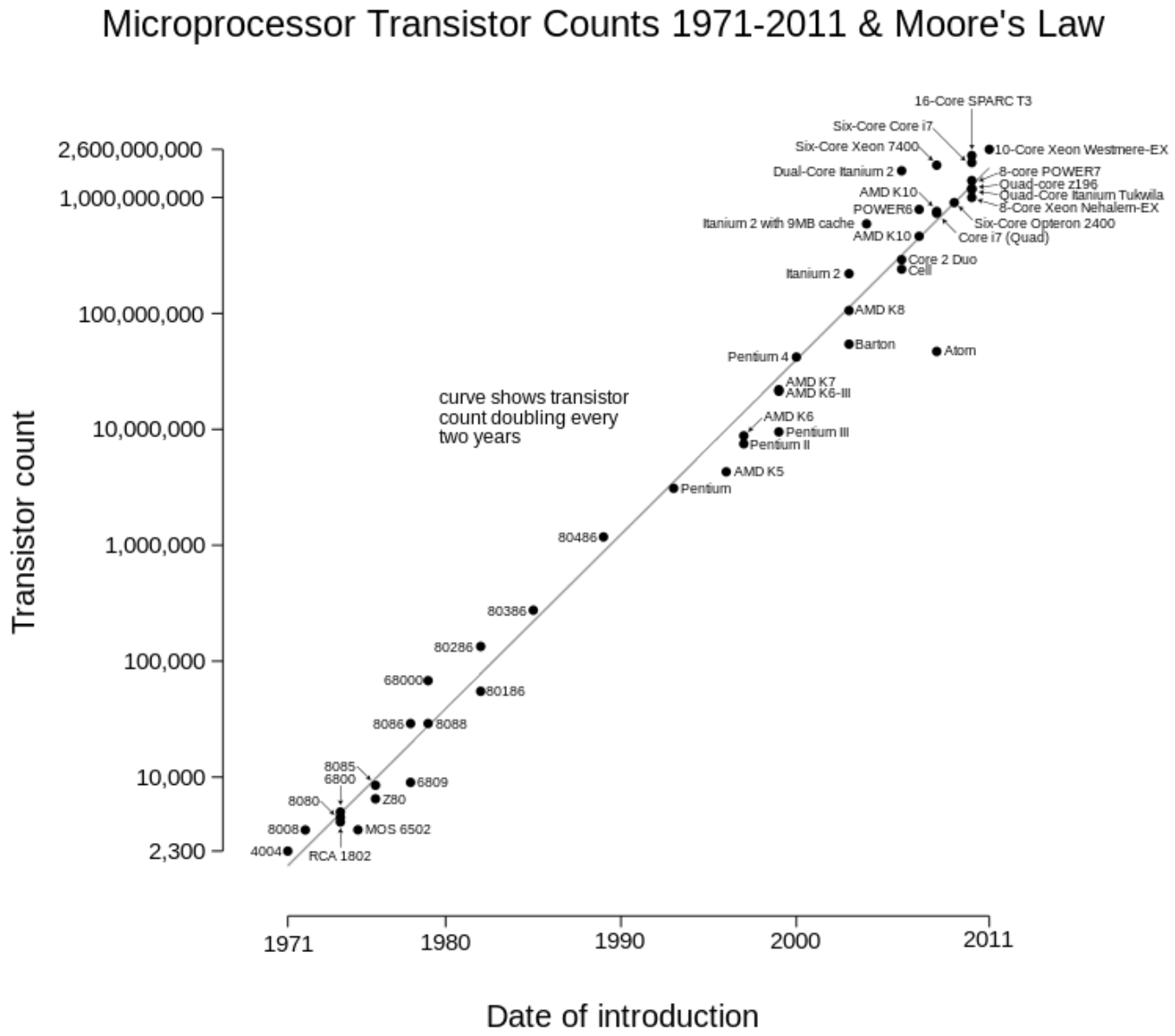
Sequencing Technologies

BIOL 432

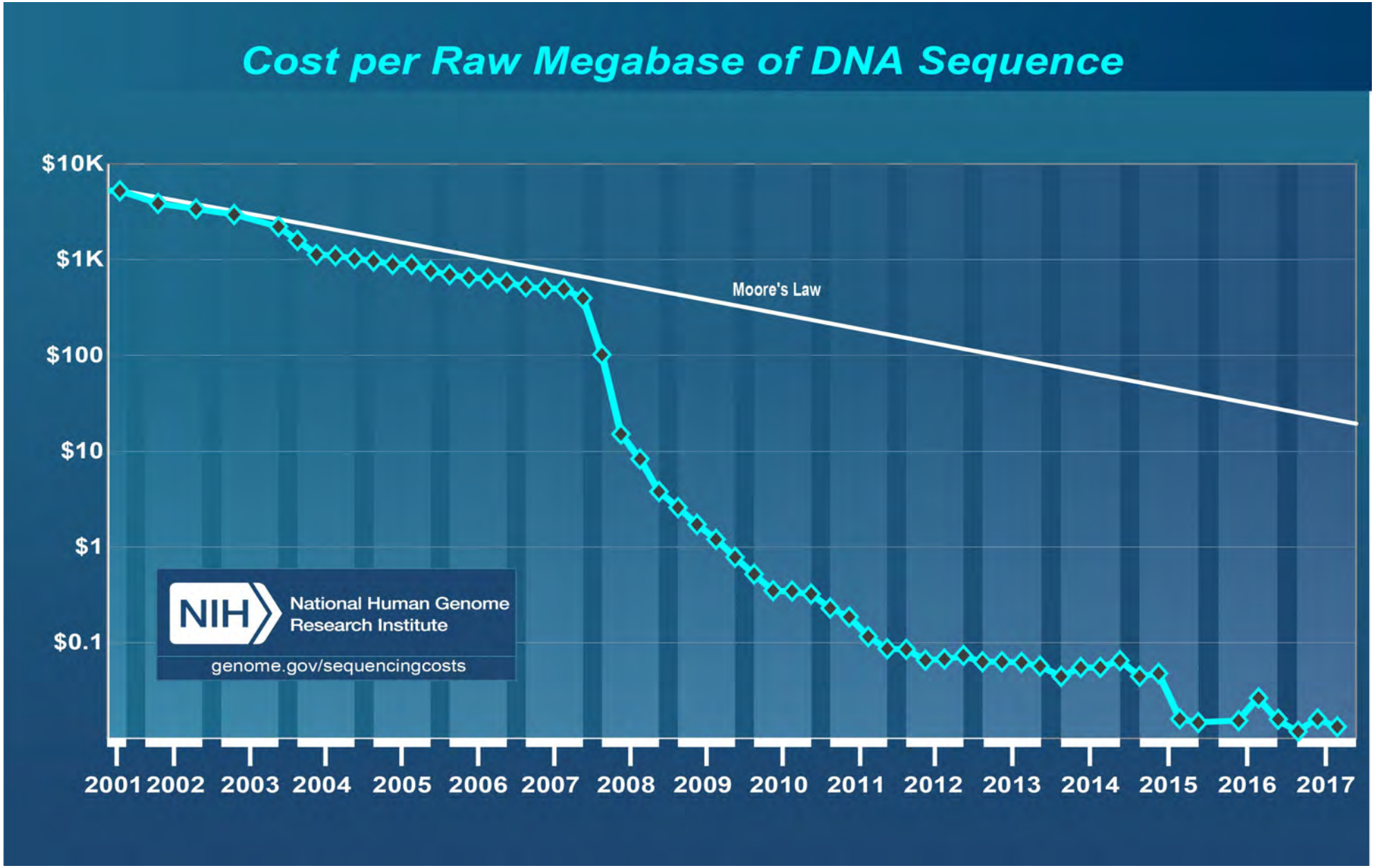
Moore’s Law

@ColauttiLab



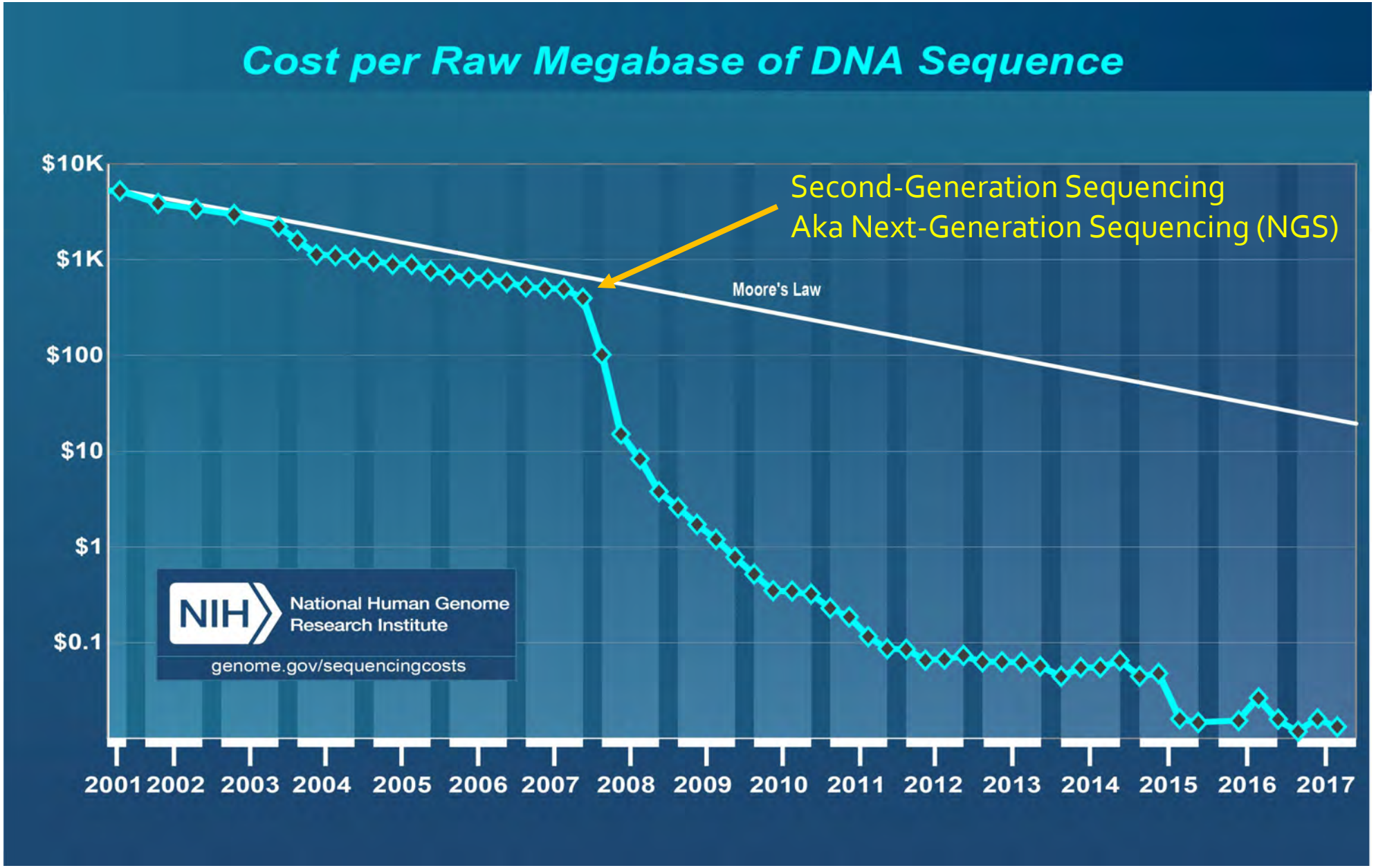
Moore's Law vs. Sequencing Technology

@ColauttiLab



Moore's Law vs. Sequencing Technology

@ColauttiLab



Sequencing Overview – Match platform with description

@ColauttiLab

Sanger

Illumina

PacBio

Nanopore

(Also: IonTorrent, 454, SOLID)

First-generation sequencing

Second-generation sequencing

