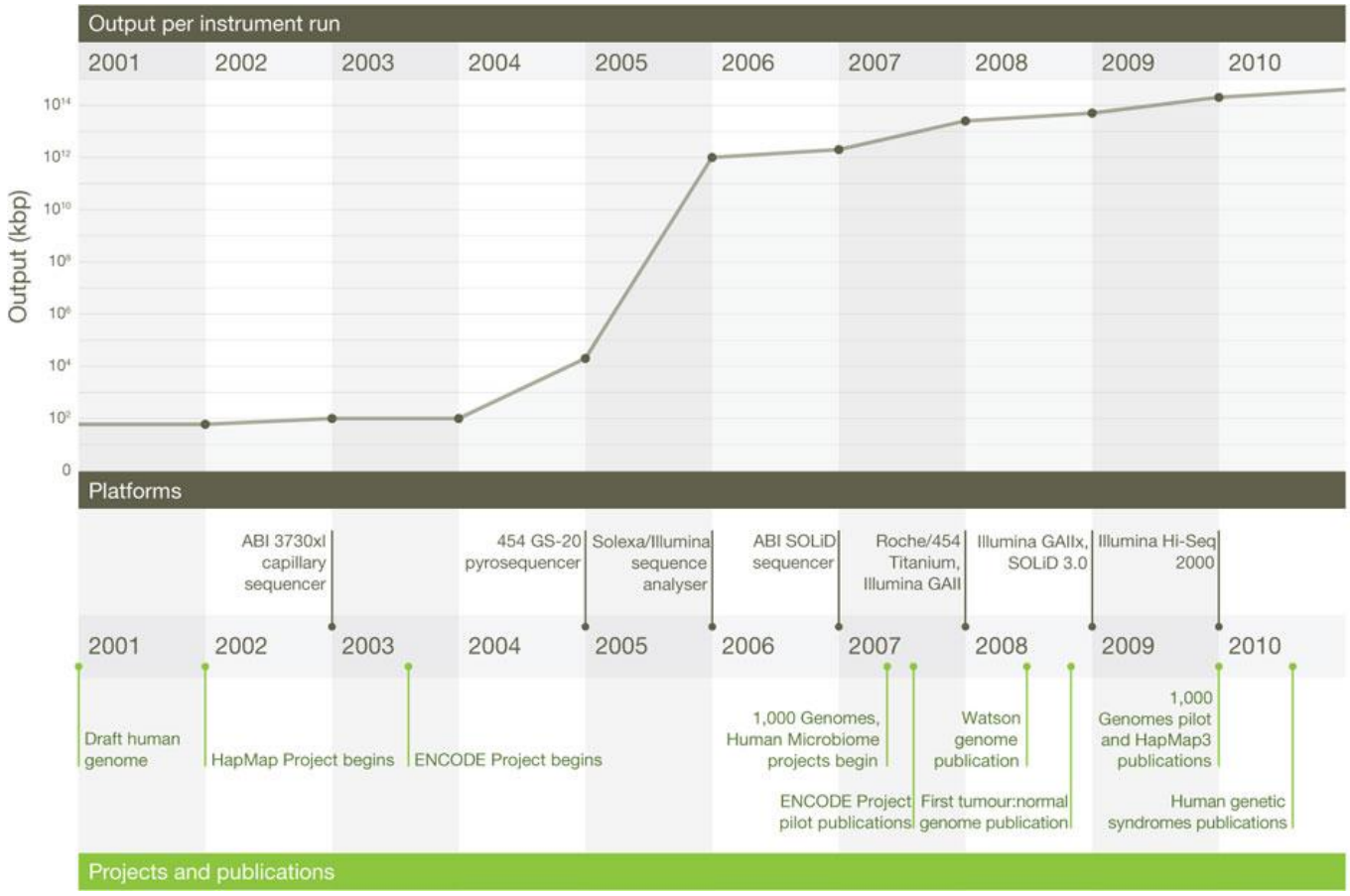
Introduction to Next-generation Sequencing Technology

Peng Liu

Next-Generation Sequencing Instrumentation

1. Roche 454 sequencer
2. Illumina (Solexa sequencing) most popular
3. SOLiD sequencer (Applied Biosystems / Life Technologies)
4. Pacific Biosciences (single molecule sequencing)
5. Ion Torrent (pH detection, Life Technologies)
6. Oxford Nanopore (current detection)
7. Many other emerging technologies

Changes in instrument capacity over the past decade, and the timing of major sequencing projects



NGS data output has reached terabase (Tb, 10^{12}) of data in a single sequencing run—nearly a 1000× increase since 2007 to 2011.

It took 10 years to sequence the first human genome. Now 5 human genomes in a week (Illumina).

ER Mardis. *Nature* 470, 198-203 (2011) doi:10.1038/nature09796

Next-generation DNA sequencing instruments

- All commercially-available sequencers have the following shared attributes:
  - Random fragmentation of starting DNA, ligation with custom linkers/adapters = “a library”
  - Library amplification on a solid surface (either bead or flat surface) eg. flow cell eg. 454 sequencing
  - Direct step-by-step detection of each nucleotide base incorporated during the sequencing reaction
  - Hundreds of thousands to hundreds of millions of reactions imaged per instrument run = “massively parallel sequencing”

From E. Mardis lecture about Next-generation sequencing technologies

Next-generation DNA sequencing instruments

- Most of current commercially-available sequencers have the following shared attributes (continued):
  - Shorter read lengths than capillary sequencers (Sanger sequencing)
  - A “digital” read type that enables direct quantitative comparisons
  - A sequencing mechanism that samples both ends of every fragment sequenced (“paired end” reads)
- From E. Mardis lecture about Next-generation sequencing technologies

Illumina Sequencing Platforms
http://www.illumina.com/systems/sequencing-platforms.html

Sequencing Platforms | C

www.illumina.com/systems/sequencing-platforms.html

Apps | Google Translate | Rafael Irizarry | edX | Tomorrow's Professor | Campus Parent Portal | ISU Stat Dept. (@ISU_ | ssize | Department of Statist | GCCC 2016-2017 | Other bookmarks

illumina®

SIGN IN | VIEW CART | CONTACT US | SEARCH

ALL SYSTEMS | Sequencing Platforms | Microarray Scanners | NGS PLATFORM SELECTOR | QUESTIONS

