Post-Lab Review, Etc

Josh Granek

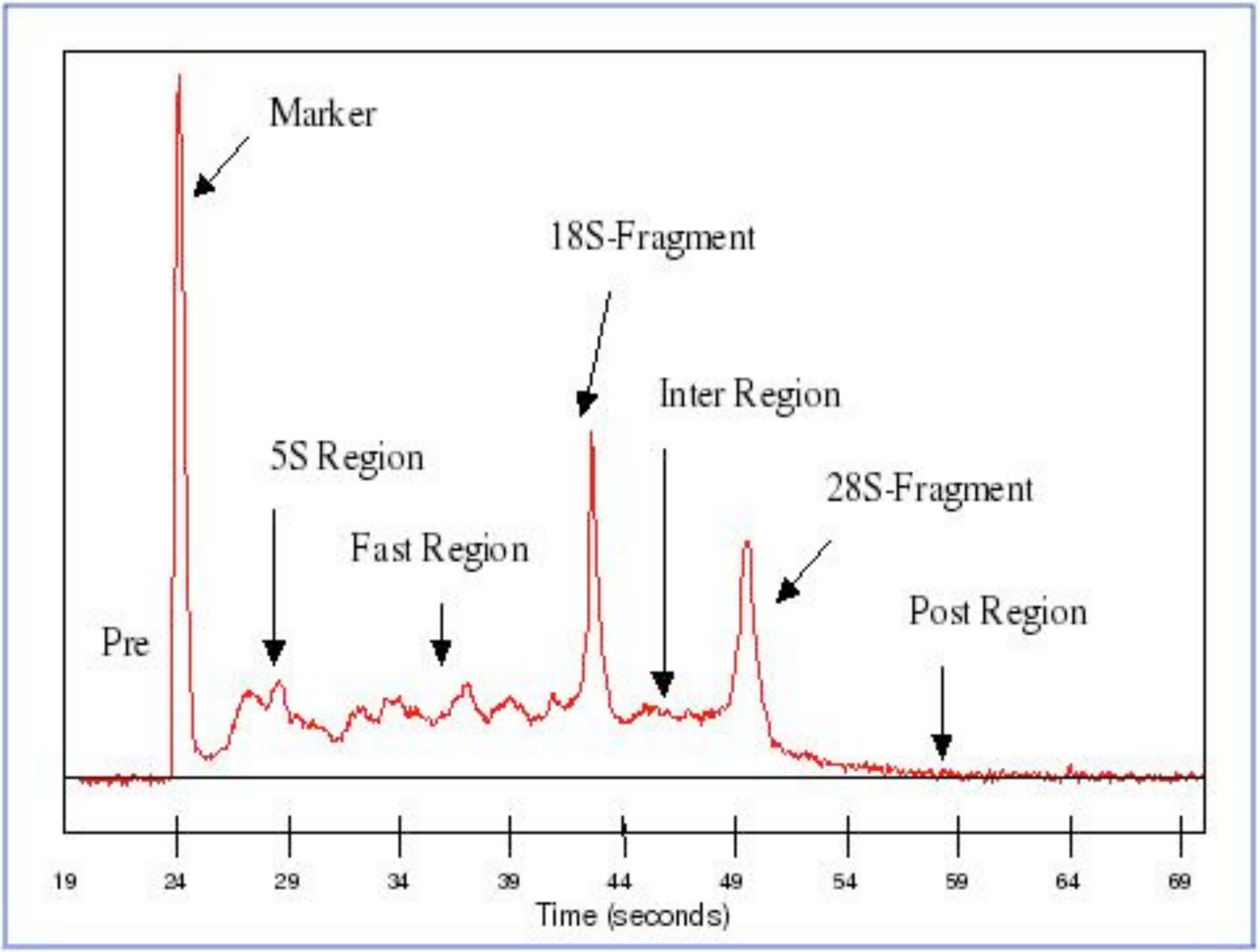
RNA Quality?

