KSBi-BIML 2021

Bioinformatics & Machine Learning (BIML)
Workshop for Life Scientists

생물정보학 & 머쉰러닝 워크샵(온라인)

Transcriptome Profiling by Nanopore Direct RNA Sequencing

장혜식

SBi 한국생명정보학회 Korean Society for Bioinformatics
서울대학교 생물정보연구소 SNU BIOINFORMATICS INSTITUTE
Gene·2·us

Bioinformatics & Machine Learning for Life Scientists

BIML-2021

안녕하십니까?

한국생명정보학회의 동계 워크샵인 BIML-2021을 2월 15부터 2월 19일까지 개최합니다. 생명정보학 분야의 융합이론 보급과 실무역량 강화를 위해 도입한 전문 교육 프로그램인 BIML 워크샵은 2015년에 시작하였으며 올해로 7차를 맞이하게 되었습니다. 유례가 없는 코로나 대유행으로 인해 올해의 BIML 워크숍은 온라인으로 준비했습니다. 생생한 현장 강의에서만 느낄 수 있는 강의자와 수강생 사이의 상호교감을 가질 수 없다는 단점이 있지만, 온라인 강의의 여러 장점을 살려서 최근 생명정보학에서 주목받고 있는 거의 모든 분야를 망라한 강의를 준비했습니다. 또한 온라인 강의의 한계를 극복하기 위해서 실시간 Q&A 세션 또한 마련했습니다.

BIML 워크샵은 전통적으로 크게 생명정보학과 AI, 두 개의 분야로 구성되어오고 있으며 올해 역시 유사한 방식을 채택했습니다. AI 분야는 Probabilistic Modeling, Dimensionality Reduction, SVM 등과 같은 전통적인 Machine Learning부터 Deep Learning을 이용한 신약개발 및 유전체 연구까지 다양한 내용을 다루고 있습니다. 생명정보학 분야로는, Proteomics, Chemoinformatics, Single Cell Genomics, Cancer Genomics, Network Biology, 3D Epigenomics, RNA Biology, Microbiome 등 거의 모든 분야가 포함되어 있습니다. 연사들은 각 분야 최고의 전문가들이라 자부합니다.

이번 BIML-2021을 준비하기까지 너무나 많은 수고를 해주신 BIML-2021 운영위원회의 김태민 교수님, 류성호 교수님, 남진우 교수님, 백대현 교수님께 커다란 감사를 드립니다. 또한 재정적 도움을 주신, 김선 교수님 (AI-based Drug Discovery), 류성호 교수님, 남진우 교수님께 감사를 표시하고 싶습니다. 마지막으로 부족한 시간에도 불구하고 강의 부탁을 흔쾌히 허락하시고 훌륭한 강의자료를 만드는데 노력하셨을 뿐만 아니라 실시간 온라인 Q&A 세션까지 참여해 수고해 주시는 모든 연사분들께 깊이 감사드립니다.

2021년 2월

한국생명정보학회장 김동섭

강의개요

Transcriptome Profiling by Nanopore Direct RNA Sequencing

짧은 리드를 이용한 RNA-seq은 RNA 발현을 살펴볼때 저렴한 가격에 높은 해상도를 얻을 수 있어 널리 활용되고 있다. 다만, 상보DNA로 변환하는 과정에서 RNA수식 정보가 사라지고, 짧은 리드로 인해 멀리 떨어진 스플라이싱 등의 정보를 얻는 데 한계가 있다. 나노포어 시퀀싱은 DNA 중합효소를 전혀 사용하지 않는 새로운 RNA 시퀀싱 기법을 제공한다. 직접 RNA 시퀀싱(Direct RNA Sequencing; DRS)은 실험기법이 단순하고 결과가 수시간 안에 빠르게 나올 뿐만 아니라, 긴 리드에서 RNA 수식을 살펴볼 수 있는 기회를 제공하고 있어 기존 방법으로 해결 불가능했던 문제들에 활용될 수 있는 잠재력이 있다.

본 강의에서는 나노포어 시퀀싱과 DRS이 나오게 된 과정, 작동원리, 데이터 특성, 사용상 주의점들을 익히고, 전사체 분석에 활용할 수 있는 방법들을 소개한다. 아직 표준화된 도구들이 등장하지 않고 다양한 다른 접근법과 프로그램이 혼용되어 사용되는 DRS의 특성을 고려해서 응용 방법들의 사용되고 있는 프로그램들이 기반으로 하고 있는 생화학적, 알고리즘적 원리들에 대해 간략히 설명한다. 이를 통해, DRS를 연구에 활용하고자 할 때 당시에 새로 나오게 될 도구들의 특성을 쉽게 파악하고 익힐 수 있도록 하는 것을 목표로 한다.

강의는 다음의 내용을 포함한다:

- 나노포어 시퀀싱의 원리와 역사
- 나노포어 시퀀싱 실험과 데이터의 특성
- 직접 RNA 시퀀싱의 활용 용도별 주요 도구 및 그 원리
- 나노포어 시퀀싱 신호 처리법 개요

*** 교육생준비물:**

Zoom 수강 가능 노트북, 인터넷

*** 강의: 장혜식 교수 (서울대학교 생명과학부)**

Curriculum Vitae

Speaker Name: Hyeshik Chang, Ph.D.

▶Personal Info

Name	Hyeshik Chang
Title	Assistant Professor
Affiliation	Seoul National University

▶Contact Information

Address: 1 Gwanak-ro Gwanak-gu, Seoul, 08826

Email: hyeshik@snu.ac.kr

Website: https://qbio.io

Research interest: High-throughput sequencing, post-transcriptional regulation, RNA-protein interaction

Educational Experience

1998–2007	B.S.E. in Information and Industrial Engineering, Yonsei University, Korea
2007–2009	M.S.E. in Bio and Brain Engineering, KAIST, Korea
2009–2014	Ph.D. in Biological Sciences, Seoul National University, Korea

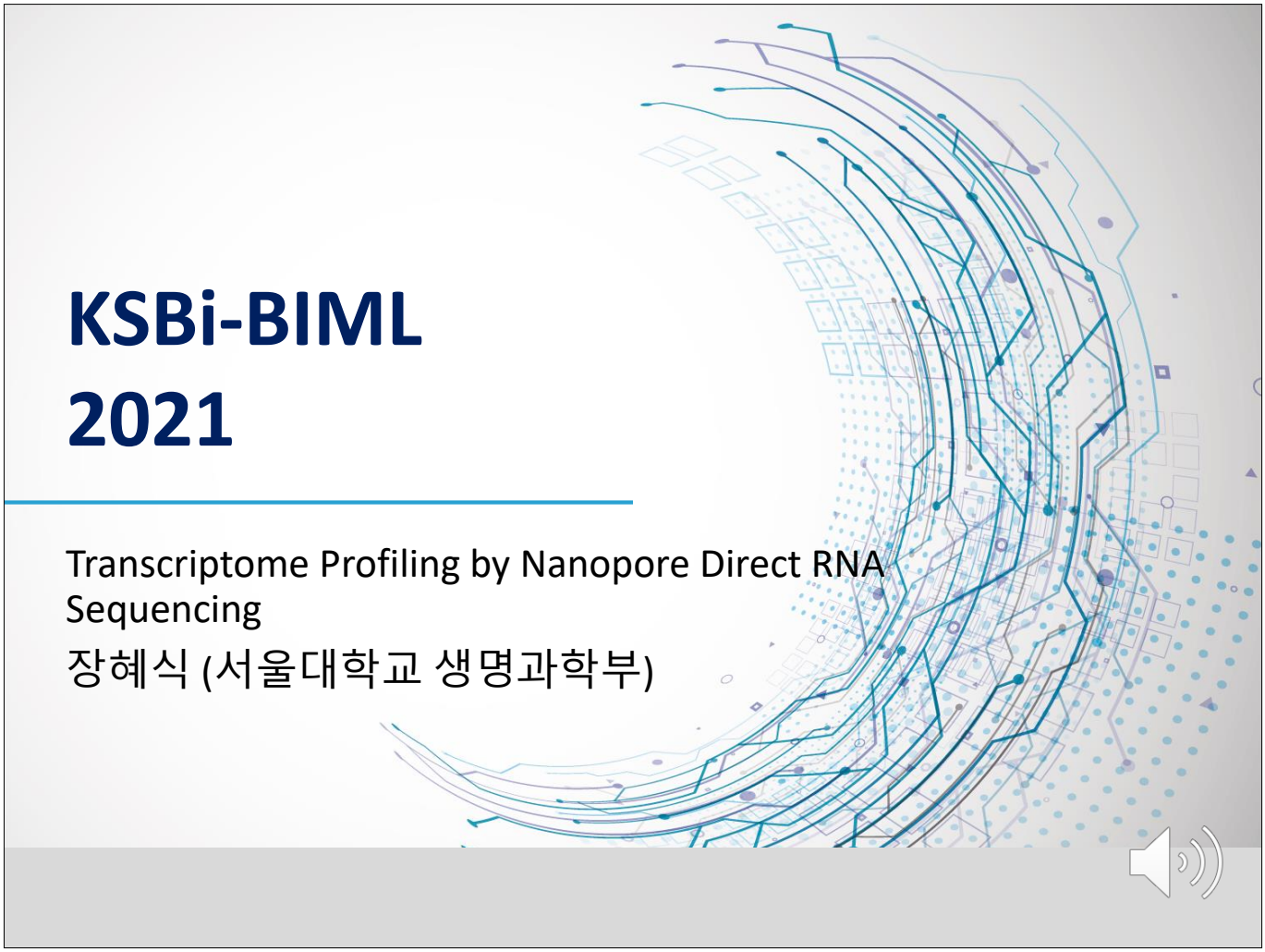
Professional Experience

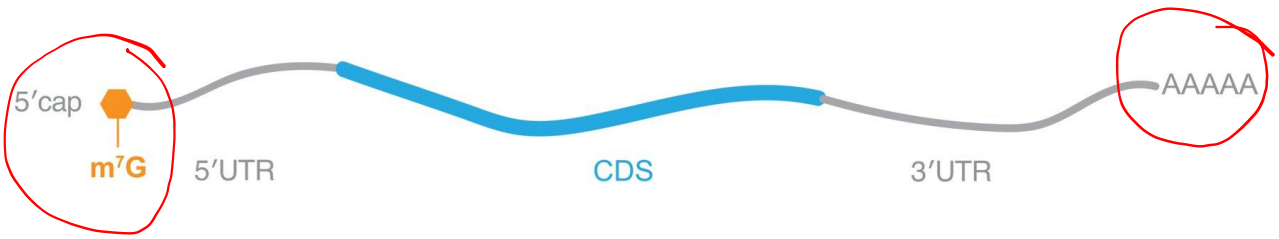
2001–2005	Software Developer, Solution Development Team, LinuxKorea, Inc.
2014–2019	Research Assistant Professor, IBS Center for RNA Research, Seoul National University
2018–	Research Fellow, Center for RNA Research, Institute for Basic Science
2019–	Assistant Professor, School of Biological Sciences, Seoul National University

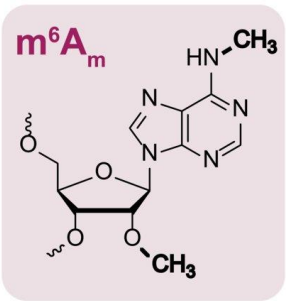
Selected Publications (5 maximum)

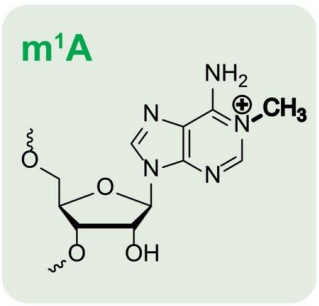
1. D. Kim, J.-Y. Lee, J.-S. Yang, J. W. Kim, V. N. Kim, and H. Chang. (2020) "The Architecture of SARS-CoV-2 Transcriptome." Cell, 181(4):914–921.
2. H. Chang[1], J. Yeo[1], J.-G. Kim, H. Kim, M. Lee, J. Lim, H. H. Kim, J. Ohk, H.-Y. Jeon, H. Lee, H. Jung, K.-W. Kim, and V. N. Kim. (2018) "Terminal uridylyltransferases execute programmed clearance of maternal transcriptome in vertebrate embryos." Molecular Cell, 70:72–82.e7.
3. J. Lim[1], M. Ha[1], H. Chang[1], S. C. Kwon, D. K. Simanshu, D. J. Patel, and V. N. Kim. (2014) "Uridylation by TUT4 and TUT7 marks mRNA for degradation." Cell, 159(6):1365–1376.
4. H. Chang[1], J. Lim[1], M. Ha, and V. N. Kim. (2014) "TAIL-seq: genome-wide determination of poly(A) tail length and 3′ end modifications." Molecular Cell, 53(6):1044–1052.
5. J. Cho[1], H. Chang[1], S. C. Kwon, B. Kim, Y. Kim, J. Choe, M. Ha, Y. K. Kim, and V. N. Kim. (2012) "LIN28A is a suppressor of ER-associated translation in embryonic stem cells." Cell, 151(4):765–777.

