

A photograph of three scientists in white lab coats standing behind a large, stylized graphic of a DNA sequence. The sequence is composed of four colored strands (yellow, green, red, blue) forming a wavy pattern. The scientists are positioned to the left of the graphic.

# Illumina sequencing technology and Sample Preparation Solutions

**Tiphaine Godefroy**  
Marketing Technical  
Specialist  
France

11-03-2014

# NGS Vocabulary

- ▶ Librairie : fragment d'acide nucléique qui contient des adaptateurs sur chacune de ses extrémités
- ▶ Flow Cell : support en verre et lieu de l'amplification clonale (formation des clusters) et du séquençage (1 ou 8 pistes, lanes ou canaux par FC)
- ▶ Cluster : 1000 molécules d'ADN (fragments de libraire). Read
- ▶ Sequencing by Synthesis : addition d'un dNTP fluorescent terminateur de chaîne réversible (1 cycle = 1 base)
- ▶ Single Read (SR) : Lecture d'une extrémité du fragment
- ▶ Paired-End (PE) : lecture des deux extrémités des fragments
- ▶ Index, Tag, Barcode : petite séquence de 6 ou 8 bases permettant de multiplexer plusieurs échantillons lors du séquençage

# Definitions-Quality Scores

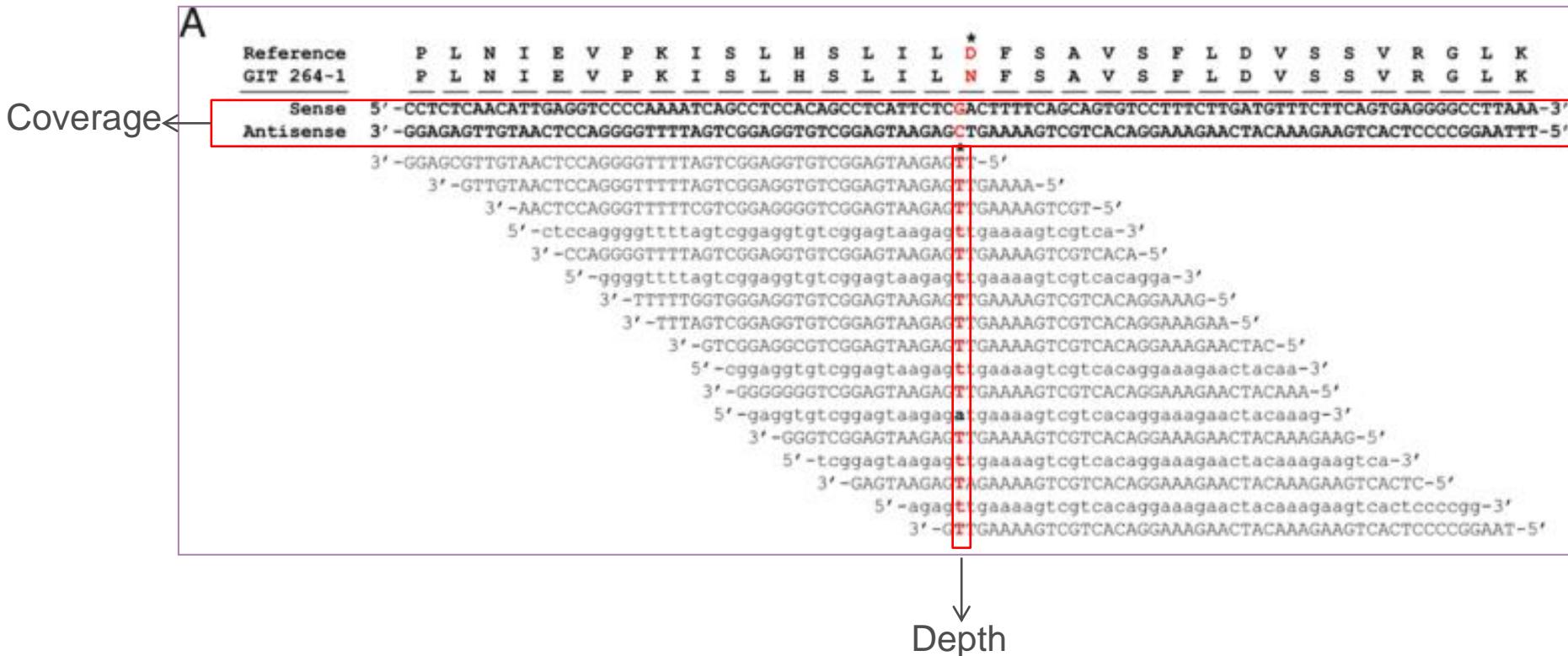
## What does a Q score mean?

- ▶ A single base call is assigned a quality score. It is an odds ratio, the probability of an incorrect call.
  - For example, a base with a Q40 score has a probability of 1 in 10,000 of an incorrect base call.

Phred Quality Score	Probability of Incorrect Base Call	Base Call Accuracy
10	1 in 10	90%
20	1 in 100	99%
30	1 in 1000	99.9%
40	1 in 10000	99.99%

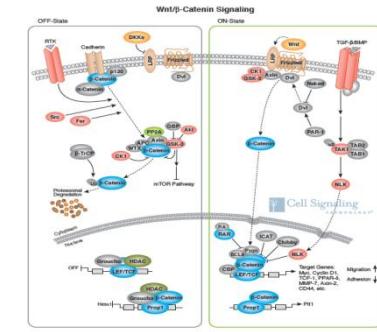
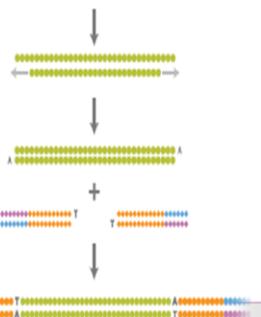
>90% bases higher than Q30 at 1x35 bp  
>>90% bases higher than Q30 at 2x25 bp  
>85% bases higher than Q30 at 2x100 bp  
>80% bases higher than Q30 at 2x150 bp  
>70% bases higher than Q30 at 2x250 bp

# Difference between Depth and Coverage



- **Coverage** : zone du génome couverte lors du séquençage
  - **Profondeur** : nombre de fois que chaque base sera séquencée

# The Illumina Workflow



Sample Preparation

Clustering

Sequencing

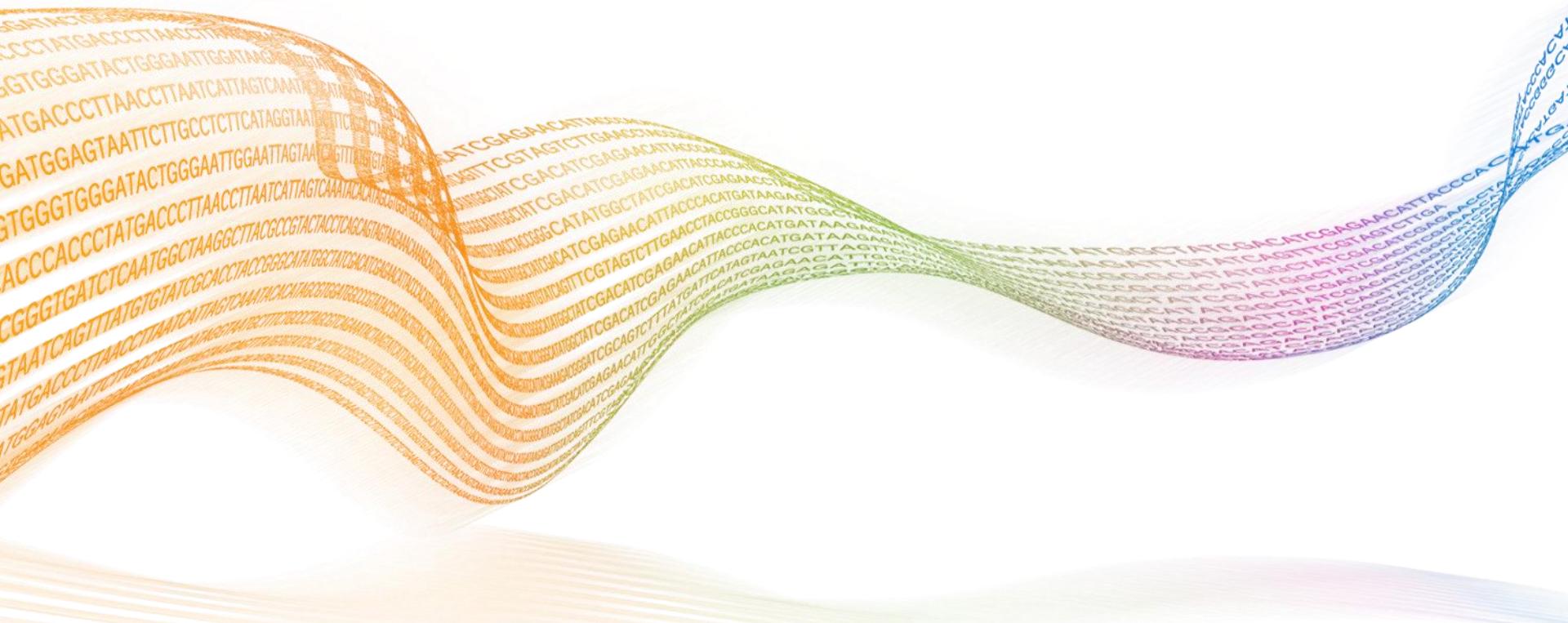
Analysis

Biological Meaning

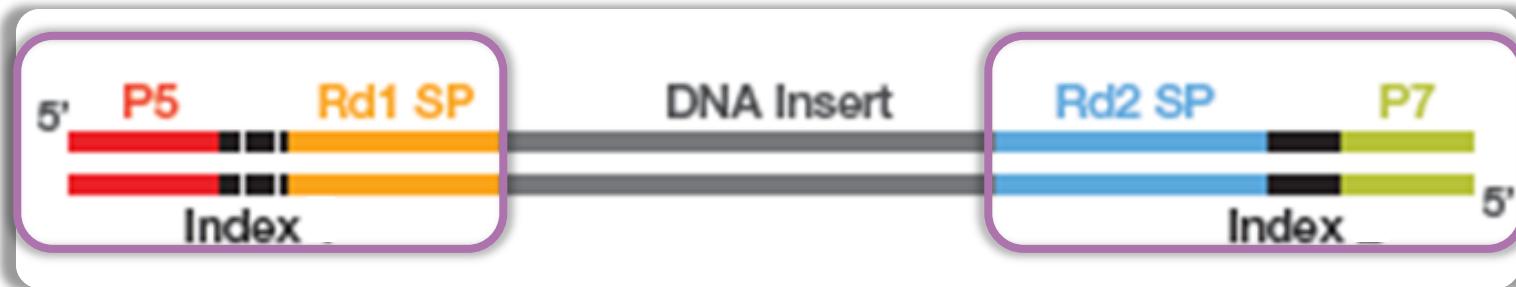
TruSeq Sample Prep Reagents/  
Nextera

TruSeq Chemistry

# Library preparation



# Sample Prep is Critical for Successful Sequencing



Dual Index Library shown

The aim of the sample prep step is to obtain nucleic acid fragments with adapters attached on both ends

