Sequencing Technologies

Joann Mudge, Ph.D.
Anitha Sundararajan, Ph.D.
Callum J. Bell, Ph.D.

New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

NCGR
National Center for Genome Resources

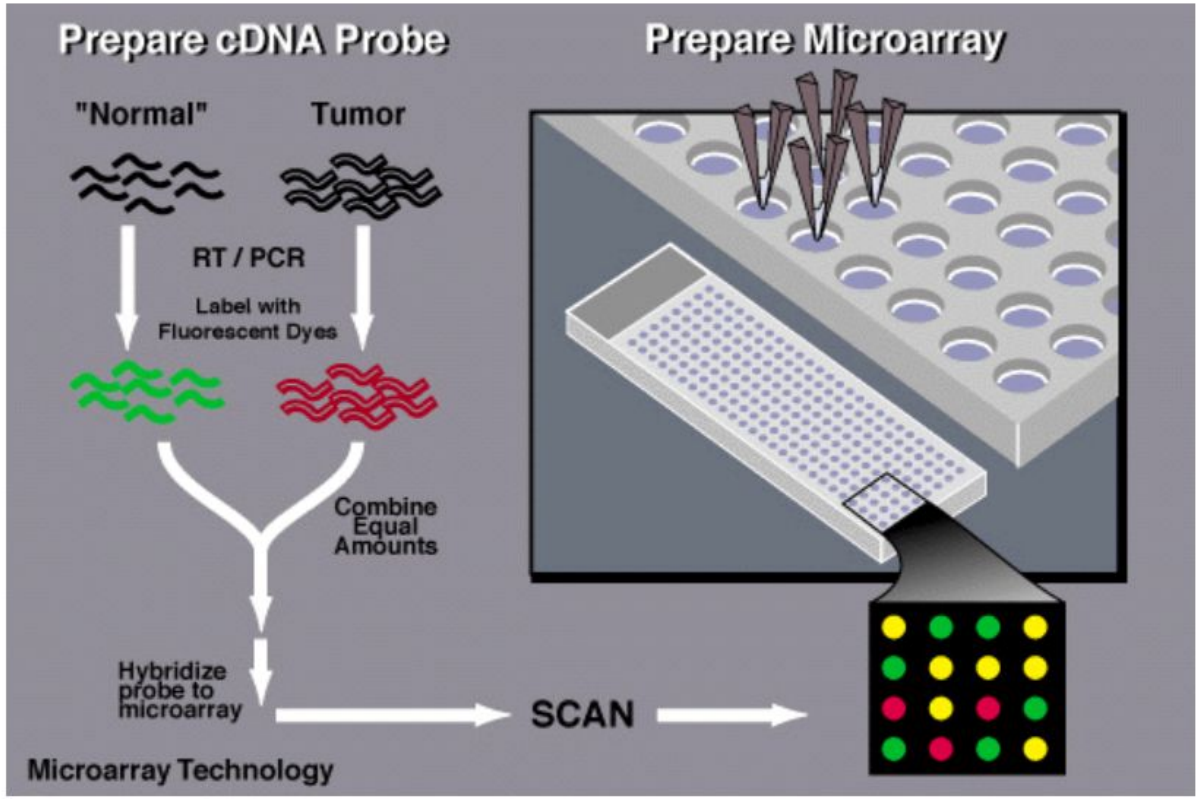
This is bioinformatics

https://www.smithsonianmag.com/games/spot-difference-180968040/

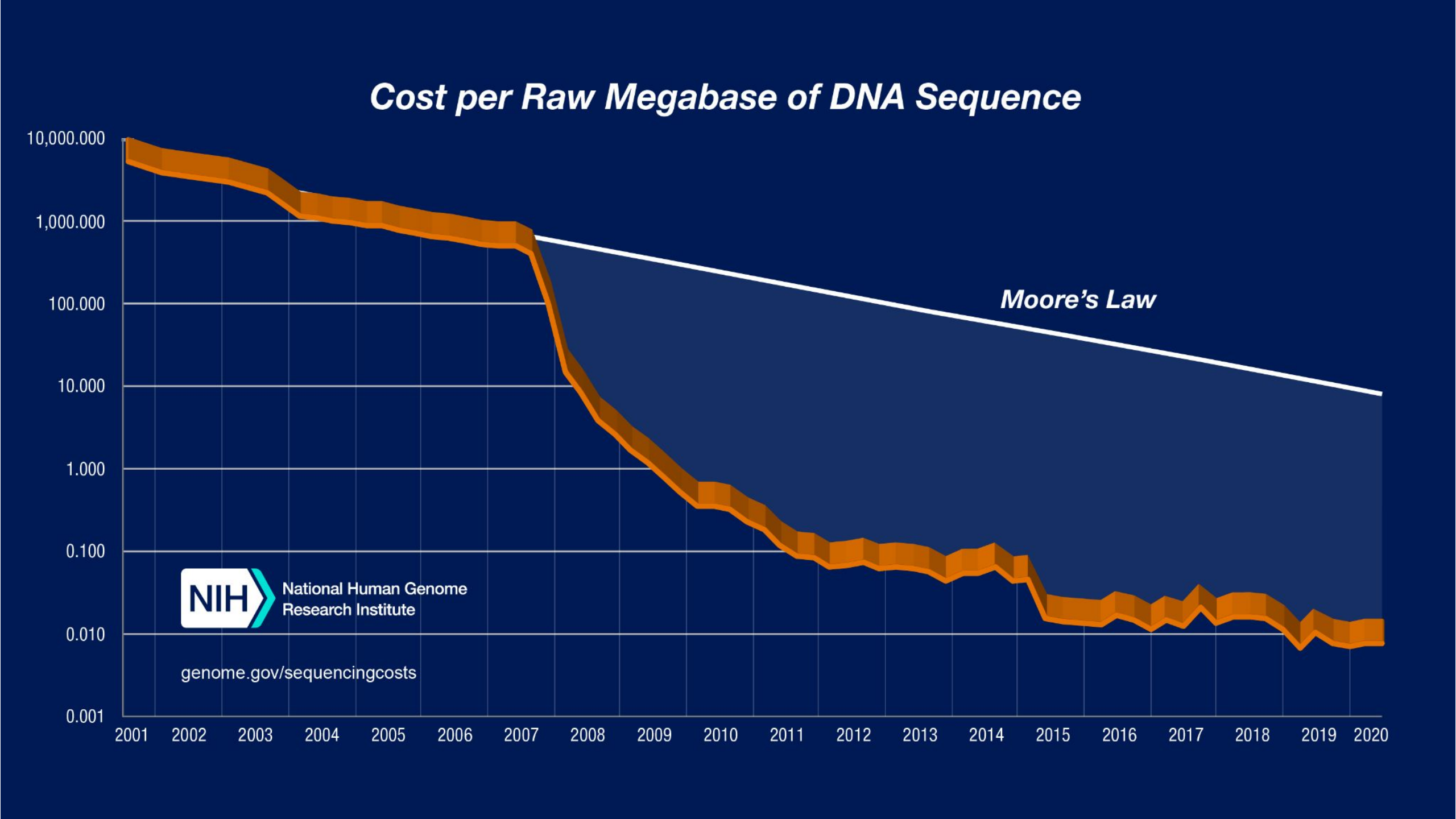
Expression profiling with sequencing vs microarrays

What are the advantages and disadvantages?

What applications are best for each platform?



https://www.biocompare.com/Editorial-Articles/518117-Microarrays-Still-Have-Their-Uses-and-Fans/


Cost per Raw Megabase of DNA Sequence
10,000.000
1,000.000
100.000
10.000
1.000
0.100
0.010
0.001
Moore's Law
NIH
National Human Genome
Research Institute
genome.gov/sequencingcosts
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020


Global Next-Generation Sequencing (NGS) Market Size, 2015-2026 (USD Million)
31,411.30
6,335.20
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
www.fortunebusinessinsights.com

Barriers

Over 99% of physicians, 20.5% of whom worked in an academic setting, reported using NGS in the past 12 months, and 73.0% used NGS always or most of the time. Despite this high utilization, 80.1% of physicians experienced at least one barrier to testing.

Survey study of barriers to evidence-based decision-making in oncology care using next-generation sequencing.

Check for updates

Elizabeth A. Szamreta, Allysen Kaminski, Ruchit Shah, Ning Ning, Jyoti Aggarwal, Arif Hussain, ...

Percentage of Physicians Reporting Each Barrier to Optimizing Clinical Impact and Utility of NGS in Routine Clinical Practice by Physician Specialty, %.

Barrier	Total (N= 201)	Oncology/Hematology (N=100)	Pathology (N=51)	Surgery (N=50)
Reimbursement Challenges	87.5	85.0	90.2	90.0
Knowledge/Awareness	81.0	75.0	82.4	92.0
Evidence of clinical utility	80.1	79.0	78.4	84.0
Availability of Supportive Resources	79.6	73.0	82.4	90.0
Logistical Barriers	77.2	73.0	72.6	90.0

© 2021 by American Society of Clinical Oncology

What barriers to utilizing next generation sequencing and analysis have you experienced or seen others experience?

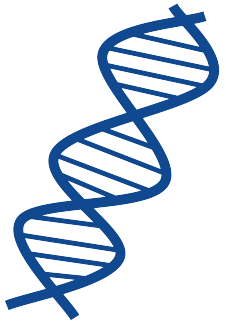
Photo by Art Wave on Unsplash

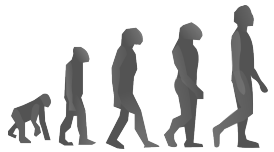
New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

NCGR
National Center for Genome Resources

NGS applications

- Genetic and genomic medicine
- Microbiology and the environment (microbiomes and microbiota)
- Pharmacogenetics (genetic variation and drug responses)
- Forensics
- Agriculture (improvement, protection)
- Physical performance
- Evolution
- Extinction
- Ecology
- Geneology and phylogeny





New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

NCGR
National Center for Genome Resources

Sequencing related to expression

mRNA
Full-length mRNA
Isoforms
Nascent transcription
miRNA or small RNAs
Single cell
Spatial transcriptomics
Methylation and other DNA modifications
Histones and histone modifications
Open chromatin
Transcription Factor Binding Sites
Stranded RNA
...