- RIN: RNA Integrity Number

http://cartoon-yeah.blogspot.com/2011/03/ren-and-stimpy-picture-4.html

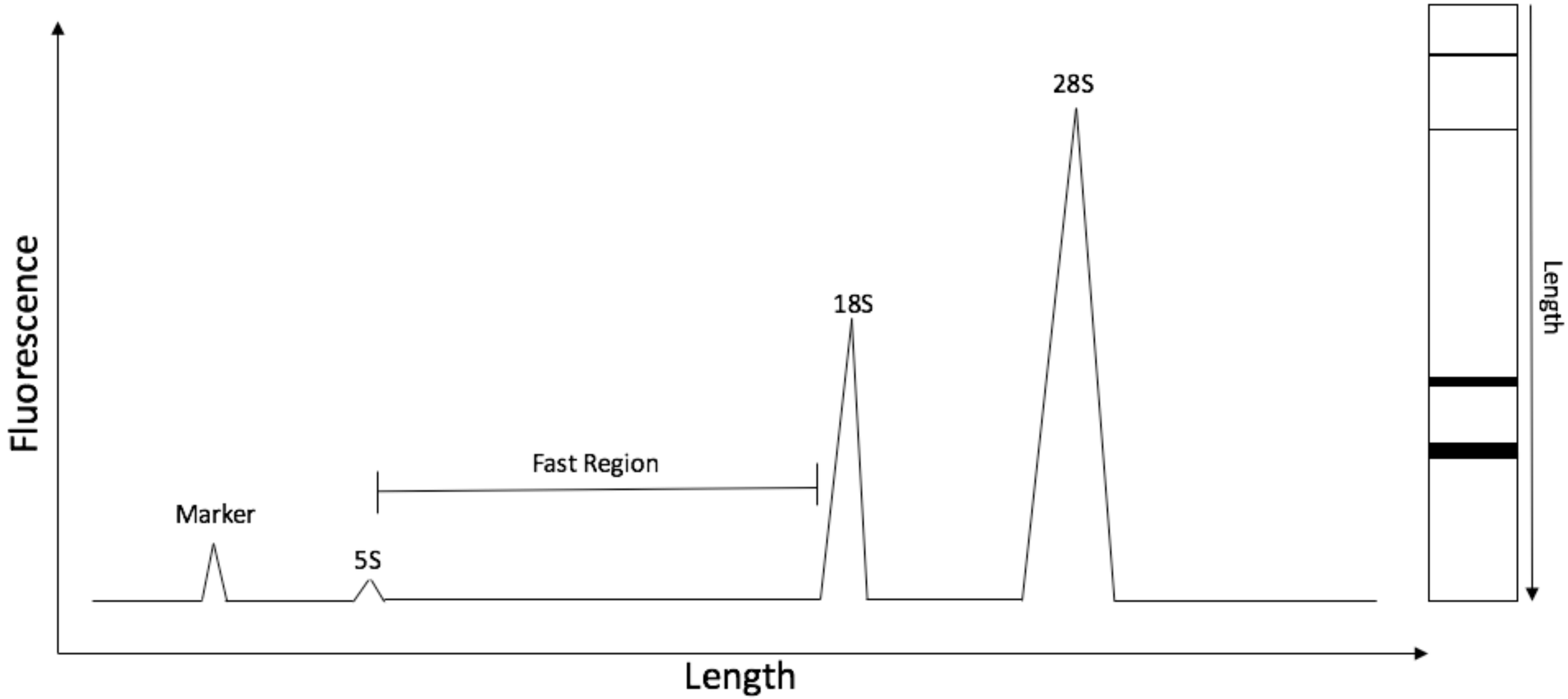
Ratio of 28S to18S ribosomal RNA

- Ratio o

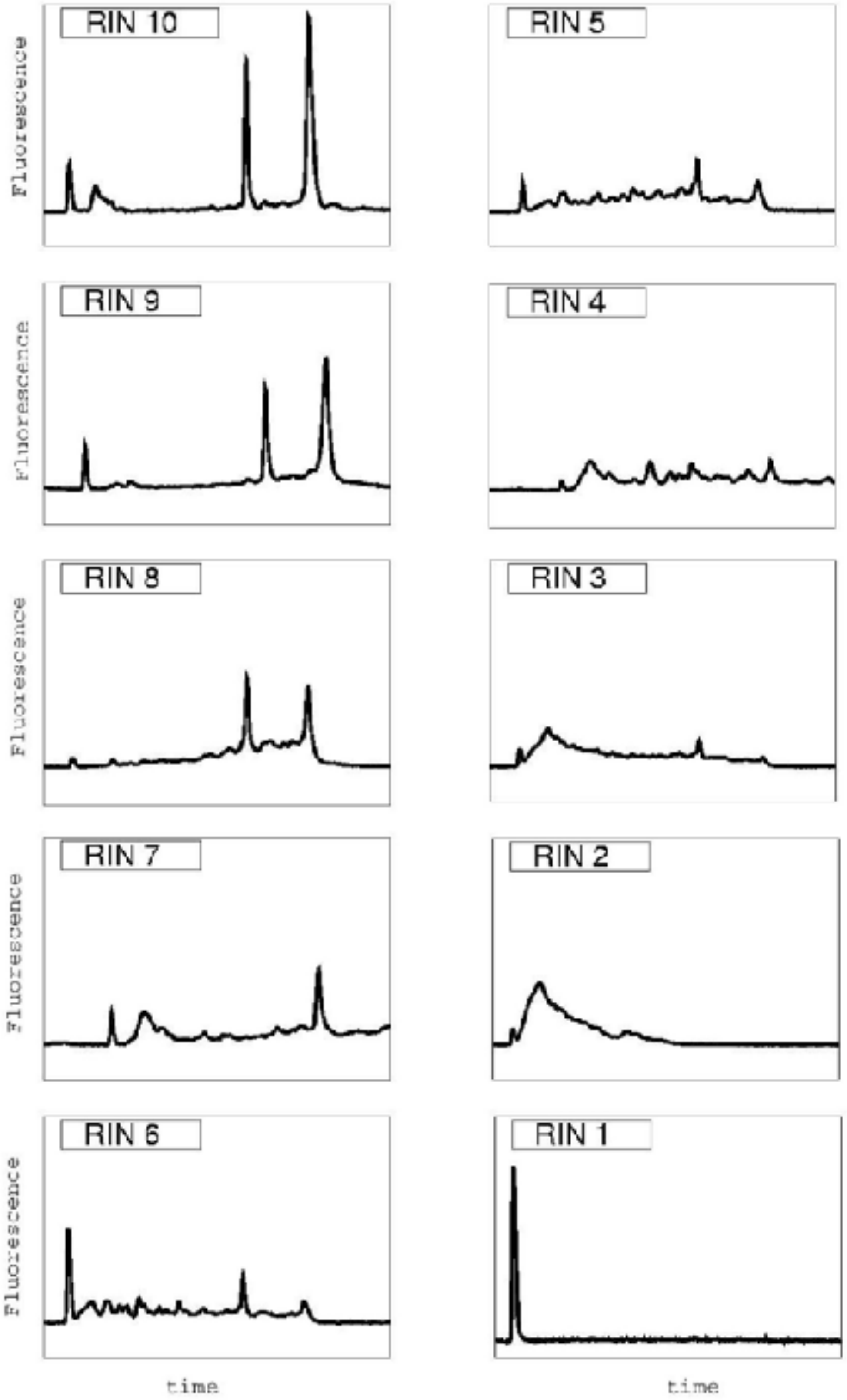


http://dx.doi.org/10.1186/1471-2199-7-3

Electropherogram

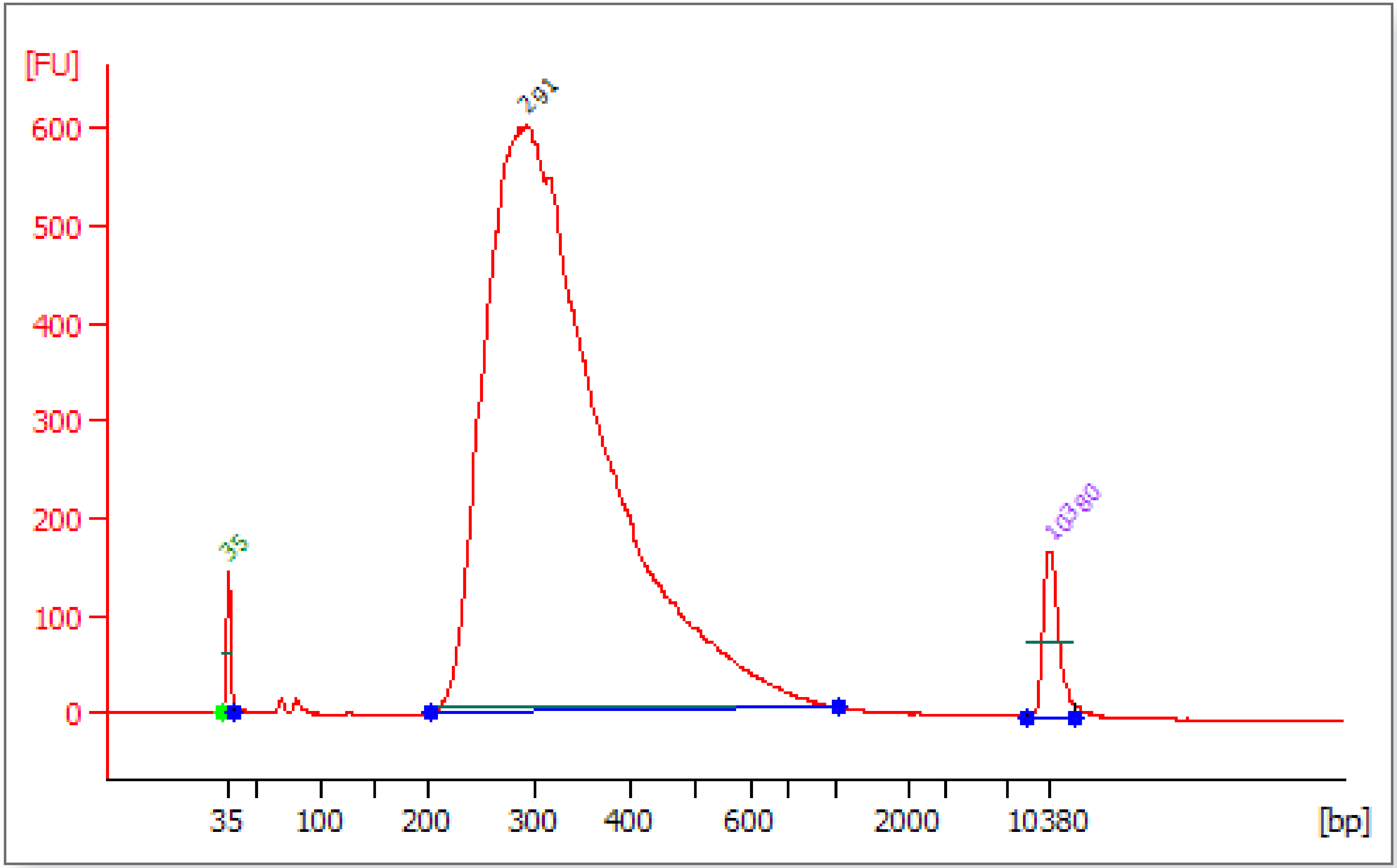


https://upload.wikimedia.org/wikipedia/commons/a/a2/Sample_RNA_electropherogram.png



http://dx.doi.org/10.1186/1471-2199-7-3

RNA Library Size Distribution



NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina®Instruction Manual, Revision 7.0

Assessment of RNA/DNA Quantity and Quality

- Advanced Analytical: Fragment Analyzer
- PerkinElmer: LabChip GX Touch
- Agilent: Bioanalyzer
- Agilent: TapeStation

Barcode Combinations

- Excitation Frequency