	NextSeq Series	HiSeq Series	NovaSeq Series	HiSeq X Series†
Popular Applications & Methods	Key Application	Key Application	Key Application	Key Application
Large Whole-Genome Sequencing (human, plant, animal)	●	●	●	●
Small Whole-Genome Sequencing (microbe, virus)	●	●	●	
Exome Sequencing	●	●	●	
Targeted Gene Sequencing (amplicon, gene panel)	●	●	●	
Whole-Transcriptome Sequencing	●	●	●	
Gene Expression Profiling with mRNA-Seq	●	●	●	
miRNA & Small RNA Analysis	●	●	●	
DNA-Protein Interaction Analysis	●	●	●	
Methylation Sequencing	●	●	●	
Shotgun Metagenomics	●	●	●	

Optimized NGS Sample Tracking and Workflows

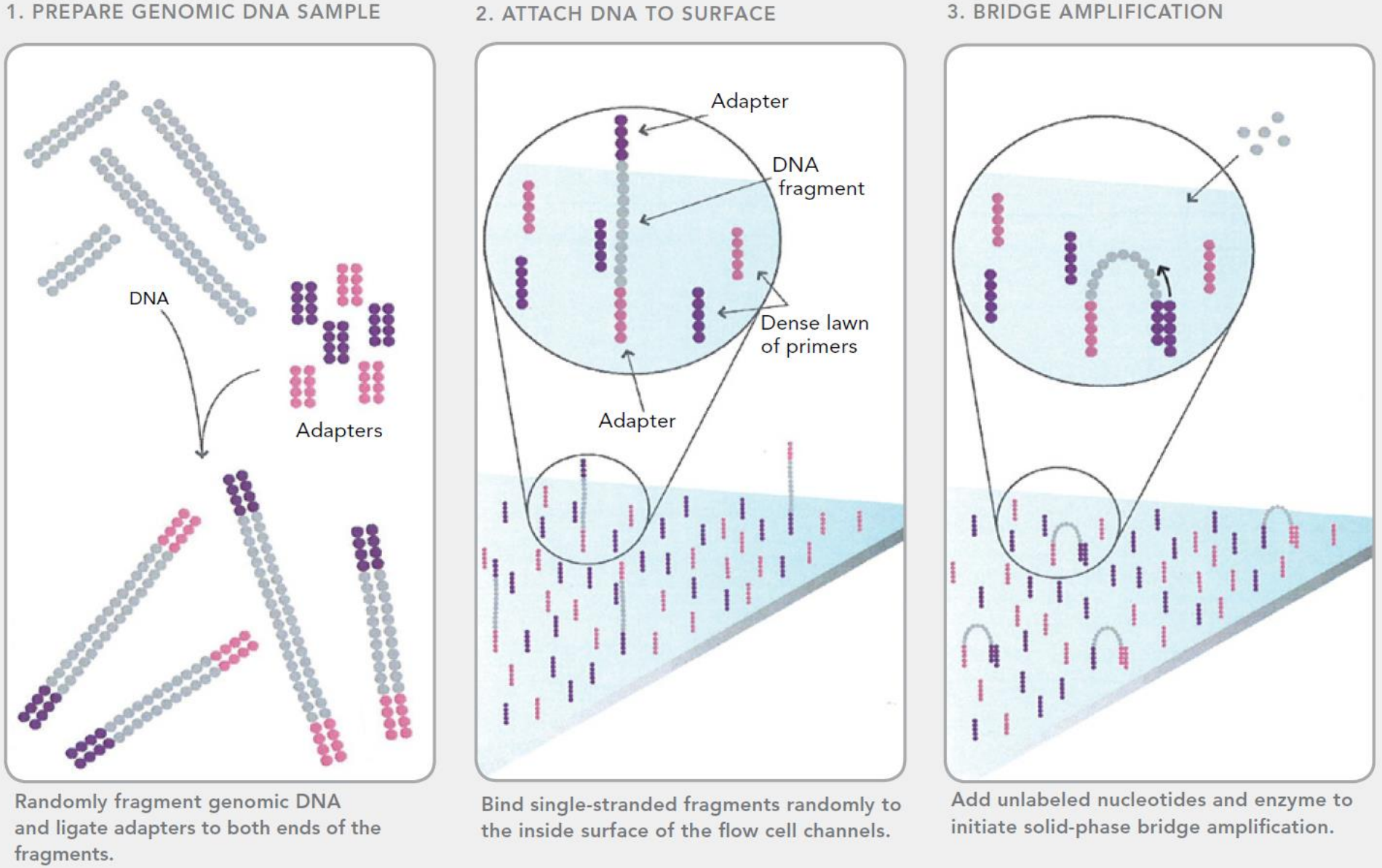
Illumina Instruments at ISU
http://www.dna.iastate.edu/illumina-sequencing.html

Table 1 - Output of Available Genome Sequencing Instruments		
Instrument	Average # Reads (Millions) Per Lane[d]	Available Read Lengths (bp)
Illumina HiSeq 3000		
	320	50, 100 and 150
Illumina MiSeq		
v2	15	50, 300 and 500
v3	25	150 and 600