# Sample Prep Options

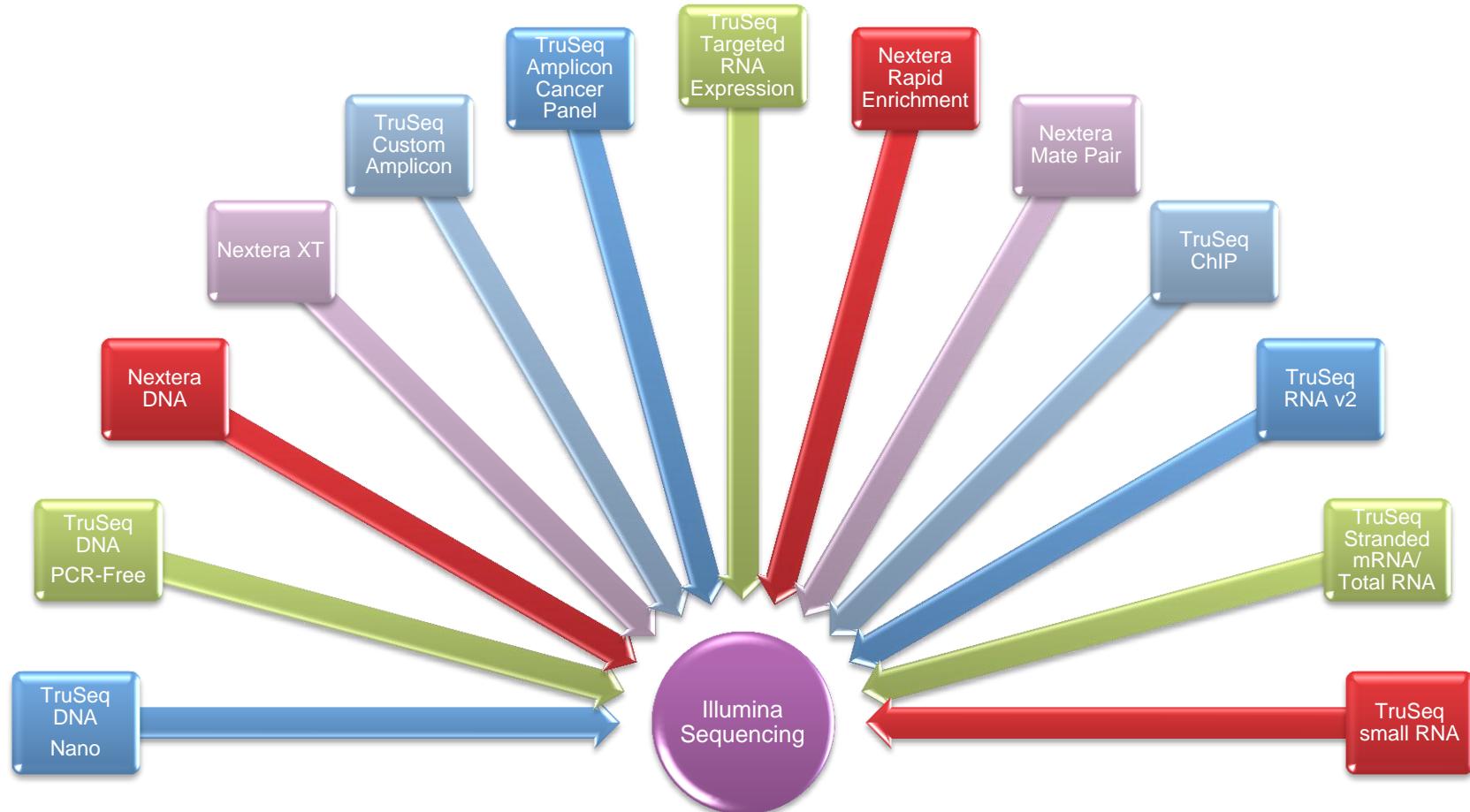
DNA Sample Prep  
Options

Amplicon and Enrichment  
Options for Targeted  
Resequencing

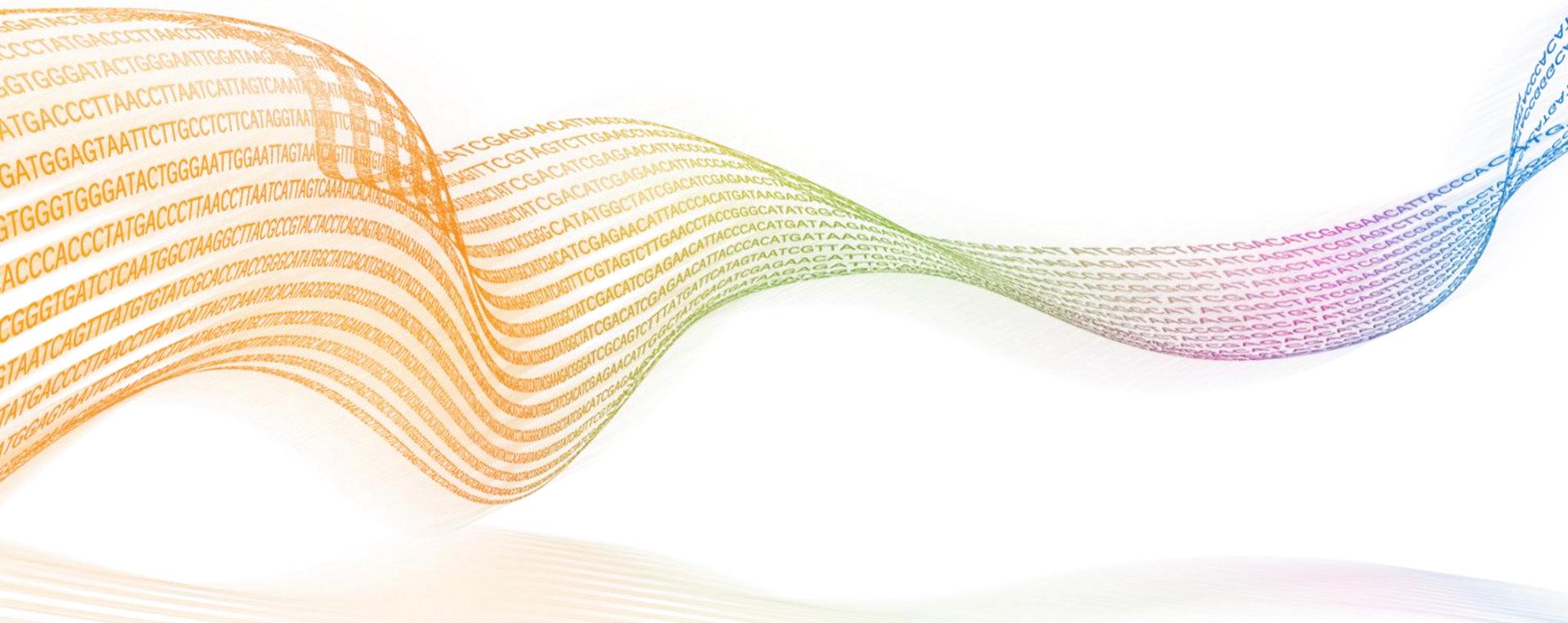
RNA Sample Prep  
Options

# Selecting the Right Kit

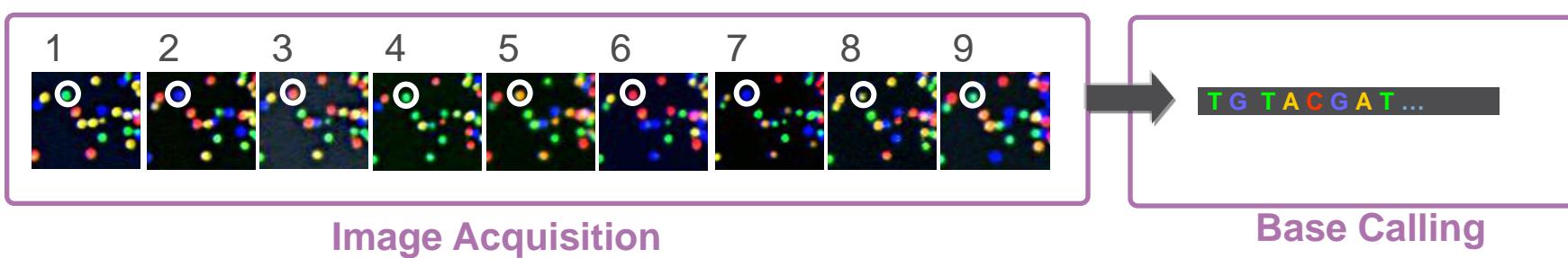
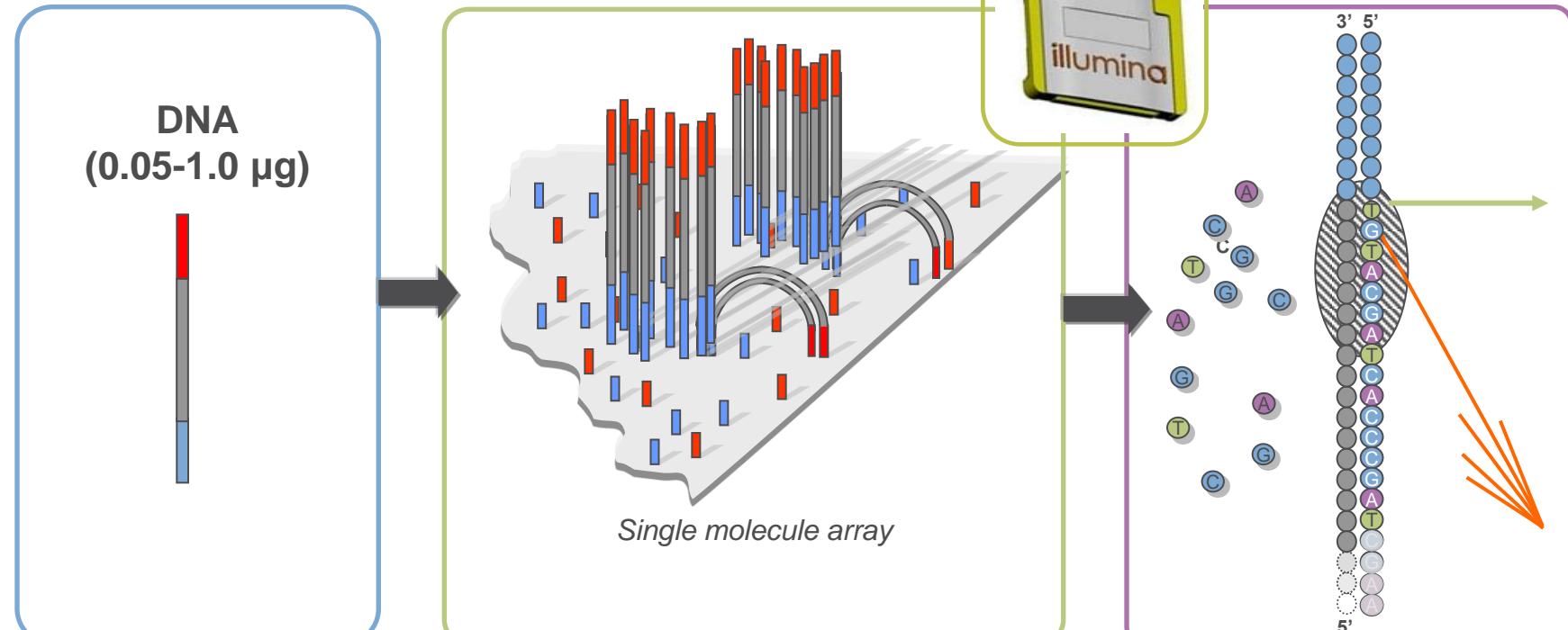
Go to: <http://support.illumina.com/sequencing-sample-prep-kit-selector.ilmn>



# Cluster Generation and Sequencing steps



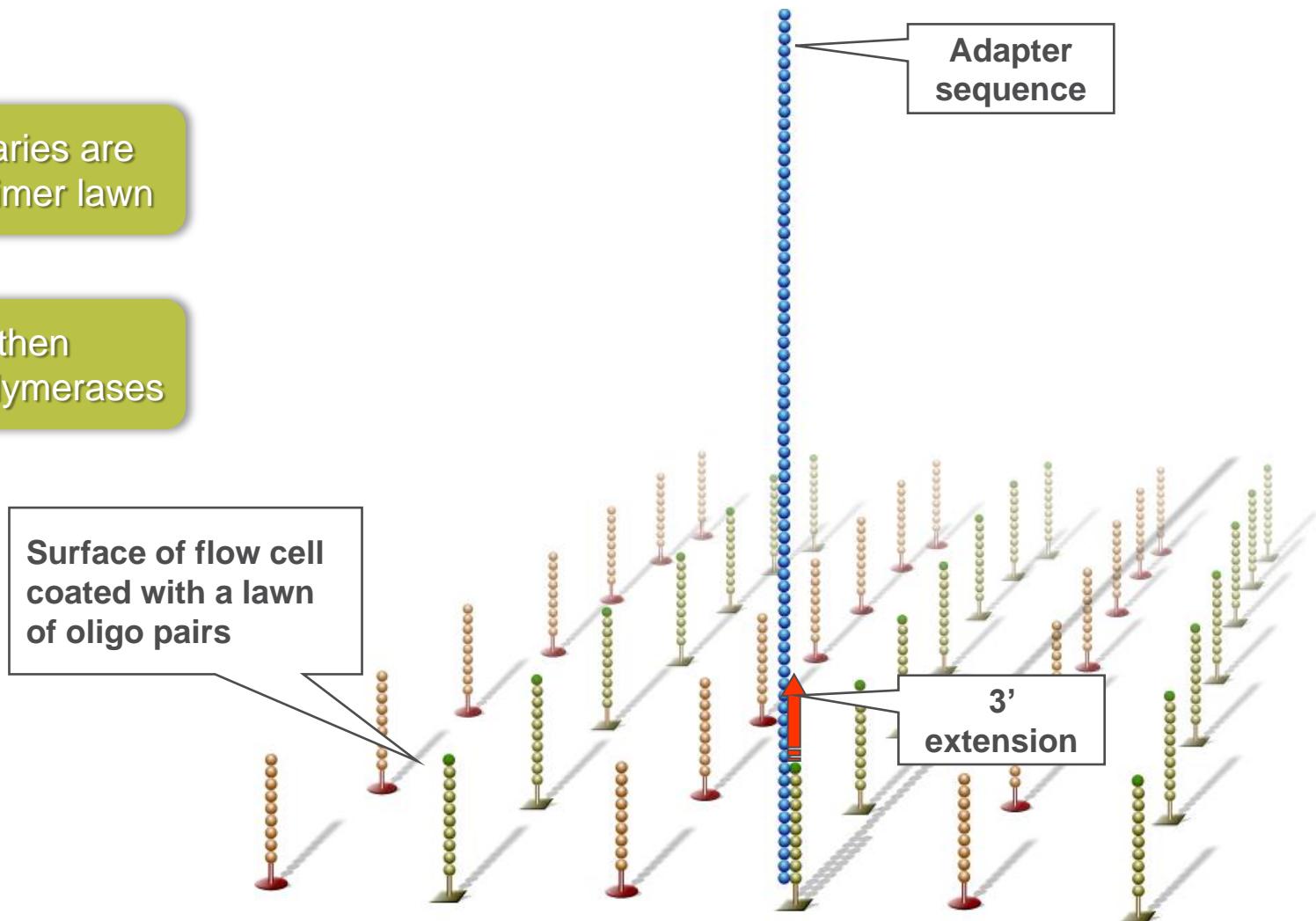
# Illumina Sequencing Workflow



# Hybridize Fragment & Extend

Single DNA libraries are hybridized to primer lawn

Bound libraries then extended by polymerases

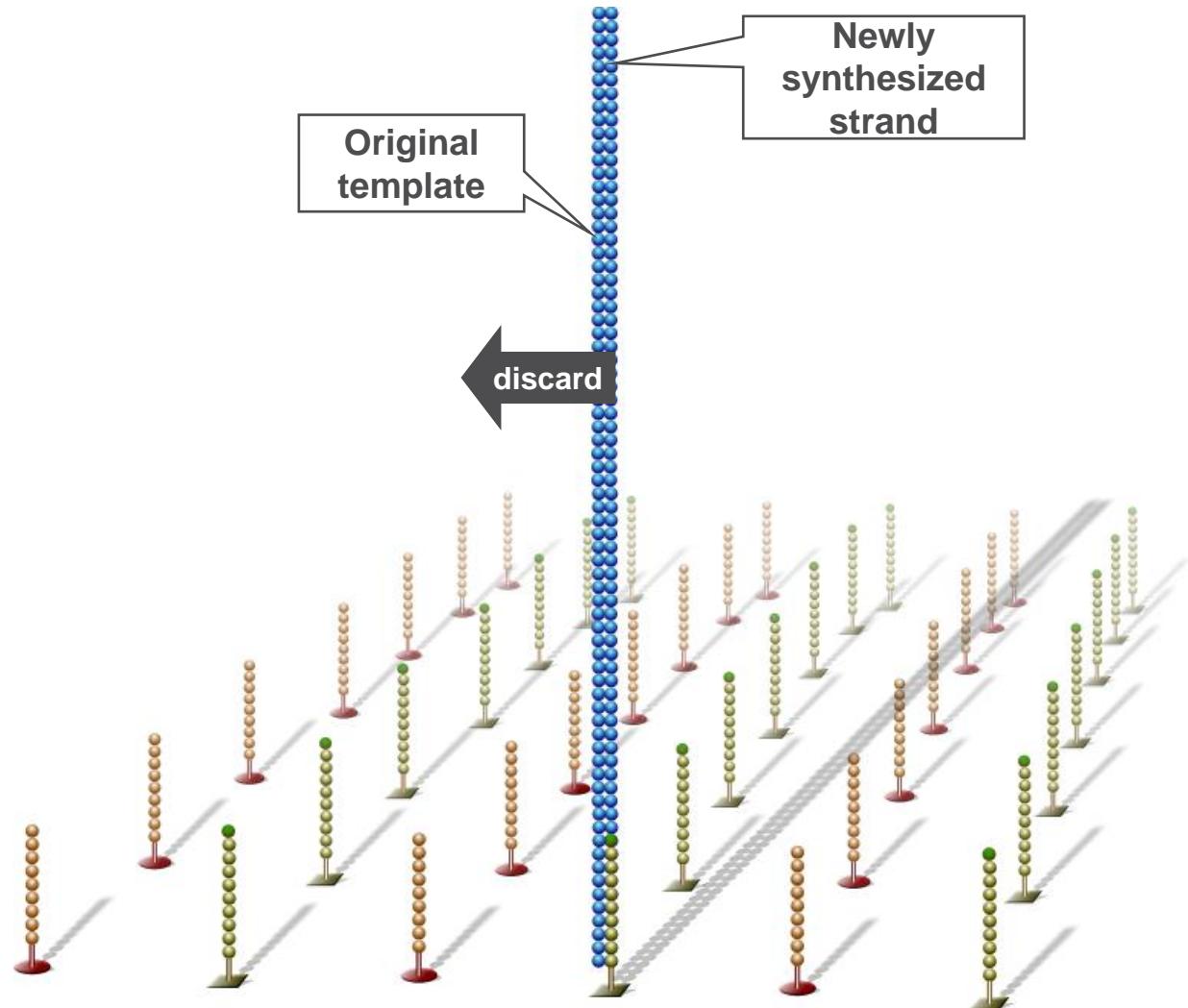


# Denature Double-Stranded DNA

Double-stranded  
molecule is denatured

Original template washed  
away

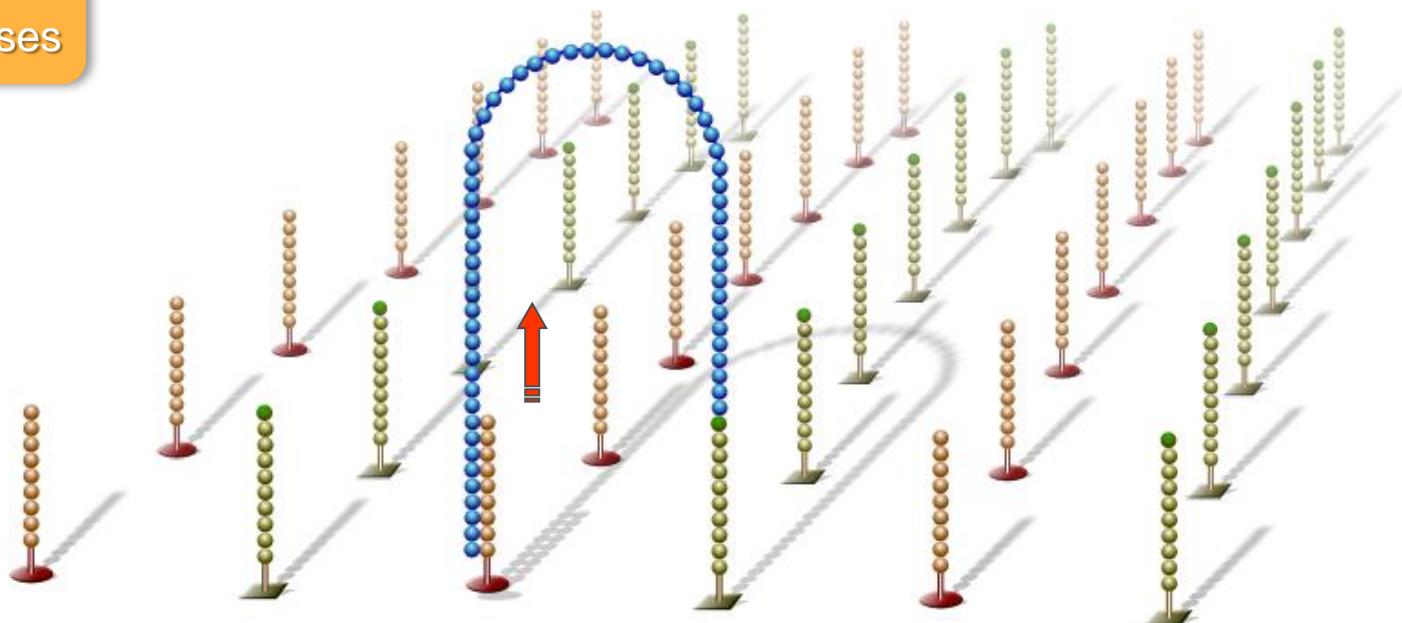
Newly synthesized strand  
is covalently attached to  
flow cell surface



