

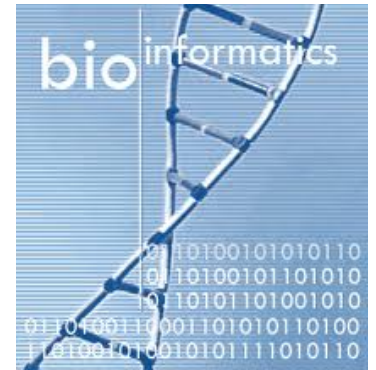
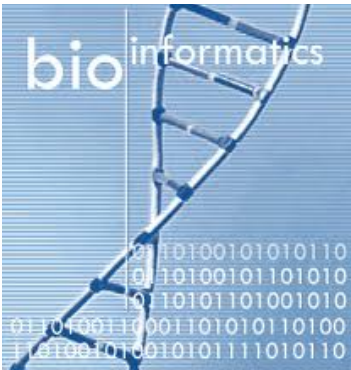
Bioinformatics

TEN

Introduction to Next Generation Sequencing

Wendy Lee

Dept of Computer Science
San José State University
Biology/CS/SE 123A
Fall 2014



Sequencing Technologies

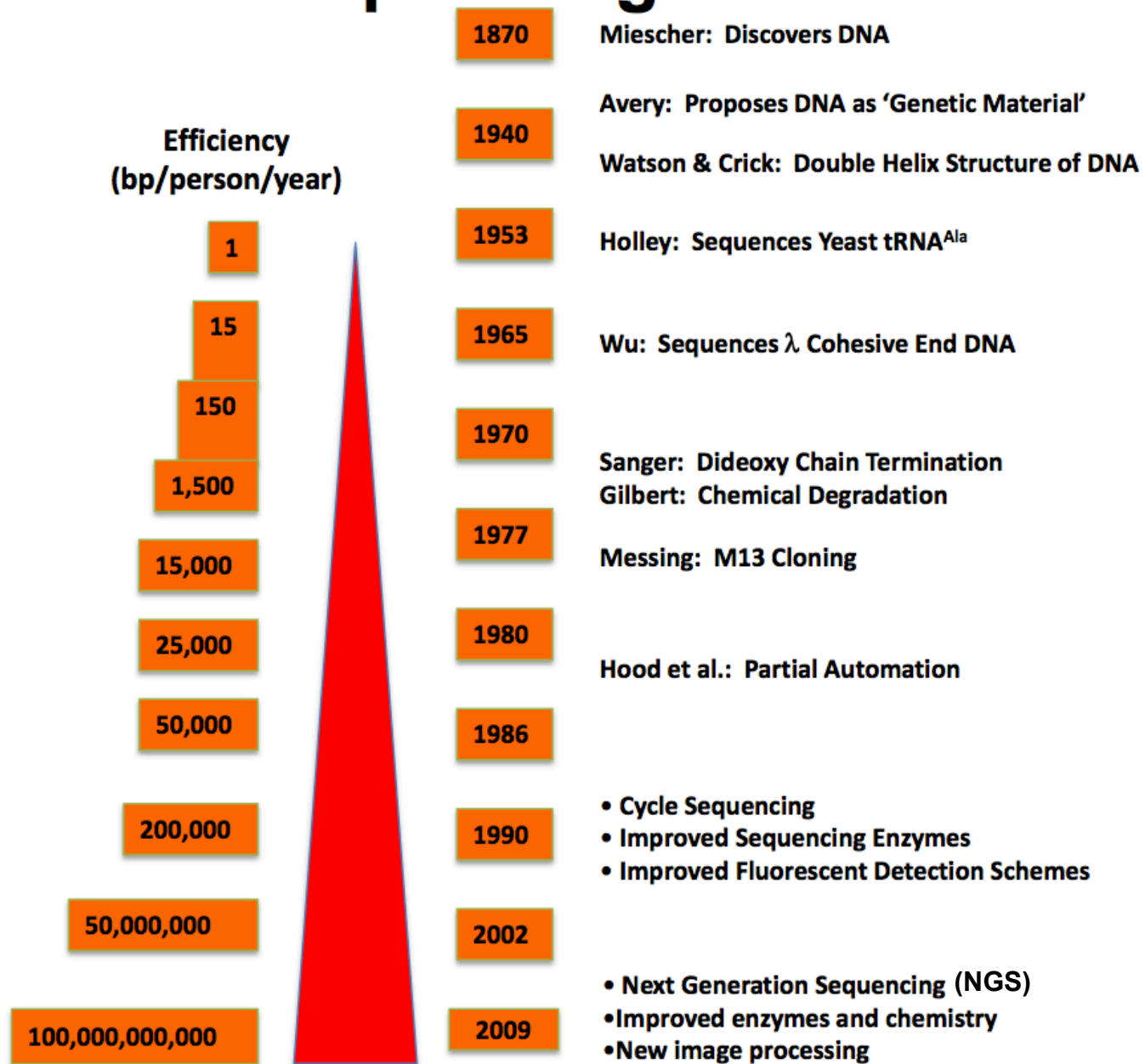
Traditional sequencing

- Sanger Sequencing

Next Generation Sequencing (NGS)

- Pyrosequencing
- Illumina/Solexa
- Ion Torrent (charge based detection)
- Helicos
- Pacific Biosciences
- Oxford Nanopore

History of DNA Sequencing



Adapted from Eric Green, NIH; Adapted from Messing & Liaca, *PNAS* (1998)

Sanger vs NGS

Sanger sequencing' has been the only DNA sequencing method for 30 years but...

...hunger for even greater sequencing throughput and more economical sequencing technology...

NGS has the ability to process millions of sequence reads in parallel rather than 96 at a time (1/6 of the cost)

NGS Platforms

- Illumina GAI, HiSeq, MiSeq
- Life Technologies Ion Torrent
- Helicos Heliscope™
- Pacific Biosciences SMRT
- Oxford Nanopore Technologies
- Roche/454 FLX
- Applied Biosystems SOLiD™ System

