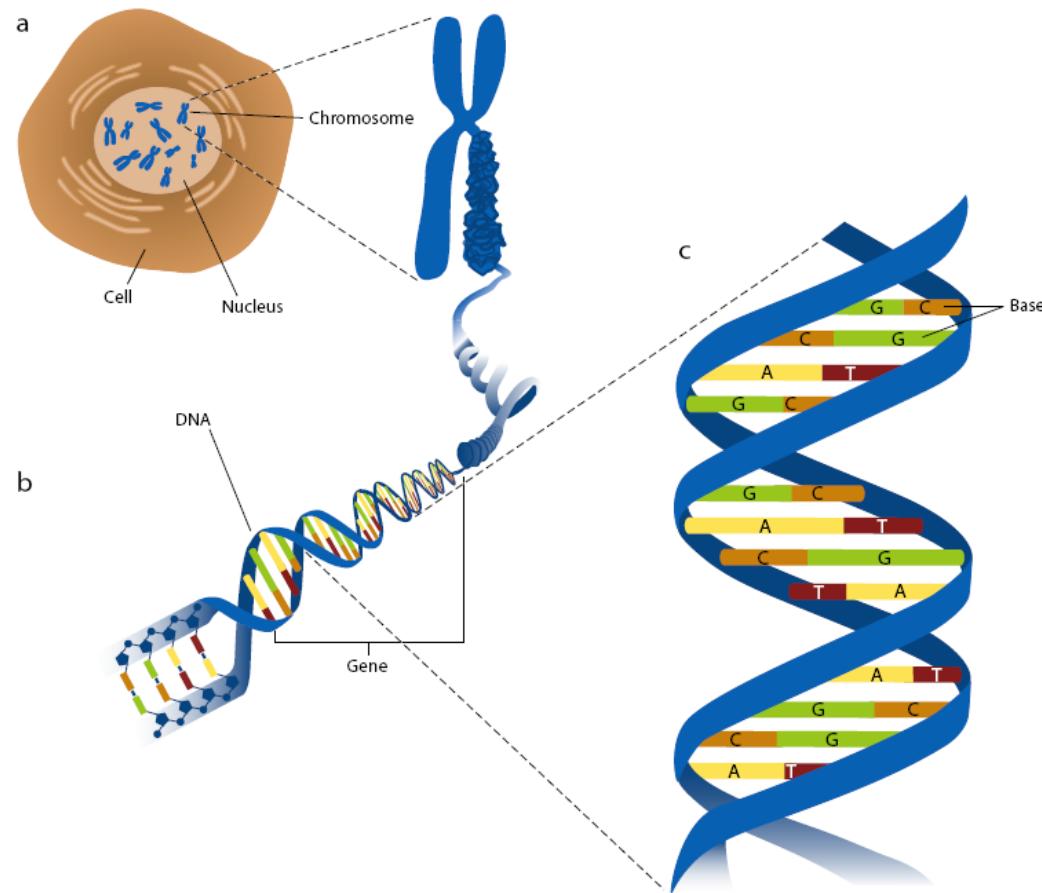


# Overview of NGS sequencing: methods and technologies



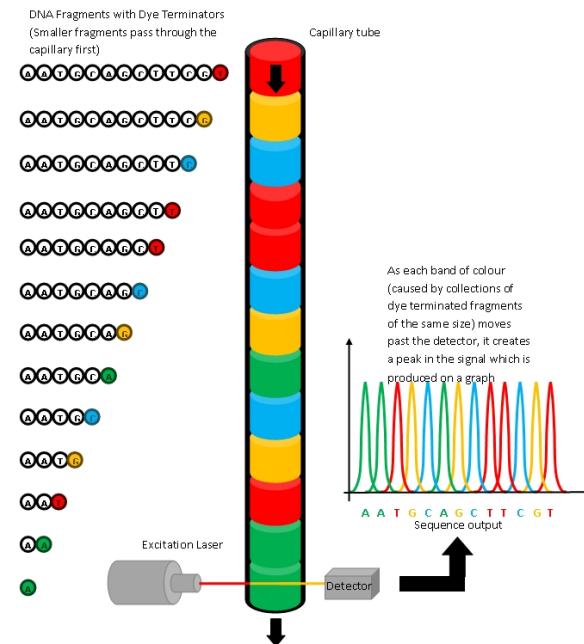
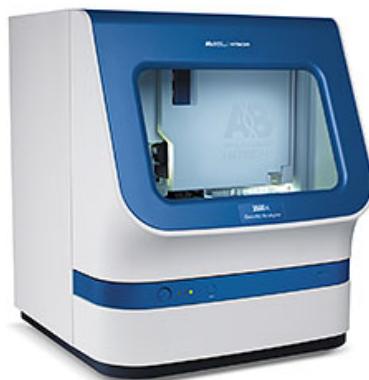
Genomics Core  
Workshop intro NGS 08/10/2019  
Céline Helsmoortel

- Next Generation Sequencing



# NGS

- Next Generation Sequencing
- “First generation sequencing”
  - Sanger sequencing



# NGS technologies

- Second generation



454/Roche



SoliD



Complete Genomics



Illumina



Life Technologies

- Third generation

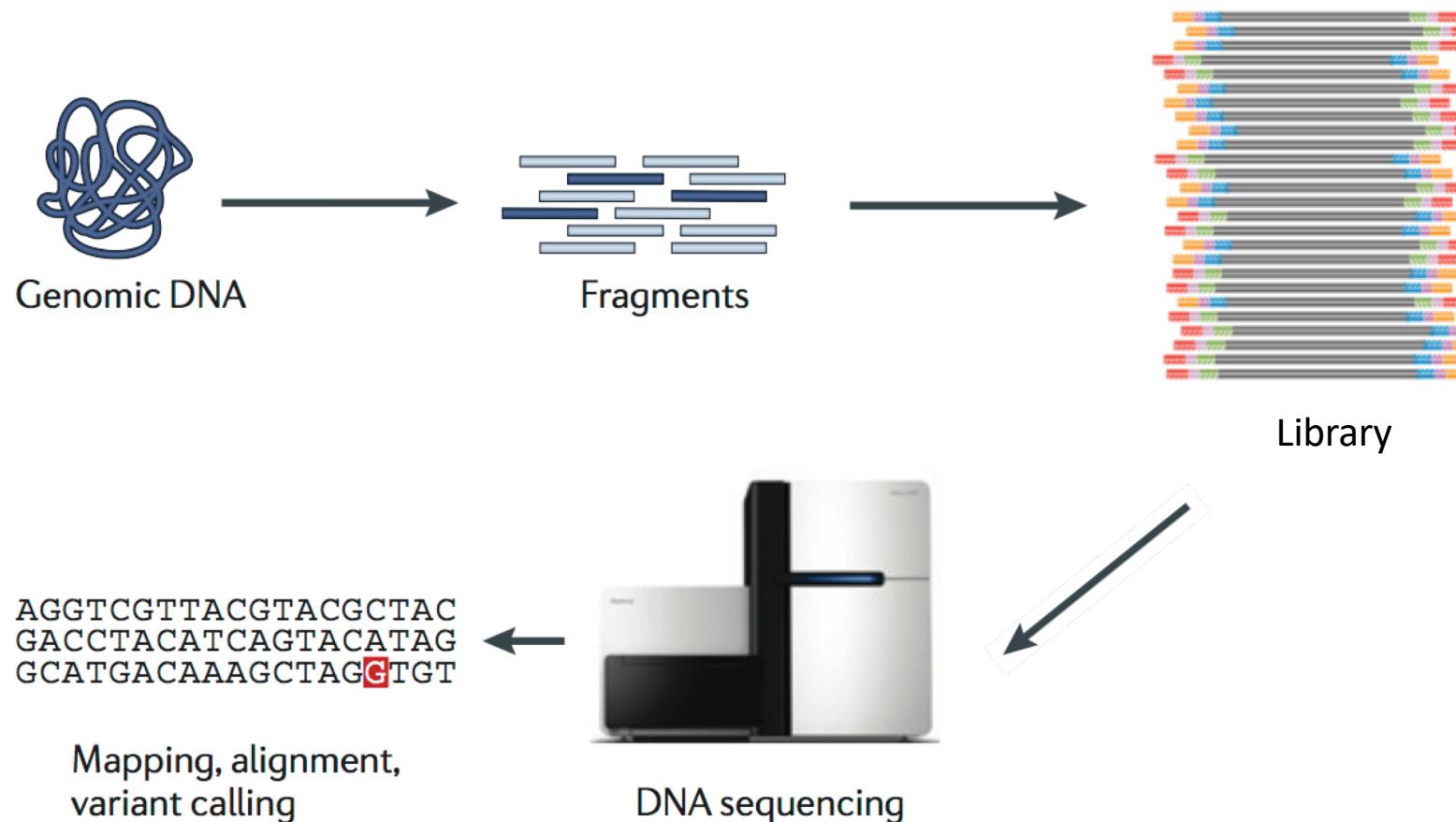


Pacific BioSciences



Oxford Nanopore

# NGS workflow



# NGS workflow

---

- Library preparation



- Sequencing

- clustering
  - sequencing



- Data-analysis

AGGTCGTTACGTACGCTAC  
GACCTACATCAGTACATAG  
GCATGACAAAGCTAG**G**TGT

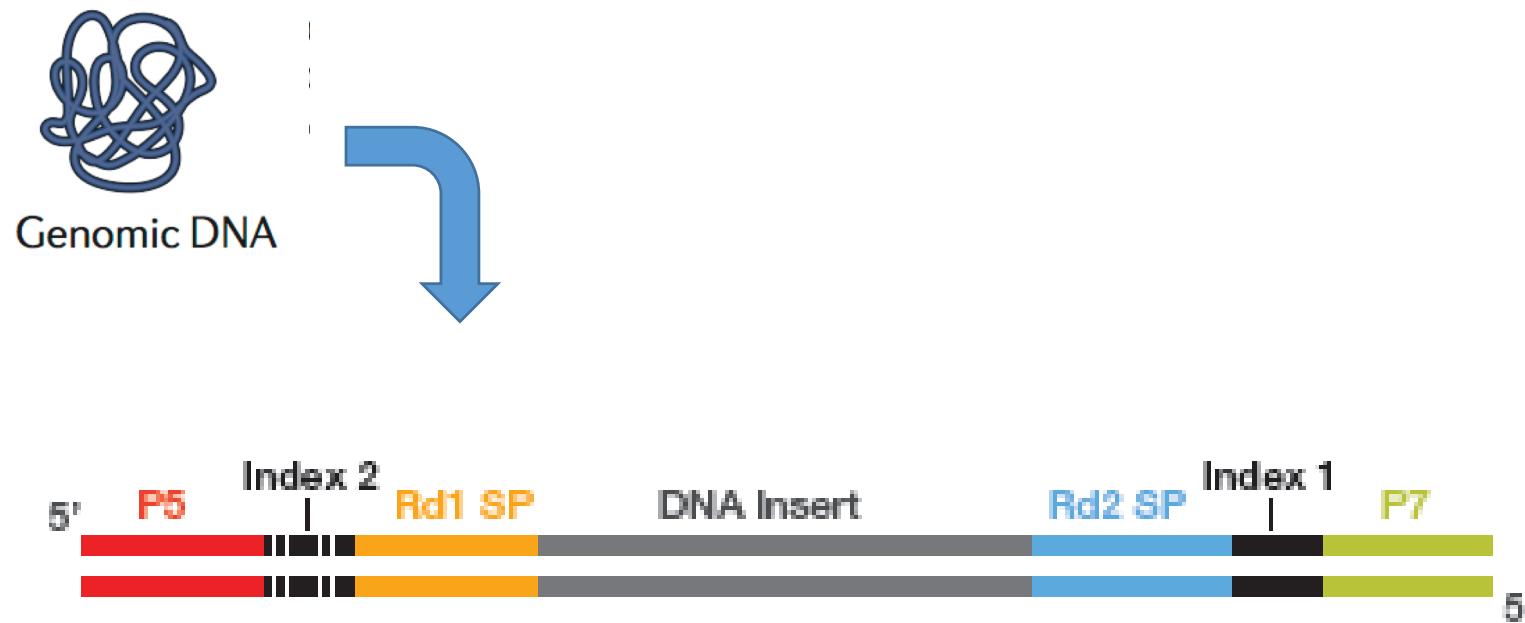
# NGS workflow

---

- Library preparation
- Sequencing
  - clustering
  - sequencing
- Data-analysis



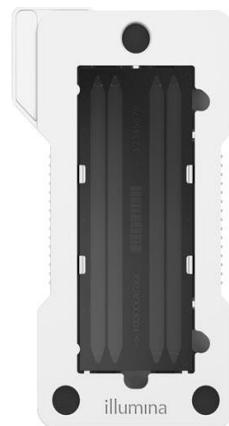
# NGS Library Prep



# Illumina flowcells



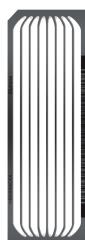
MiSeq



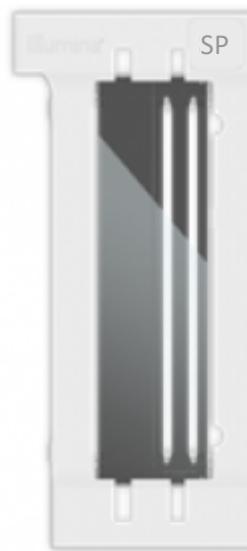
NextSeq



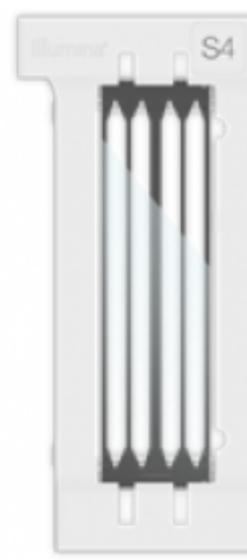
Rapid  
HiSeq



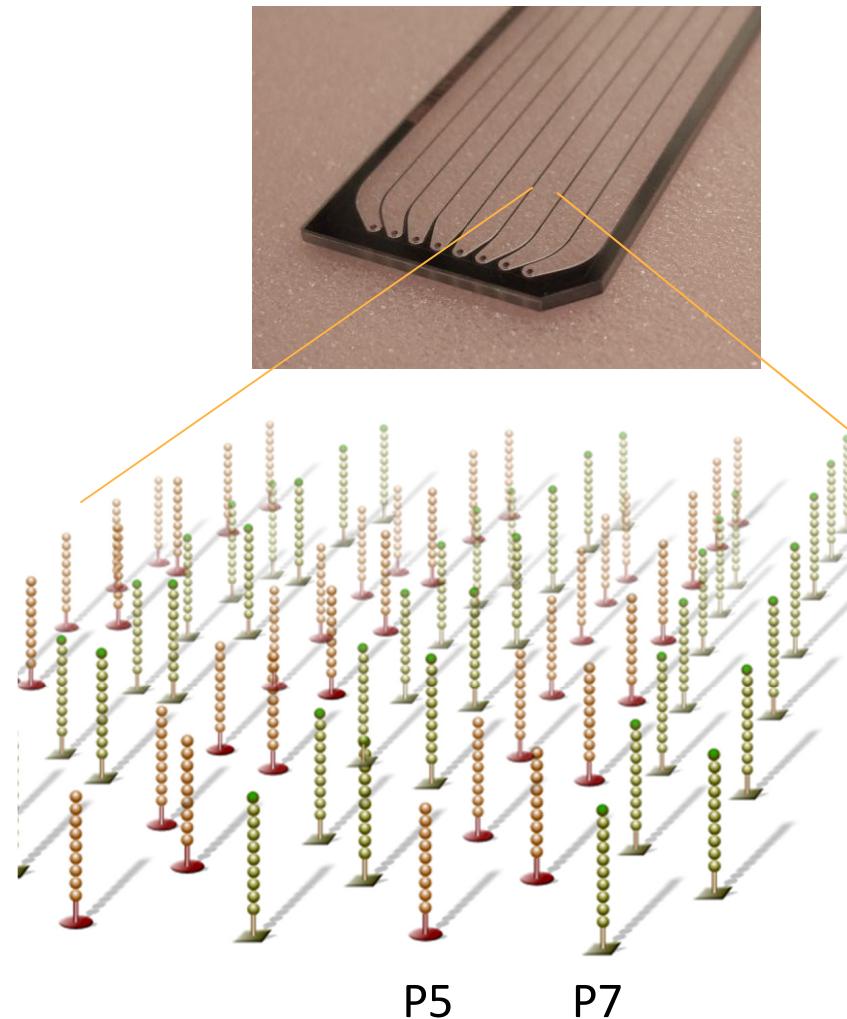
High Output



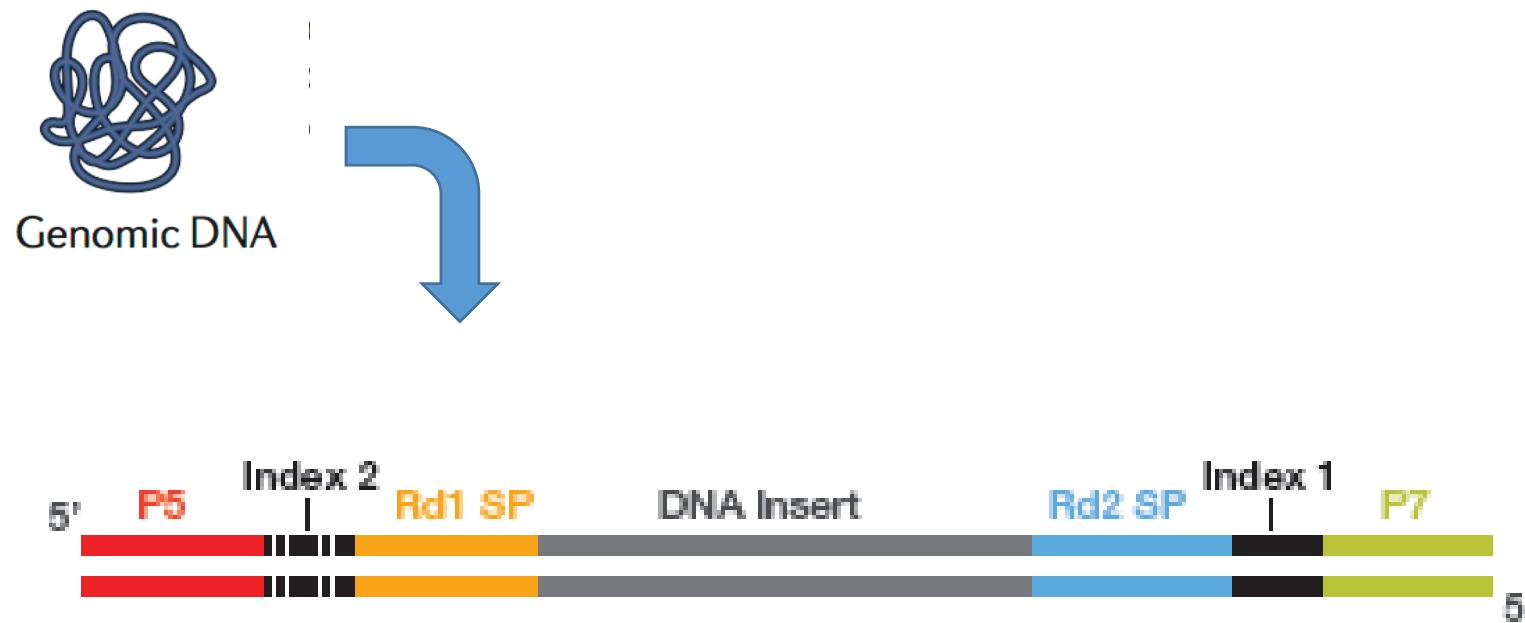
NovaSeq



# NGS Library Prep

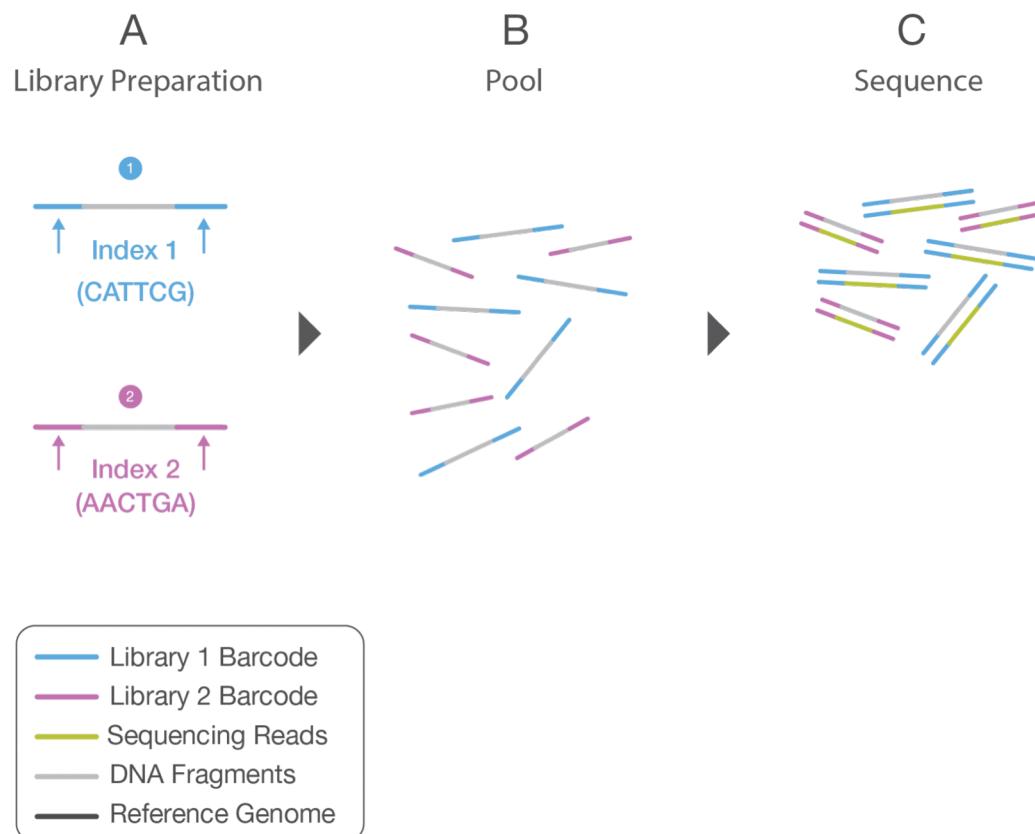


# NGS Library Prep

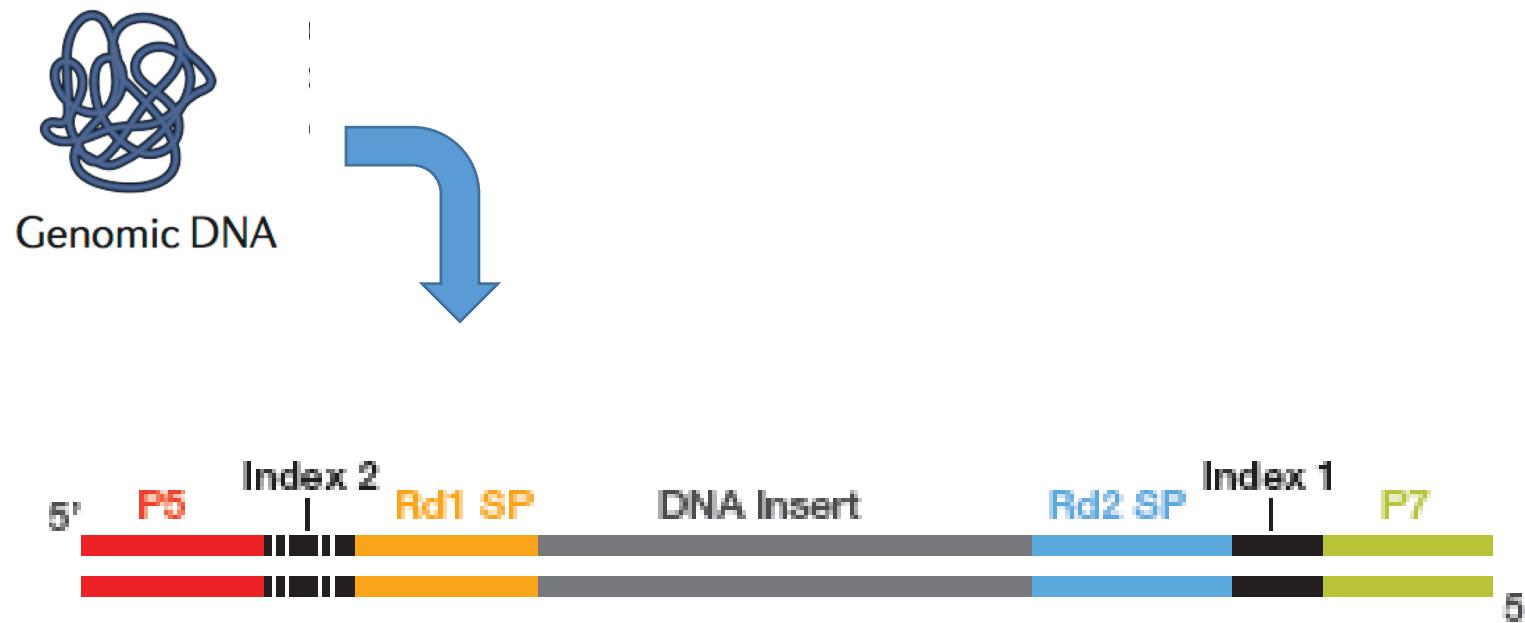


# NGS Library Prep

- Indexing



# NGS Library Prep



# NGS Library Prep

---

- General DNA library prep
- General RNA library prep
- Targeted library prep approaches

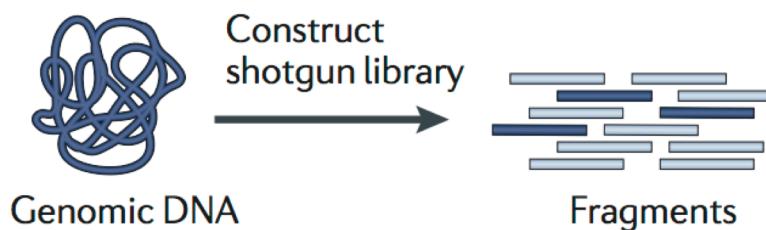
# NGS Library Prep

---

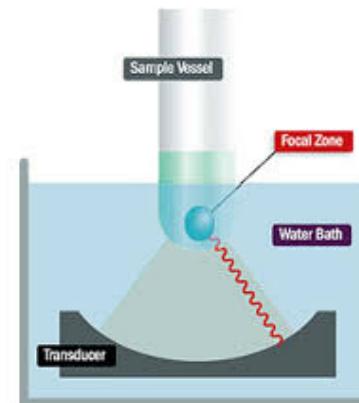
- General DNA library prep
- General RNA library prep
- Targeted library prep approaches