# Bridge Amplification

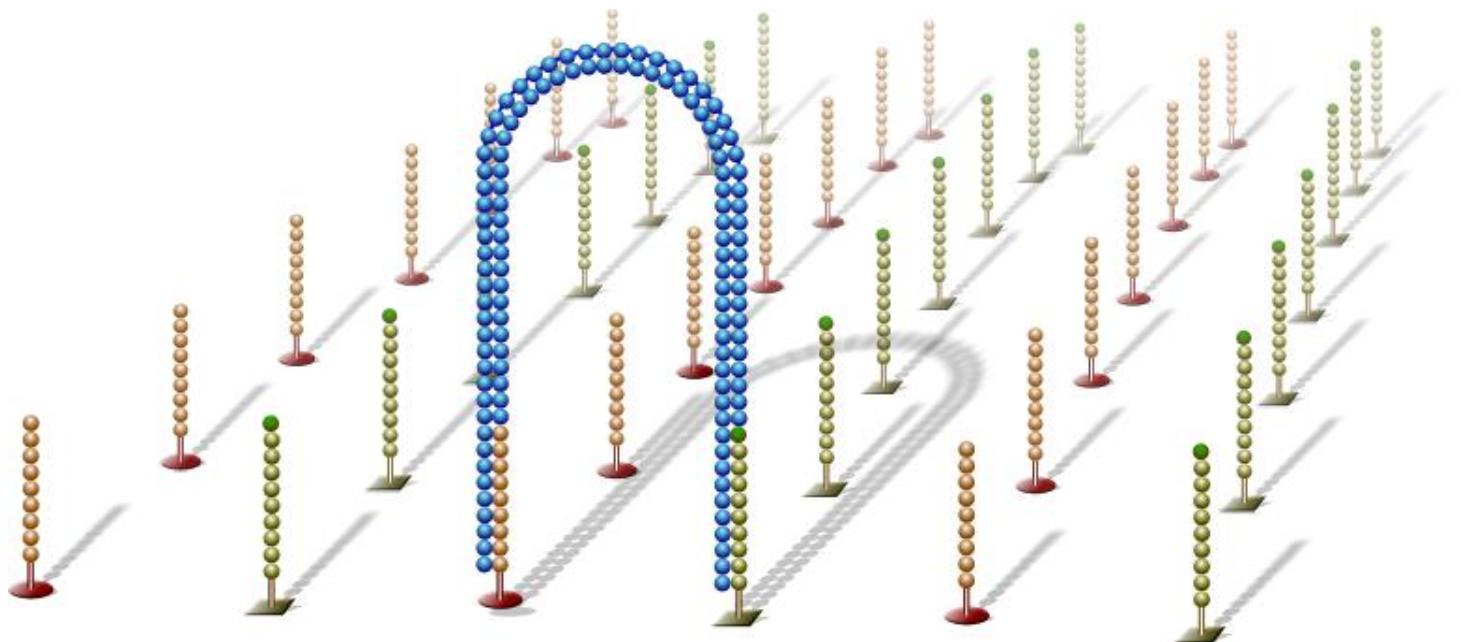
Single-stranded molecule flips over and forms a bridge by hybridizing to adjacent, complementary primer

Hybridized primer is extended by polymerases



# Bridge Amplification

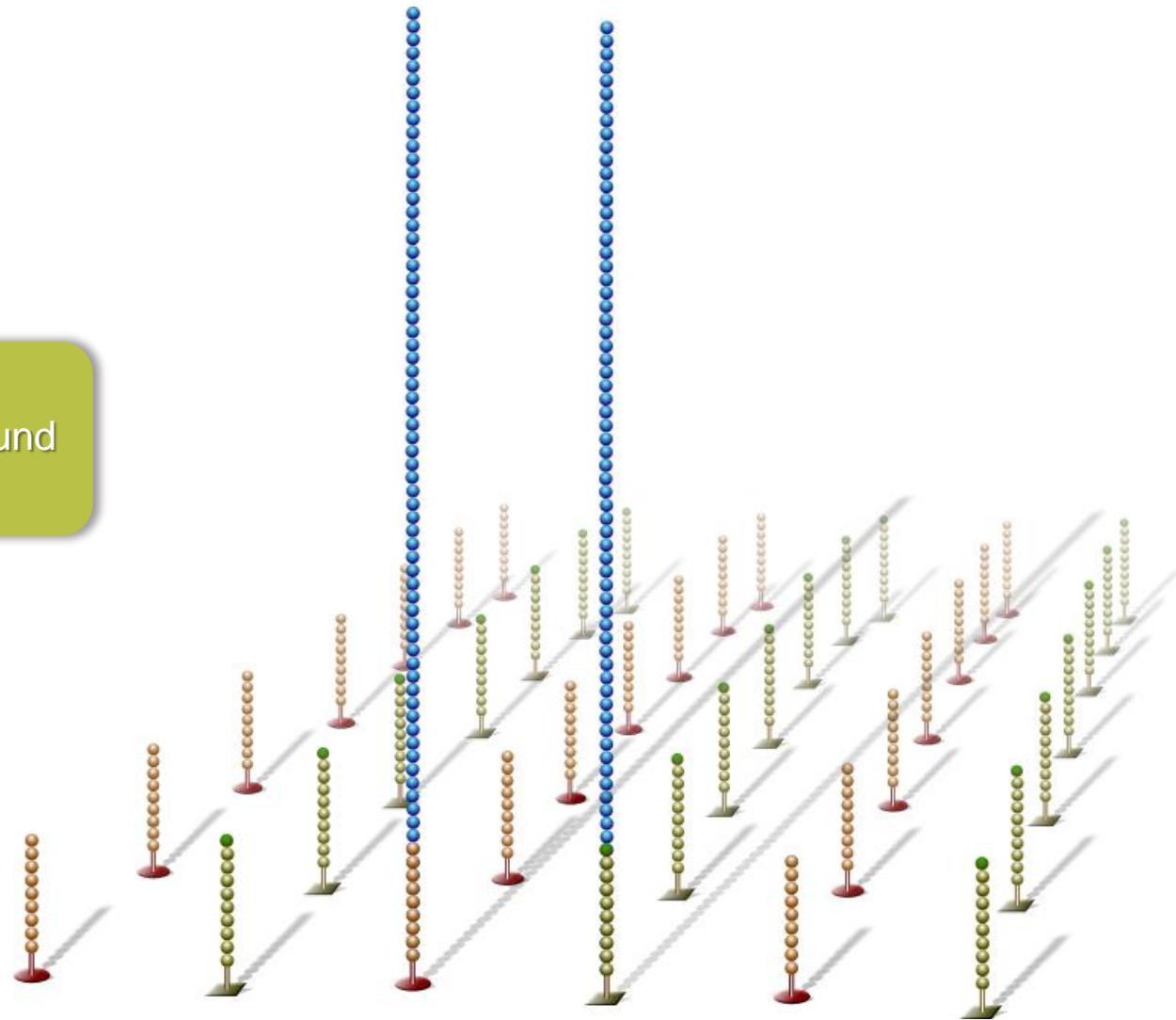
Double-stranded bridge is formed



# Denature Double-Stranded Bridge

Double-stranded bridge is denatured

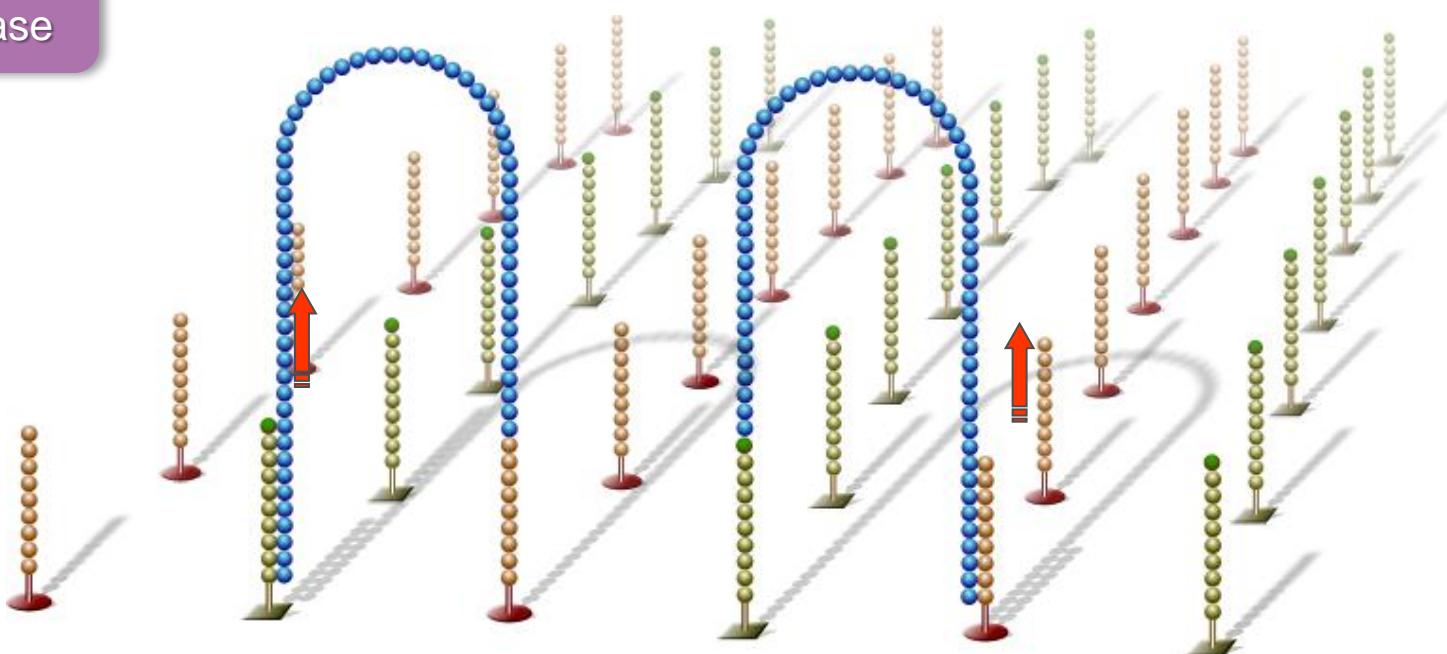
Result:  
Two copies of covalently bound single-stranded templates



# Bridge Amplification

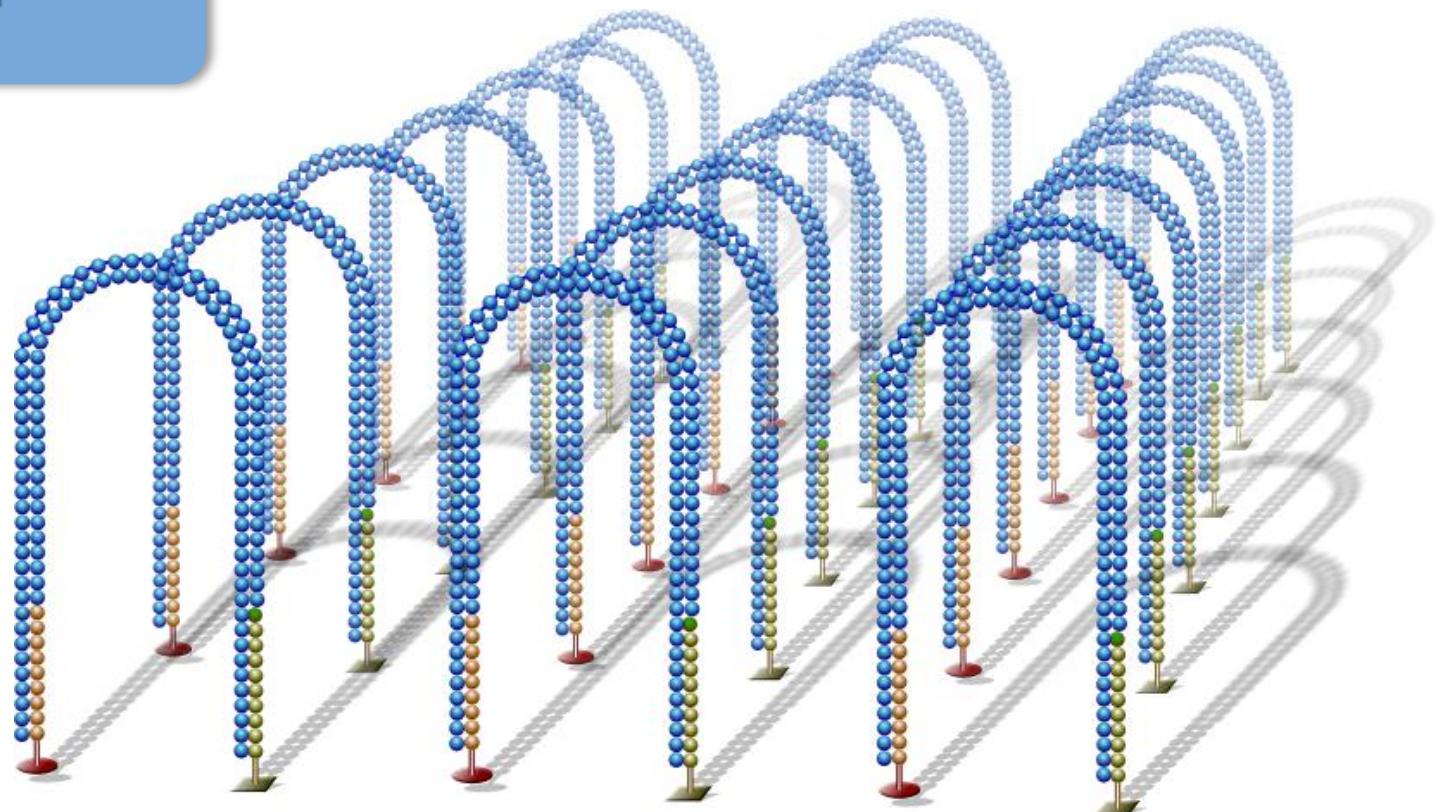
Single-stranded molecules flip over  
to hybridize to adjacent primers