RNA is usually converted first to cDNA then sequenced

Next Generation (NGS) Sequencing Technologies

Short Read

Sanger

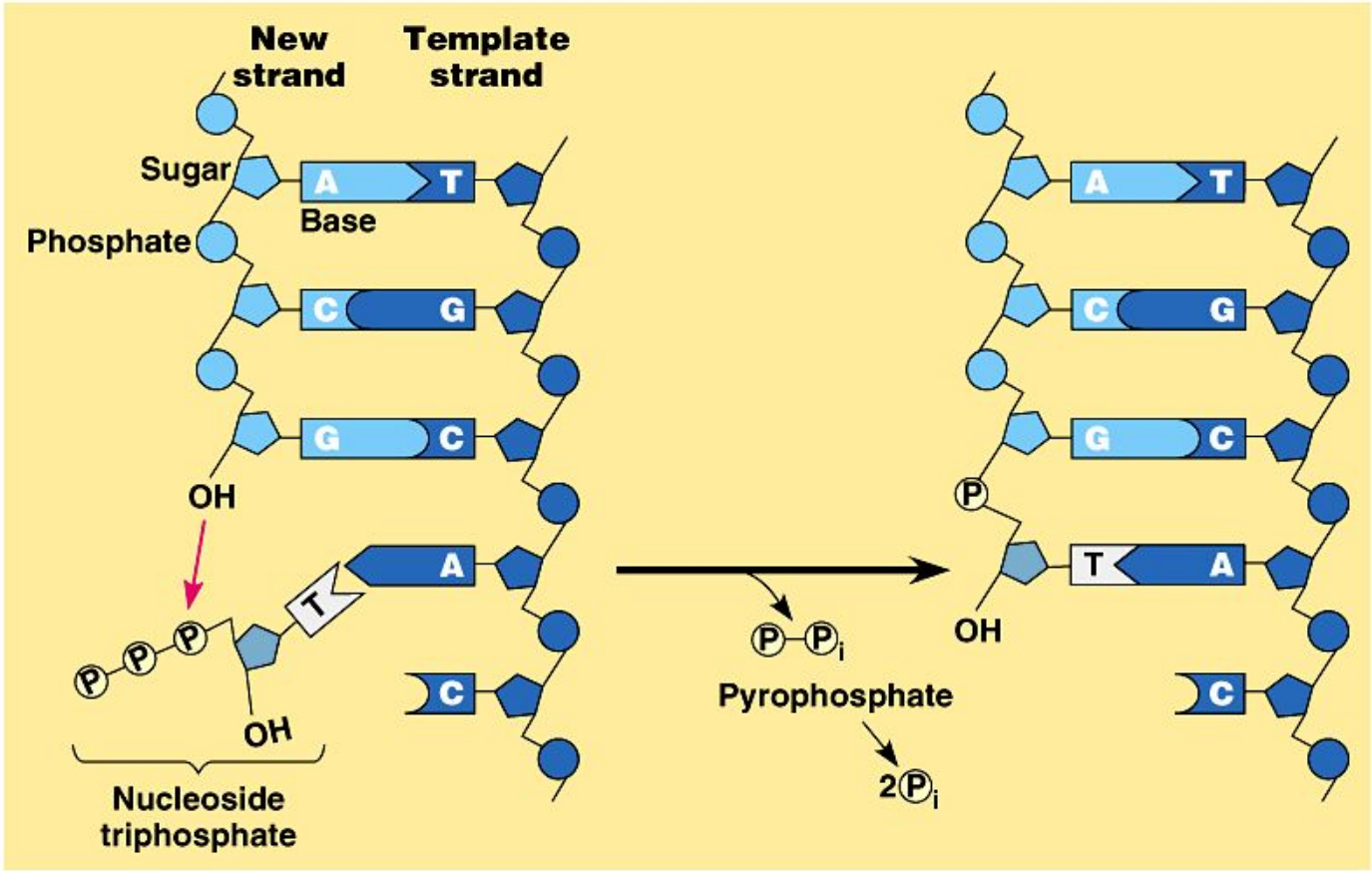
Illumina

Long Read

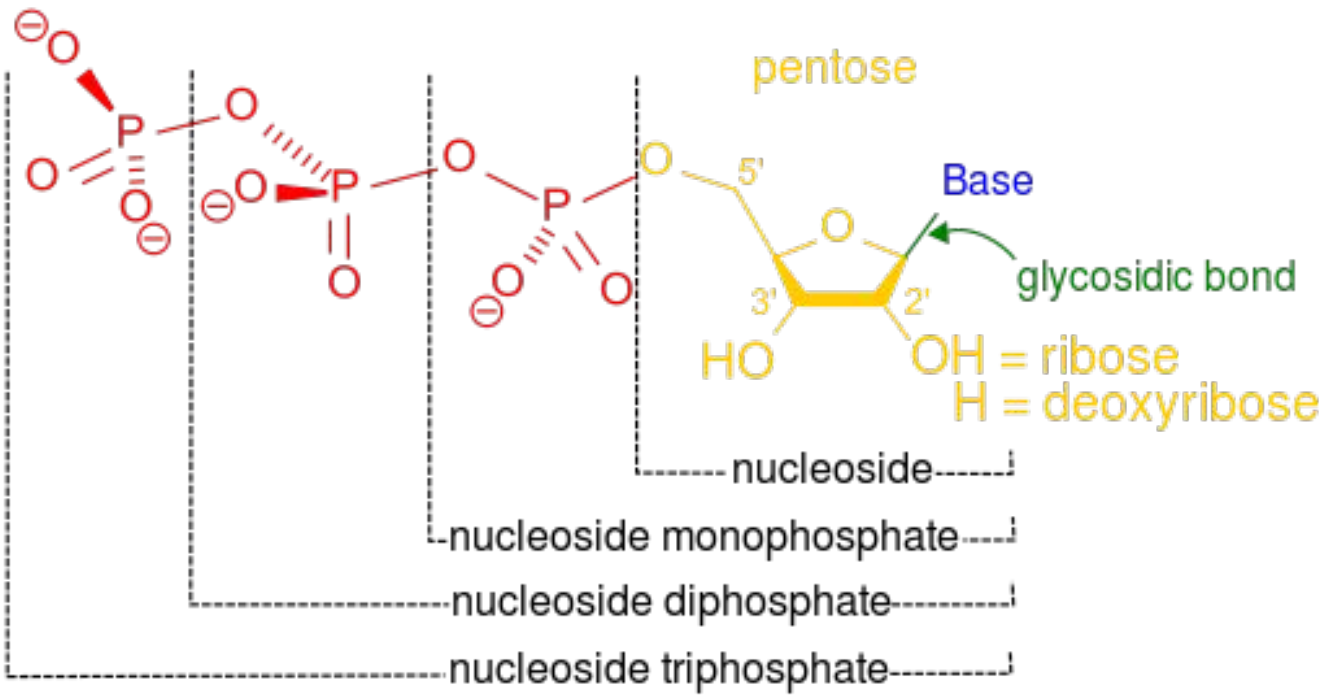
PacBio

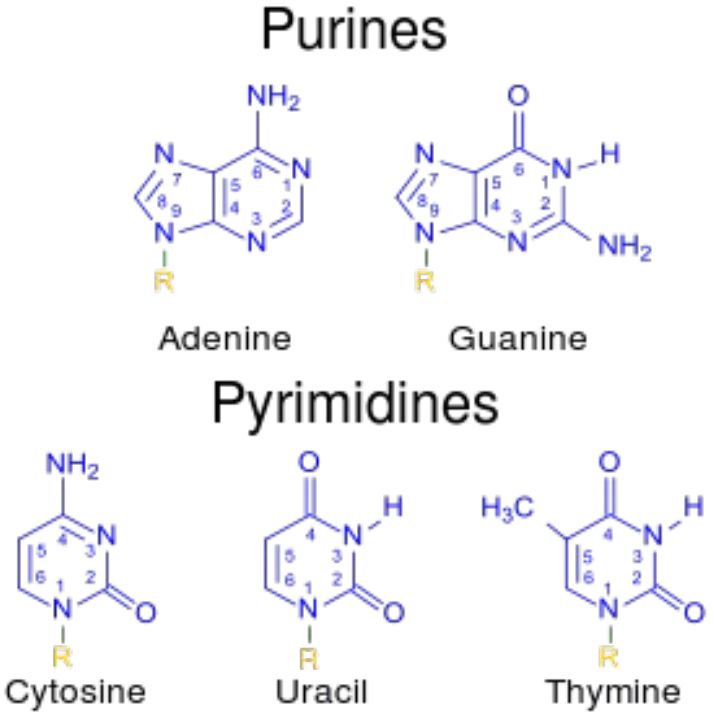
Oxford Nanopore Technologies

Most sequencing technologies are based on DNA replication



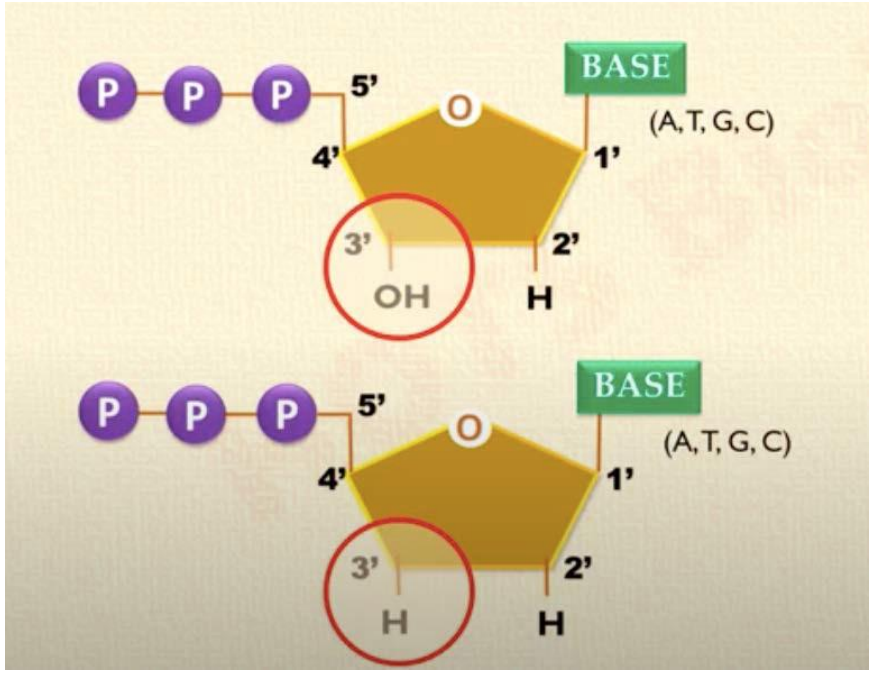
DNA chemistry



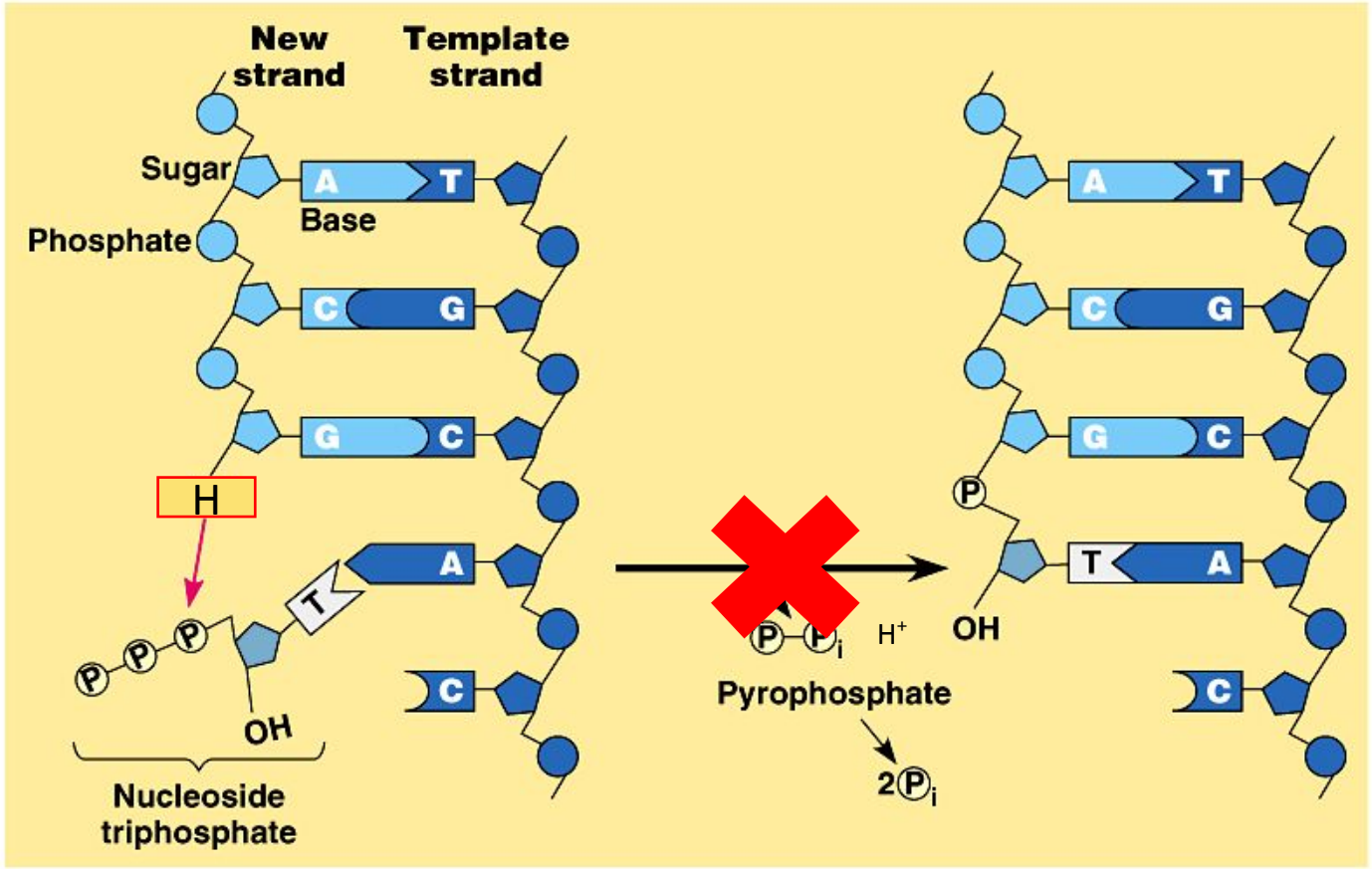


http://en.wikipedia.org/wiki/File:Nucleotides_1.svg

Sanger sequencing uses dideoxy nucleotides

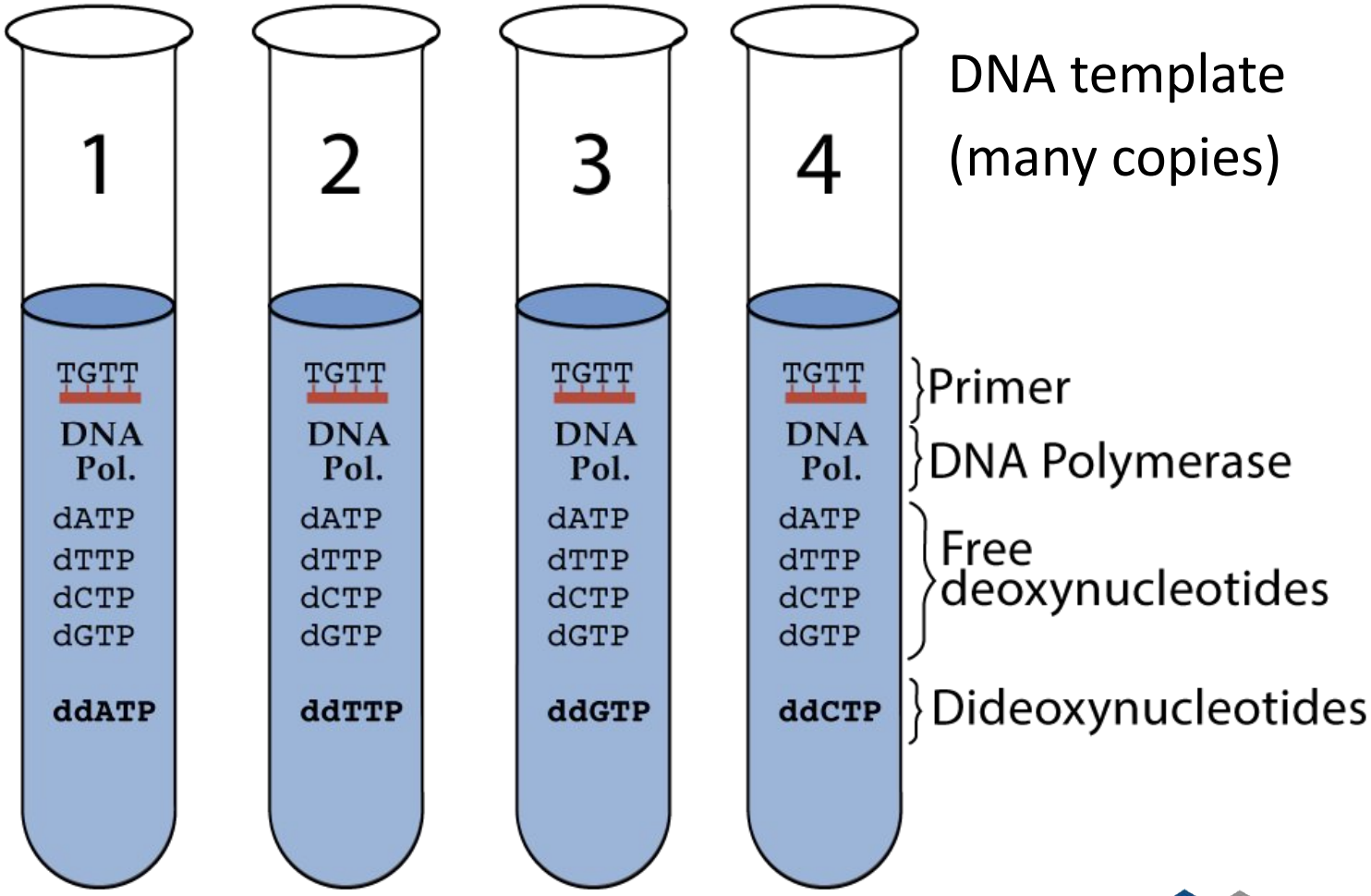


ddNTPs prevent polymerization



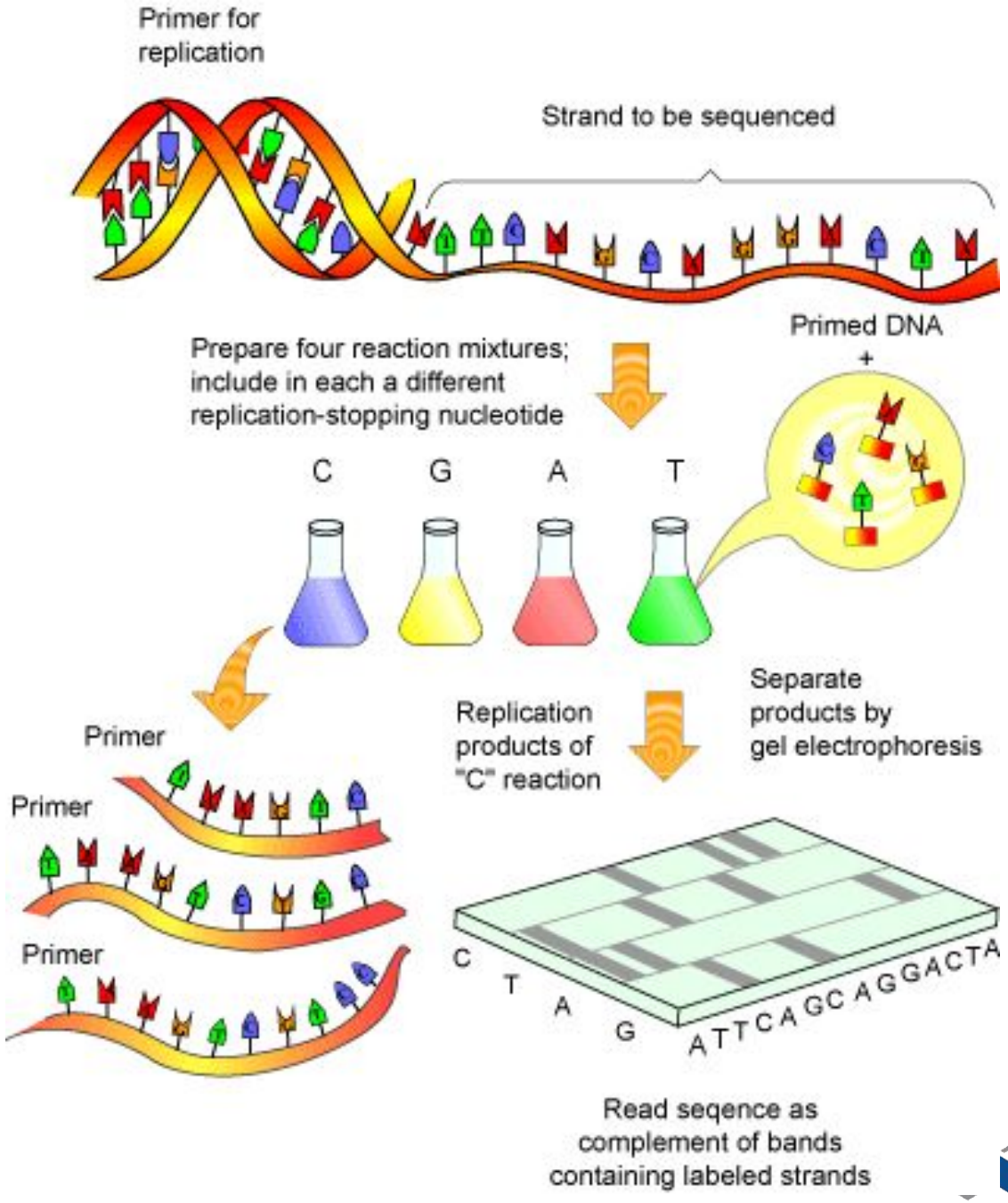
The basis of Sanger sequencing

Different type of ddNTP's in each reaction (small amounts)



New Mexico INBRE
IDeA Networks of Biomedical Research Excellence
NCGR
National Center for Genome Resources