# NGS Library prep

- Fragmentation

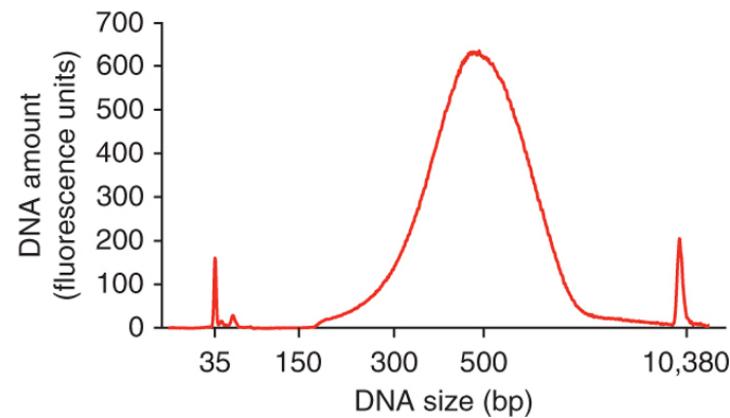


- acoustic  
covaris
- enzymatic



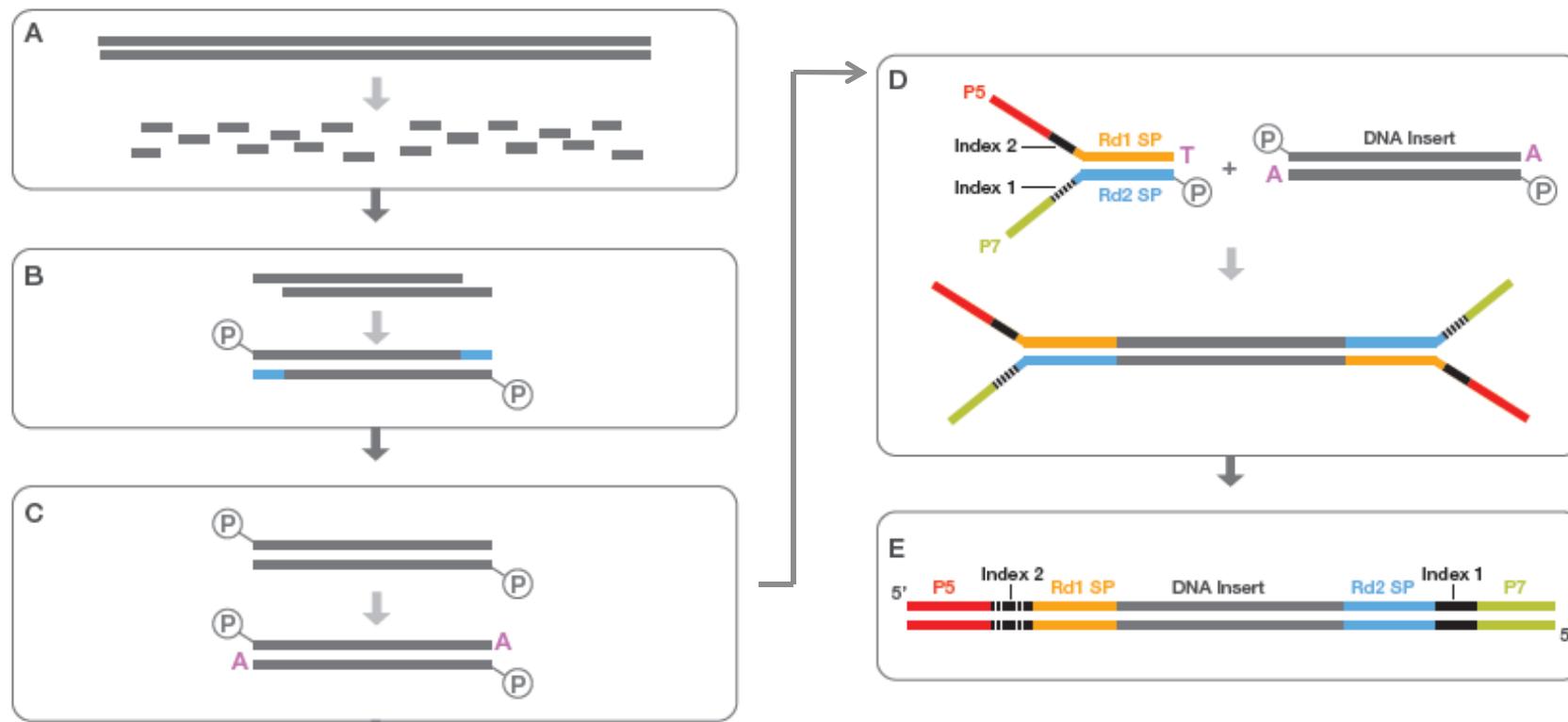
# NGS Library prep

- Fragmentation
  - quality check: BioAnalyser / FragmentAnalyser



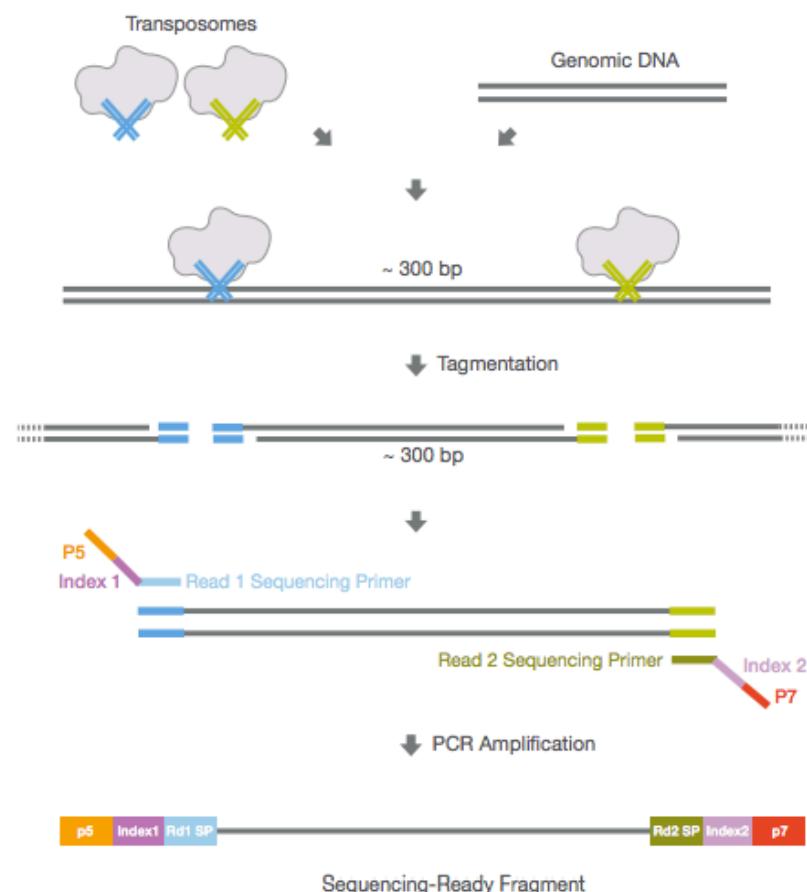
# NGS Library prep

- Standard A-tailing & adaptor-ligation (DNA/RNA)



# NGS Library Prep

- Illumina Nextera tagmentation



# NGS Library Prep

---

- General DNA library prep
- General RNA library prep
- Targeted library prep approaches

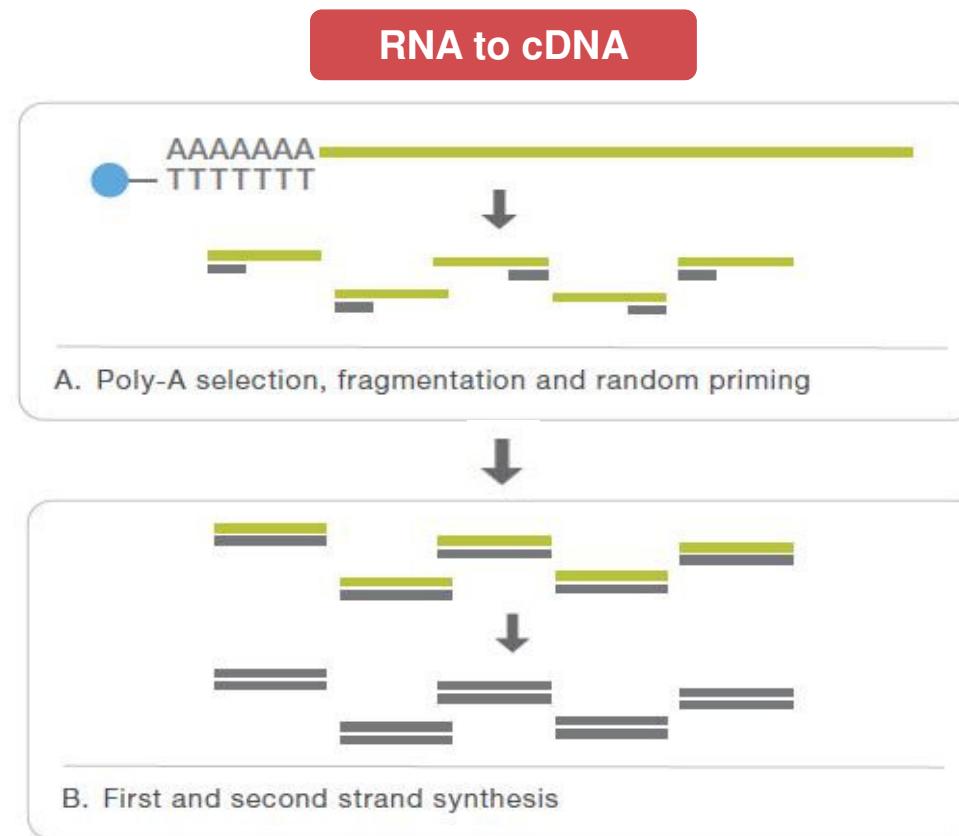
# NGS Library Prep

---

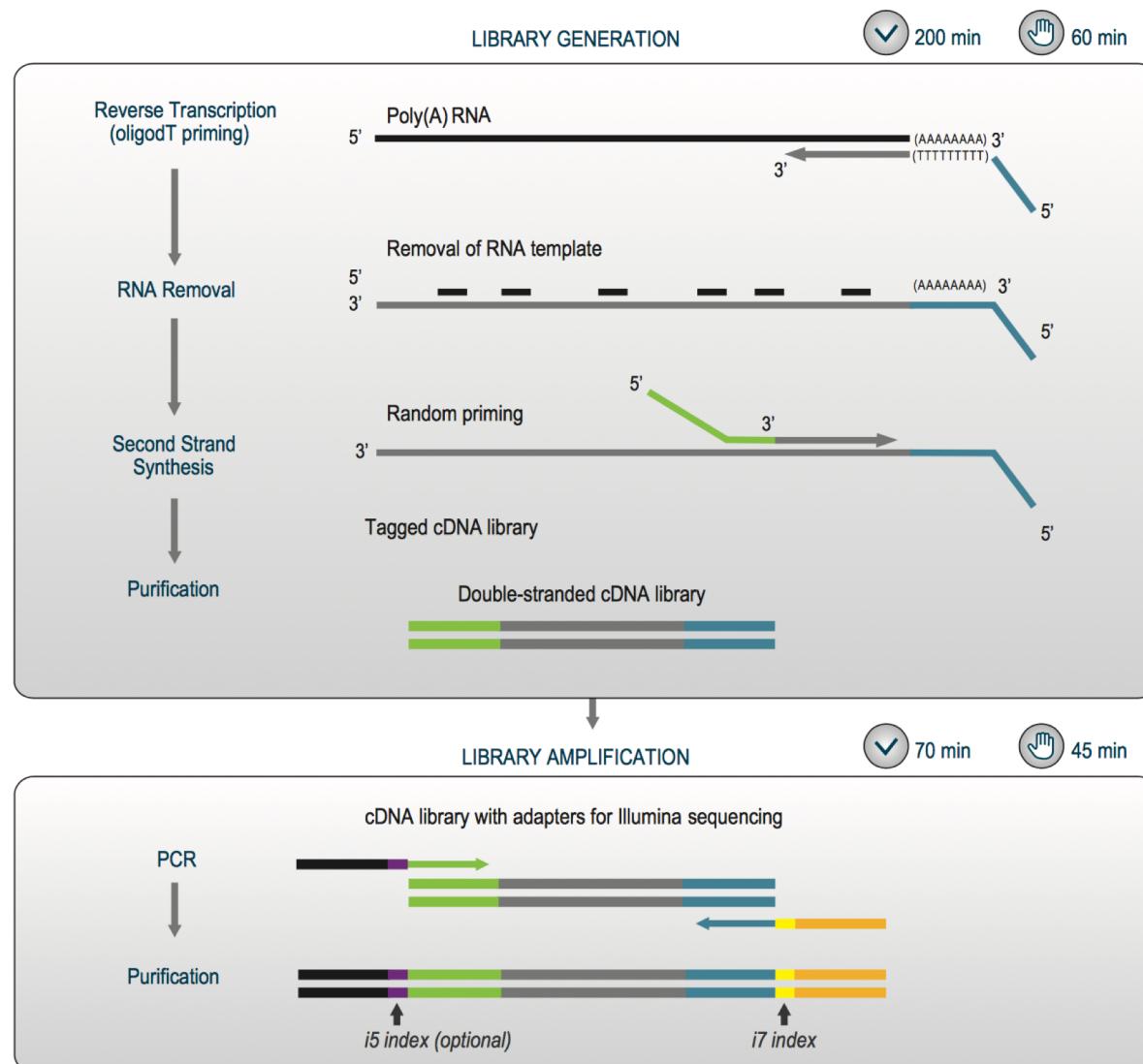
- General DNA library prep
- General RNA library prep
- Targeted library prep approaches

# NGS Library Prep

- Illumina TruSeq Stranded mRNA



# LexoGen QuantSeq (RNA)



# NGS Library Prep

---

- General DNA library prep
- General RNA library prep
- Targeted library prep approaches

# NGS Library Prep

---

- General DNA library prep
- General RNA library prep
- Targeted library prep approaches
  - amplicon based
  - ligation based
  - enrichment based

# NGS Library Prep

---

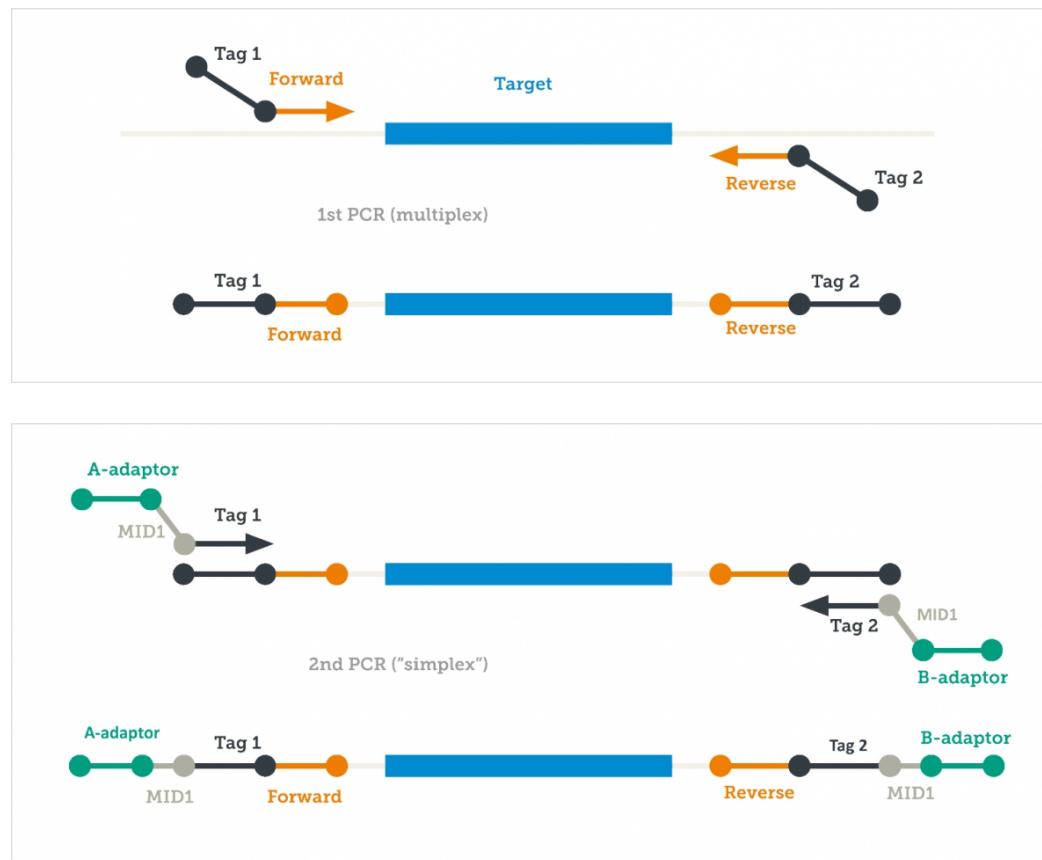
- Custom two step PCR



Performed by you

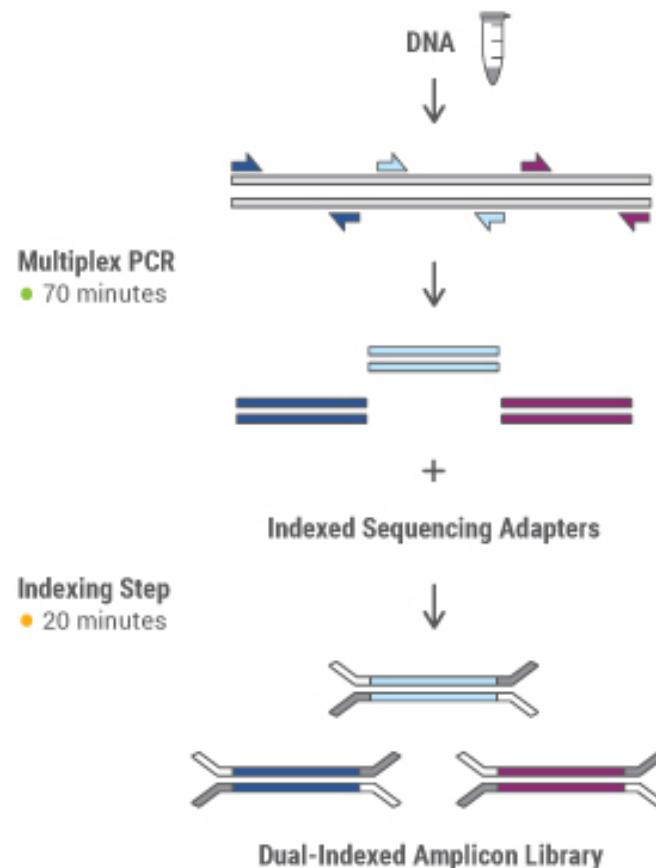
# NGS Library Prep

- Multiplicom MASTR assay



# NGS Library Prep

- Accel amplicon



# NGS Library Prep

- HaloPlex

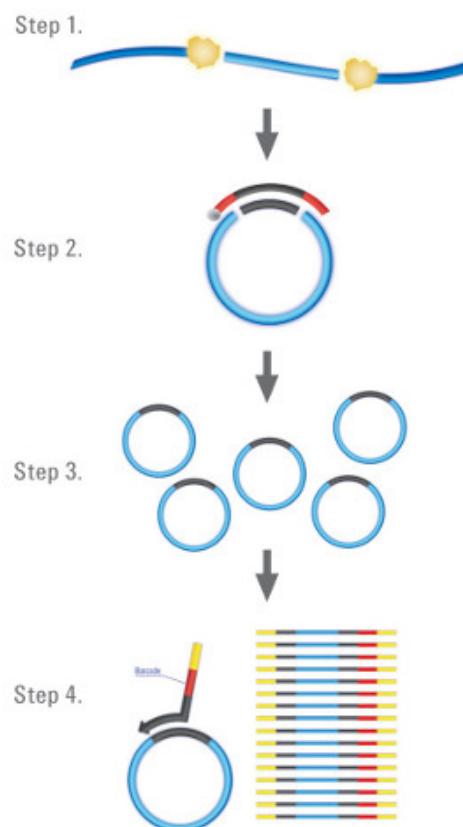
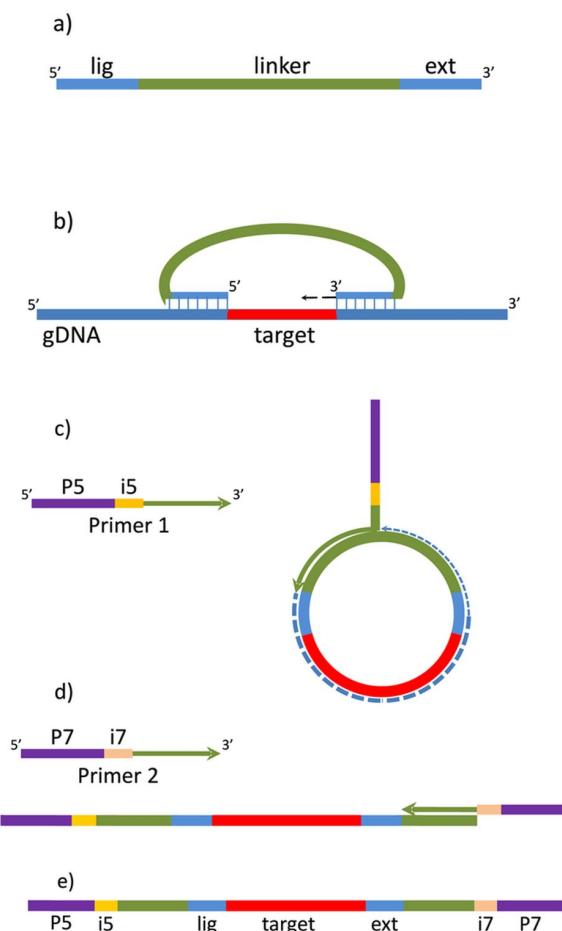


Figure 1. HaloPlex workflow.