Hybridized primer is  
extended by polymerase

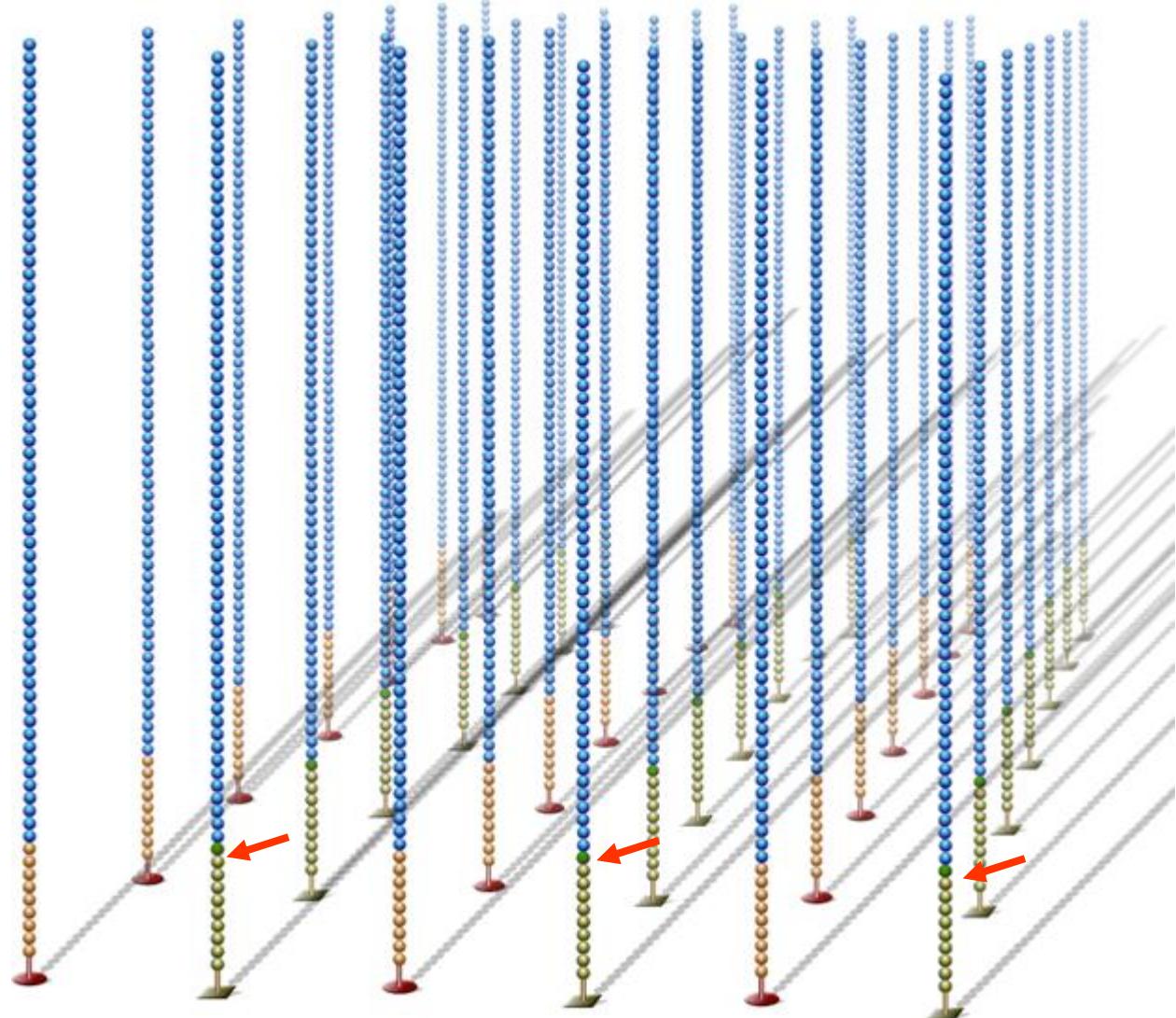


# Bridge Amplification

Bridge amplification cycle  
repeated until multiple  
bridges are formed

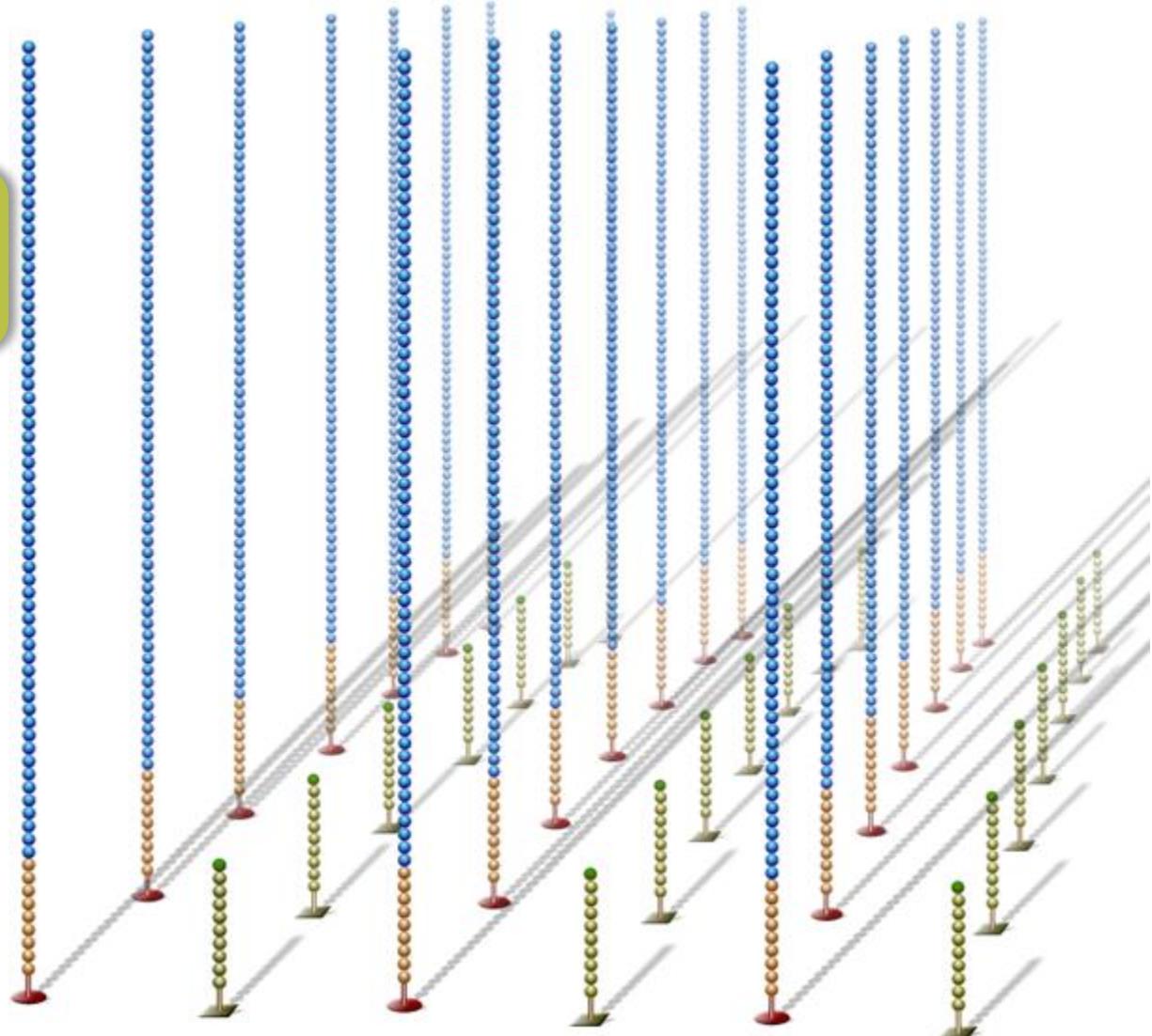


# Linearization



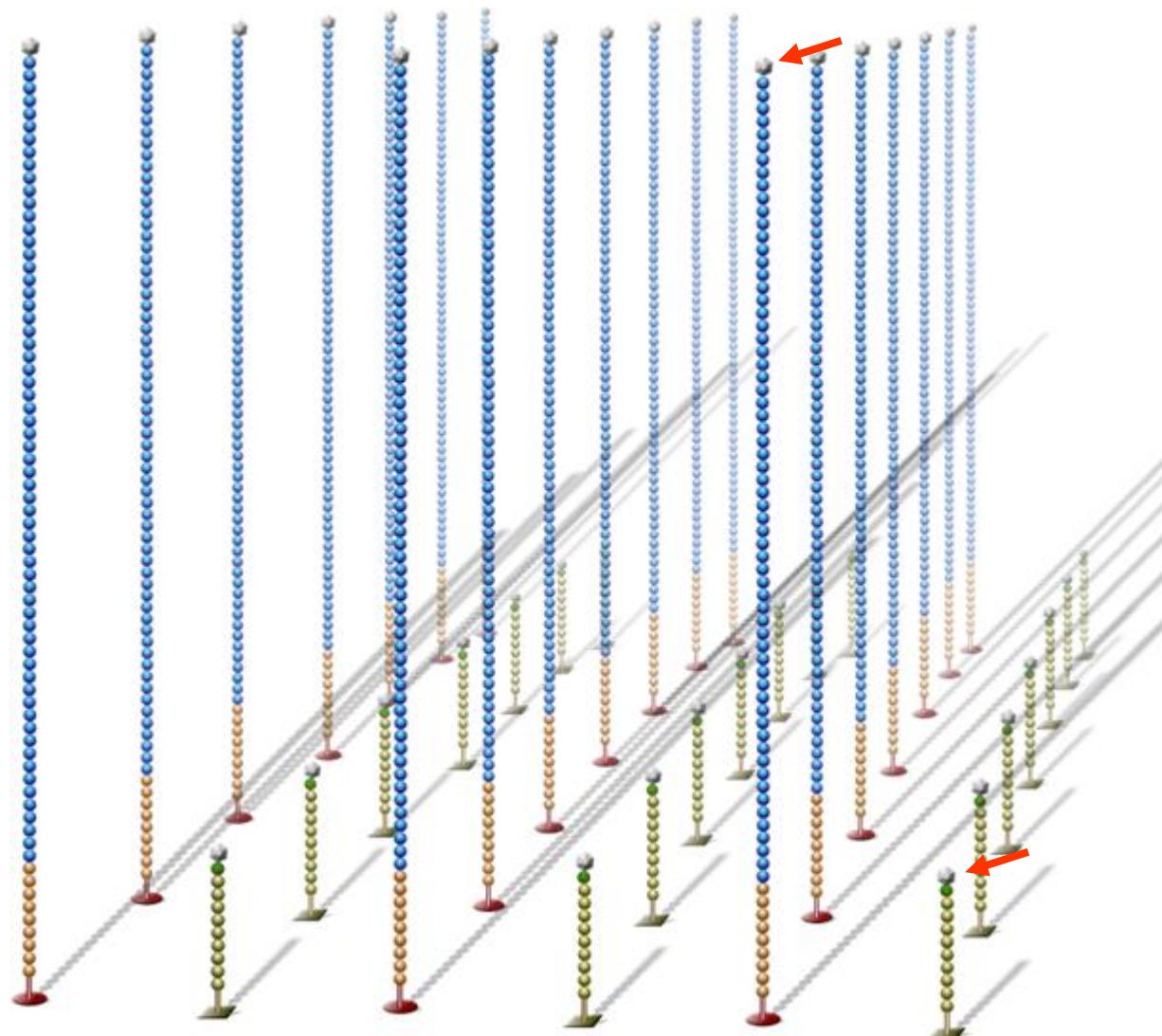
# Reverse Strand Cleavage

Reverse strands cleaved and washed away, leaving a cluster with forward strands only