Third-generation sequencing

Next-generation sequencing

Sequencing-by-synthesis

Nanopore sequencing

High-throughput sequencing

Short-read sequencing

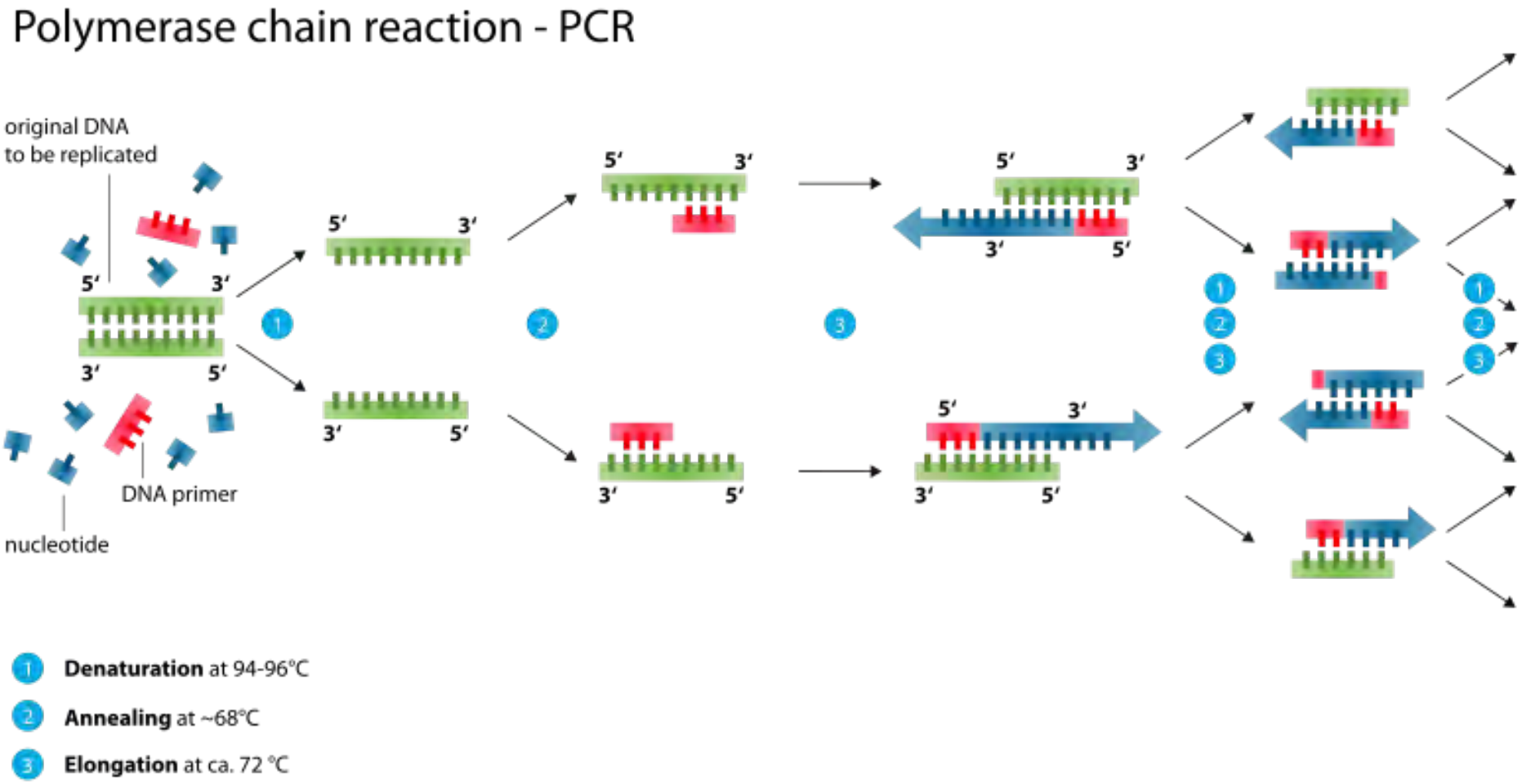
Long-read sequencing

Sanger Method

First-generation sequencing

Review: Polymerase chain reaction

@ColauttiLab



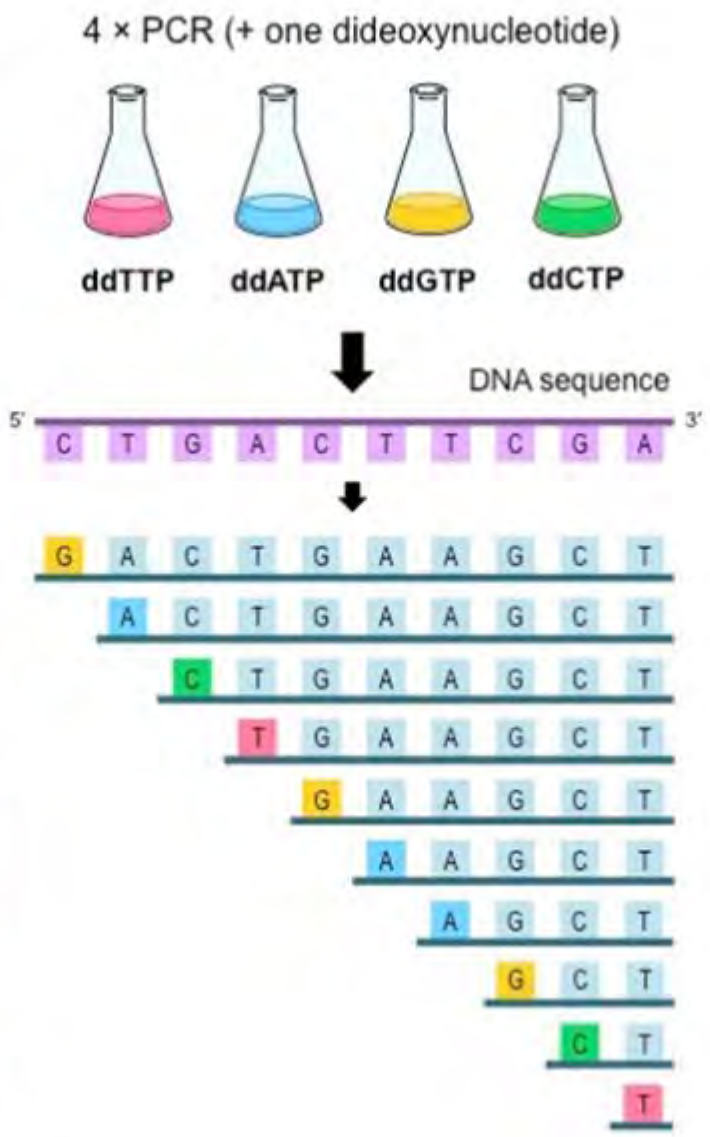
Source: Wikipedia

1st Generation Sequencing: Sanger Method

@ColauttiLab

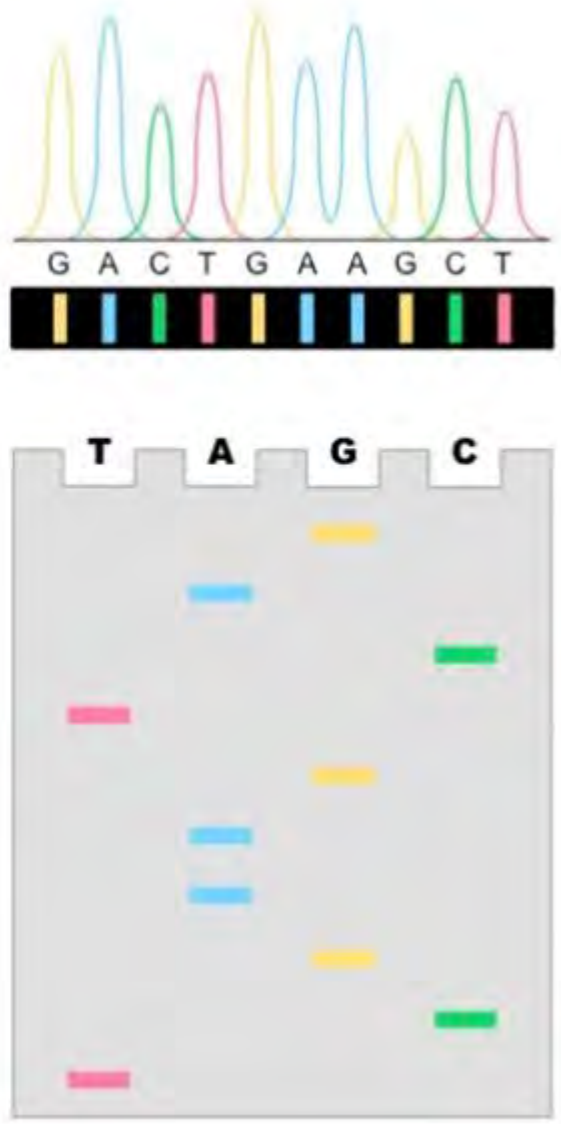
= 'Sequencing-by-synthesis'

