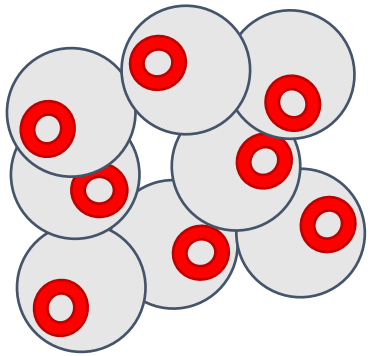
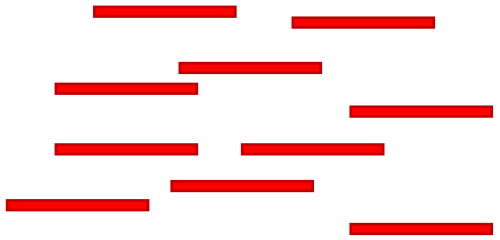
Introduction to Genome Sequencing

Dr Liz Batty

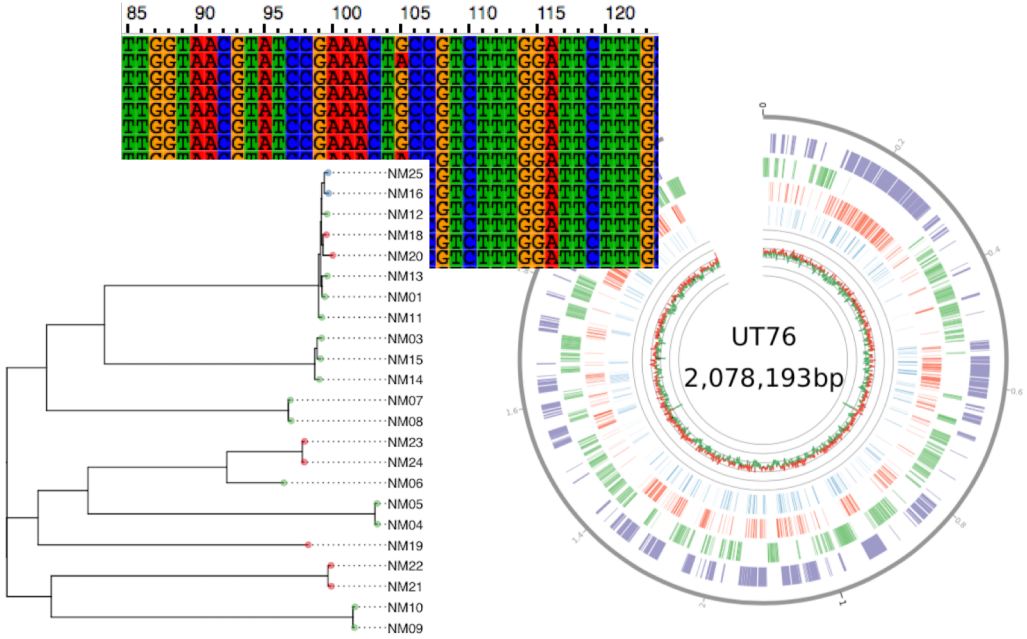
MORU Tropical Health Network

How do we sequence bacterial genomes?





```
@SEQ_ID
GATTTGGGGTTCAAAGCAGTATCGATCAA
+
!''*((((***+))%%%++)(%%%%).1*
```



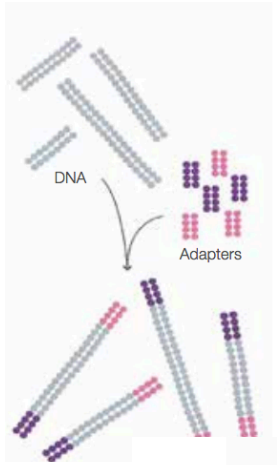
Technology	Machine	Read length	Output from one run	Error Rate
Illumina	HiSeq	75-150bp	250-300Gb	<1%
	MiSeq	75-300bp	15Gb	
Ion Torrent		200-400bp	2Gb	~1%
PacBio		Tens of kb	Up to 10Gb	~15%
Oxford Nanopore		Tens of kb, up to 1Mb	~15Gb	~15%