Illumina

Video: <https://www.youtube.com/watch?v=womKfikWlxM>

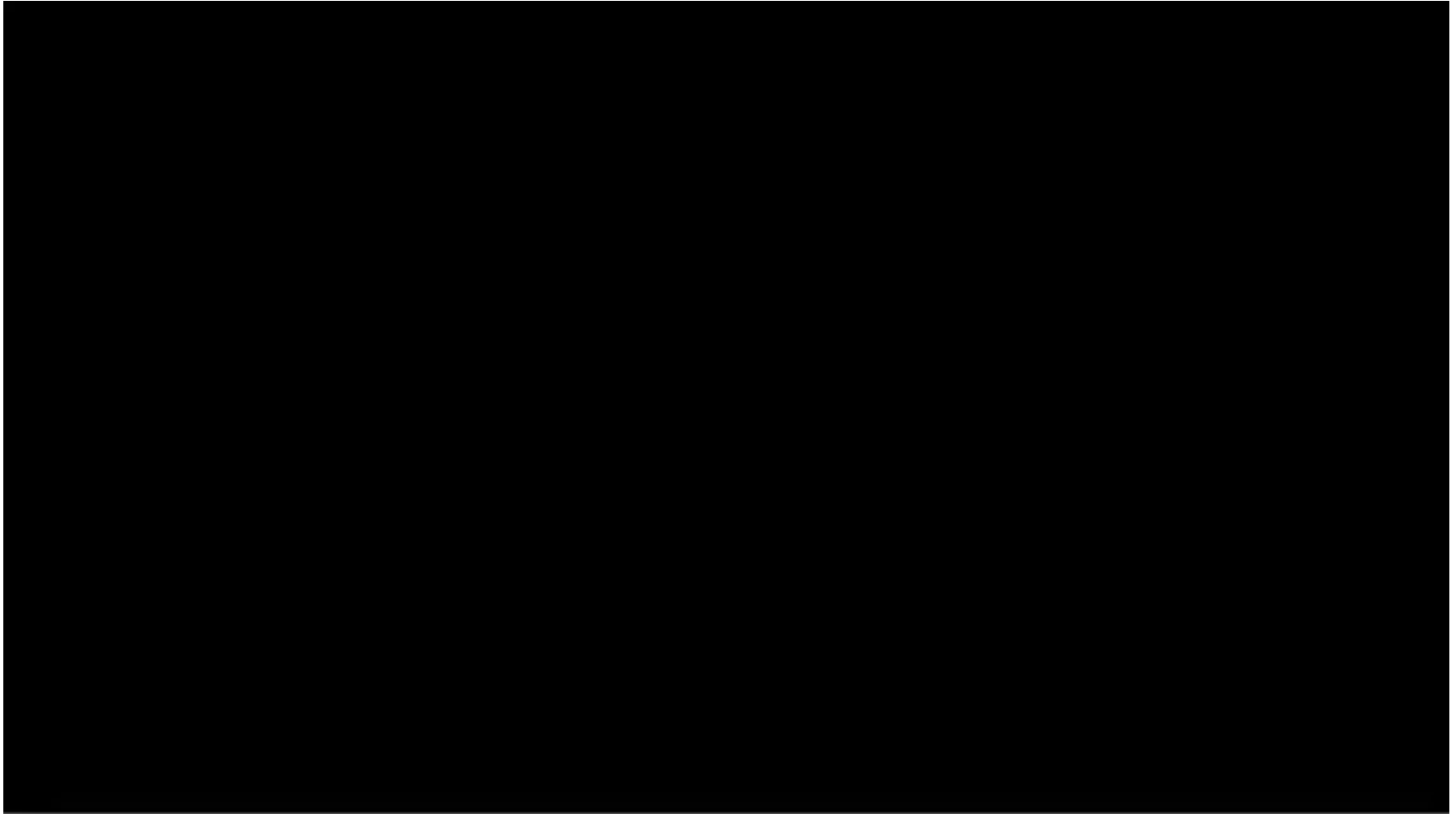
© 2014 Wendy Lee