Beckman Coulter CEQ 8000



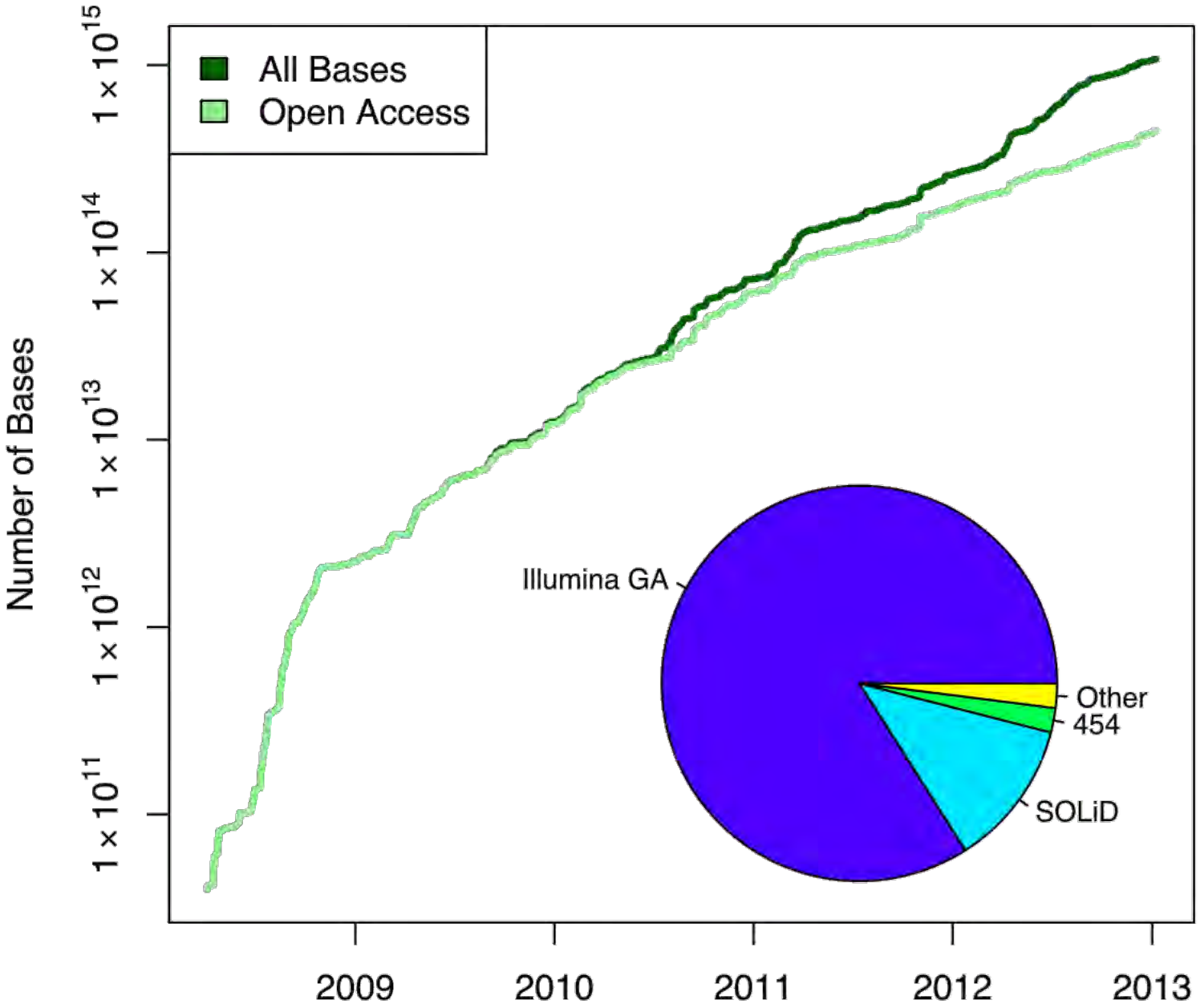
Use a sequencing machine

Separate with a gel



Sequencing Read Archive (NCBI)

@ColauttiLab



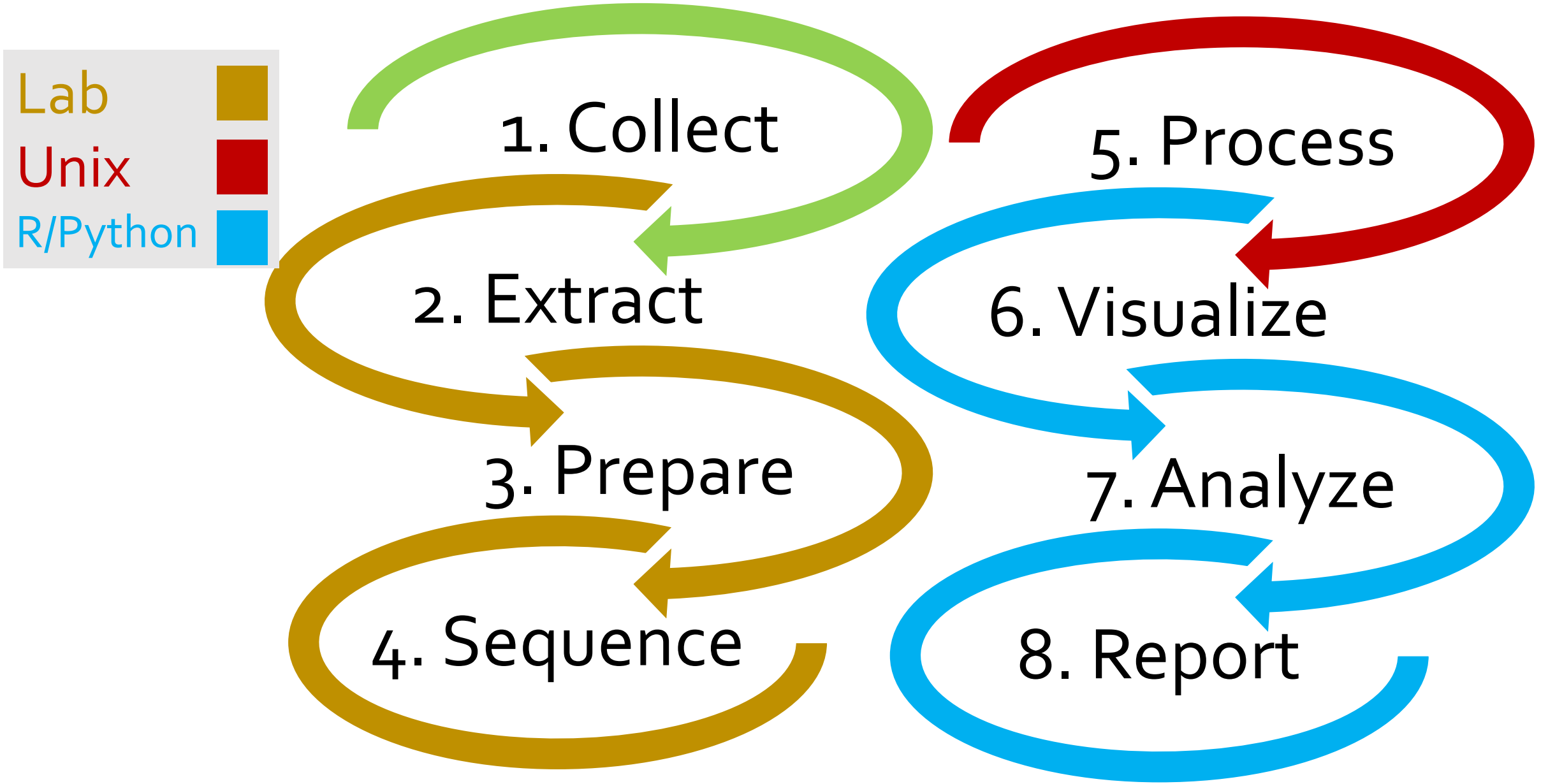
Blue and green =
'Next Generation Sequencing'

Second-generation sequencing

Short-read, high-throughput sequencing

Next-Generation Sequencing: Typical Workflow

@ColauttiLab

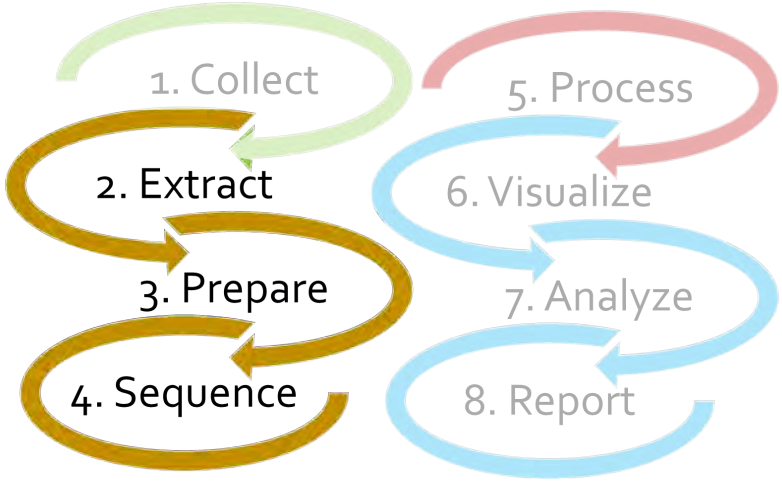


2nd Gen: Common Elements

@ColauttiLab

Sequencing Library Preparation

1. Extract & purify DNA
2. Fragment to target size (75-750 bp)
3. Strand isolation
4. Clonal Amplification
5. Nucleotide detection