Illumina and Ion Torrent

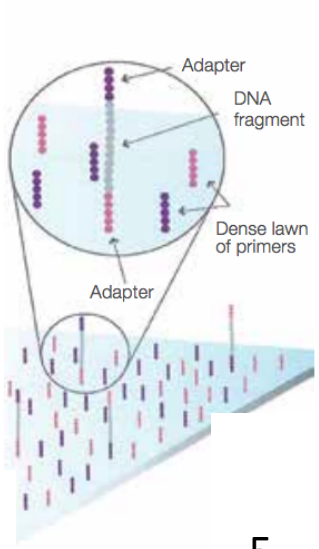
- Cut DNA into short fragments
- Amplify
- Add bases and detect incorporation

Illumina sequencing - amplification

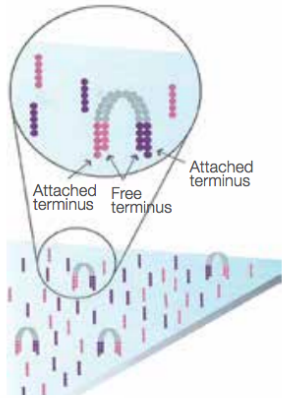
1.



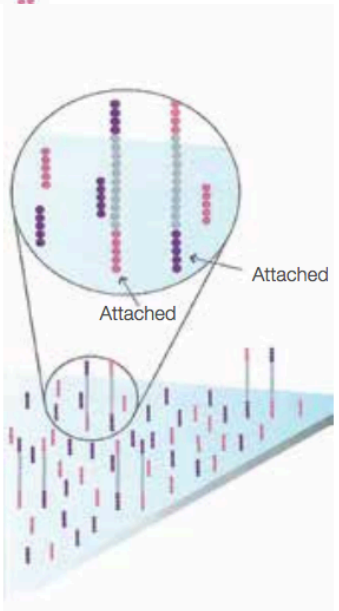
2.



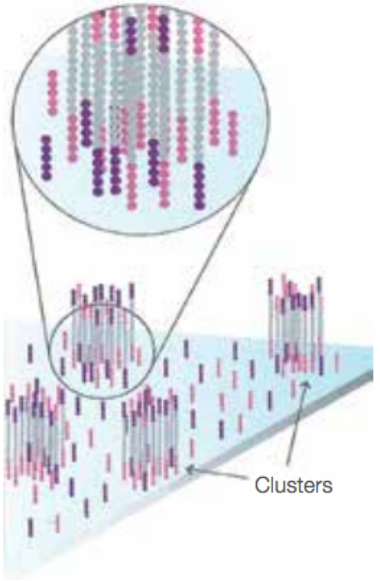
3.



4.