# NGS Library Prep

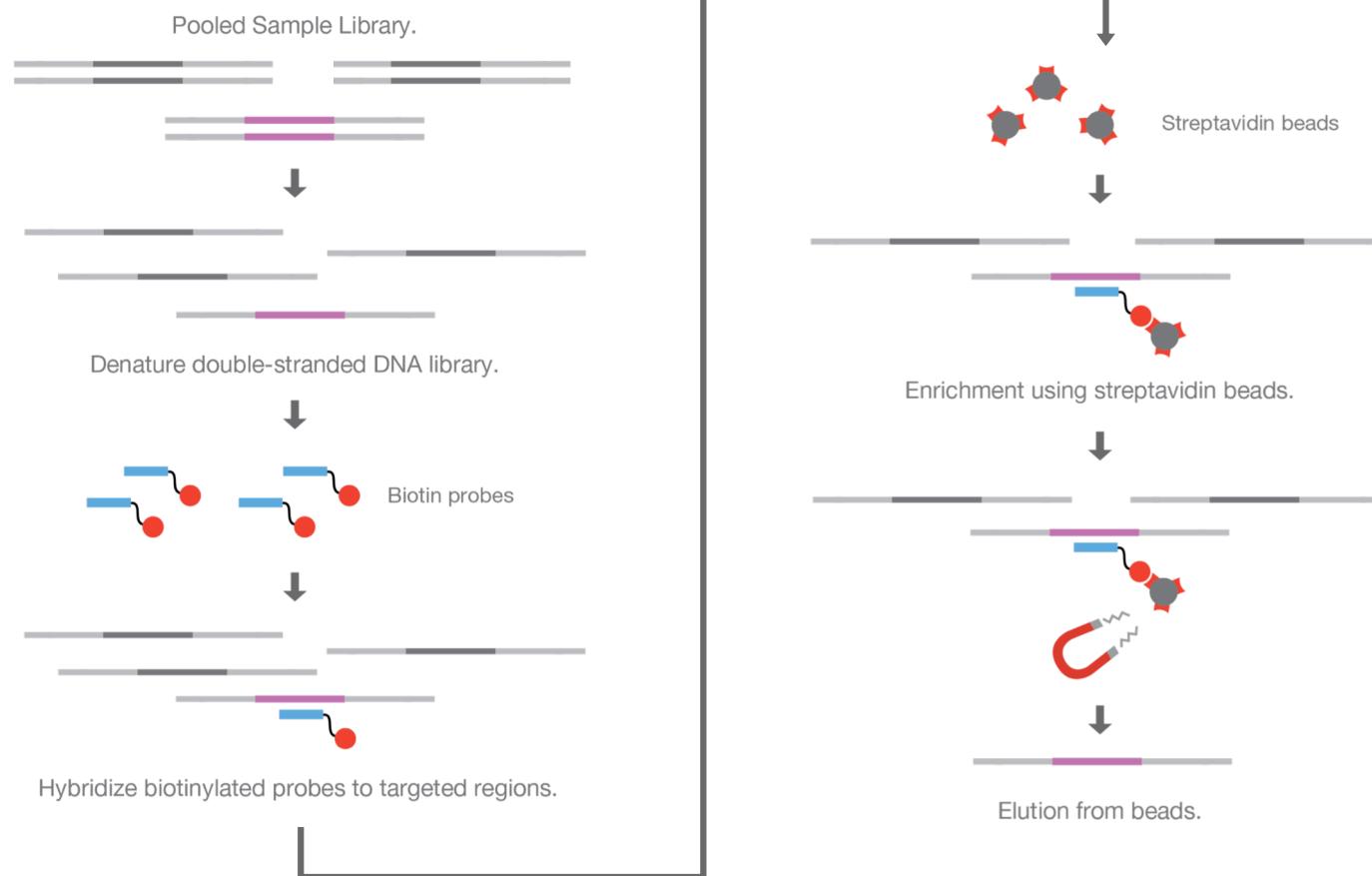
- MIPs



Niedzicka *et al.* 2016

# NGS Enrichment

- Sequence capture



# NGS workflow

---

- Library preparation

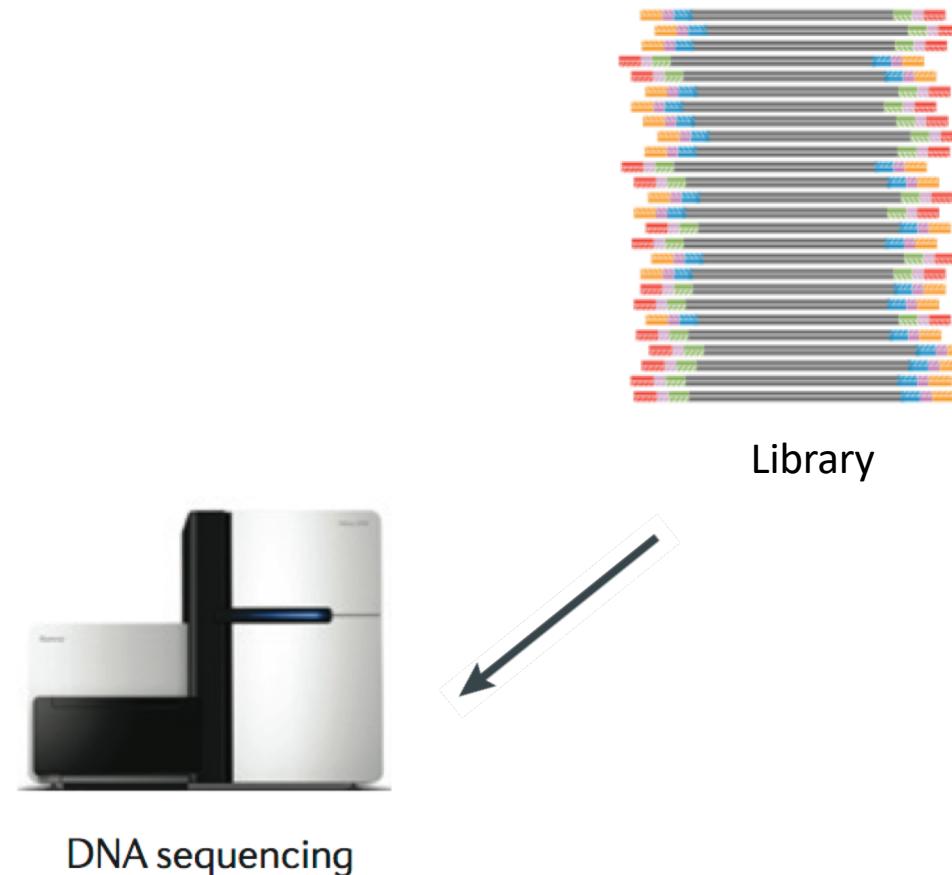
- Sequencing

- clustering
- sequencing

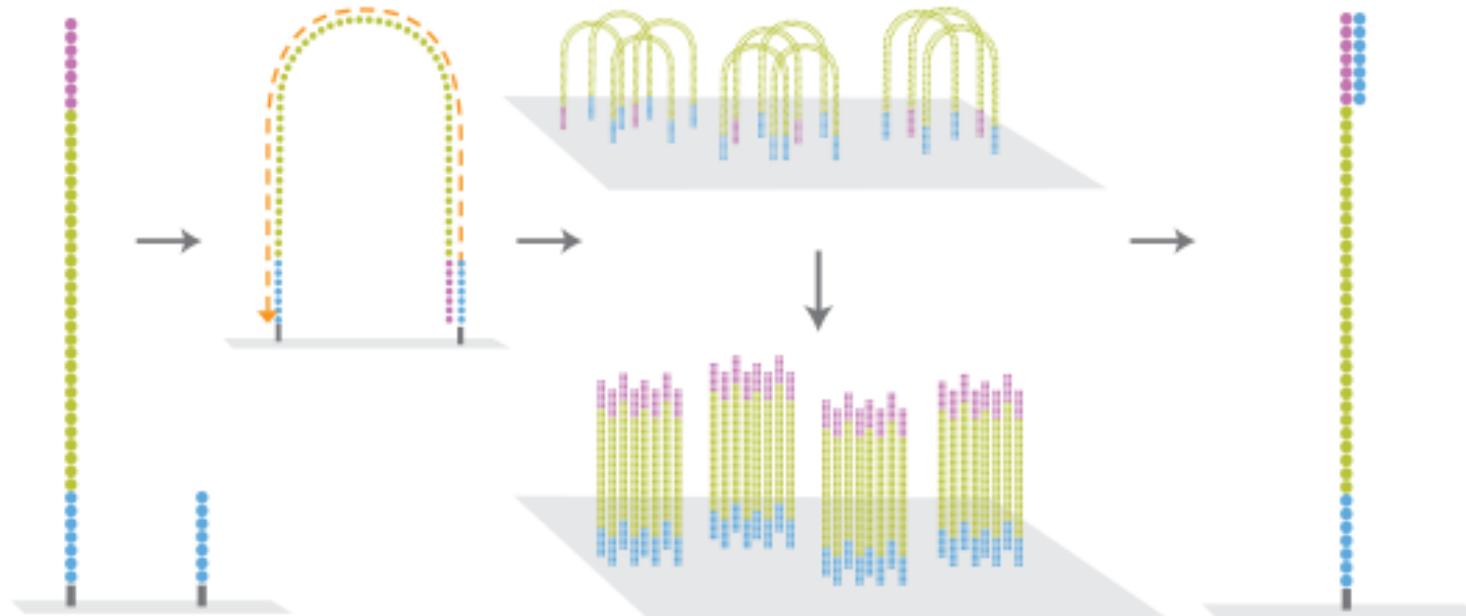


- Data-analysis

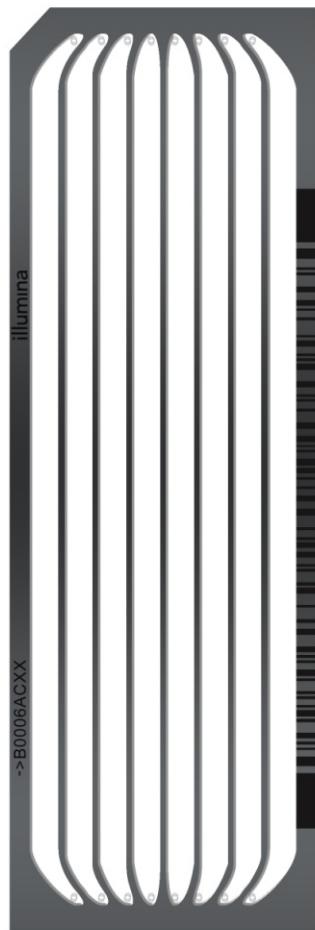
# NGS workflow



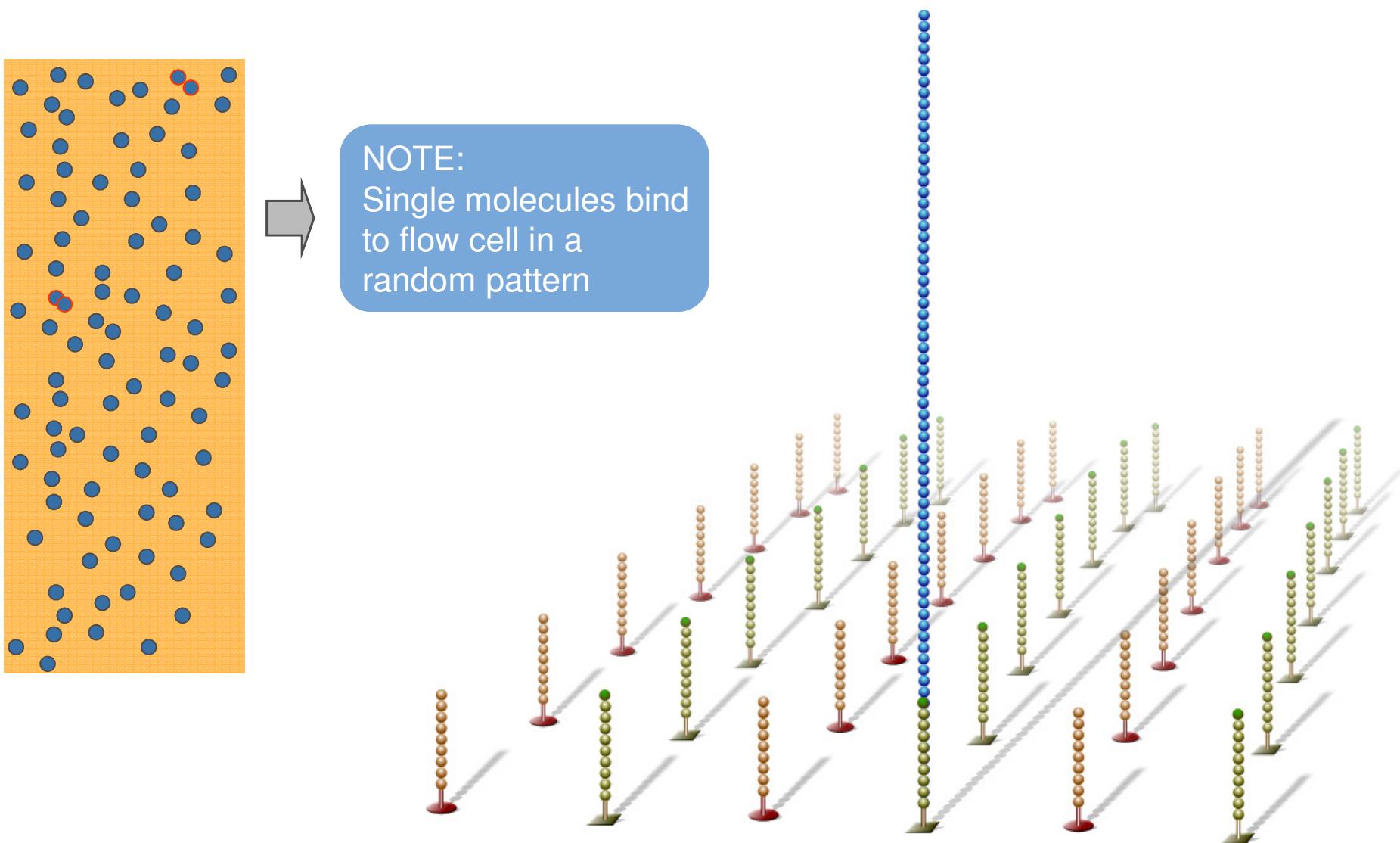
# NGS Illumina Clustering



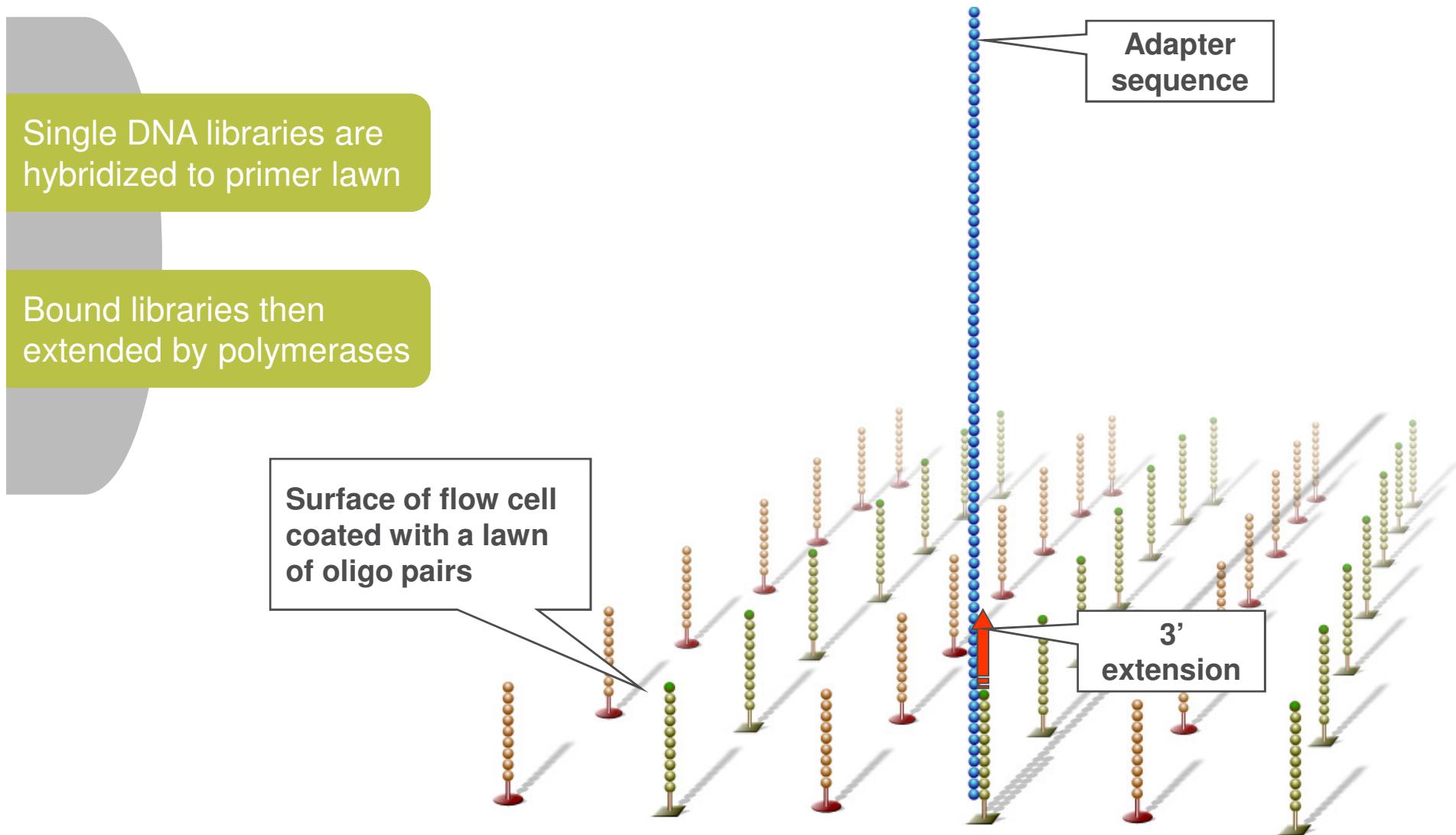
# NGS Illumina Clustering



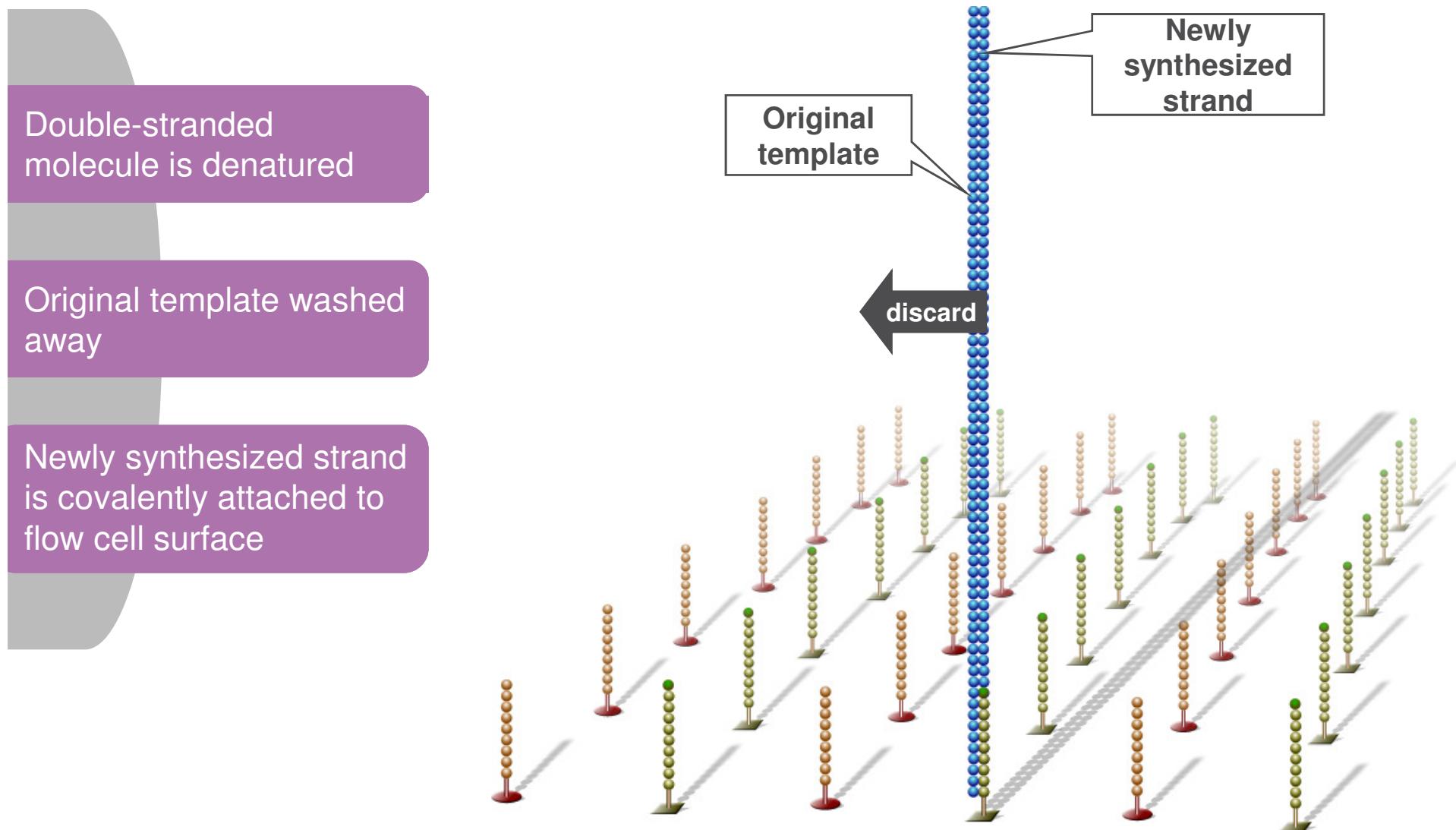
# NGS Illumina Clustering



# NGS Illumina Clustering



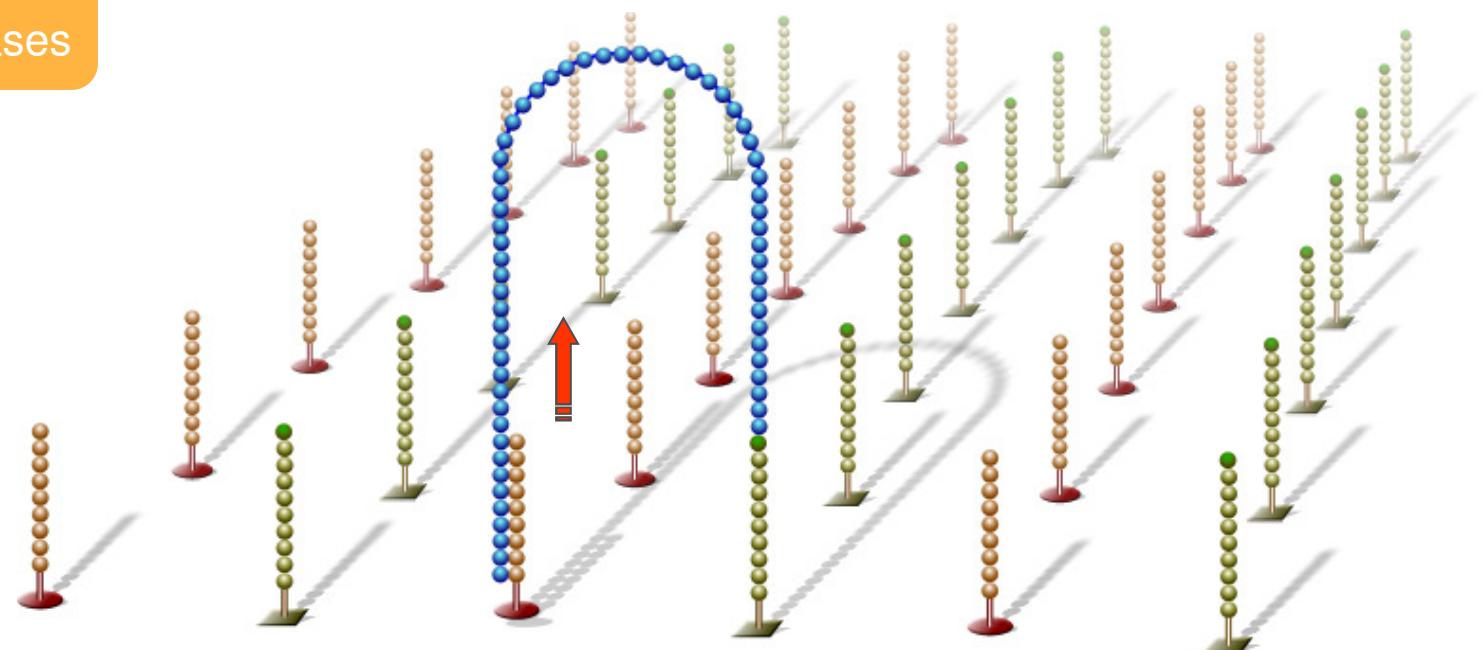
# NGS Illumina Clustering



# NGS Illumina Clustering

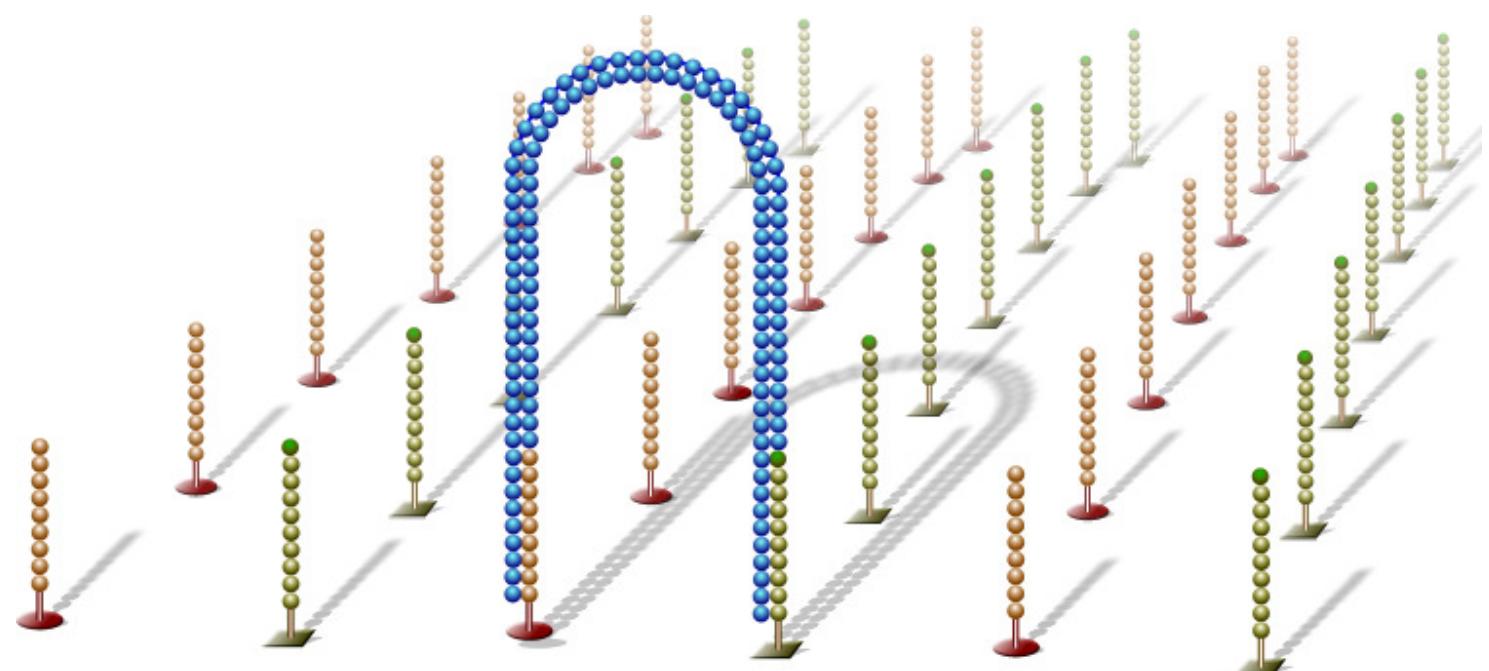
Single-stranded molecule flips over  
and forms a bridge by hybridizing to  
adjacent, complementary primer