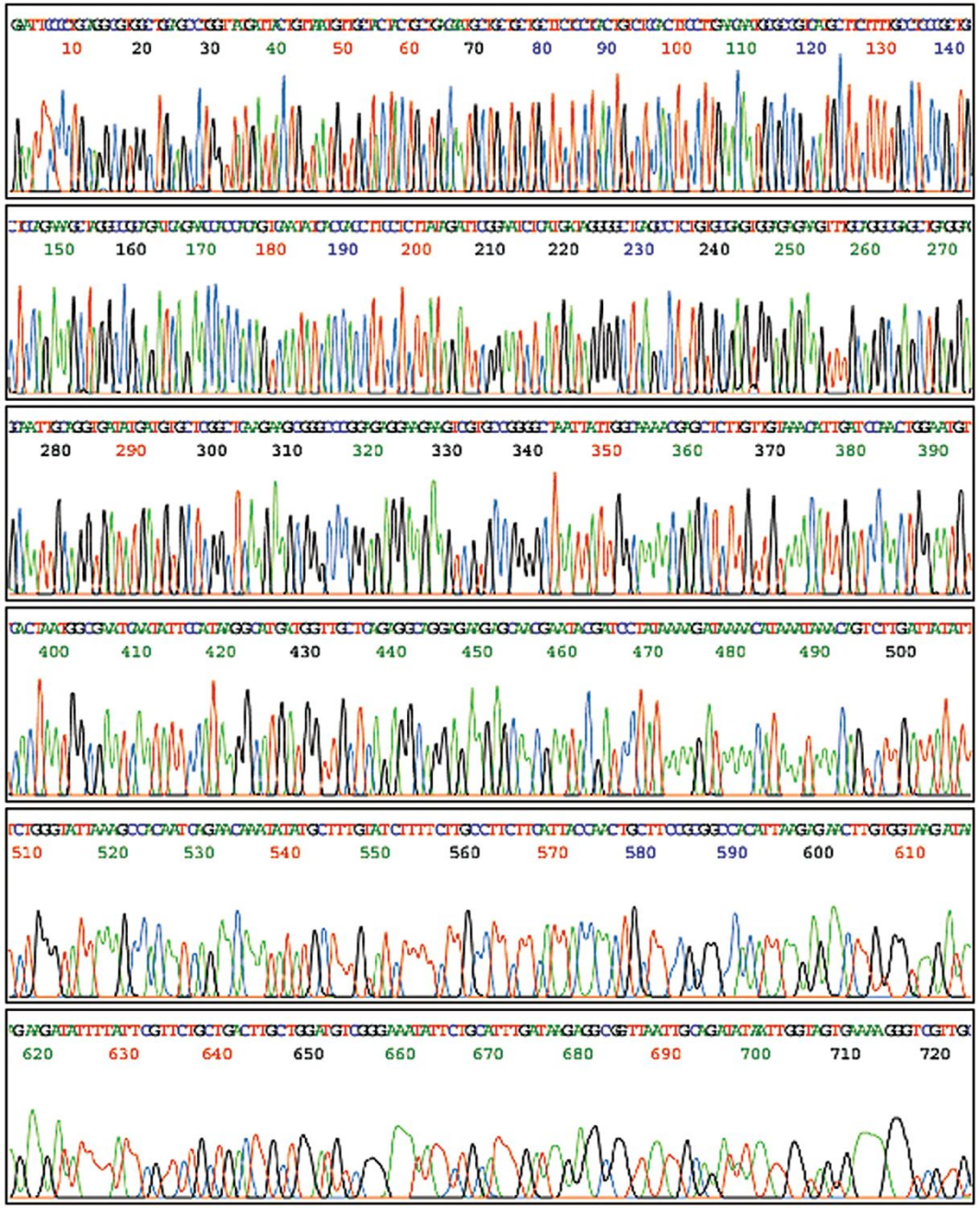
Sanger sequencing



New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

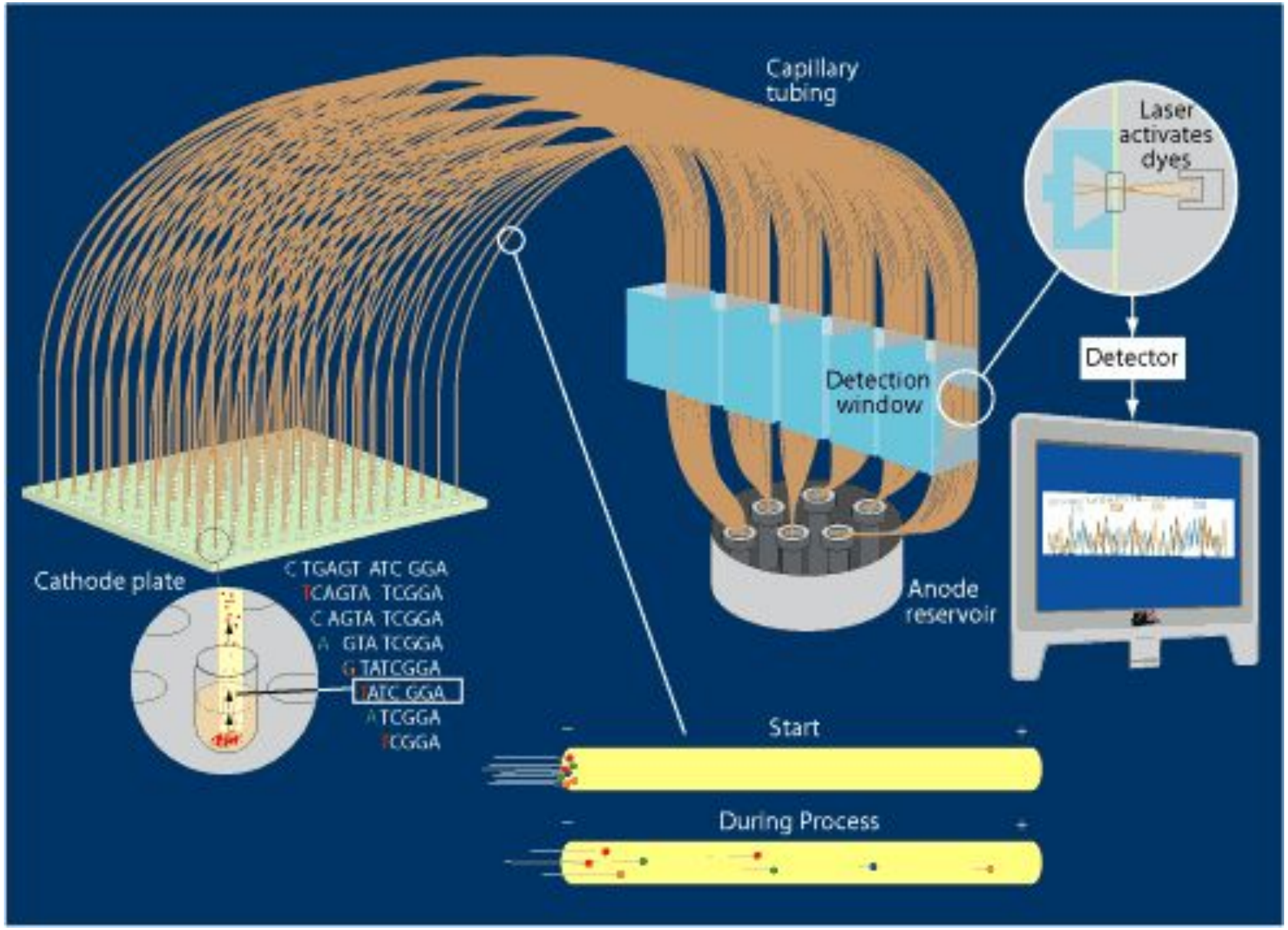
NCGR
National Center for Genome Resources

Sanger output



New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

NCGR
National Center for Genome Resources


Capillary tubing
Laser activates dyes
Detector
Detection window
Anode reservoir
Cathode plate
C TGAGT ATC GGA
TCAGTA TCGGA
C AGTA TCGGA
A GTA TCGGA
G TATCGGA
TATC GGA
A TCGGA
TCGGA
Start
During Process

Human Genome Sequence

- 13 years (1990-2003)
- $ 2.7 B

New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

NCGR
National Center for Genome Resources

Illumina Sequencing Machines

iSeq 100

MiniSeq

MiSeq Series ⊕

NextSeq 550 Series ⊕

NextSeq 1000 & 2000

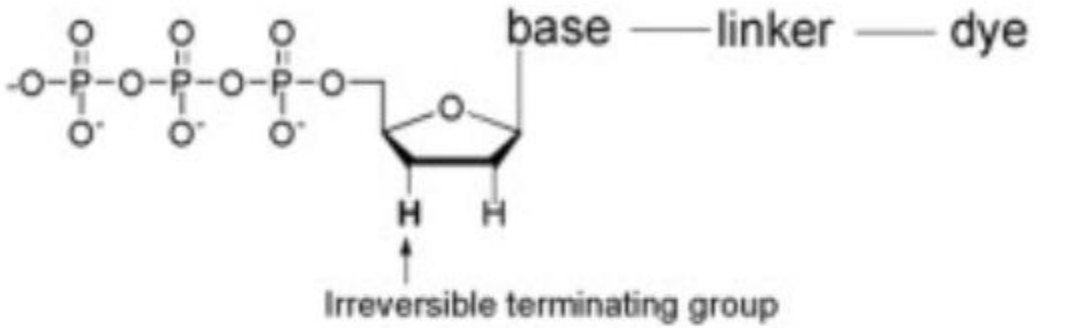
NextSeq 550 Series ⊕

NextSeq 1000 & 2000

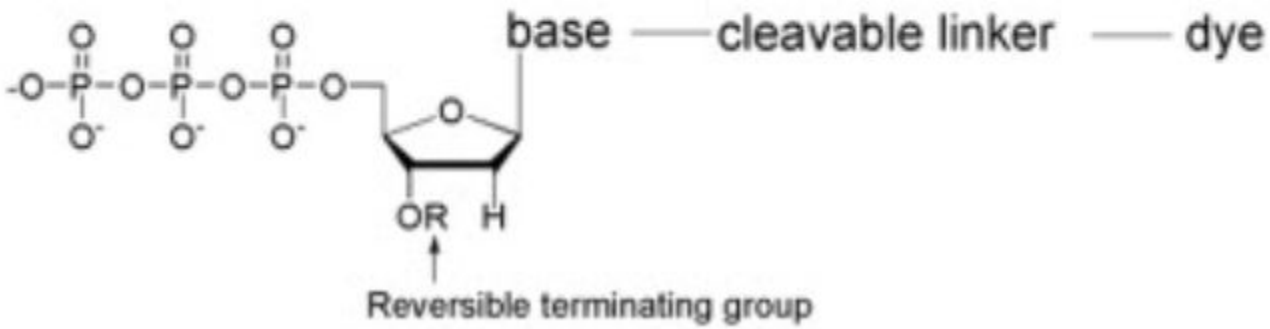
NovaSeq 6000

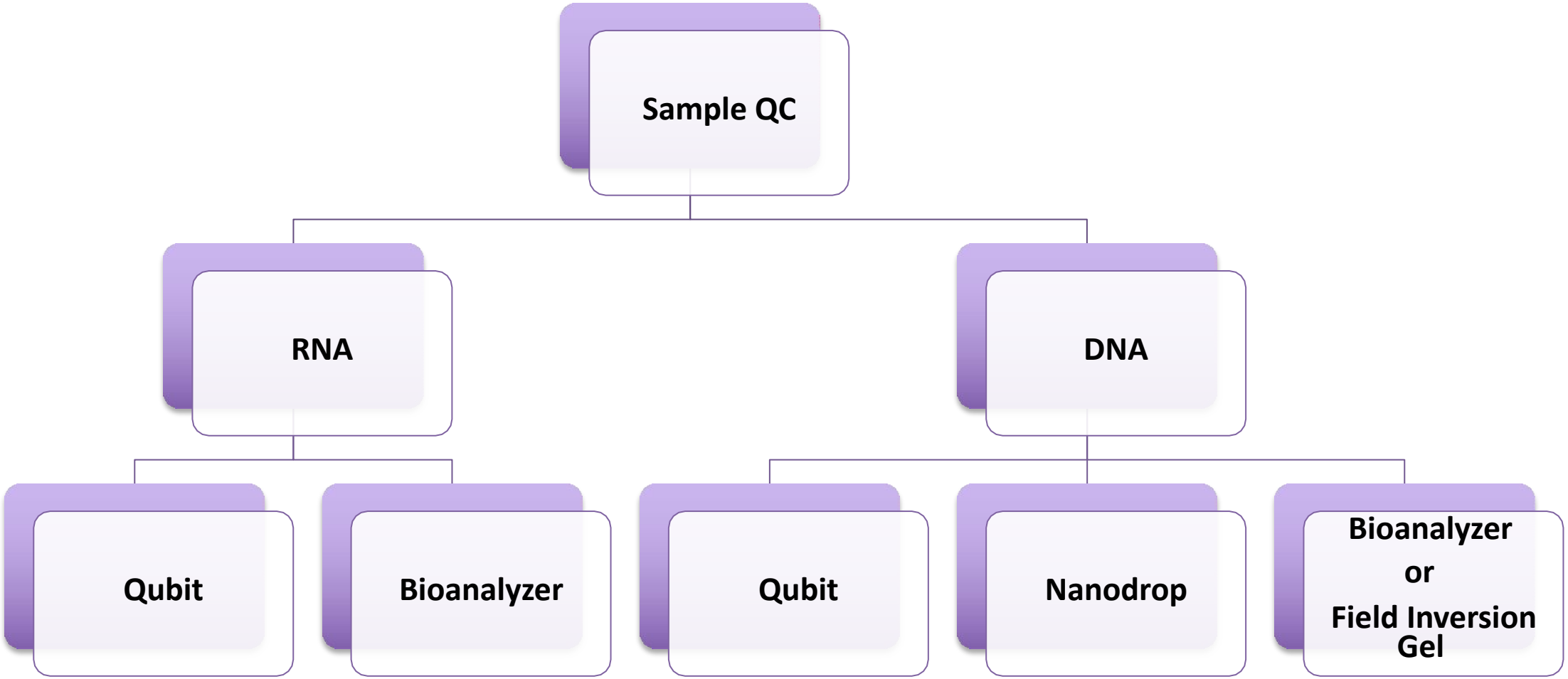
Reversible terminator

A Sanger cleavable fluorescent dideoxynucleotide



B 3'-*O*-blocked reversible terminator

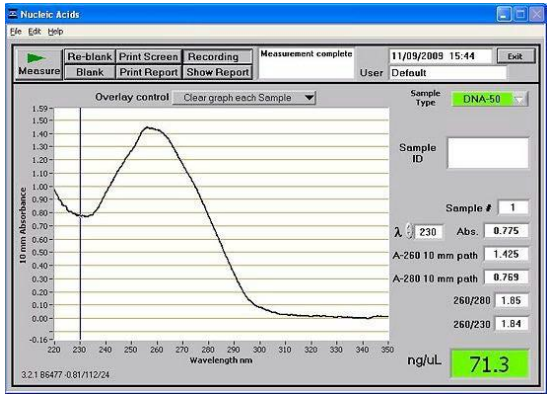



Sample QC
RNA
DNA
Qubit
Bioanalyzer
Qubit
Nanodrop
Bioanalyzer
or
Field Inversion
Gel
Determines amount of sample
Integrity of sample (RIN)
Determines amount of sample
Presence of contaminants (260/280 ratio)
Size estimation

Qubit

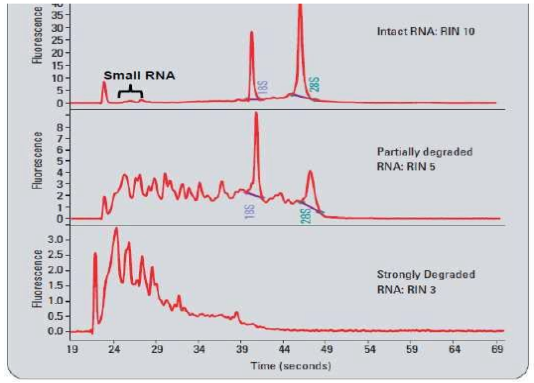
Fluorometer: Uses intercalating dye and is more accurate than UV absorbance based measurements such as with nanodrop

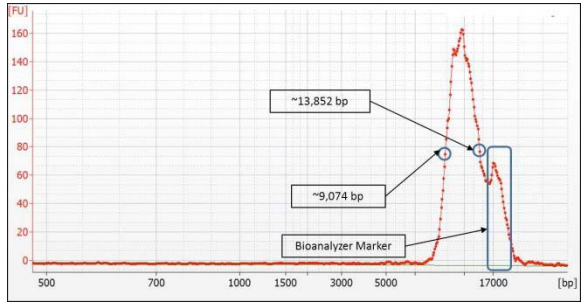
Nanodrop



260/230 and 260/280 ratios.