  - Red: A and C
  - Green: G and T
- Need both frequencies in each cycle for image registration

Barcode Combinations

GOOD

PRIMER	INDEX SEQUENCE							
P1-A1	T	T	A	C	C	G	A	C
P2-A2	A	G	T	G	A	C	C	T
P3-A3	T	C	G	G	A	T	T	C
P4-A4	C	A	A	G	G	T	A	C
	✔	✔	✔	✔	✔	✔	✔	✔

PRIMER	INDEX SEQUENCE							
P41-D5	G	A	C	G	T	C	A	T
P42-D6	C	T	T	A	C	A	G	C
P43-D7	T	C	C	A	T	T	G	C
P44-D8	A	G	C	G	A	G	A	T
	✔	✔	✔	✔	✔	✔	✔	✔

BAD

PRIMER	INDEX SEQUENCE							
P9-A9	C	G	C	A	A	C	T	A
P10-A10	C	G	T	A	T	C	T	C
P11-A11	G	T	A	C	A	C	C	T
P12-A12	C	G	G	C	A	T	T	A
	✔	✖	✔	✖	✔	✔	✔	✔

PRIMER	INDEX SEQUENCE							
P56-E8	T	A	T	G	G	C	A	C
P57-E9	C	T	C	G	A	A	C	A
P58-E10	C	A	A	C	T	C	C	A
P59-E11	G	T	C	A	T	C	G	T
	✔	✔	✔	✔	✔	✖	✔	✔

NEBNext® Multiplex Oligos for Illumina® (96 Index Primers) Instruction Manual NEB #E6609S/L