Hybridized primer is  
extended by polymerases



# NGS Illumina Clustering

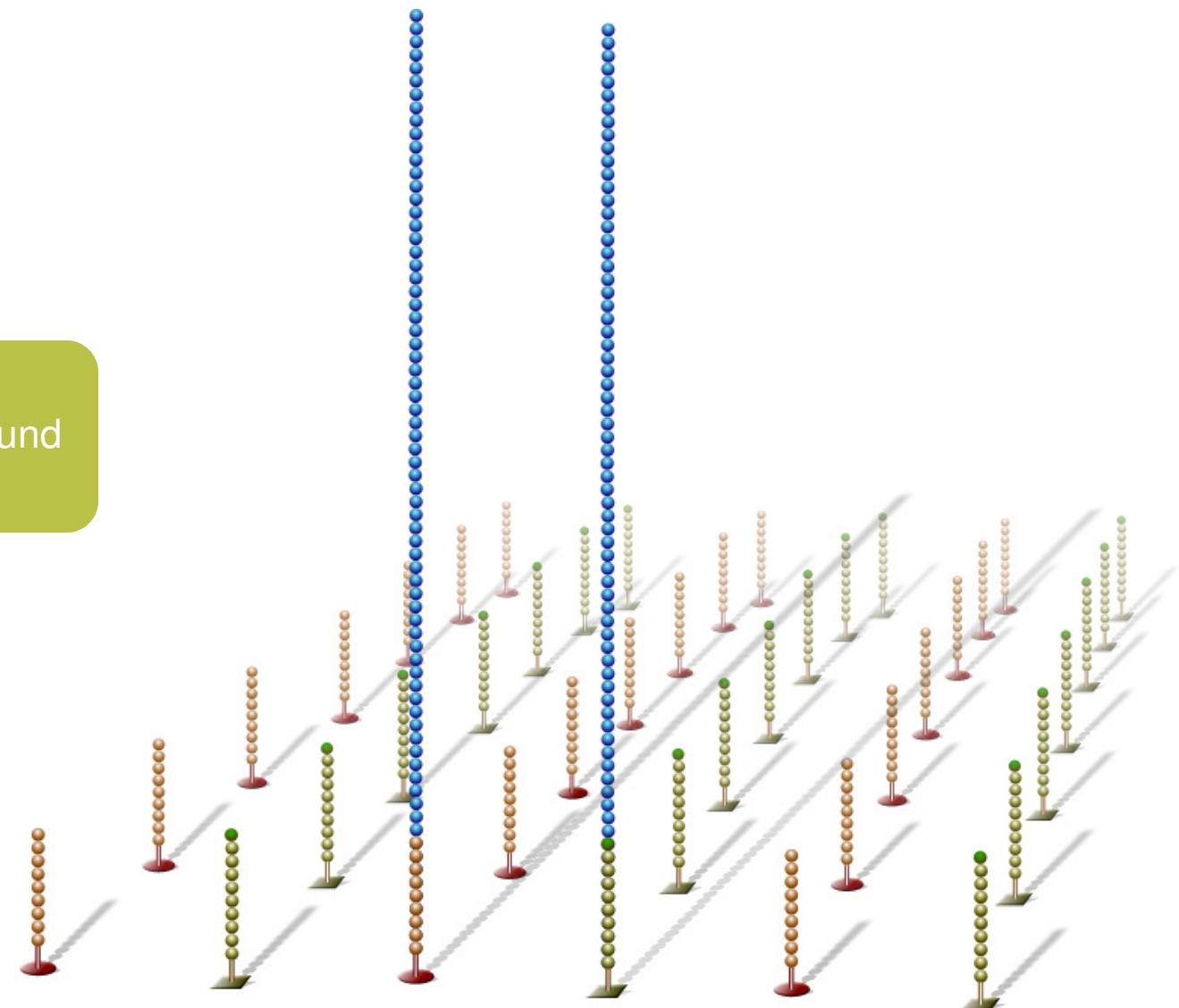
Double-stranded bridge is formed



# NGS Illumina Clustering

Double-stranded bridge is denatured

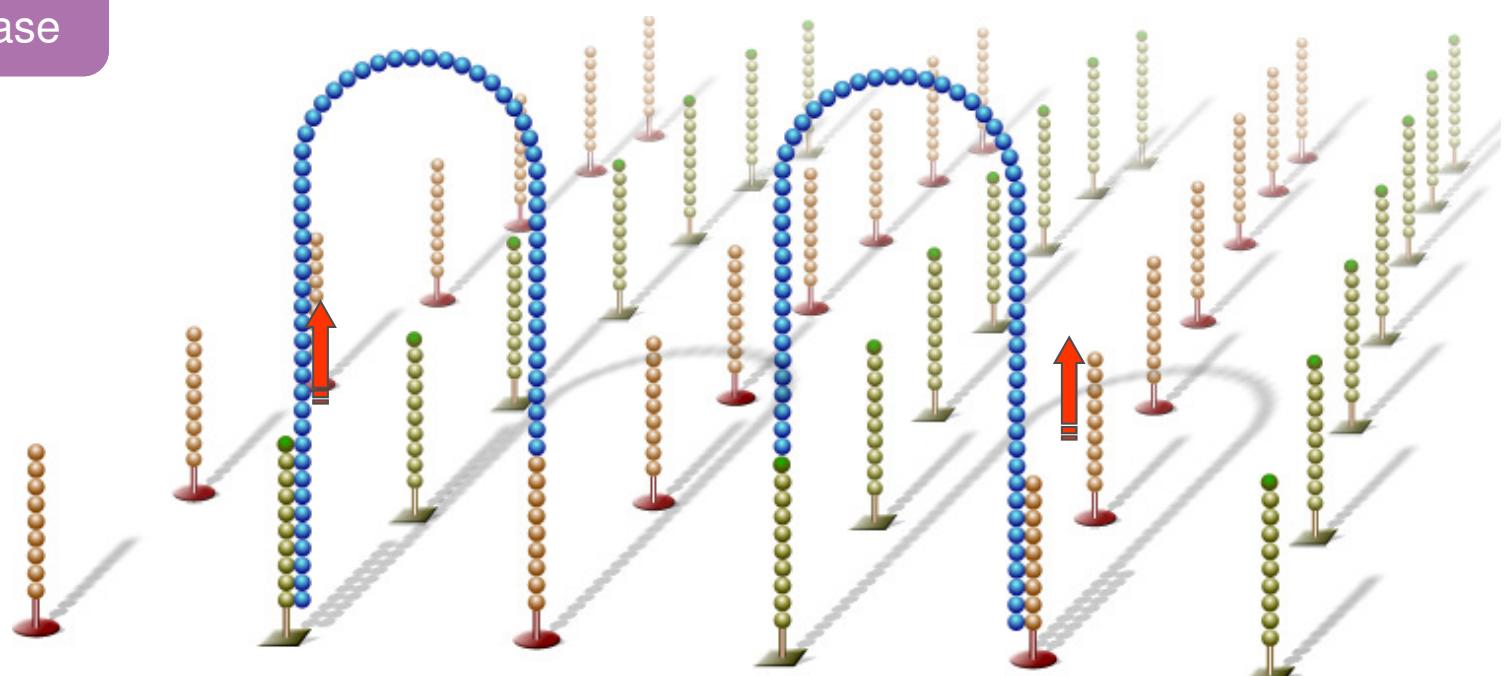
Result:  
Two copies of covalently bound single-stranded templates



# NGS Illumina Clustering

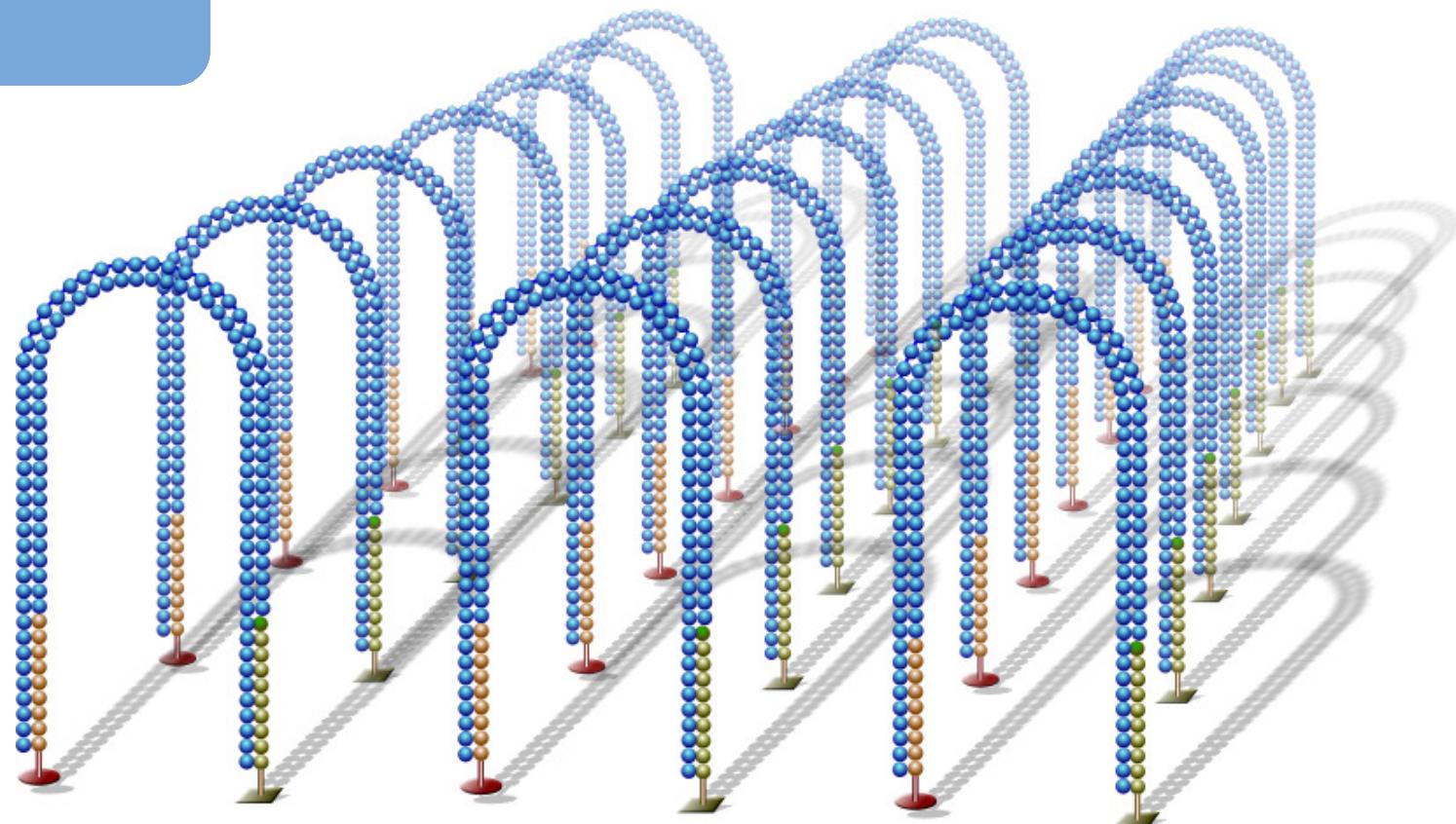
Single-stranded molecules flip over to hybridize to adjacent primers

Hybridized primer is extended by polymerase



# NGS Illumina Clustering

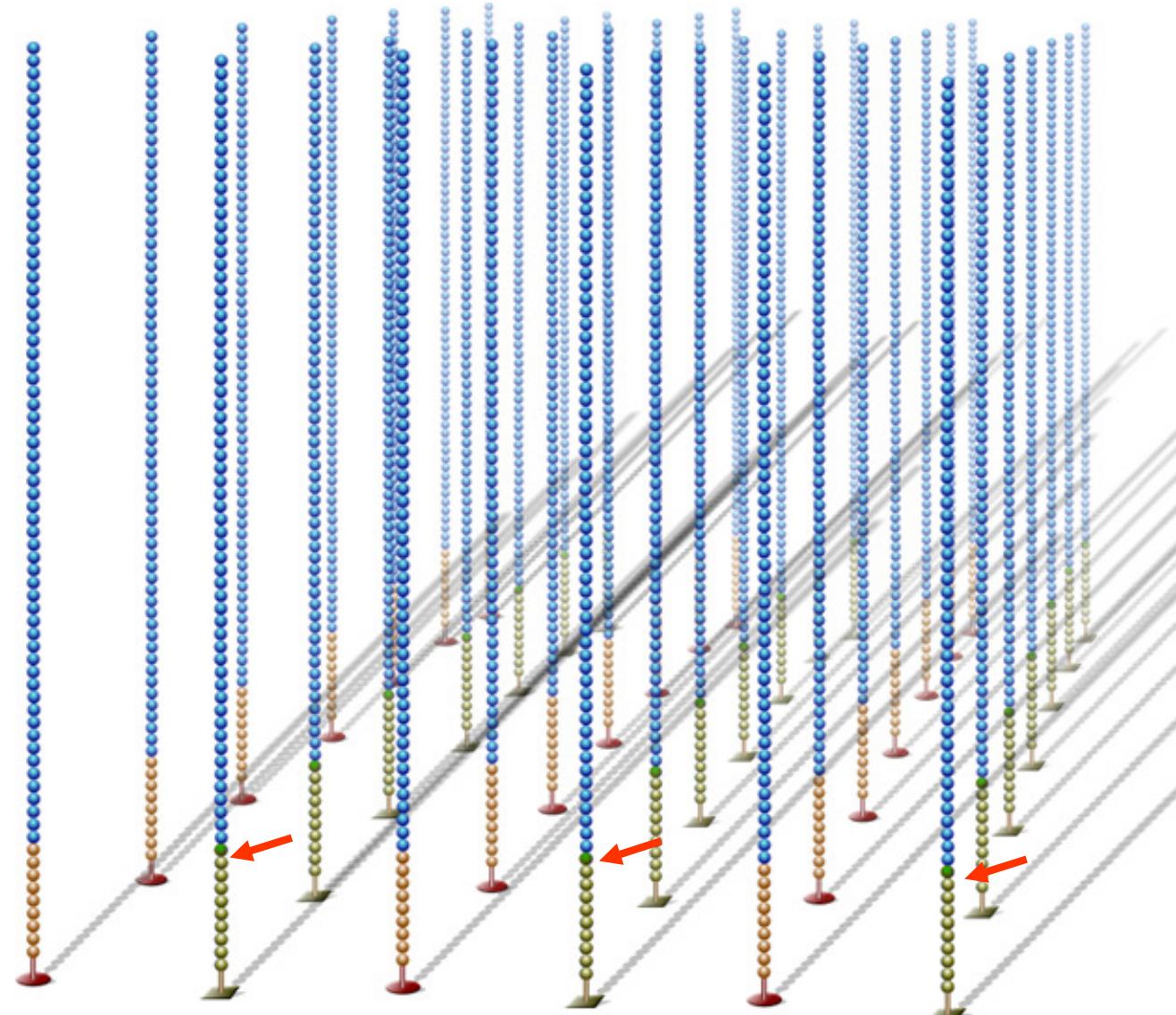
Bridge amplification cycle  
repeated until multiple  
bridges are formed



# NGS Illumina Clustering



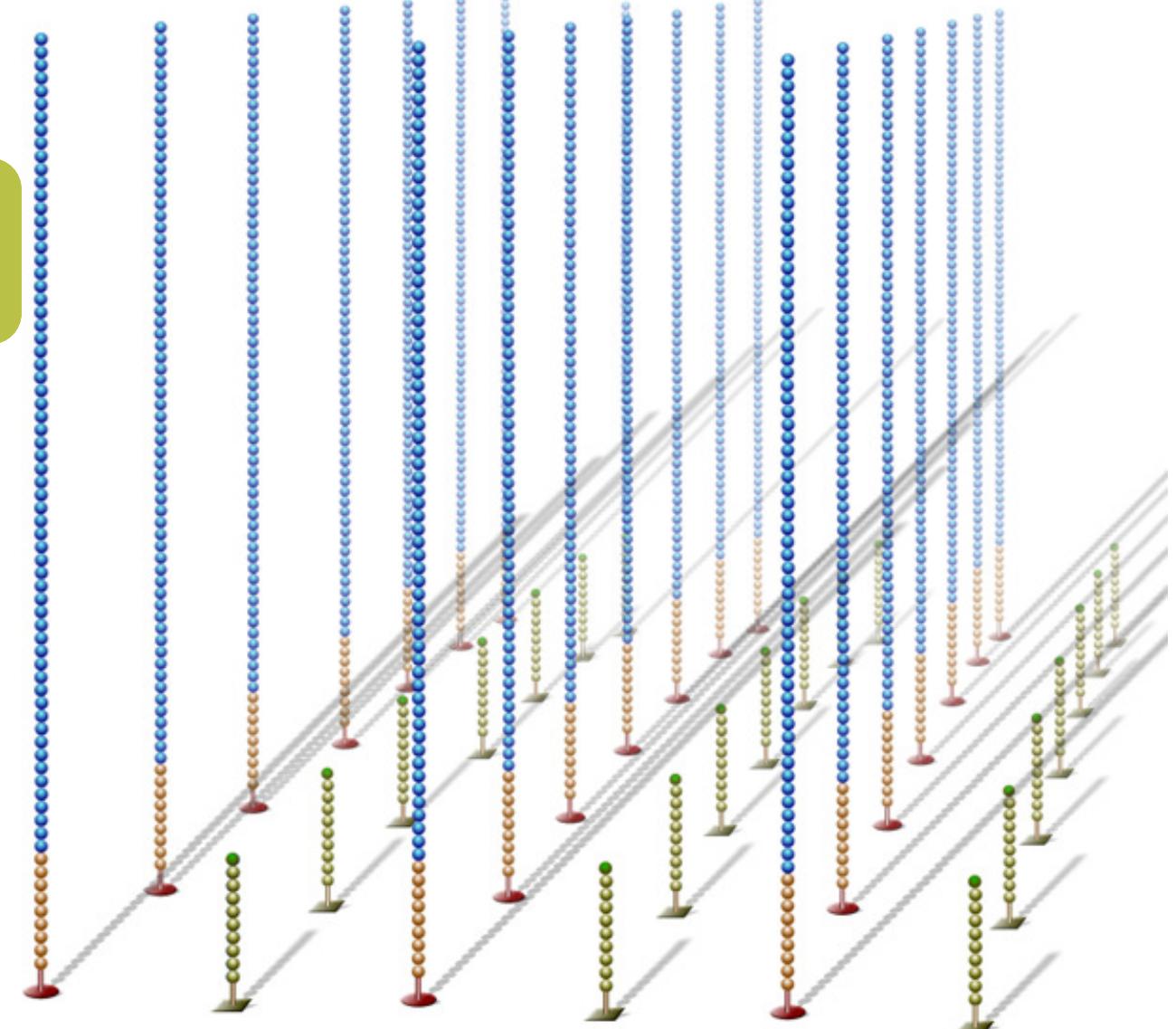
dsDNA bridges are denatured



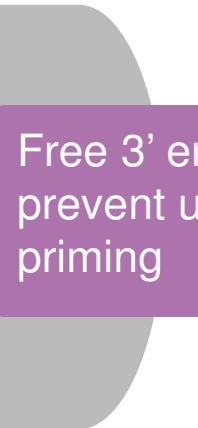
# NGS Illumina Clustering



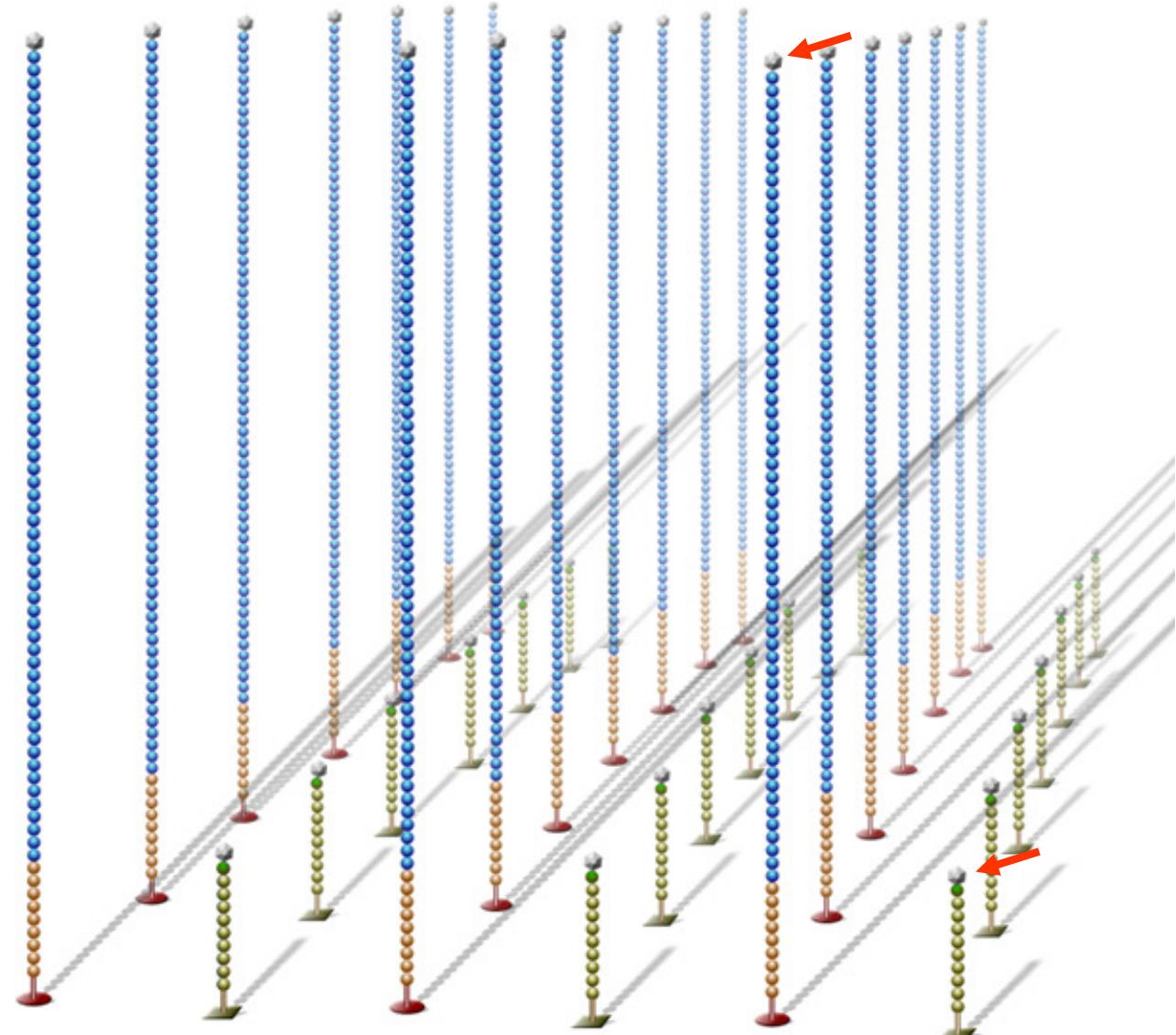
Reverse strands cleaved and washed away, leaving a cluster with forward strands only