Ion Torrent



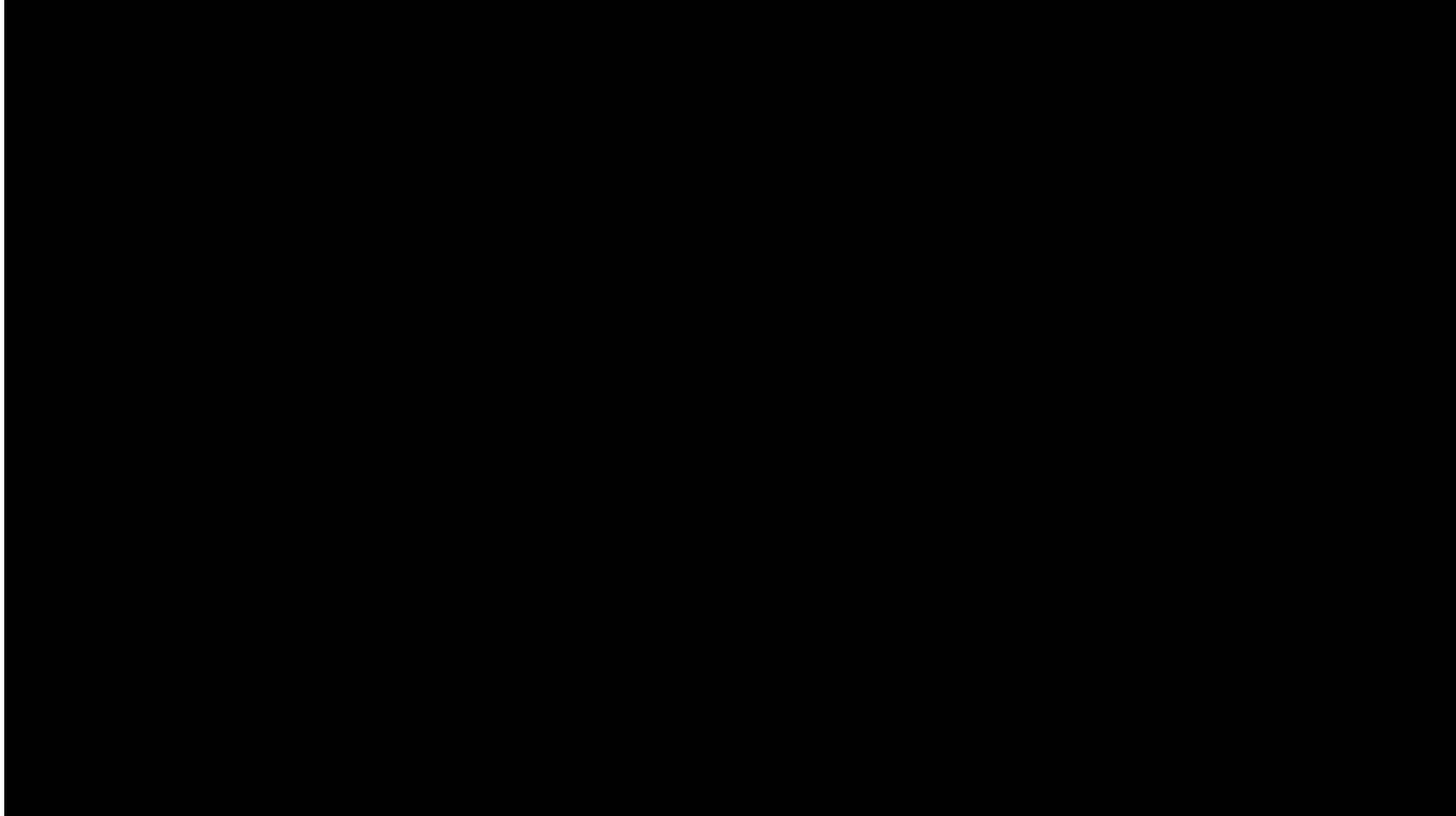
Video: <https://www.youtube.com/watch?v=WYBzbXlfuKs>