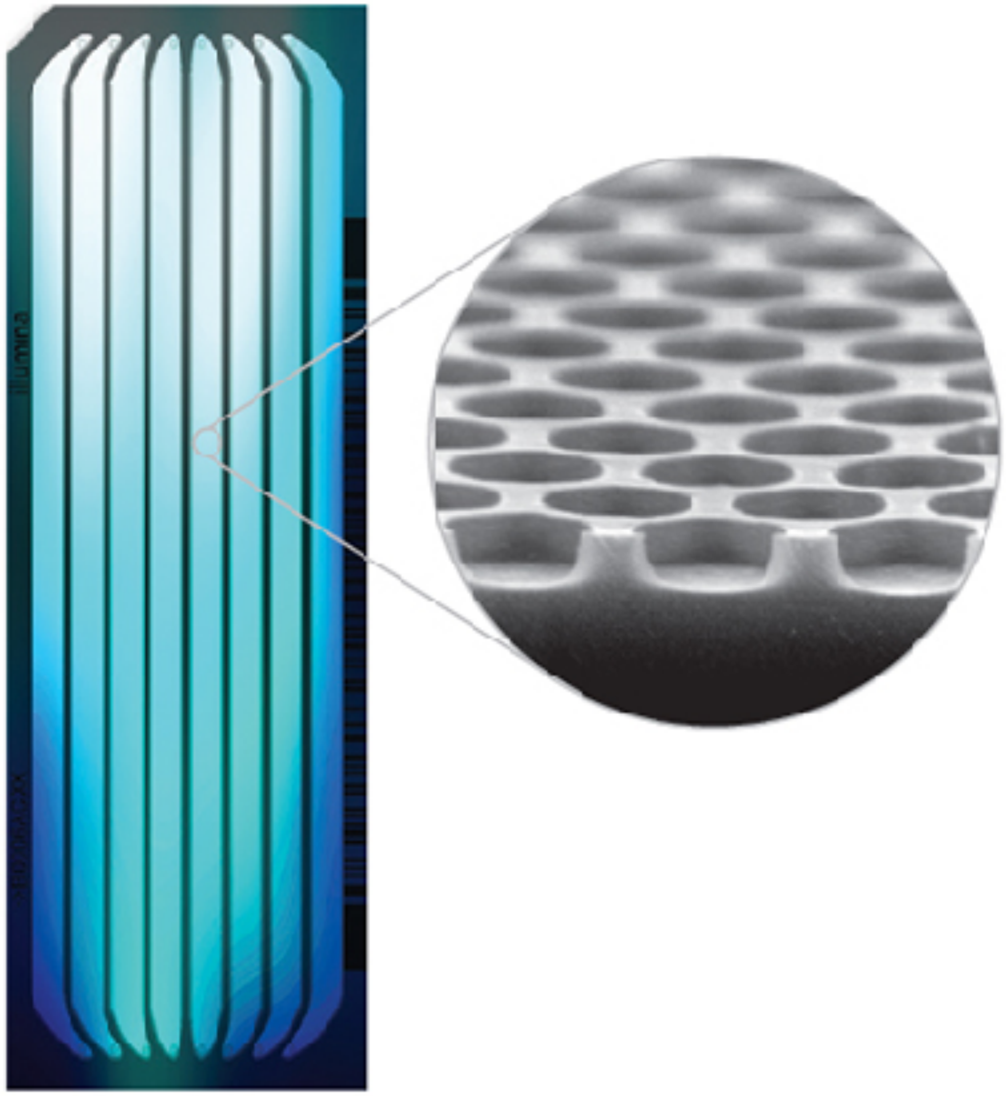
MiSeq, NextSeq, and More Seqs

	MiSeq	**NextSeq**	**HiSeq 4000**	**NovaSeq 6000**
Maximum Output	15 Gb	120 Gb	750 Gb	3000 Gb
Maximum Reads per Run	25 million	400 million	2.5 billion	10 billion
Maximum Read Length	2 × 300 bp	2 x 150 bp	2 × 150 bp	2 × 150 bp
Run Time	4-56 hours	15-29 hours	< 1–3.5 days	13-45 hours
Cost*	$1,787	$4,695	$19,206	$35,538
Cost/Mbp*	$0.119	$0.039	$0.026	$0.012

* Duke Sequencing and Genomic Technologies Shared Resource, July 2018

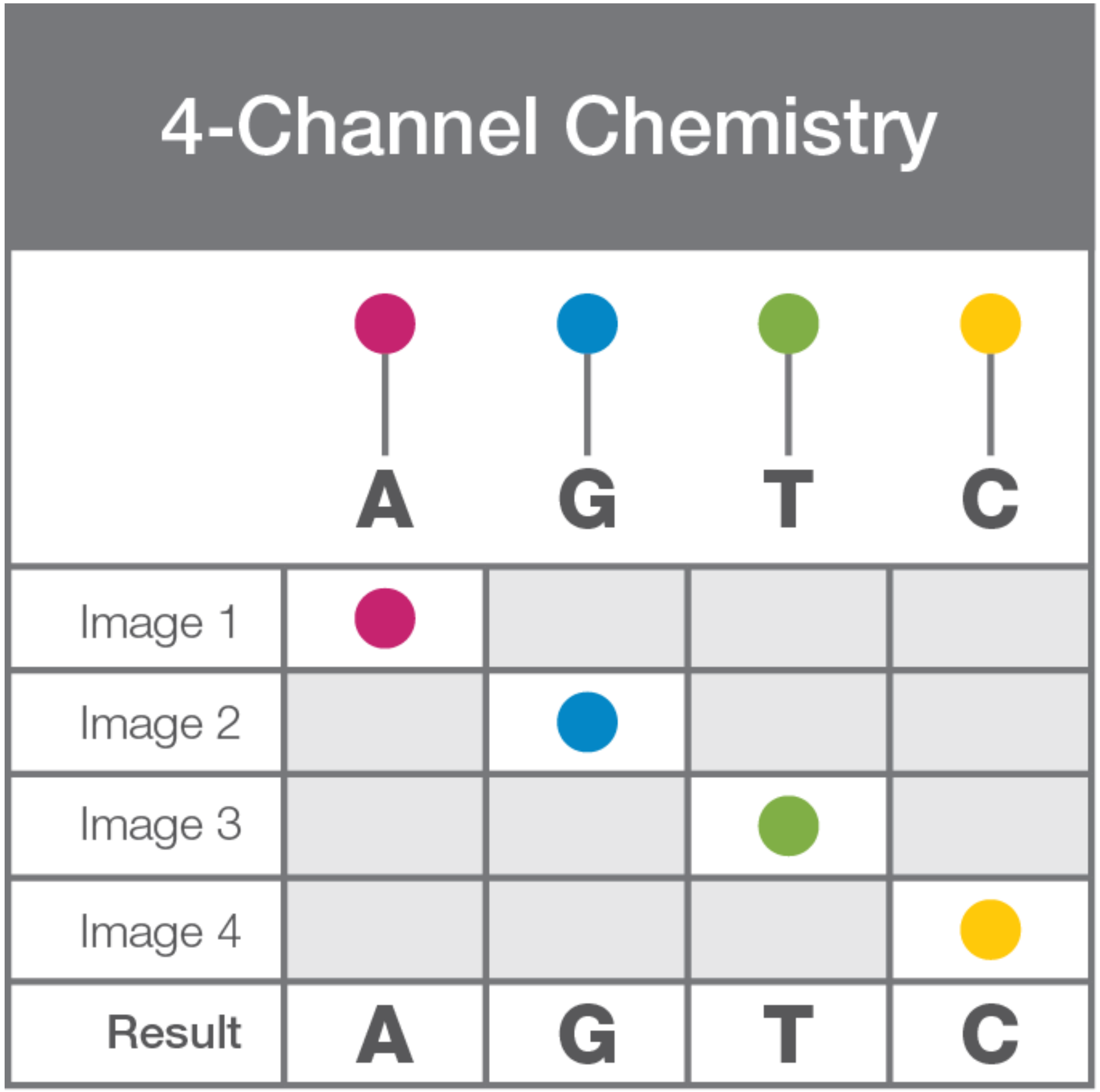
Patterned Flow Cells

- ExAmp
- Machines
  - HiSeq X
  - HiSeq 3000/4000
  - NovaSeq 6000



https://assets.illumina.com/content/dam/illumina-marketing/images/science/v2/web-graphic/nano-well-technology-web-graphic.jpg