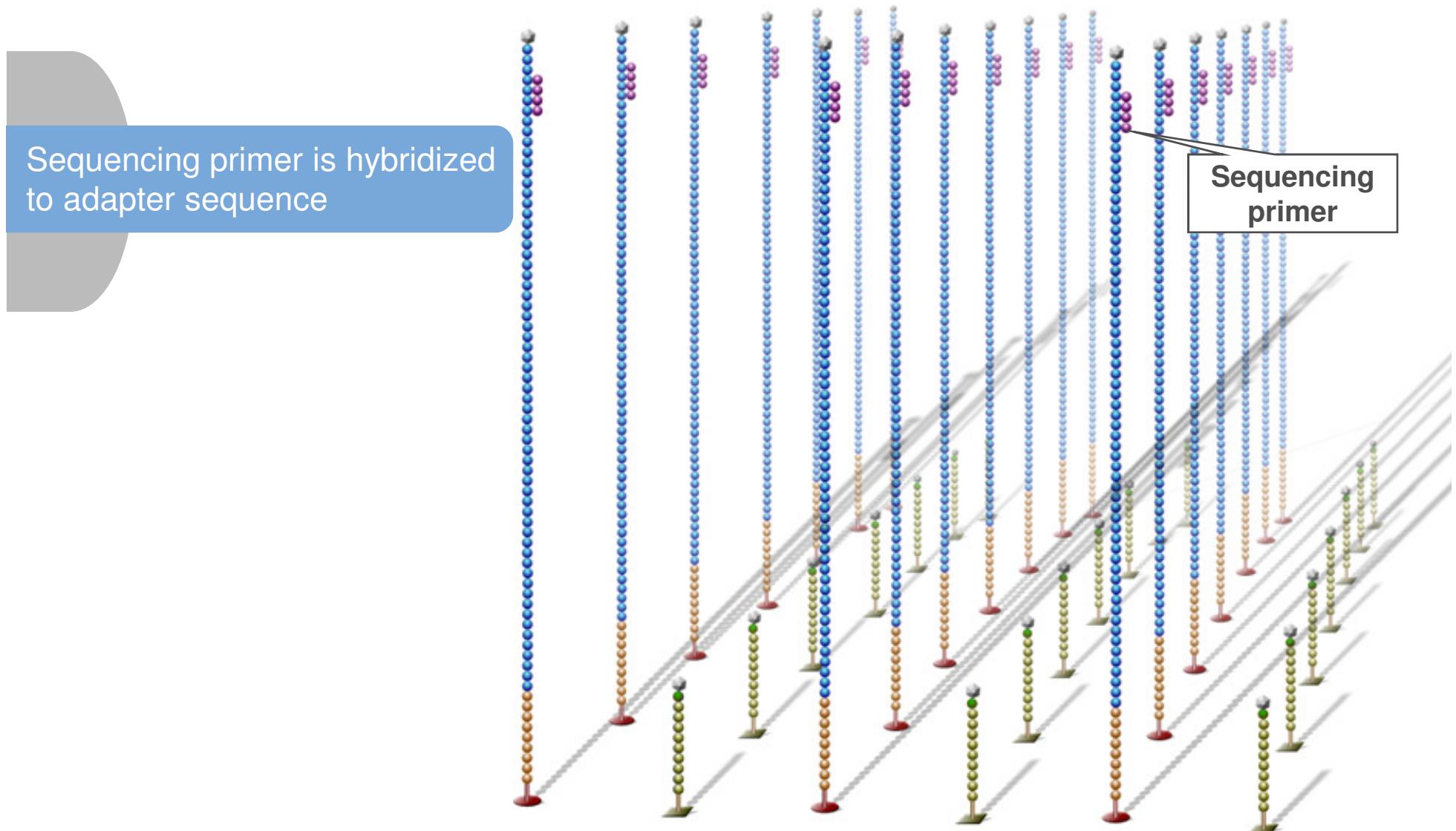
# NGS Illumina Clustering



Free 3' ends are blocked to prevent unwanted DNA priming



# NGS Illumina Clustering



# NGS workflow

---

- Library preparation

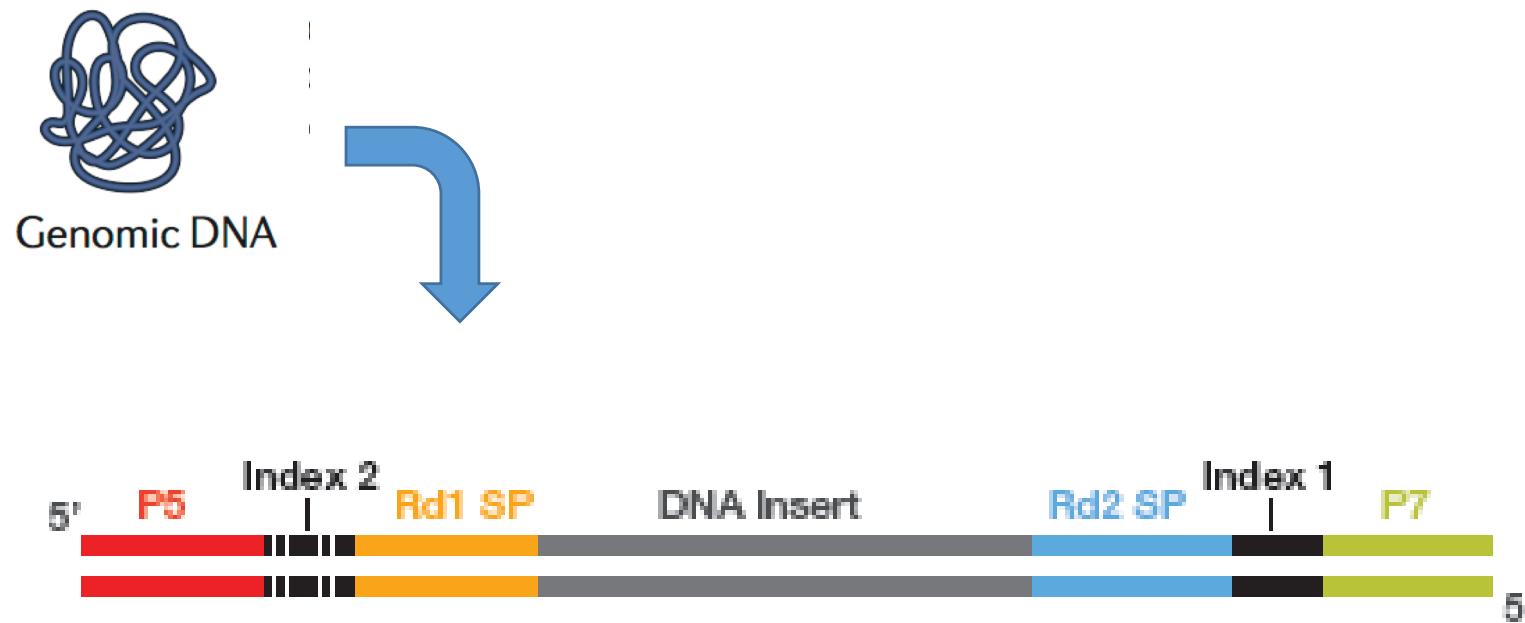
- Sequencing

- clustering
- sequencing



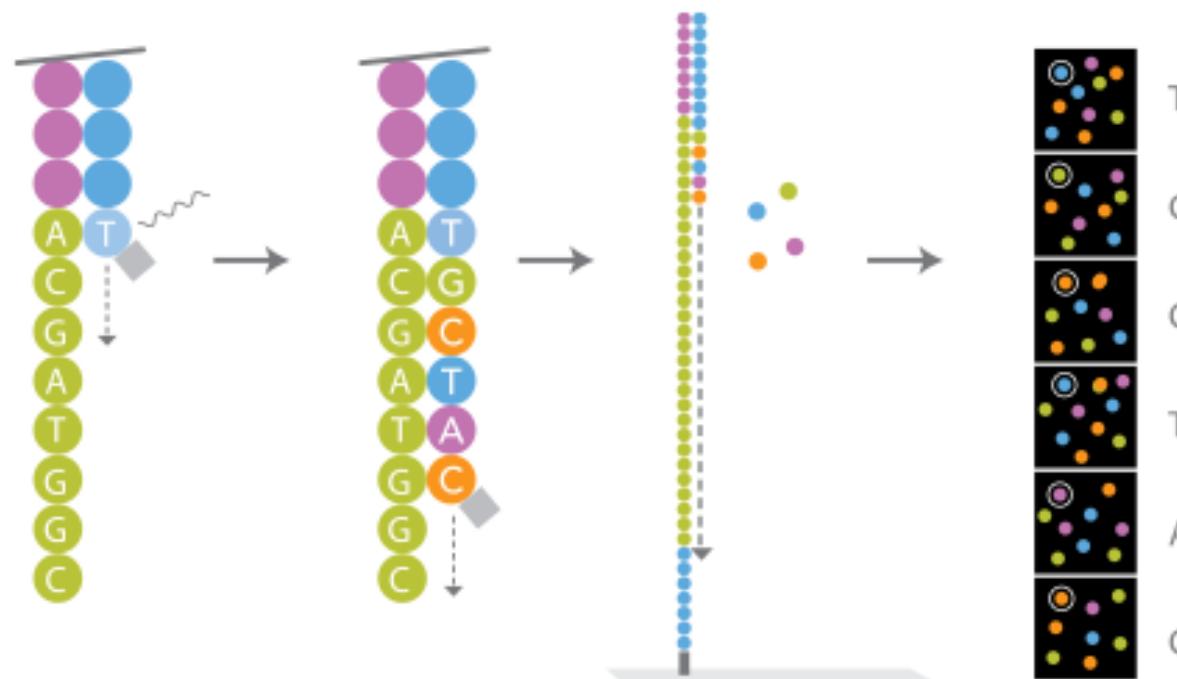
- Data-analysis

# NGS Library Prep



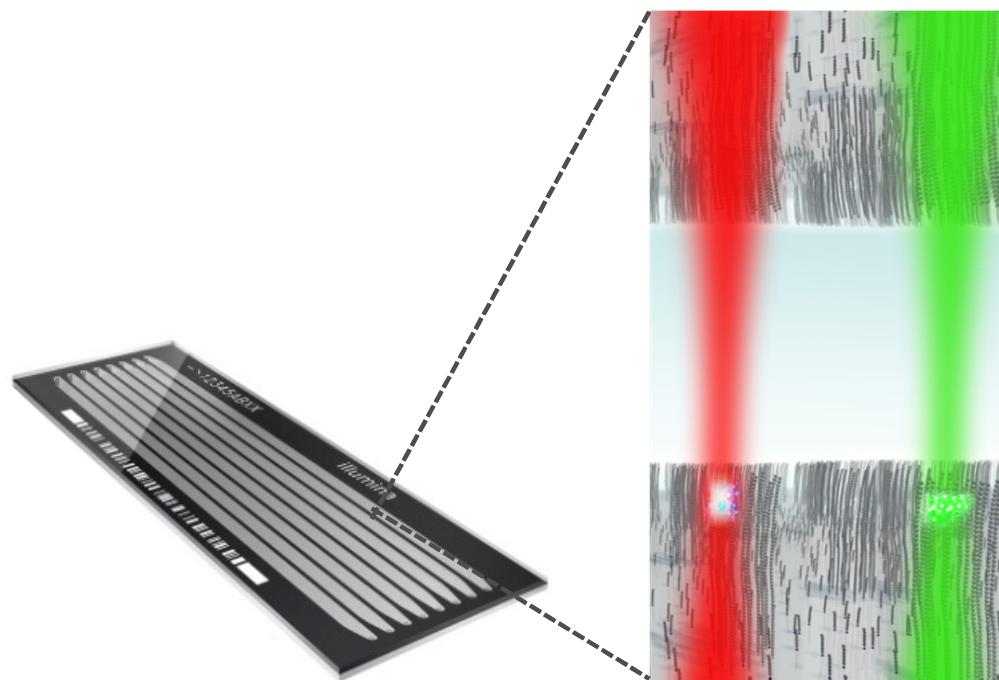
# NGS Illumina Sequencing

- Sequencing By Synthesis (SBS)



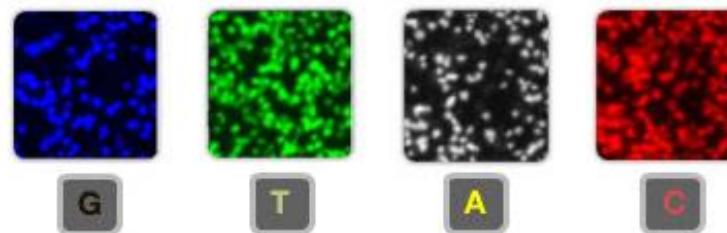
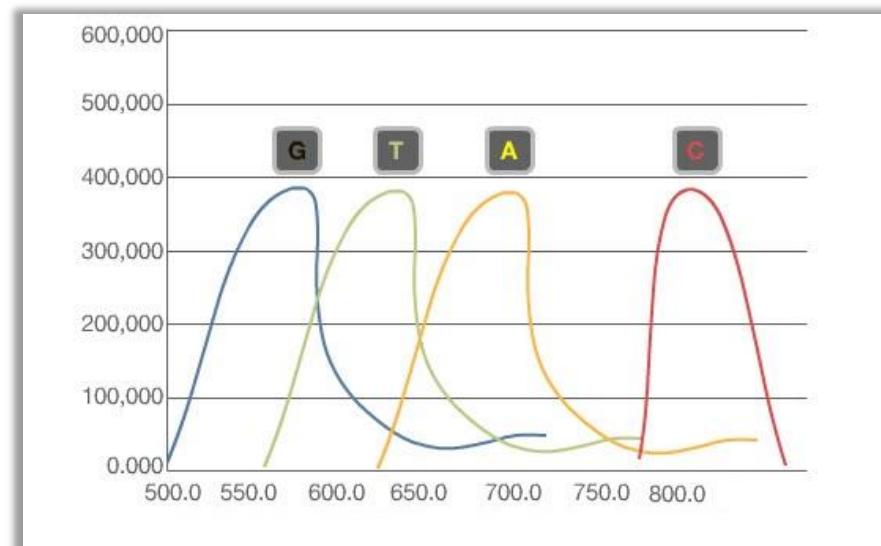
# NGS Illumina Sequencing

- Imaging HiSeq/MiSeq



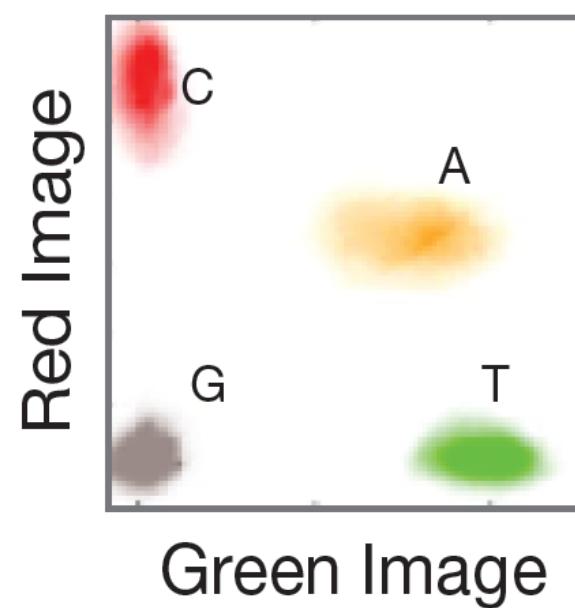
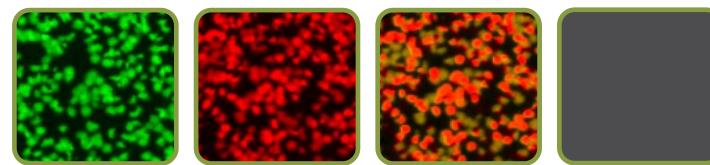
# NGS Illumina Sequencing

- Imaging MiSeq, HiSeq2500, HiSeq4000



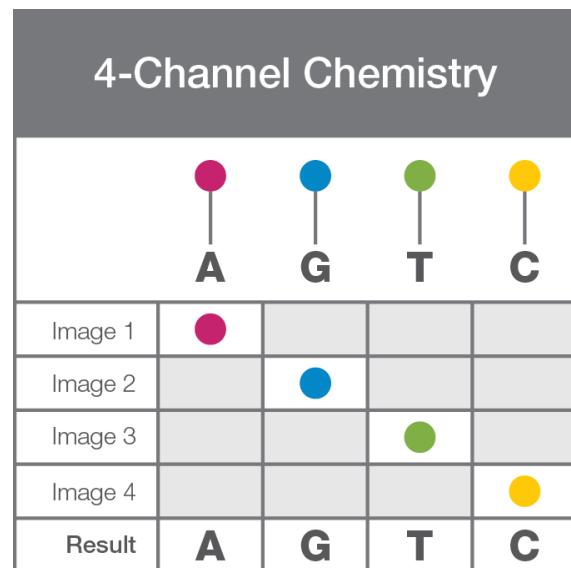
# NGS Illumina Sequencing

- Imaging NextSeq, NovaSeq



# NGS Illumina Sequencing

- Different SBS dyes



# Overview of Illumina sequencers



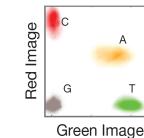
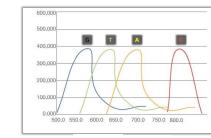
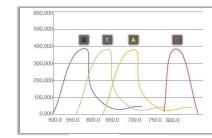
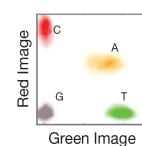
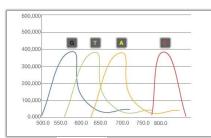
MiSeq