Bioanalyzer

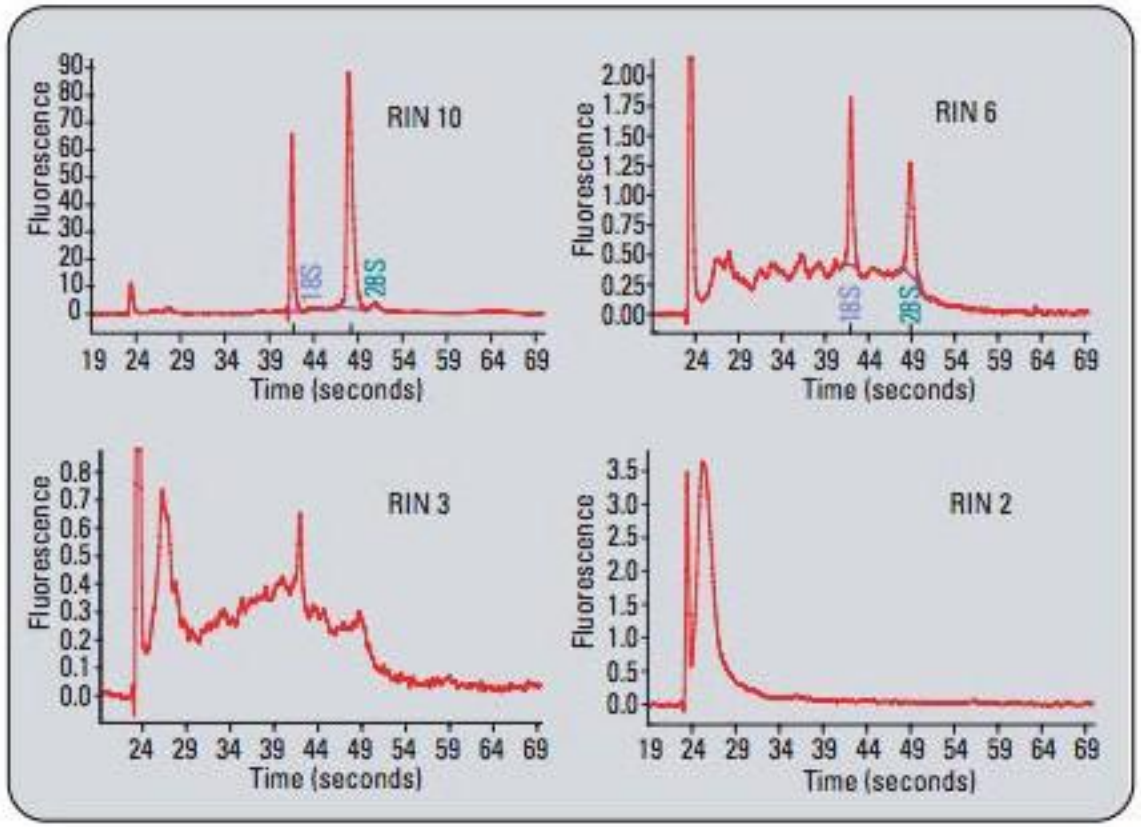




New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

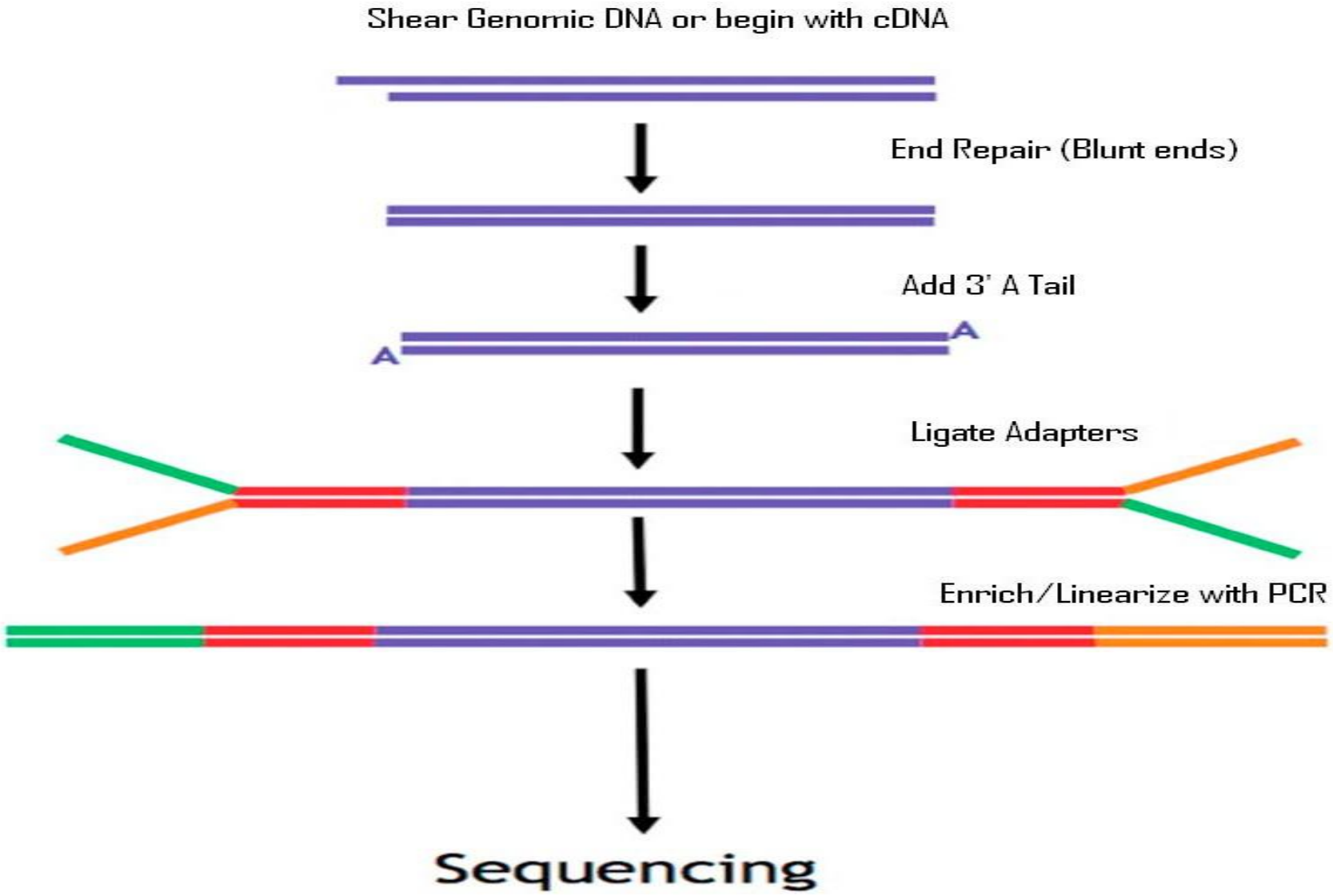
NCGR
National Center for Genome Resources

Agilent Bioanalyzer examples

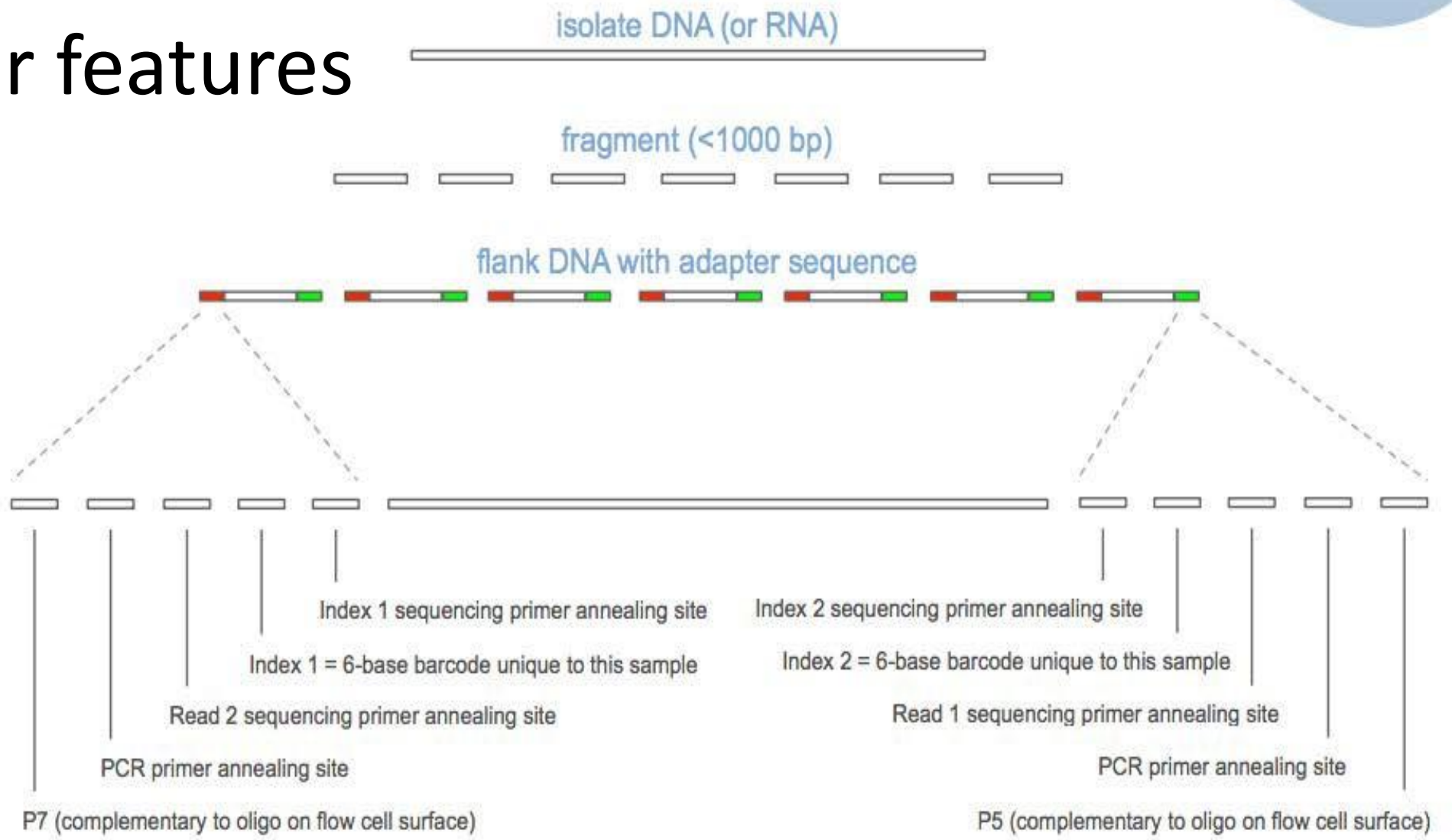


http://www.agilent.com/cs/library/applications/5989-1165EN.pdf

Illumina librarv architecture

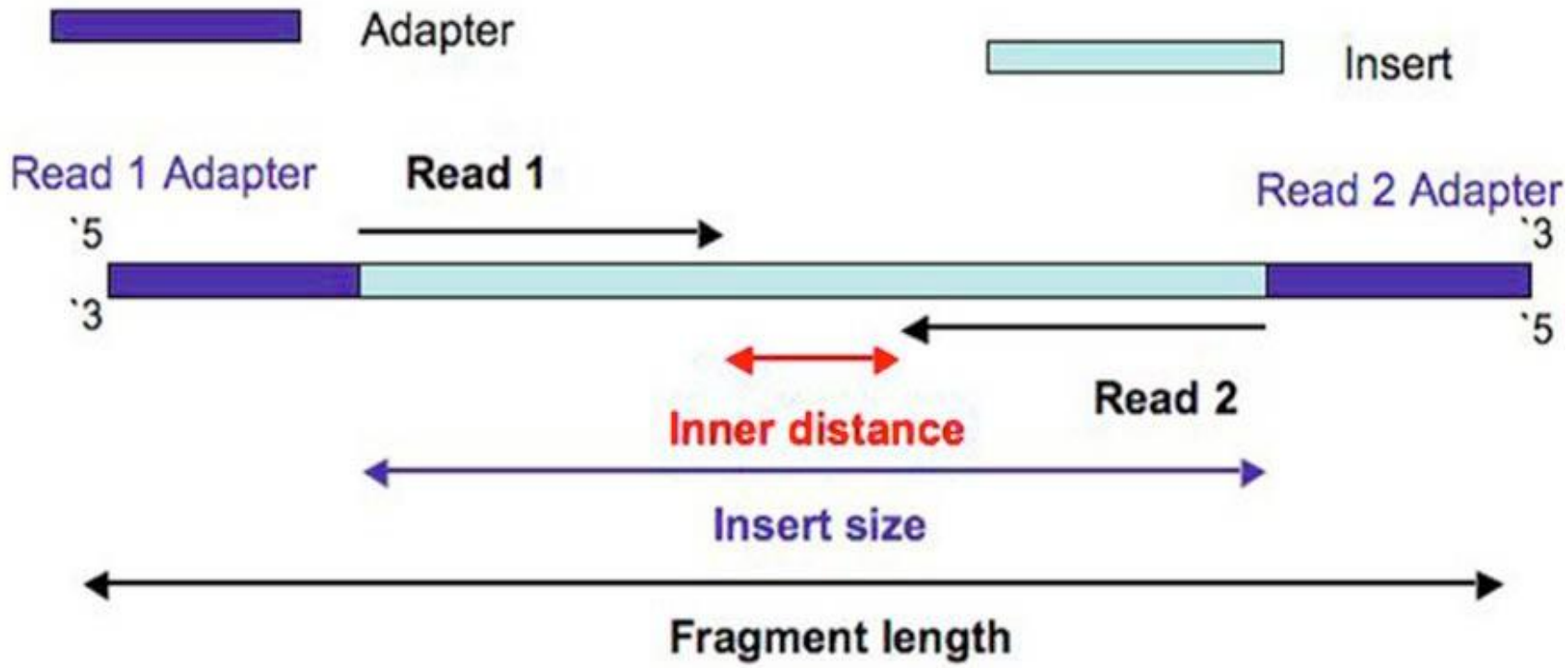


Adapter features



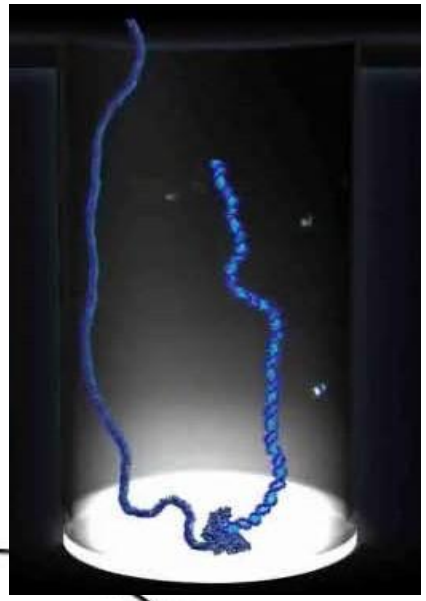
https://www.youtube.com/watch?v=fCd6B5HRaZ8