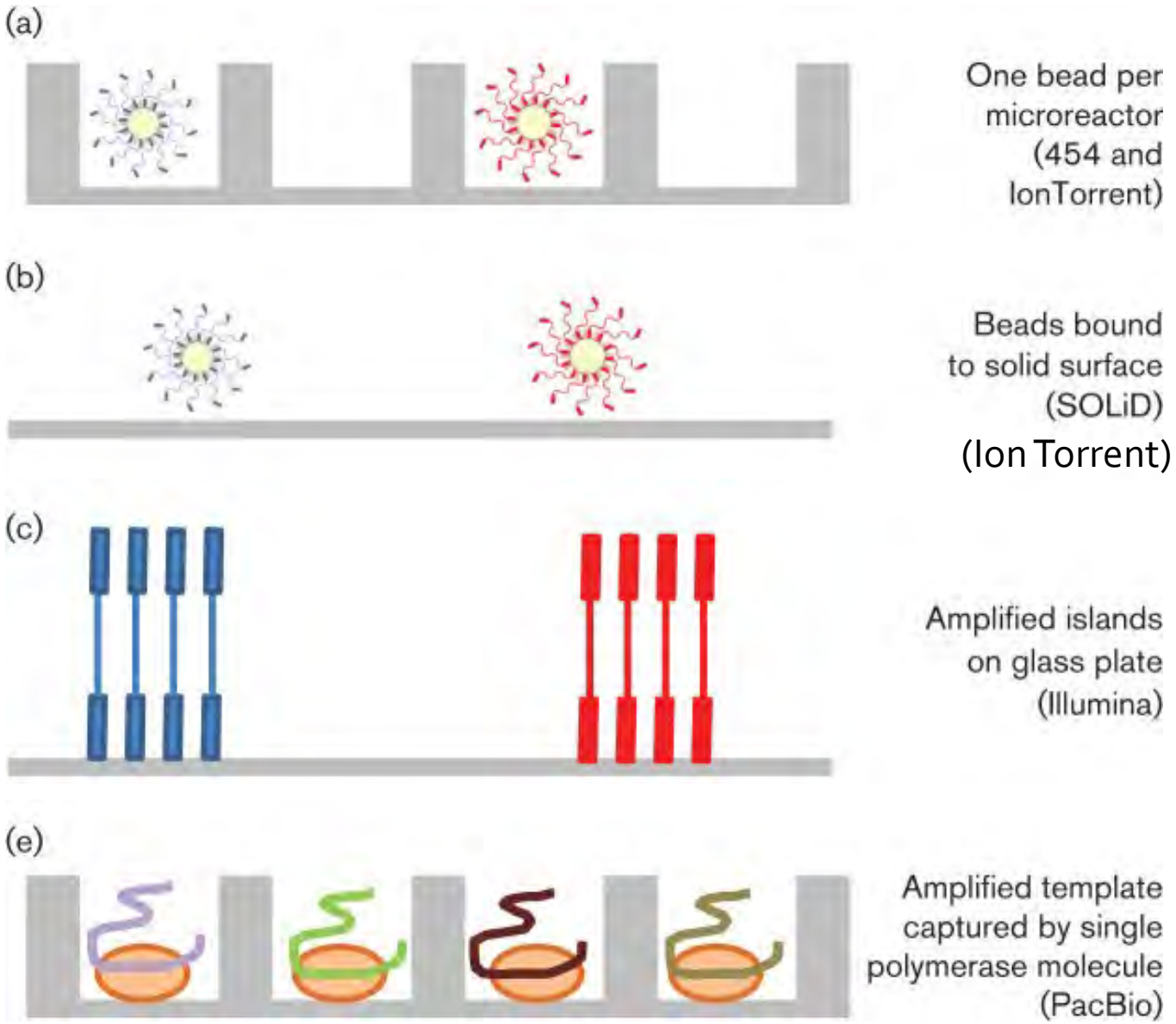
Fragment Sizes (partial list)

@ColauttiLab

Platform	Instrument	Mreads	Length (bp)	Gbp	Type
Illumina	NovaSeq 6000 S4	10000	300	3000	SR & PE
Illumina	NextSeq 500 High-Output	400	300	120	SR & PE
Illumina	HiSeq X	375	300	112.5	PE
Illumina	HiSeq High-Output v4	250	250	62.5	SR & PE
Illumina	MiSeq v3	25	600	15	SR & PE
Illumina	MiniSeq High-Output	25	300	7.5	SR & PE
Ion	Proton I	60	200	12	SR
Ion	PGM 318	4	400	1.6	SR
Ion	PGM 316	2	400	0.8	SR
Ion	PGM 314	0.4	400	0.16	SR
PacBio	PacBio Sequel	0.37	20000	7.4	SR
PacBio	PacBio RS II (P6)	0.055	15000	0.825	SR
Roche 454	GS FLX+ / FLX	0.7	700	0.49	SR
SOLiD	5500xl W	267	100	26.7	SR & PE

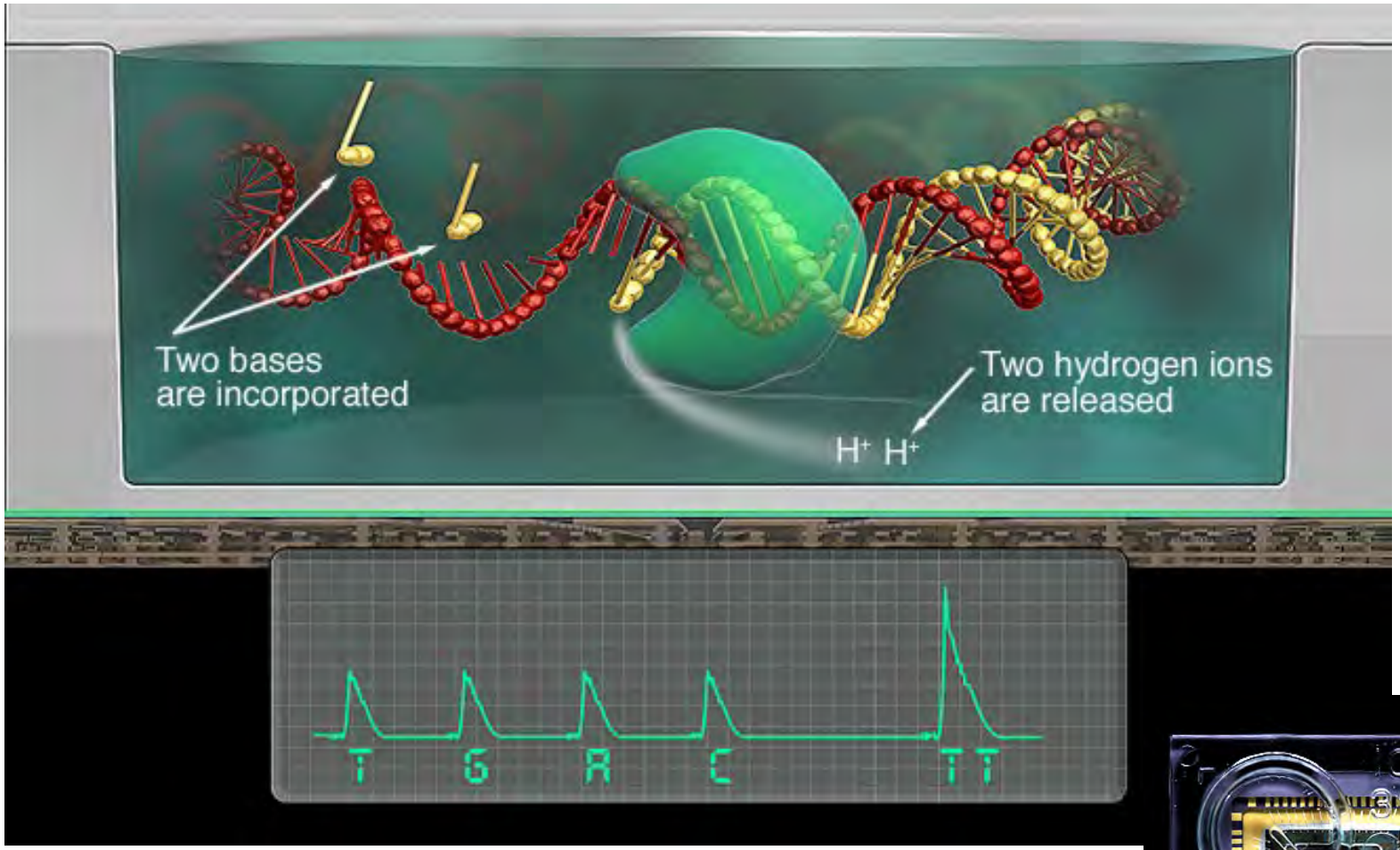
Sequence isolation (and cloning)

