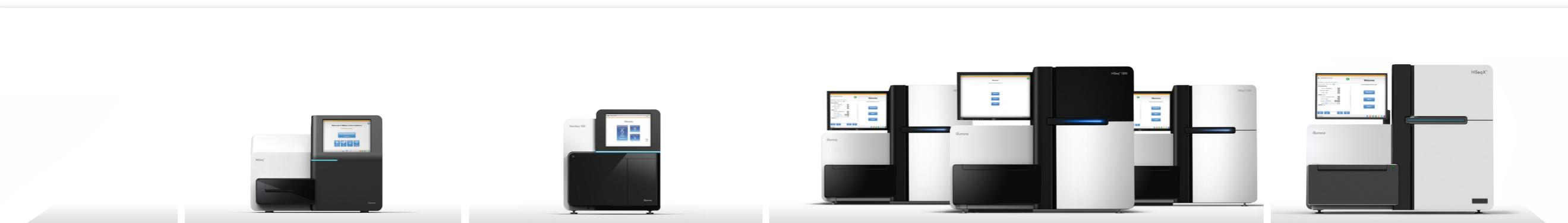
							
Product	MiSeq	NextSeq 500*	HiSeq 2500	HiSeq 3000	HiSeq 4000	HiSeq X Five†	HiSeq X Ten†
Description	Focused power Speed and simplicity for targeted and small-genome sequencing	Flexible power Speed and simplicity for everyday genomics	Production power Power and efficiency for large-scale genomics	Production power Maximum throughput and lowest cost for production-scale genomics	Population power Maximum throughput and lowest cost for population- and production-scale human whole-genome sequencing		
Key methods	Small genome, amplicon, targeted gene panel sequencing	Everyday genome, exome, transcriptome sequencing, and more	Production-scale genome, exome, transcriptome sequencing, and more	Production-scale genome, exome, transcriptome sequencing, and more	Population-scale human whole-genome sequencing		
Run mode	—	Mid-output      High-output	Rapid run      High-output	—      —	—      —	—	—
Flow cells processed per run	1	1      1	1 or 2      1 or 2	1      1 or 2	1 or 2	1 or 2	1 or 2      1 or 2
Output range	0.3–15 Gb	20–39 Gb      30–120 Gb	10–300 Gb      50–1000 Gb	125–750 Gb      125–1500 Gb	900–1800 Gb	900–1800 Gb	
Run time	5–55 hours	15–26 hours      12–30 hours	7–60 hours      < 1–6 days	< 1–3.5 days      < 1–3.5 days	< 3 days	< 3 days	
Reads per flow cell <sup>‡</sup>	25 million <sup>§</sup>	130 million      400 million	300 million      2 billion	2.5 billion      2.5 billion	3 billion	3 billion	
Maximum read length	2 × 300 bp	2 × 150 bp      2 × 150 bp	2 × 250 bp      2 × 125 bp	2 × 150 bp      2 × 150 bp	2 × 150 bp	2 × 150 bp	

# Illumina's sequencing systems



Product	MiSeq	NextSeq 500*	HiSeq 2500	HiSeq 3000	HiSeq 4000	HiSeq X Five†	HiSeq X Ten†
Description	Focused power Speed and simplicity for targeted and small-genome sequencing	Flexible power Speed and simplicity for everyday genomics	Production power Power and efficiency for large-scale genomics	Production power Maximum throughput and lowest cost for production-scale genomics		Population power Maximum throughput and lowest cost for population- and production-scale human whole-genome sequencing	
Key methods	Small genome, amplicon, targeted gene panel sequencing	Everyday genome, exome, transcriptome sequencing, and more		Production-scale genome, exome, transcriptome sequencing, and more		Population-scale human whole-genome sequencing	
Run mode	—	Mid-output High-output	Rapid run High-output	— —	— —	— —	— —
Flow cells processed per run	1	1	1	1 or 2	1 or 2	1	1 or 2
Output range	0.3–15 Gb	20–39 Gb	30–120 Gb	10–300 Gb	50–1000 Gb	125–750 Gb	125–1500 Gb
Run time	5–55 hours	15–26 hours	12–30 hours	7–60 hours	< 1–6 days	< 1–3.5 days	< 1–3.5 days
Reads per flow cell <sup>‡</sup>	25 million <sup>§</sup>	130 million	400 million	300 million	2 billion	2.5 billion	2.5 billion
Maximum read length	2 × 300 bp	2 × 150 bp	2 × 150 bp	2 × 250 bp	2 × 125 bp	2 × 150 bp	2 × 150 bp

# Illumina's sequencing systems



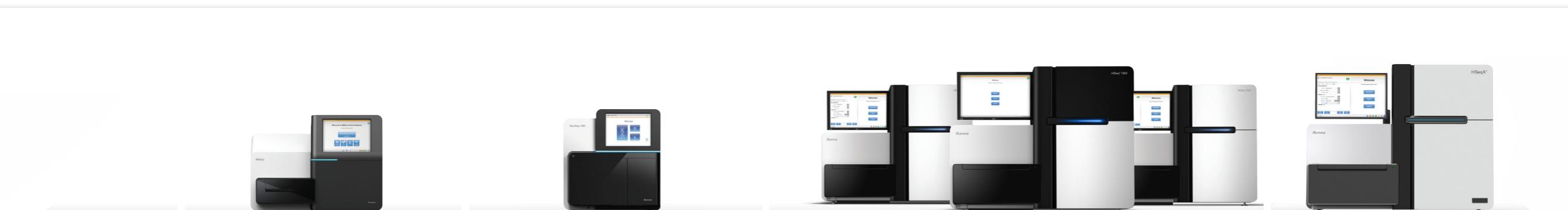
Product	MiSeq	NextSeq 500*	HiSeq 2500	HiSeq 3000	HiSeq 4000	HiSeq X Five†	HiSeq X Ten†	
Description	Focused power Speed and simplicity for targeted and small-genome sequencing	Flexible power Speed and simplicity for everyday genomics	Production power Power and efficiency for large-scale genomics	Production power Maximum throughput and lowest cost for production-scale genomics		Population power Maximum throughput and lowest cost for population- and production-scale human whole-genome sequencing		
Key methods	Small genome, amplicon, targeted gene panel sequencing	Everyday genome, exome, transcriptome sequencing, and more		Production-scale genome, exome, transcriptome sequencing, and more		Population-scale human whole-genome sequencing		
Run mode	—	Mid-output High-output	Rapid run High-output	— —	— —	— —	— —	
Flow cells processed per run	1	1	1	1 or 2	1 or 2	1	1 or 2	1 or 2
Output range	0.3–15 Gb	20–39 Gb	30–120 Gb	10–300 Gb	50–1000 Gb	125–750 Gb	125–1500 Gb	900–1800 Gb
Run time	5–55 hours	15–26 hours	12–30 hours	7–60 hours	< 1–6 days	< 1–3.5 days	< 1–3.5 days	< 3 days
Reads per flow cell <sup>‡</sup>	25 million <sup>§</sup>	130 million	400 million	300 million	2 billion	2.5 billion	2.5 billion	3 billion
Maximum read length	2 × 300 bp	2 × 150 bp	2 × 150 bp	2 × 250 bp	2 × 125 bp	2 × 150 bp	2 × 150 bp	2 × 150 bp