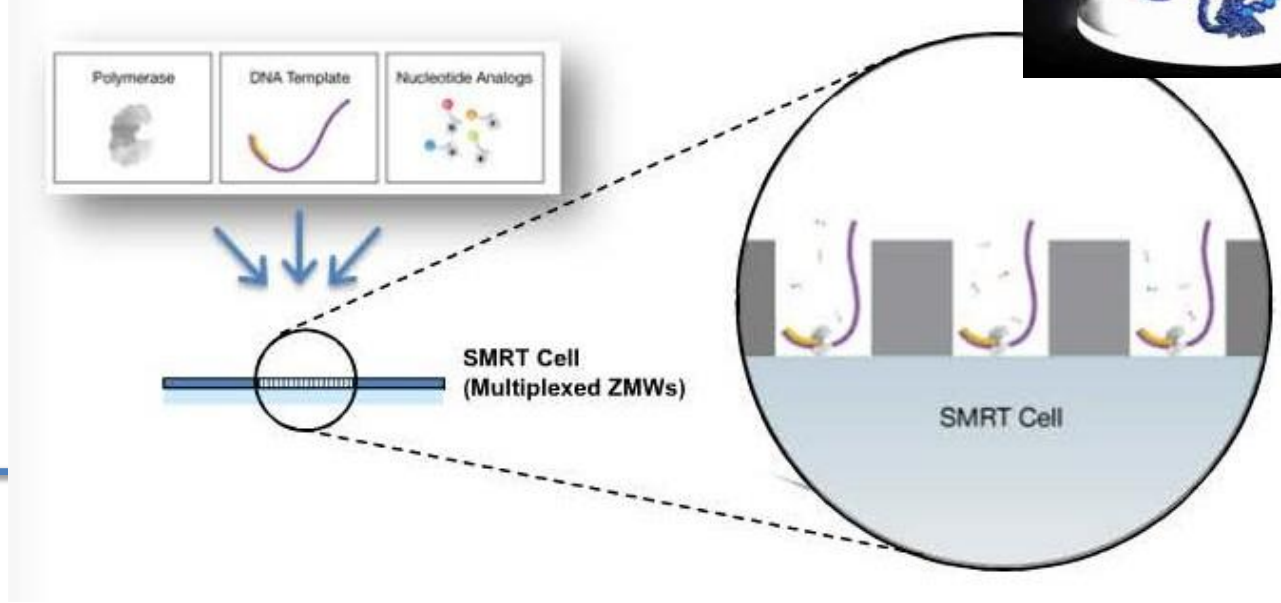
Commonly used terms



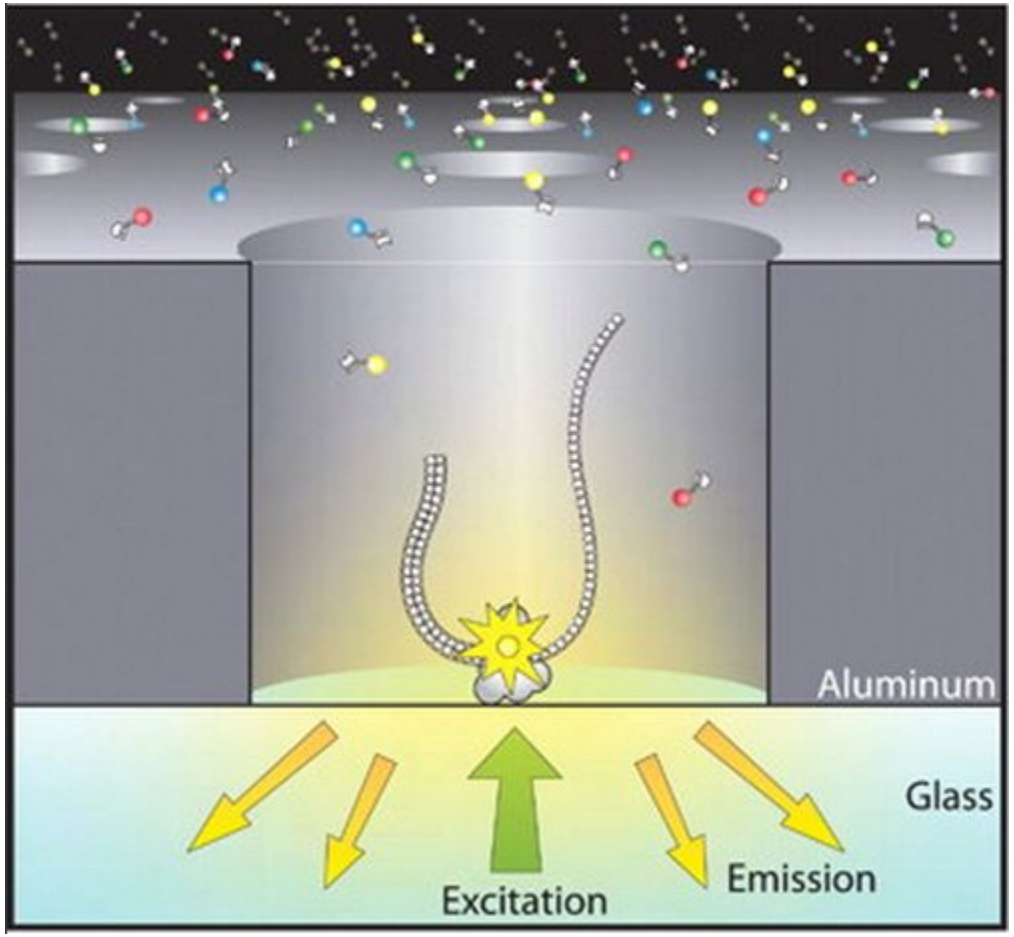
Pacific Biosciences

Single Molecule Real Time (SMRT)