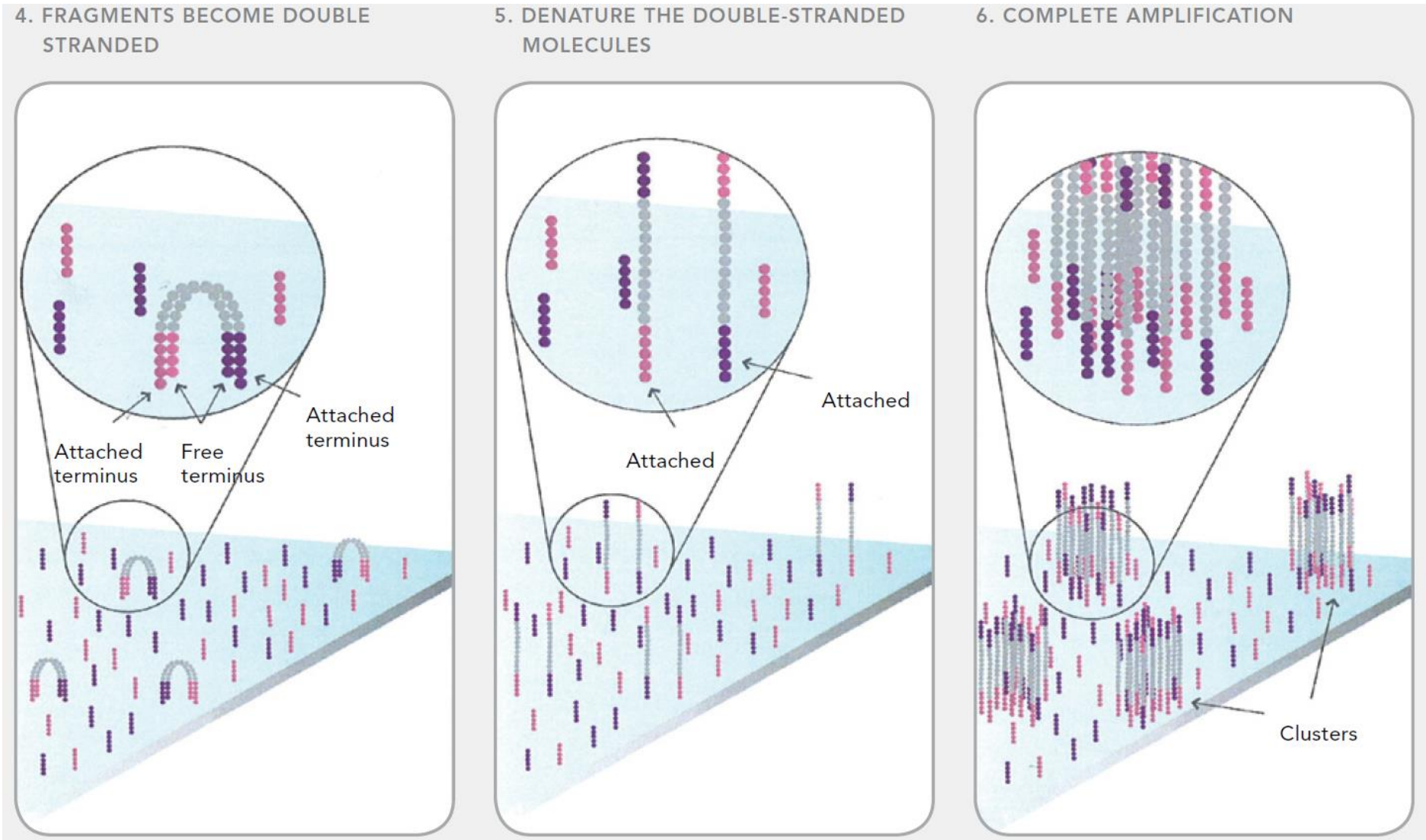
Illumina (Solexa) sequencing by synthesis procedure

- Documents and video from Illumina©
- Introductions from Illumina (required reading):
  - http://www.illumina.com/documents/products/techspotlights/techspotlight_sequencing.pdf
- Sequencing Technology Video: http://www.illumina.com/technology/next-generation-sequencing/sequencing-technology.html

Illumina (Solexa) sequencing by synthesis

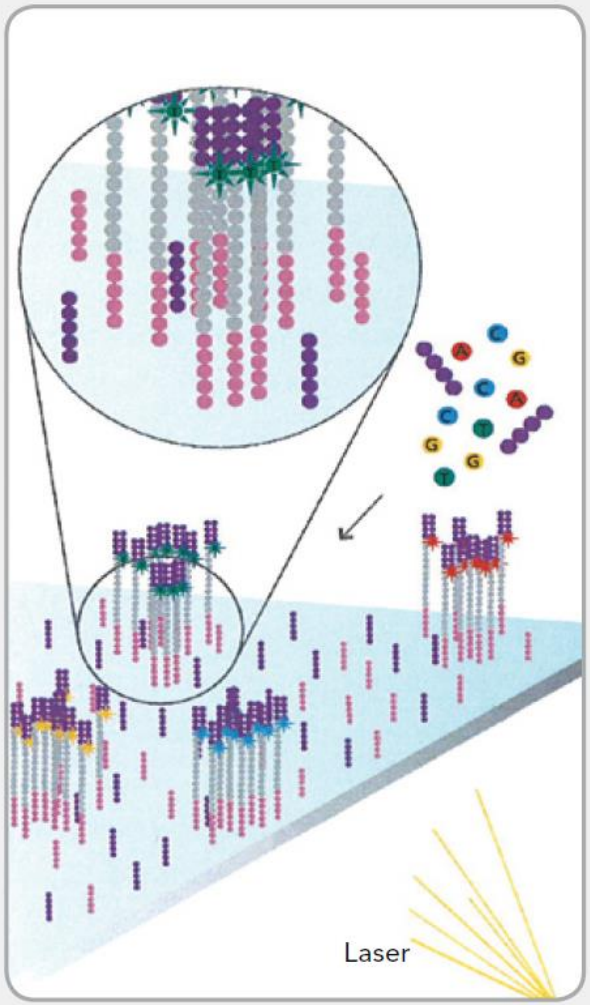


Illumina (Solexa) sequencing by synthesis



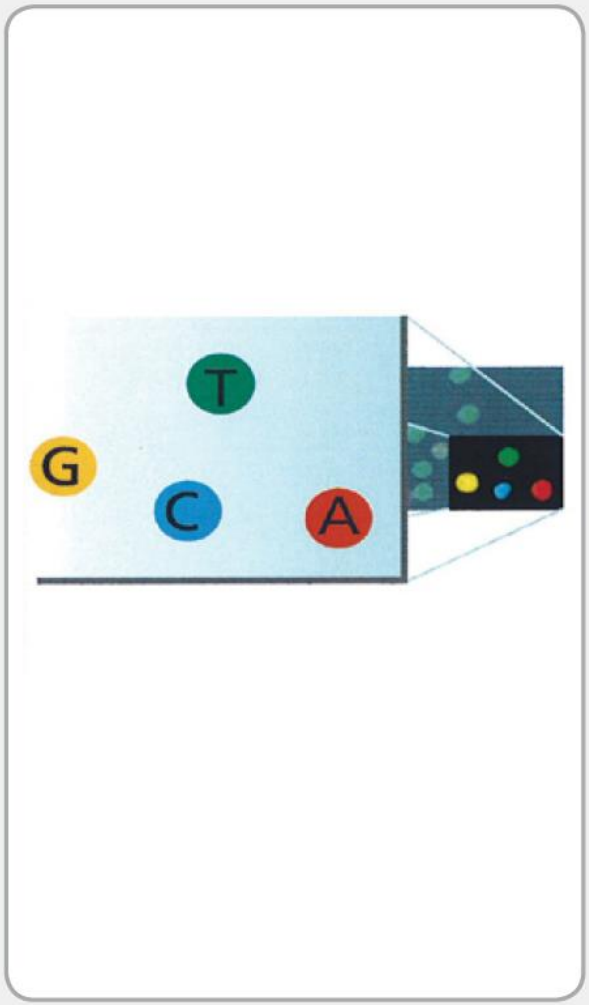
Illumina (Solexa) sequencing by synthesis