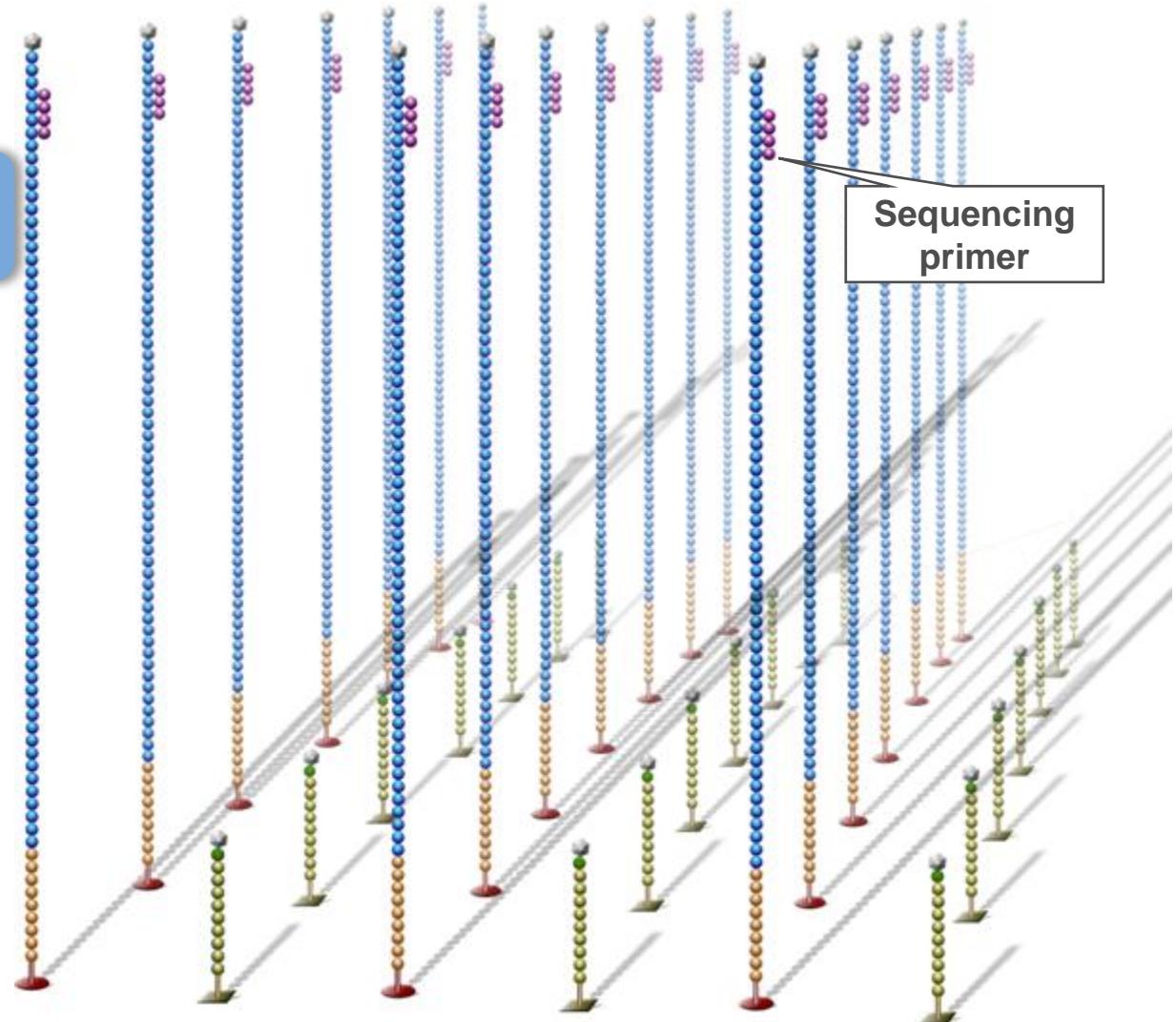
# Blocking

Free 3' ends are blocked to prevent unwanted DNA priming

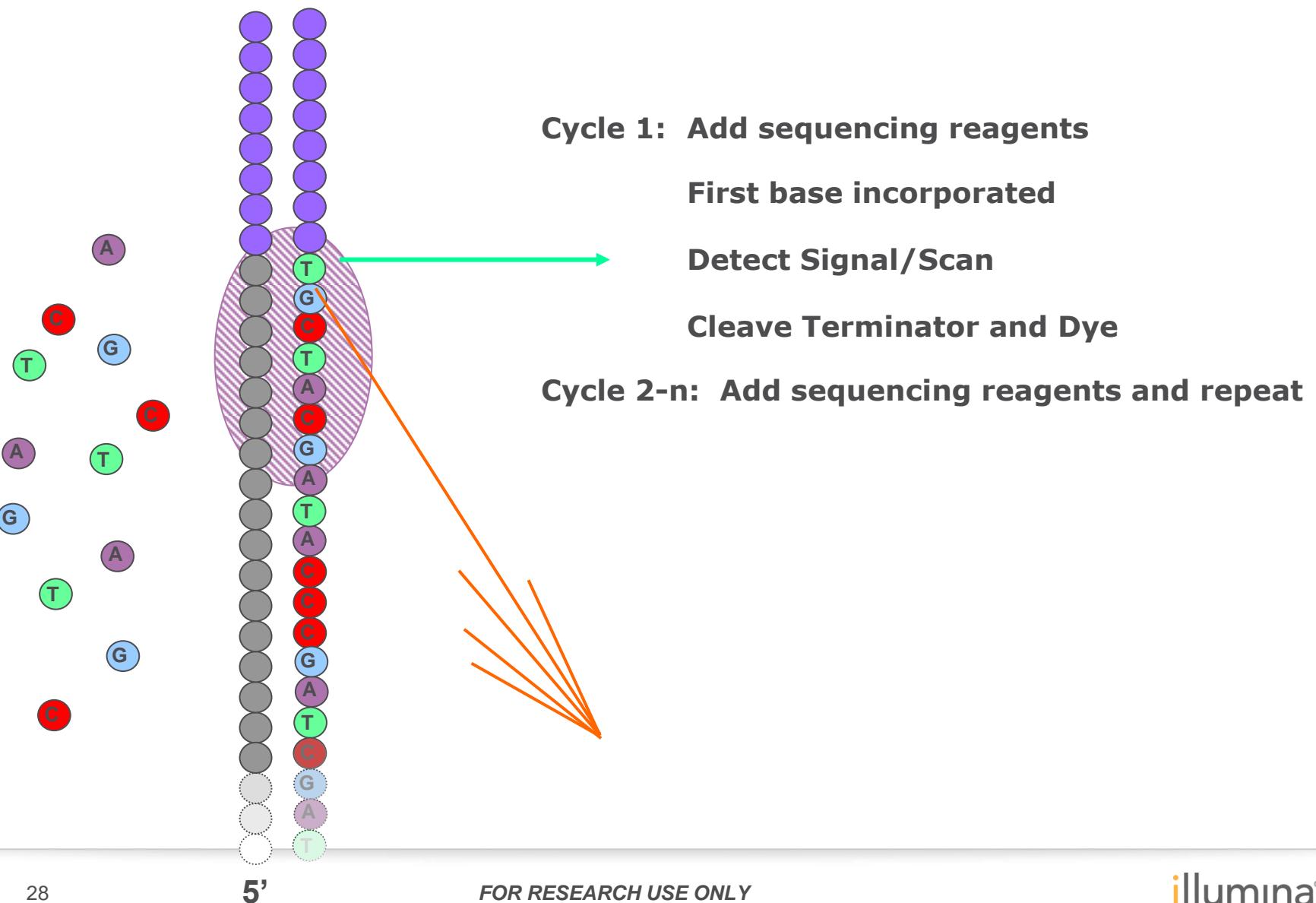


# Read 1 Primer Hybridization

Sequencing primer is hybridized  
to adapter sequence



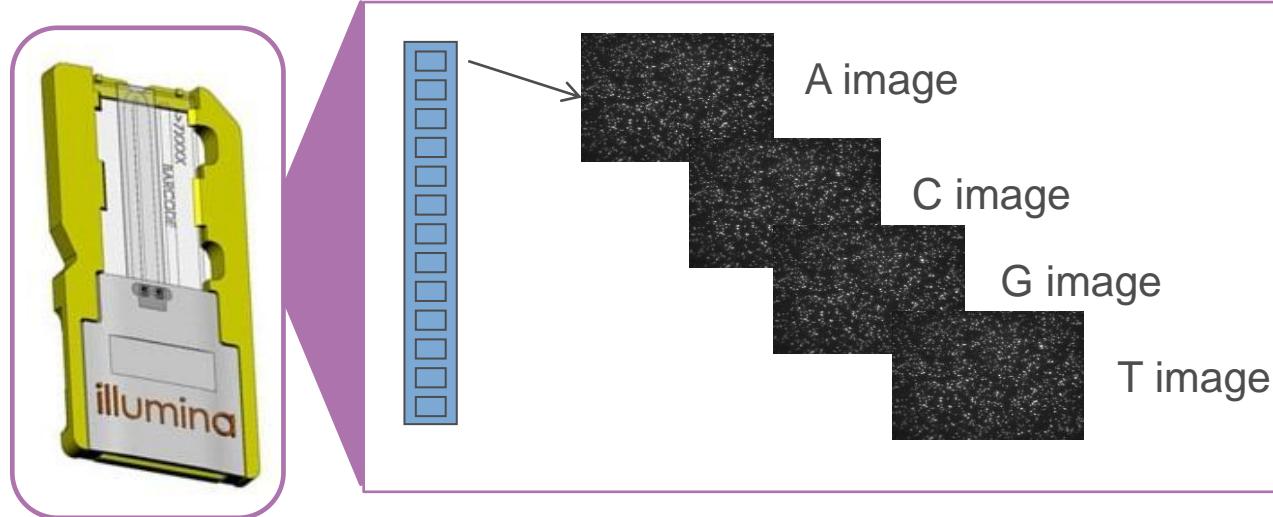
# Sequencing By Synthesis (SBS)



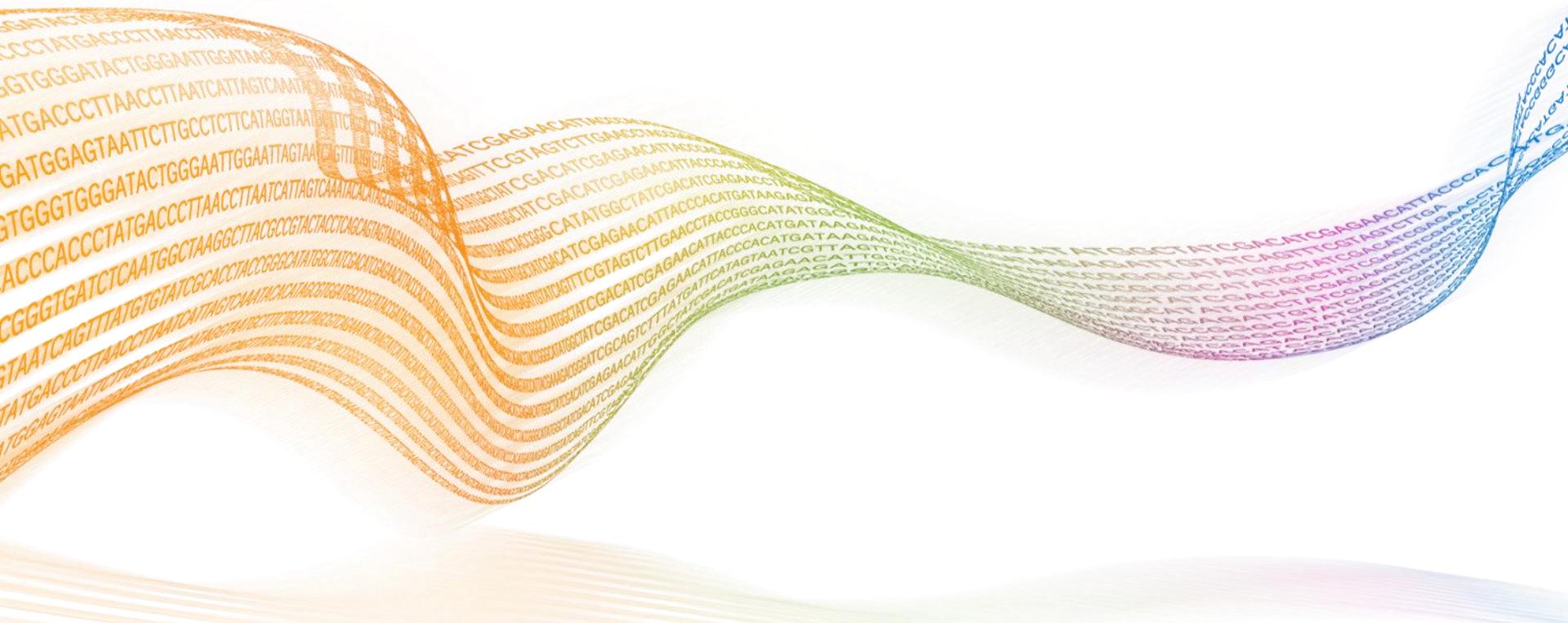
# Sequencing

Clusters are images using LED and filter combinations specific for each fluorescently-labeled nucleotide

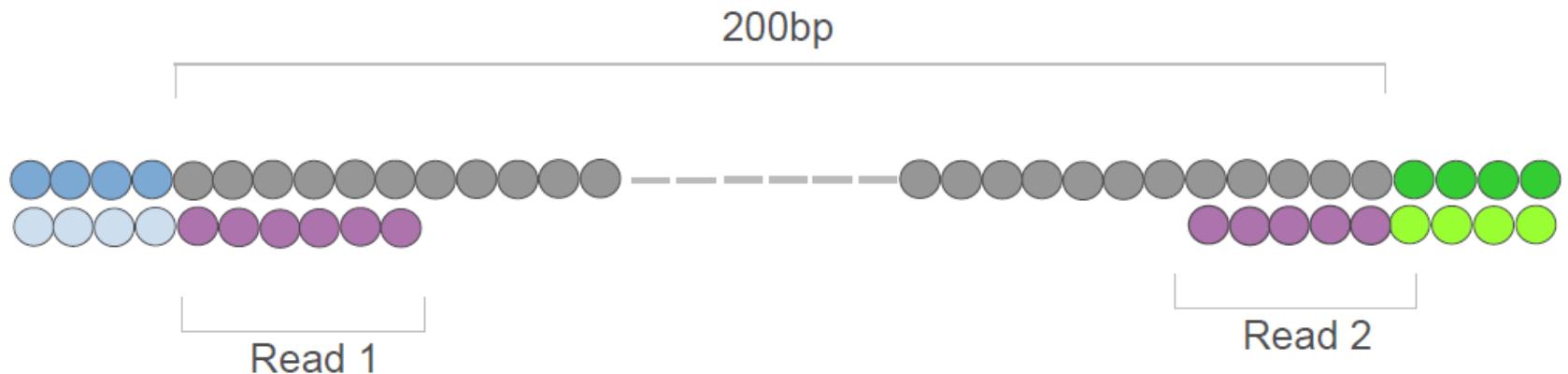
After imaging is complete for one section (tile), the flow cell is moved to the next tile and the process is repeated



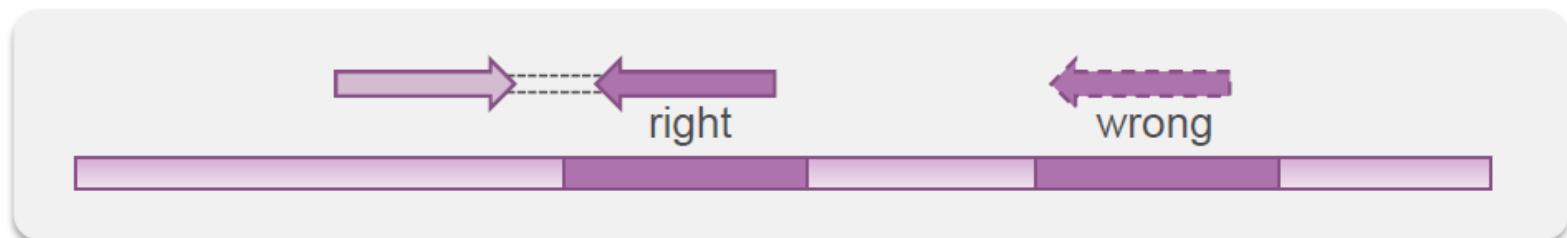
# Paired End Sequencing



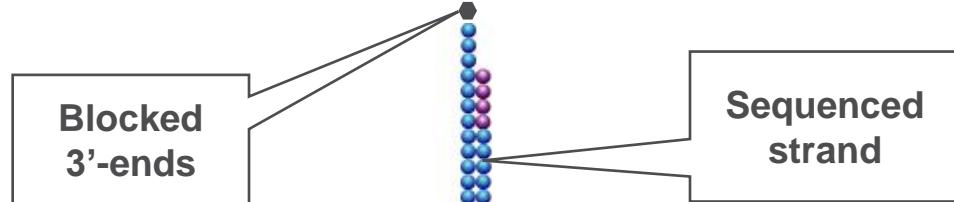
# Paired End Sequencing



- ▶ Paired End sequencing allow sequencing of both extremities of DNA fragment
- ▶ It provides more information than single reads, where just one extremity is sequenced

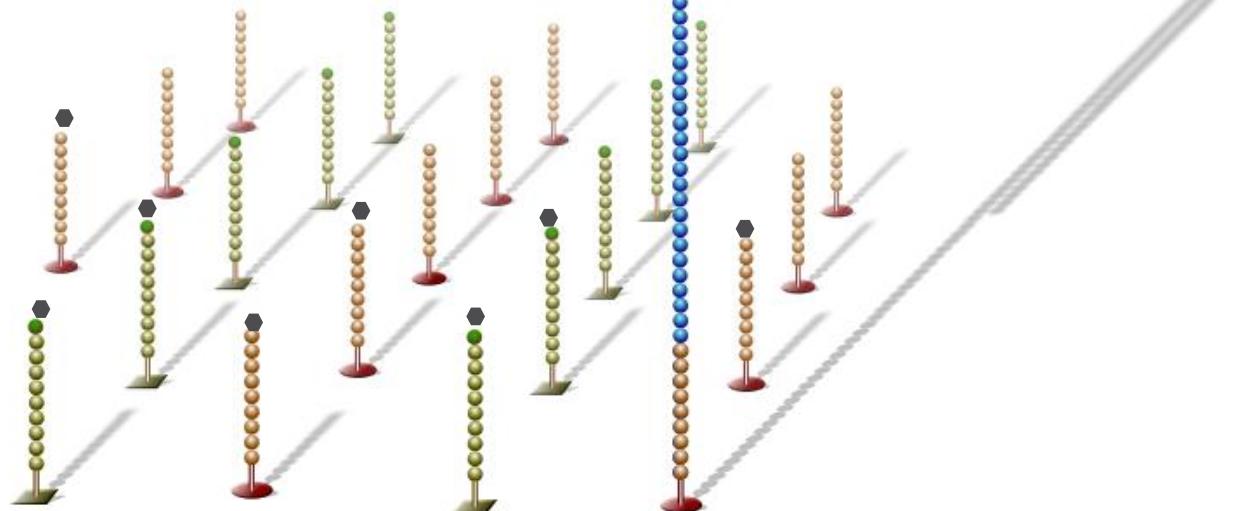


# Paired End Sequencing



Sequenced strand is stripped off

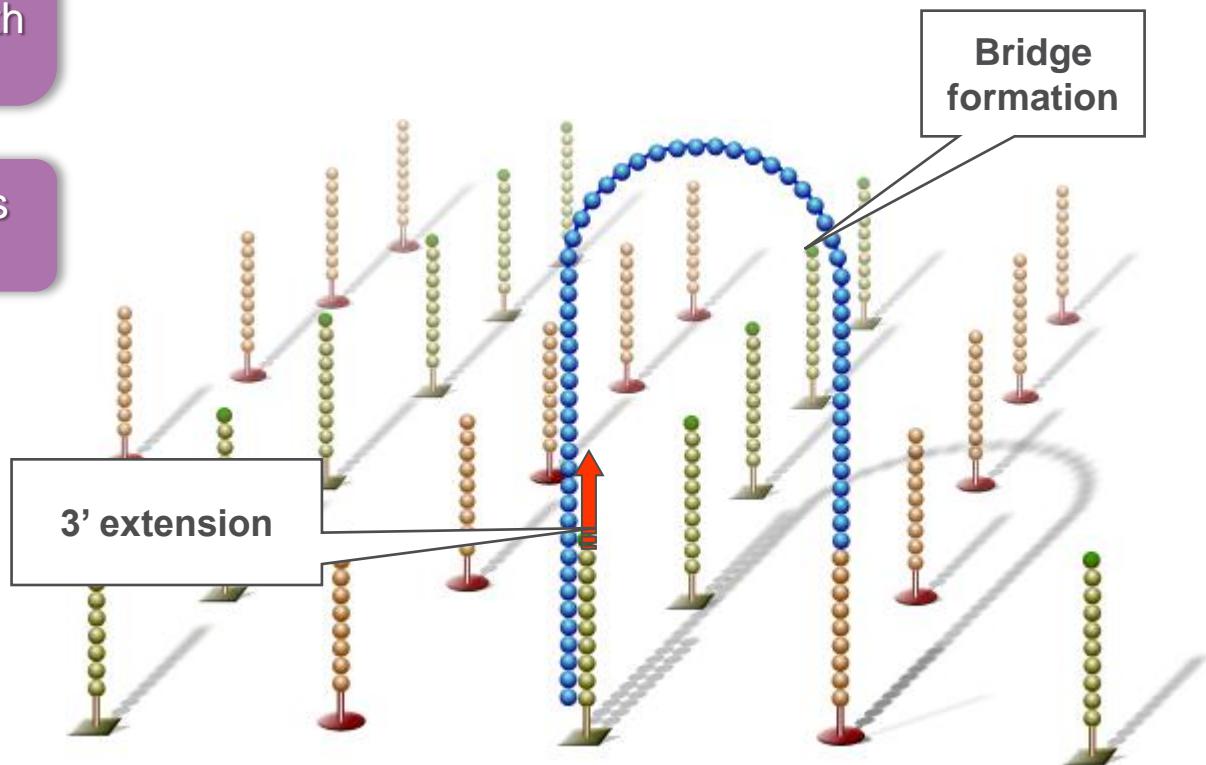
3'-ends of template strands and lawn primers are unblocked