Helicos



Video: <https://www.youtube.com/watch?v=TboL7wODBj4>

Pacific Biosciencics SMRT



Video: <https://www.youtube.com/watch?v=v8p4ph2MAvI>

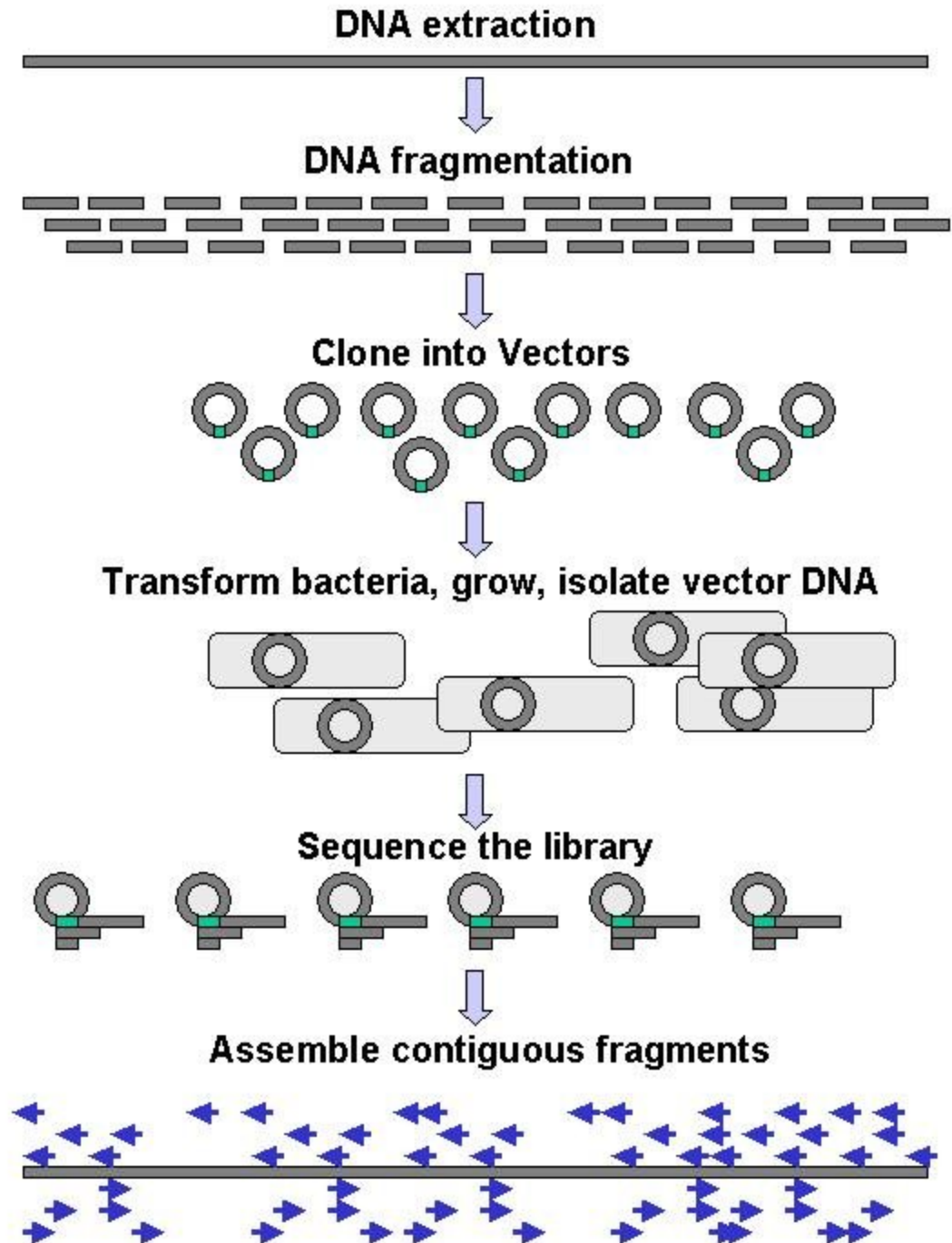
Oxford Nanopore Technology

Video: <https://www.youtube.com/watch?v=3UHw22hBpAk>

Next Generation Sequencing: Why Now?