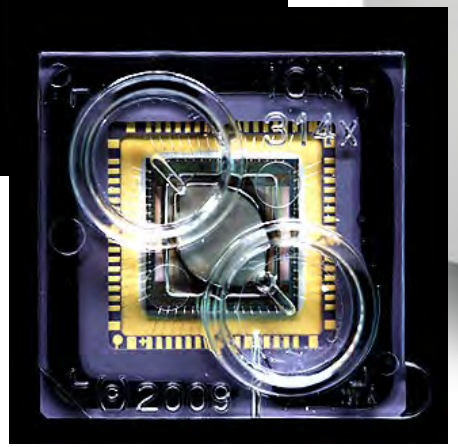
@ColauttiLab



Ion Torrent

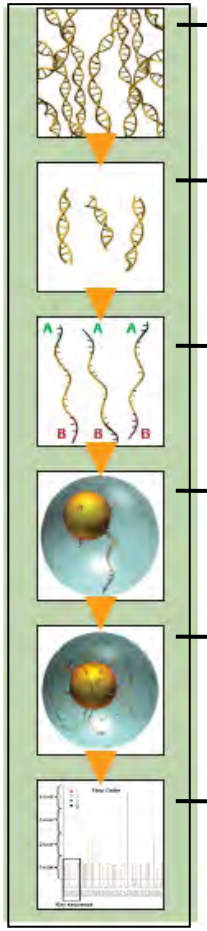
@ColauttiLab



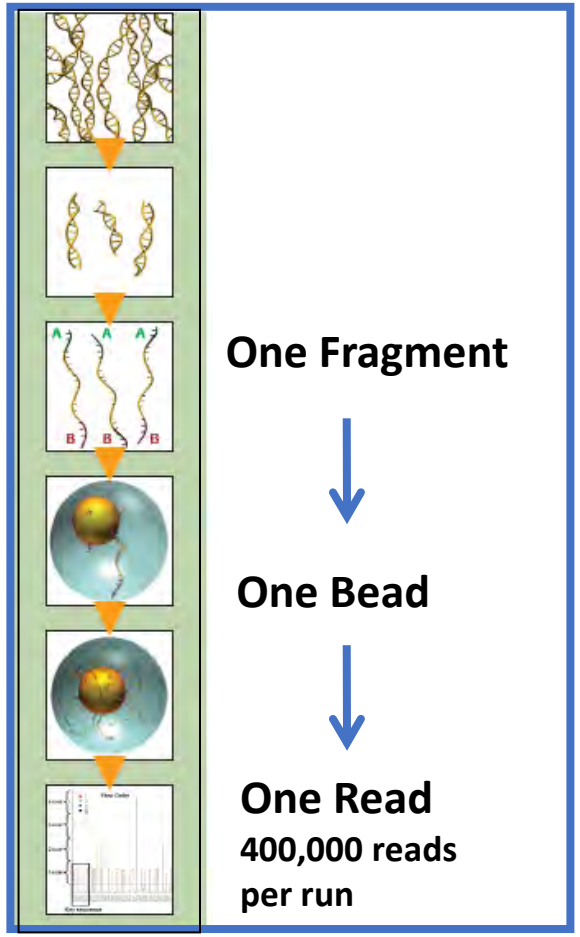


2nd Gen: 454 Sequencing (Roche; deprecated)

@ColauttiLab



- Sample input: Genomic **DNA**, BACs, amplicons, cDNA
- Generation of small **DNA fragments** via nebulization
- Ligation of A/B-Adaptors flanking **single-stranded DNA fragments**
- Emulsification of beads and fragments in water-in-oil **microreactors**
- **Clonal amplification** of fragments bound to beads in microreactors
- **Sequencing** and **base calling**

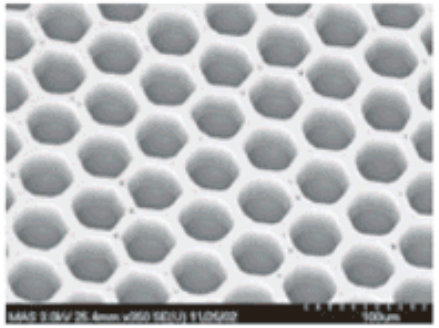


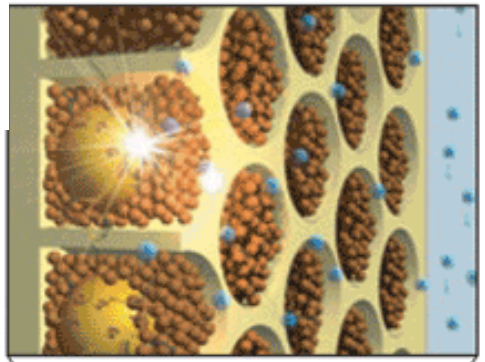
CSB2008 August 2008

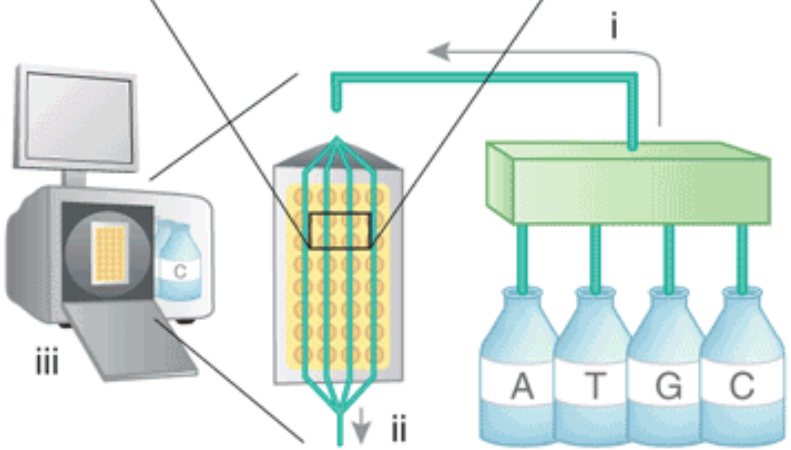
Rothberg & Leomon 2008

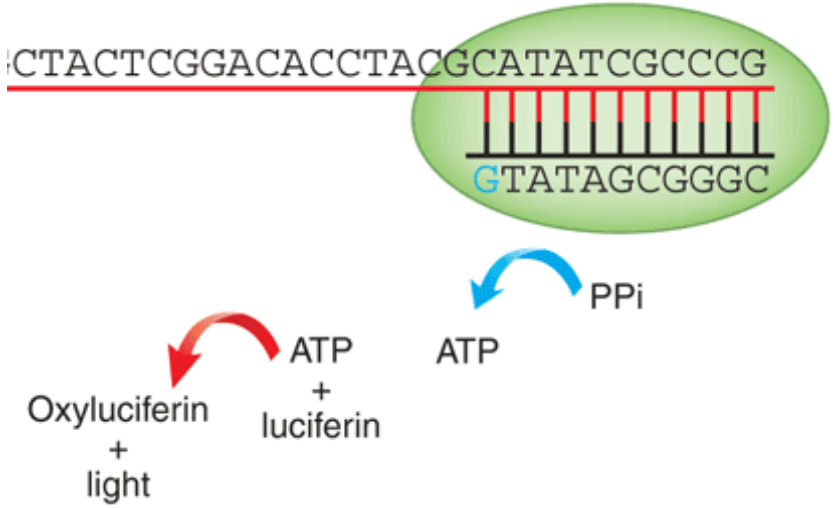
2nd Gen: 454 Sequencing (Roche; deprecated)