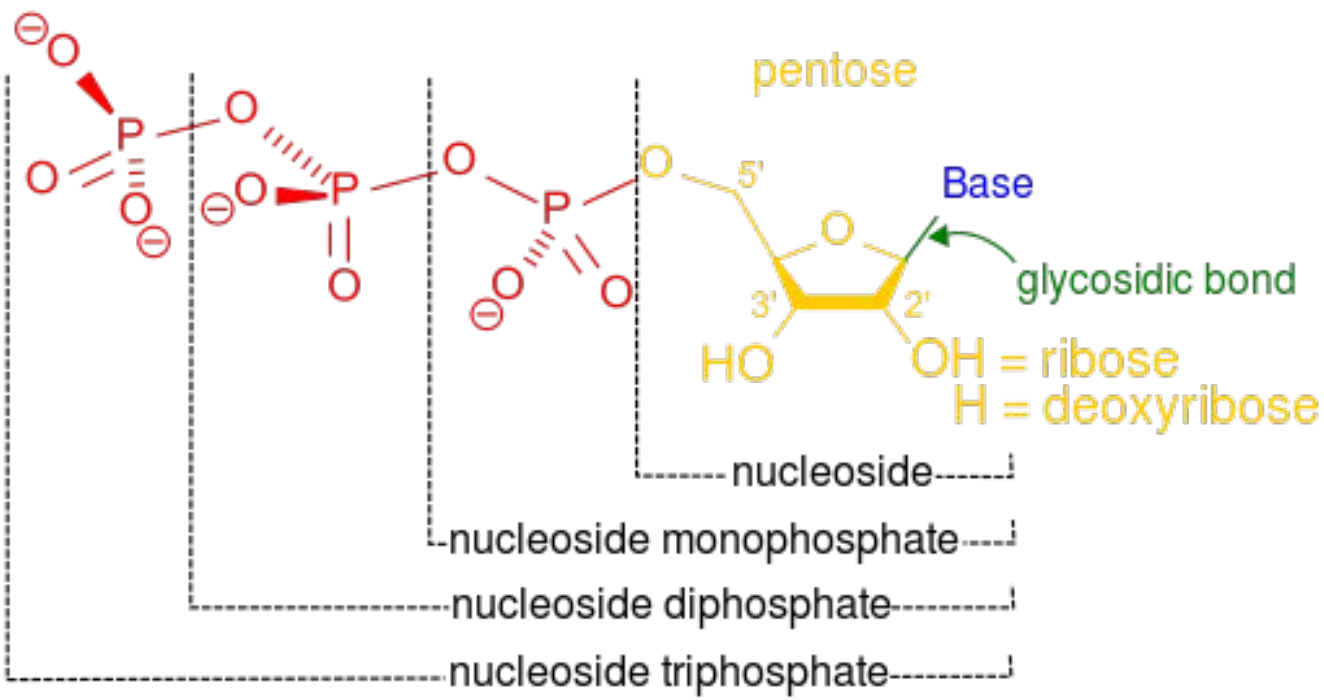


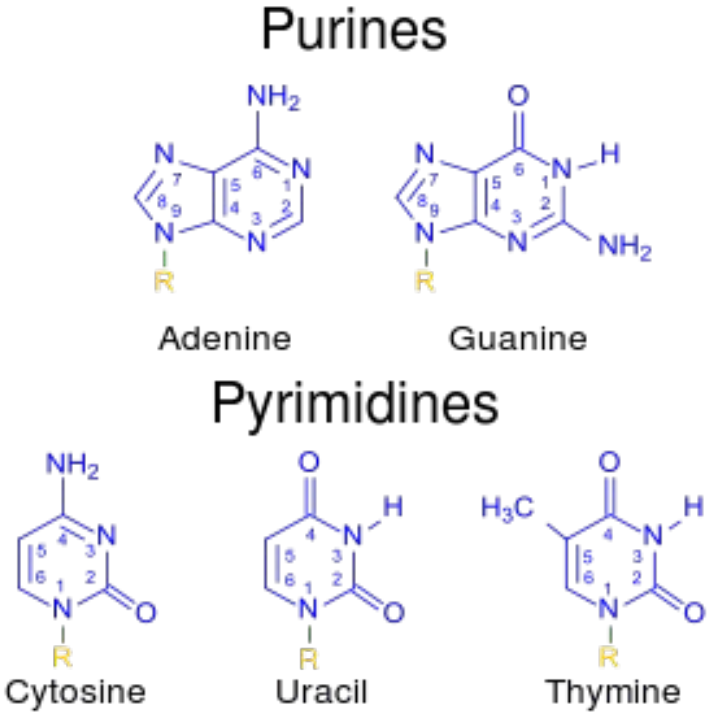


Zero Mode Waveguide

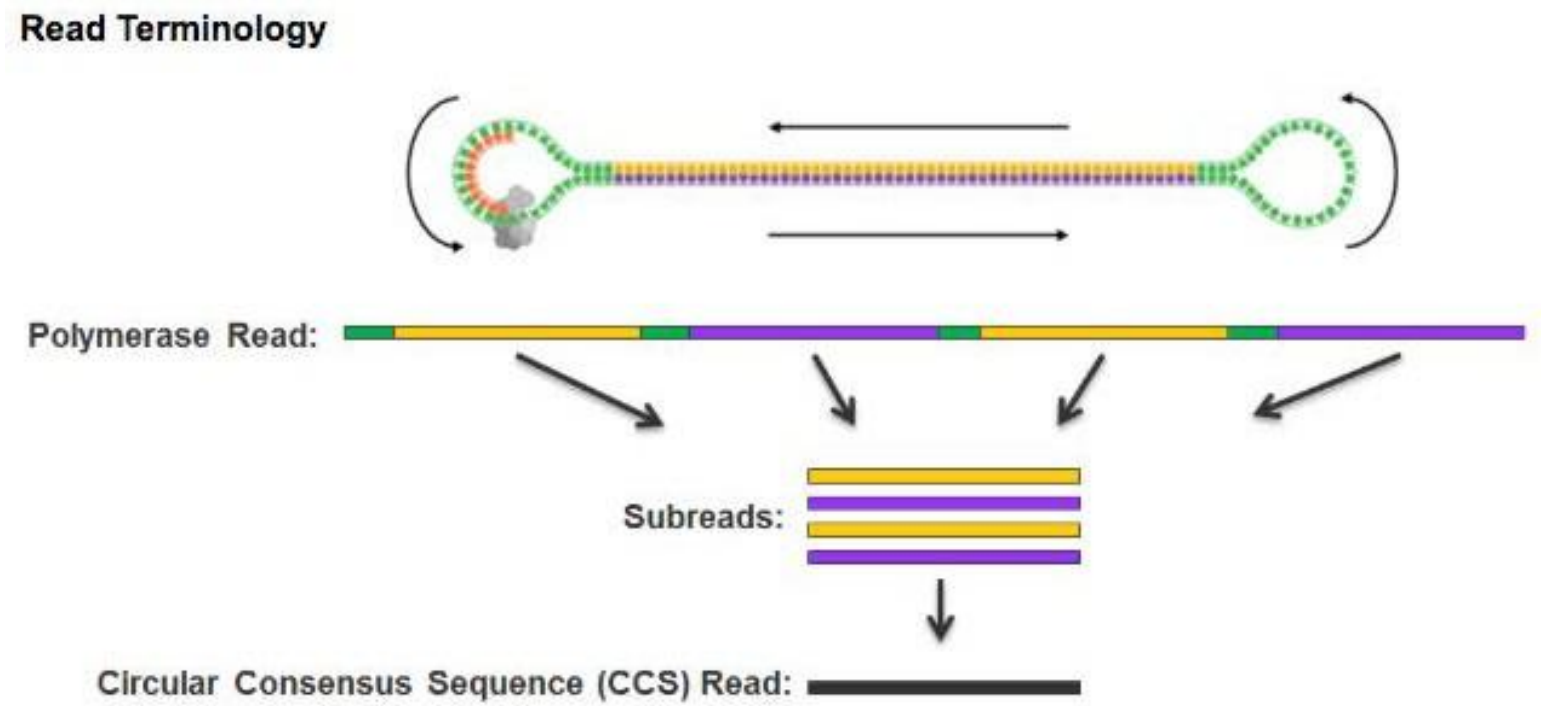


Fluorescence

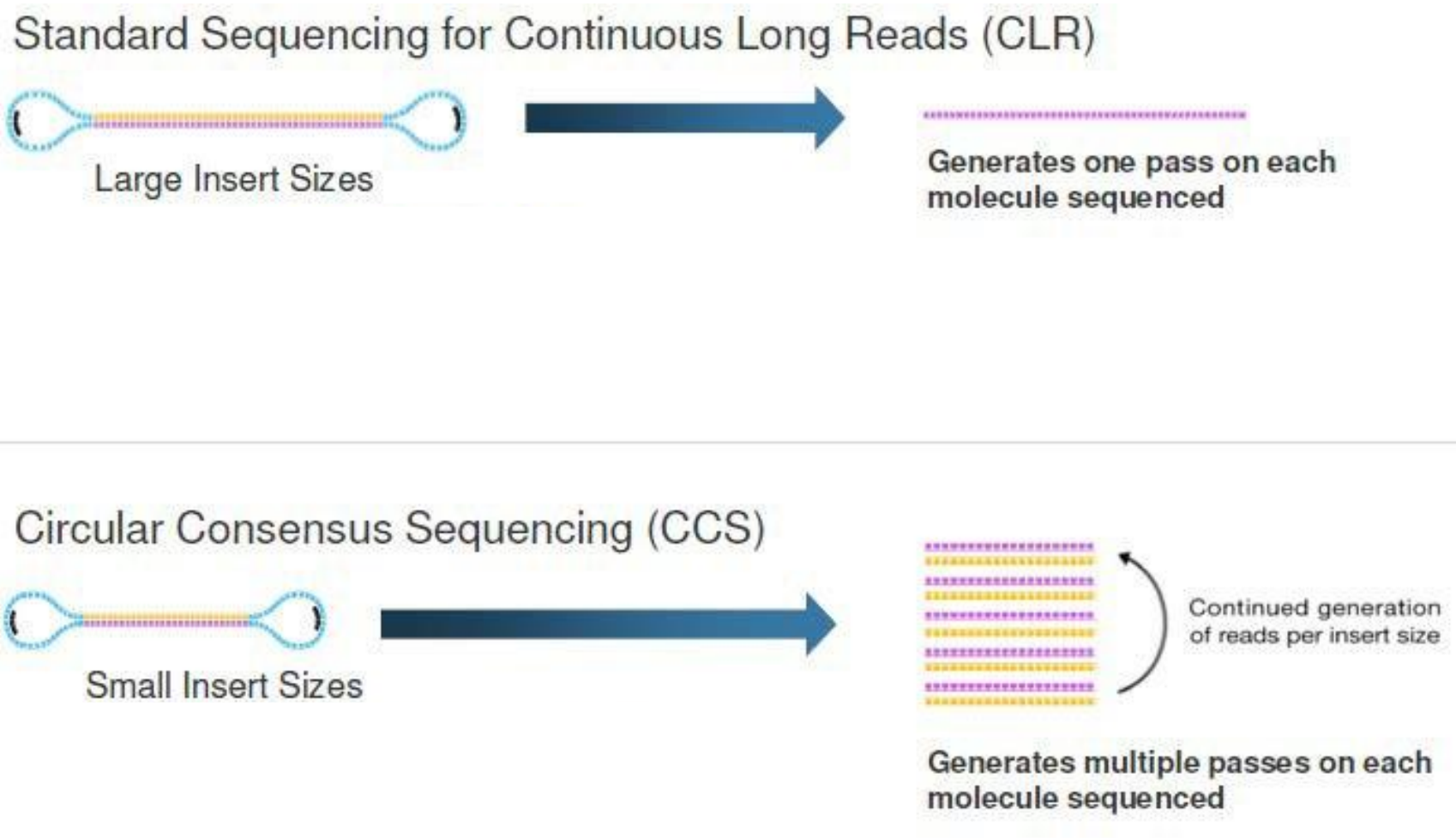




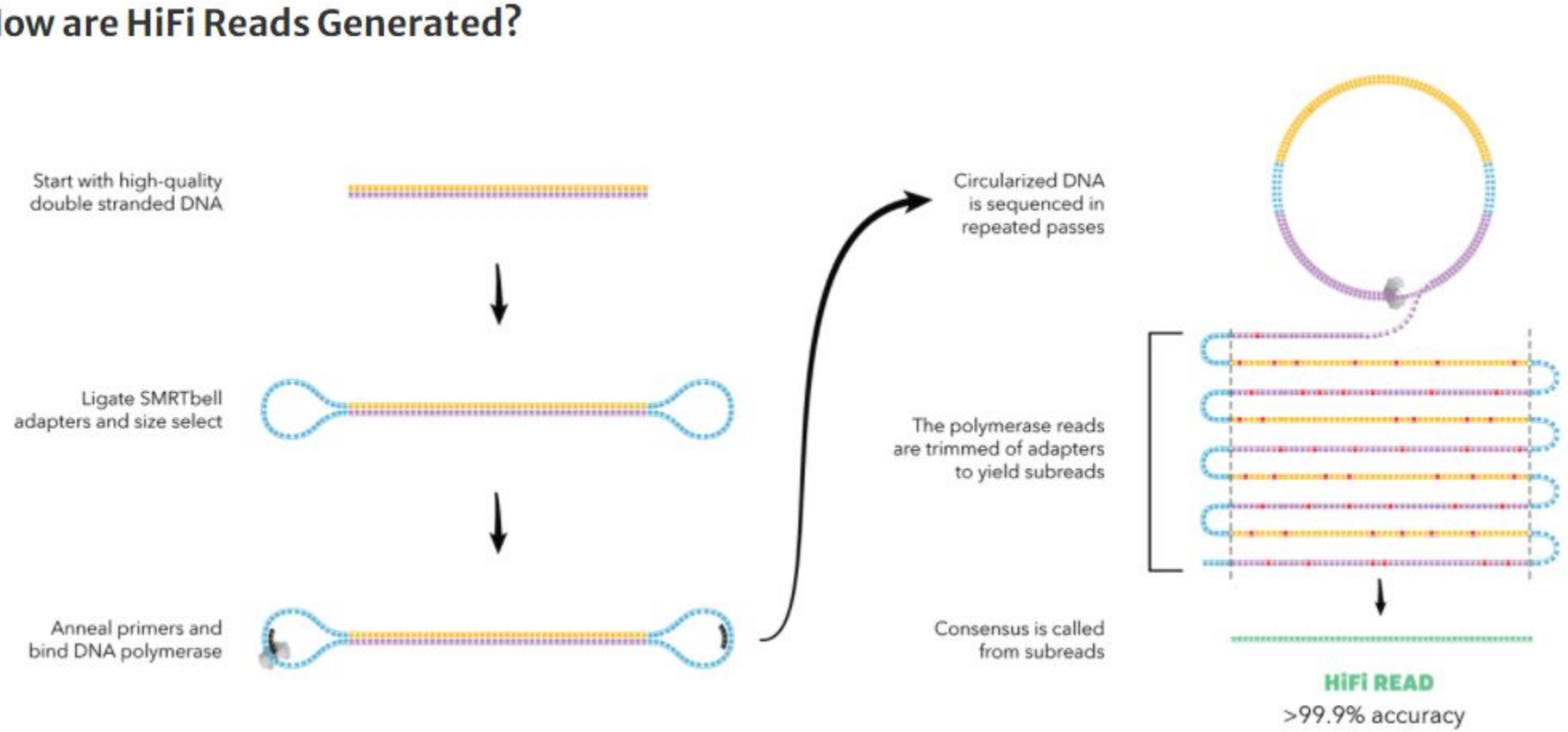
PacBio Terminology



PacBio Terminology

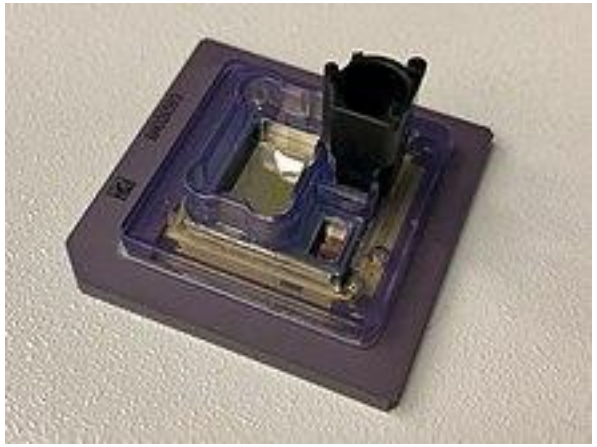


PacBio Terminology



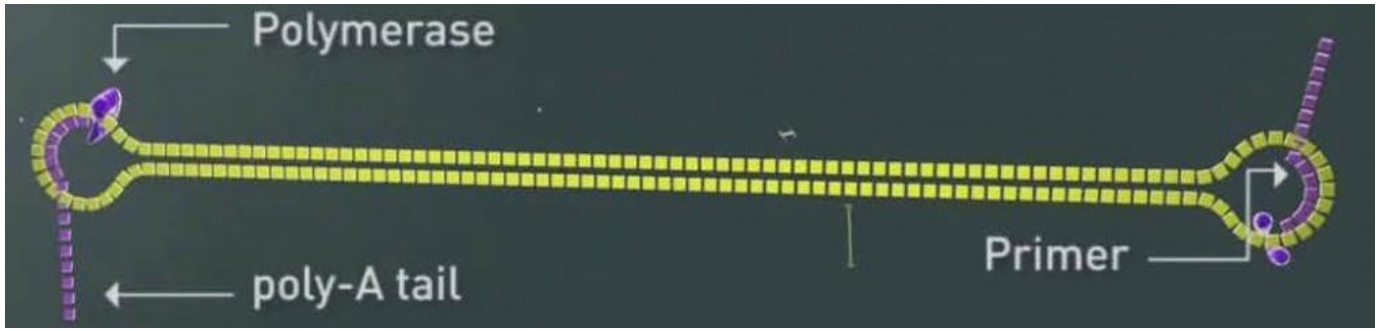
PacBio development

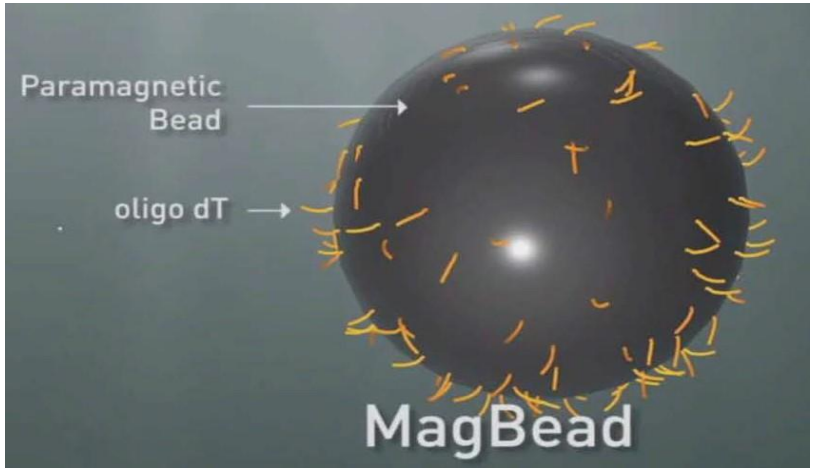
RS II



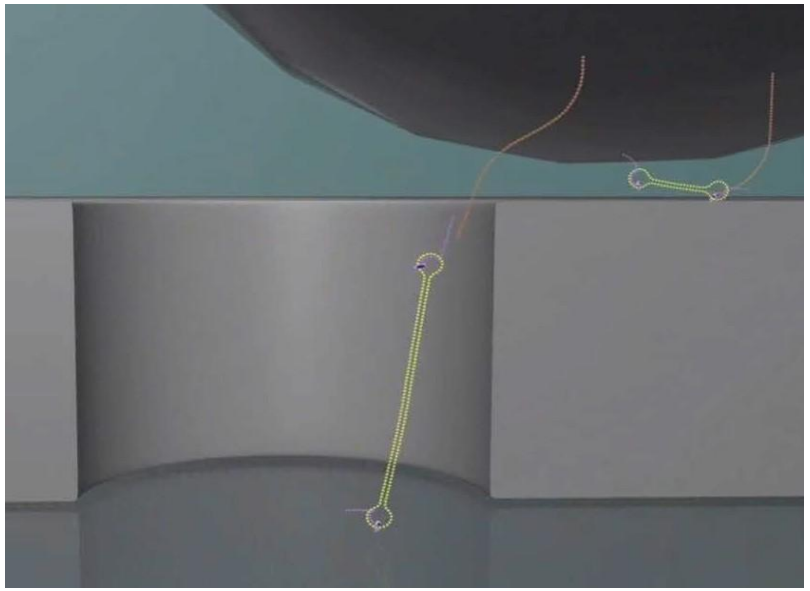
Sequel

Magnetic bead loading

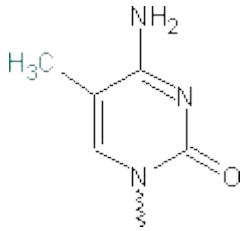




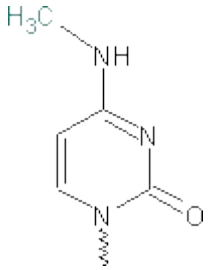
http://www.youtube.com/watch?v=1b7UeGu9xa8



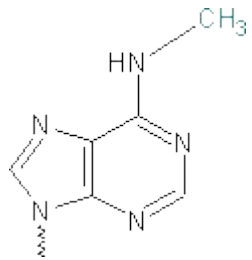
Base modification detection



5mC



4mC



6mA

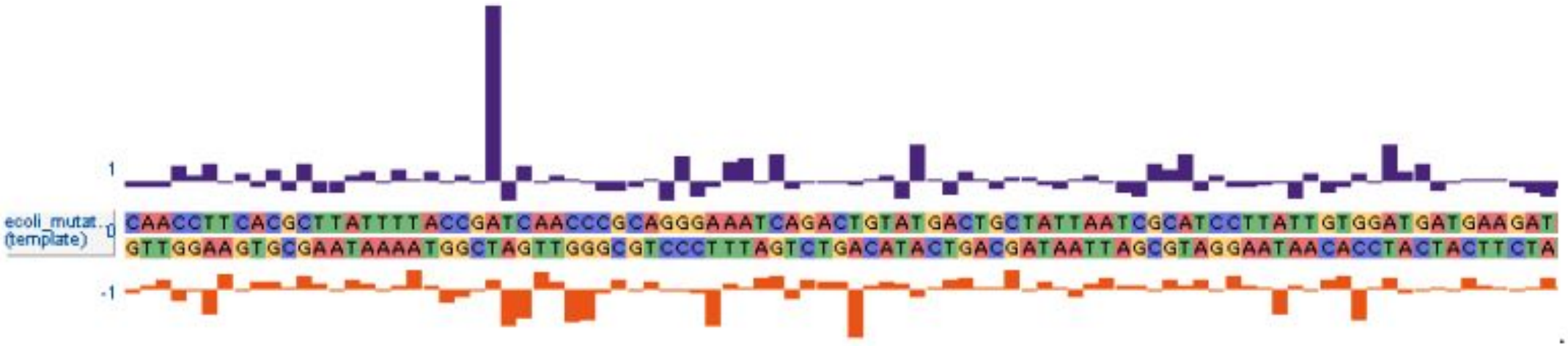
Chemical Modifications:

- 5mC
- 5hmC
- 4mC
- 5fC
- 6mA

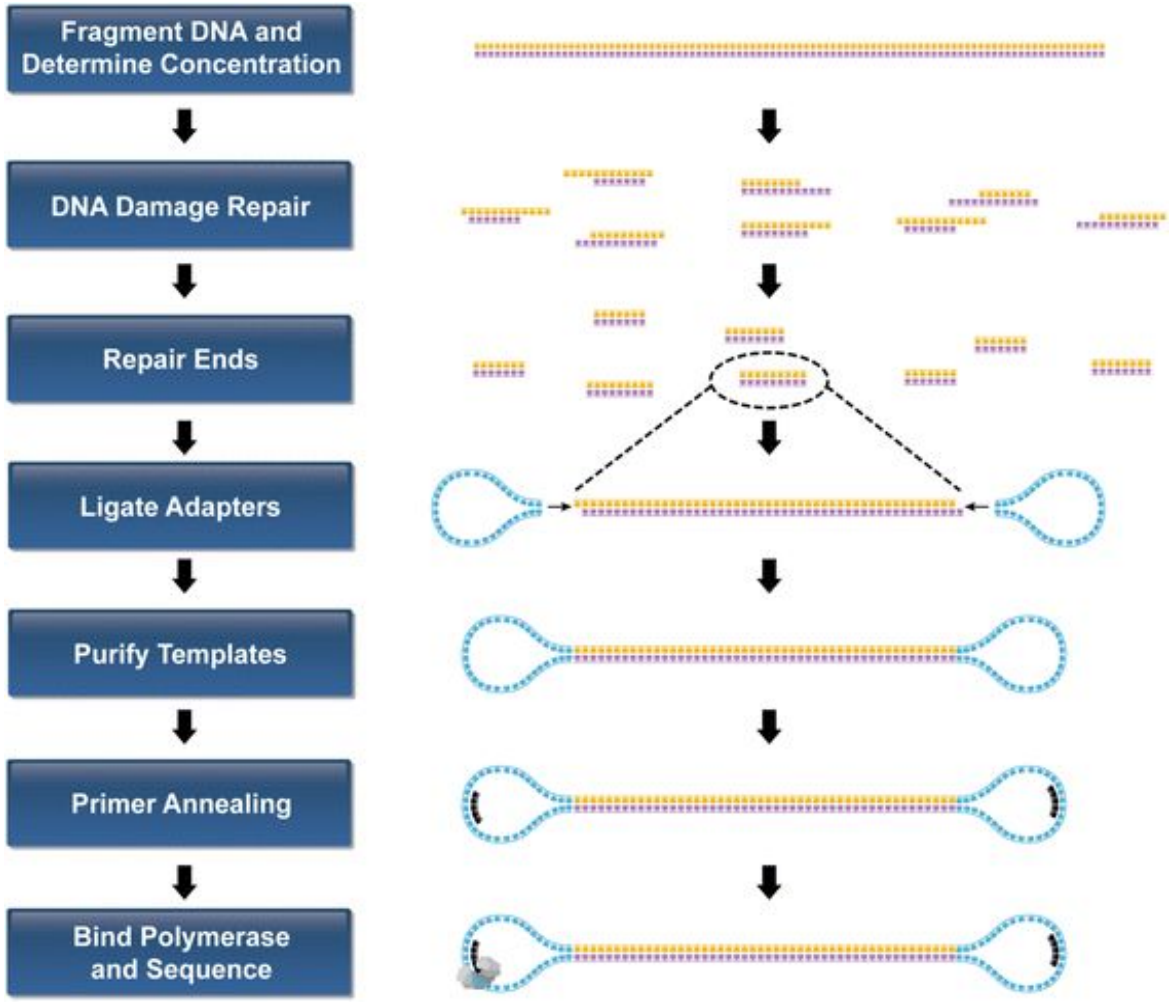
DNA Damage:

- 8oxoG
- O6mG
- 5hC
- 8oxoA
- O4mT

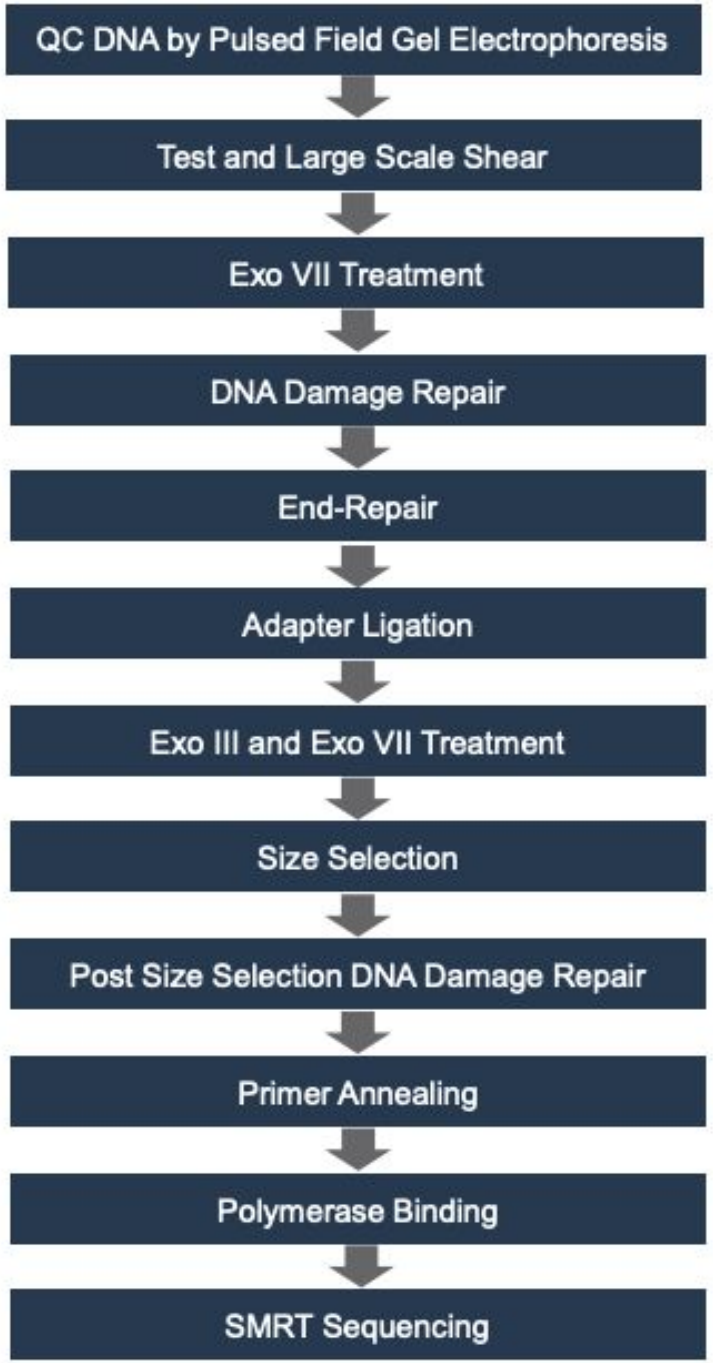
Inter pulse distance



Library construction

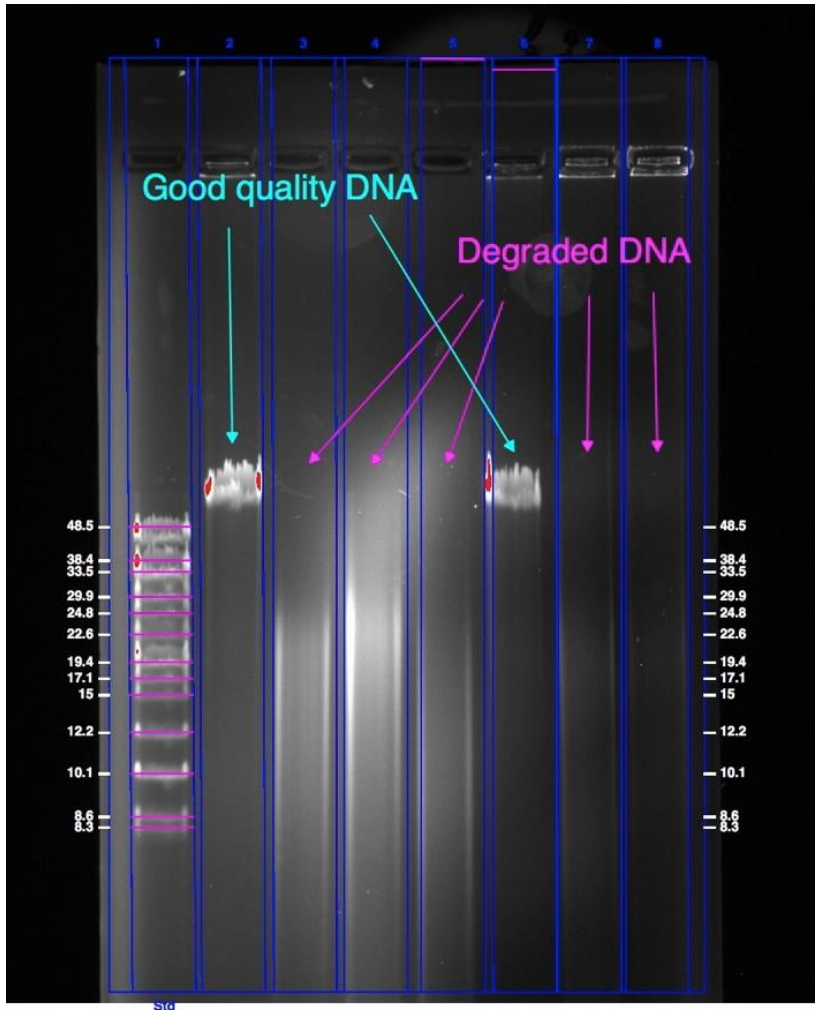


More recent example



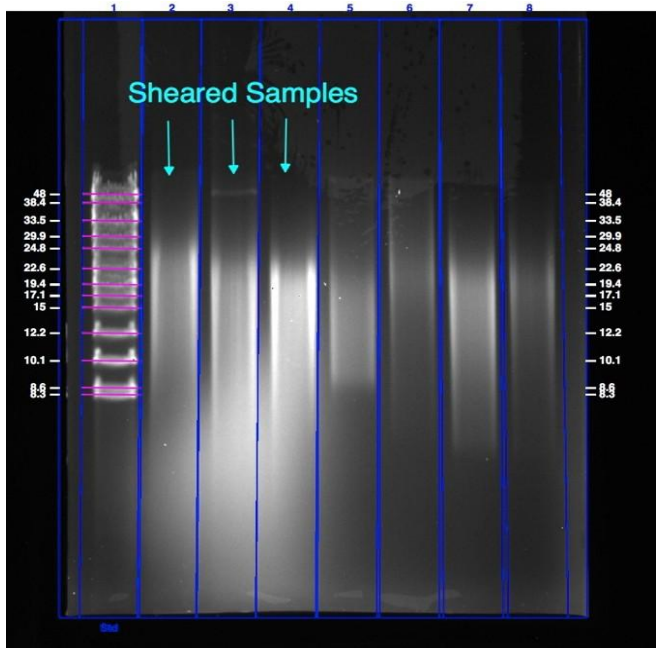
Verify quality of DNA

- Qubit
  - Accurate quantification
- Field Inversion Gel
  - Verify DNA integrity



DNA Shearing

- Covaris g-tube
- Centrifugation
- Synthetic ruby pore
- 4-20 kb

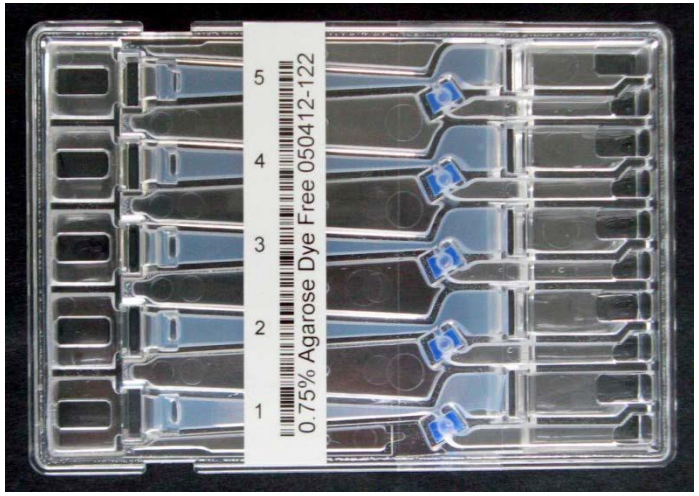


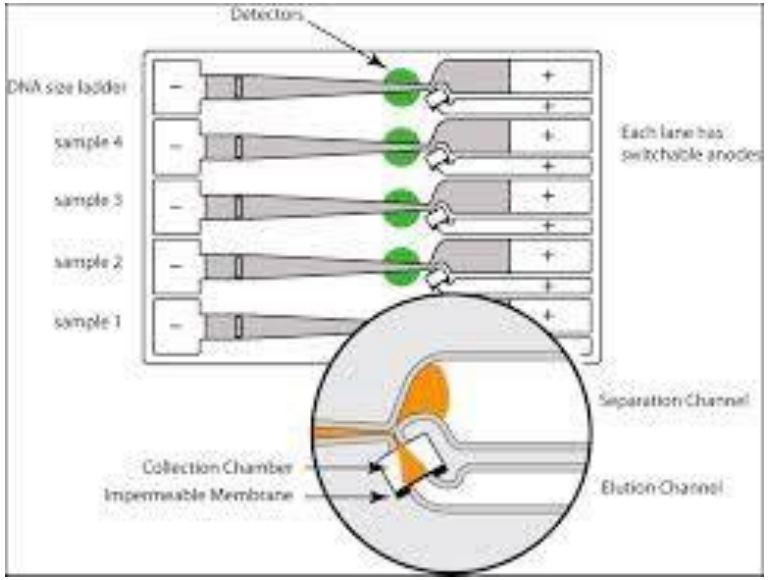
DNA Shearing

- Diagenode megaruptor
- Hydropore
- Controlled flow rate
- 20-75 kb

Size Selection

- Sage Sciences Blue Pippin
- Automated sample collection
- Replaced gel purification