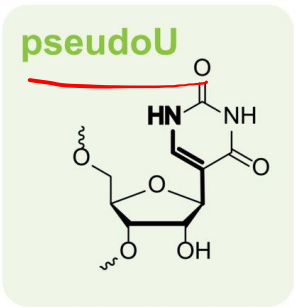
[1] Co-first authors

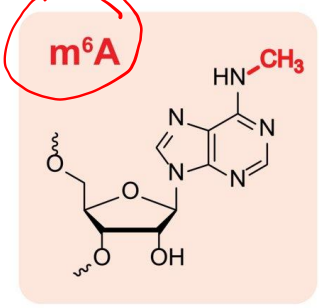


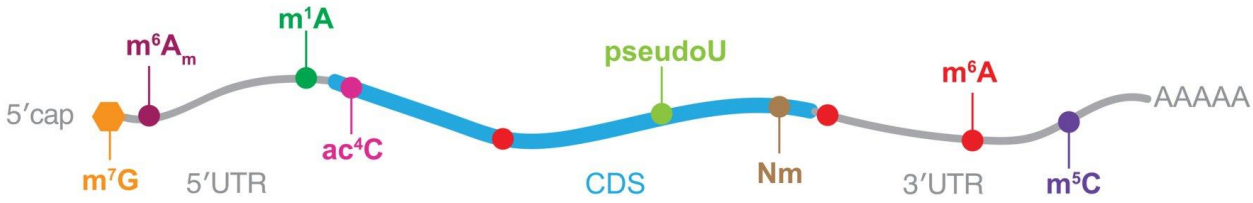


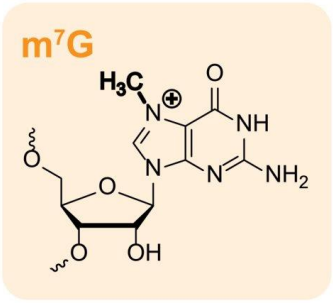


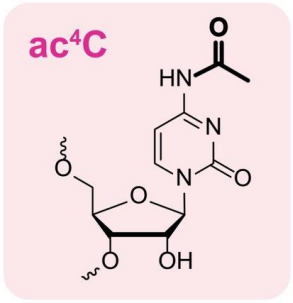


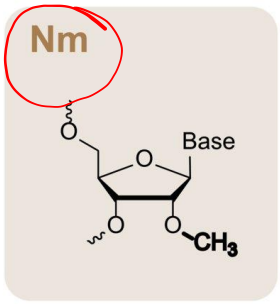


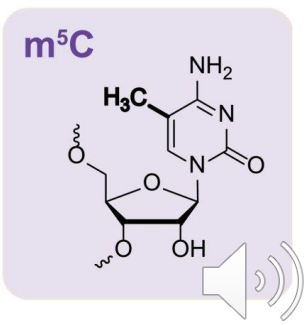


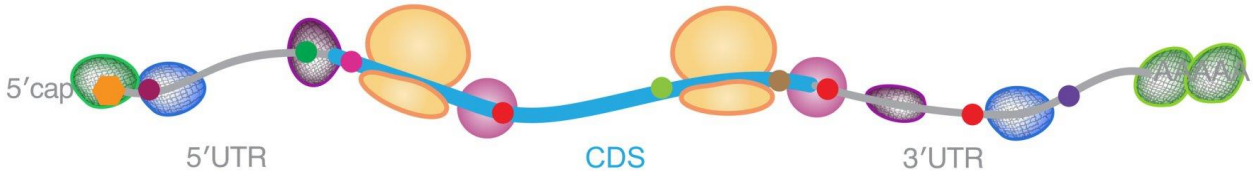










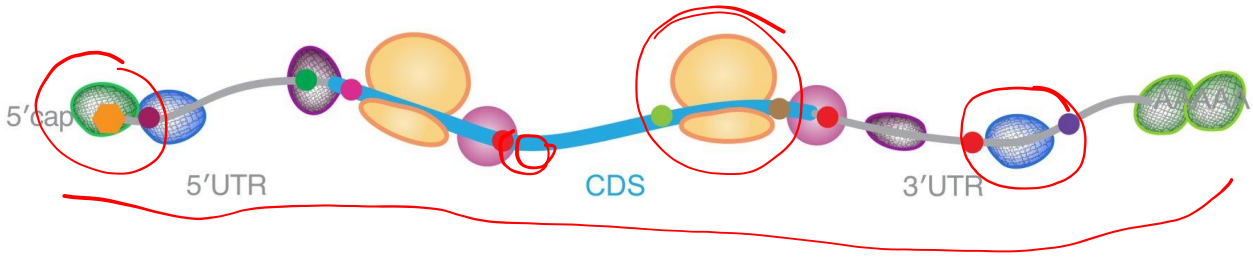


Genomic DNA	Messenger RNA
Epigenome	Epitranscriptome *(controversial)*

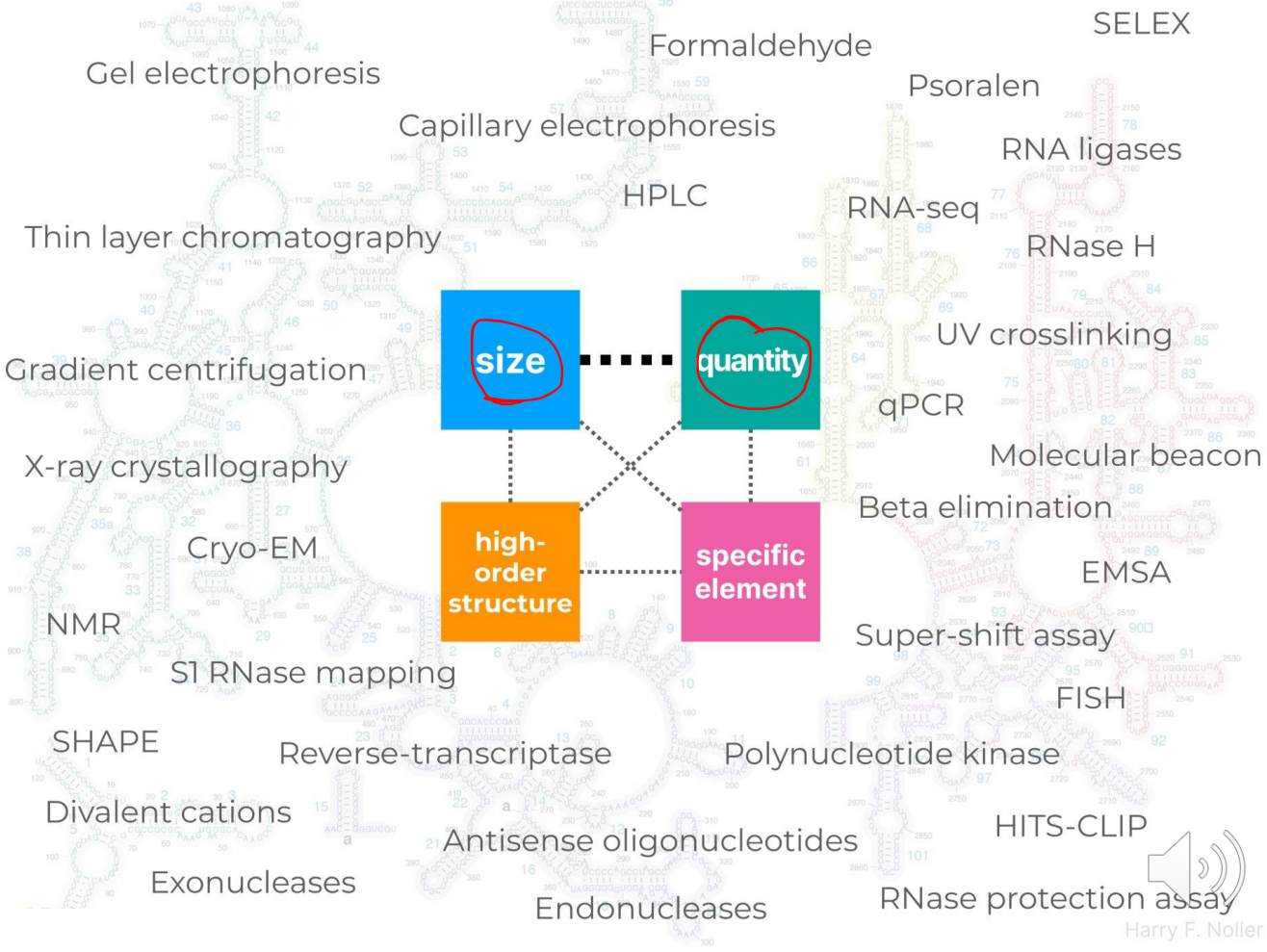
Genomic DNA	Messenger RNA
Epigenome	Epitranscriptome *(controversial)*
DNA Methylation	RNA Methylation RNA Editing

Genomic DNA	Messenger RNA
Epigenome	Epitranscriptome *(controversial)*
DNA Methylation	RNA Methylation RNA Editing
Histone Modification	Combinations of RBP Binding

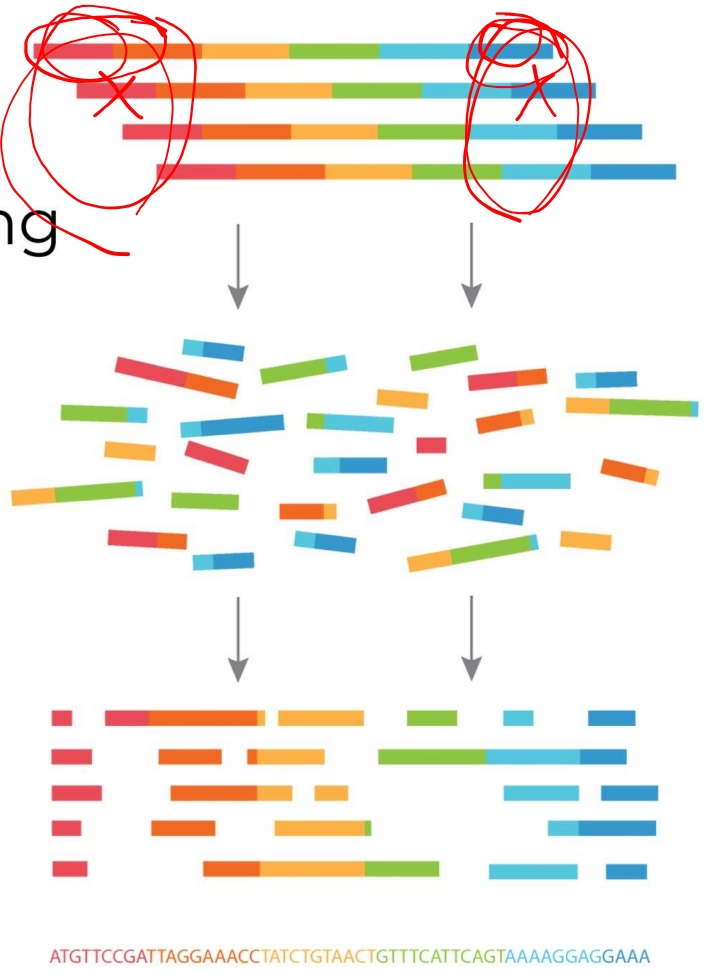
Genomic DNA	Messenger RNA
Epigenome	Epitranscriptome *(controversial)*
DNA Methylation	RNA Methylation RNA Editing
Histone Modification	Combinations of RBP Binding
Controls Transcription	Controls Splicing, Localization, Translation, and mRNA degradation