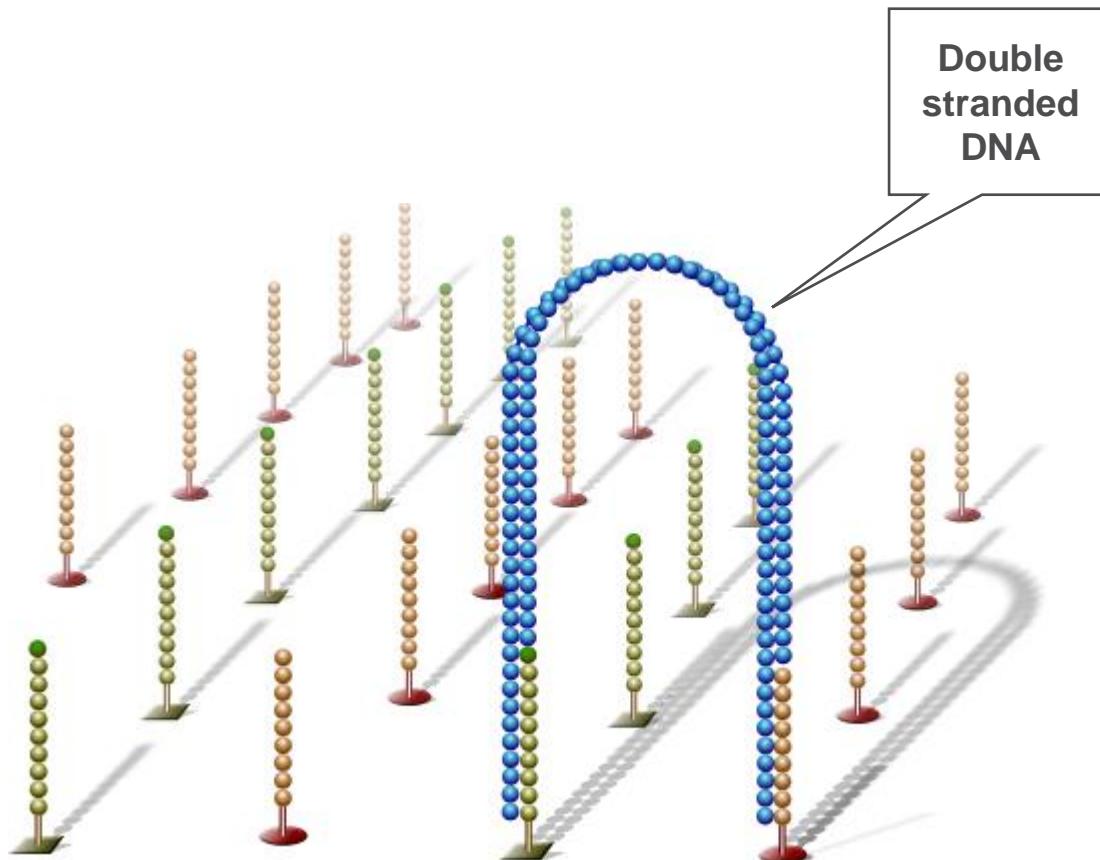
# Paired End Sequencing

Single-stranded template loops over to form a bridge by hybridizing with a lawn primer

3'-ends of lawn primer is extended

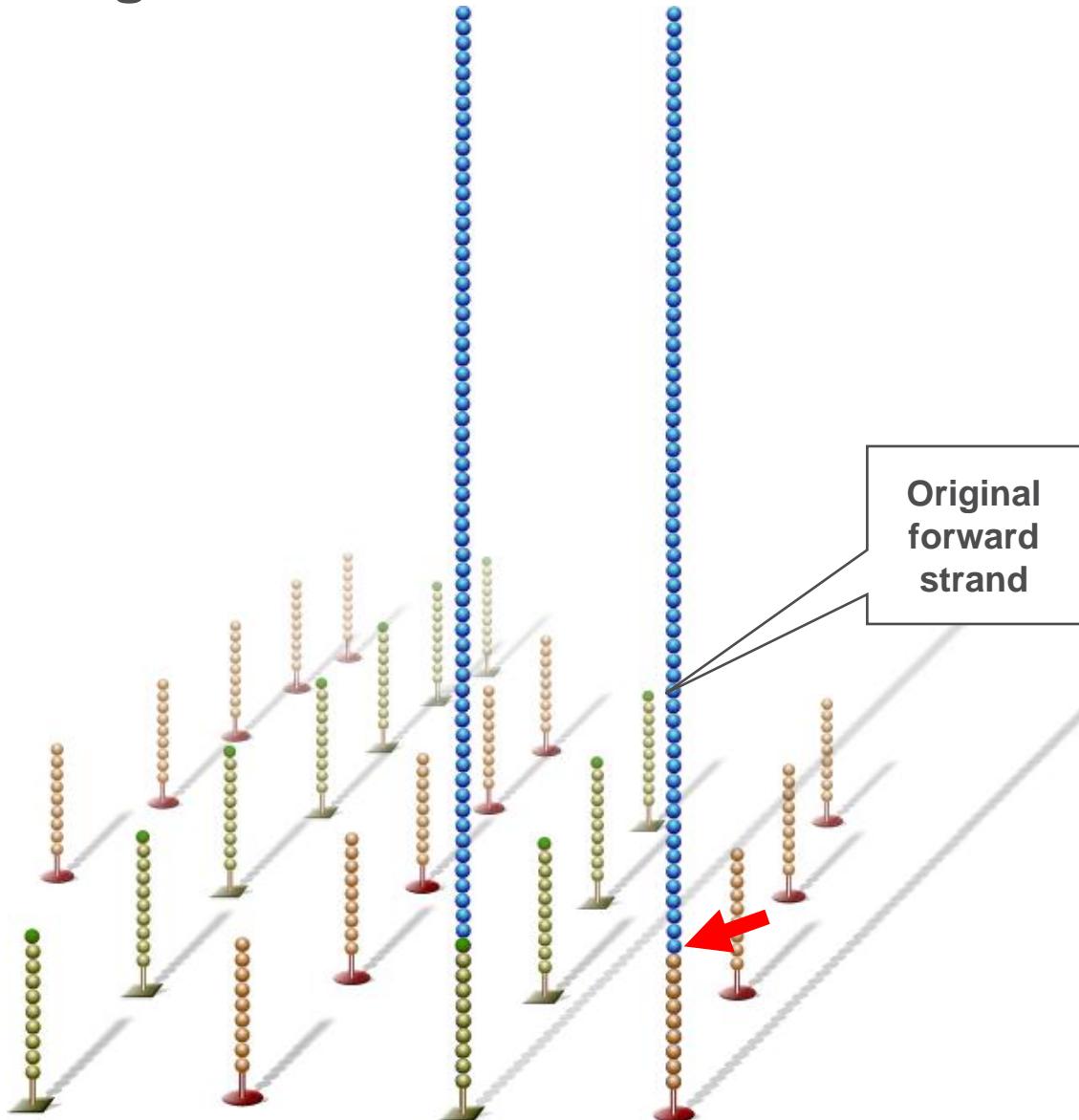


# Paired End Sequencing



# Paired End Sequencing

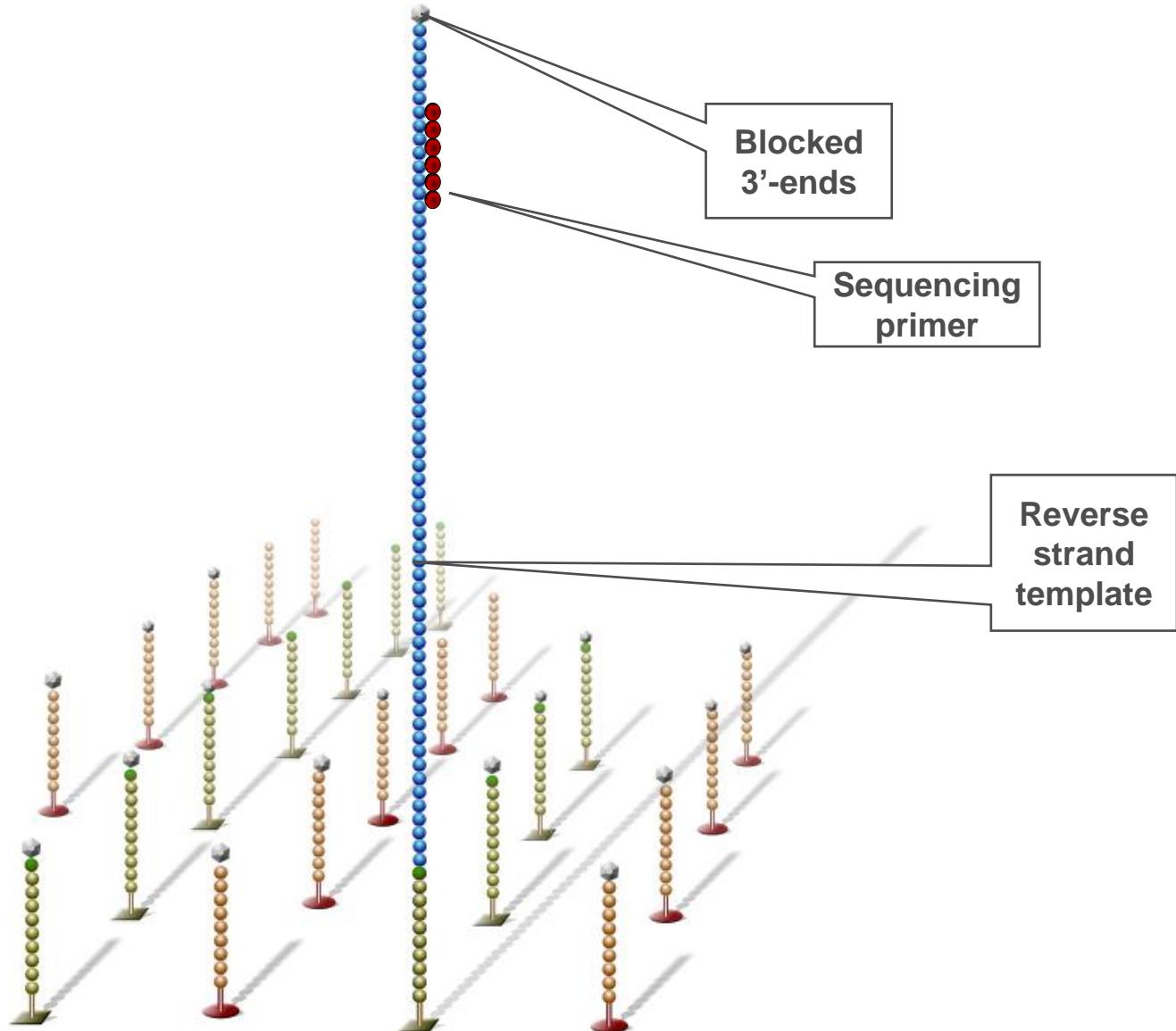
Bridges are linearized and  
the original forward template  
is cleaved



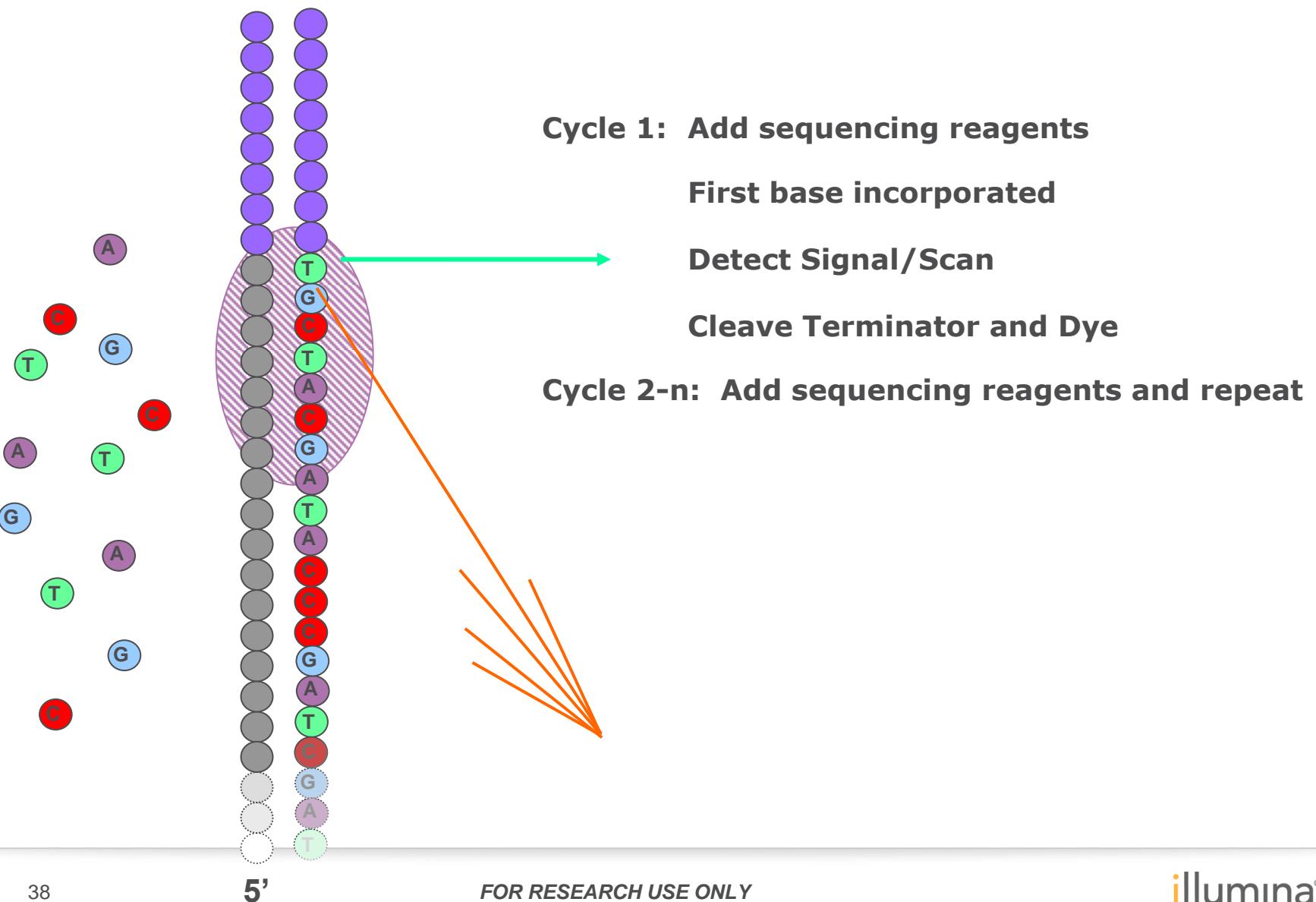
# Paired End Sequencing

Free 3' ends of the reverse template and lawn primers are blocked to prevent unwanted DNA priming

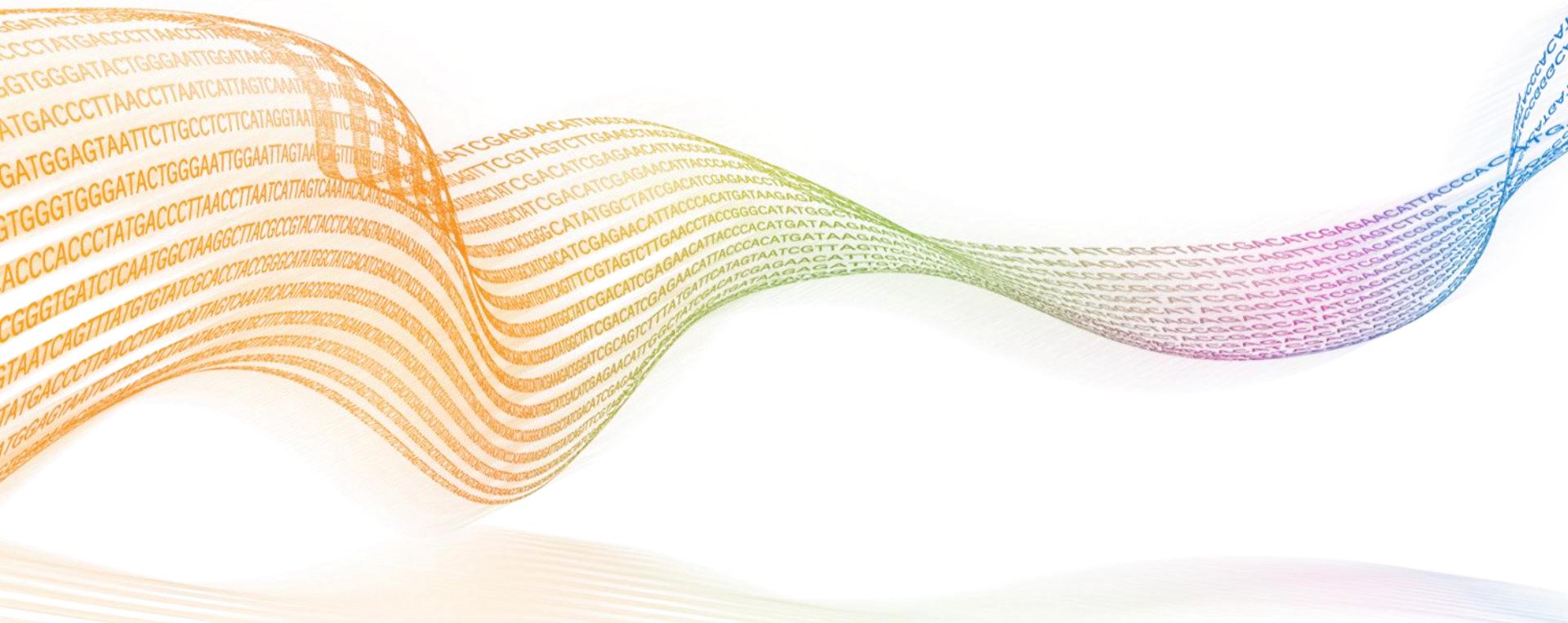
Sequencing primer is hybridized to adapter sequence