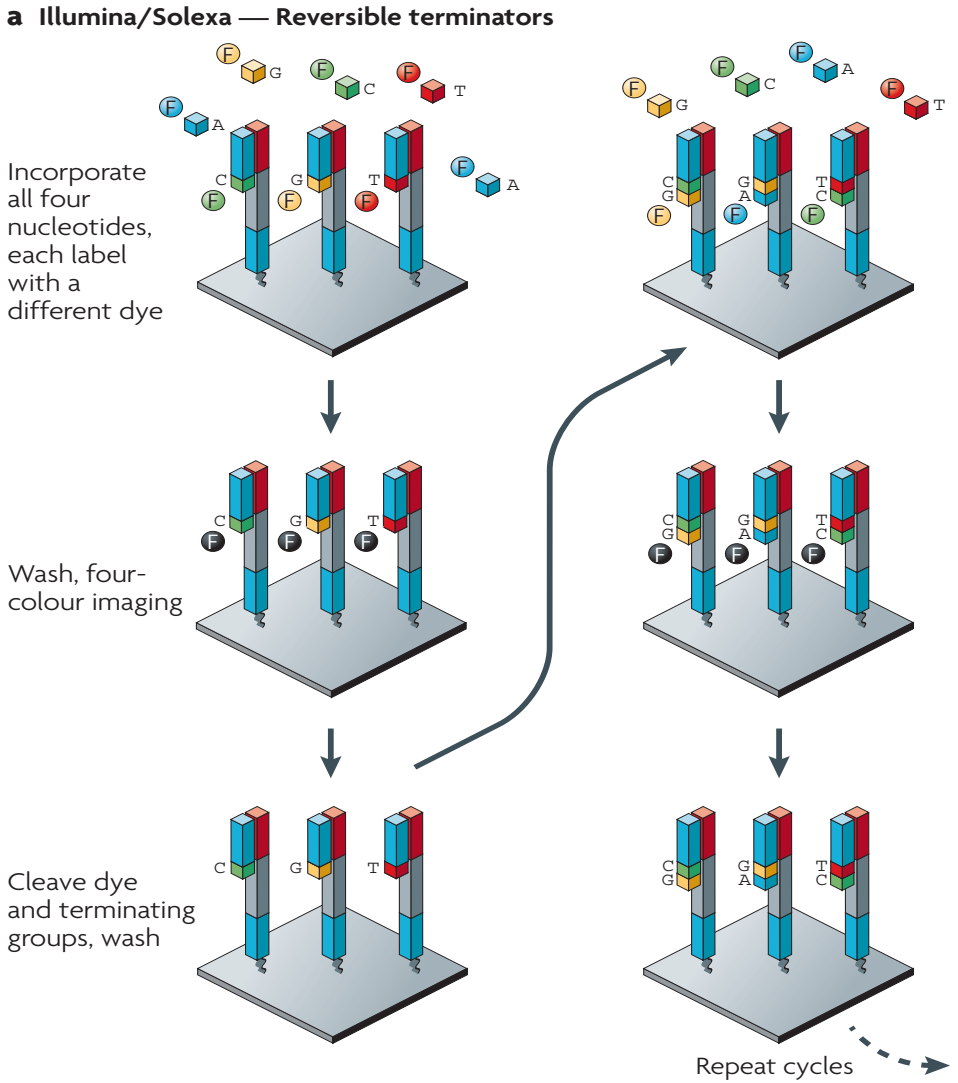


5.



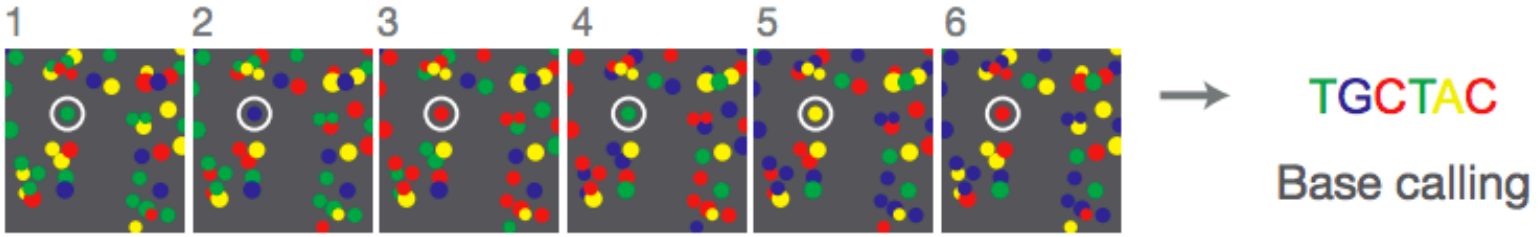
Images from www.illumina.com

Illumina sequencing – detection



Images from Metzker, M.L. (2010). Sequencing technologies - the next generation Nat. Rev. Genet. 11, 31–46.