5′cap
5′UTR
CDS
3′UTR

© Hailing Shi

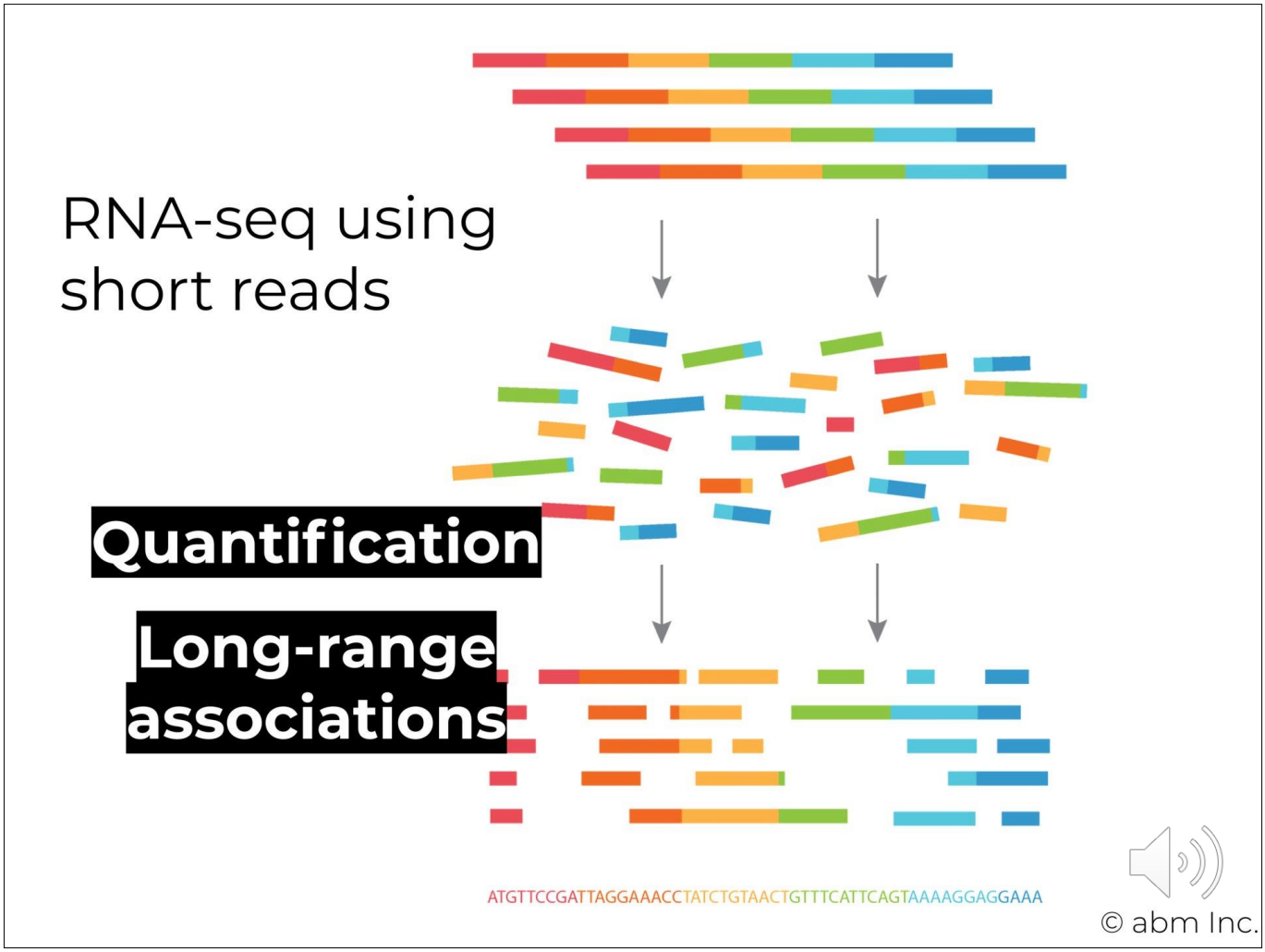

SELEX
Formaldehyde
Gel electrophoresis
Psoralen
Capillary electrophoresis
RNA ligases
HPLC
RNA-seq
Thin layer chromatography
RNase H
size
quantity
UV crosslinking
Gradient centrifugation
qPCR
Molecular beacon
X-ray crystallography
Beta elimination
high-order structure
specific element
Cryo-EM
EMSA
NMR
Super-shift assay
S1 RNase mapping
FISH
SHAPE
Reverse-transcriptase
Polynucleotide kinase
Divalent cations
HITS-CLIP
Antisense oligonucleotides
Exonucleases
RNase protection assay
Endonucleases
Harry F. Noller

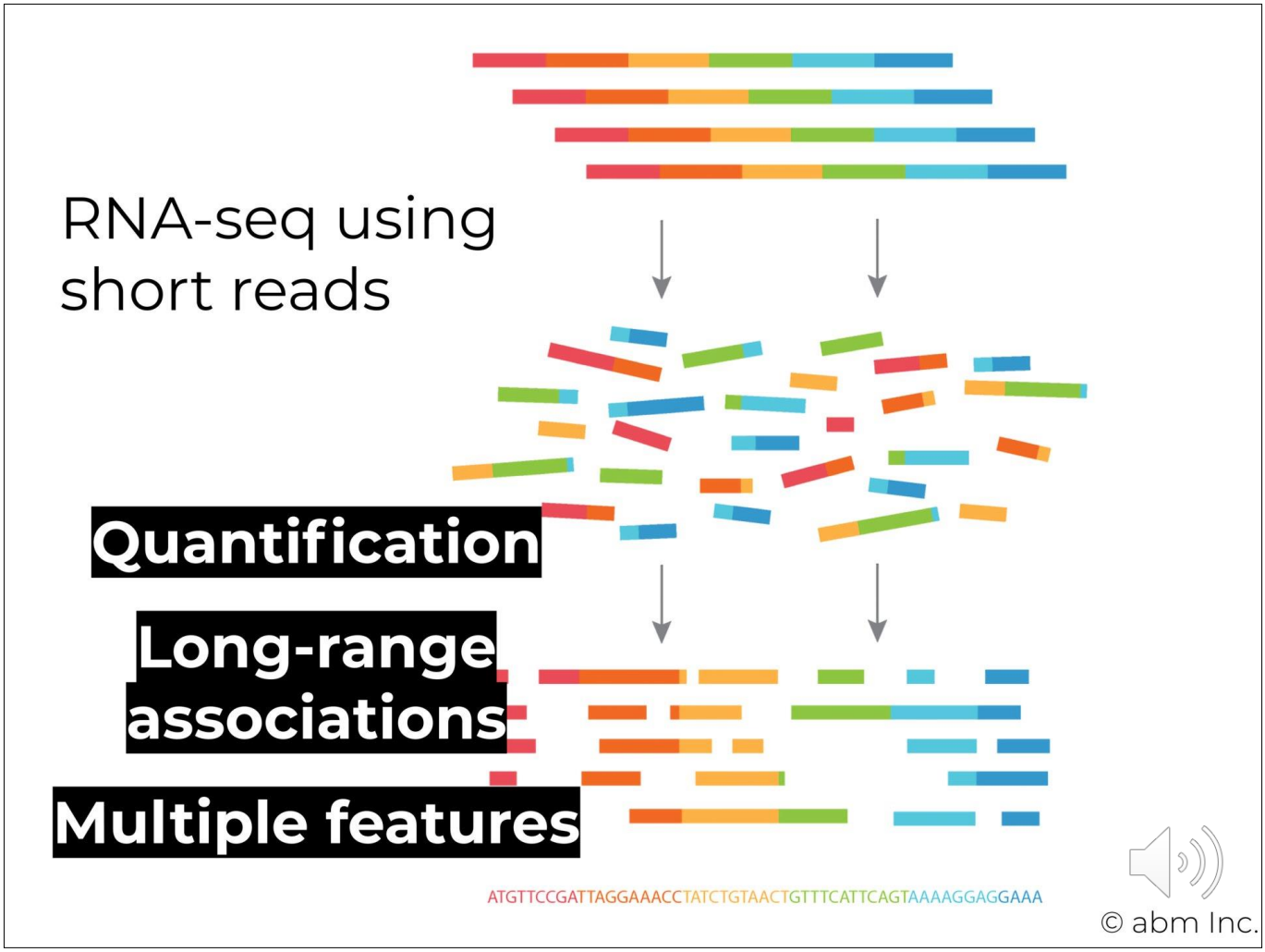
RNA-seq using short reads



RNA-seq using short reads

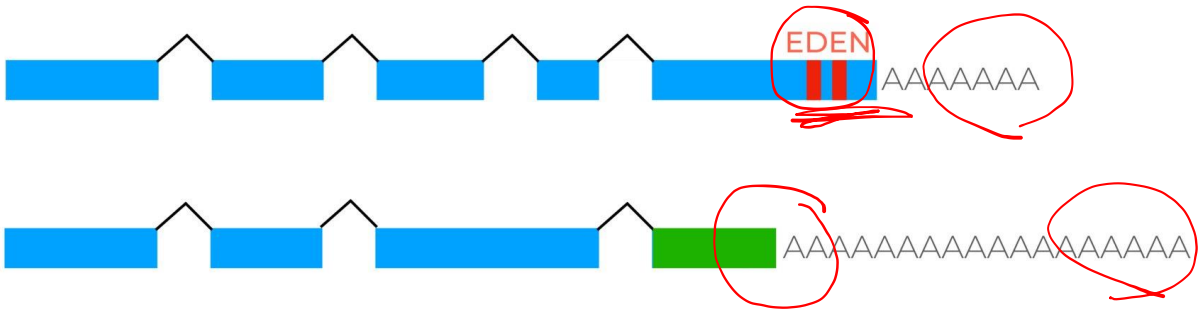
Quantification

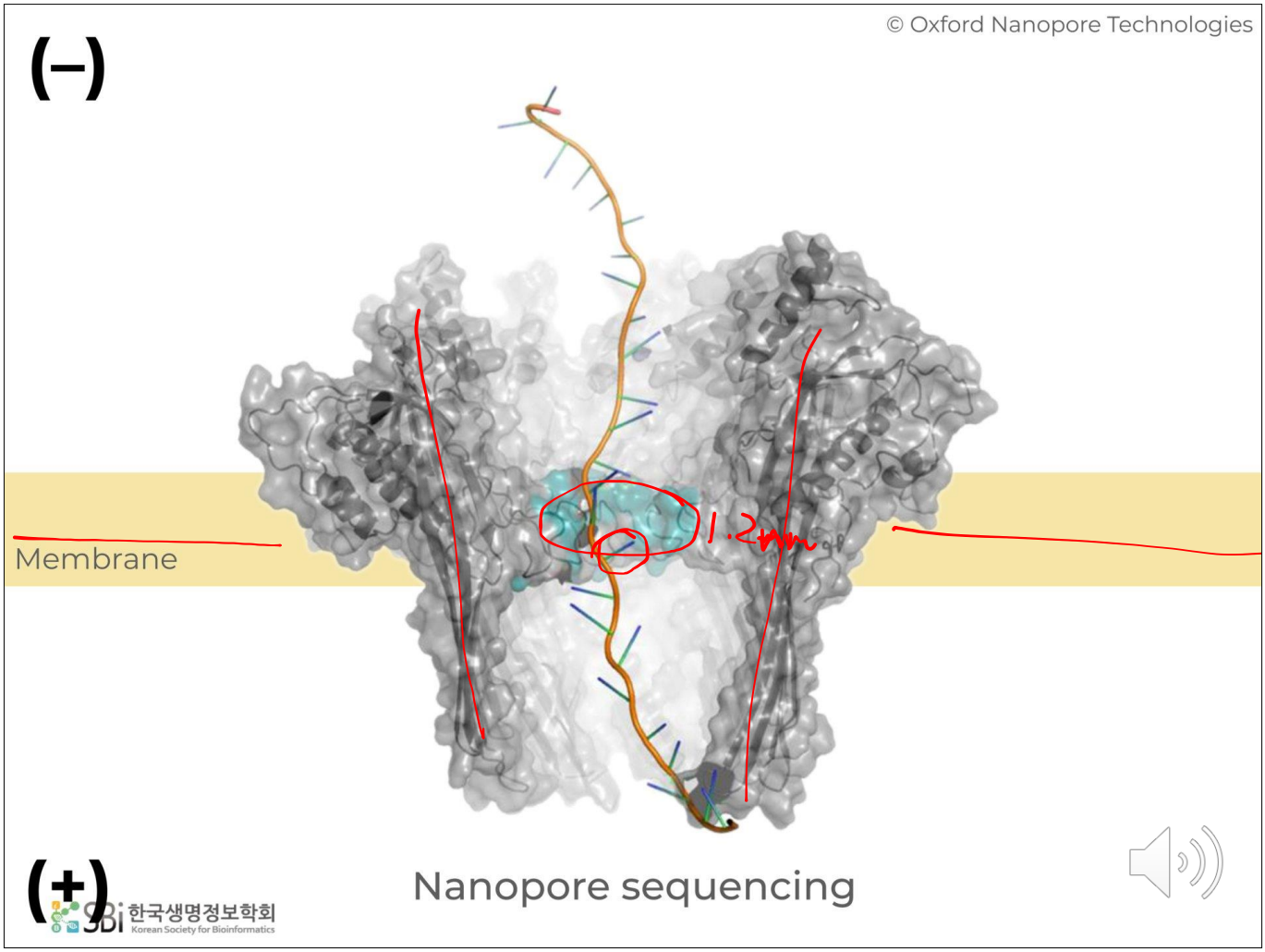

RNA-seq using short reads
Quantification
Long-range associations
ATGTTCCGATTAGGAAACCTATCTGTAACTGTTTCATTCAGTAAAAGGAGGAAA
© abm Inc.