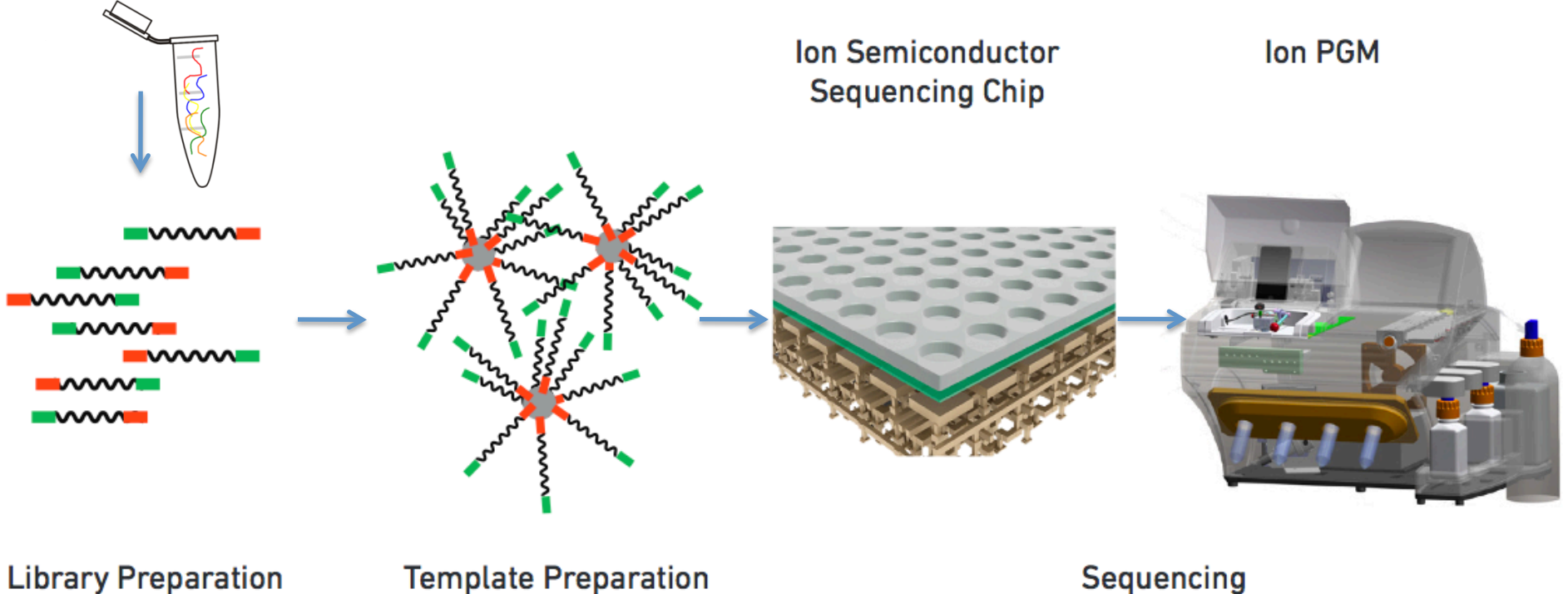
- **Motivation:** HGP and its derivatives, personalized medicine
- **Short reads applications:** (re-)sequencing, other methods (e.g. gene expression)
- Advancements in technology

Genomic DNA Sequencing using Sanger



Generation of Polony array: DNA Beads (Ion Torrent)