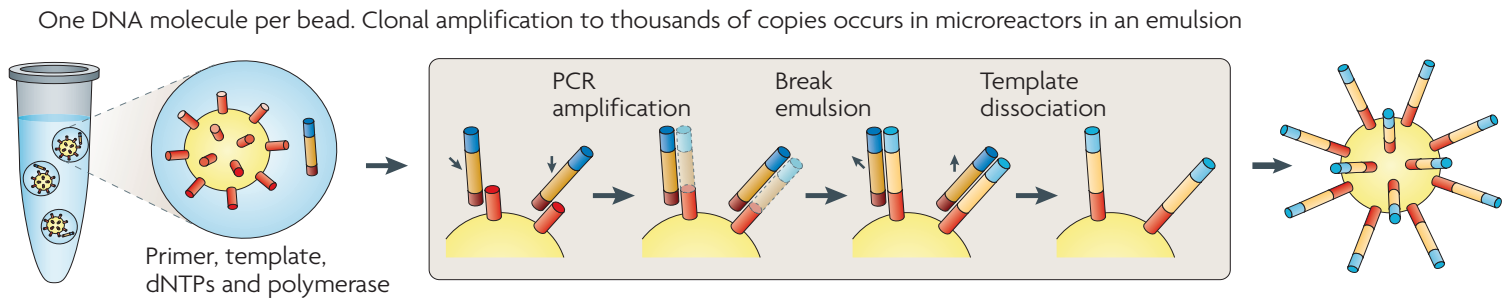
Illumina sequencing - detection



Images from illumina.com

https://www.flickr.com/photos/97397973@N05/19092162353/

Ion Torrent Sequencing – amplification



Images from Metzker, M.L. (2010). Sequencing technologies - the next generation Nat. Rev. Genet. 11, 31–46.