NextSeq

HiSeq2500

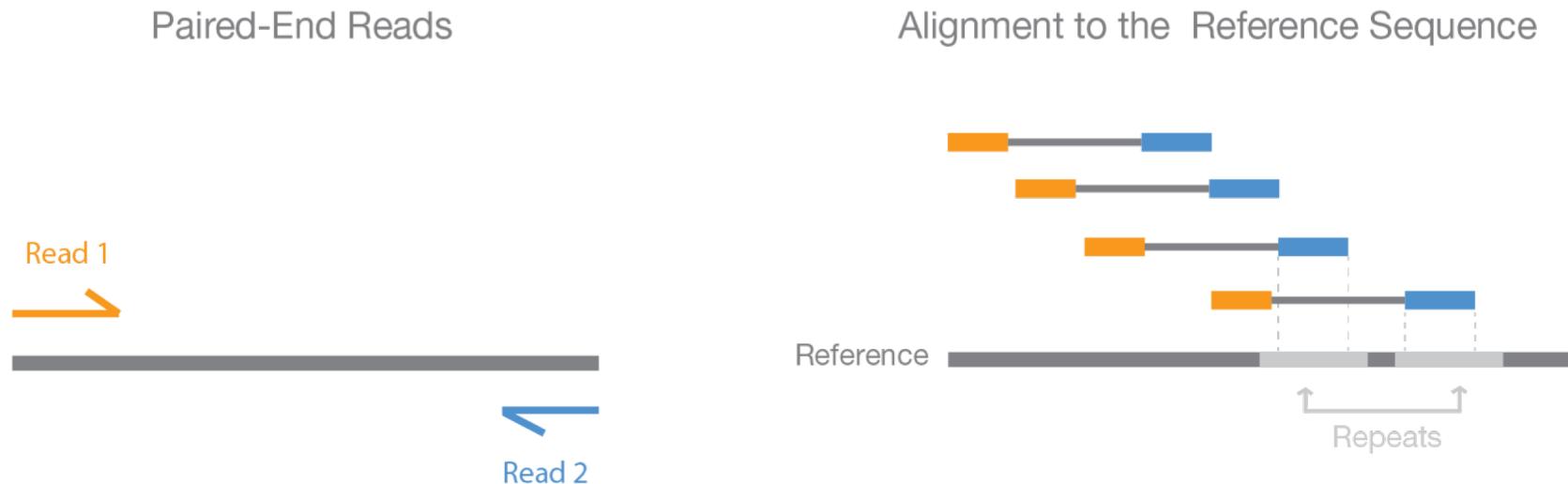
HiSeq4000

NovaSeq

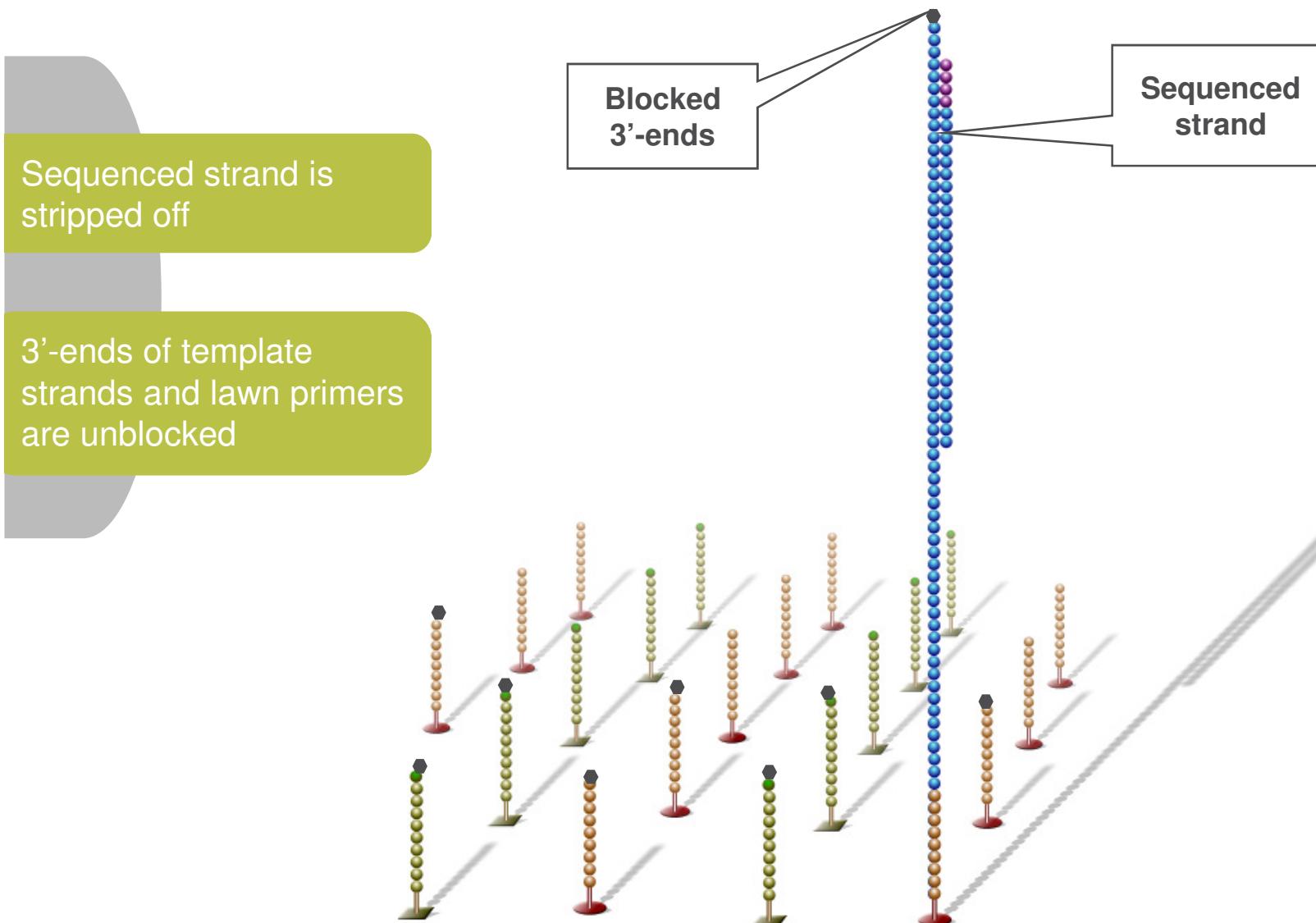


# NGS Illumina Sequencing

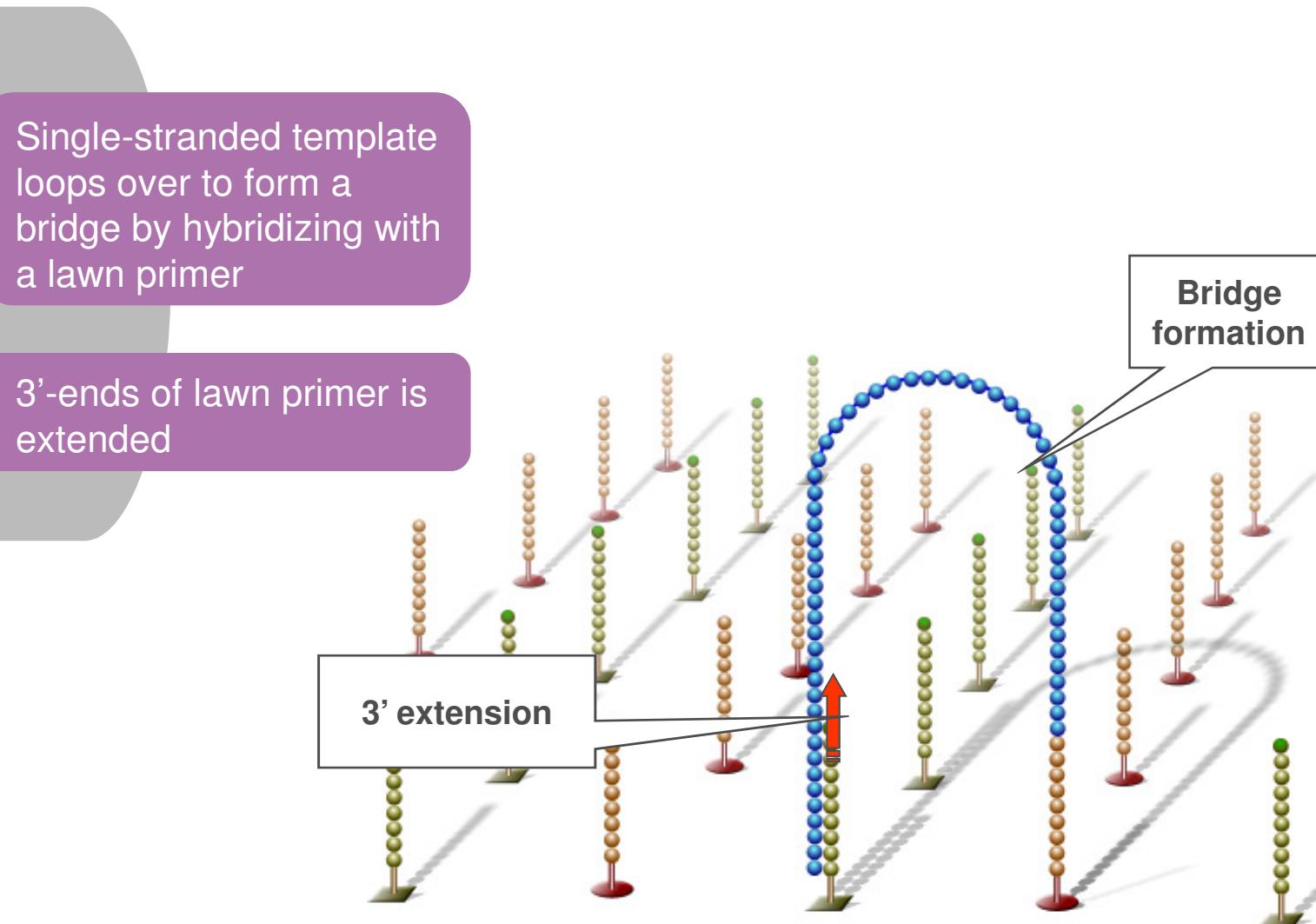
- Paired-end sequencing



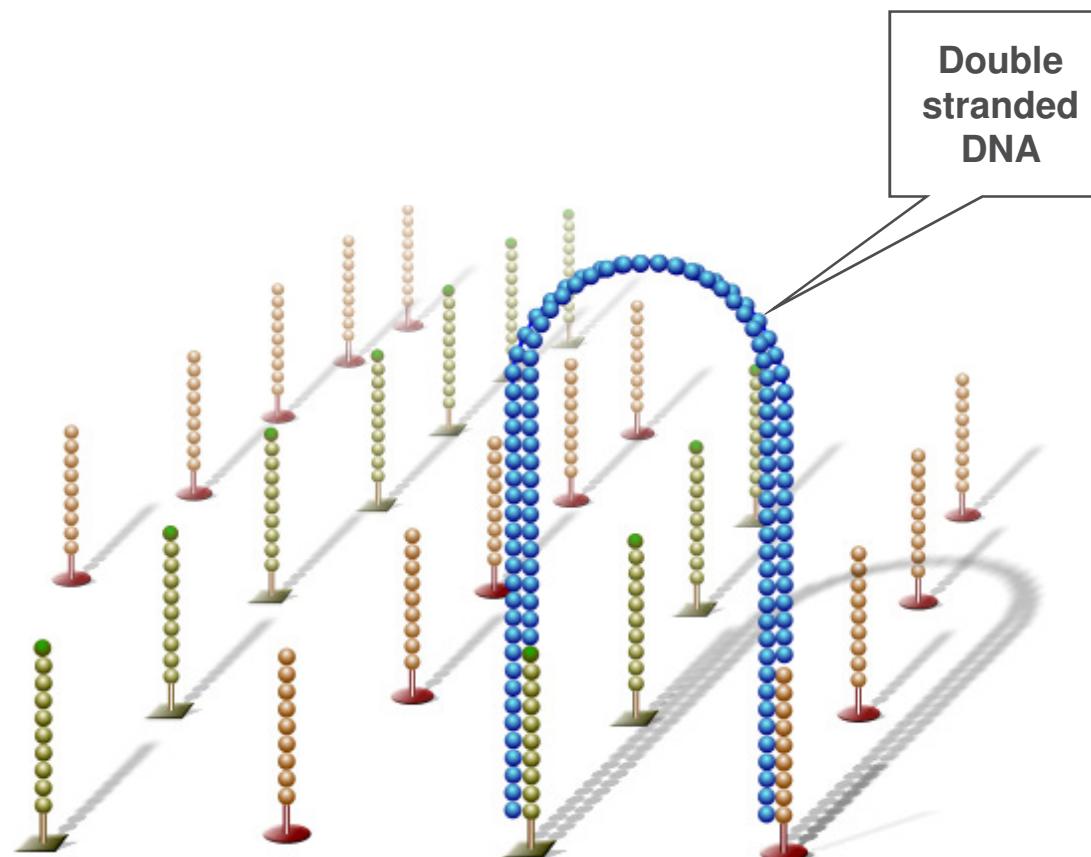
# NGS Illumina Sequencing



# NGS Illumina Sequencing

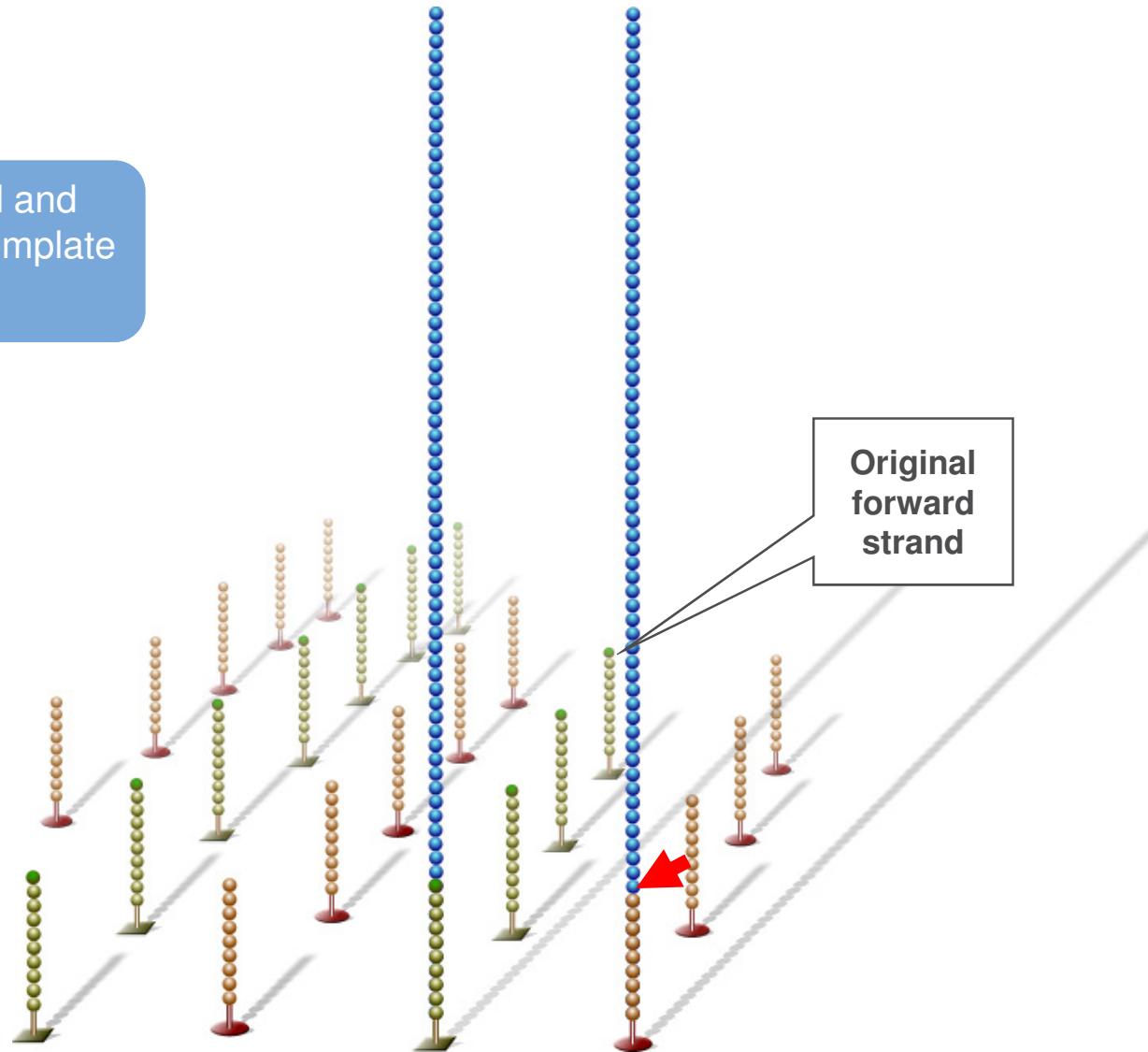


# NGS Illumina Sequencing

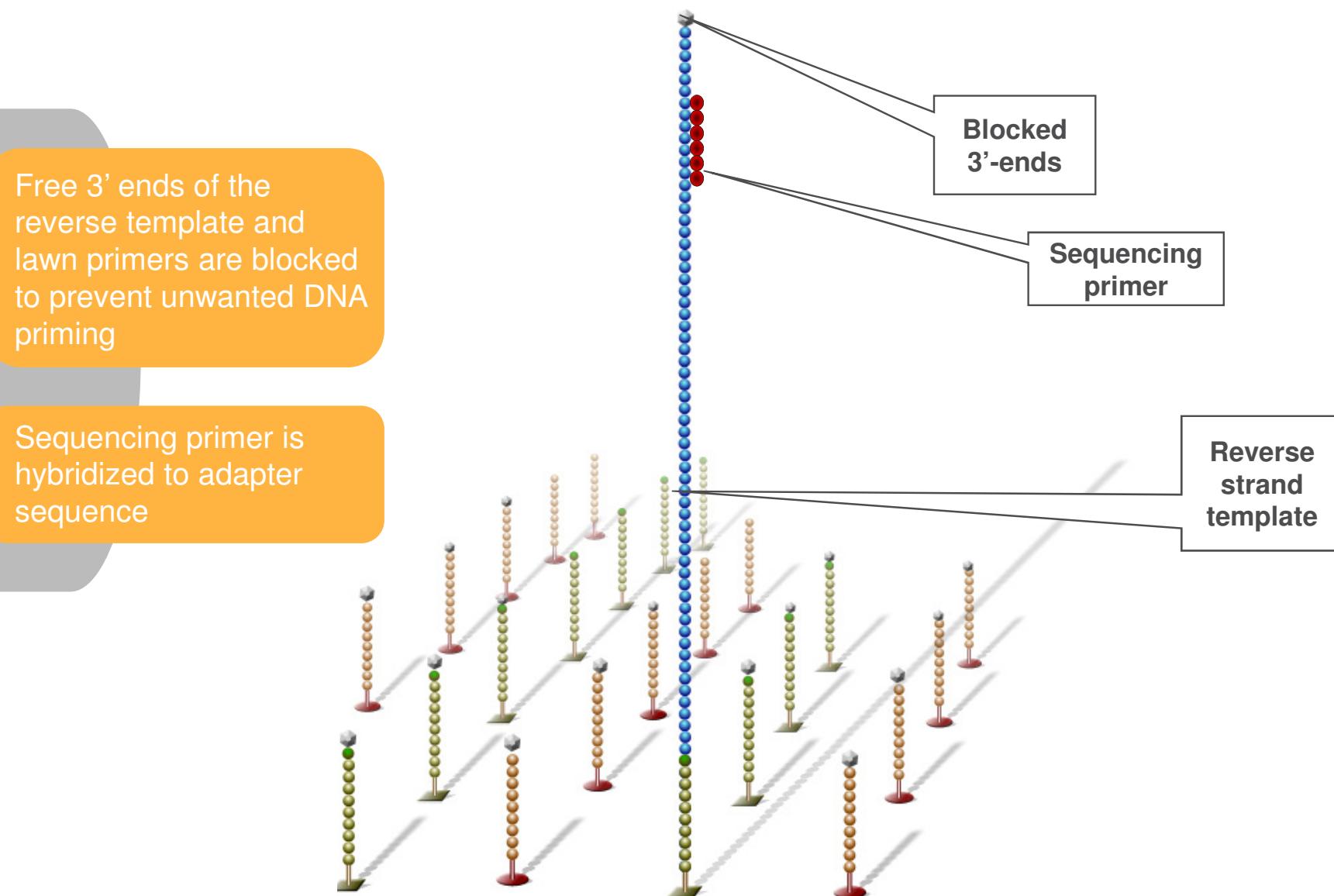


# NGS Illumina Sequencing

Bridges are linearized and the original forward template is cleaved

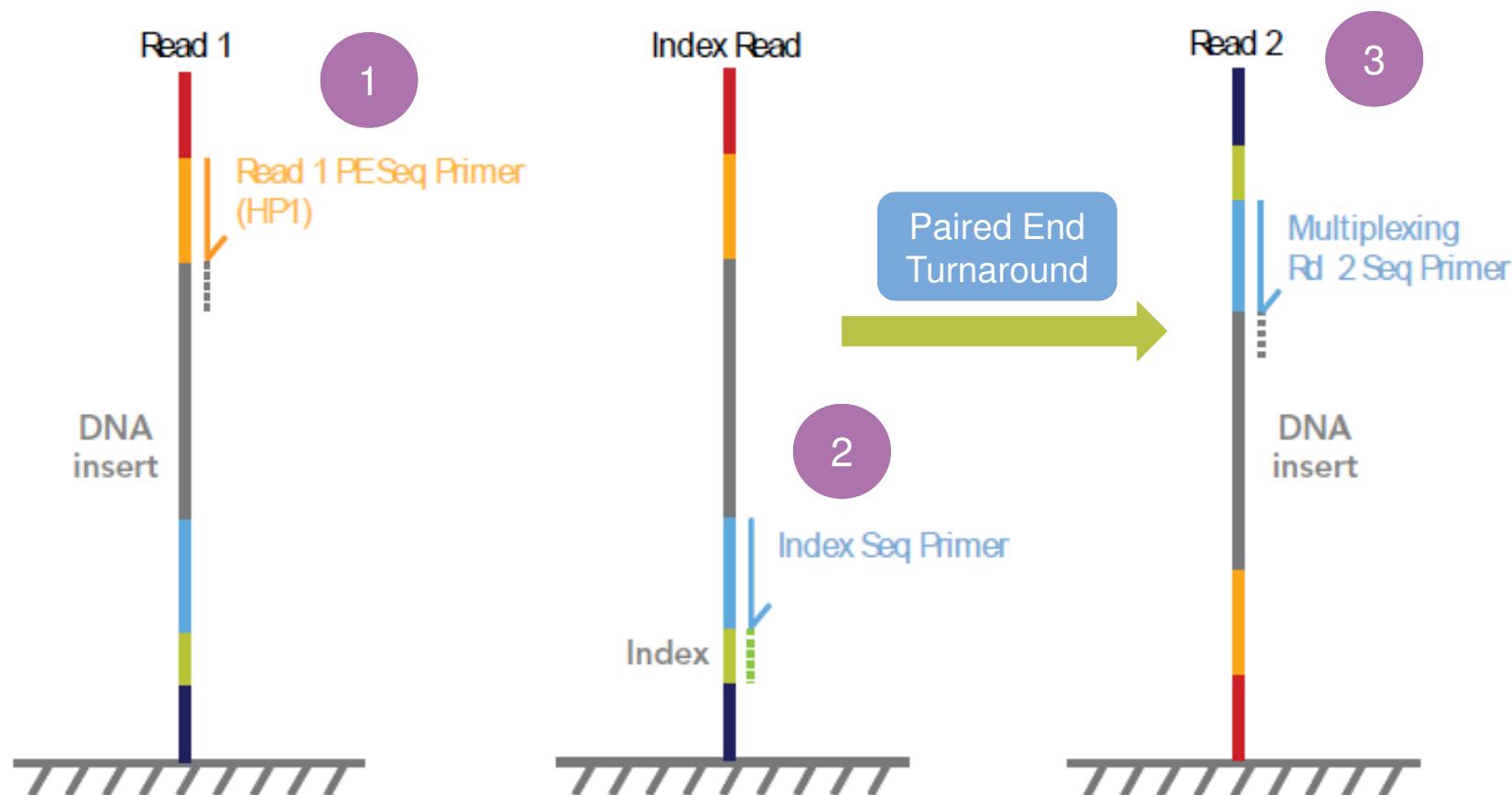


# NGS Illumina Sequencing



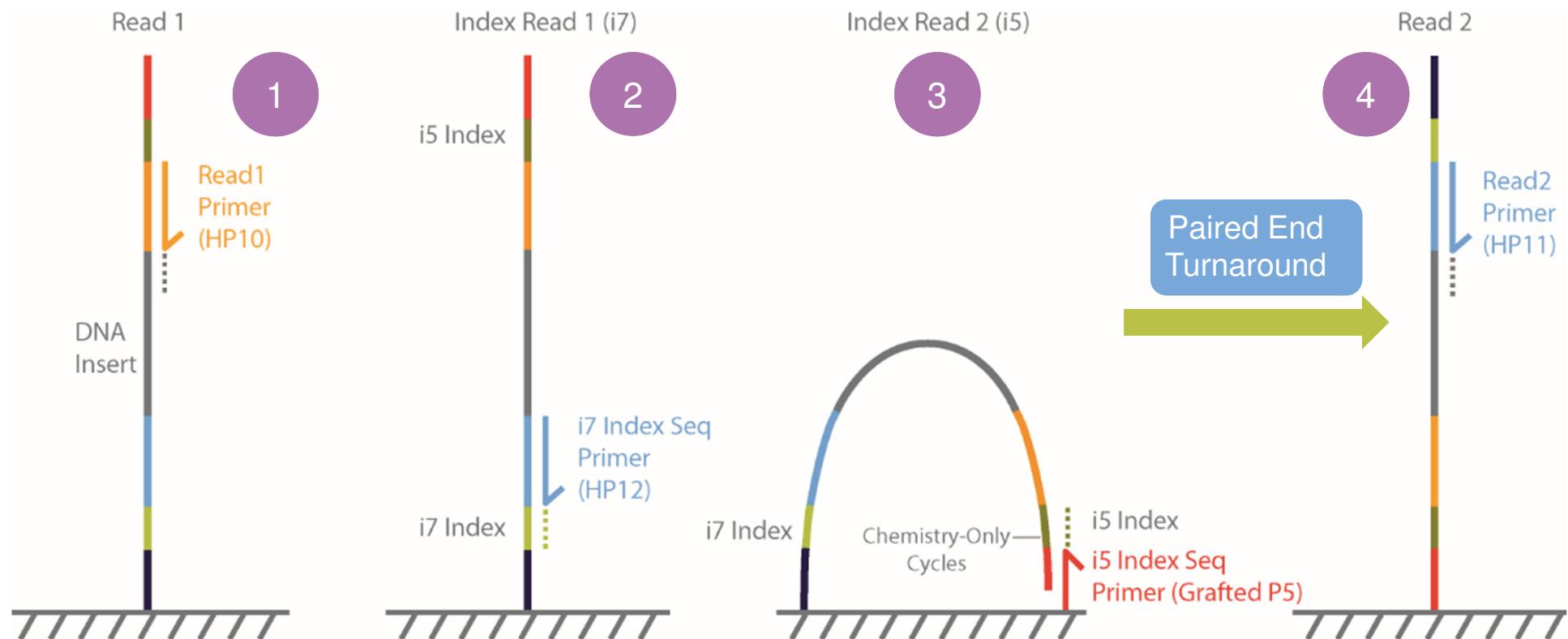
# NGS Illumina Sequencing

- Single index sequencing



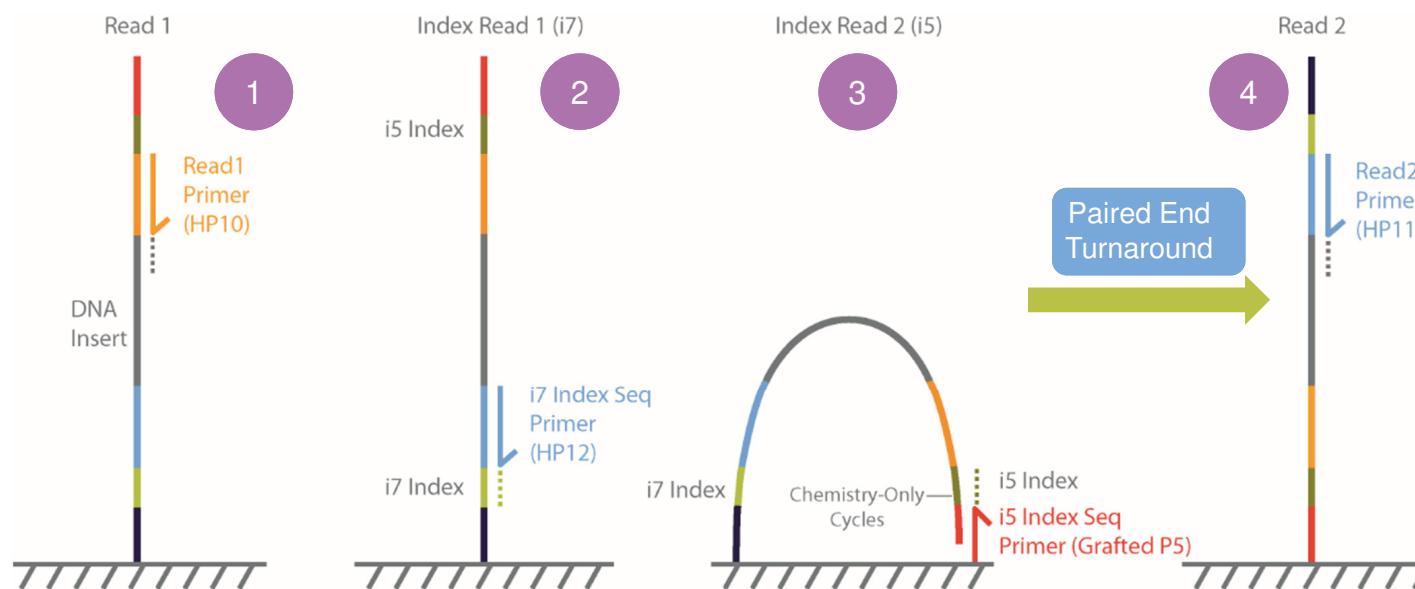
# NGS Illumina Sequencing

- Dual index sequencing



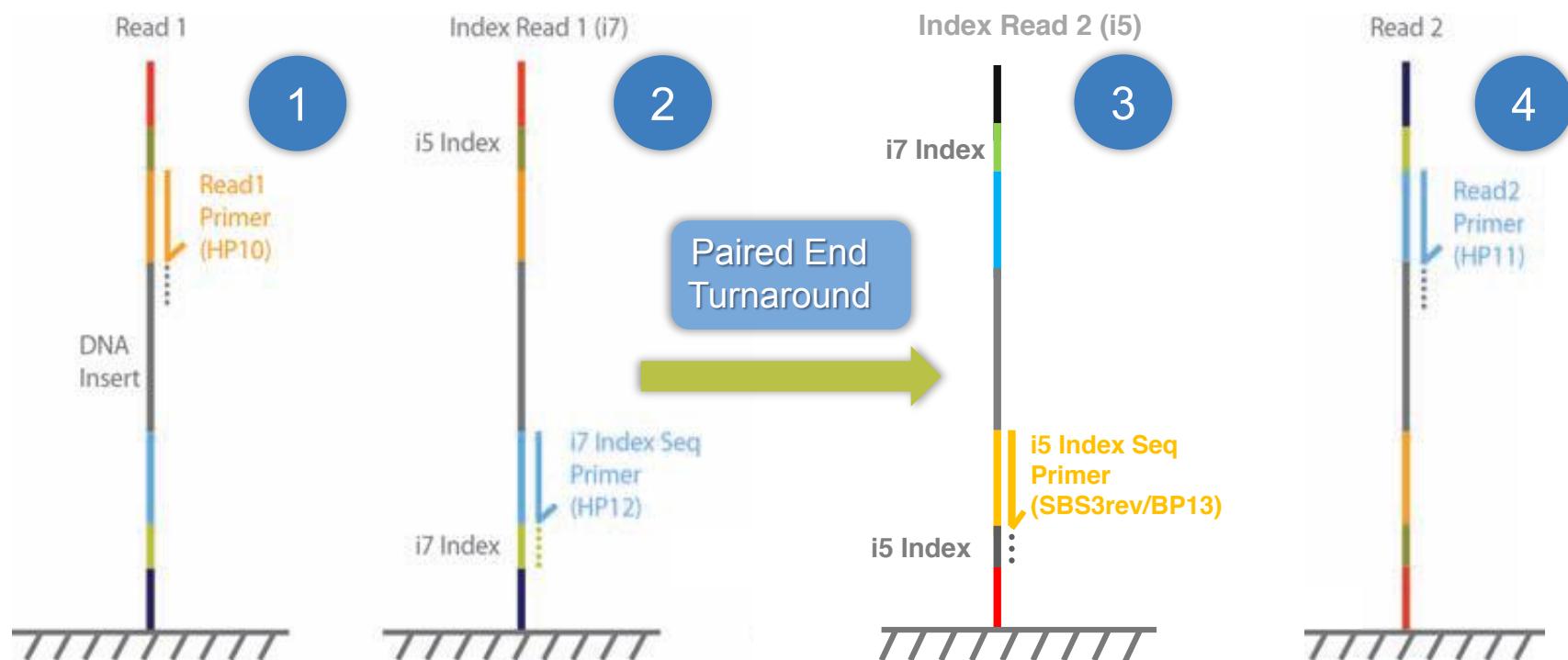
# NGS Illumina Sequencing

- Custom Sequencing Primers



# NGS Illumina Sequencing

- Dual index sequencing **NextSeq & HiSeq4000 PE**



# NGS Illumina Sequencing

## • Quality Control

### Run Summary

Level	Yield Total (G)	Projected Total Yield (G)	Aligned (%)	Error Rate (%)	Intensity Cycle 1	% >= Q30
Read 1	0,3	0,3	33,59	1,76	76	95,3
Read 2 (I)	0,0	0,0	0,00	0,00	77	64,9
Read 3 (I)	0,0	0,0	0,00	0,00	379	97,2
Read 4	0,3	0,3	33,07	1,79	100	88,3
Total	0,6	0,6	33,33	1,78	158	91,5

### Read 1

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,108 / 0,078	1,23	1,16	95,3	0,3	250	33,59 +/- 0,22	1,76 +/- 0,03	0,16 +/- 0,08	0,20 +/- 0,08	0,26 +/- 0,06	76 +/- 8

### Read 2 (I)

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,000 / 0,000	1,23	1,16	64,9	0,0	0	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	77 +/- 0

### Read 3 (I)

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,000 / 0,000	1,23	1,16	97,2	0,0	0	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	379 +/- 6

### Read 4

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,147 / 0,068	1,23	1,16	88,2	0,3	250	33,07 +/- 0,61	1,79 +/- 0,01	0,20 +/- 0,10	0,25 +/- 0,08	0,30 +/- 0,06	100 +/- 5

# NGS Illumina Sequencing

- Quality Control

Phred Quality Score	Probability of Incorrect Base Call	Base Call Accuracy
10	1 in 10	90%
20	1 in 100	99%
30	1 in 1,000	99.9%
40	1 in 10,000	99.99%
50	1 in 100,000	99.999%

# NGS Illumina Sequencing

## • Quality Control

### Run Summary

Level	Yield Total (G)	Projected Total Yield (G)	Aligned (%)	Error Rate (%)	Intensity Cycle 1	% >= Q30
Read 1	0,3	0,3	33,59	1,76	76	95,3
Read 2 (I)	0,0	0,0	0,00	0,00	77	64,9
Read 3 (I)	0,0	0,0	0,00	0,00	379	97,2
Read 4	0,3	0,3	33,07	1,79	100	88,3
Total	0,6	0,6	33,33	1,78	158	91,5

### Read 1

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,108 / 0,078	1,23	1,16	95,3	0,3	250	33,59 +/- 0,22	1,76 +/- 0,03	0,16 +/- 0,08	0,20 +/- 0,08	0,26 +/- 0,06	76 +/- 8

### Read 2 (I)

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,000 / 0,000	1,23	1,16	64,9	0,0	0	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	77 +/- 0

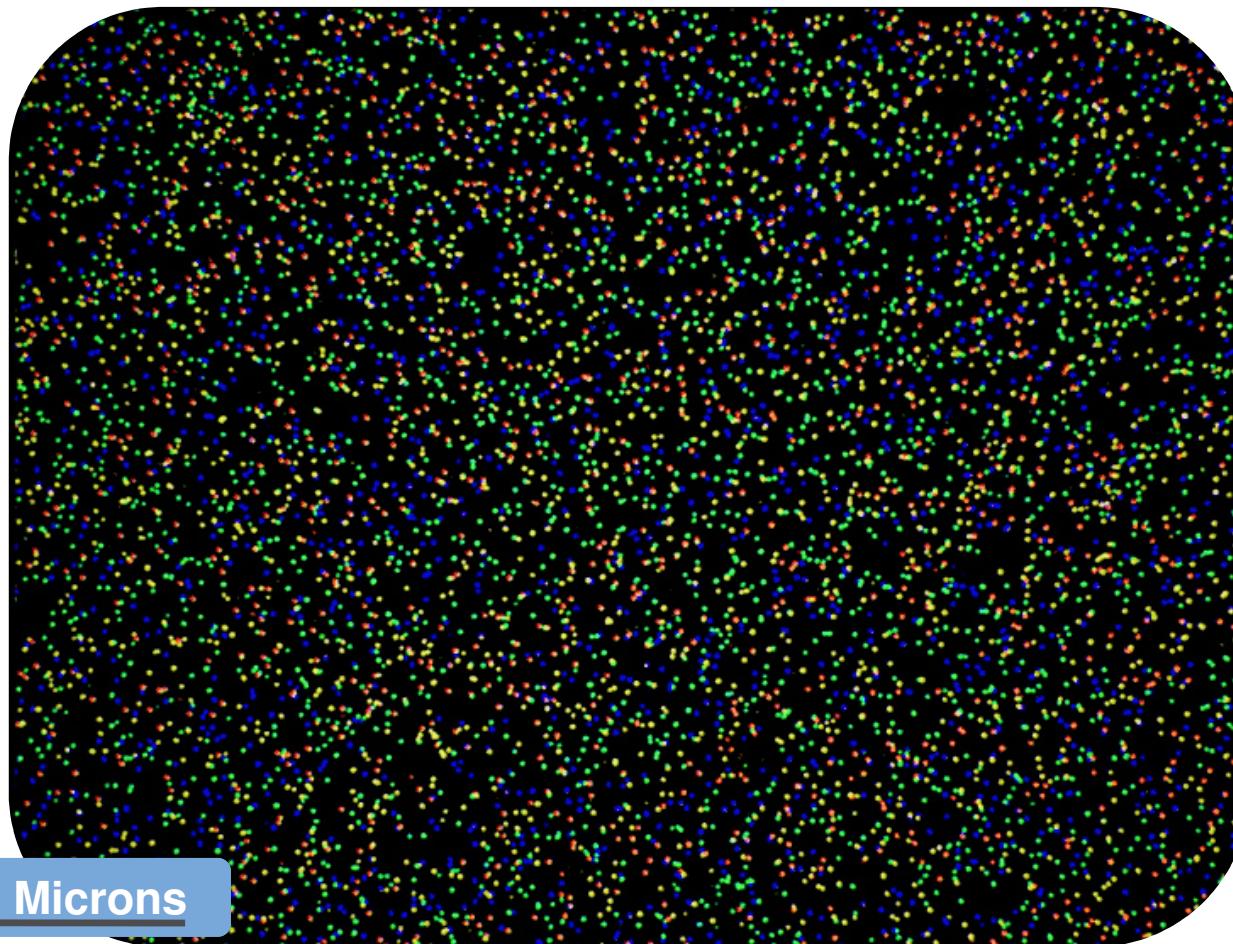
### Read 3 (I)

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,000 / 0,000	1,23	1,16	97,2	0,0	0	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	379 +/- 6

### Read 4

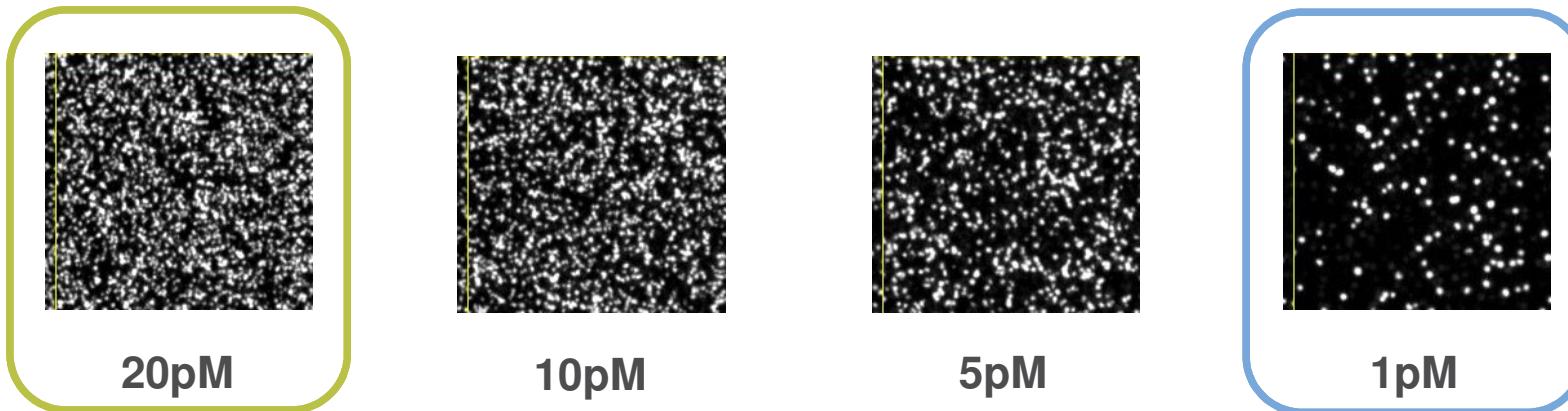
Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,147 / 0,068	1,23	1,16	88,3	0,3	250	33,07 +/- 0,61	1,79 +/- 0,01	0,20 +/- 0,10	0,25 +/- 0,08	0,30 +/- 0,06	100 +/- 5

# NGS Illumina Clustering



# NGS Illumina Clustering

- Clusterdensity – non patterned flow cells



# NGS Illumina Sequencing

## • Quality Control

### Run Summary

Level	Yield Total (G)	Projected Total Yield (G)	Aligned (%)	Error Rate (%)	Intensity Cycle 1	% >= Q30
Read 1	0,3	0,3	33,59	1,76	76	95,3
Read 2 (I)	0,0	0,0	0,00	0,00	77	64,9
Read 3 (I)	0,0	0,0	0,00	0,00	379	97,2
Read 4	0,3	0,3	33,07	1,79	100	88,3
Total	0,6	0,6	33,33	1,78	158	91,5