7. DETERMINE FIRST BASE



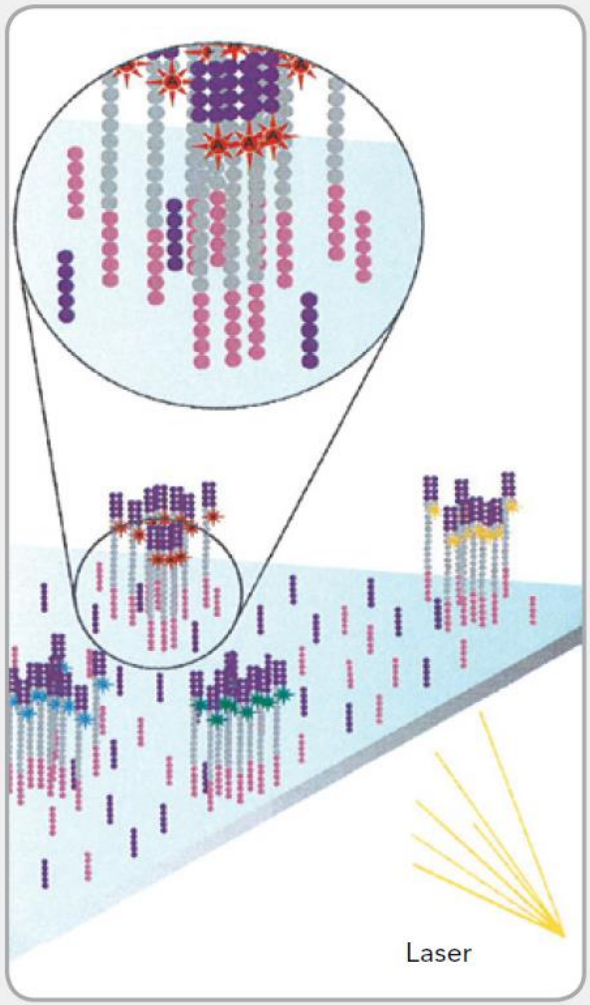
First chemistry cycle: to initiate the first sequencing cycle, add all four labeled reversible terminators, primers and DNA polymerase enzyme to the flow cell.

8. IMAGE FIRST BASE



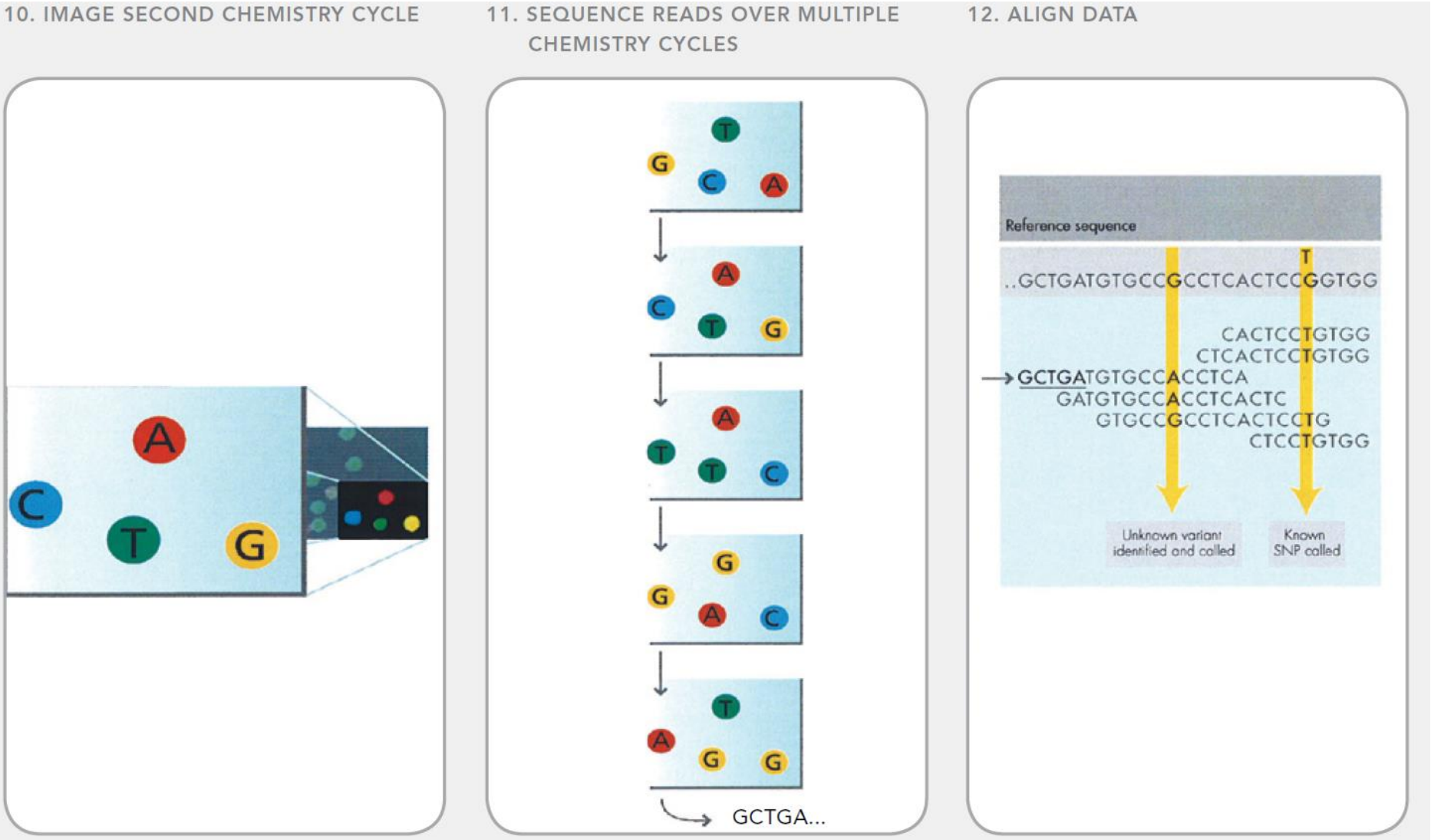
After laser excitation, capture the image of emitted fluorescence from each cluster on the flow cell. Record the identity of the first base for each cluster.

9. DETERMINE SECOND BASE



Second chemistry cycle: to initiate the next sequencing cycle, add all four labeled reversible terminators and enzyme to the flow cell.