RNA-seq using short reads
Quantification
Long-range associations
Multiple features
ATGTTCCGATTAGGAAACCTATCTGTAACTGTTTCATTCAGTAAAAGGAGGAAA
© abm Inc.

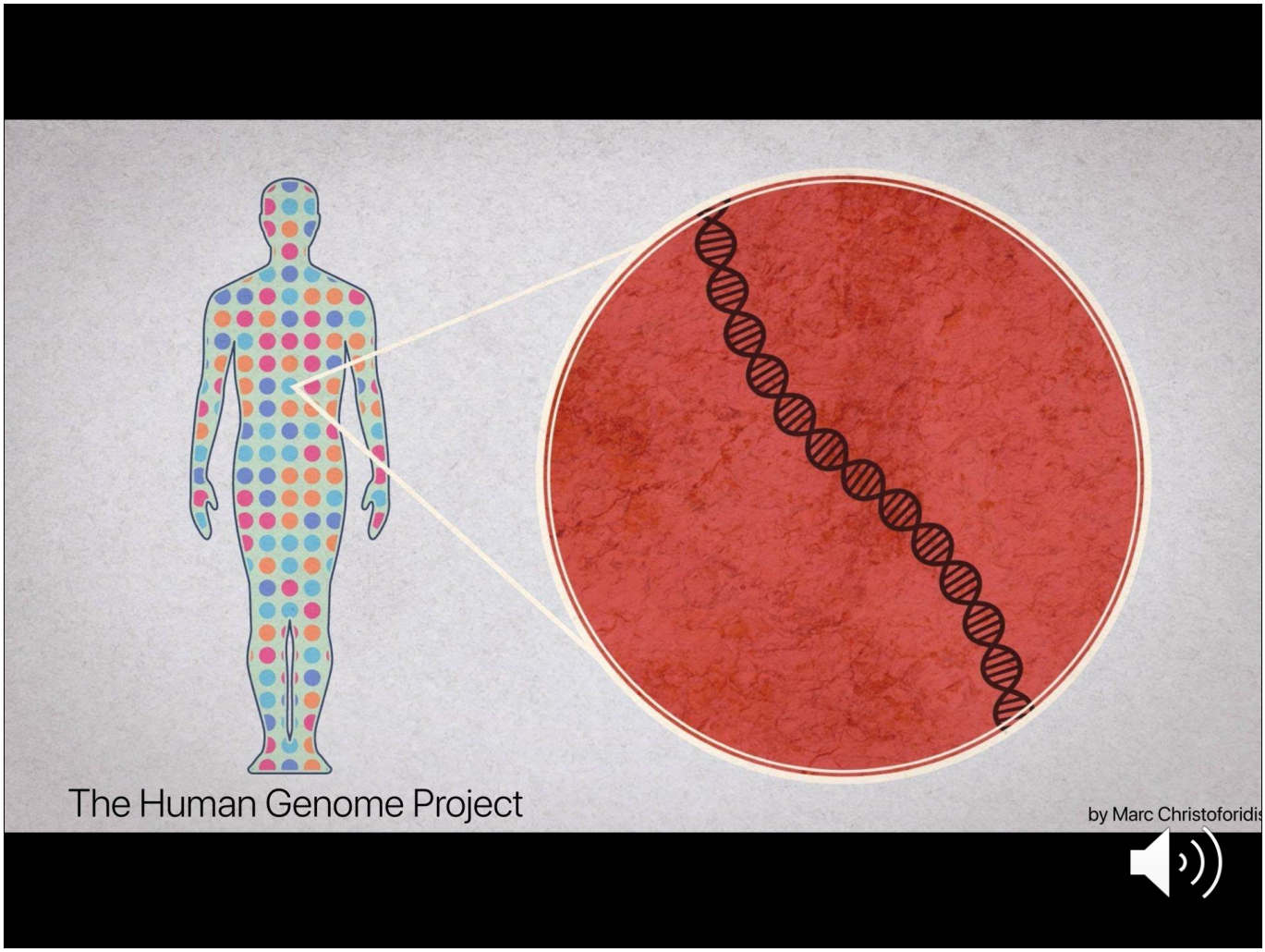
Does the **m6A** modifications near the stop codon affect the **secondary structure** of 3' UTR?

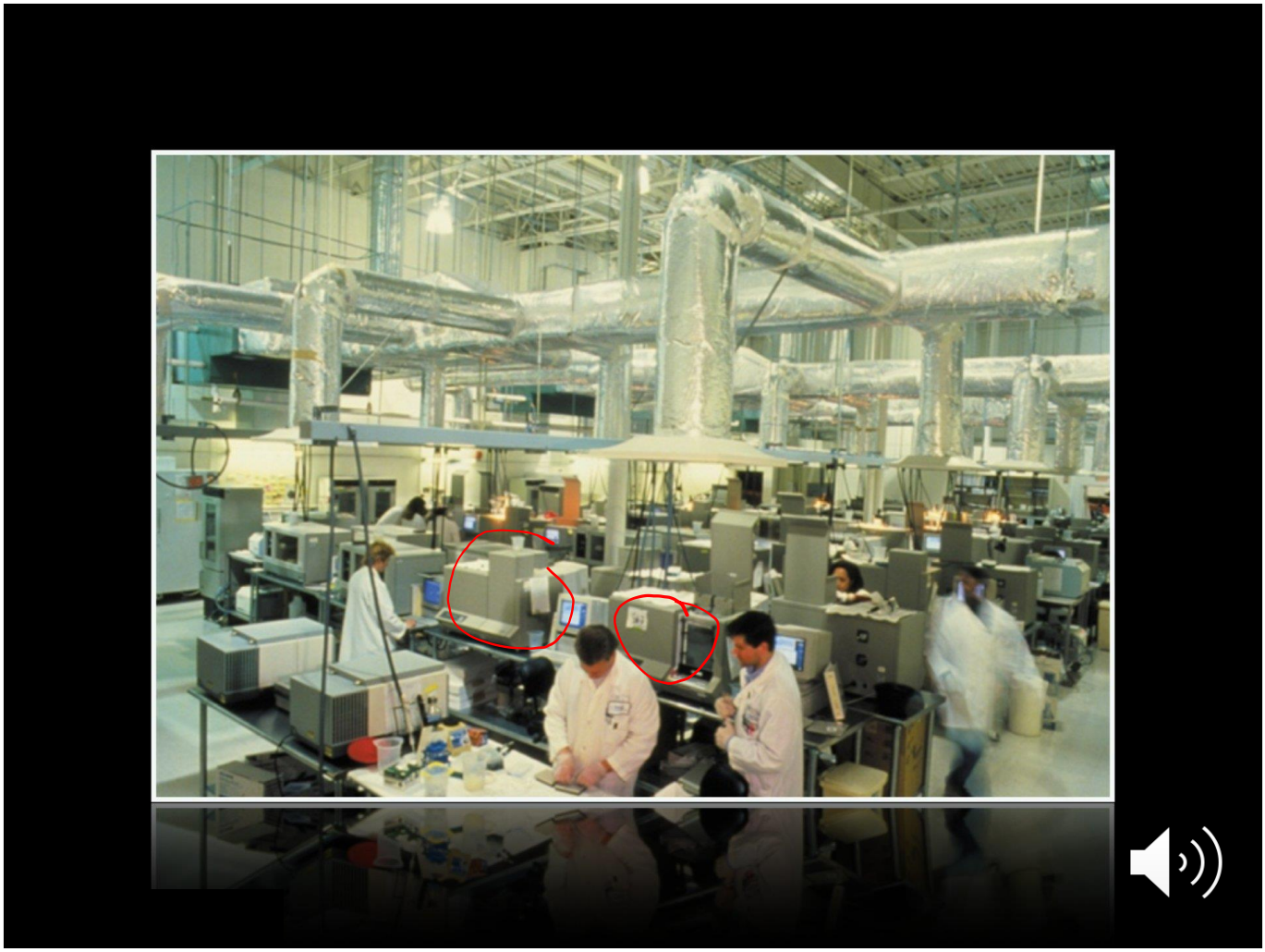
What **elements** make poly(A) lengths different for transcripts of the same gene?