@ColauttiLab







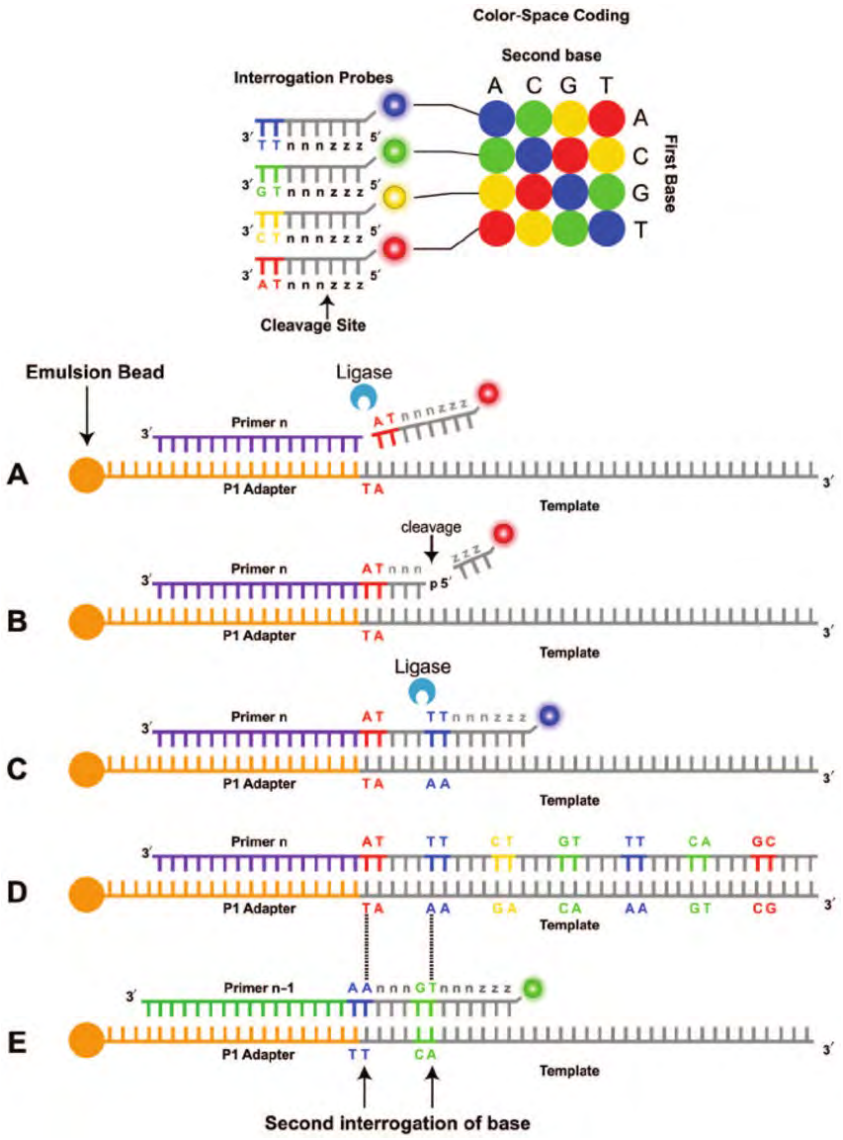


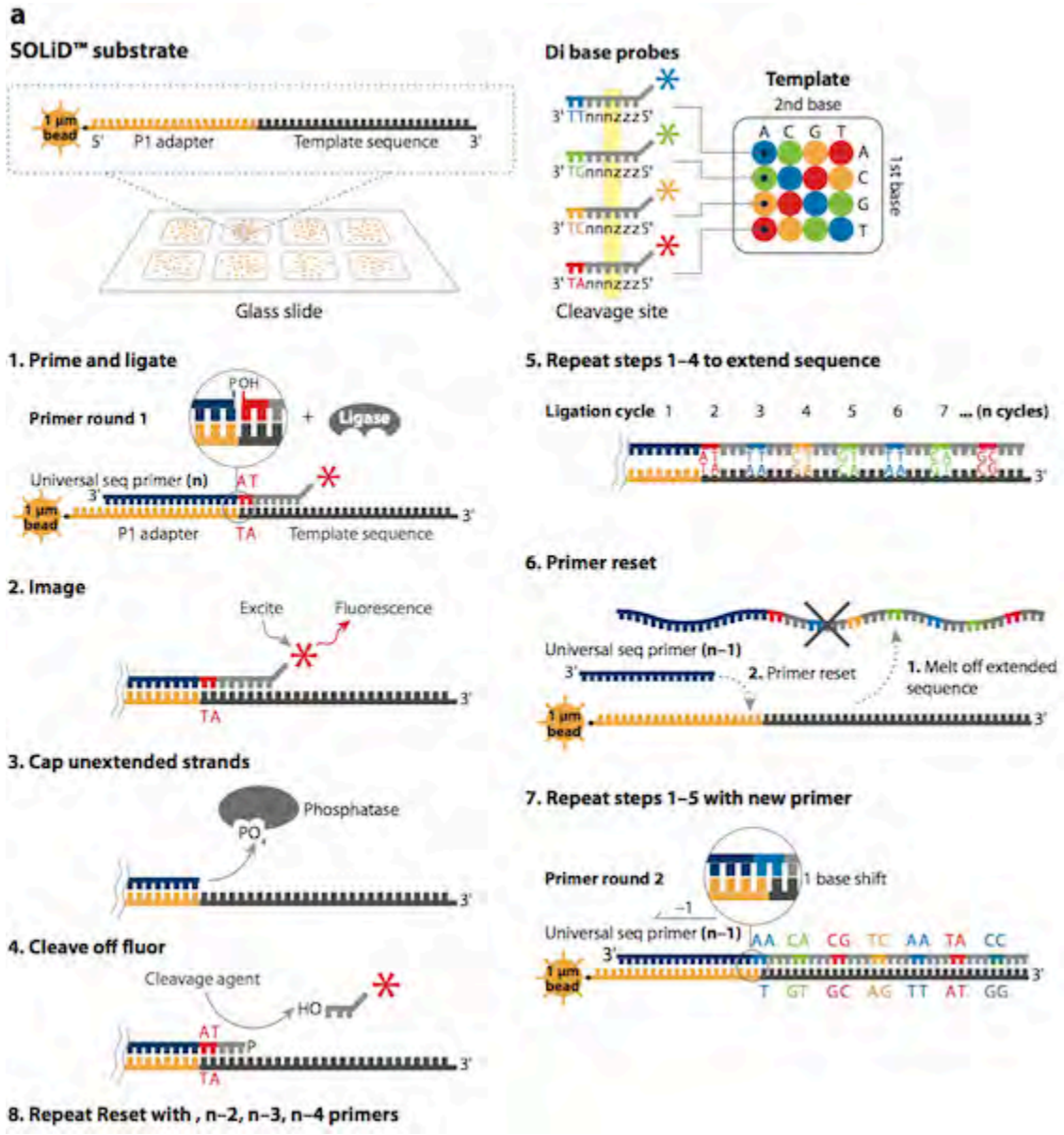
CSB2008 August 2008

Rothberg & Leomon 2008

SOLiD Sequencing (ABI)

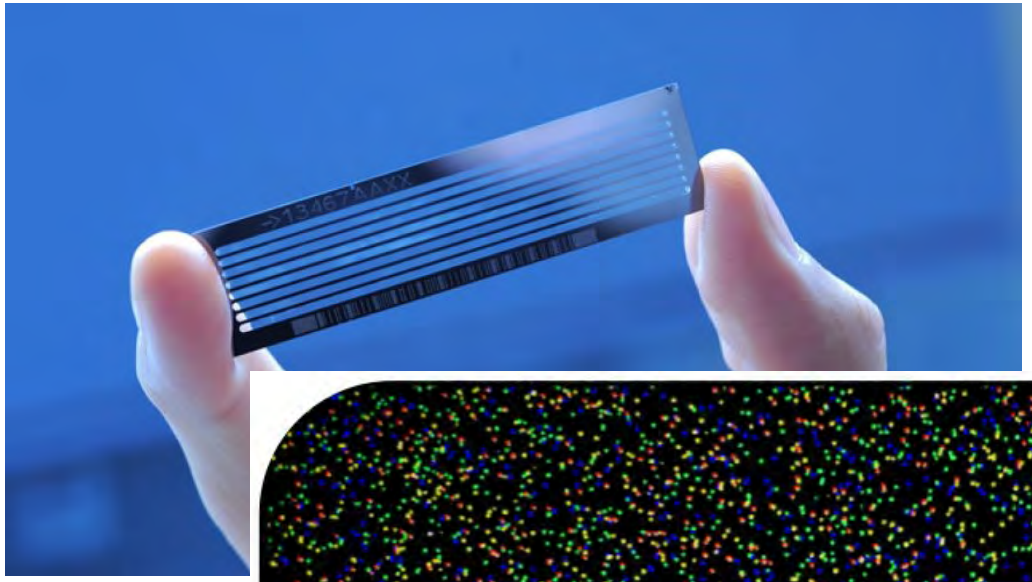
@ColauttiLab

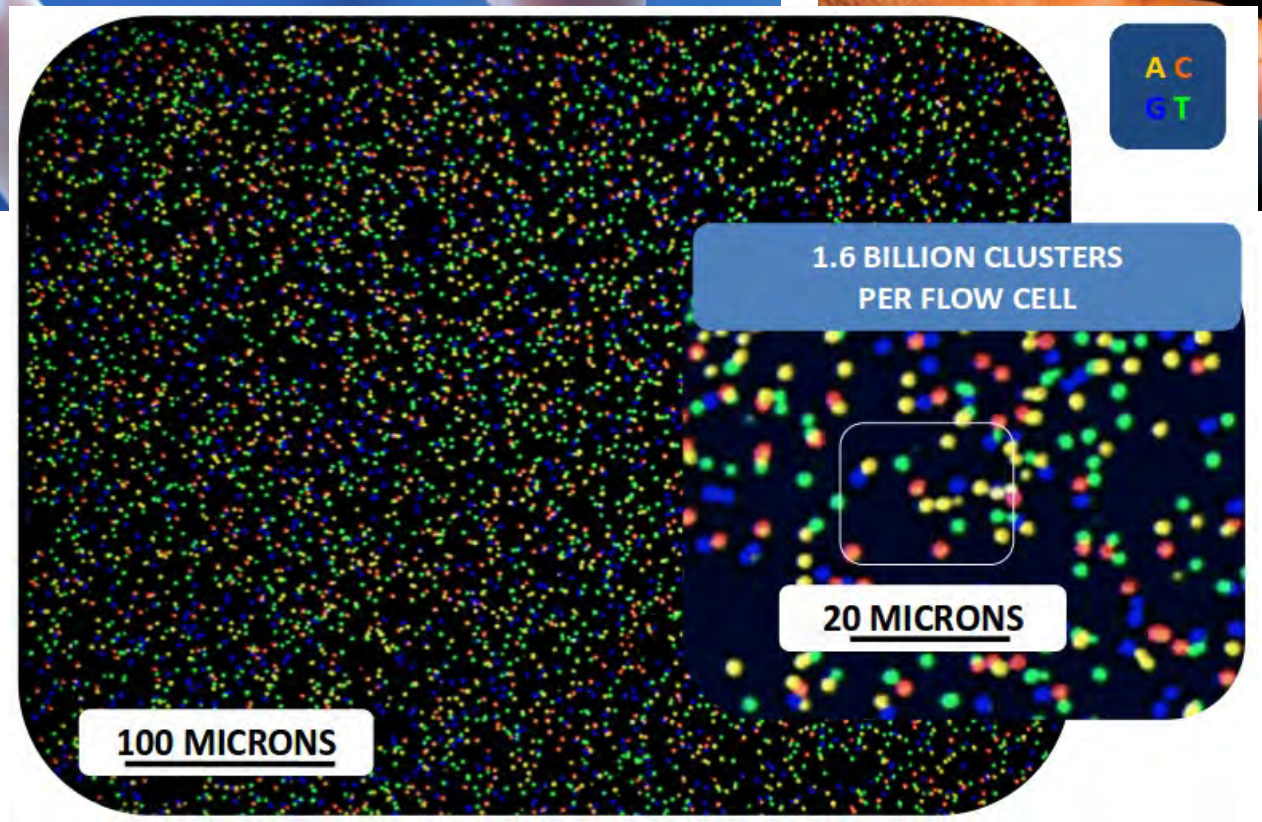




Illumina flow cells (a micro-array technology)