Ion Torrent Sequencing – detection

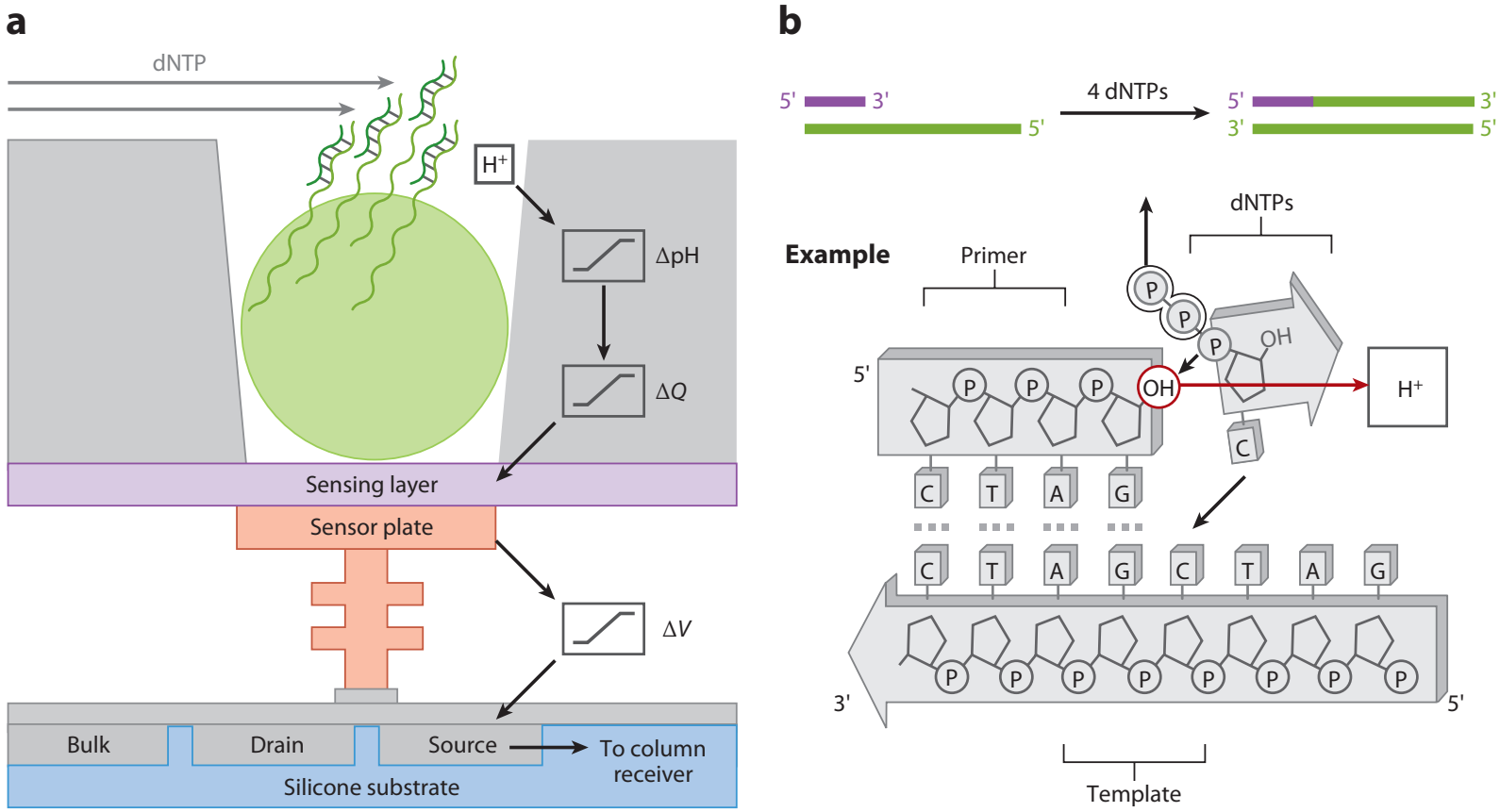


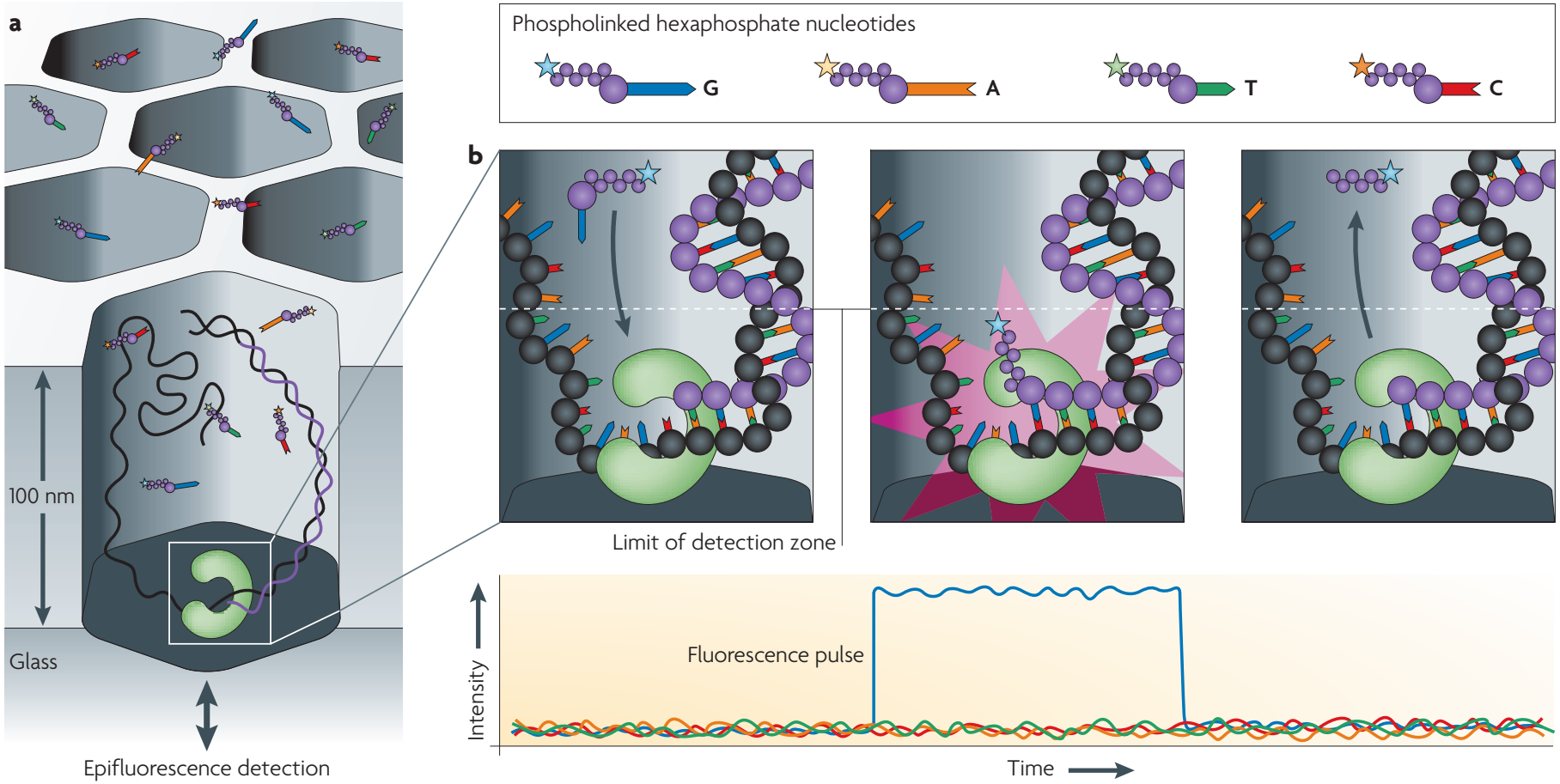
Image from Mardis, E.R. Next-Generation Sequencing Platforms. Annu. Rev. Anal. Chem. 2013. 6:287–303