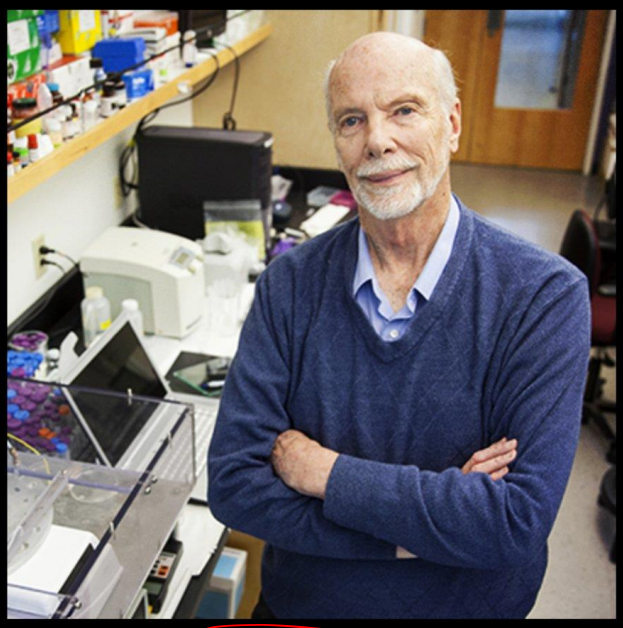




The birth of nanopore sequencing

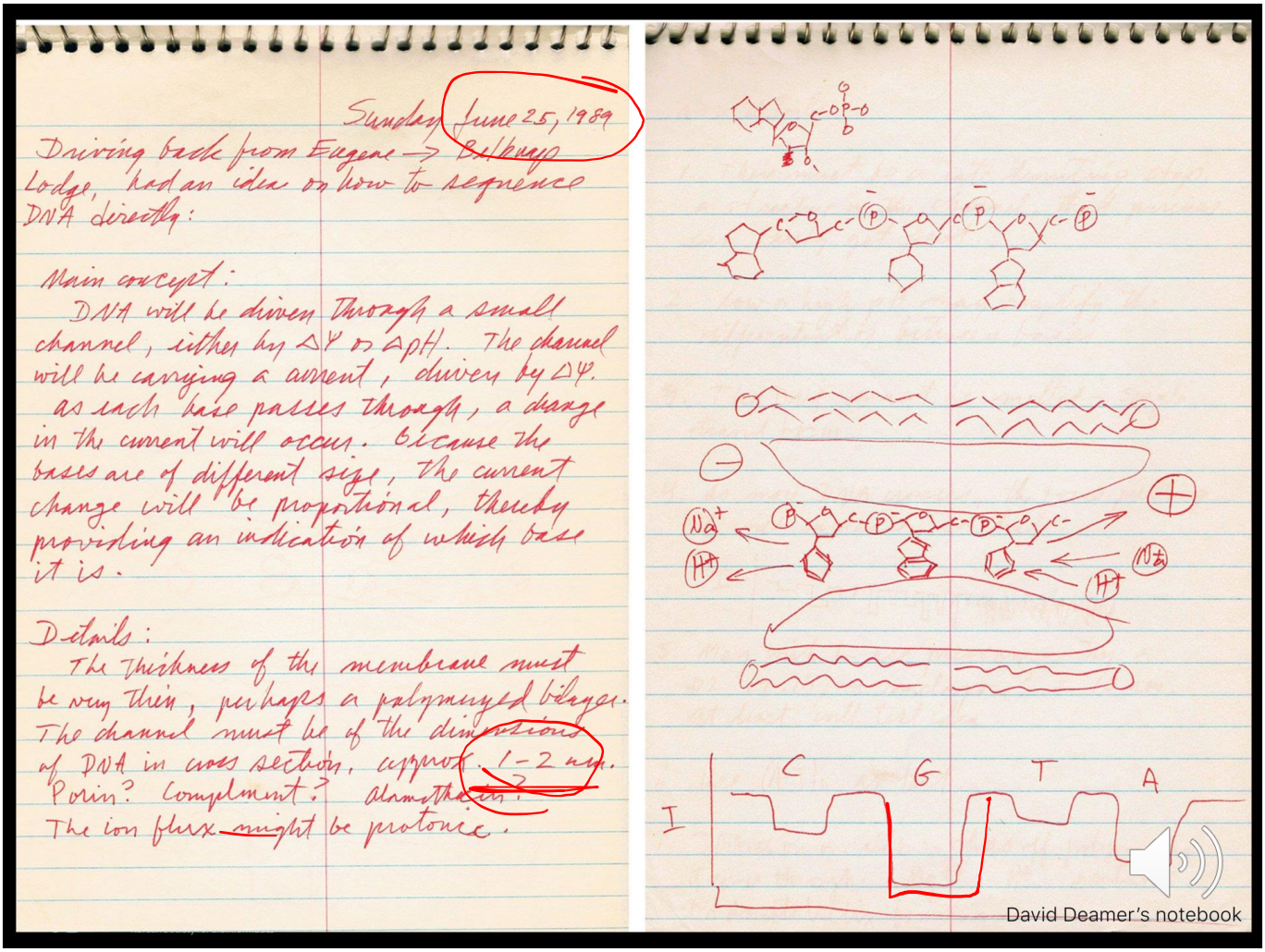

The Human Genome Project
by Marc Christoforidis

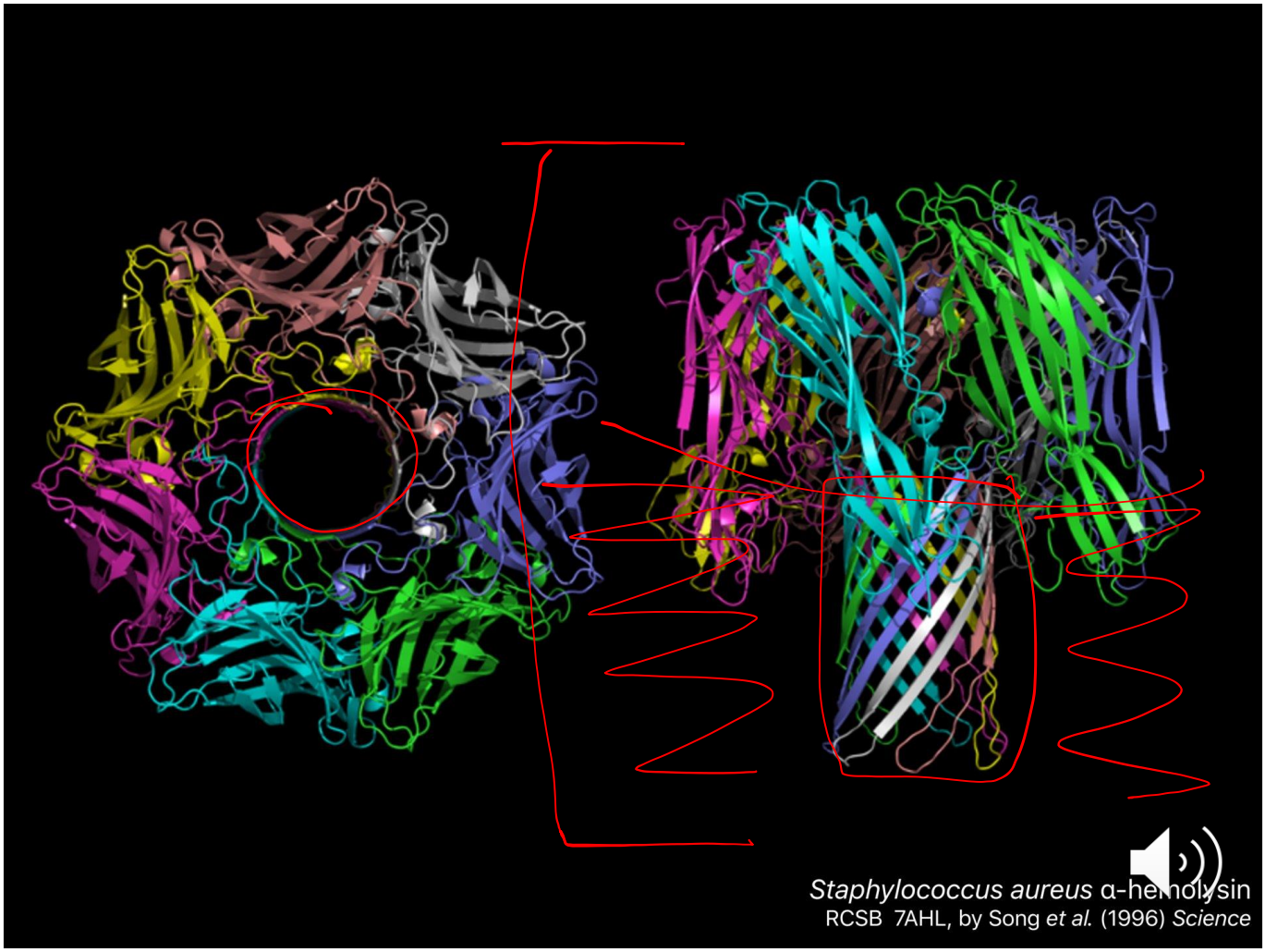


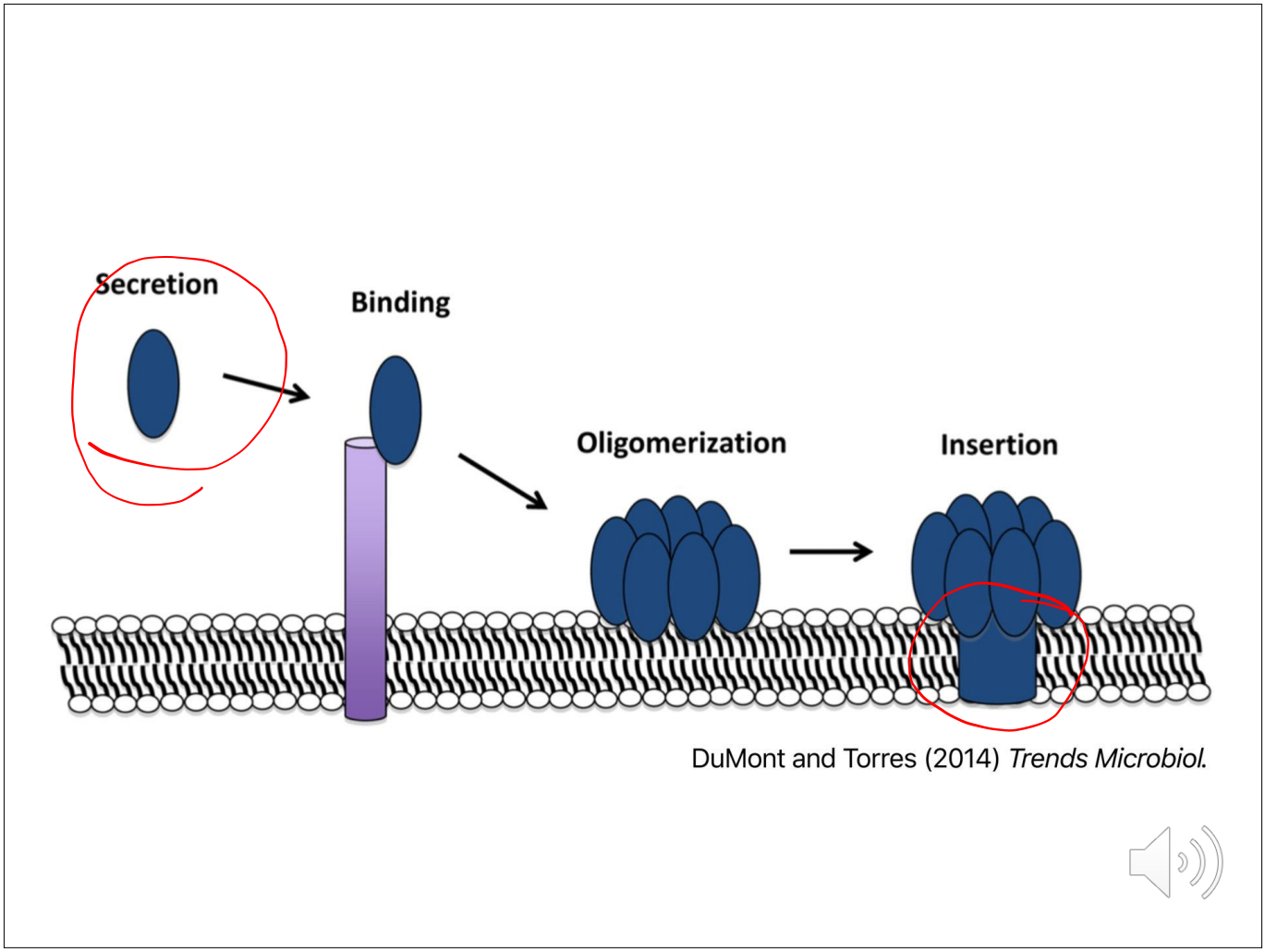


David Deamer

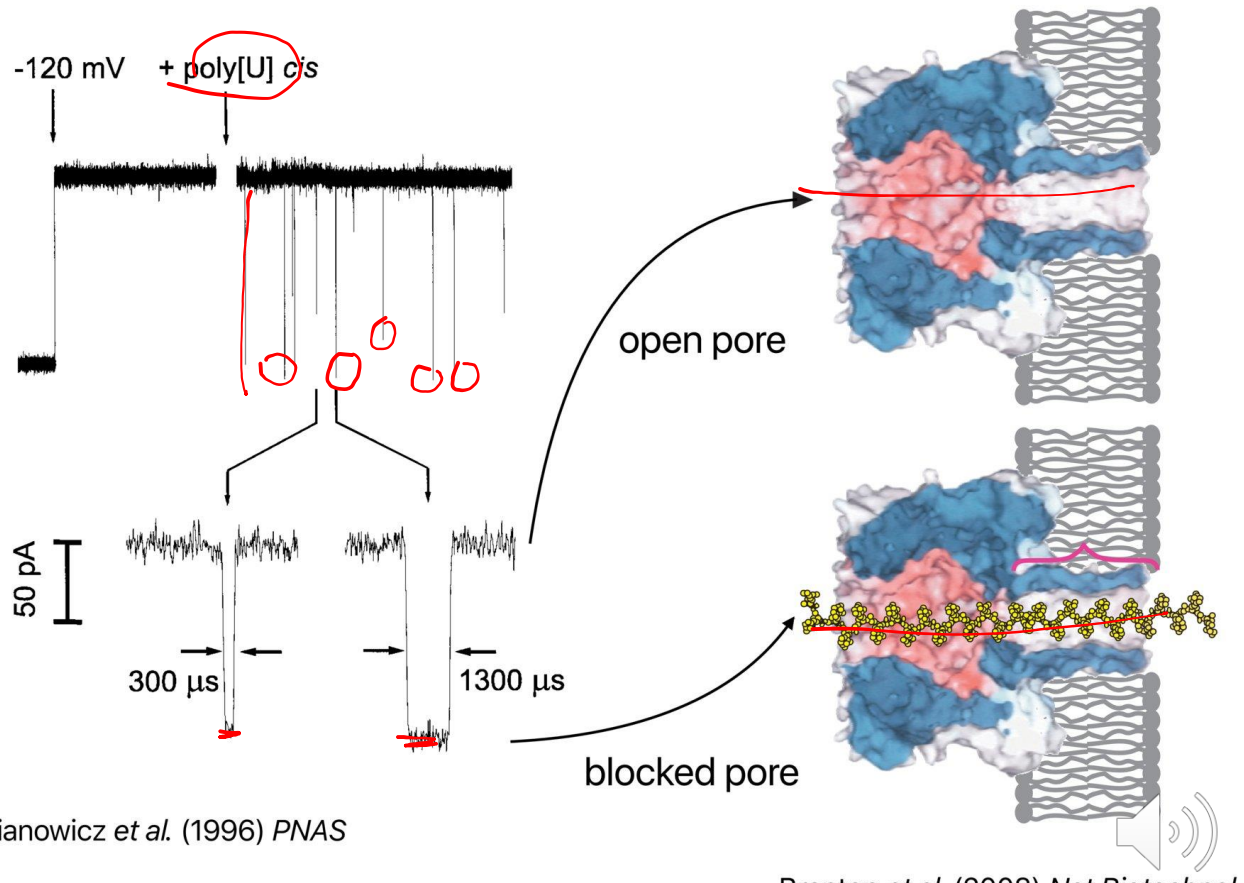
George Church




Staphylococcus aureus α-hemolysin
RCSB 7AHL, by Song *et al.* (1996) *Science*

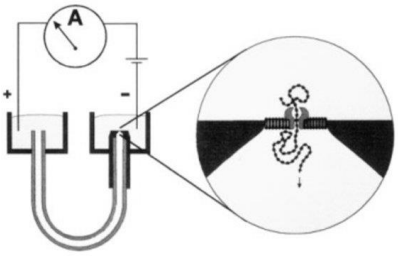

Secretion
Binding
Oligomerization
Insertion
DuMont and Torres (2014) *Trends Microbiol.*

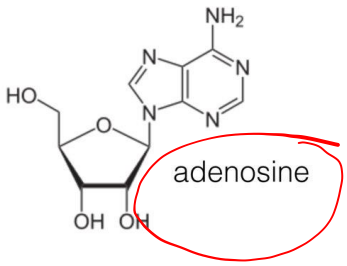
What happens when a DNA passes through a nanopore?