4-Channel Chemistry

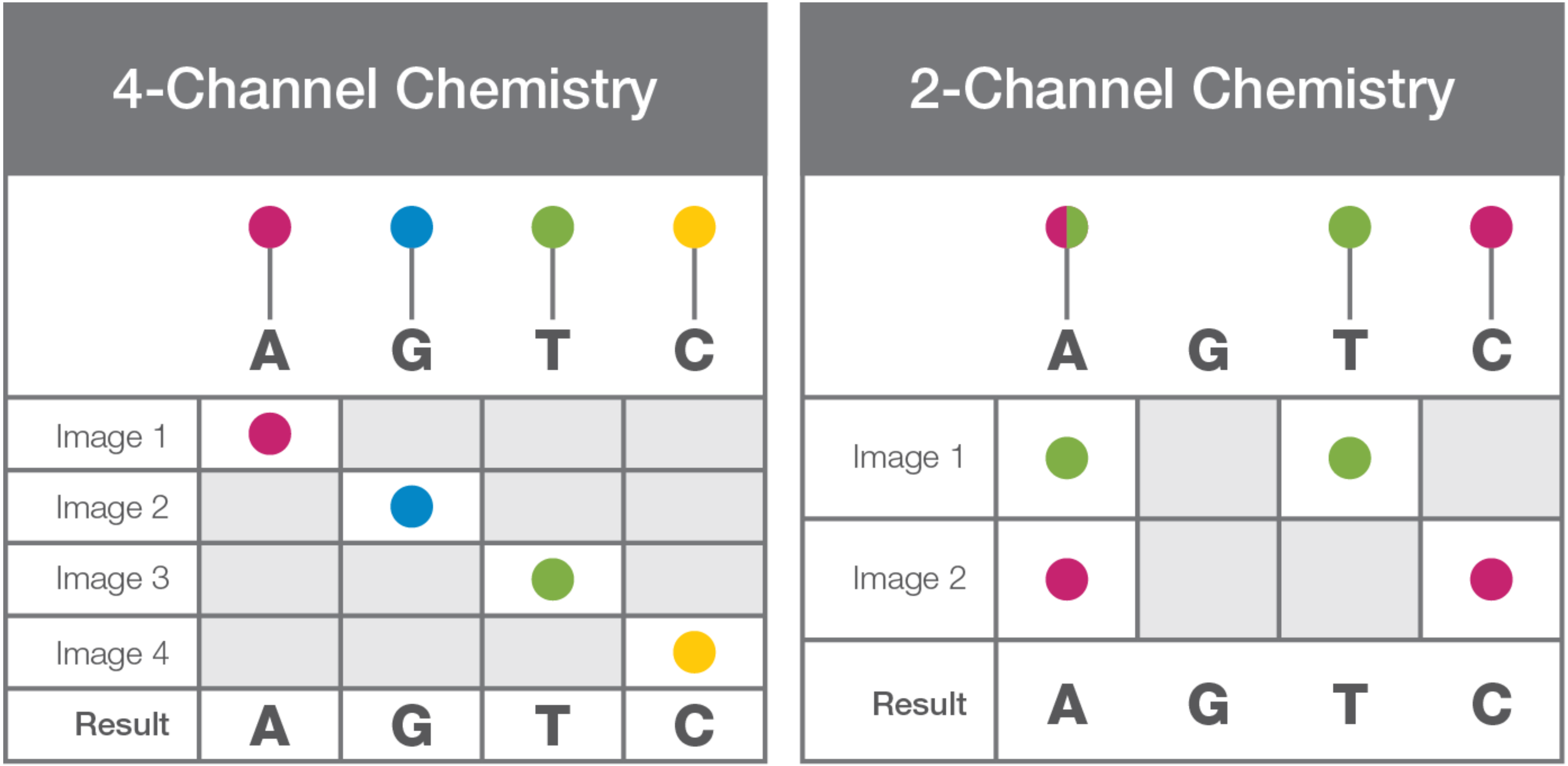


4-Channel Chemistry

	● A	● G	● T	● C
Image 1	●			
Image 2		●		
Image 3			●	
Image 4				●
Result	**A**	**G**	**T**	**C**

https://www.illumina.com/content/dam/illumina-marketing/documents/products/techspotlights/cmos-tech-note-770-2013-054.pdf

2-Channel Chemistry



https://www.illumina.com/content/dam/illumina-marketing/documents/products/techspotlights/cmos-tech-note-770-2013-054.pdf

Uracil DNA glycosylase
and
DNA lyase

Uracil DNA glycosylase: What

- Remove Uracil base from DNA

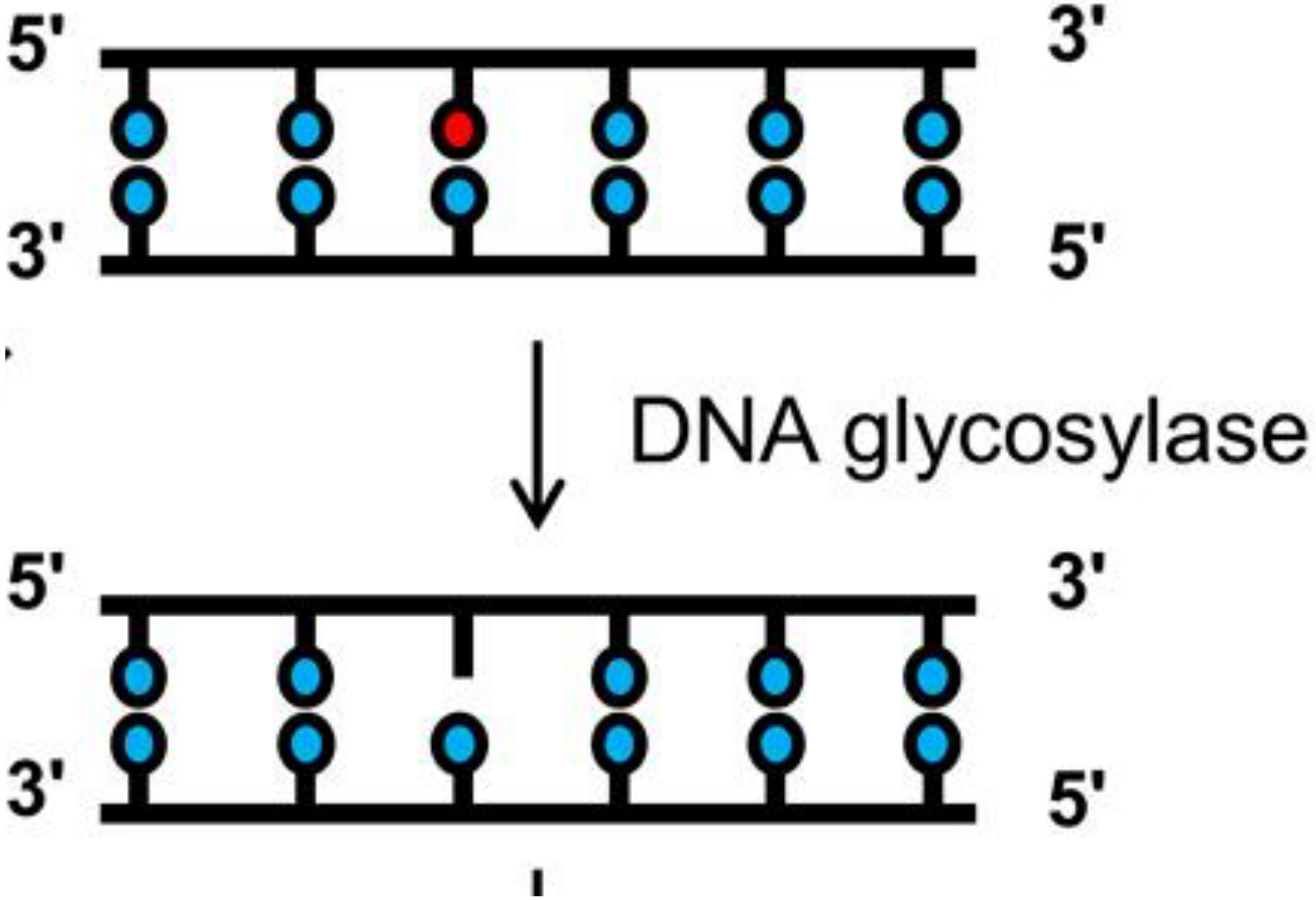
Uracil DNA glycosylase: What

```
5′-CTGATCUGACTGATG-3'
3′-GACTAGACTGACTAC-5′
```