@ColauttiLab

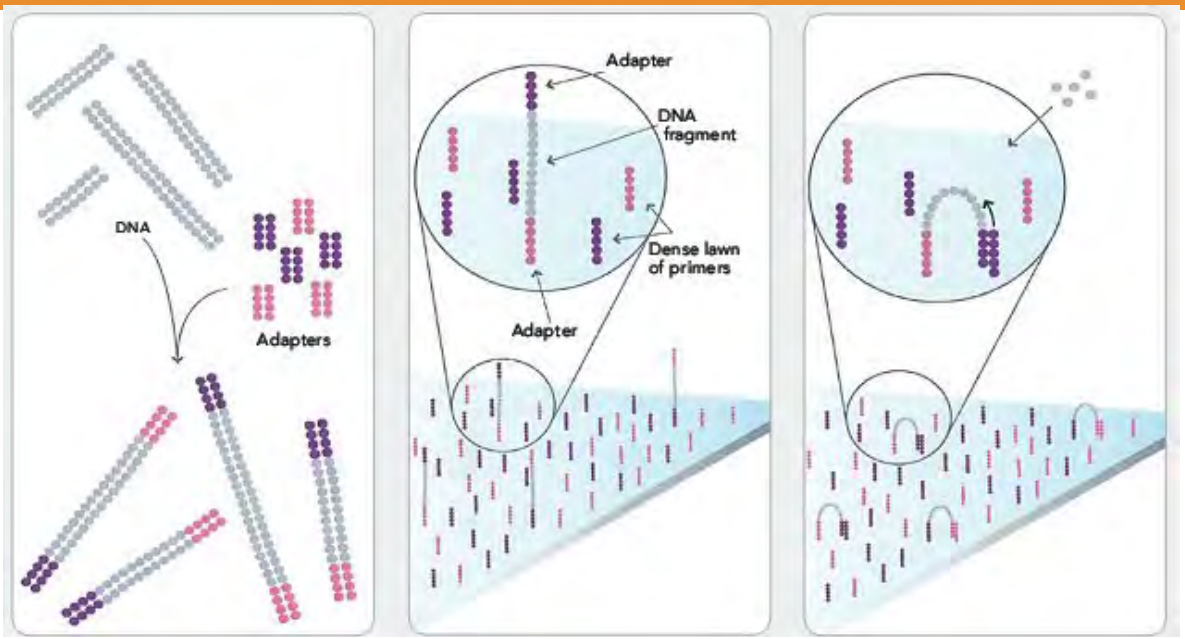


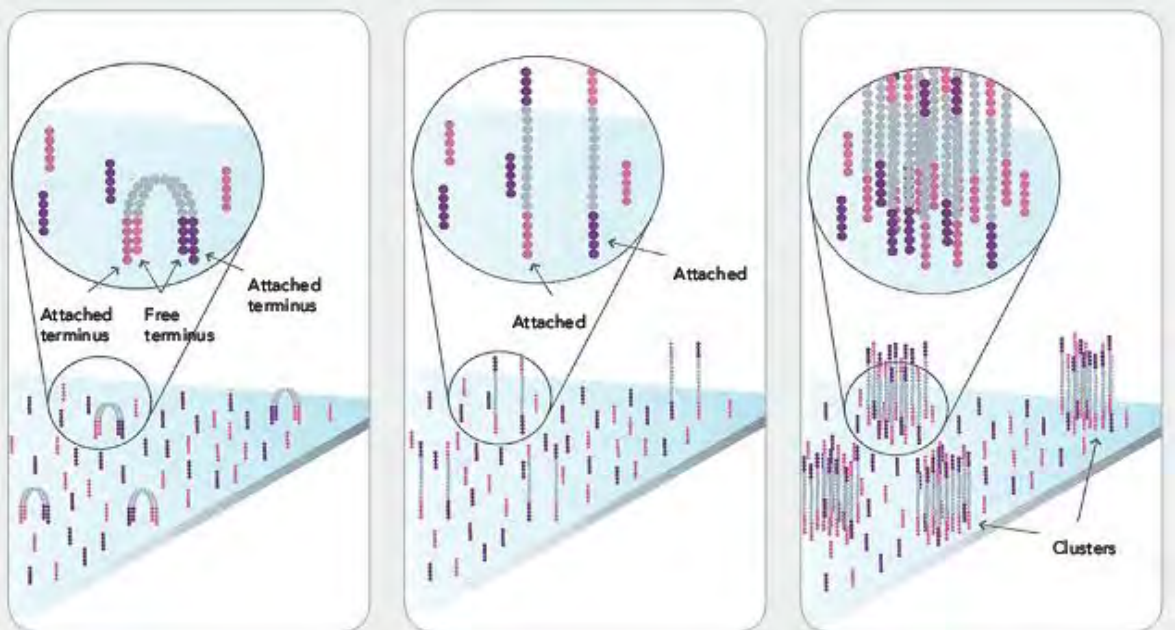


Illumina.com

Illumina sequencing (formerly Solexa)

@ColauttiLab



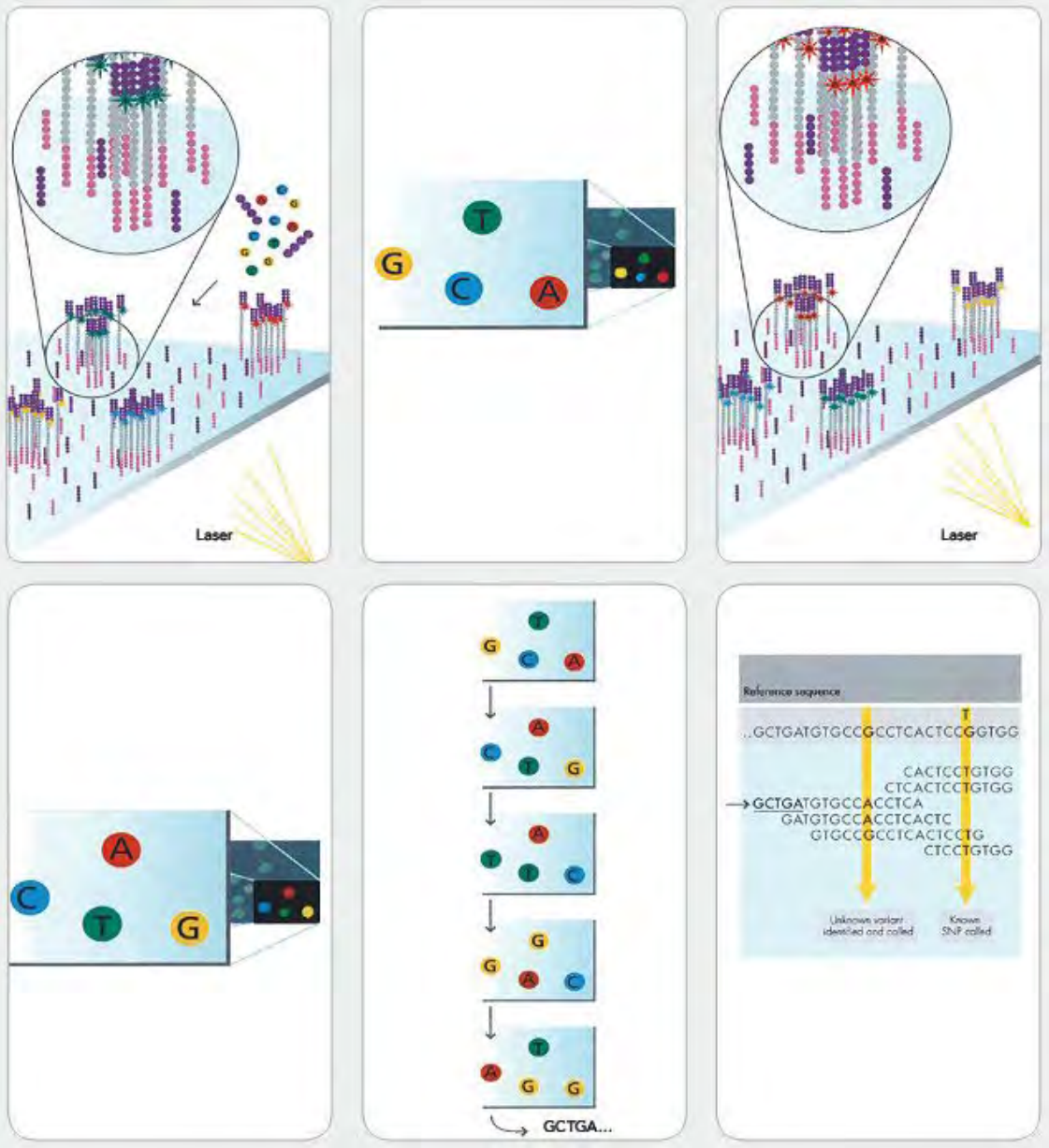


1. Prepare genomic DNA
2. Attach DNA to surface
3. Bridge amplification
4. Fragment become double stranded
5. Denature the double stranded molecules
6. Complete amplification

Illumina.com

Illumina sequencing

@ColauttiLab



7. Determine first base
8. Image first base
9. Determine second base
10. Image second base
11. Sequence reads over multiple cycles
12. Align data

Illumina Devices (benchtop)

@ColauttiLab

iSeq 100

MiniSeq

MiSeq Series

NextSeq 550 Series

NextSeq 1000 & 2000

Popular Applications & Methods	Key Application	Key Application	Key Application	Key Application	Key Application
Large Whole-Genome Sequencing (human, plant, animal)					
Small Whole-Genome Sequencing (microbe, virus)	●	●	●	●	●
Exome & Large Panel Sequencing (enrichment-based)				●	●
Targeted Gene Sequencing (amplicon-based, gene panel)	●	●	●	●	●
Single-Cell Profiling (scRNA-Seq, scDNA-Seq, oligo tagging assays)				●	●
Transcriptome Sequencing (total RNA-Seq, mRNA-Seq, gene expression profiling)				●	●
Targeted Gene Expression Profiling	●	●	●	●	●
miRNA & Small RNA Analysis	●	●	●	●	●
DNA-Protein Interaction Analysis (ChIP-Seq)			●	●	●
Methylation Sequencing				●	●
16S Metagenomic Sequencing		●	●	●	●
Metagenomic Profiling (shotgun metagenomics, metatranscriptomics)				●	●
Cell-Free Sequencing & Liquid Biopsy Analysis				●	●

Illumina.com

Illumina Devices (industrial scale)

@ColauttiLab

NextSeq 550 Series

NextSeq 1000 & 2000

NovaSeq 6000

Popular Applications & Methods	Key Application	Key Application	Key Application
Large Whole-Genome Sequencing (human, plant, animal)			●
Small Whole-Genome Sequencing (microbe, virus)	●	●	●
Exome & Large Panel Sequencing (enrichment-based)	●	●	●
Targeted Gene Sequencing (amplicon-based, gene panel)	●	●	●
Single-Cell Profiling (scRNA-Seq, scDNA-Seq, oligo tagging assays)	●	●	●
Transcriptome Sequencing (total RNA-Seq, mRNA-Seq, gene expression profiling)	●	●	●
Chromatin Analysis (ATAC-Seq, ChIP-Seq)	●	●	●
Methylation Sequencing	●	●	●
Metagenomic Profiling (shotgun metagenomics, metatranscriptomics)	●	●	●
Cell-Free Sequencing & Liquid Biopsy Analysis	●	●	●