Illumina (Solexa) sequencing by synthesis

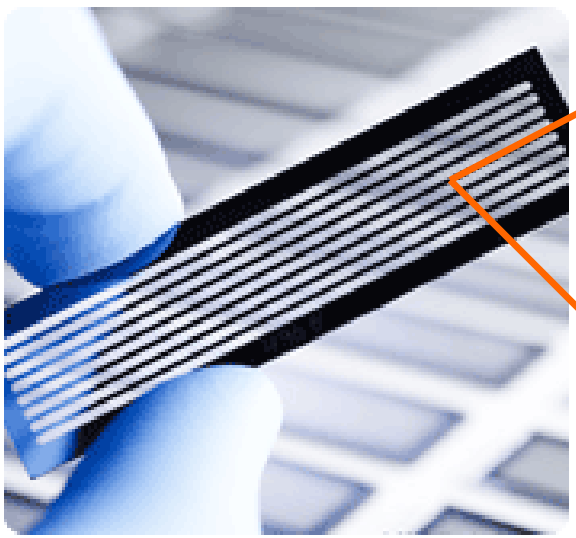


After laser excitation, collect the image data as before. Record the identity of the second base for each cluster.

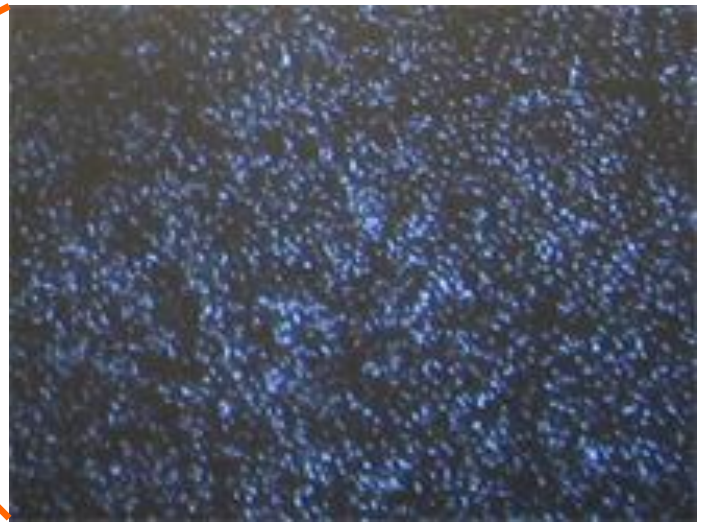
Repeat cycles of sequencing to determine the sequence of bases in a given fragment a single base at time.

Align data, compare to a reference, and identify sequence differences.