# Illumina's sequencing systems



Product	MiSeq	NextSeq 500*	HiSeq 2500	HiSeq 3000	HiSeq 4000	HiSeq X Five†	HiSeq X Ten†		
Description	Focused power Speed and simplicity for targeted and small-genome sequencing	Flexible power Speed and simplicity for everyday genomics	Production power Power and efficiency for large-scale genomics	Production power Maximum throughput and lowest cost for production-scale genomics		Population power Maximum throughput and lowest cost for population- and production-scale human whole-genome sequencing			
Key methods	Small genome, amplicon, targeted gene panel sequencing	Everyday genome, exome, transcriptome sequencing, and more		Production-scale genome, exome, transcriptome sequencing, and more		Population-scale human whole-genome sequencing			
Run mode	—	Mid-output High-output	Rapid run High-output	— —	— —	— —	— —		
Flow cells processed per run	1	1	1	1 or 2	1 or 2	1	1 or 2	1 or 2	
Output range	0.3–15 Gb	20–39 Gb	30–120 Gb	10–300 Gb	50–1000 Gb	125–750 Gb	125–1500 Gb	900–1800 Gb	900–1800 Gb
Run time	5–55 hours	15–26 hours	12–30 hours	7–60 hours	< 1–6 days	< 1–3.5 days	< 1–3.5 days	< 3 days	< 3 days
Reads per flow cell <sup>‡</sup>	25 million <sup>§</sup>	130 million	400 million	300 million	2 billion	2.5 billion	2.5 billion	3 billion	3 billion
Maximum read length	2 × 300 bp	2 × 150 bp	2 × 150 bp	2 × 250 bp	2 × 125 bp	2 × 150 bp	2 × 150 bp	2 × 150 bp	2 × 150 bp

# Illumina's sequencing systems



Product	MiSeq	NextSeq 500*	HiSeq 2500	HiSeq 3000	HiSeq 4000	HiSeq X Five†	HiSeq X Ten†		
Description	Focused power Speed and simplicity for targeted and small-genome sequencing	Flexible power Speed and simplicity for everyday genomics	Production power Power and efficiency for large-scale genomics	Production power Maximum throughput and lowest cost for production-scale genomics		Population power Maximum throughput and lowest cost for population- and production-scale human whole-genome sequencing			
Key methods	Small genome, amplicon, targeted gene panel sequencing	Everyday genome, exome, transcriptome sequencing, and more		Production-scale genome, exome, transcriptome sequencing, and more		Population-scale human whole-genome sequencing			
Run mode	—	Mid-output High-output	Rapid run High-output	— —	— —	— —	— —		
Flow cells processed per run	1	1	1	1 or 2	1 or 2	1	1 or 2	1 or 2	
Output range	0.3–15 Gb	20–39 Gb	30–120 Gb	10–300 Gb	50–1000 Gb	125–750 Gb	125–1500 Gb	900–1800 Gb	900–1800 Gb
Run time	5–55 hours	15–26 hours	12–30 hours	7–60 hours	< 1–6 days	< 1–3.5 days	< 1–3.5 days	< 3 days	< 3 days
Reads per flow cell <sup>‡</sup>	25 million <sup>§</sup>	130 million	400 million	300 million	2 billion	2.5 billion	2.5 billion	3 billion	3 billion
Maximum read length	2 × 300 bp	2 × 150 bp	2 × 150 bp	2 × 250 bp	2 × 125 bp	2 × 150 bp	2 × 150 bp	2 × 150 bp	2 × 150 bp

# Illumina's sequencing systems

## [More info?](#)



**MiniSeq System**



**MiSeq Series**



**NextSeq Series**



**HiSeq Series**



**HiSeq X Series\***

Amplicon, targeted RNA,  
small RNA, and targeted  
gene panel sequencing.

Small genome,  
amplicon, and  
targeted gene panel  
sequencing.

Everyday exome,  
transcriptome, and  
targeted resequencing.

Production-scale  
genome, exome,  
transcriptome  
sequencing, and more.

Population- and  
production-scale whole-  
genome sequencing.

# Illumina's sequencing systems

# Overview

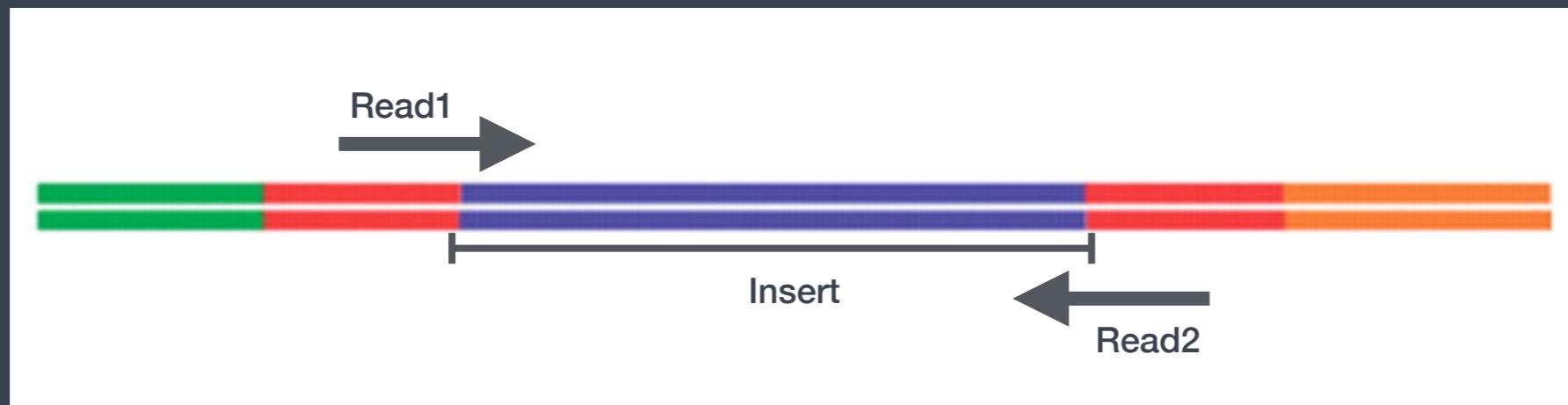
- ✓ Illumina's Sequencing systems
- ✓ Standard library prep (Tru-Seq)
- ✓ Tagmentation-based approach (Nextera)
  - Long-insert library prep (mate pairs)
- ✓ Modifications/Strategies for application-specific libraries
- ✓ Multiplexing