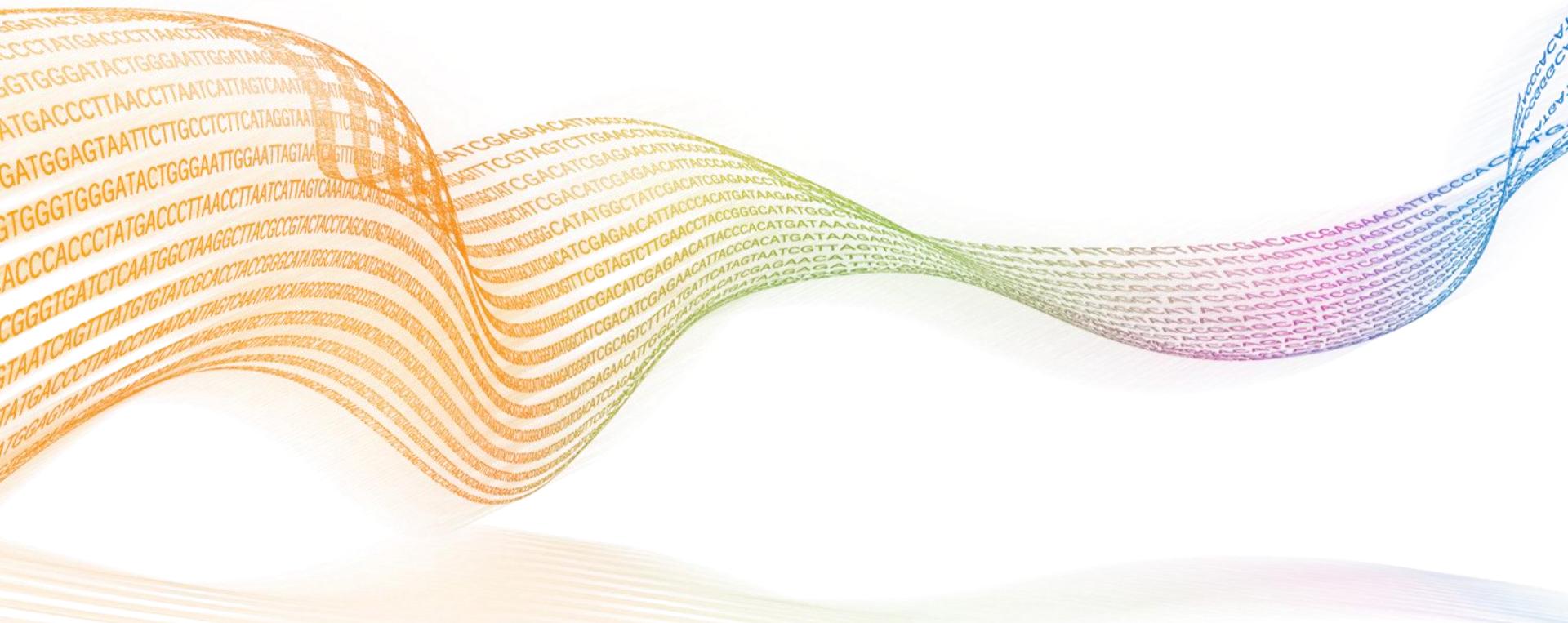
# Sequencing By Synthesis (SBS)



# Dual Indexing



# Data Analysis Overview



# Data Analysis Overview

MiSeq Control  
Software (MCS)  
(on MiSeq)

- Images
- Base calling
- Q Score

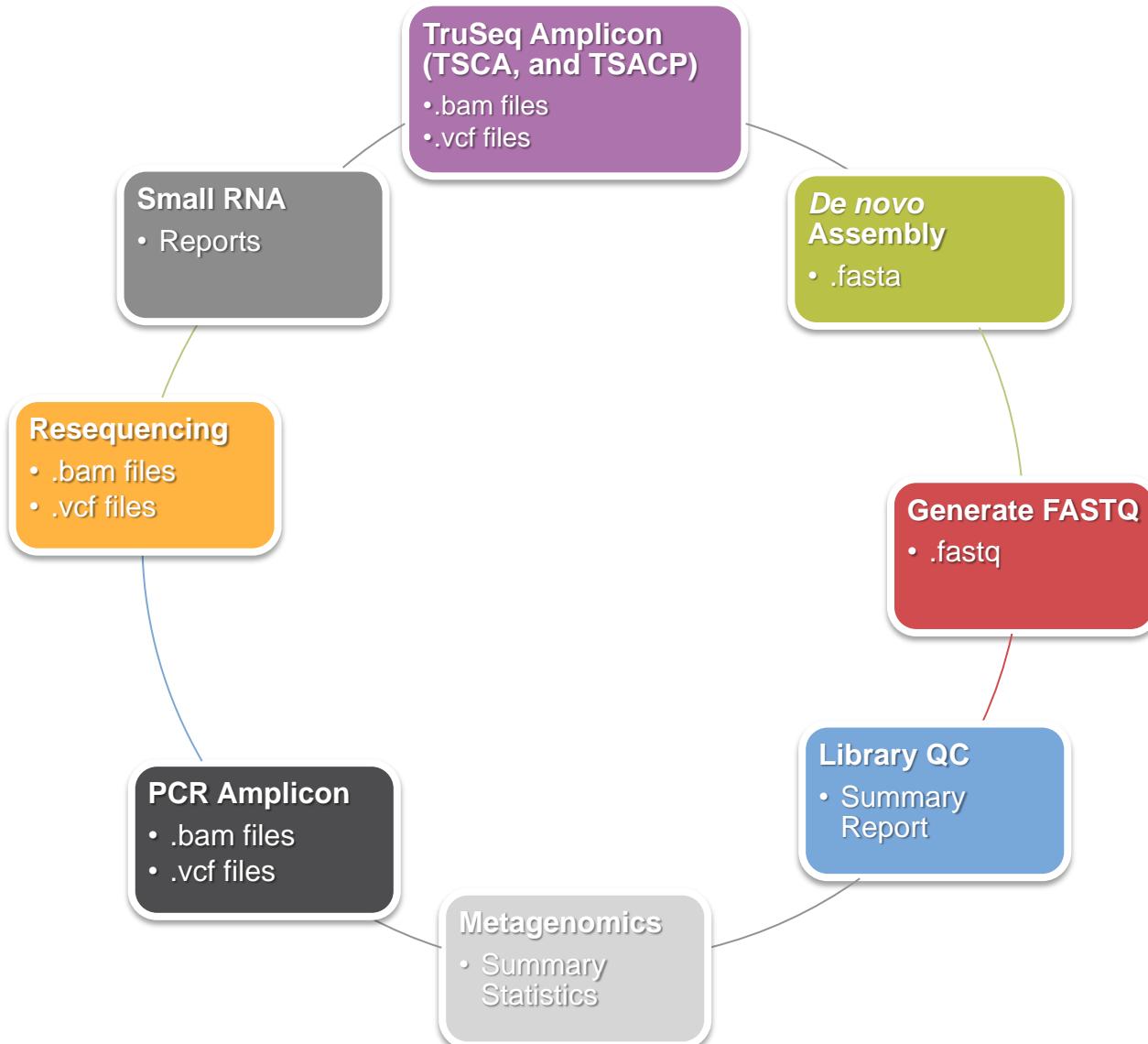
Sequence Analysis  
Viewer (SAV)  
(external computer)

- Monitor Run Data

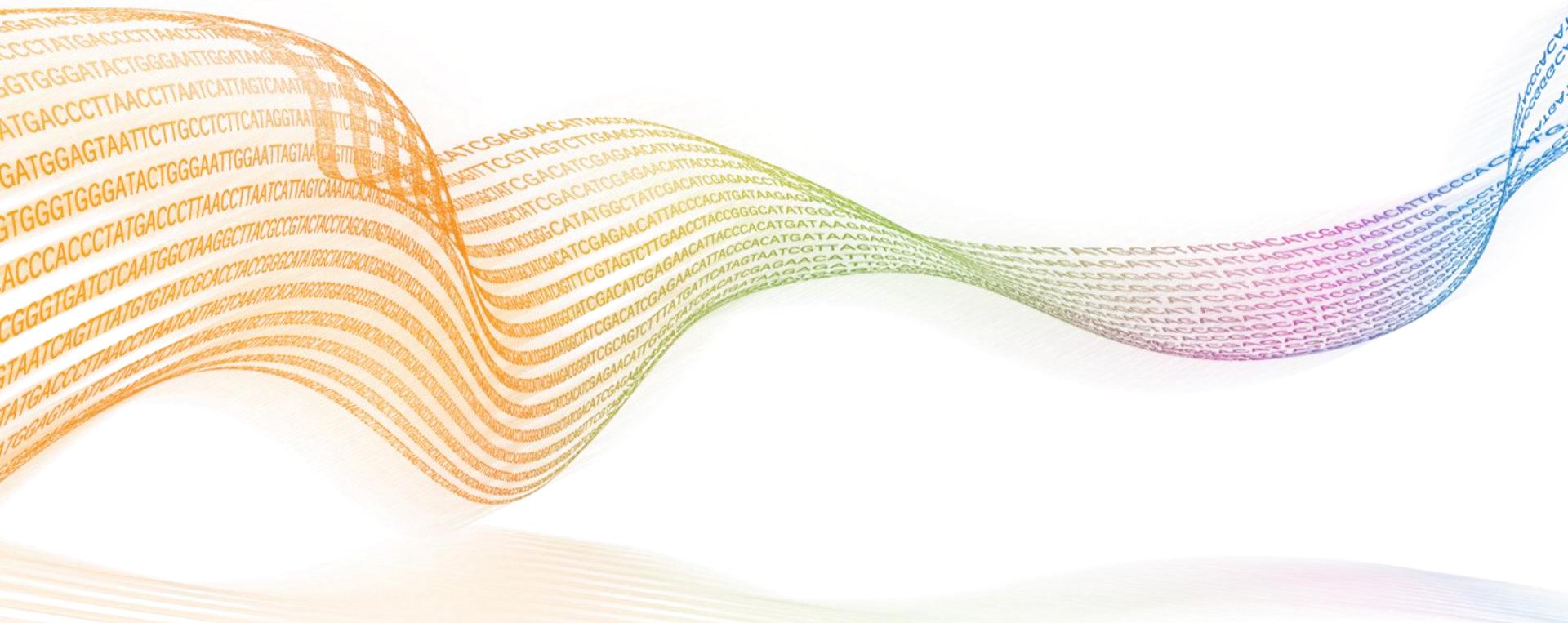
MiSeq Reporter  
Software (MSR)  
(on MiSeq)  
BaseSpace

- Alignment
- Assembly

# Supported Workflows and MSR Output Format

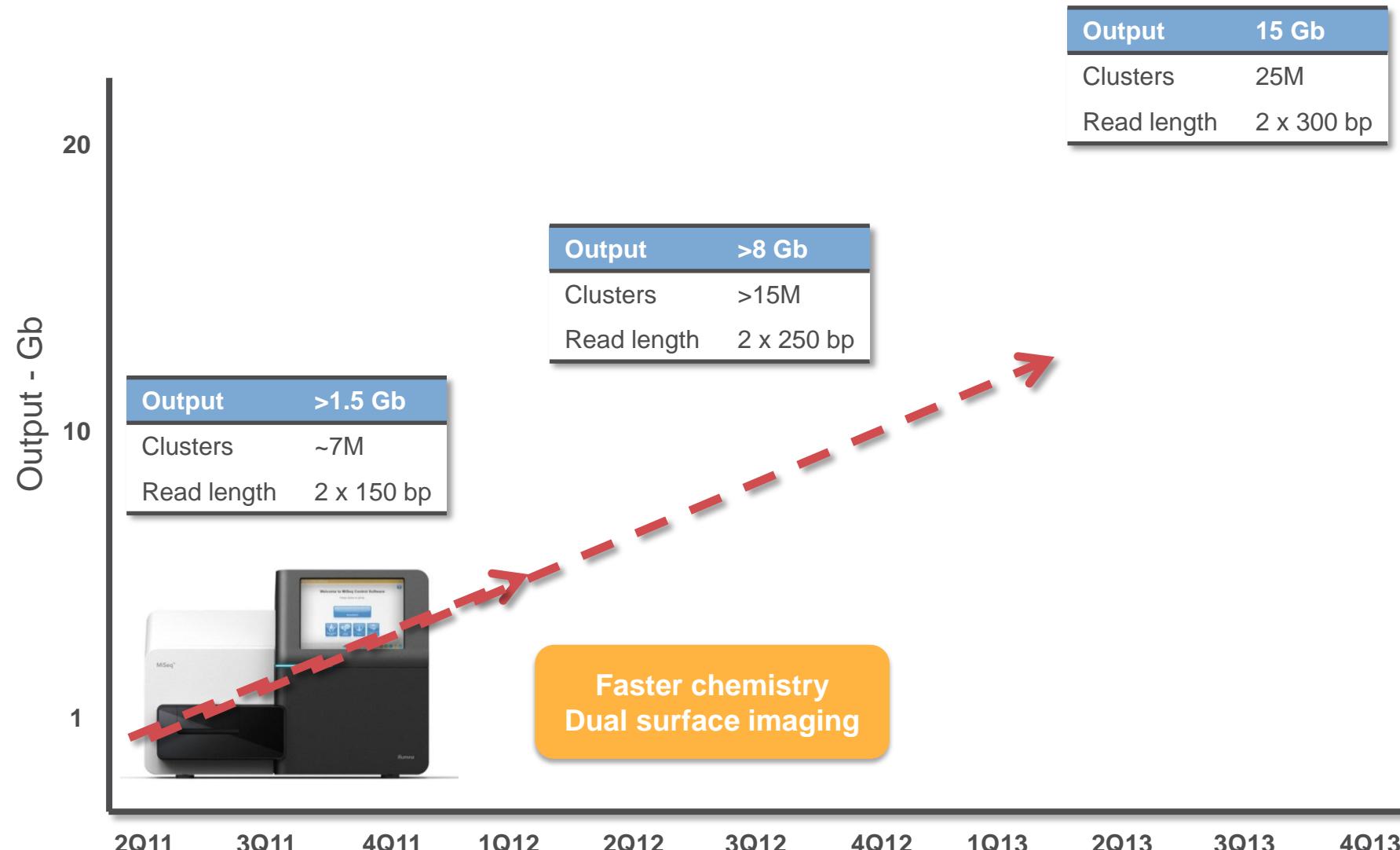


# MiSeq Reagents



# MiSeq – Continuous Performance Improvements

*Path towards 15Gb per run; enabling broader range of applications*



# MiSeq Reagent Kits

Kit Type	Imaging	Available kit sizes
MiSeq Reagent Kit v3	19 tiles, dual surface (38 tiles total )	600 (625 cycles of sequencing) 150 (175 cycles of sequencing)
MiSeq Reagent Kit v2	14 tiles, dual surface (28 tiles total )	500 (525 cycles of sequencing) 300 (325 cycles of sequencing) 50 (75 cycles of sequencing)
MiSeq Reagent Micro Kit v2	4 tiles, dual surface (8 tiles total )	300 (325 cycles of sequencing)
MiSeq Reagent Nano Kit v2	2 tiles, single surface (2 tiles total )	500 (525 cycles of sequencing) 300 (325 cycles of sequencing)