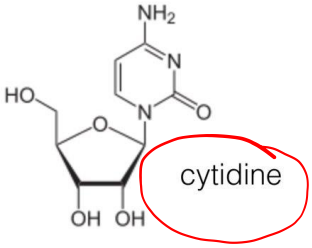


Kasianowicz *et al.* (1996) *PNAS*

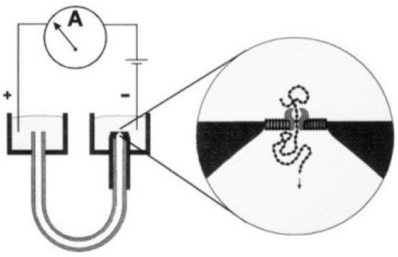
Branton *et al.* (2008) *Nat Biotechnol.*

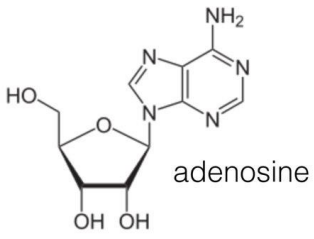


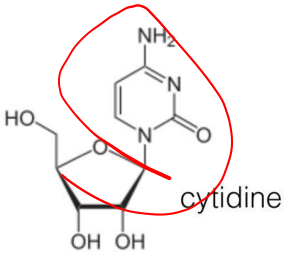


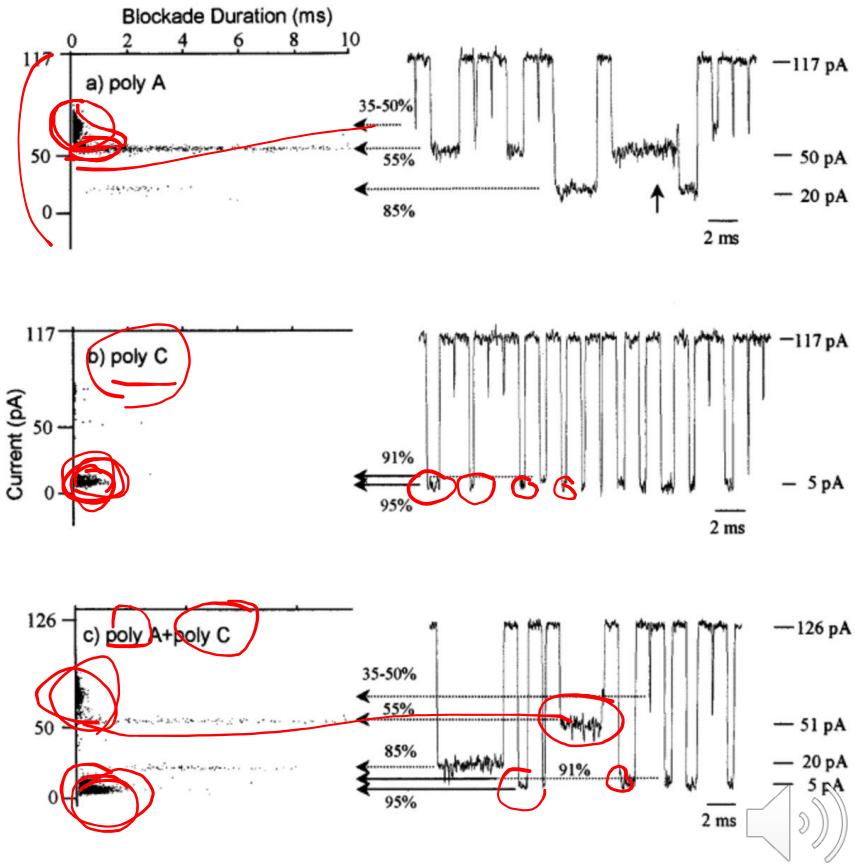


Akeson *et al.* (1999) *Biophys J*

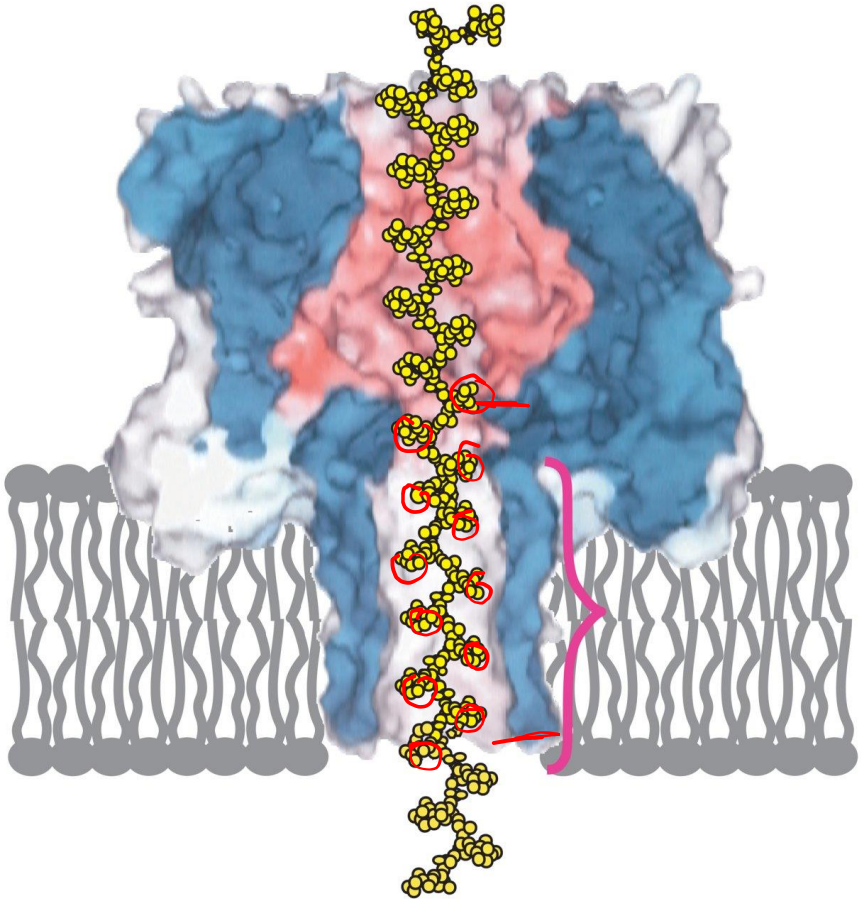






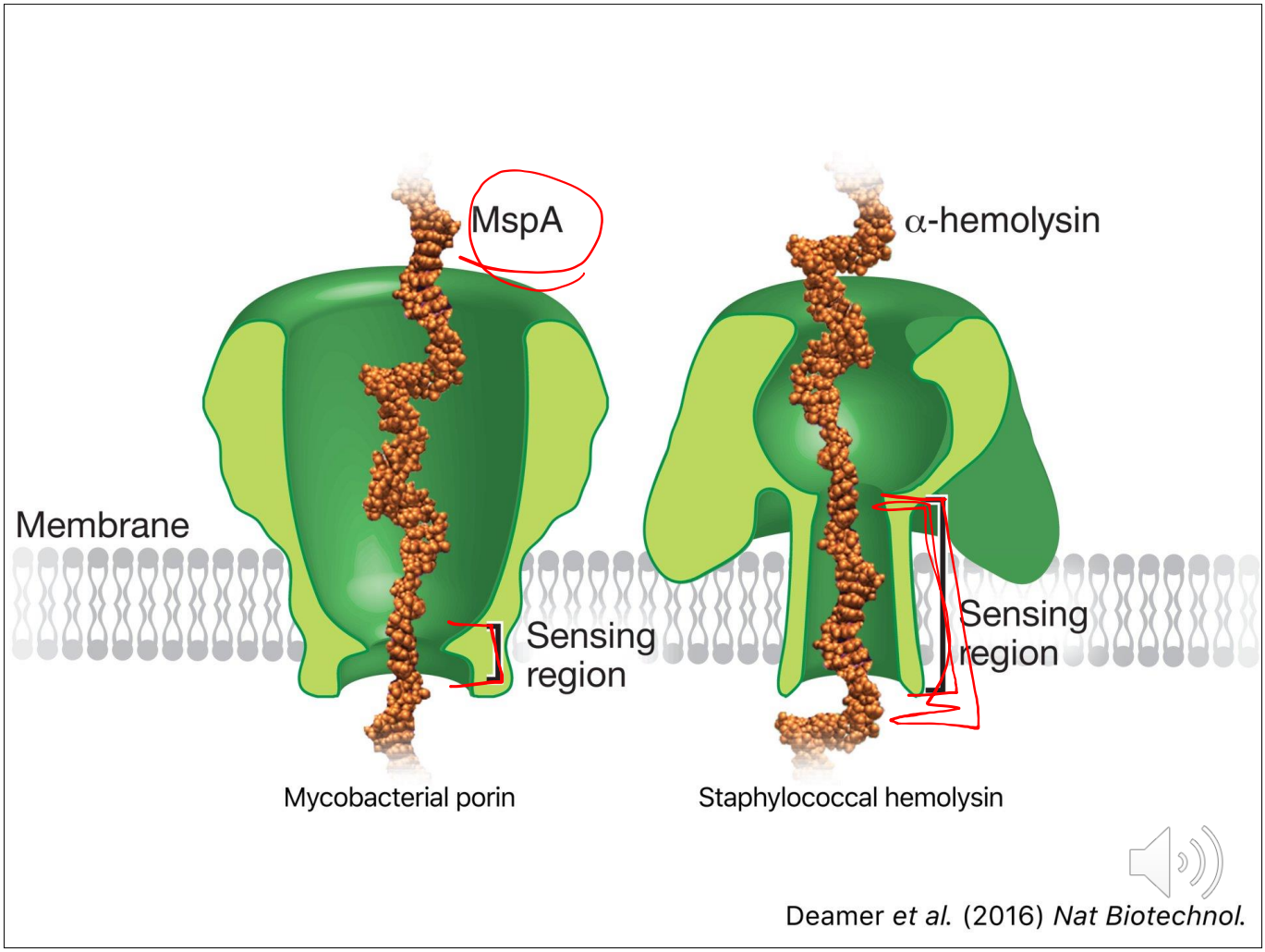


Akeson *et al.* (1999) *Biophys J*

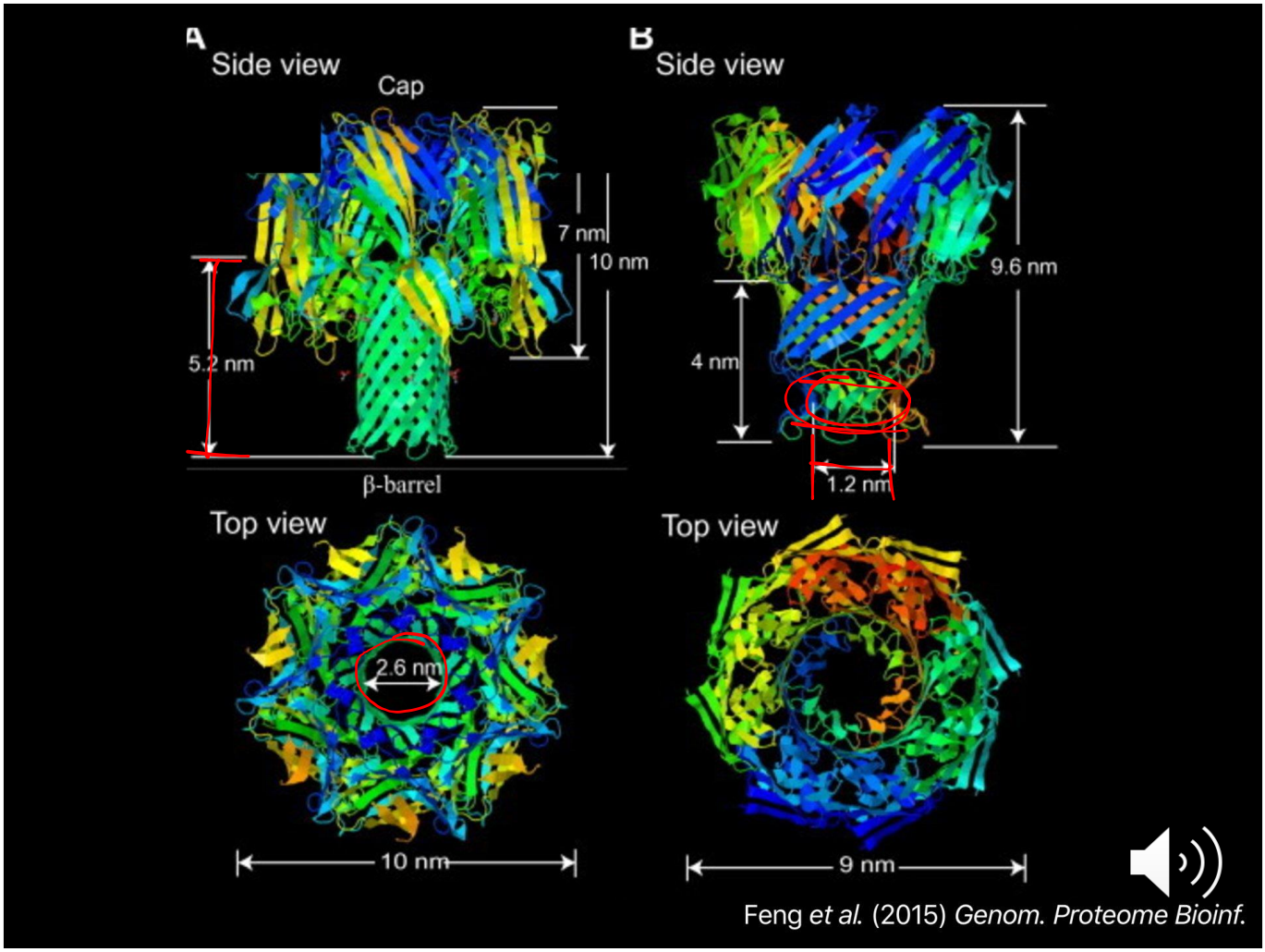


Branton *et al.* (2008) *Nat Biotechnol.*

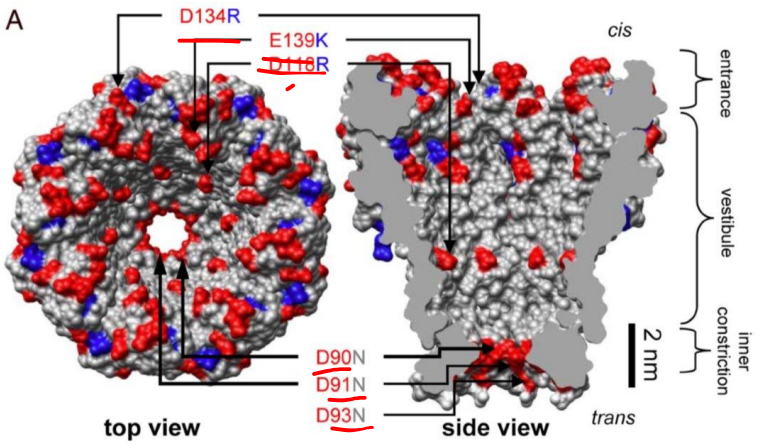