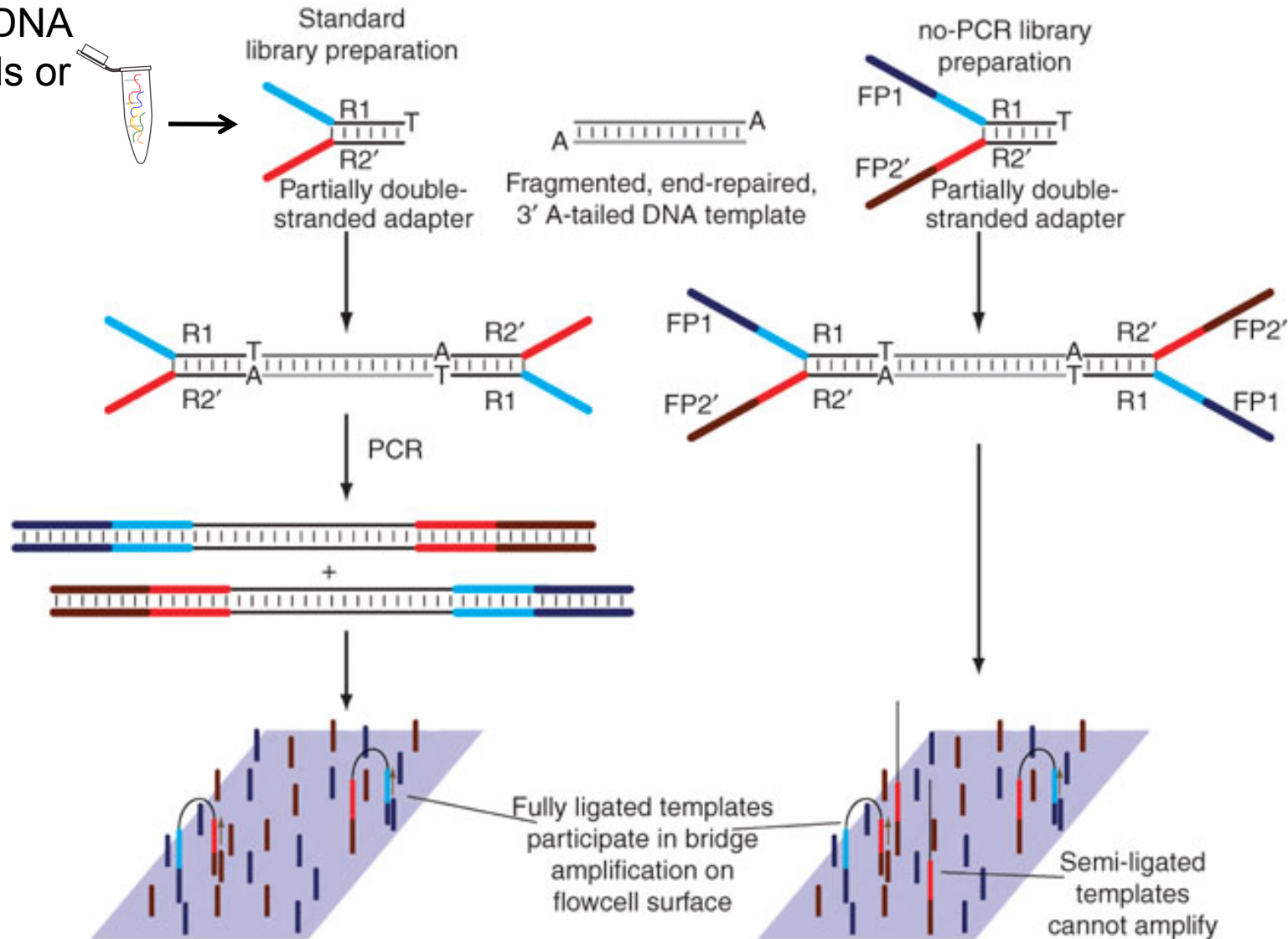
Extract DNA from cells or tissue



DNA Beads are placed in wells

Generation of Polony array: Bridge-amplification (Illumina)

Extract DNA
from cells or
tissue



Single Molecule Sequencing

(Helicos, PacBio SMRT, Oxford Nanopore)

- Direct sequencing of DNA molecules: no amplification stage
- DNA fragments are attached to array (Helicos & PacBio SMRT)
- Potential benefits: higher throughput, less errors (DNA amplification can introduce errors)

Comparing Different Technologies

Sanger Sequencing

Advantages	Disadvantages
<p data-bbox="386 716 747 760">Lowest error rate</p> <p data-bbox="386 829 963 878">Long read length (~750 bp)</p> <p data-bbox="386 943 800 992">Can target a primer</p>	<p data-bbox="1115 716 1514 764">High cost per base</p> <p data-bbox="1115 829 1688 878">Long time to generate data</p> <p data-bbox="1115 943 1465 992">Need for cloning</p> <p data-bbox="1115 1057 1738 1105">Amount of data per run is low</p>

Comparing Different Technologies

Ion Torrent Sequencing

Advantages	Disadvantages
Low startup costs	Relatively higher error rate than Illumina
Scalable (10 – 1000 Mb of data per run)	Difficult to enumerate long repeats such as homopolymer repeats of the same nucleotide (e.g. GGGGG)
Medium/low cost per base	Read lengths only ~100-400 bp so far
Low error rate	Relatively low throughput
Fast runs (<3 hours)	