Typical steps in standard library preparation

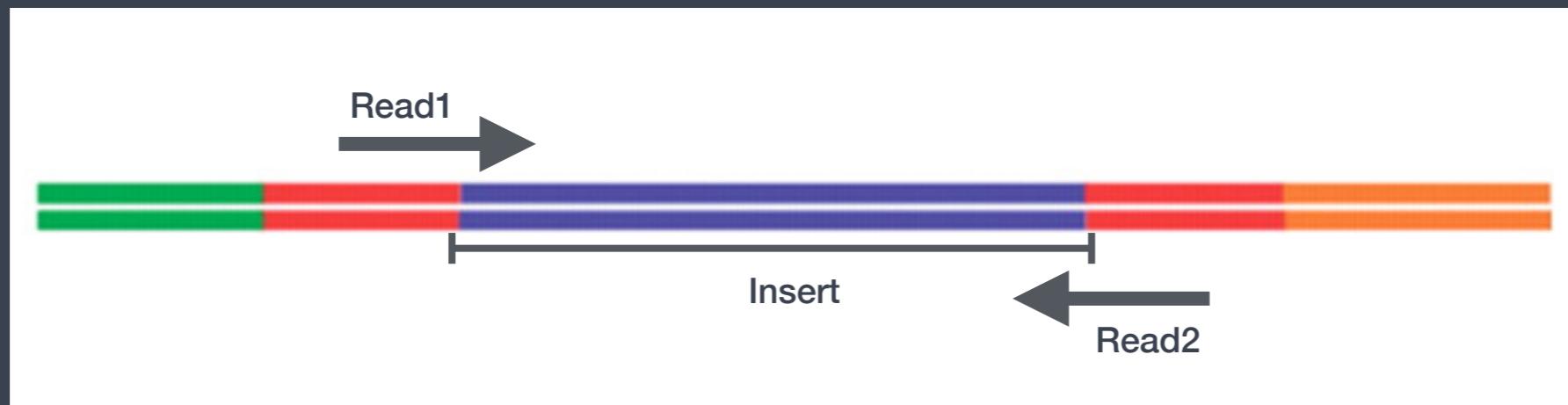


Options for sequencing



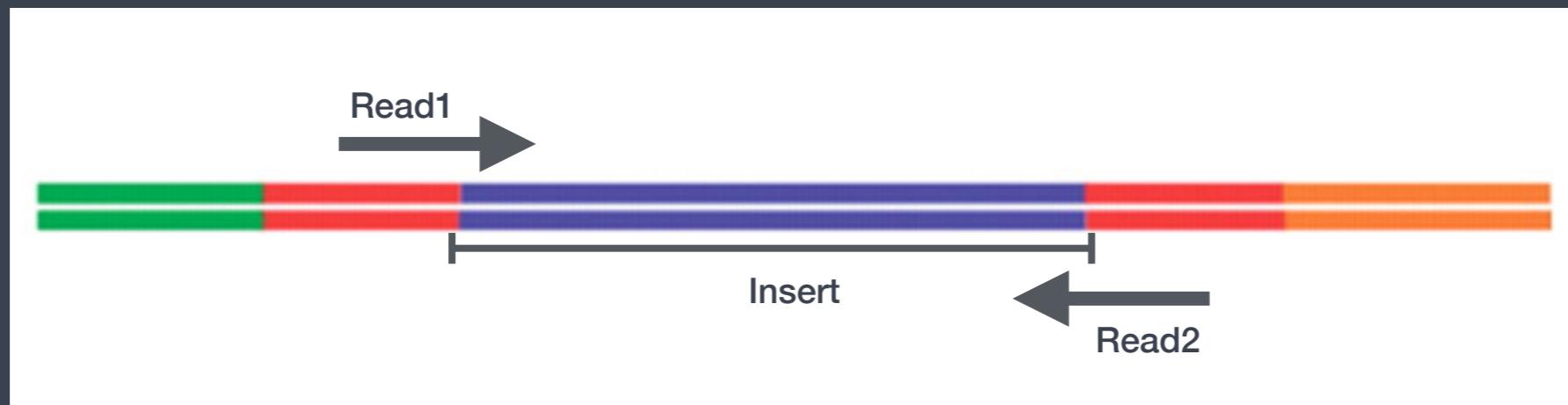
- ✓ SE - Single end dataset => Only Read1

## Options for sequencing



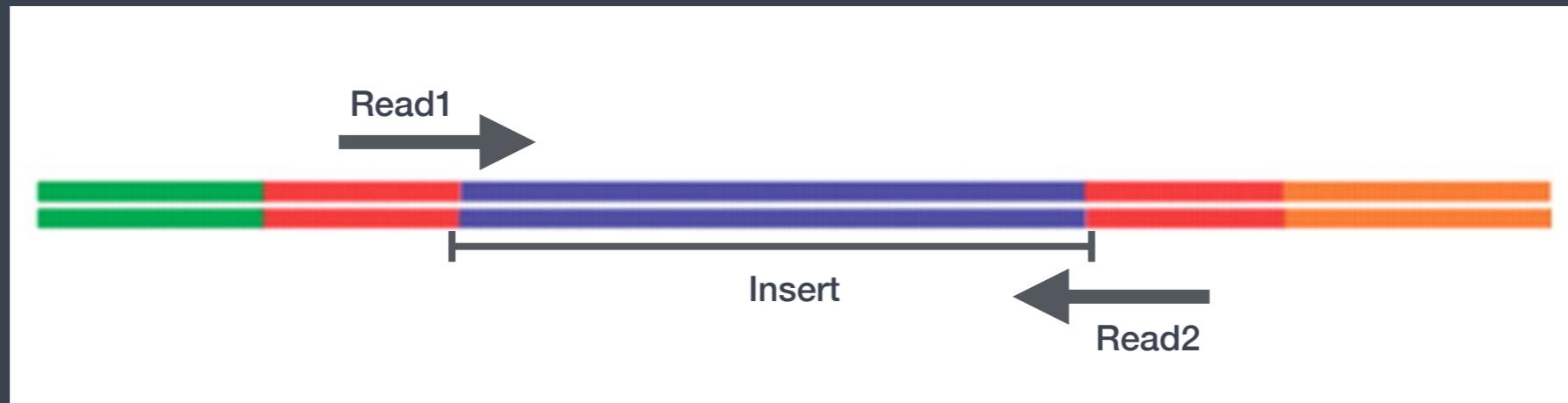
- ✓ SE - Single end dataset => Only Read1
- ✓ PE - Paired-end dataset => Read1 + Read2
  - can be 2 separate FastQ files or just one with interleaved pairs
  - insert refers to the DNA fragment\*\* flanked by the adapters

## Options for sequencing



- ✓ SE - Single end dataset => Only Read1
- ✓ PE - Paired-end dataset => Read1 + Read2
  - can be 2 separate FastQ files or just one with interleaved pairs
  - insert refers to the DNA fragment\*\* flanked by the adapters
- ✓ Read length - 50bp - 250bp, depends on the sequencer

## Options for sequencing



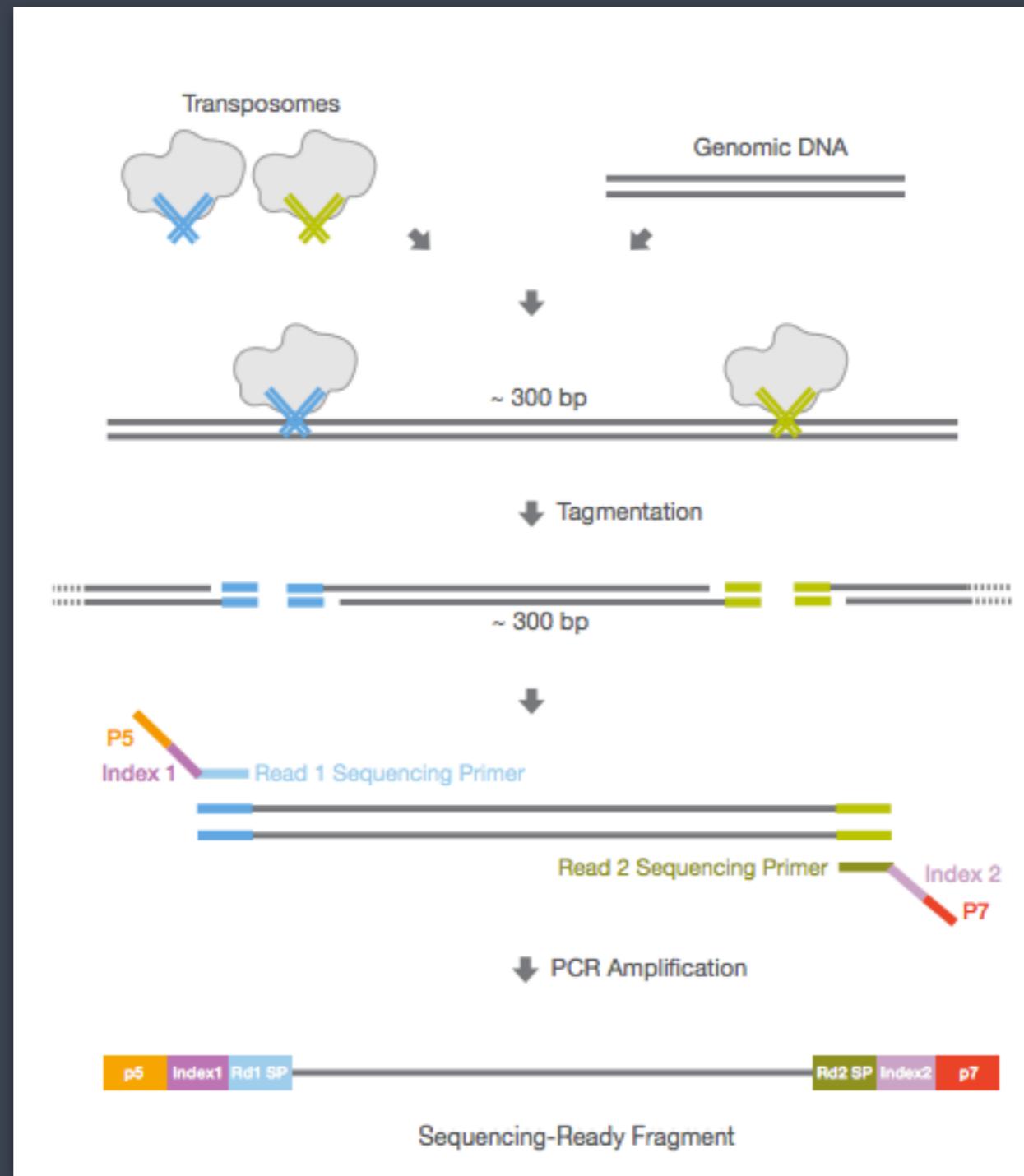
- ✓ SE - Single end dataset => Only Read1
- ✓ PE - Paired-end dataset => Read1 + Read2
  - can be 2 separate FastQ files or just one with interleaved pairs
  - insert refers to the DNA fragment\*\* flanked by the adapters
- ✓ Read length - 50bp - 250bp, depends on the sequencer