### Read 1

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,108 / 0,078	1,23	1,16	95,3	0,3	250	33,59 +/- 0,22	1,76 +/- 0,03	0,16 +/- 0,08	0,20 +/- 0,08	0,26 +/- 0,06	76 +/- 8

### Read 2 (I)

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,000 / 0,000	1,23	1,16	64,9	0,0	0	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	77 +/- 0

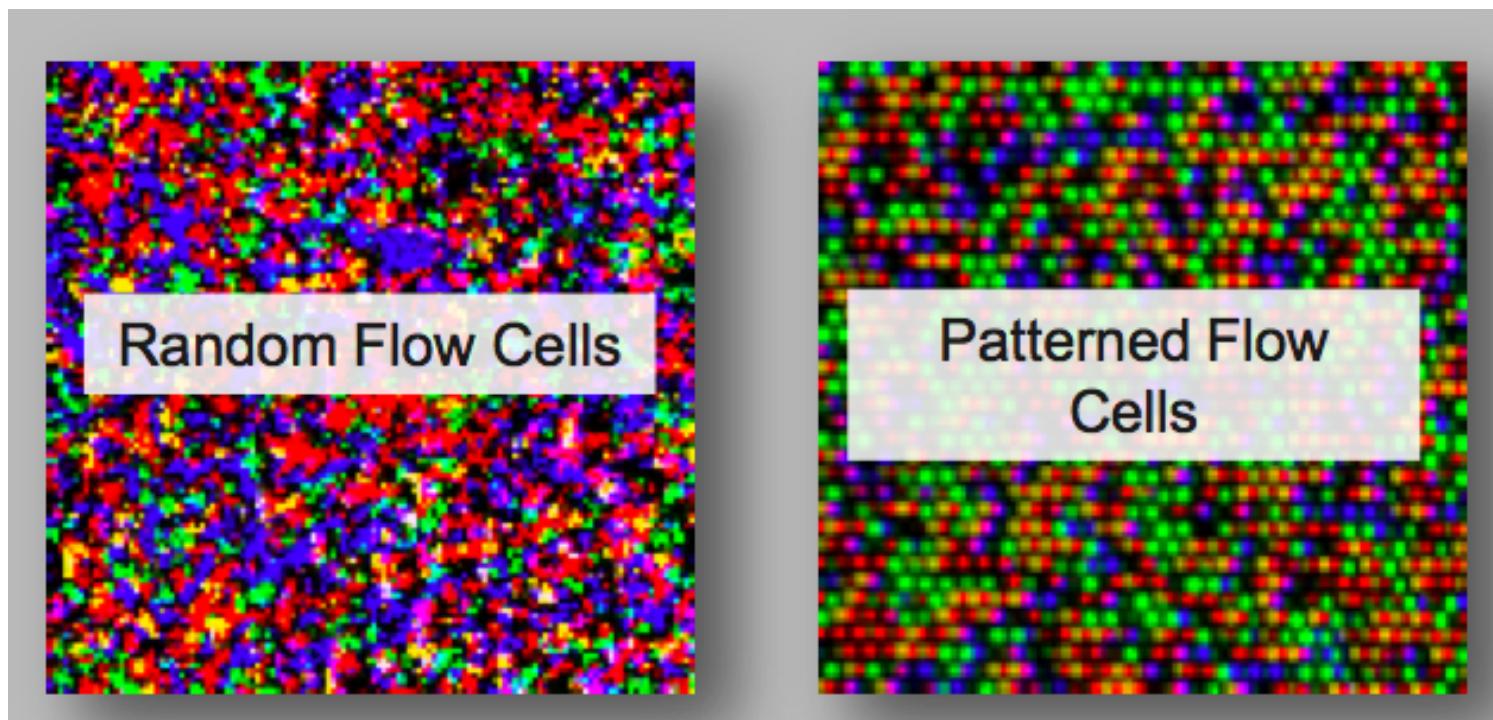
### Read 3 (I)

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,000 / 0,000	1,23	1,16	97,2	0,0	0	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	379 +/- 6

### Read 4

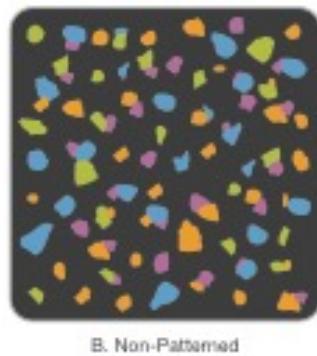
Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,147 / 0,068	1,23	1,16	88,3	0,3	250	33,07 +/- 0,61	1,79 +/- 0,01	0,20 +/- 0,10	0,25 +/- 0,08	0,30 +/- 0,06	100 +/- 5

# NGS Clusters Passing Filter

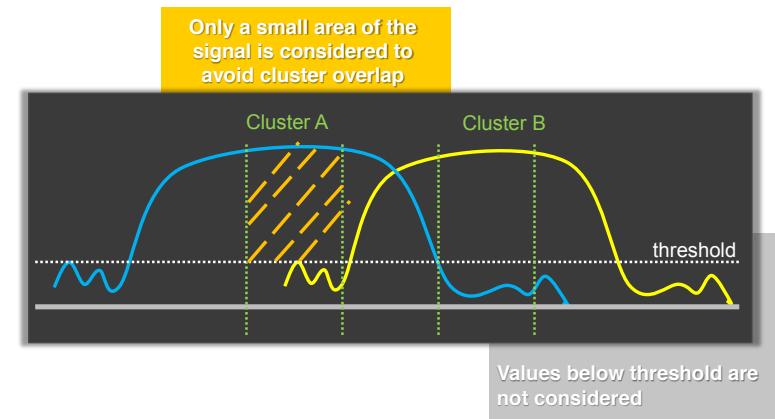


# NGS Clusters Passing Filter

- Cluster generation MiSeq, NextSeq, HiSeq2500



B. Non-Patterned



- Chastity filter



Template Generation



Chastity Filtration

# NGS Clusters Passing Filter

- Cluster generation HiSeq4000 & NovaSeq

- patterned flowcell

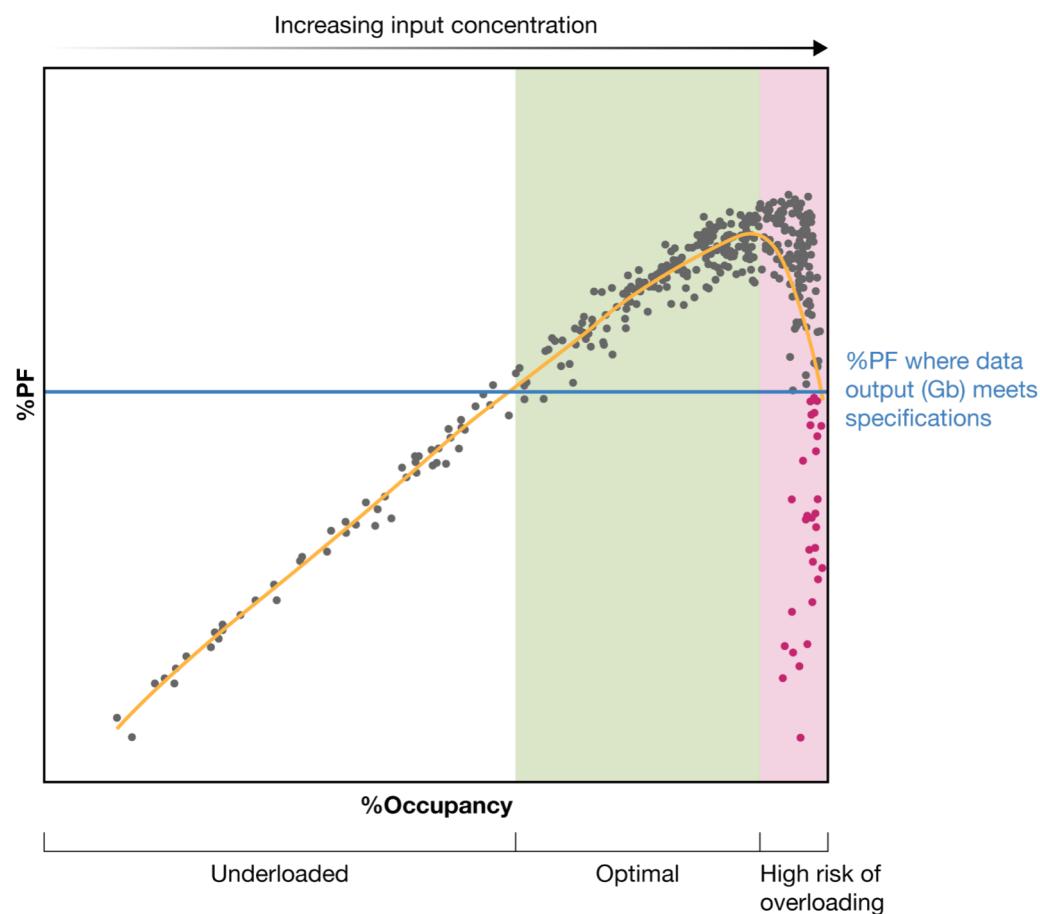
- empty well
- monoclonal well
- polyclonal well

} % occupied  
} % clusters passing filter

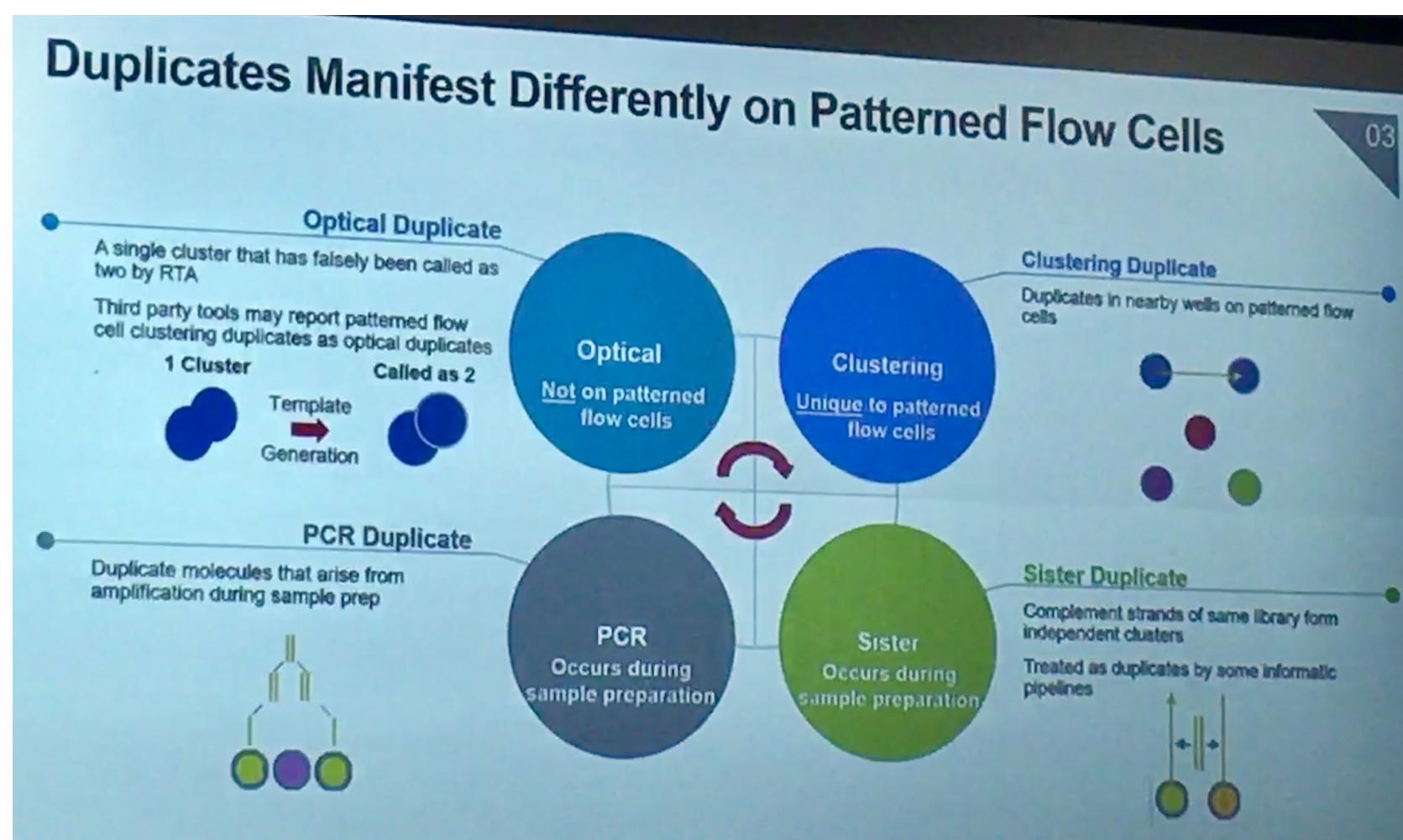


# NGS Illumina Clustering

- Clusterdensity – patterned flow cells



# PCR duplicates



# Overview of Illumina sequencers



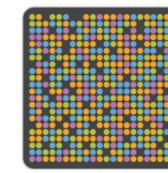
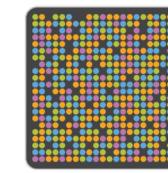
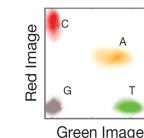
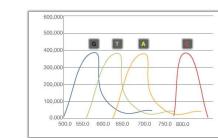
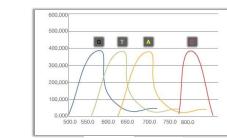
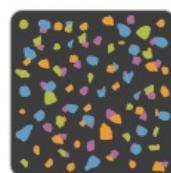
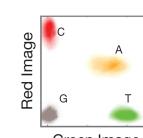
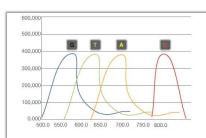
MiSeq

NextSeq

HiSeq2500

HiSeq4000

NovaSeq



# NGS Illumina Sequencing

## • Quality Control

### Run Summary

Level	Yield Total (G)	Projected Total Yield (G)	Aligned (%)	Error Rate (%)	Intensity Cycle 1	% >= Q30
Read 1	0,3	0,3	33,59	1,76	76	95,3
Read 2 (I)	0,0	0,0	0,00	0,00	77	64,9
Read 3 (I)	0,0	0,0	0,00	0,00	379	97,2
Read 4	0,3	0,3	33,07	1,79	100	88,3
Total	0,6	0,6	33,33	1,78	158	91,5

### Read 1

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,108 / 0,078	1,23	1,16	95,3	0,3	250	33,59 +/- 0,22	1,76 +/- 0,03	0,16 +/- 0,08	0,20 +/- 0,08	0,26 +/- 0,06	76 +/- 8

### Read 2 (I)

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,000 / 0,000	1,23	1,16	64,9	0,0	0	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	77 +/- 0

### Read 3 (I)

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,000 / 0,000	1,23	1,16	97,2	0,0	0	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	0,00 +/- 0,00	379 +/- 6

### Read 4

Lane	Tiles	Density (K/mm2)	Cluster PF (%)	Phas/Prephas (%)	Reads (M)	Reads PF (M)	% >= Q30	Yield (G)	Cycles Err Rated	Aligned (%)	Error Rate (%)	Error Rate 35 cycle (%)	Error Rate 75 cycle (%)	Error Rate 100 cycle (%)	Intensity Cycle 1
1	2	877 +/- 6	94,89 +/- 1,62	0,147 / 0,068	1,23	1,16	88,3	0,3	250	33,07 +/- 0,61	1,79 +/- 0,01	0,20 +/- 0,10	0,25 +/- 0,08	0,30 +/- 0,06	100 +/- 5

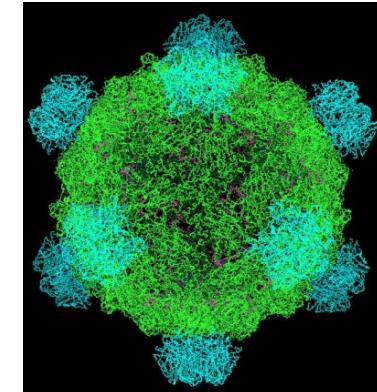
# NGS Illumina Sequencing

- PhiX

- artificial phage
- small genome: quick sequencing & alignment

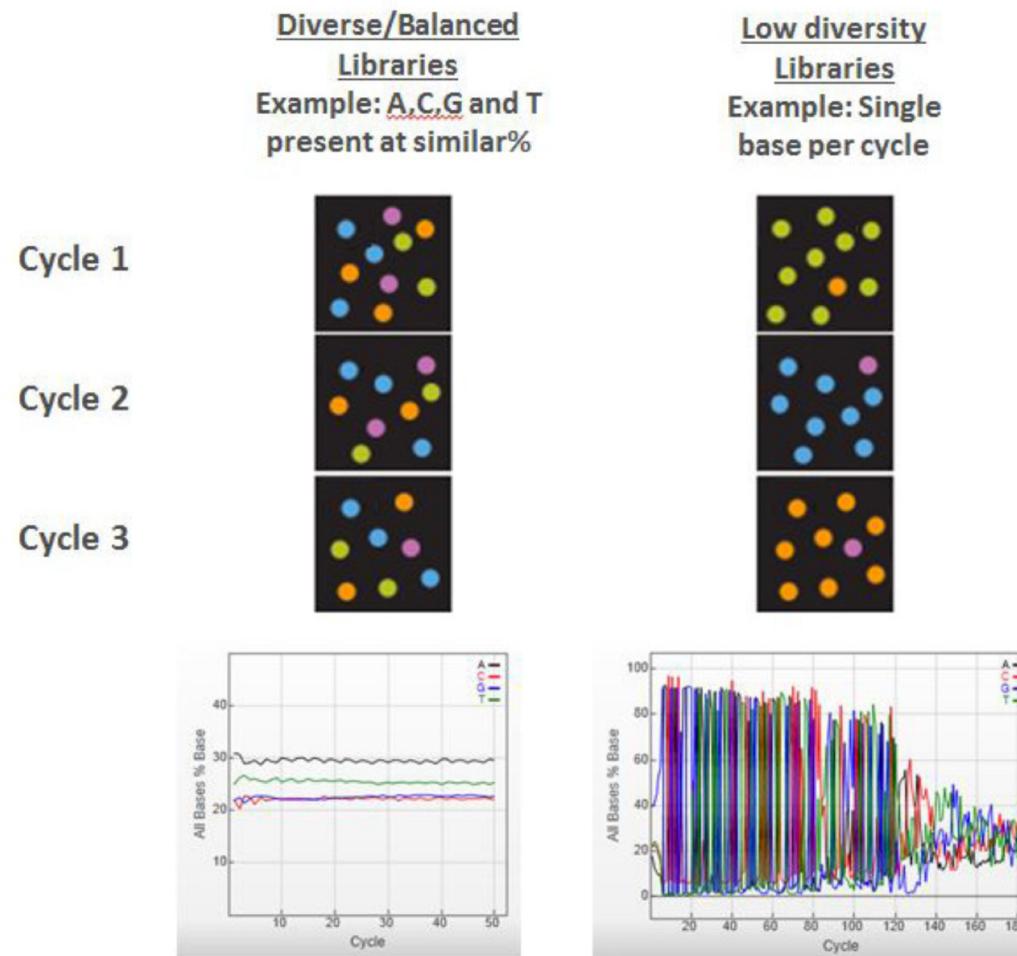
→ estimation of error rate

- diverse: ~45% GC and 55% AT
- well known genome



# NGS Illumina Sequencing

- PhiX



# NGS workflow

---

- Library preparation
- Sequencing
  - clustering
  - sequencing
- Data-analysis

AGGTCGTTACGTACGCTAC  
GACCTACATCAGTACATAG  
GCATGACAAAGCTAG**G**TGT

# NGS workflow

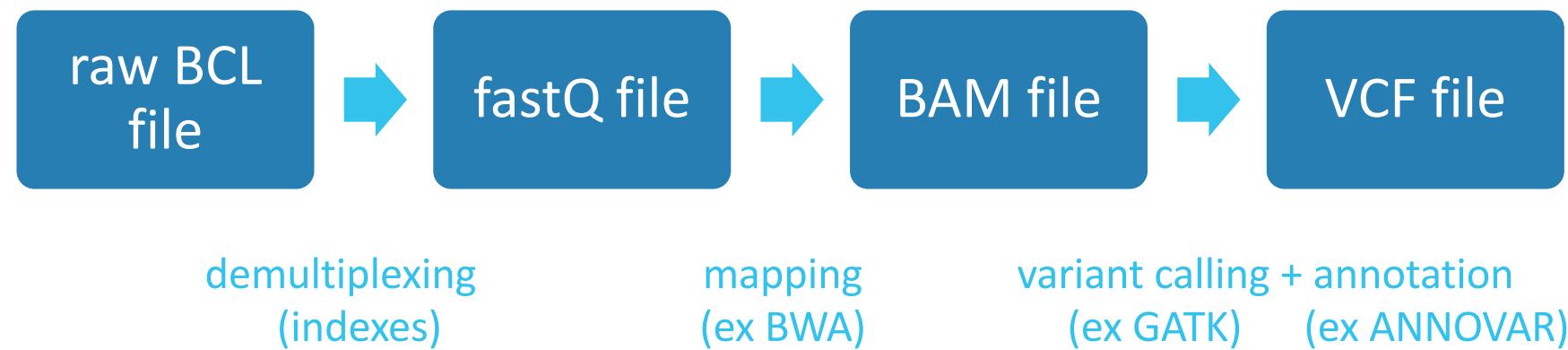
---

AGGTCGTTACGTACGCTAC  
GACCTACATCAGTACATAG  
GCATGACAAAGCTAG**GTGT**

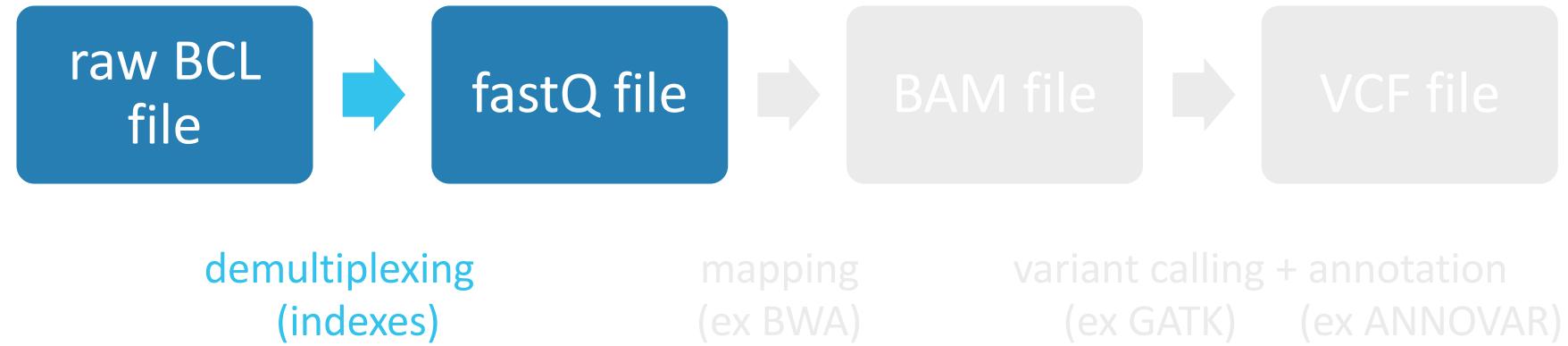
Mapping, alignment,  
variant calling

# Data-analysis

---



# Data-analysis



# Data-analysis

- FastQ file

Divided into blocks of **4 lines**

Machine ID	Run ID	Lane	Tile	X pos	Y pos	
@ILMN-GA001_3_208HWAAXX	1	1	110	812		1. ID
ATACAAGCAAAGTATAAAGTTCTGATGCCGTCTT						2. Sequence
+ILMN-GA001_3_208HWAAXX	1	1	110	812		
hhhhYhh]NYhhhhhhYIhhazT[hhYHNSPKXR						4. Quality
@ILMN-GA001_3_208HWAAXX	1	1	111	879		
GGAGGGCTGGAGTTGGGGACGTATGCCGCATAG						
+ILMN-GA001_3_208HWAAXX	1	1	111	879		
hSWhRNJ\hFhLdhVOhAIB@NFKD@PAB?N?						