Comparing Different Technologies

Illumina Sequencing

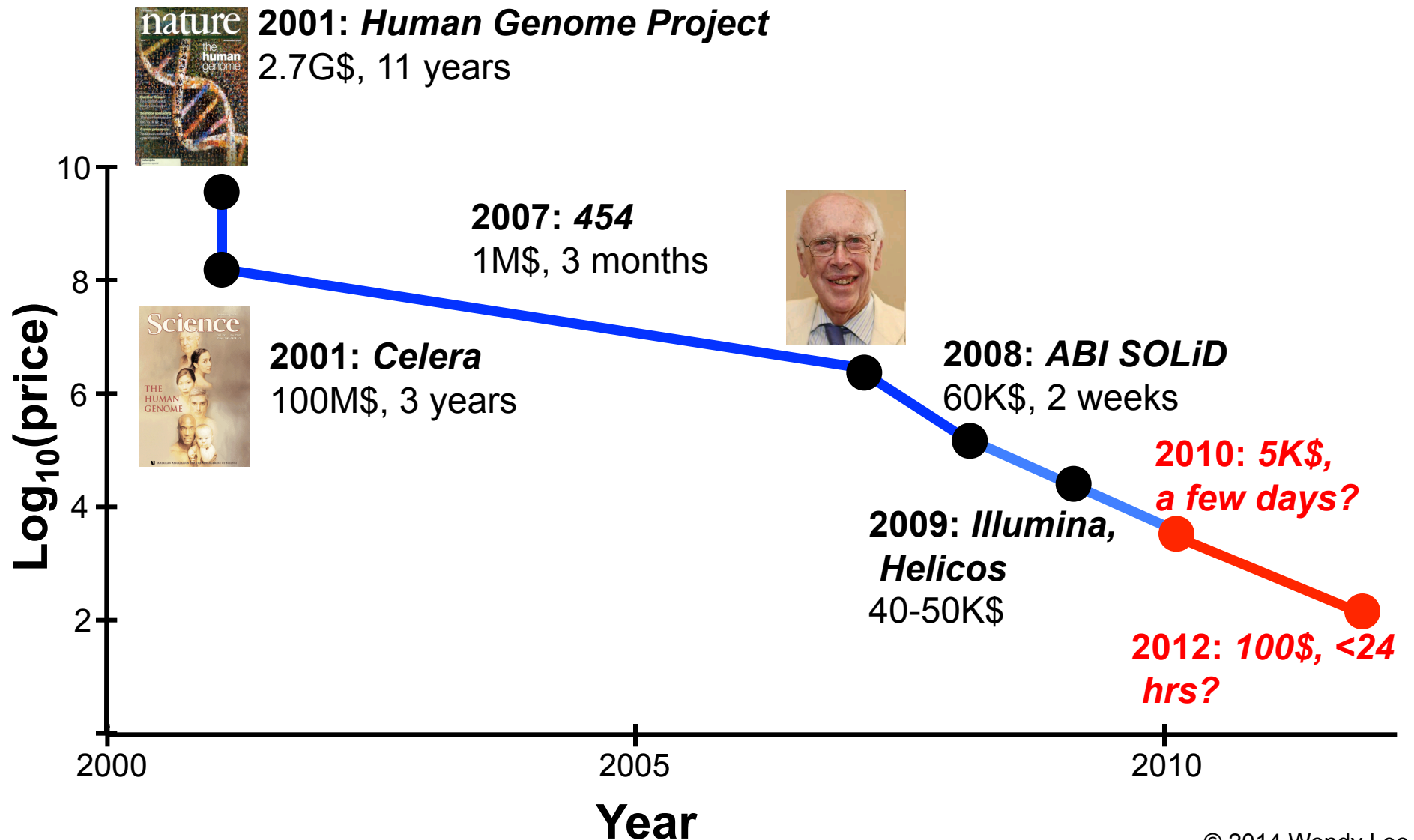
Advantages	Disadvantages
Low error rate (< 1%) Lowest cost per base Tons of data (15 Gb to 1.8 Tb)	Short read length (50-300 bp) Runs take multiple days High startup costs Difficult for De Novo assembly

Comparing Different Technologies

PacBio Sequencing

Advantages	Disadvantages
<p>Can use single molecule as template</p> <p>Potential for very long reads (several kb+)</p>	<p>High error rate (~10-15%)</p> <p>Medium/high cost per base</p> <p>High startup costs</p>

Sequencing the Human Genome



The interpretation bottleneck