*\*\* “fragment” during library prep (Illumina) refers to the whole piece of DNA (insert + adapters). But, during downstream processing steps “fragment” can sometime refer to only the insert.*

## Options for sequencing

# Overview

- ✓ Illumina's Sequencing systems
- ✓ Standard library prep (Tru-Seq)
- ✓ Tagmentation-based approach (Nextera)
  - Long-insert library prep (mate pairs)
- ✓ Modifications/Strategies for application-specific libraries
- ✓ Multiplexing



# Tagmentation-based approach

(DNA fragmentation facilitated by transposon activity)

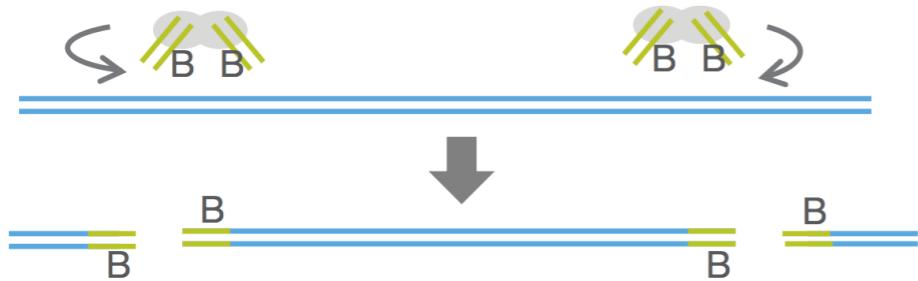
- ✓ *Nextera* from Illumina
- ✓ Transposomes contain the transposase and an oligo (transposon +adapter)
- ✓ Very fast and efficient for DNA library preps
- ✓ Works with small amounts of DNA
- ✓ Needs very precise DNA quantitation (Qubit)
- ✓ Can be used effectively for long inserts (3Kb - ~15Kb)

Tagmentation-based approach  
(DNA fragmentation facilitated by transposon activity)

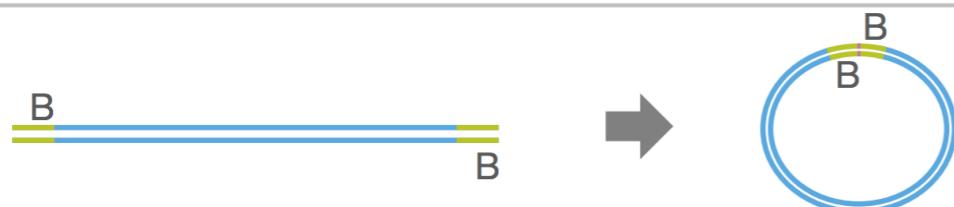
- ✓ Paired data from long fragments (3Kb to ~15Kb)
  - typically used for genome assemblies
  - span large areas of repeats
  - join contigs to build scaffolds
  - discover large genomic variations
- ✓ PE data from these libraries are referred to as “mate pairs”



Long-insert libraries



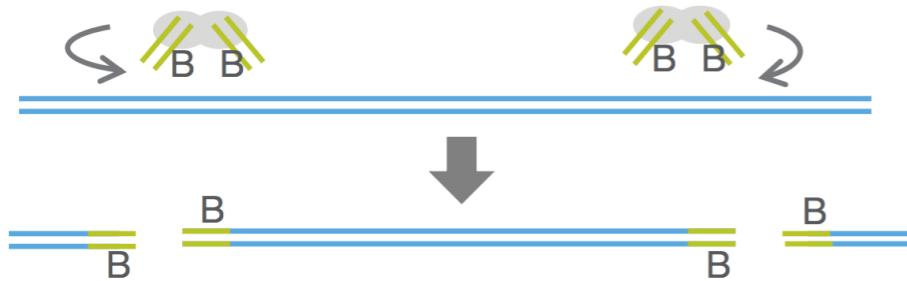
Genomic DNA (blue) is fragmented by the tagment enzyme, which attaches a biotinylated junction adapter (green) to both ends of the fragmented molecule.



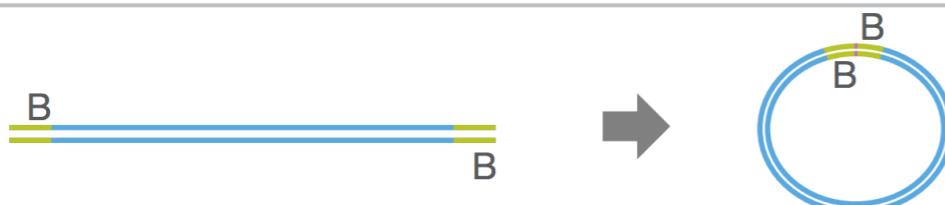
The fragmented DNA molecules are then circularized and the ends of the genomic fragment are linked by two copies of the biotin junction adapter.



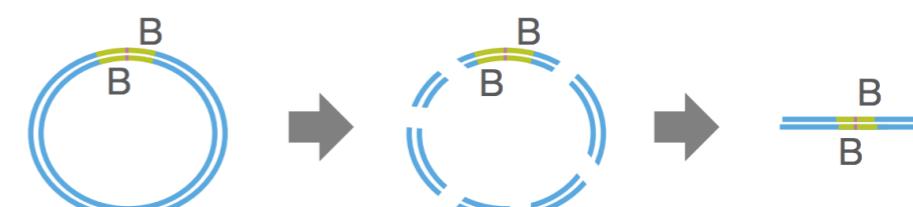
# Long-insert libraries



Genomic DNA (blue) is fragmented by the tagment enzyme, which attaches a biotinylated junction adapter (green) to both ends of the fragmented molecule.



The fragmented DNA molecules are then circularized and the ends of the genomic fragment are linked by two copies of the biotin junction adapter.



Circularized molecules are fragmented again, yielding smaller fragments. Fragments containing the original junction are enriched via the biotin tag (B) in the junction adapter.



After end repair and A-tailing, adapters (gray and purple) are added, enabling amplification and sequencing.

## Long-insert libraries



Genomic DNA (blue) is fragmented by the fragment enzyme, which attaches a biotinylated junction adapter (green) to both ends of the fragmented molecule.

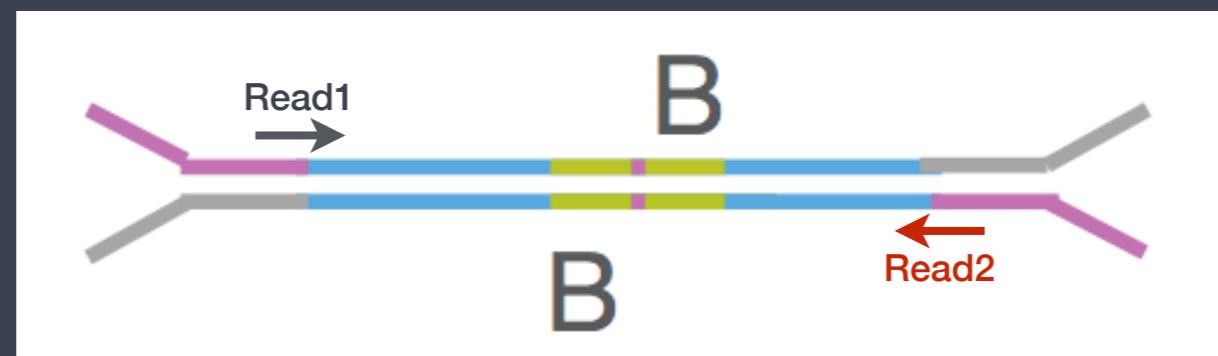


After end repair and A-tailing, adapters (gray and purple) are added, enabling amplification and sequencing.