Image analysis



Flow cells

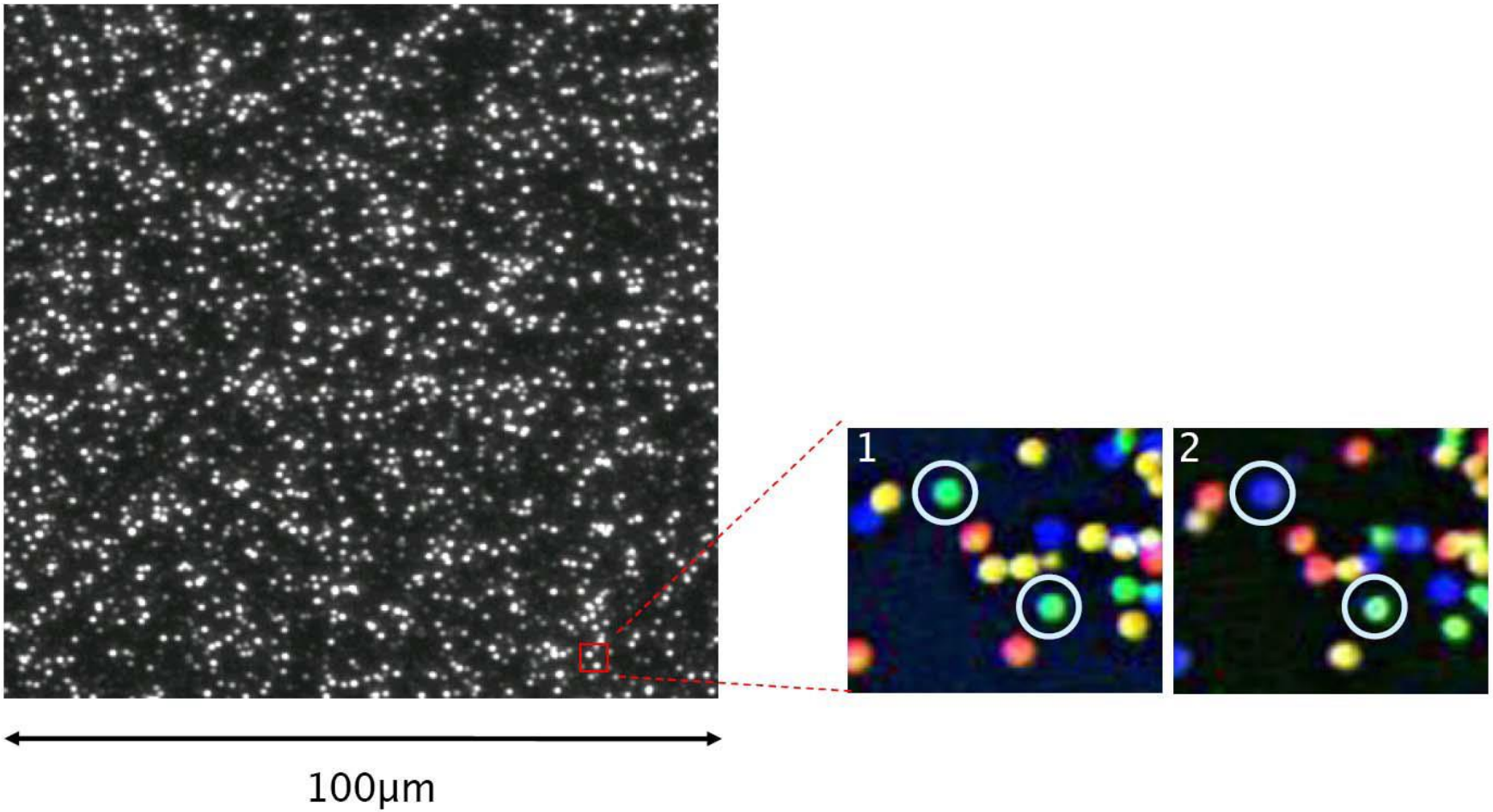


Clusters

Images copied from http://www.vmsr.net/Sequencing_slides-Intro_Seminar.pdf

Solexa

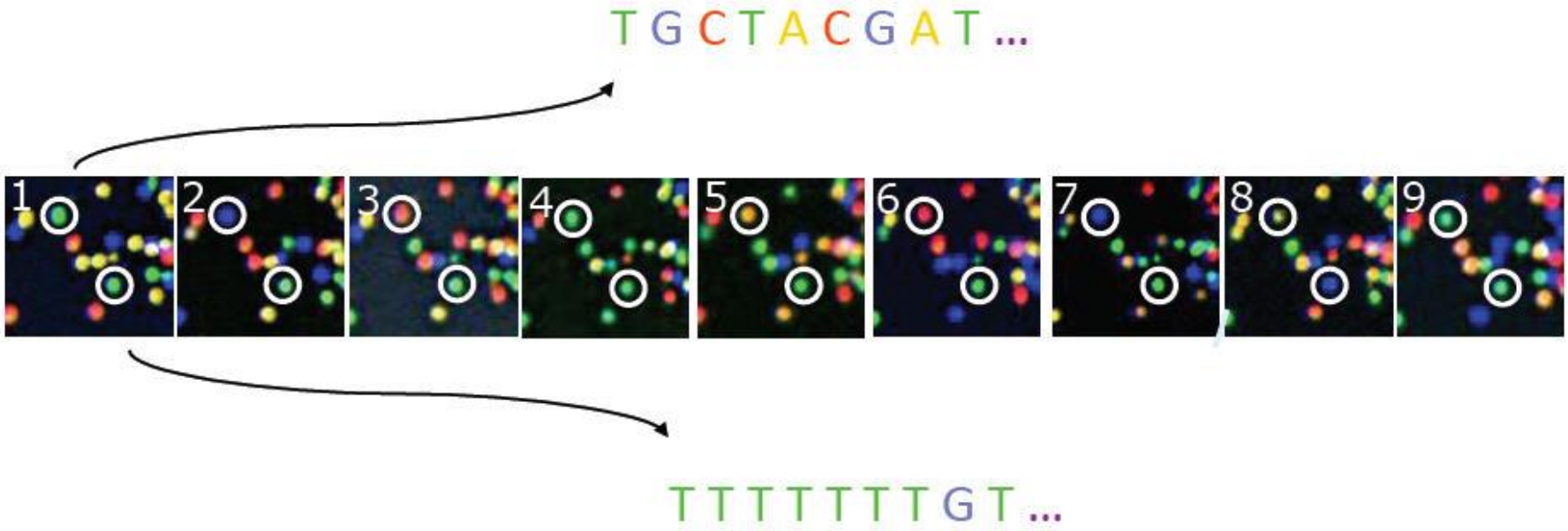
Sequencing by synthesis (SBS)



colony of ≈1000 single-stranded DNA templates

Copied from http://www.vmsr.net/Sequencing_slides-Intro_Seminar.pdf

Base Calling from Raw Data



The identity of each base of a cluster is read off from sequential images.

Slide (though modified) Courtesy of Illumina Inc.

Paired-end sequencing

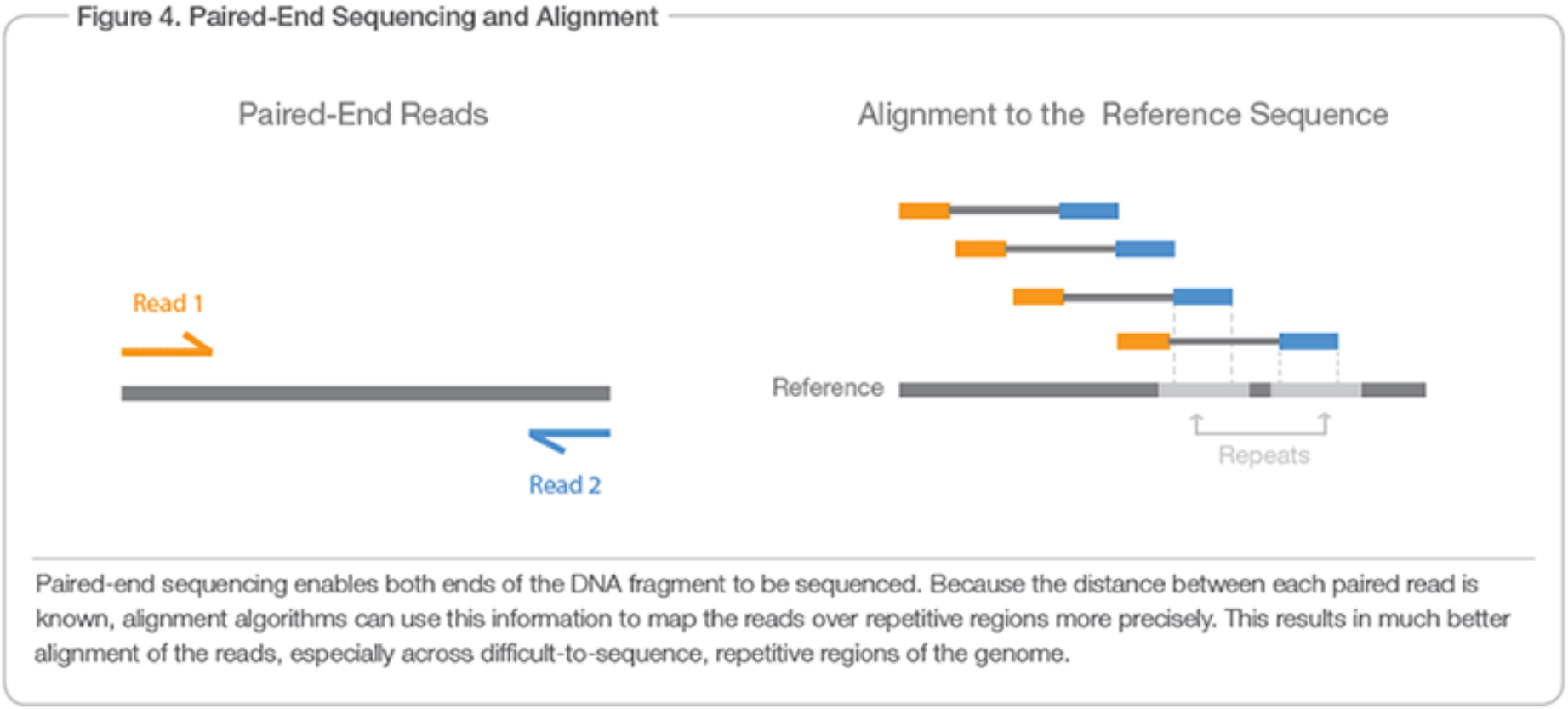
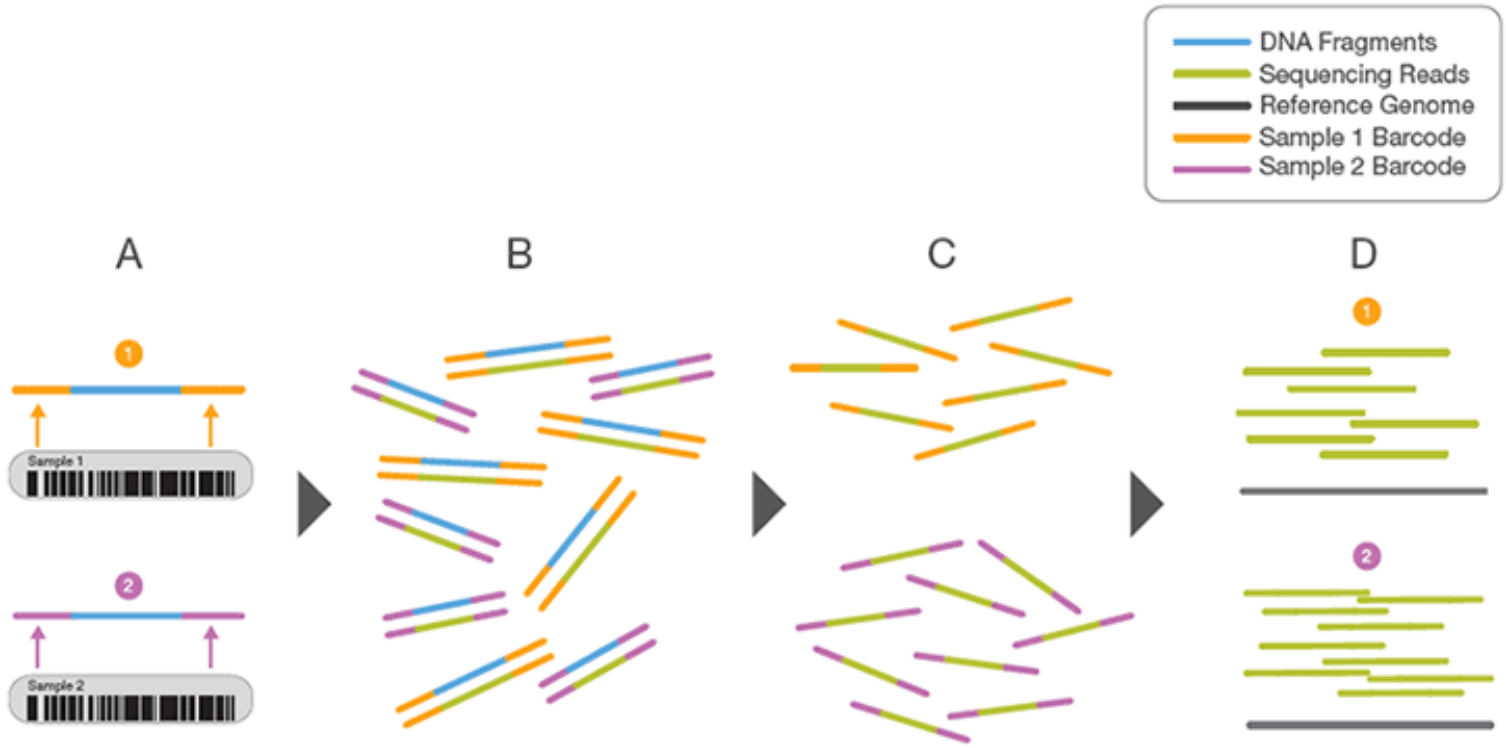