# Data-analysis

- FastQ file

**A**

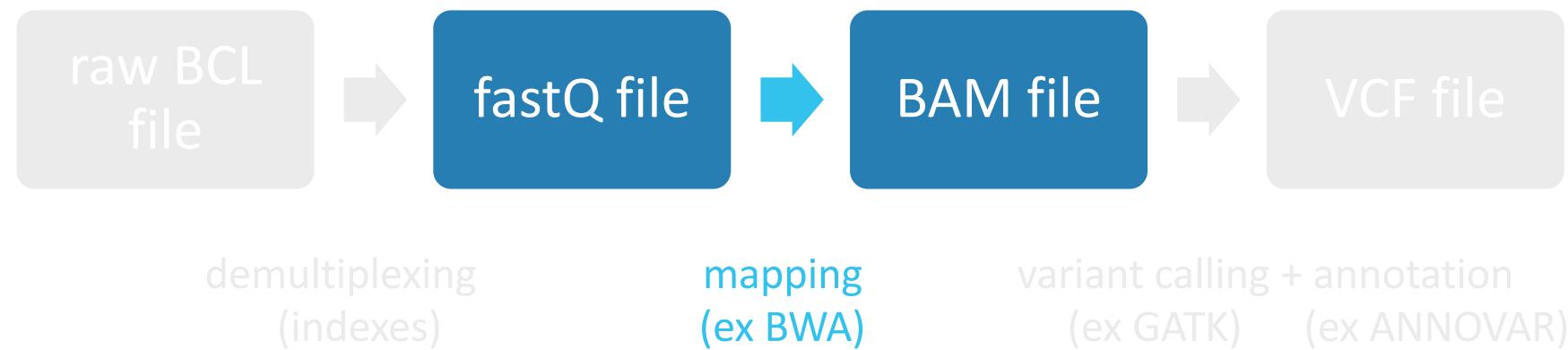
<i>Machine ID</i>	<i>Run ID</i>	<i>Lane:Tile</i>	<i>x:y coord.</i>	<i>Read pair #</i>
@HWI-ST395_0083:3:1:3429:2628#0/1				
<b>SEQ</b> AAAGAATGTACAGCTCGGAAATCACTGACTTGCT				
+HWI-ST395_0083:3:1:3429:2628#0/1				
<b>QUAL</b> GGFDDGGGBGEEGGEGGGDDG>GGHHEHDDEGGG				

Here quality scores are represented as ASCII characters (ASCII offset 33)

-> You have to add 33 to the quality Phred score

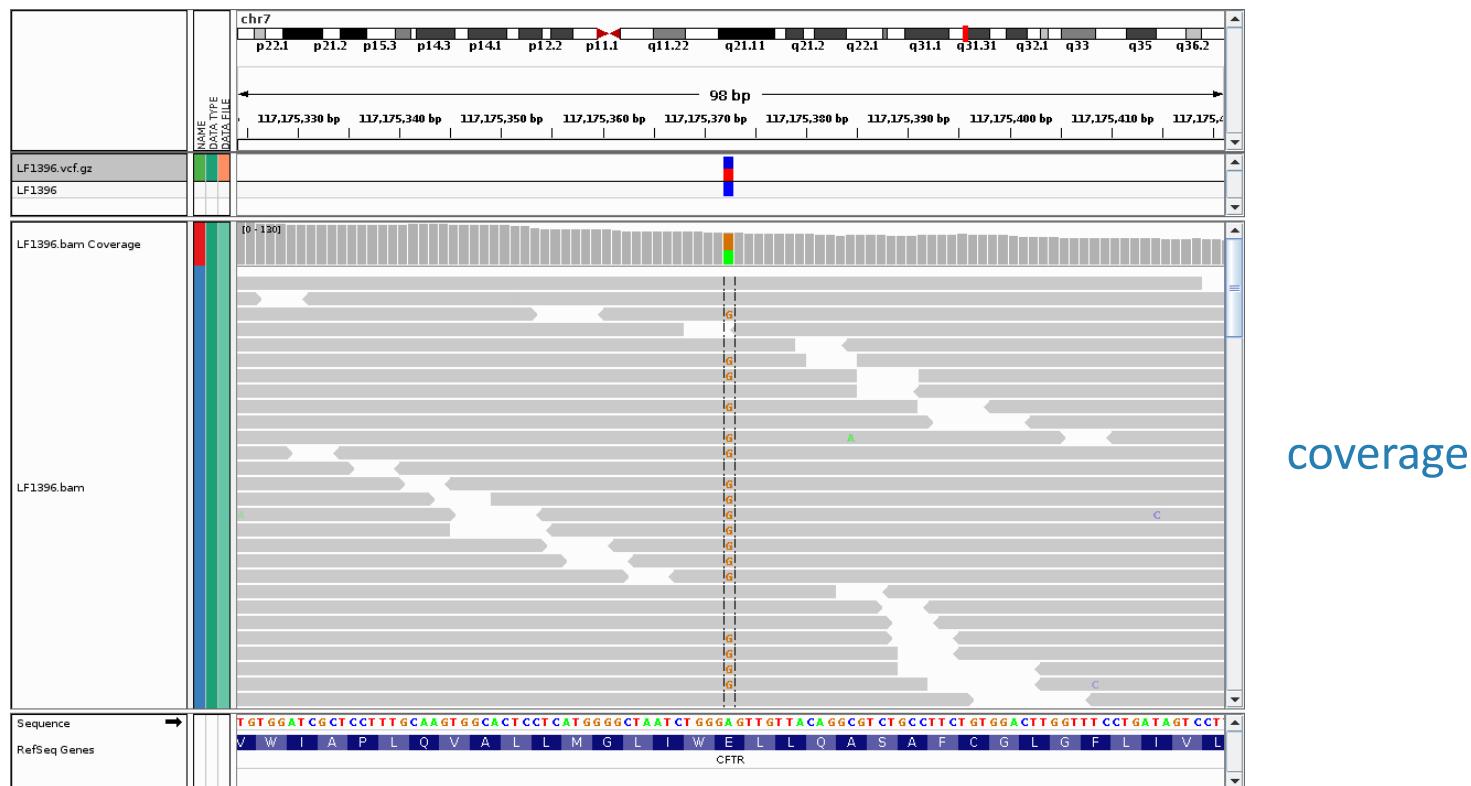
e.g. quality score = 38 ;  $38 + 33 = 71$  ; character 71 in ASCII code = G

# Data-analysis

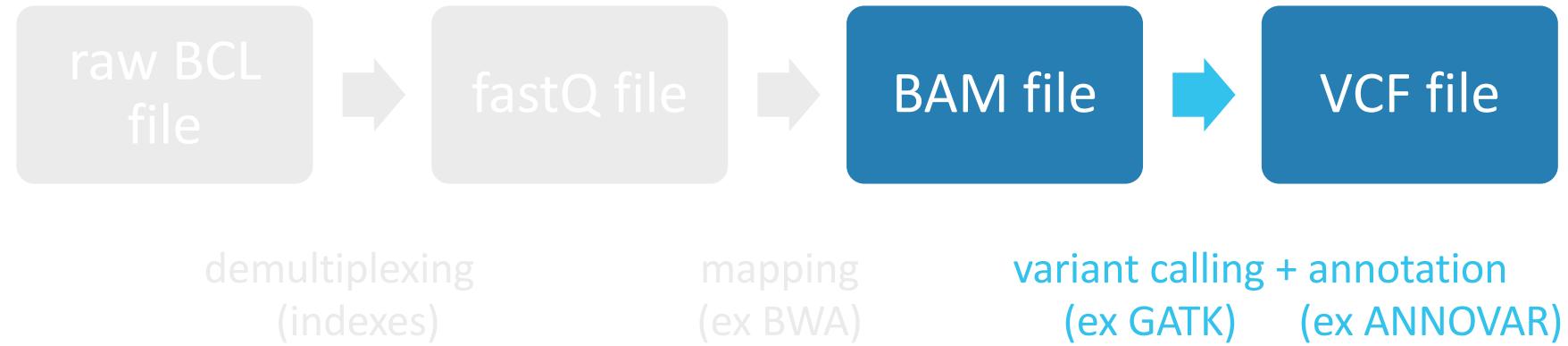


# Data-analysis

- BAM file (binary alignment/map)
  - visualisation in IGV



# Data-analysis



# Data-analysis

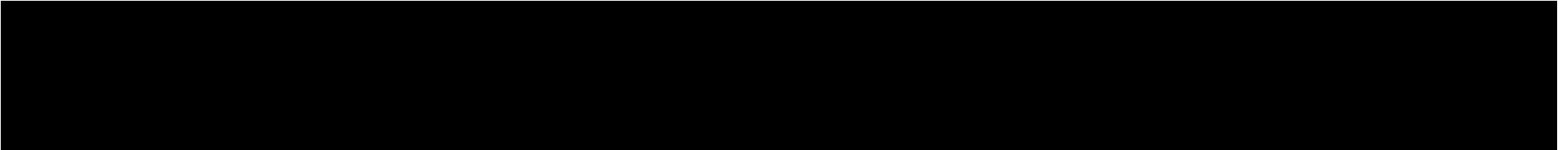
- VCF file (variant call format)

VCF header									
<pre>##fileformat=VCFv4.0 ##fileDate=20100707 ##source=VCFtools ##reference=NCBI36 ##INFO=&lt;ID=AA,Number=1,Type=String,Description="Ancestral Allele"&gt; ##INFO=&lt;ID=H2,Number=0,Type=Flag,Description="HapMap2 membership"&gt; ##FORMAT=&lt;ID=GT,Number=1,Type=String,Description="Genotype"&gt; ##FORMAT=&lt;ID=GQ,Number=1,Type=Integer,Description="Genotype Quality (phred score)"&gt; ##FORMAT=&lt;ID=GL,Number=3,Type=Float,Description="Likelihoods for RR,RA,AA genotypes (R=ref,A=alt)"&gt; ##FORMAT=&lt;ID=DP,Number=1,Type=Integer,Description="Read Depth"&gt; ##ALT=&lt;ID=DEL,Description="Deletion"&gt; ##INFO=&lt;ID=SVTYPE,Number=1,Type=String,Description="Type of structural variant"&gt; ##INFO=&lt;ID=END,Number=1,Type=Integer,Description="End position of the variant"&gt;</pre>									
<pre>#CHROM POS ID REF ALT QUAL FILTER INFO</pre>					<pre>FORMAT</pre>		<b>SAMPLE1</b>	<b>SAMPLE2</b>	<b>Reference alleles (GT=0)</b>
<pre>1 1 . 1 2 rs1 ACG A,AT . 1 5 . 1 100</pre>					<pre>GT:DP</pre>	<pre>1/2:13</pre>	<pre>0/0:29</pre>		
<pre>REF ALT C T,CT A G T &lt;DEL&gt;</pre>					<pre>GT:GQ</pre>	<pre>0 1:100</pre>	<pre>2/2:70</pre>		
Body									
<pre>QUAL FILTER . PASS . . PASS . . PASS SVTYPE=DEL;END=300</pre>					<pre>GT:GQ</pre>	<pre>1 0:77</pre>	<pre>1/1:95</pre>		
<pre>INFO</pre>					<pre>GT:GQ:DP</pre>	<pre>1/1:12:3</pre>	<pre>0/0:20</pre>		
<b>Other event</b>									
<b>Phased data (G and C above are on the same chromosome)</b>									
<b>Deletion</b> <b>SNP</b> <b>Large SV</b> <b>Insertion</b> <b>Other event</b>									
<b>Mandatory header lines</b>									
<b>Optional header lines (meta-data about the annotations in the VCF body)</b>									
<b>Alternate alleles (GT&gt;0 is an index to the ALT column)</b>									

CS1 calls

# NGS workflow

---



# NGS technologies

- Second generation



454/Roche



SoliD



Complete Genomics



Illumina



Life Technologies

- Third generation



Pacific BioSciences



Oxford Nanopore

# Pacific BioSciences

Observe **single molecules** in **real time** to provide high-throughput SMRT® Sequencing of DNA and base modifications simultaneously

- Generate finished genomes
- Discover a broad spectrum of base modifications
- Characterize complex variations



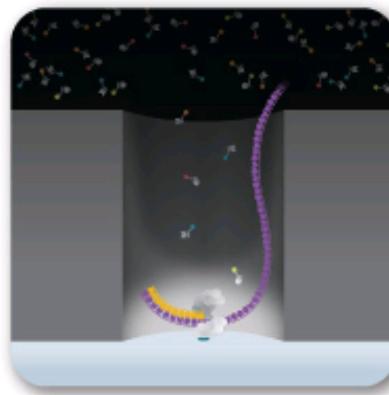
- Long read lengths
- High accuracy
- Ultra sensitivity
- Shortest run time
- Least GC bias
- No amplification bias

# Pacific BioSciences

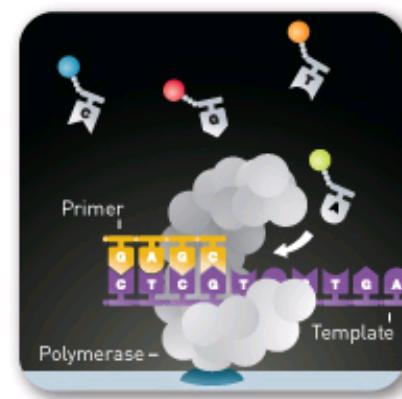
## SMRT® Cells



## Zero-Mode Waveguides



## Phospholinked Nucleotides



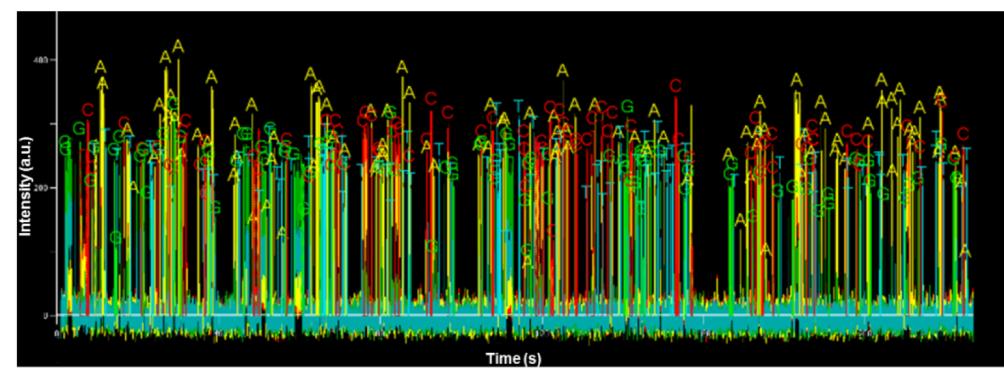
PacBio® RS II



PacBio® Sequel



## Trace



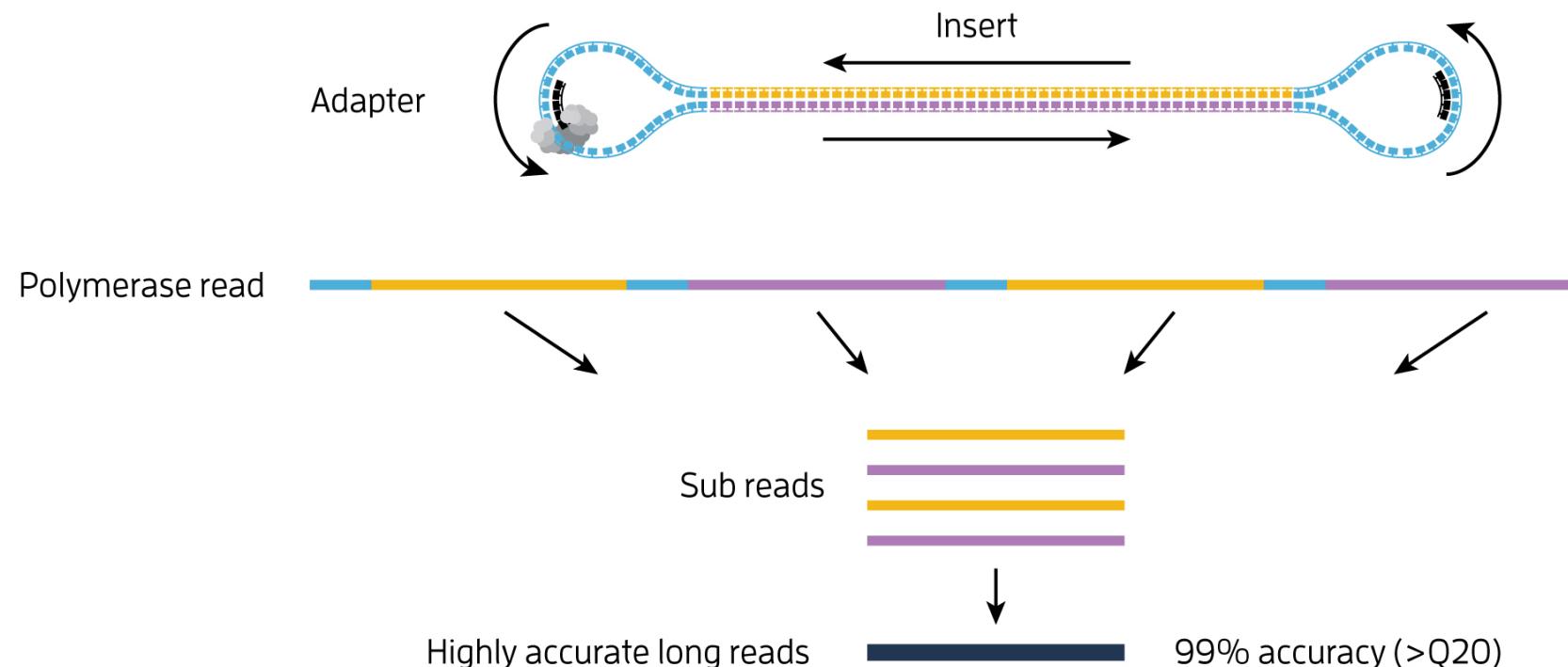
# Pacific BioSciences

---



SMRT® Cell

- Very long reads
- Accurate Hifi “shorter” reads



# NGS @ Genomics Core

---

- Create a project on <http://www.genomicscore.be/request>
- Specify
  - project description
  - sample type (DNA / RNA / library / ...)
  - sample origin
  - concentration, average fragment size (libraries @ 4nM)
  - indexes
  - sequencing requirements (PE100, SR50, ...)
  - bioinformatics
  - transfer of results

# NGS @ Genomics Core

Instrument / kit	Sequencing Capacity (# reads)
MiSeq v2 Nano	1 M
MiSeq v2 Micro	4 M
MiSeq v2 Standard	15 M
MiSeq v3	25 M
NextSeq Mid Output	130 M
NextSeq High Output	400 M
HiSeq2500 Rapid run v2	150 M / lane (2 lanes)
HiSeq2500 High Output run v2	250 M / lane (8 lanes)
HiSeq4000	320 M / lane (8 lanes)
NovaSeq SP	400 M / lane (2 lanes)
NovaSeq S1	800 M / lane (2 lanes)
NovaSeq S2	1650 M / lane (2 lanes)
NovaSeq S3	1650 M / lane (4 lanes)
NovaSeq S4	2500 M / lane (4 lanes)
PacBio RSII	40.000 – 70.000 reads
PacBio Sequel	350.000 - 600.000 reads

# NGS @ Genomics Core

---

- Aangeleverde concentratie staal: .....nM
- Eindconcentratie staal op sequencer: .....pM
- Percentage PhiX: .....%
  
- Custom primers: JA / NEEN
  - *Naam en eindconc.:* .....μM
  - *Enkel custom primers of spike-in bij illumina primers?*
  
- Aantal cycli  
(read1 – index1 – optioneel index2 – read2):  
..... - ..... - ..... - .....

# The Genomics Core

---



[info@genomicscore.be](mailto:info@genomicscore.be)

# NGS technologies

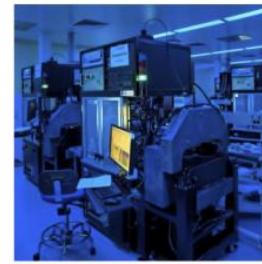
- Second generation



454/Roche



SoliD



Complete Genomics



Illumina



Life Technologies

- Third generation



Pacific BioSciences



Oxford Nanopore

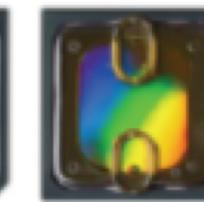
# Life Technologies



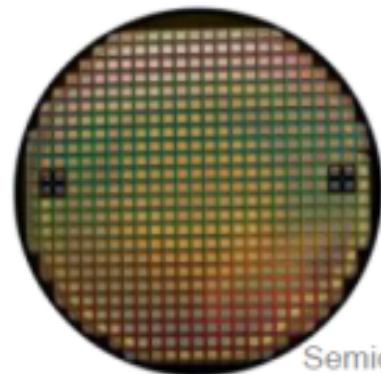
Ion PGM™ Sequencer



Ion Proton™ Sequencer

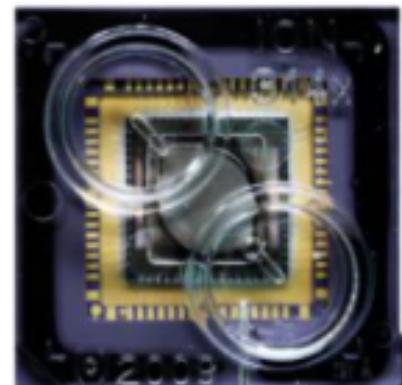


# Life Technologies



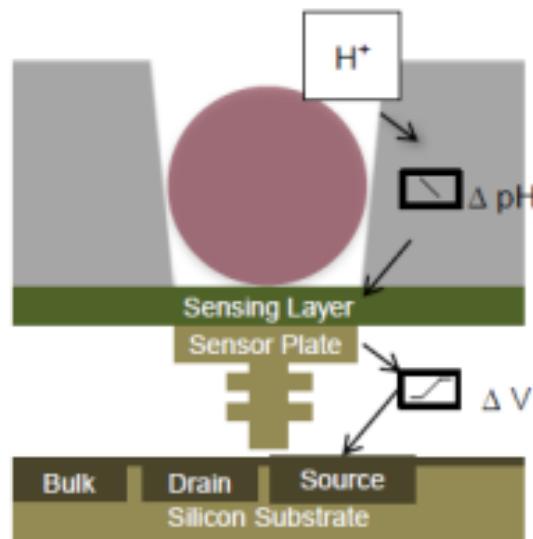
Wafer

Semiconductor Manufacturing

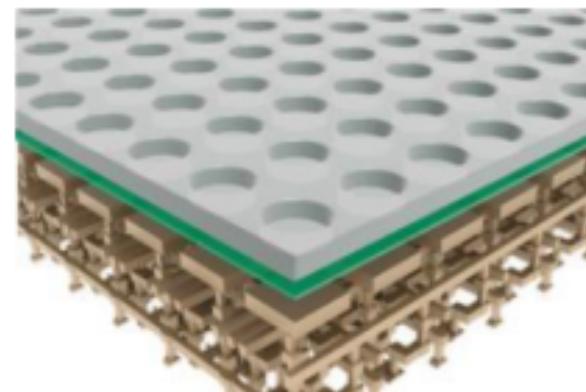


Chip

Semiconductor Packaging



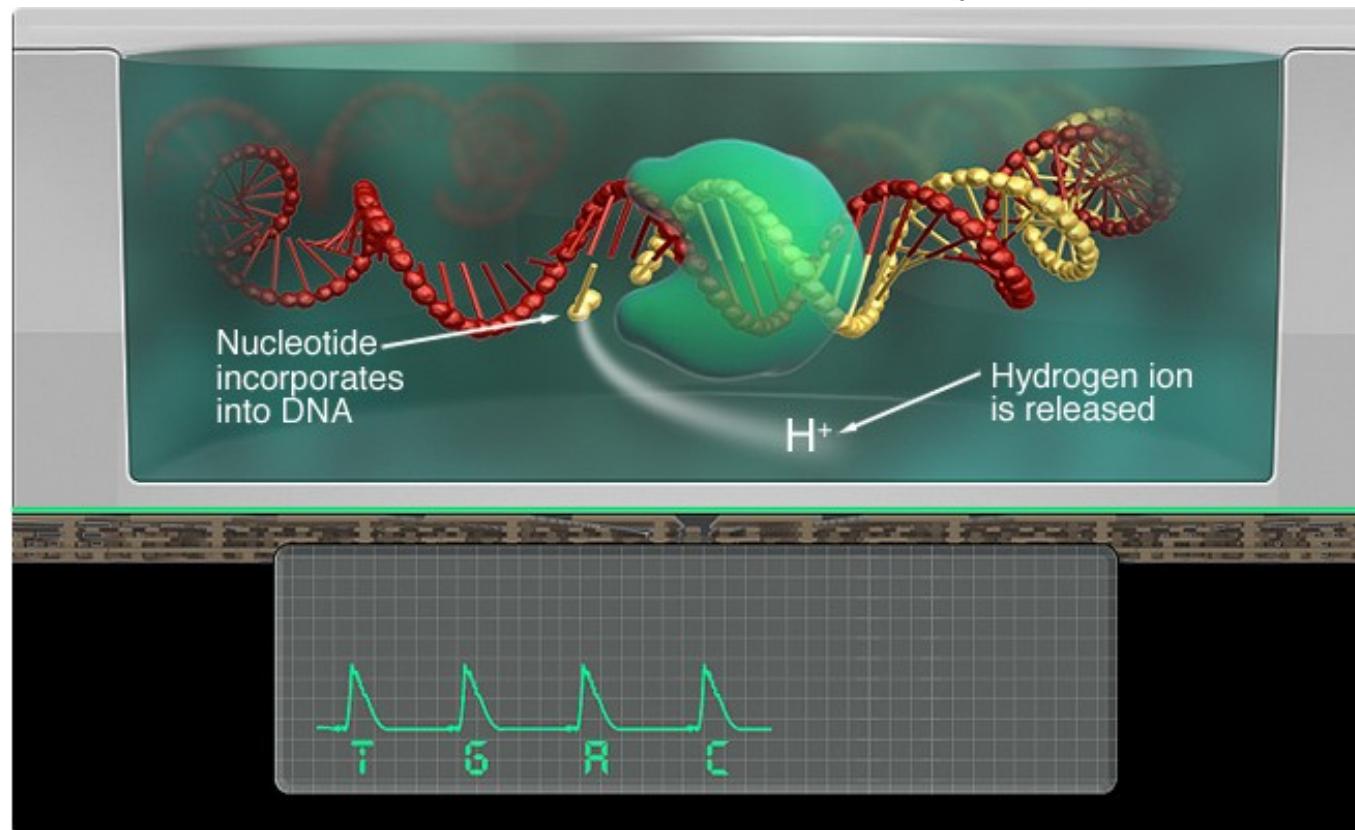
Single Sensor  
Chemical to Digital Sequence



Millions of Sensors  
Semiconductor Design

# Life Technologies

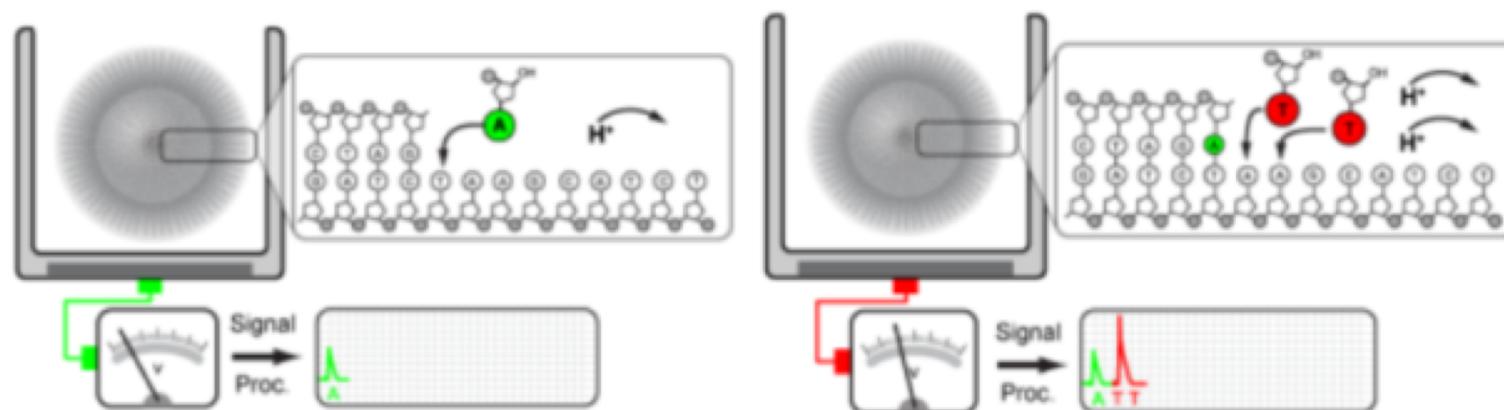
4 nucleotides flow sequentially



No camera, just a pH sensor

# Life Technologies

- Measurement of  $\Delta\text{pH}$  @ base incorporation



# NGS technologies

- Second generation



454/Roche



SoliD



Complete Genomics



Illumina



Life Technologies

- Third generation



Pacific BioSciences



Oxford Nanopore

# Oxford Nanopore