↓

```
5′-CTGATC-GACTGATG-3′
3′-GACTAGACTGACTAC-5′
```

Uracil DNA glycosylase: What



http://www.frontiersin.org/files/Articles/98174/fgene-05-00232-r2/image_m/fgene-05-00232-g001.jpg

DNA Lyase: What

- Cleave DNA backbone at abasic site

DNA Lyase: What

```
5′-CTGATC-GACTGATG-3′
3′-GACTAGACTGACTAC-5′
```

↓

```
5′-CTGATC GACTGATG-3′
3′-GACTAGACTGACTAC-5′
```

Comparing Technologies

Method	Read length	Accuracy	Reads per run	Max Output	Cost ($/Mb)	Pros	Cons
Sanger	400-900 bp	99.9%	1	900 bp	$2400	Longer reads.	Expensive. Low Output
Illumina	600 bp (300bp PE)	99.9%	20×10^9	6000 Gb	$0.01	High yield per base cost	Equipment expense. Short reads
PacBio	>10kb ave. >40kb max	99%	5×10^5	10 Gb	$0.08	Very long reads	Homopolymer errors. Moderate Output. Equipment expense.
Nanopore	>100 kb N50 >1Mb Max	92%	1×10^6	5 Gb	$0.10	Very long reads Portable Cheap Equipment	Homopolymer errors. Moderate Output.

https://en.wikipedia.org/wiki/DNA_sequencing
https://blog.genohub.com/2017/06/16/pacbio-vs-oxford-nanopore-sequencing/

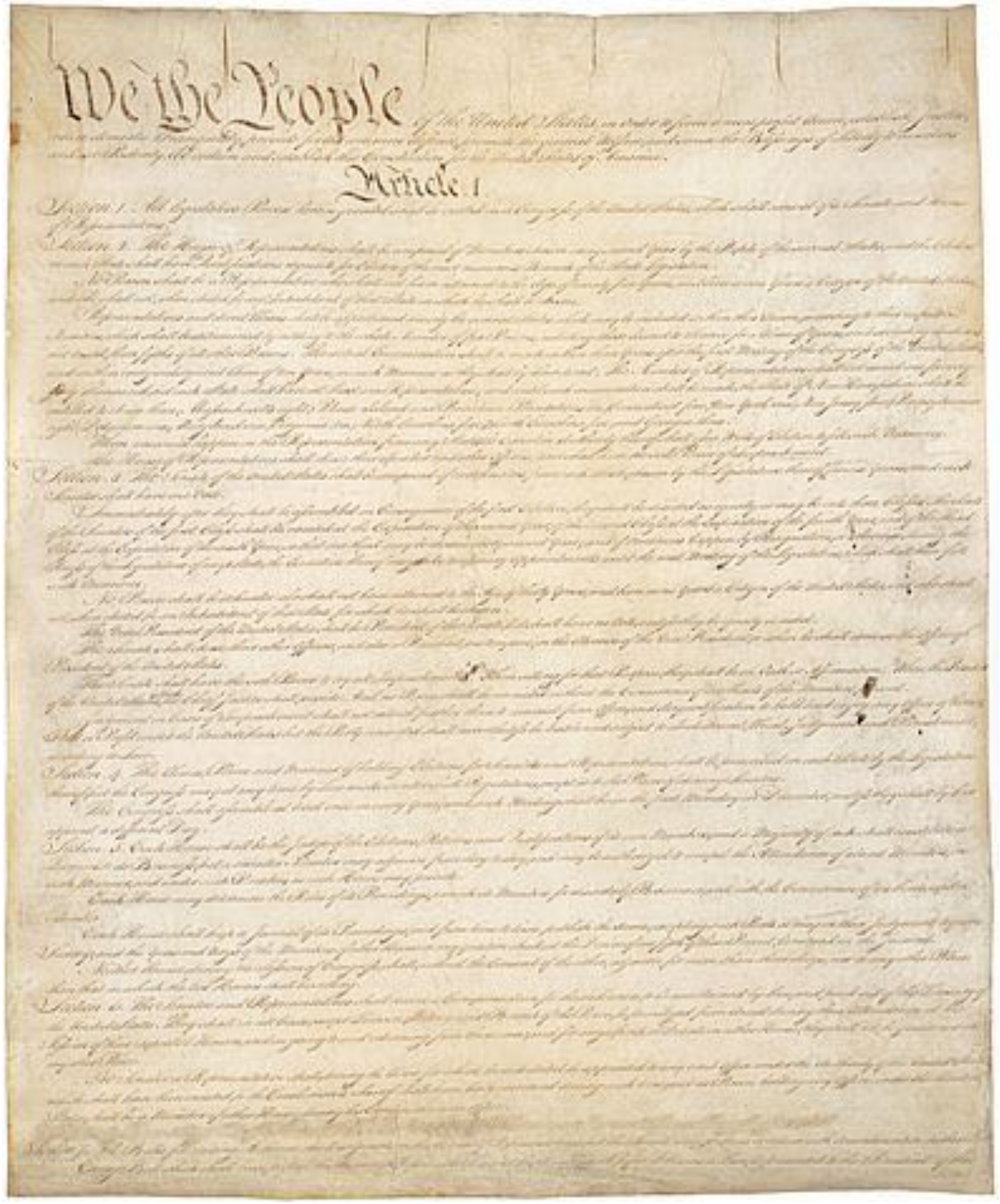
Why Long Reads?