Each Kit is Single use and Contains:

Reagent Cartridge (1)

HT1 (1)

PR2 bottle (1)

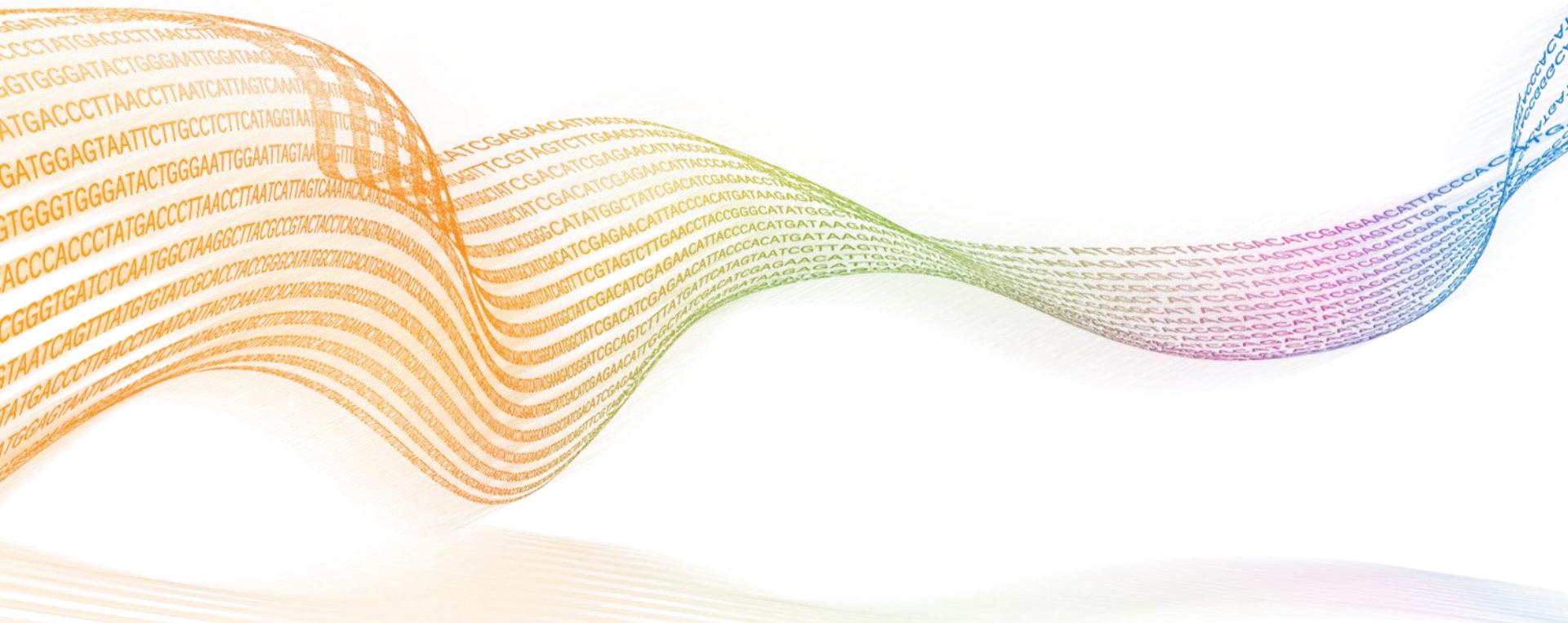
Flow Cell

# Flow Cell Output



Flow cell	# of Reads	Read length	2 x 75 Output	2 x 150 Output	2 x 250 Output	2 x 300 Output
Standard v3 FC	22-25M	Up to 2 x 300	4.5 Gb	----	----	~15 Gb
Standard v2 FC	15 M	Up to 2 x 250	2.25 Gb	4.5 Gb	7.5 Gb	----
Micro FC	4 M	Up to 2 x 150	600 Mb	1.2 Gb	----	----
Nano FC	1 M	Up to 2 x 250	150 Mb	300 Mb	500 Mb	----

# Targeted Resequencing



# Targeted Resequencing

- ▶ Few amplicons
  - Direct PCR
  - Double Step PCR
  - Amplicon > 300bp + Nextera XT
- ▶ From 2 to 650kb
  - TruSeq Custom Amplioch (TSCA)
- ▶ From 0.5 to 15MB
  - Nextera Rapid Capture Custom Enrichment (NRCCE)
- ▶ Exome Sequencing
  - TruSight Exome
  - TruSight One
  - Nextera Rapid Capture Exome
  - Nextera Rapid Capture Extended Exome



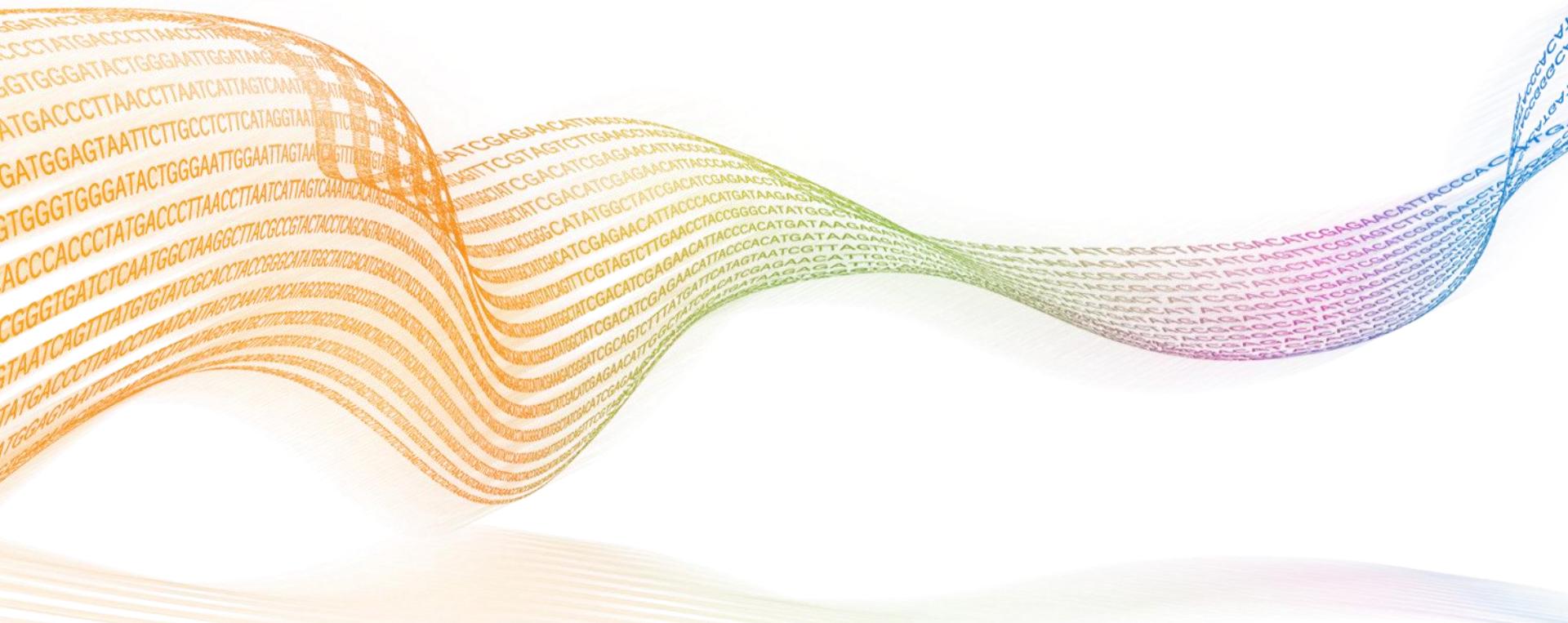
# Targeted Resequencing

- ▶ Few amplicons
  - Direct PCR
  - **Double Step PCR**
  - Amplicon > 300bp + Nextera XT
- ▶ From 2 to 650kb
  - TruSeq Custom Amplioch (TSCA)
- ▶ From 0.5 to 15MB
  - Nextera Rapid Capture Custom Enrichment (NRCCE)
- ▶ Exome Sequencing
  - TruSight Exome
  - TruSight One
  - Nextera Rapid Capture Exome
  - Nextera Rapid Capture Extended Exome



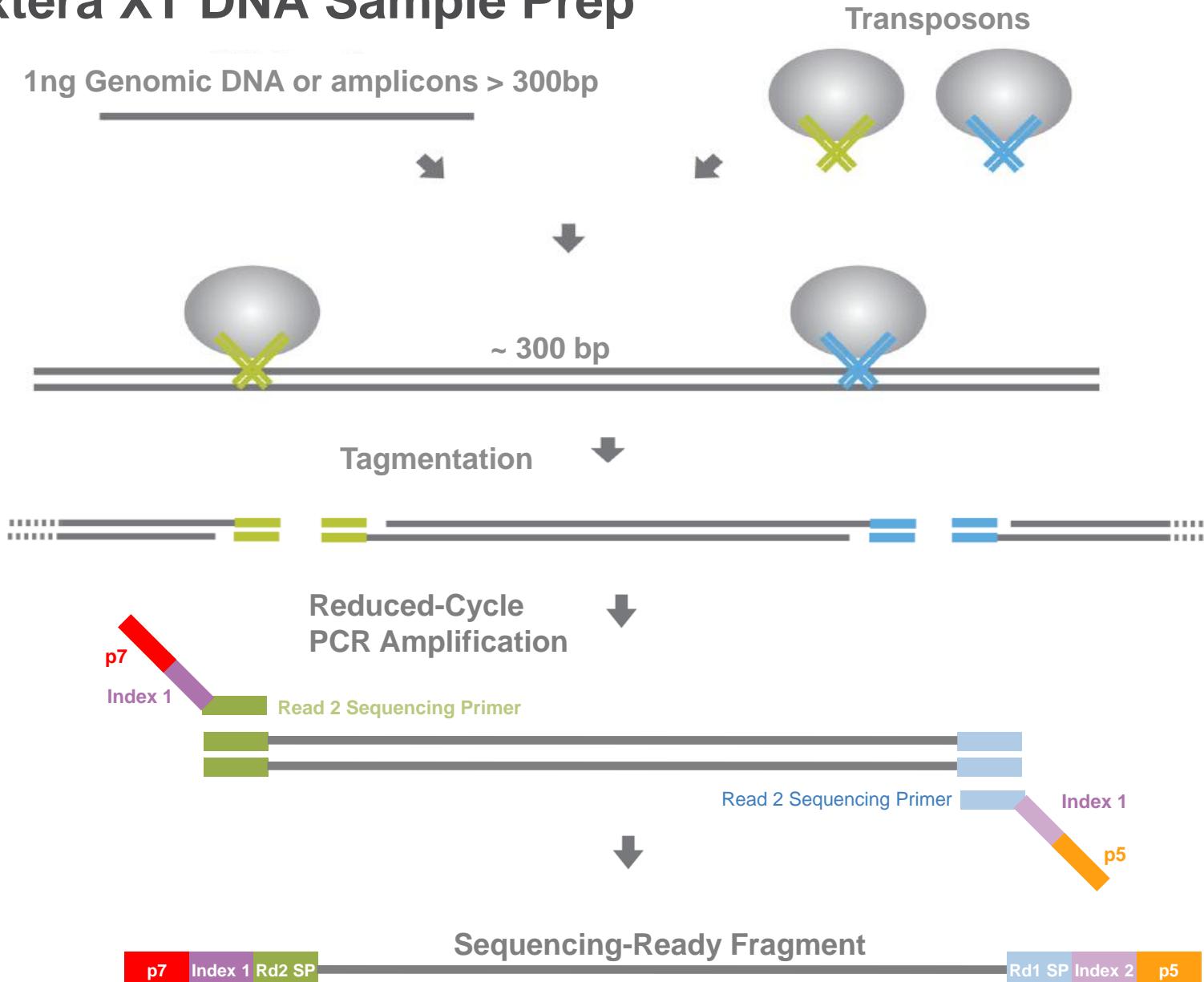
# Targeted Resequencing

## *Two-steps PCR*

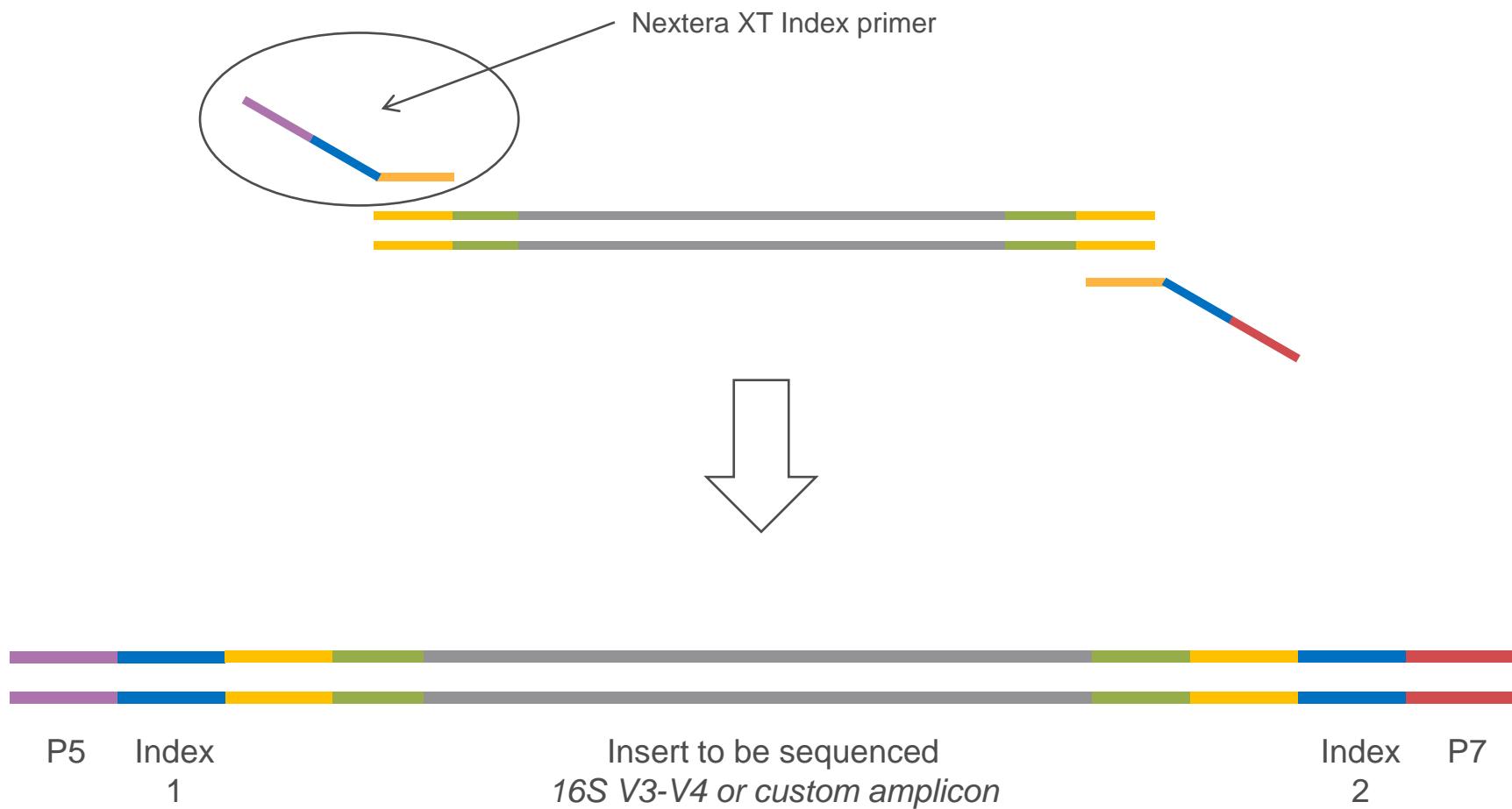


# NexTera XT DNA Sample Prep

1ng Genomic DNA or amplicons > 300bp



## Step 2: 2<sup>nd</sup> round of PCR adds ILMN indices and adapters



# Questions?