https://www.flickr.com/photos/fdaphotos/8754348757/

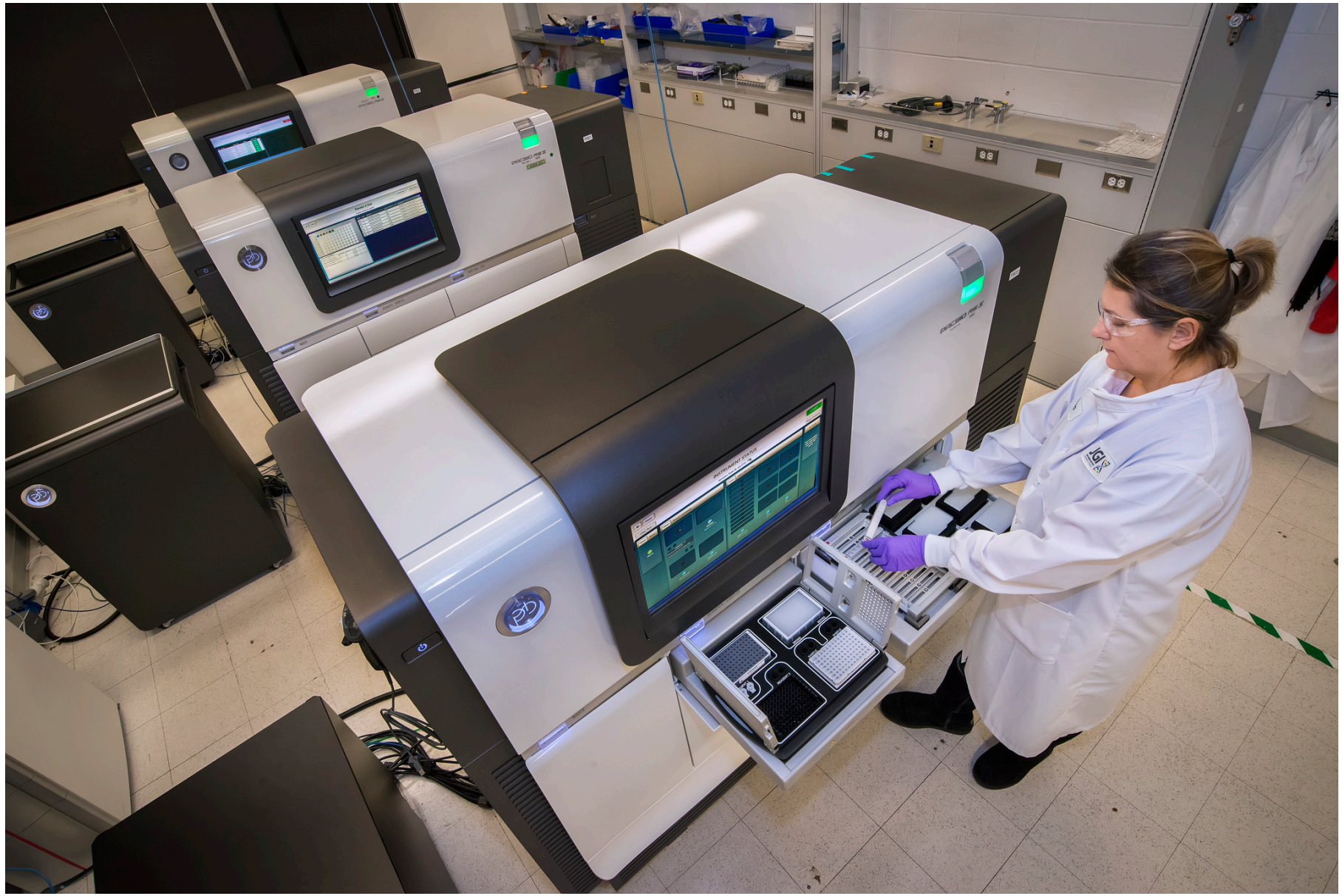
PacBio and Nanopore - single molecule sequencing