Figure from Illumina webpage

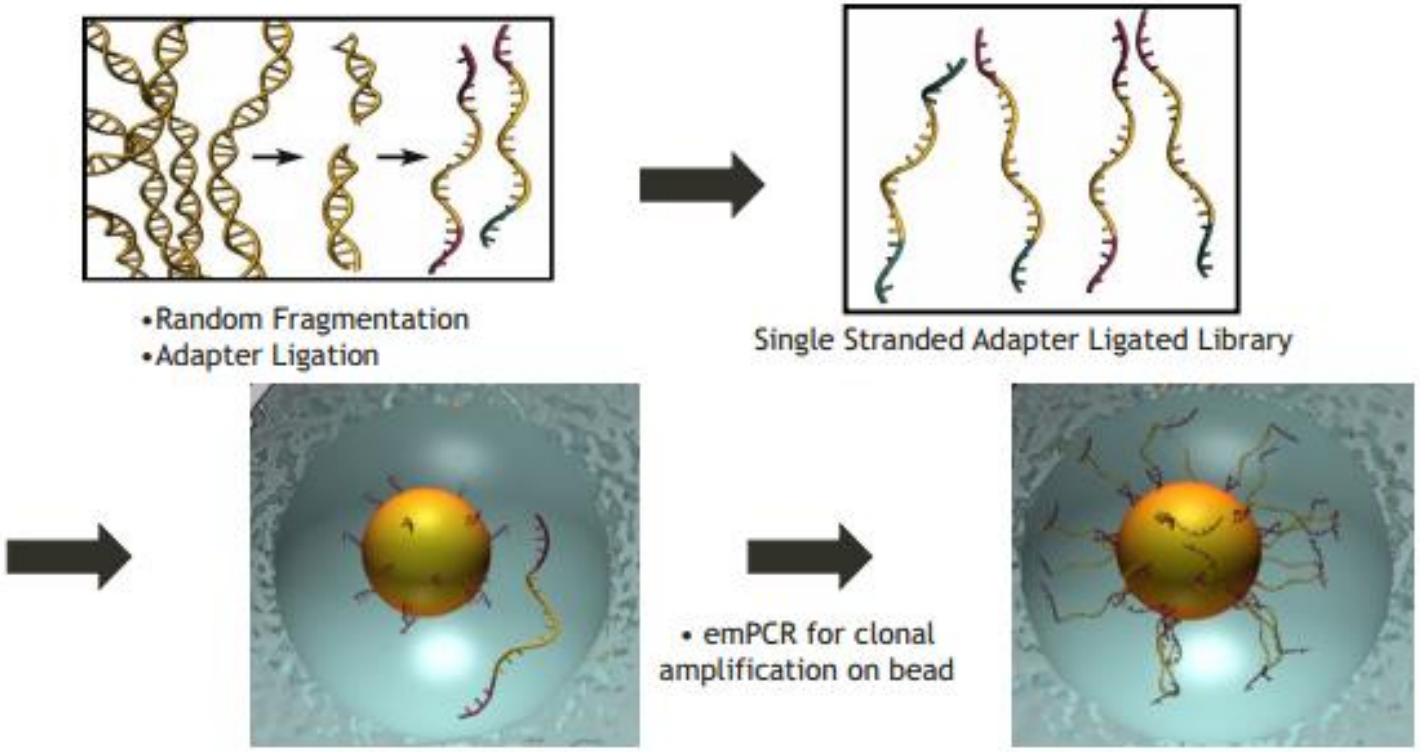
Multiplex sequencing: individual “barcode” sequences are added to each sample so they can be differentiated during the data analysis

Figure 2: Conceptual Overview of Sample Multiplexing



A. Two representative DNA fragments from two unique samples, each attached to a specific barcode sequence that identifies the sample from which it originated.
B. Libraries for each sample are pooled and sequenced in parallel. Each new read contains both the fragment sequence and its sample-identifying barcode.
C. Barcode sequences are used to de-multiplex, or differentiate reads from each sample.
D. Each set of reads is aligned to the reference sequence.