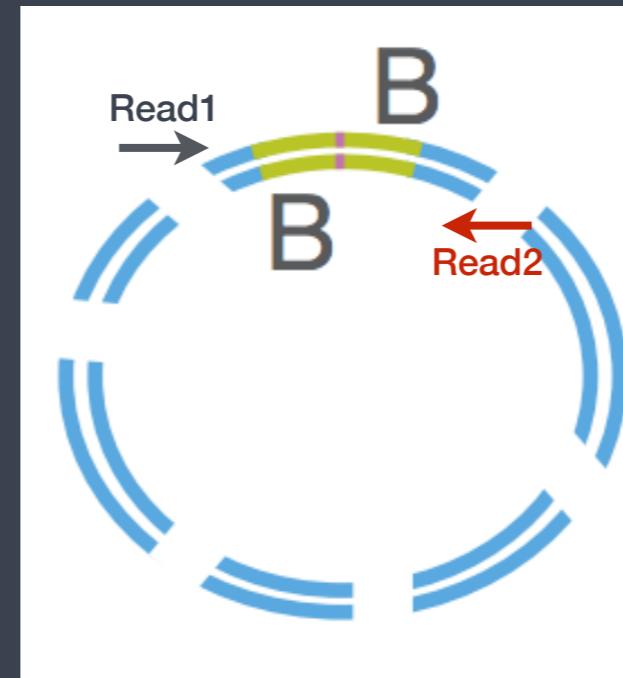
After end repair and A-tailing, adapters (gray and purple) are added, enabling amplification and sequencing.