Could it be a bit more distinguishable?



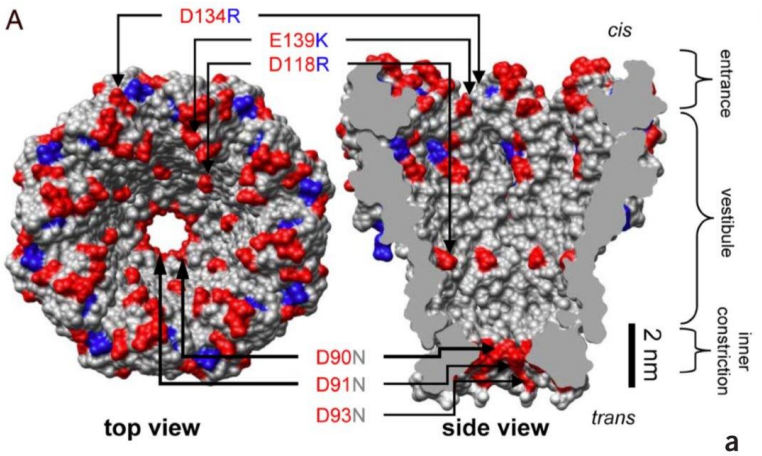
Deamer *et al.* (2016) *Nat Biotechnol.*



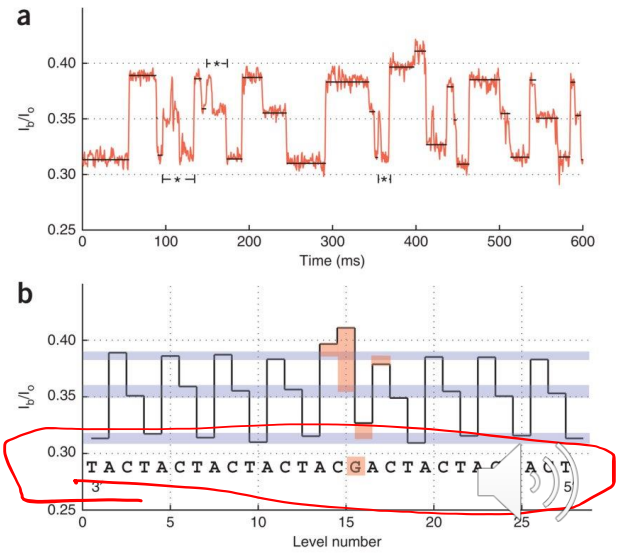
Feng *et al.* (2015) *Genom. Proteome Bioinf.*



Butler *et al.* (2008) *PNAS*

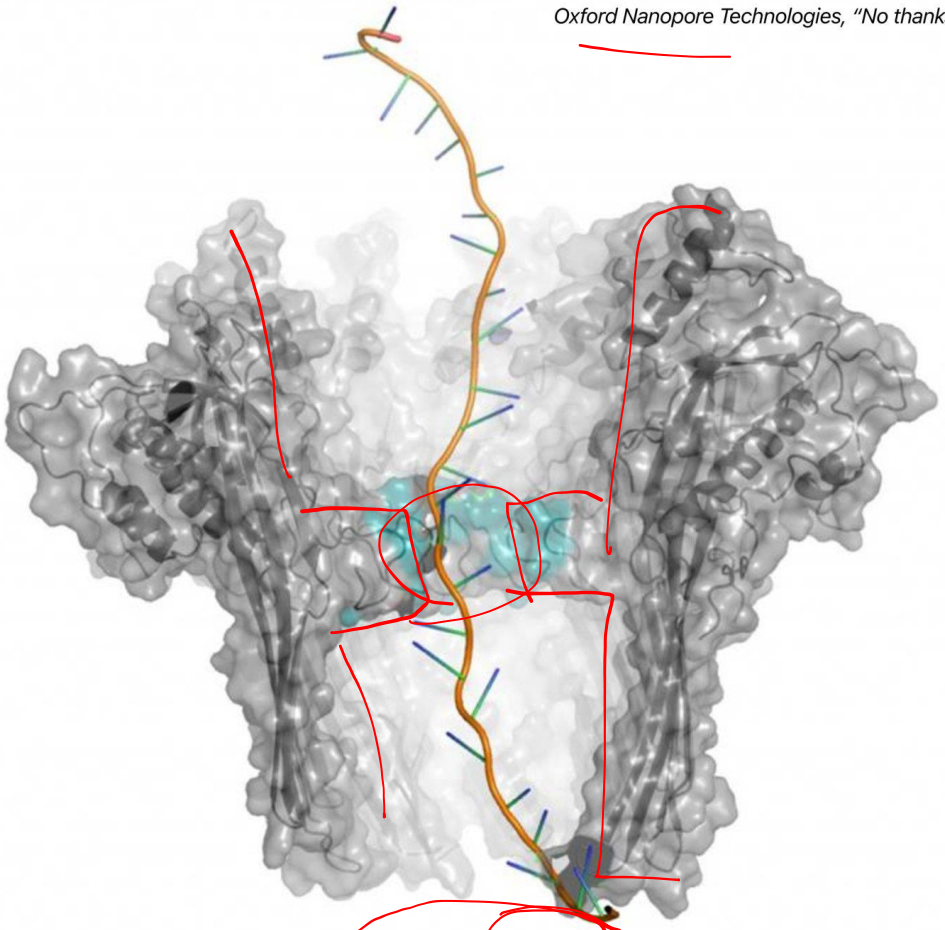


Butler *et al.* (2008) *PNAS*



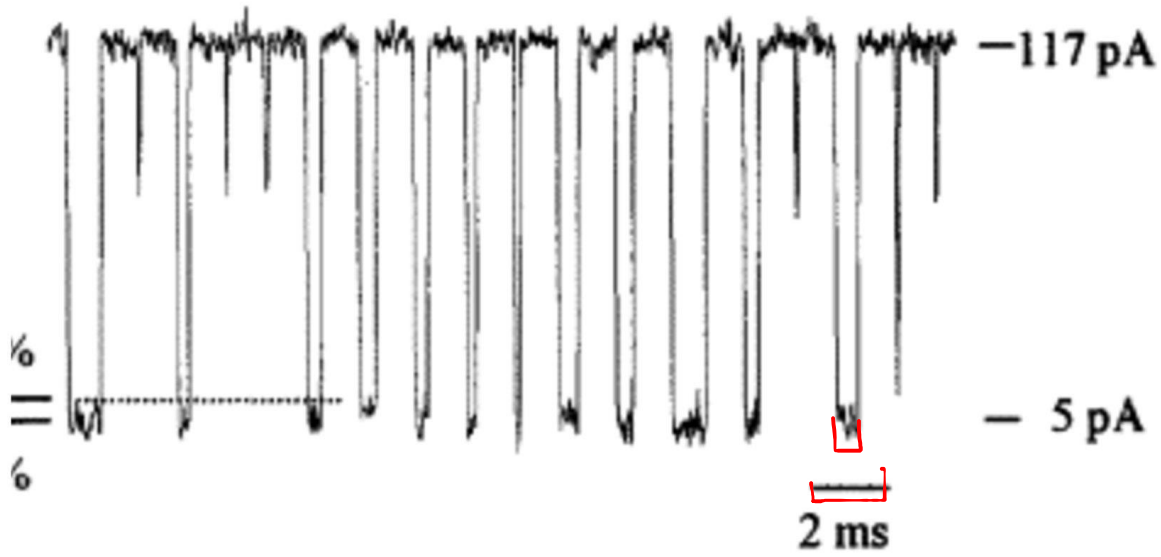
Manrao *et al.* (2012) *Nat. Biotechnol.*

Oxford Nanopore Technologies, "No thanks, I've already got one"

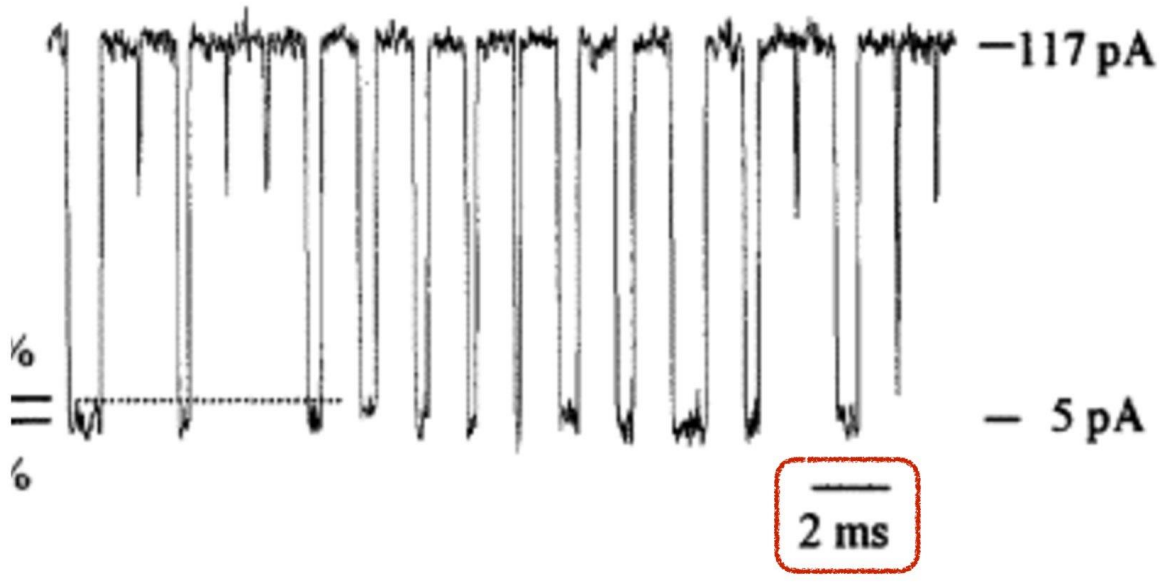


E. coli CsgG

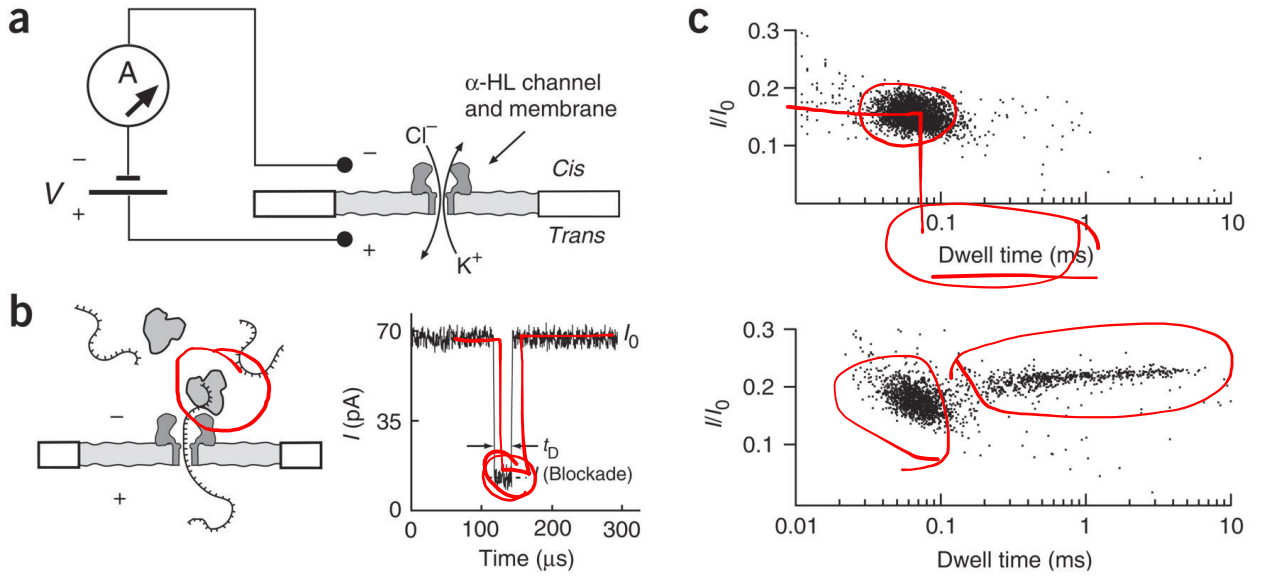
Could it be a bit more readable?