- No amplification needed
- Reduces GC bias
- Allows for longer reads
- Currently higher error rates

Pacific Biosciences (PacBio)

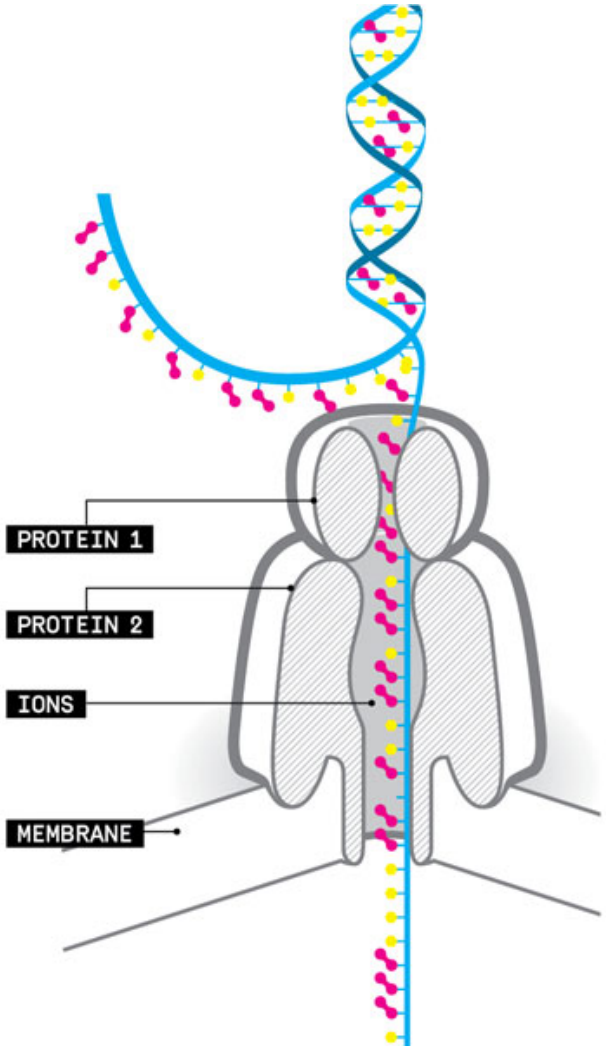


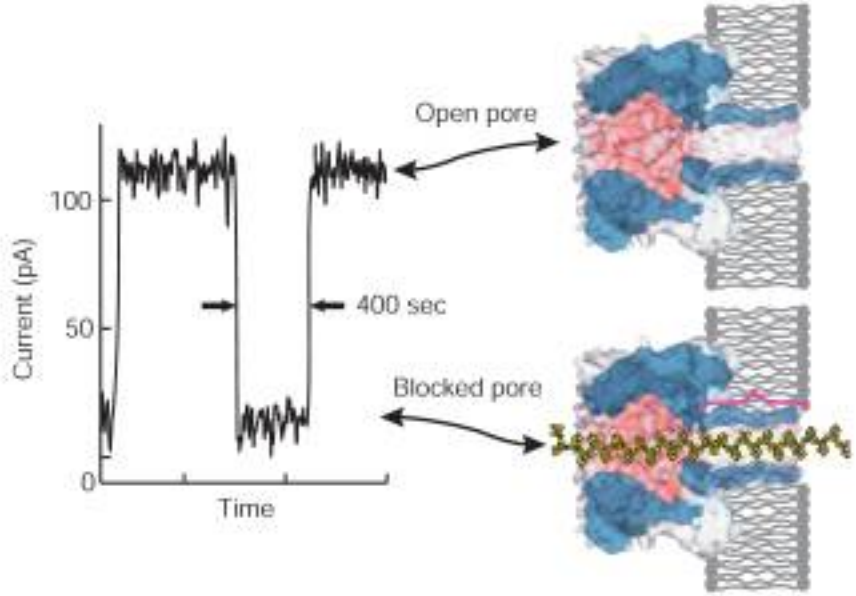
Images from Metzker, M.L. (2010). Sequencing technologies - the next generation Nat. Rev. Genet. 11, 31–46.