## Long-insert libraries



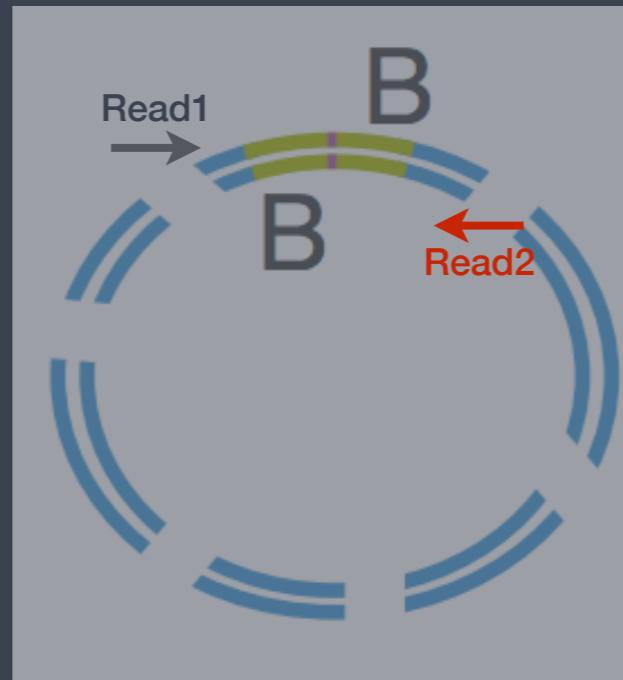
Long-insert libraries

Going backwards  
↓



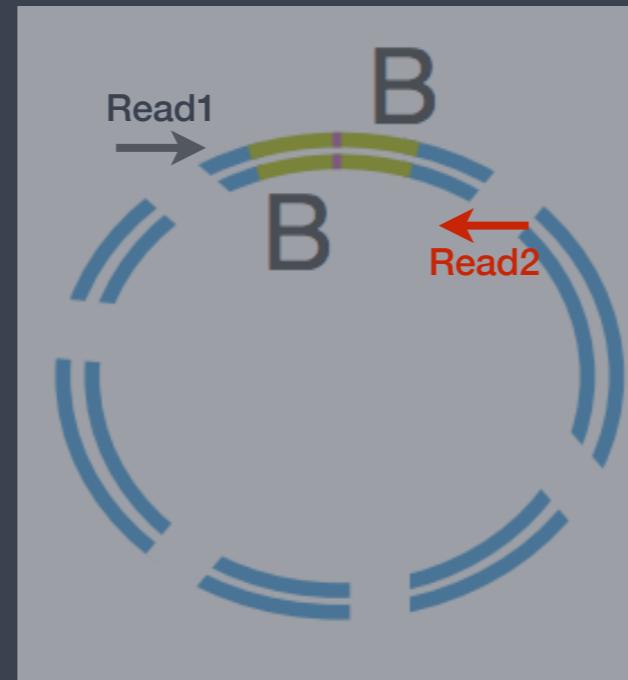
Long-insert libraries

Going backwards



Long-insert libraries

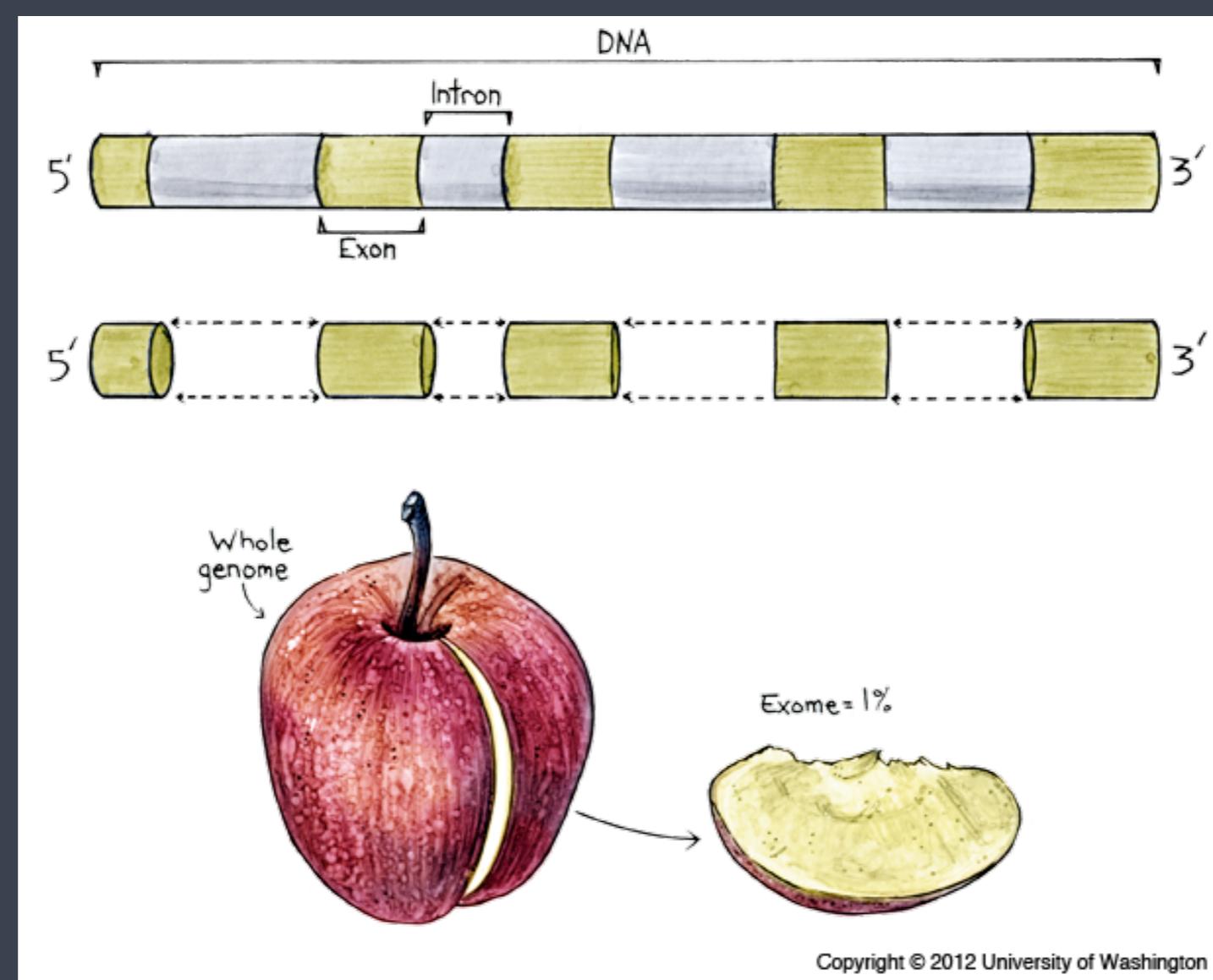
Going backwards



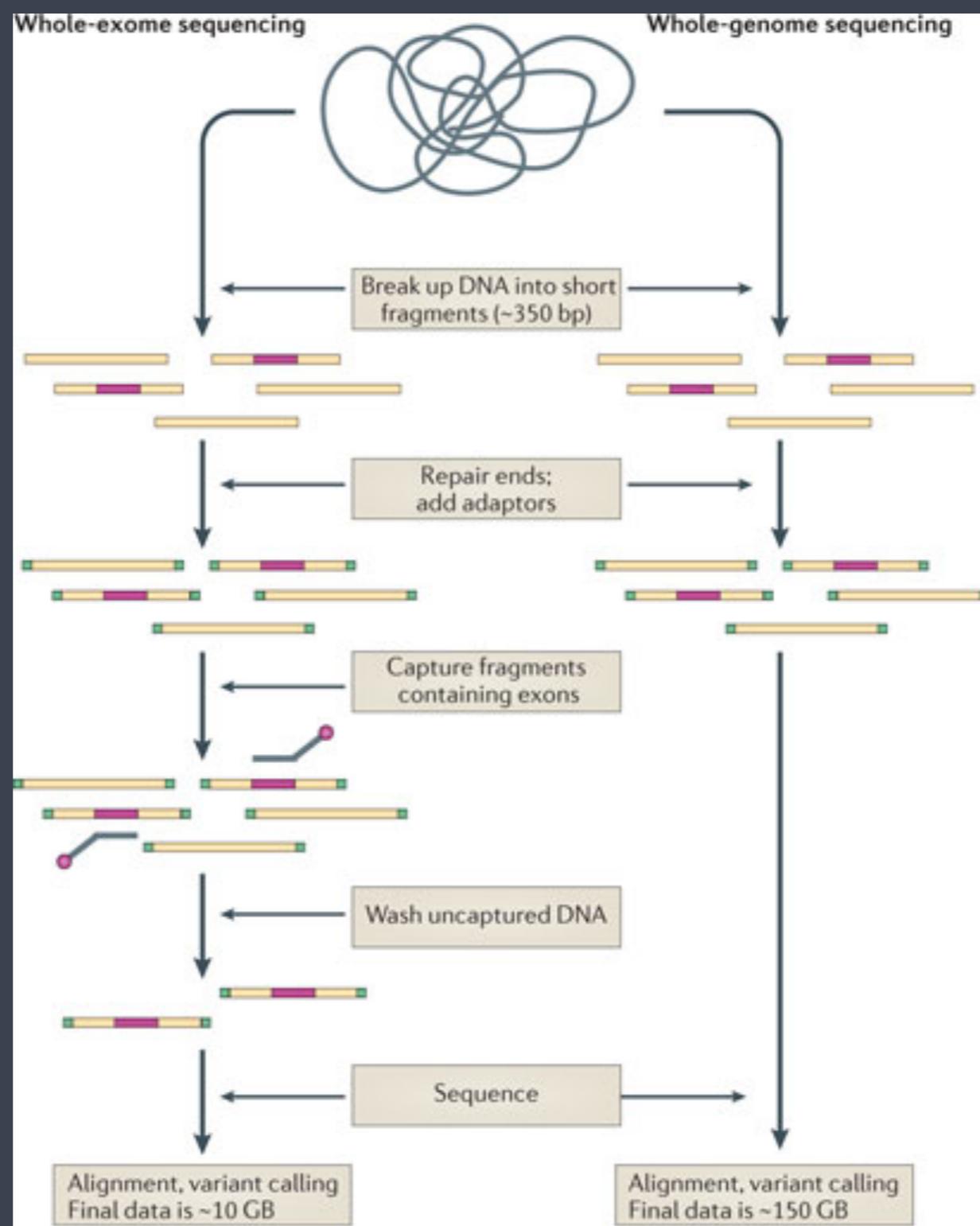
Long-insert libraries

# Overview

- ✓ Illumina's Sequencing systems
- ✓ Standard library prep (Tru-Seq)
- ✓ Tagmentation-based approach (Nextera)
  - Long-insert library prep (mate pairs)
- ✓ Modifications/Strategies for application-specific libraries
- ✓ Multiplexing

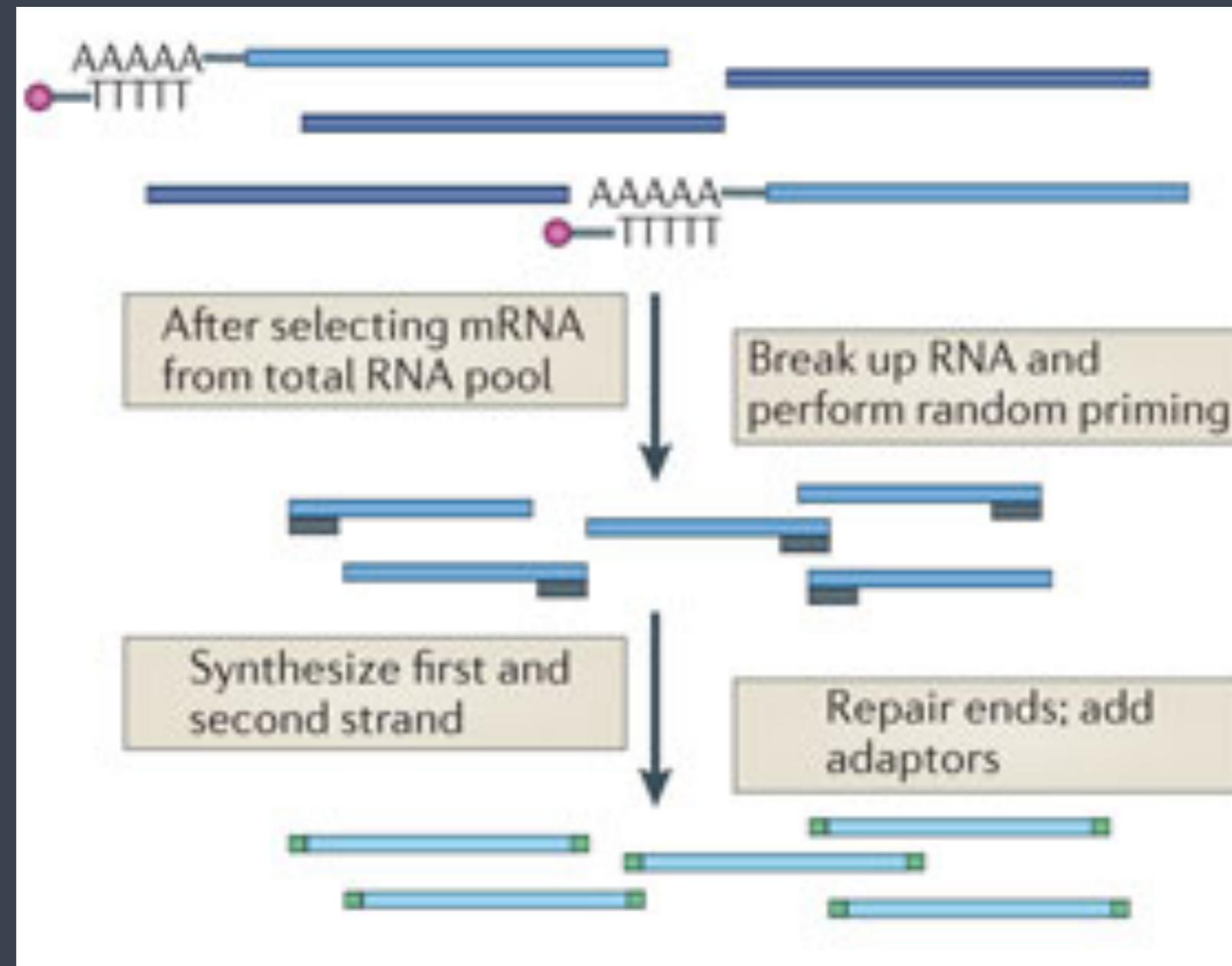


# Whole Genome Sequencing (WGS) vs Whole Exome Sequencing (WES)



(Exome)

(Genome)



# Transcriptomics

- ✓ Variant Detection
  - Whole genomes
  - Exomes
- ✓ Expression analysis: Transcriptomes
- ✓ Immunoprecipitation-based studies
  - ChIP-Seq
  - CLIP-Seq
- ✓ Environmental samples
  - 16S rRNA-based metagenomics
  - Whole metagenomes and metatranscriptomes
- ✓ many more....