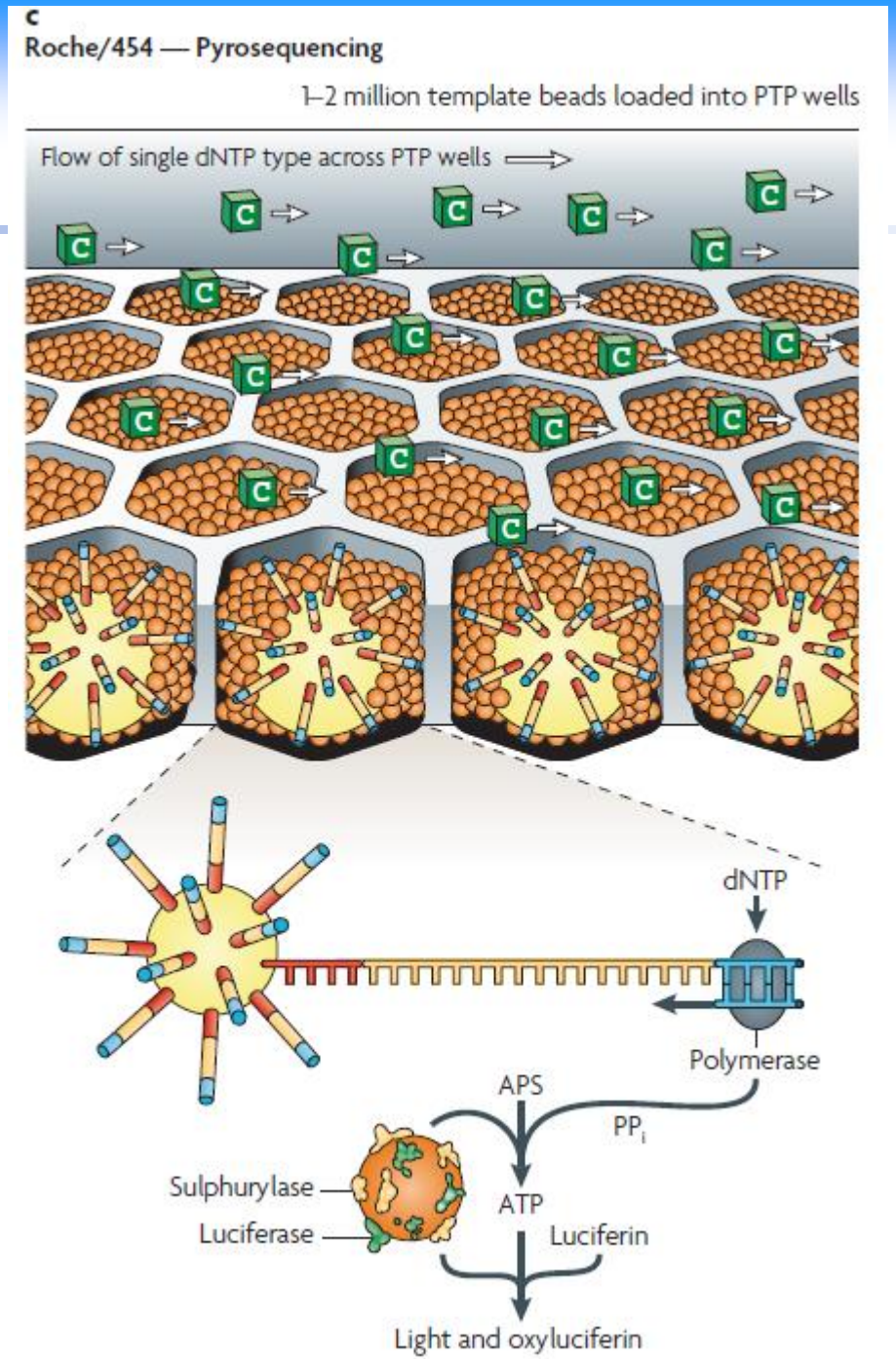
454 sequencing



E. Mardis lecture slide for the Current Topics in Genome Analysis 2012

454 sequencing

Metzker, 2010,
Nature Reviews Genetics



454 sequencing

- Earliest next-gen sequencing
- Run time 10-20 hours
- Output 100-900 MB/run
- Read length: 400-1000
- Applications:
  - *de novo* sequencing
  - Variation detection
  - Gene expression
  - Metagenomics

Newer technologies are being developed

- Newer sequencing technology (3rd generation?)
  - Pacific Biosciences (Single-Molecule, Real-Time (SMRT) technology, http://www.pacificbiosciences.com/products/smrt-technology/)
  - Ion Torrent (http://www.lifetechnologies.com/us/en/home/life-science/sequencing/sequencing-education.html)
  - Oxford Nanopore (https://nanoporetech.com/technology/introduction-to-nanopore-sensing/introduction-to-nanopore-sensing)
  - Many more (http://allseq.com/knowledgebank/sequencing-platforms/illumina)
  - Faster runs, longer reads, more throughput.
  - Check each company's webpage for more details.

Comments

- Different technologies have different pros and cons.
- Different technologies differ in the way they generate clusters and sequence clusters.
- Check http://allseq.com/knowledgebank/sequencing-platforms for more details about other technologies.
- Check http://www.dna.iastate.edu/nextgensequencing.html for ISU services for NGS.

References for the technology

- Metzker, M.L. (2010) Sequencing technologies – the next generation, Nature Reviews Genetics, 11: 31
- Mardis, E. R. (2011) A decade's perspective on DNA sequencing technology, Nature 470, 198–203
- Ozsolak, F. and Milos P. M. (2011) RNA sequencing: advances, challenges and opportunities, Nature Reviews Genetics 12, 87-98
- Shapiro et al., (2013) Single-cell sequencing based technologies will revolutionize whole-organism science, Nature Reviews Genetics, 14: 618-630

References for the technology

- Current Topics in Genome Analysis 2016, Lecture: Next-Generation Sequencing Technologies by E. Mardis
- Goodwin et al. (2016) Coming of age: ten years of nextgeneration sequencing technologies, Nature 17: 333-351 (http://www.nature.com/nrg/journal/v17/n6/pdf/nrg.2016.49.pdf)
- Nature collection of reviews on NGS: http://www.nature.com/subjects/next-generation-sequencing#research-and-reviews
- Company webpages