Illumina.com

@ColauttiLab

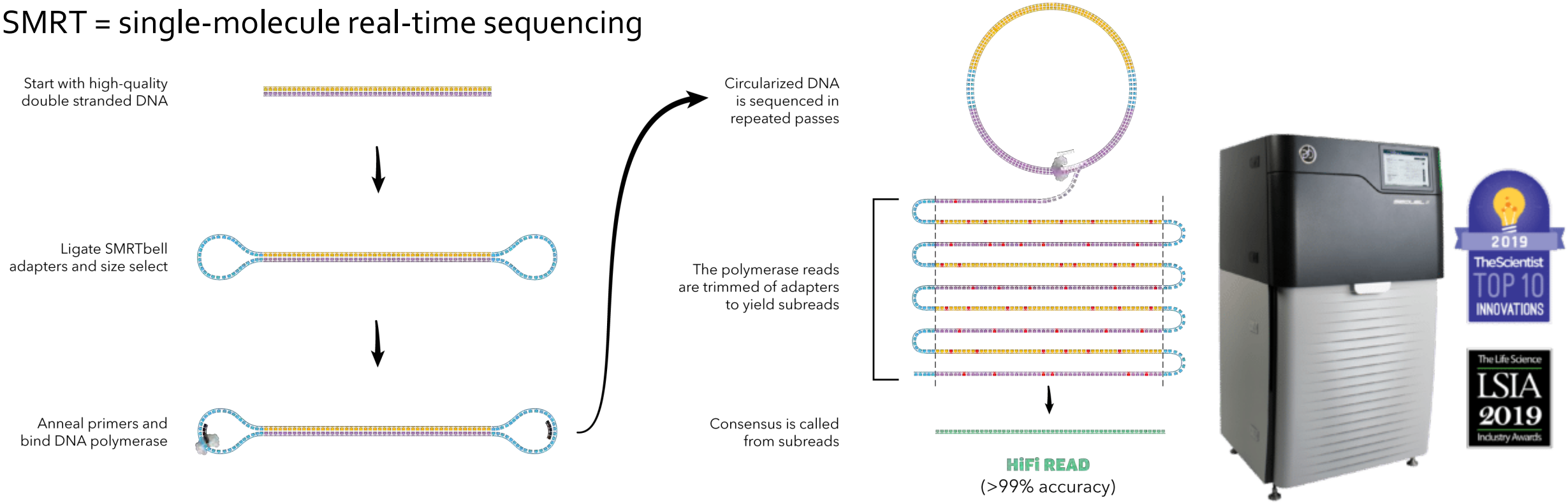
Third-generation sequencing

Long-read sequencing

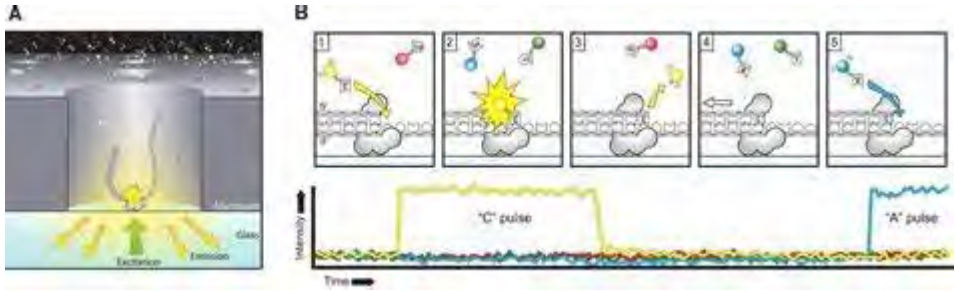
3rd Generation sequencing: SMRT

@ColauttiLab

SMRT = single-molecule real-time sequencing



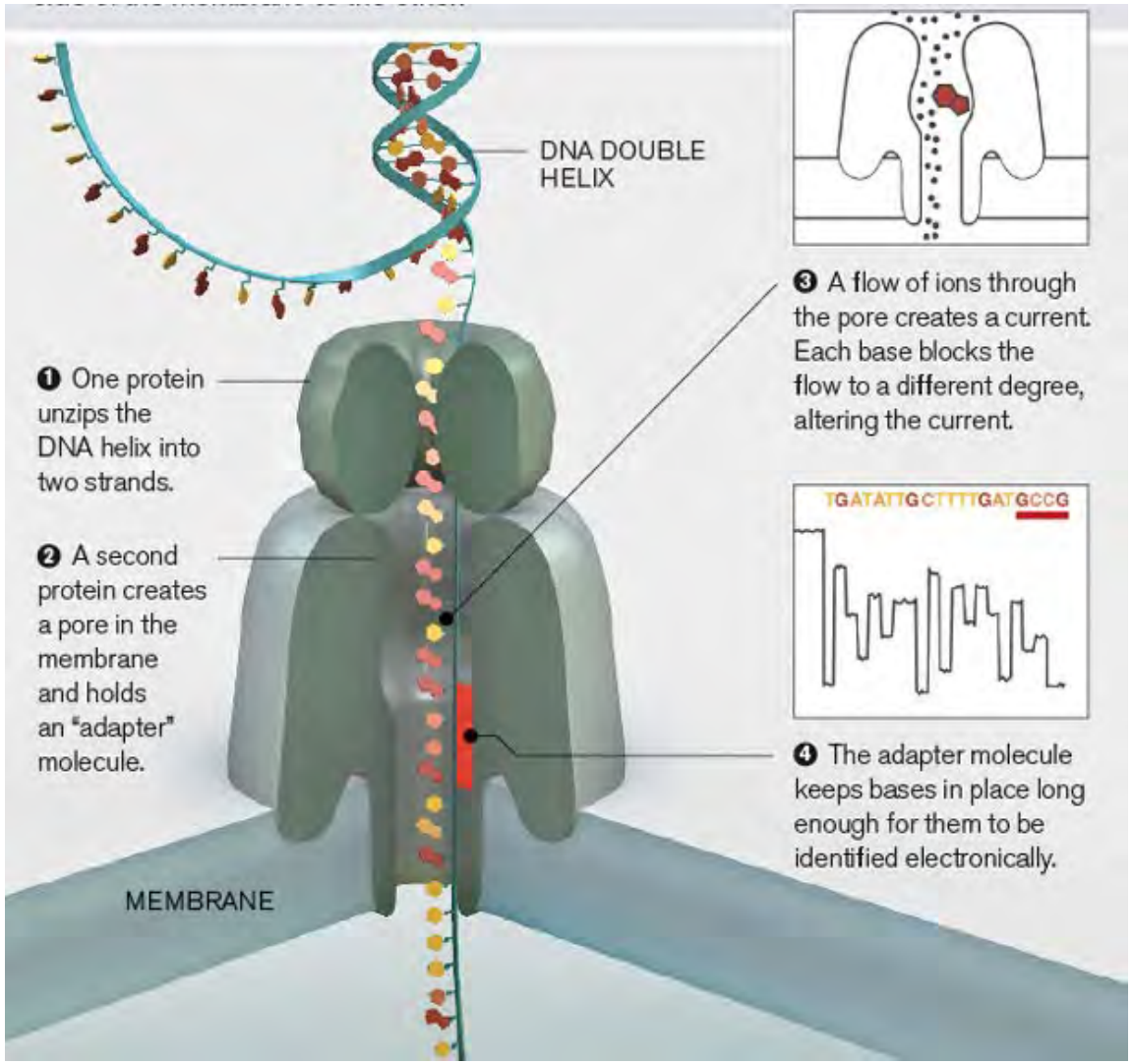
	Sequel IIe System	Sequel II System	Sequel System
Supported SMRT Cell	SMRT Cell 8M	SMRT Cell 8M	SMRT Cell 1M
Number of HiFi Reads >99%* Accuracy	Up to 4,000,000	Up to 4,000,000	Up to 500,000
Sequencing Run Time per SMRT Cell	Up to 30 hrs	Up to 30 hrs	Up to 20 hrs



Nanoporetech.com
Science.org

3rd Generation sequencing: Nanopore

@ColauttiLab



Nanoporetech.com

Nanopore Sequencing Comparison (2019)

@ColauttiLab

Platform	Instrument	Mreads	Length (bp)	Gbp	Type
Illumina	NovaSeq 6000 S4	10000	300	3000	SR & PE
Illumina	NextSeq 500 High-Output	400	300	120	SR & PE
Illumina	HiSeq X	375	300	112.5	PE
Illumina	HiSeq High-Output v4	250	250	62.5	SR & PE
Illumina	MiSeq v3	25	600	15	SR & PE
Illumina	MiniSeq High-Output	25	300	7.5	SR & PE
Oxford Nanopore	MinION		1M+	20	SR
Oxford Nanopore	PromethION			1000	SR
Ion	Proton I	60	200	12	SR
Ion	PGM 318	4	400	1.6	SR
Ion	PGM 316	2	400	0.8	SR
Ion	PGM 314	0.4	400	0.16	SR
PacBio	PacBio Sequel	0.37	20000	7.4	SR
PacBio	PacBio RS II (P6)	0.055	15000	0.825	SR
Roche 454	GS FLX+ / FLX	0.7	700	0.49	SR
SOLiD	5500xl W	267	100	26.7	SR & PE

Review

A set of bioinformatics packages for R

```
install.packages("BiocManager")
Update all/some/none [a/s/n] choose n
library(BiocManager)
install(c("sangerseqR","annotate"))
```

Working in groups (15 mins):

Stretch and divide into working groups

Summarize sequencing-by-synthesis (SBS) with Illumina

Review key concepts:

1. How a flow cell works
2. Contrast Sanger with SBS sequencing

Try flowcharts or cartoons to simplify & summarize

BRAINSTORM:

What are the main benefits & limitations of each technology?

Why is coding valuable for 2nd generation sequencing?