## Modifications/Strategies for application-specific libraries

Refer to <http://www.illumina.com/applications/sequencing/ngs-library-prep/library-prep-methods.html> for a more complete list

- ✓ Variant Detection
  - Whole genomes (HiSeq X, HiSeq, MiSeq, NextSeq)
  - Exomes (HiSeq, MiSeq, NextSeq)
- ✓ Expression analysis: Transcriptomes (HiSeq, MiSeq, NextSeq)
- ✓ Immunoprecipitation-based studies (HiSeq, NextSeq)
  - ChIP-Seq
  - CLIP-Seq
- ✓ Environmental samples
  - 16S rRNA-based metagenomics (MiSeq)
  - Whole metagenomes and metatranscriptomes (HiSeq, MiSeq, NextSeq)
- ✓ many more....

## Modifications/Strategies for application-specific libraries

Refer to <http://www.illumina.com/applications/sequencing/ngs-library-prep/library-prep-methods.html> for a more complete list

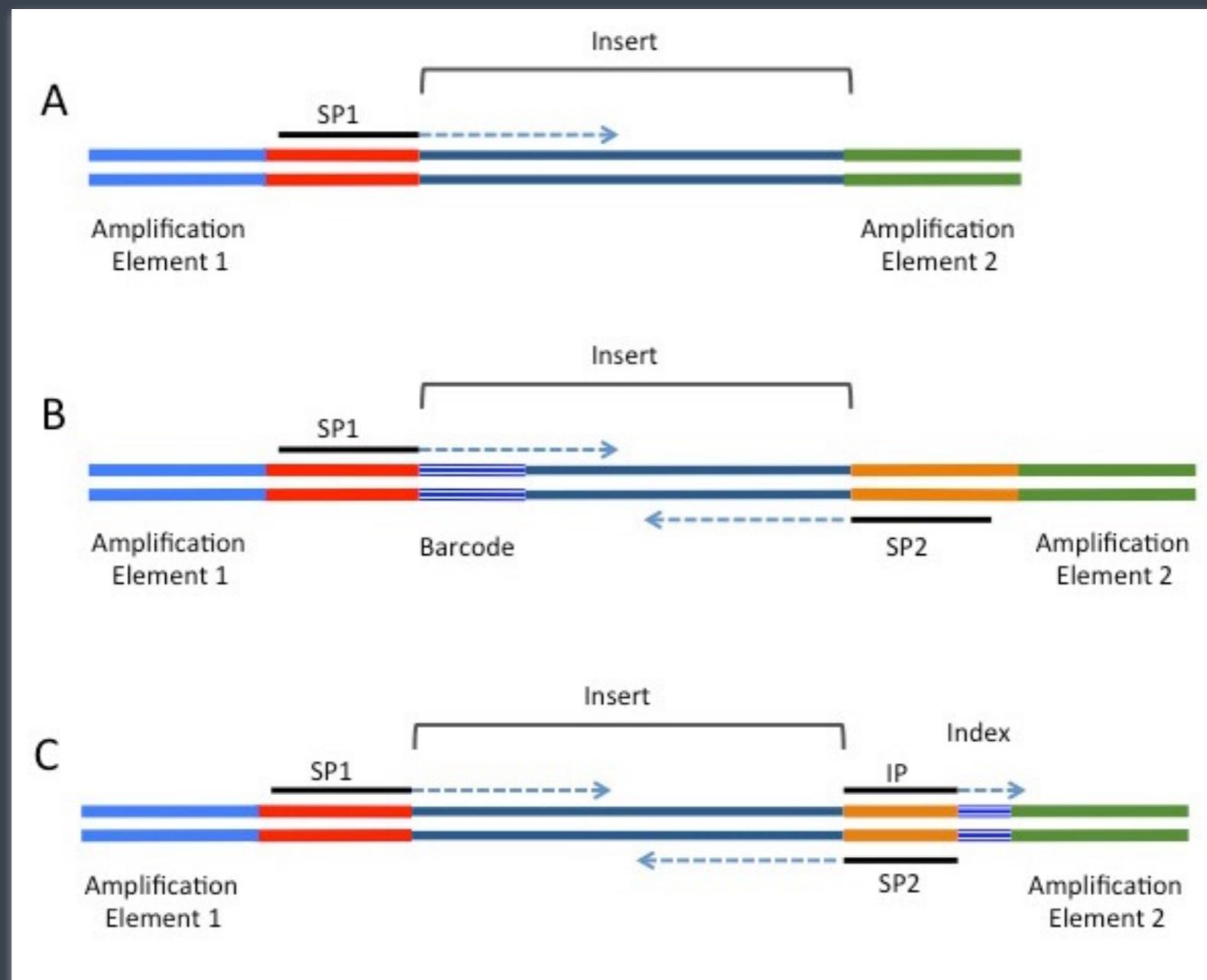
- ✓ Variant Detection
  - Whole genomes (HiSeq X, HiSeq, MiSeq, NextSeq)
  - Exomes (HiSeq, MiSeq, NextSeq)
- ✓ Expression analysis: Transcriptomes (HiSeq, MiSeq, NextSeq)
- ✓ Immunoprecipitation-based studies (HiSeq, NextSeq)
  - ChIP-Seq
  - CLIP-Seq
- ✓ Environmental samples
  - 16S rRNA-based metagenomics (MiSeq)
  - Whole metagenomes and metatranscriptomes (HiSeq, MiSeq, NextSeq)
- ✓ many more....

## Modifications/Strategies for application-specific libraries

Refer to <http://www.illumina.com/applications/sequencing/ngs-library-prep/library-prep-methods.html> for a more complete list

# Overview

- ✓ Illumina's Sequencing systems
- ✓ Standard library prep (Tru-Seq)
- ✓ Tagmentation-based approach (Nextera)
  - Long-insert library prep (mate pairs)
- ✓ Modifications/Strategies for application-specific libraries
- ✓ Multiplexing