- Structural Variation
  - Large Insertions or Deletions
  - Duplications
  - Translocations
- De Novo Genome Assembly
- Phasing

Short Reads

```
e of the U
stablish J
Union, est
nited Stat
 to form a
rder to fo
e perfect 
ion, estab
eople of t
 the Peopl
```

“Genome” Reference



We the People

Article 1
]

Reference Based Mapping

```
We the People of the United States, in Order to form a more perfect Union, establish Justice, insur
```

```
e of the U
stablish J
Union, est
nited Stat
 to form a
rder to fo
e perfect 
ion, estab
eople of t
 the Peopl
```

Reference Based Mapping

```
We the People of the United States, in Order to form a more perfect Union, establish Justice, insur
```

```
   the Peopl
        eople of t
            e of the U
                      nited Stat
                                        rder to fo
                                             to form a
                                                          e perfect
                                                                    Union, est
                                                                      ion, estab
                                                                            stablish J
```

De Novo Assembly

Overlapping Random Fragments

```
rious disg

Age.  "You

rinking Ag

uises of A

the portra

ugh the po

of every D

nking Age.

r various 

, under va
```

Assemble Contigs

```
        Age.  “You
rinking Ag
  nking Age.
```

```
          rious disg
      r various
, under va
```

```
    the portra
ugh the po
```

```
uises of A
```

```
of every D
```

Assemble Contigs

`rinking Age.  "You`

`, under various disg`

`ugh the portra`

`uises of A`

`of every D`

Assemble Contigs

rinking Age. "You

, under various disg

ugh the portra

uises of A

of every D

ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little

More Reads

```
rious disg
Age."You
rinking Ag
uises of A
the portra
ugh the po
of every D
nking Age.
r various
, under va
Age."Yo
rough the
rinking Ag
ed, under
ugh the po
ry Drinkin
sguises of
 u are a li
"You are
, under va
```

More Reads

```
rough the
  ugh the po
  ugh the po
      the portra
```

```
ed, under
  , under va
  , under va
        r various
            rious disg
                    sguises of
                      uises of A
```

```
of every D
     ry Drinkin
         rinking Ag
         rinking Ag
           nking Age.
                 Age.  "You
                 Age.  "Yo
                       "You are
                          u are a li
```

More Reads

`rough the portra`

`ed, under various disguises of A`

`of every Drinking Age.  "You are a li`

More Reads

rough the portra

ed, under various disguises of A

of every Drinking Age. "You are a li

ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little

Longer Reads

```
various disguises of
 Drinking Age."You
 every Drinking Age.
sguises of Art, thro
 ough the portraits o
through the portrai
raits of every Drink
 ery Drinking Age."
 er various disguises
, under various disg
```

Longer Reads

```
, under various disg
     er various disguises
        various disguises of
                  sguises of Art, thro
                                  through the portrai
                                     ough the portraits o
                                                  raits of every Drink
                                                           every Drinking Age.
                                                             ery Drinking Age.  "
                                                                 Drinking Age.  "You
```

Longer Reads

```
, under various disguises of Art, through the portraits of every Drinking Age.  "You
```

```
ed, under various disguises of Art, through the portraits of every Drinking Age.  "You are a little
```

Fragmentation

```
"You
Age.
Art,
Drinking
a
are
disguises
ed,
every
little
of
of
portraits
the
through
under
various
```

Problem Sequences

- Repeats
  - Transposons
  - Centromeres
- Homologs
- Duplications

De novo “Reference”

ed, under various disguises of Art, through the portraits of every Drinking Age. "You are a little

— *A Tale of Two Cities*

Single Molecule Technologies

DNA Sequencing Technologies (Abridged)

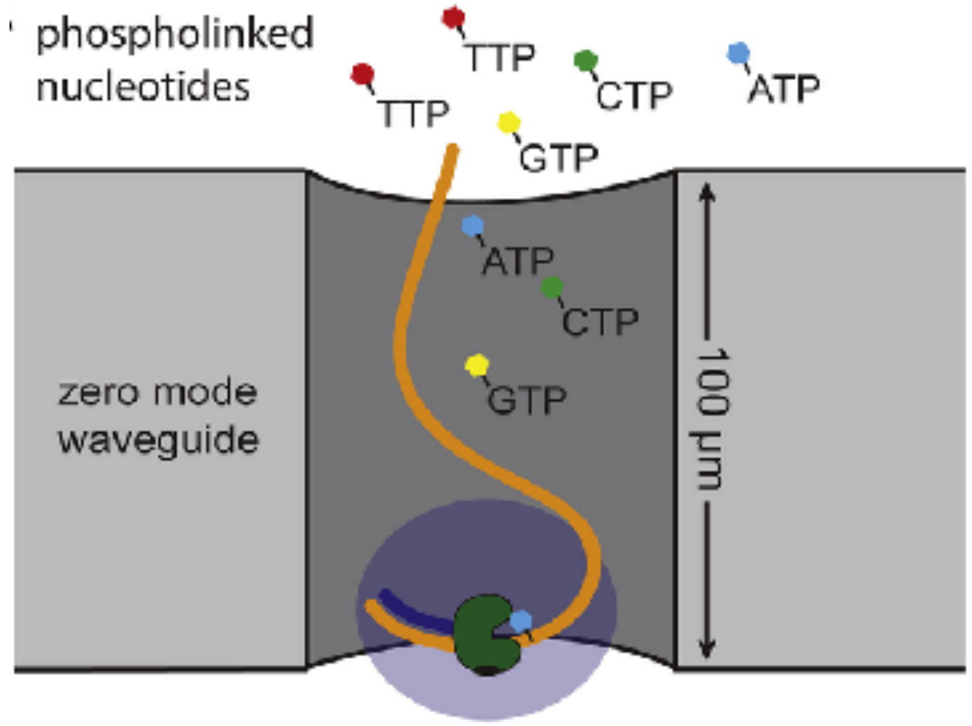
1st Generation	2nd Generation	3rd Generation
Chemical (Maxim-Gilbert)	Pyrosequencing (454)	Single molecule real time (PacBio)
Chain Termination (Sanger)	Chain Termination (Illumina)	Nanopore sequencing (Oxford Nanopore)
Pyrosequencing	Sequencing by ligation (SOLiD sequencing)	
	Ion semiconductor (Ion Torrent)	

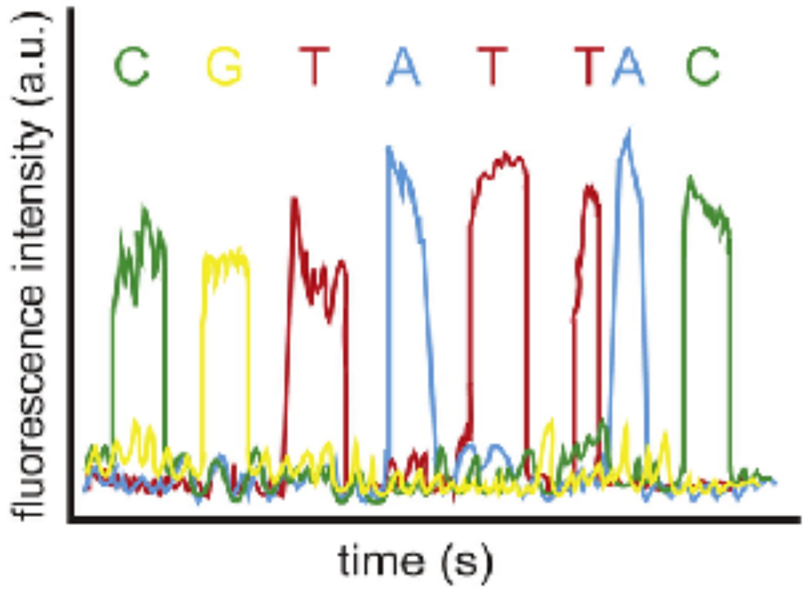
Sequencing by Synthesis

1st Generation	2nd Generation	3rd Generation
Chemical (Maxim-Gilbert)	Pyrosequencing (454)	Single molecule real time (PacBio)
Chain Termination (Sanger)	Chain Termination (Illumina)	Nanopore sequencing (Oxford Nanopore)
Pyrosequencing	Sequencing by ligation (SOLiD sequencing)	
	Ion semiconductor (Ion Torrent)	