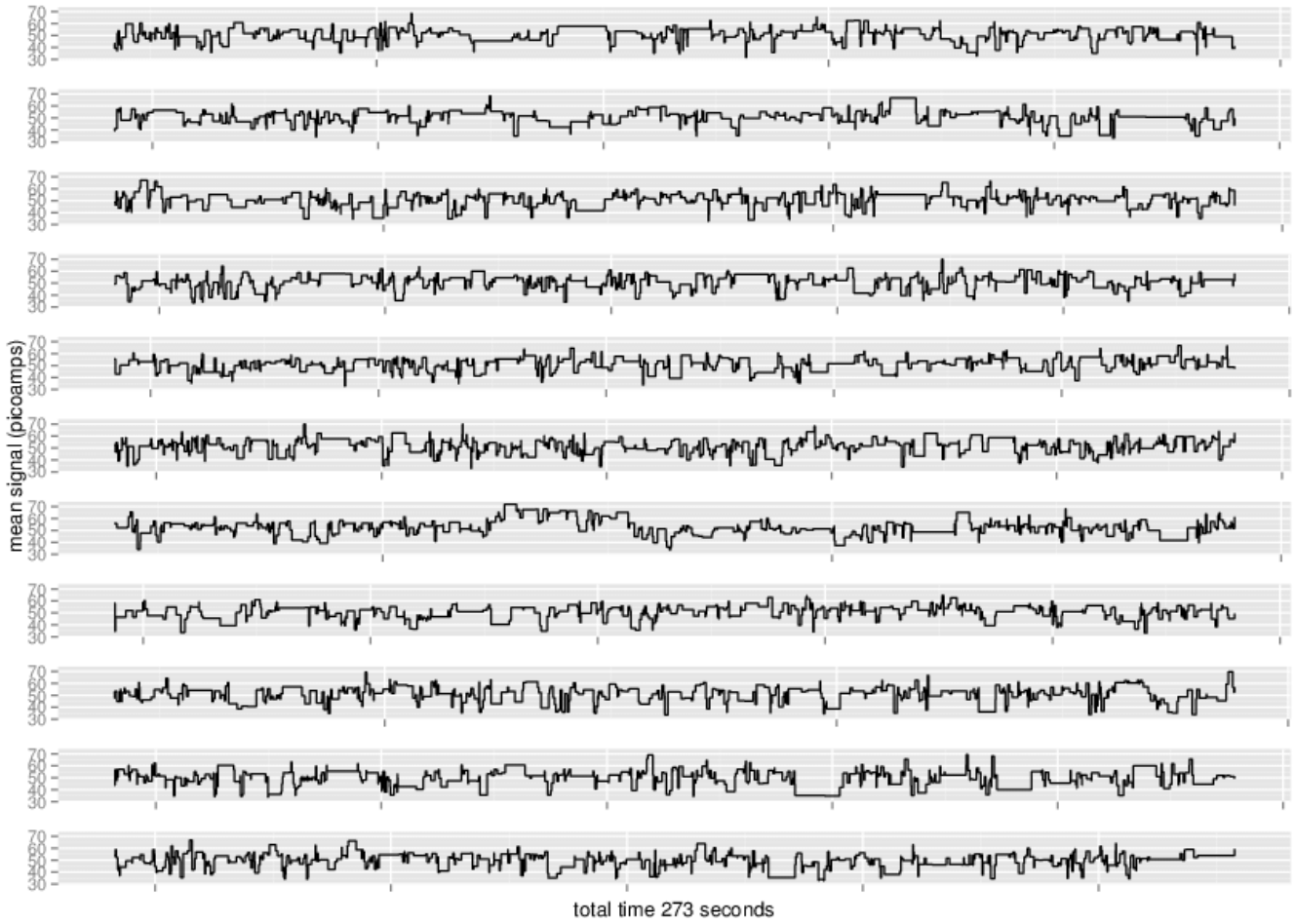
https://www.flickr.com/photos/doe_jgi/34810045553/

Nanopore Sequencing





Nanopore Sequencing - detection



Loman, Nicholas (2014): Wiggle plot showing Oxford Nanopore signal data for a P. aeruginosa read. figshare.