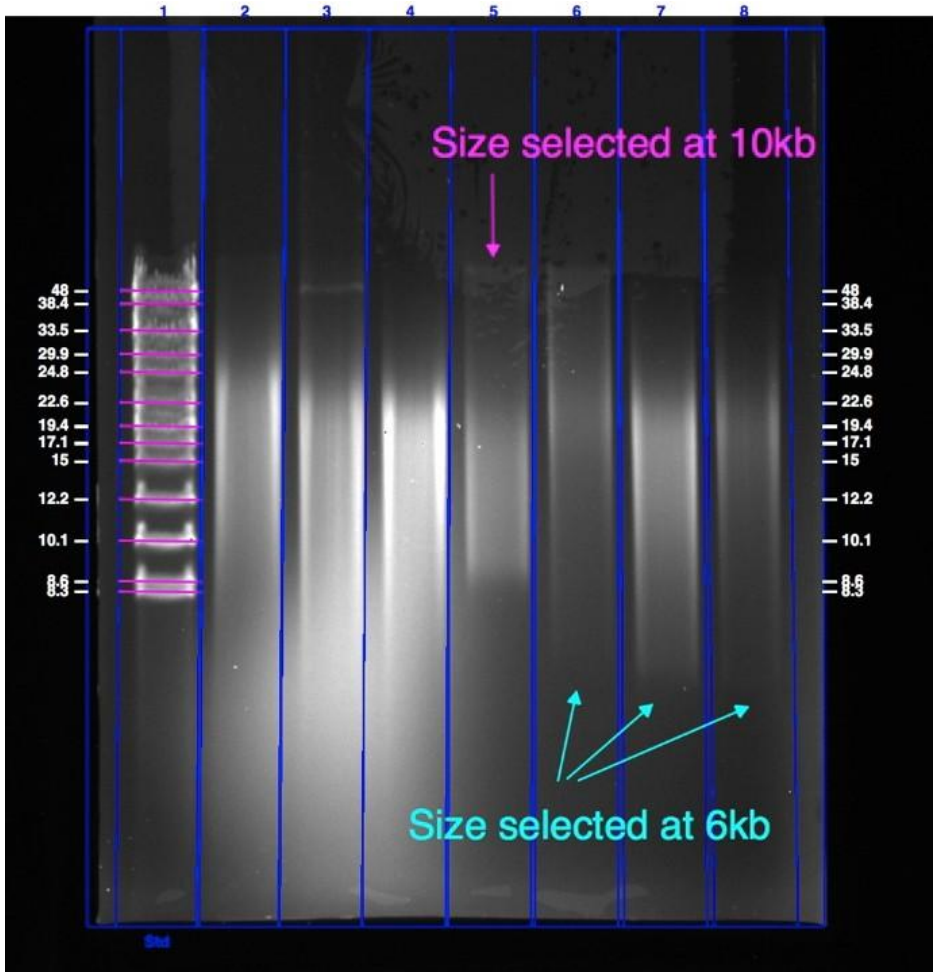




New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

NCGR
National Center for Genome Resources

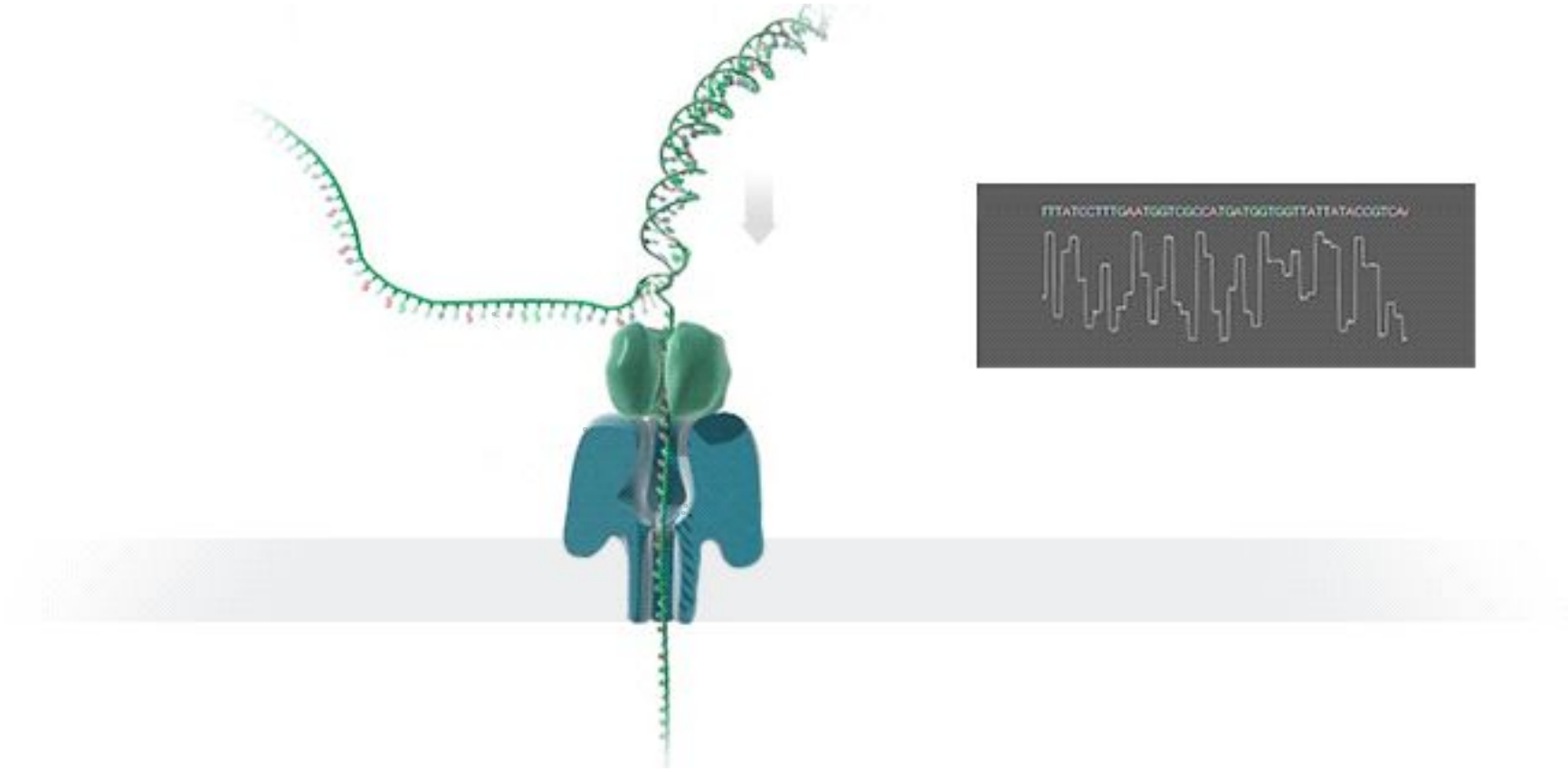
Before and after size selection

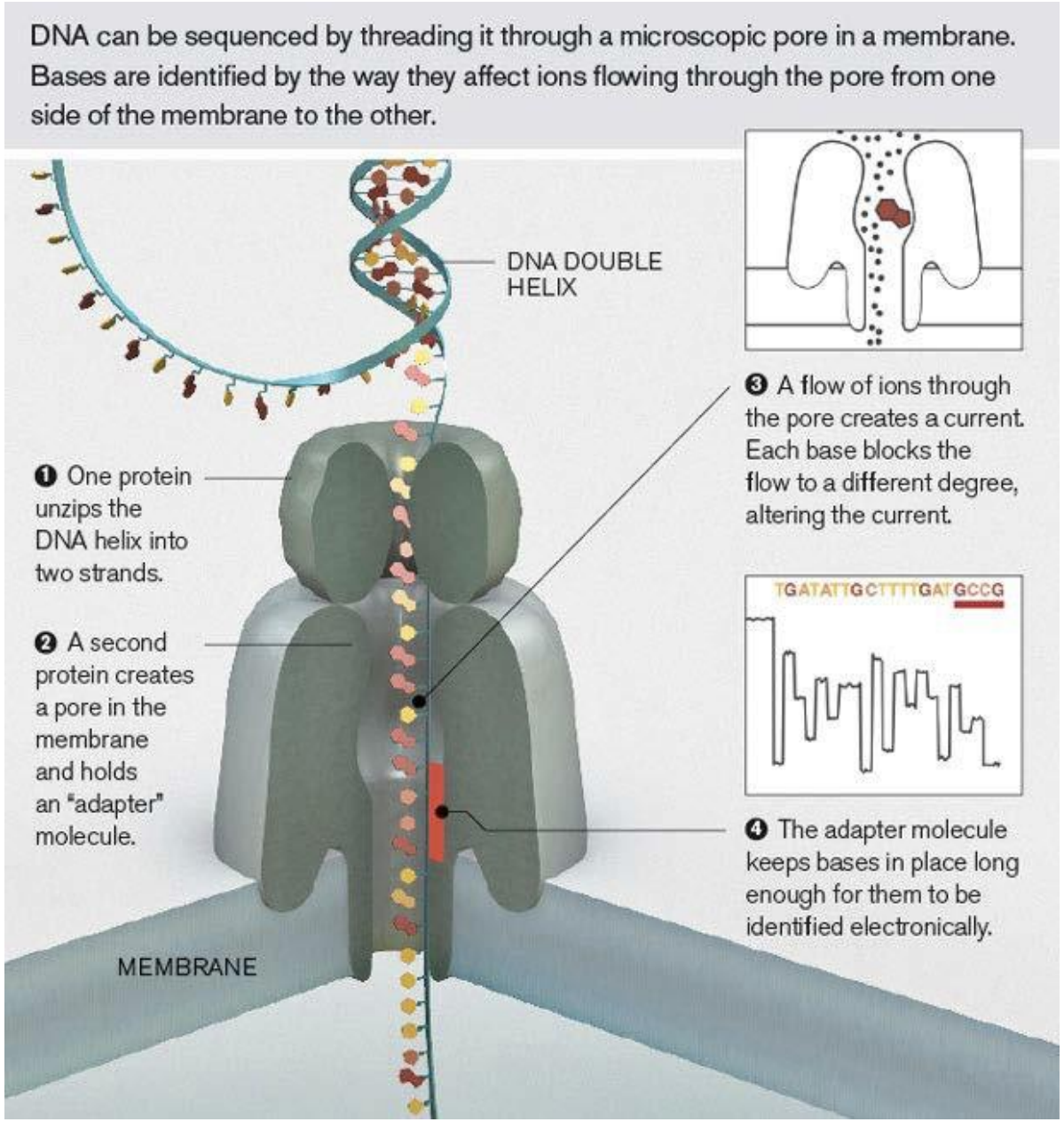


New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

NCGR
National Center for Genome Resources

Oxford Nanopore Technologies





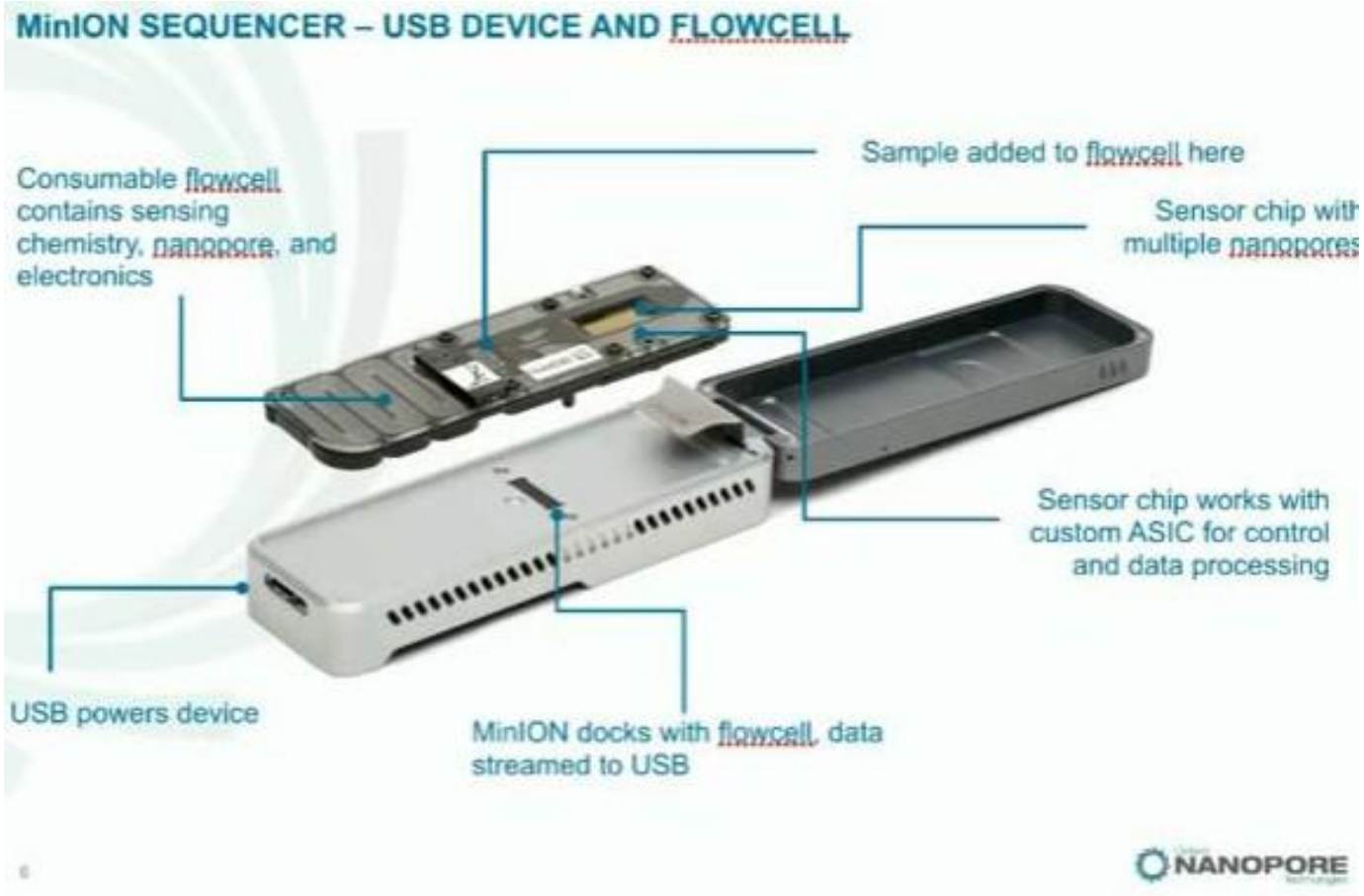
https://youtu.be/GUb1TZvMWsw

https://www.youtube.com/watch?v=CGWZvHIi3i0

New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

NCGR
National Center for Genome Resources

MinIon



"The MinION ***FlowCell*** contains the proprietary sensor array which includes the Nanopores that are needed to perform a complete single-molecule sensing experiment. It also has the ***Application-Specific Integrated Circuit (ASIC)*** used to capture the Nanopore signal used for base calling in the sequencing application."

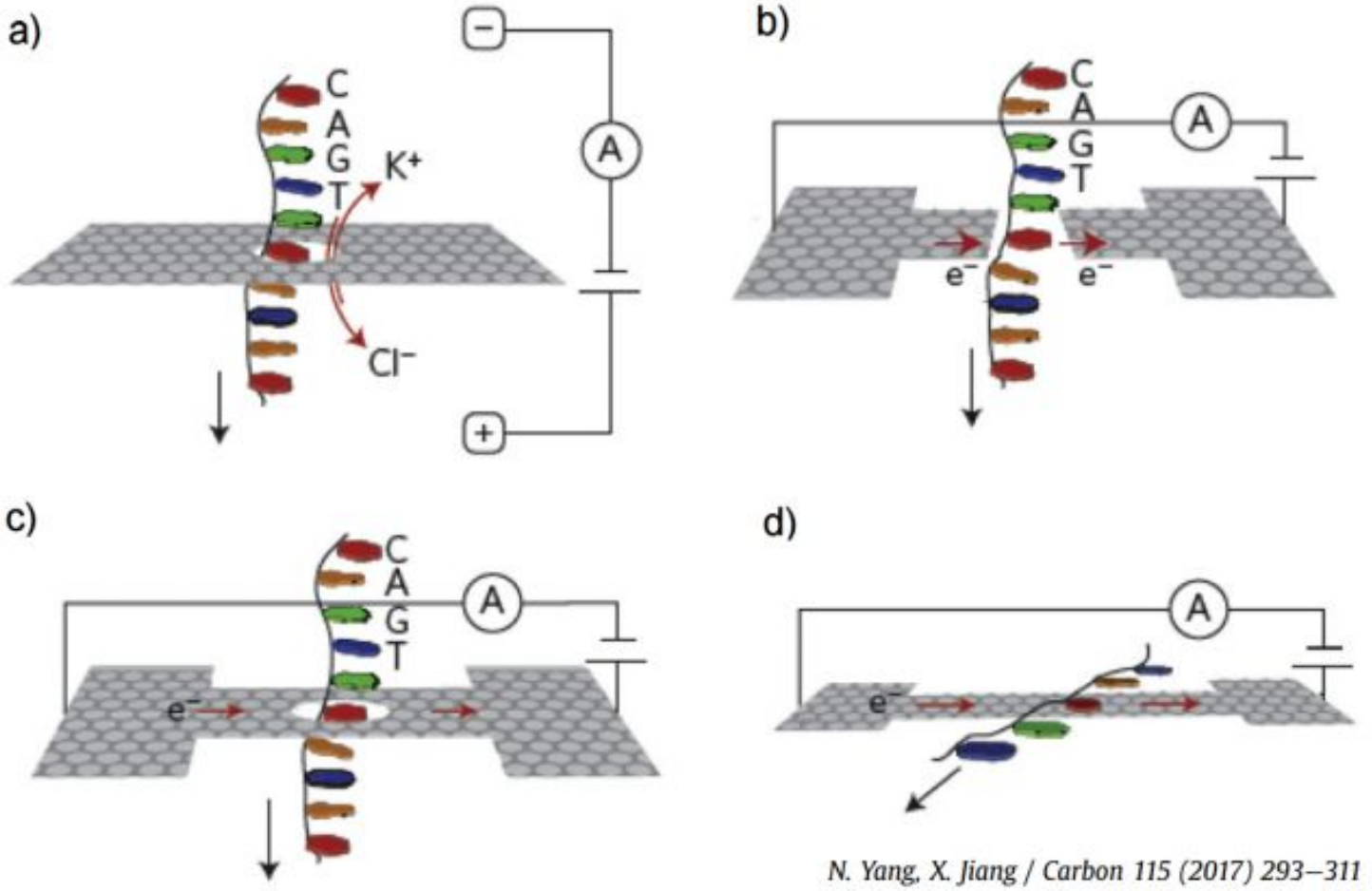
MinION MK1C

ONT features

- Introduced 2012
- Sequencing is real time. Can analyze as the sequence is read
- No amplification
- Scalability
- Use in the field
- Essentially eliminates capital cost

Scalability

Solid State Nanopores



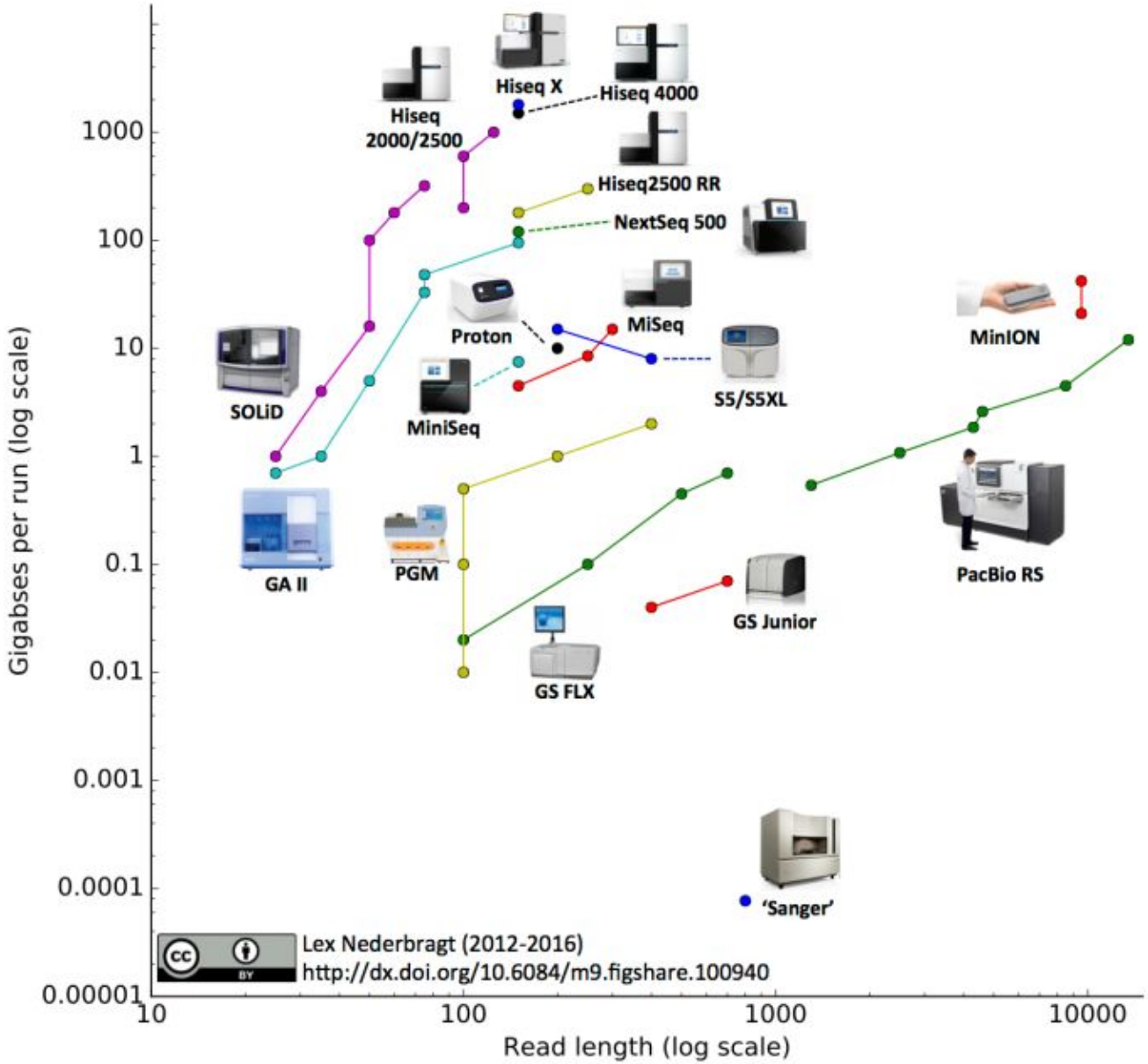
https://mappingignorance.org/2017/01/23/graphene-nanopore-dna-sequencing/

Short read vs long read sequencing

What are the advantages and disadvantages?

What applications are best for each technology?

NovaSeq 6000


Hiseq X
Hiseq 4000
Hiseq 2000/2500
Hiseq2500 RR
NextSeq 500
Proton
MiSeq
MinION
SOLiD
MiniSeq
S5/S5XL
GA II
PGM
PacBio RS
GS Junior
GS FLX
'Sanger'
Lex Nederbragt (2012-2016)
http://dx.doi.org/10.6084/m9.figshare.100940
Gigabses per run (log scale)
Read length (log scale)
1000
100
10
1
0.1
0.01
0.001
0.0001
0.00001
10
100
1000
10000

Matching the library and sequencing technology to the project

1. Expression-related
2. What are the samples?
3. How many samples?
4. How many replicates?
5. What library prep strategies?
6. What sequencing technology?
7. How much sequencing?
8. What analysis tools?
9. Conclusions?

mRNA
Full-length mRNA
Isoforms
Nascent transcription
miRNA or small RNAs
Single cell
Spatial transcriptomics
Methylation and other DNA modifications
Histones and histone modifications
Open chromatin
Transcription Factor Binding Sites
Stranded RNA
…

The bioinformatics workshop is supported by New Mexico INBRE through an Institutional Development Award (IDeA) from National Institute of General Medical Sciences of the NIH grant number P20GM103451

New Mexico INBRE
IDeA Networks of Biomedical Research Excellence

NCGR
National Center for Genome Resources