1st Generation	2nd Generation	3rd Generation
Chemical (Maxim-Gilbert)	Pyrosequencing (454)	Single molecule real time (PacBio)
Chain Termination (Sanger)	Chain Termination (Illumina)	Nanopore sequencing (Oxford Nanopore)
Pyrosequencing	Sequencing by ligation (SOLiD sequencing)	
	Ion semiconductor (Ion Torrent)	

Pacific Biosciences

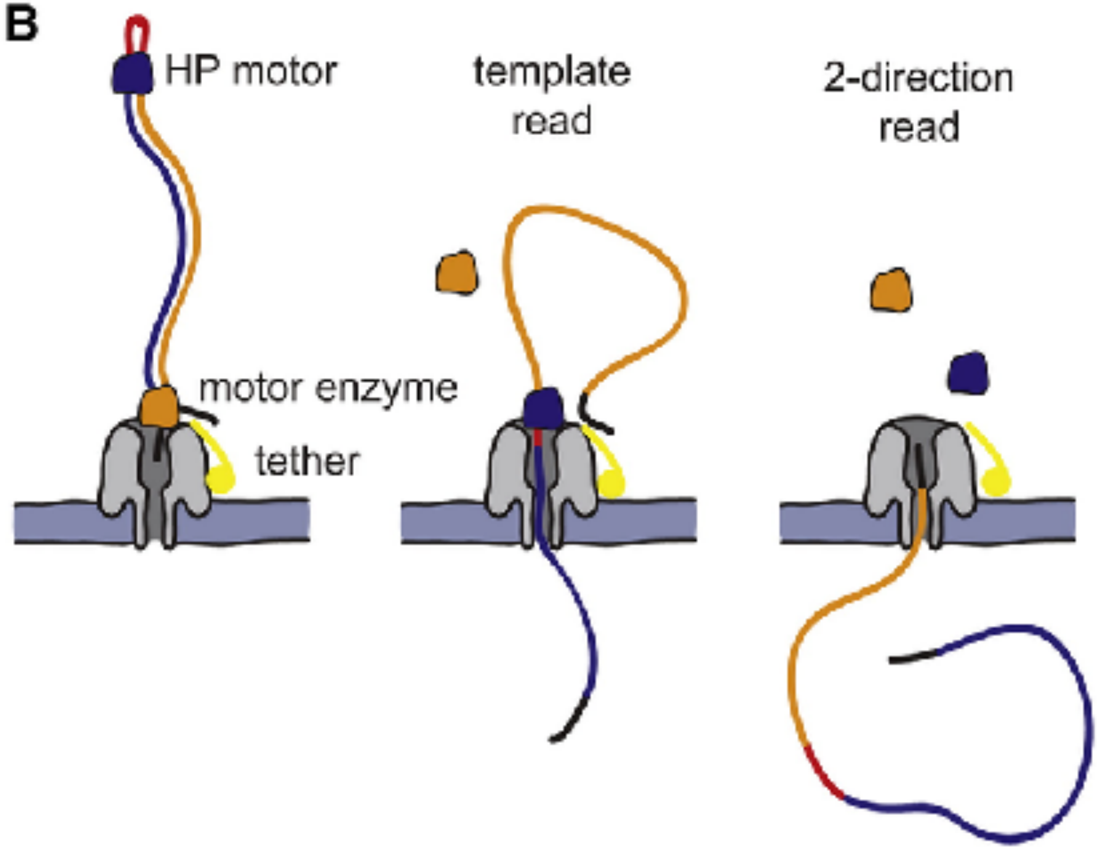


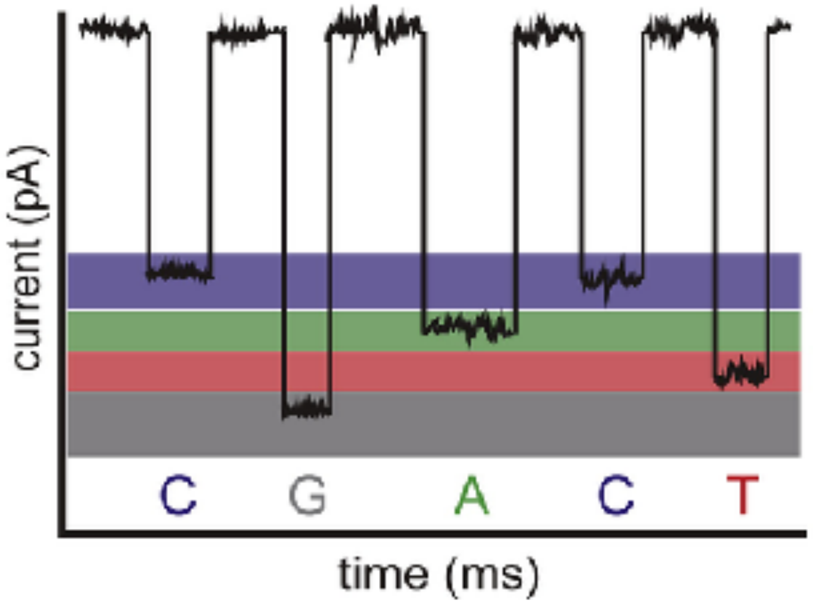


Reuter, et al 2015

1st Generation	2nd Generation	3rd Generation
Chemical (Maxim-Gilbert)	Pyrosequencing (454)	Single molecule real time (PacBio)
Chain Termination (Sanger)	Chain Termination (Illumina)	Nanopore sequencing (Oxford Nanopore)
Pyrosequencing	Sequencing by ligation (SOLiD sequencing)	
	Ion semiconductor (Ion Torrent)	

Oxford Nanopore





Reuter, et al 2015

Sequencers

https://www2.nanoporetech.com/images/product-page/MinIOn-Banner.jpg
http://www.gatc-biotech.com/en/gatc/sequencing-technologies/pacbio-rs-ii.html
http://www.dnavision.com/illumina.php

DNA-Seq

1. Purify DNA
2. Fragment
3. Size Select
4. Adapter Ligation

RNA-Seq

1. Purify RNA
2. Fragment
3. Size Select
4. Make DNA From RNA
5. Adapter Ligation