# Multiplexing (with barcodes and indices)

sample1

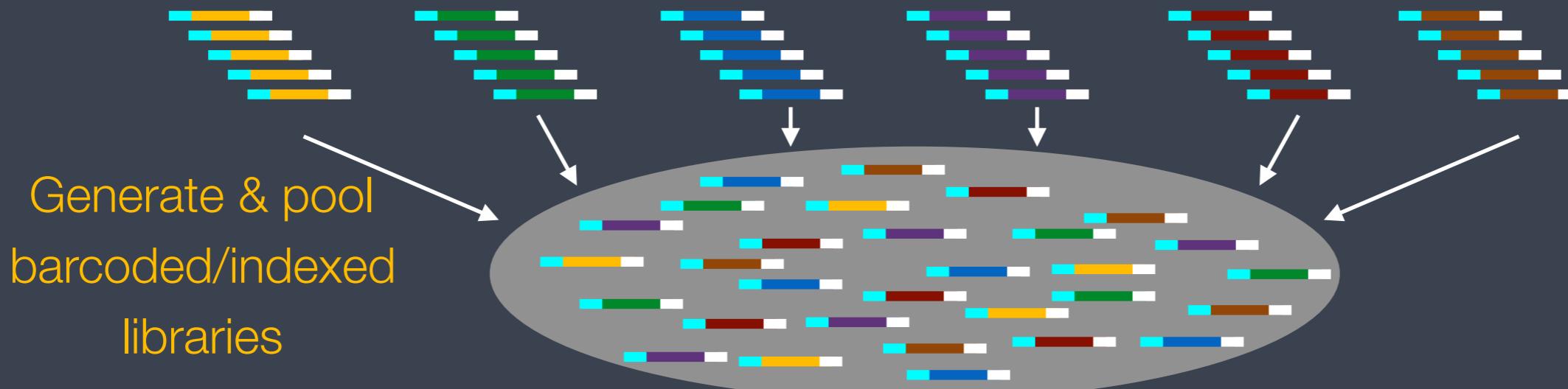
sample2

sample3

sample4

sample5

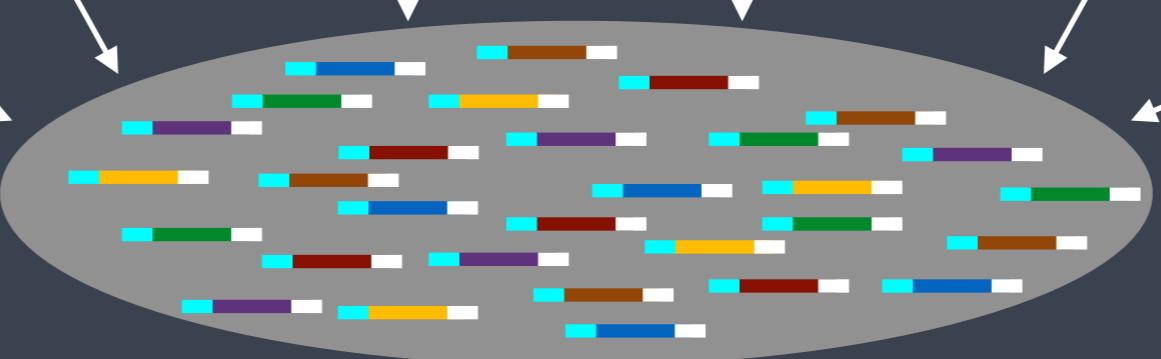
sample6



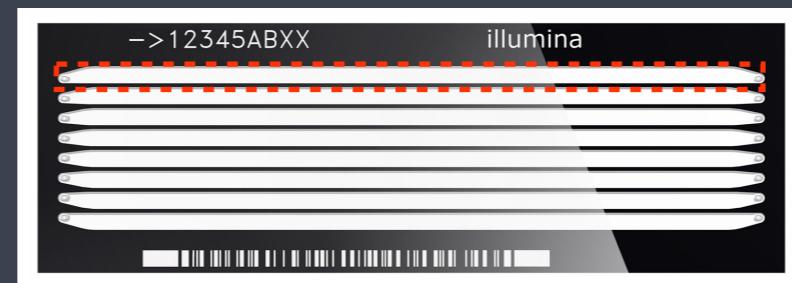
sample1 sample2 sample3 sample4 sample5 sample6

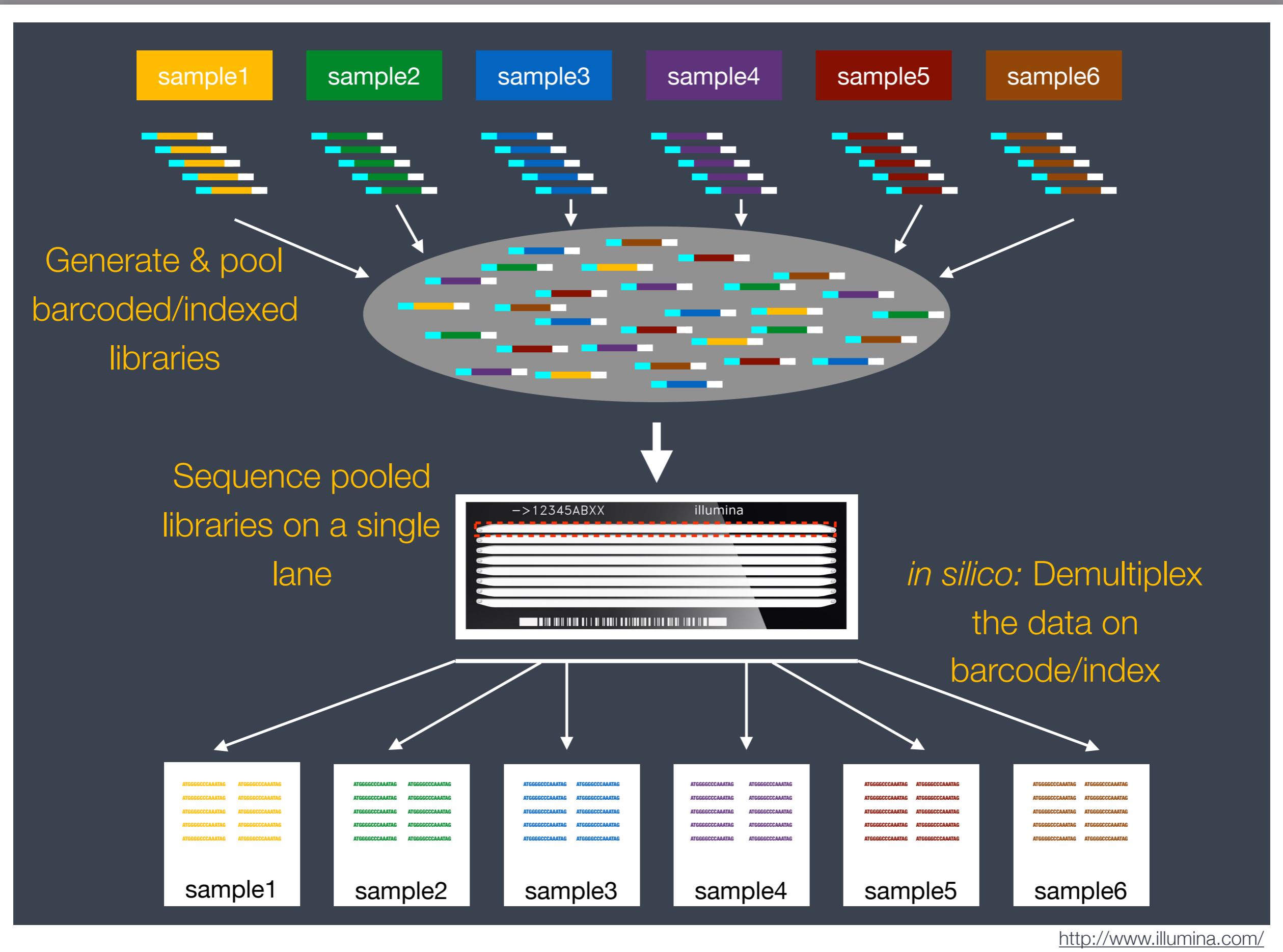


Generate & pool  
barcoded/indexed  
libraries



Sequence pooled  
libraries on a single  
lane





- <http://support.illumina.com/content/illumina-support/us/en/sequencing/literature.html>
- <http://support.illumina.com/content/dam/illumina-marketing/documents/products/other/ngs-primer-genetic-disease-cell-biology-1070-2014-006.pdf>
- [https://www.illumina.com/content/dam/illumina-marketing/documents/products/research\\_reviews/sequencing-methods-review.pdf](https://www.illumina.com/content/dam/illumina-marketing/documents/products/research_reviews/sequencing-methods-review.pdf)
- <https://www.illumina.com/content/dam/illumina-marketing/documents/applications/ngs-library-prep/ForAllYouSeqMethods.pdf>

Resources at Illumina

*These materials have been developed by members of the teaching team at the Harvard Chan Bioinformatics Core (HBC). These are open access materials distributed under the terms of the Creative Commons Attribution license (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*