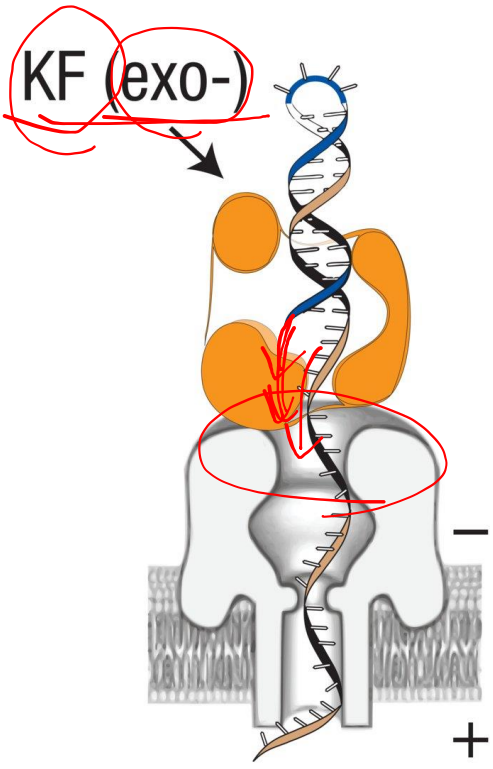
Akeson *et al.* (1999) *Biophys J*



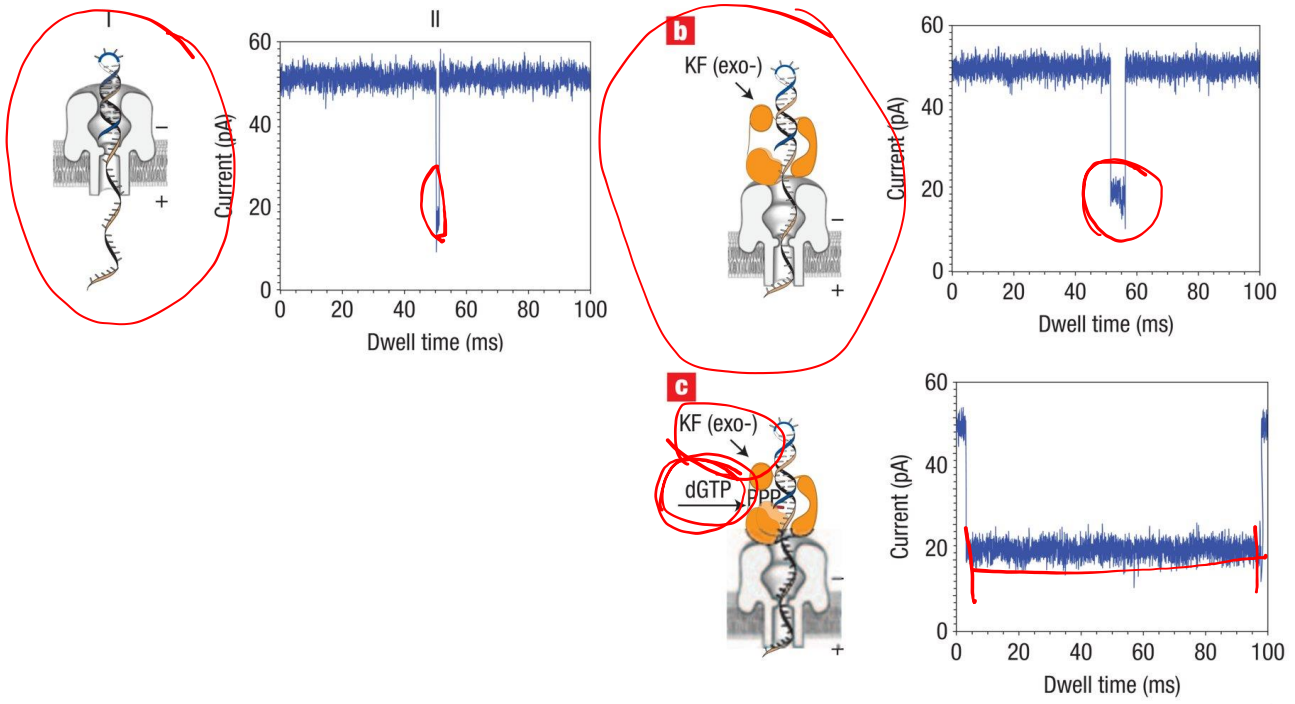
Akeson *et al.* (1999) *Biophys J*



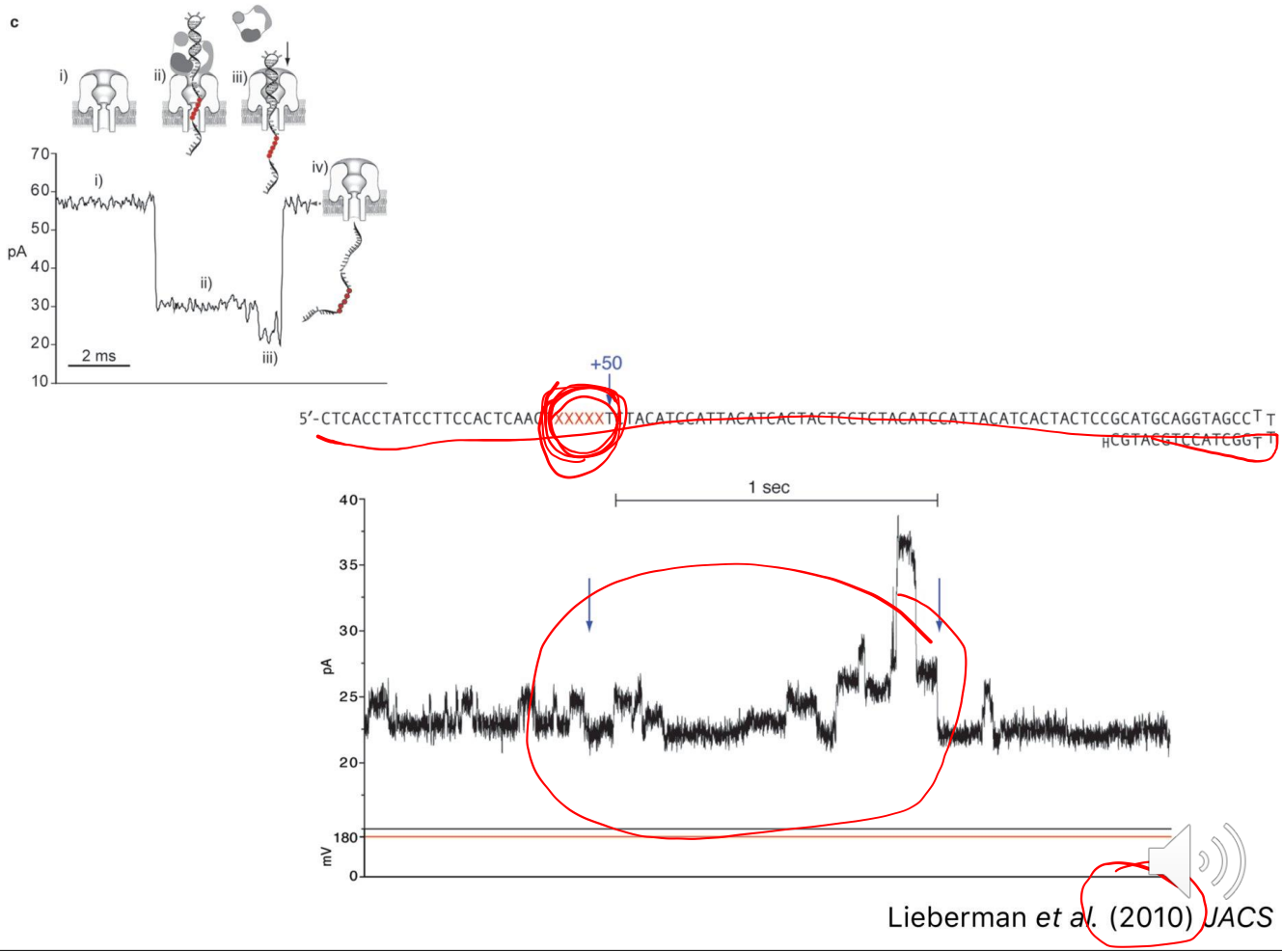
Hornblower *et al.* (2007) *Nat Methods*



Benner *et al.* (2007) *Nat Nanotechnol*



Benner *et al.* (2007) *Nat Nanotechnol*



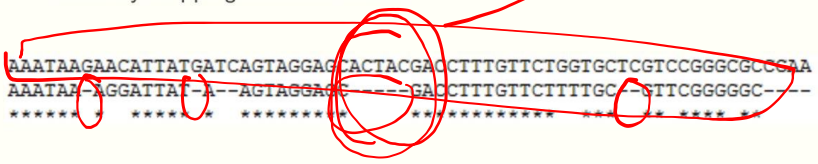
Lieberman *et al.* (2010) *JACS*

Effect of Poor Movement on Base-calling

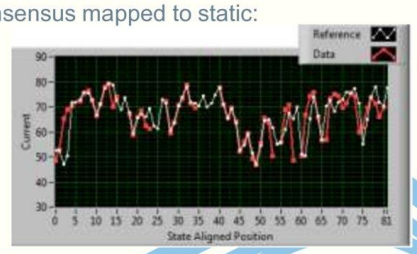
~June 2011

- Missing states from problems with movement make base calling from single molecule data difficult
- Consensus plots can be generated from multiple single molecule reads

- Base calls can be made from a consensus of single molecule reads by mapping to known model



Consensus mapped to static:



- Gaps in the sequence are evident even for known model
  - Need to make improvements to movement to drive progress

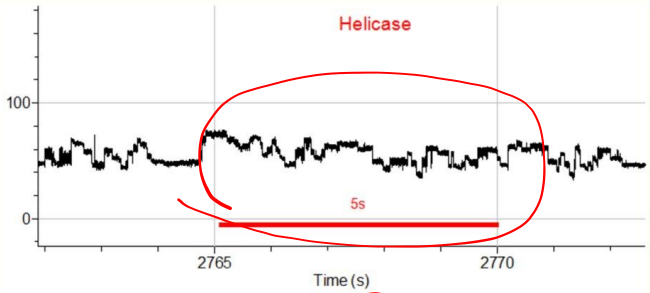
Oxford Nanopore Technologies, KeyGene Seminar (2016)

Confidential

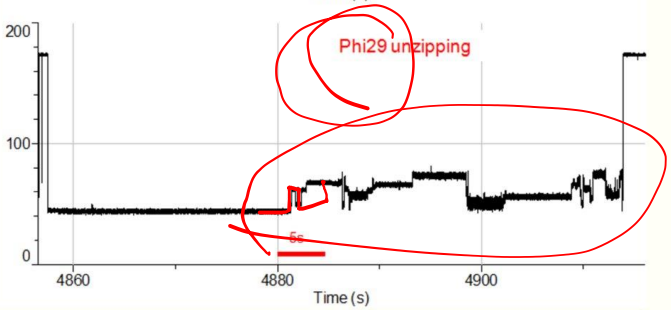
Strand Sequencing – Movement July 2011

Comparison Between Phi29 and Helicase Data

- Problems with using Phi29 as a motor include:
  - missed states, fast and slow movement regimes, backwards movement at low potential, pausing
- Initial work on helicases show that the distribution of movement is a lot more controlled



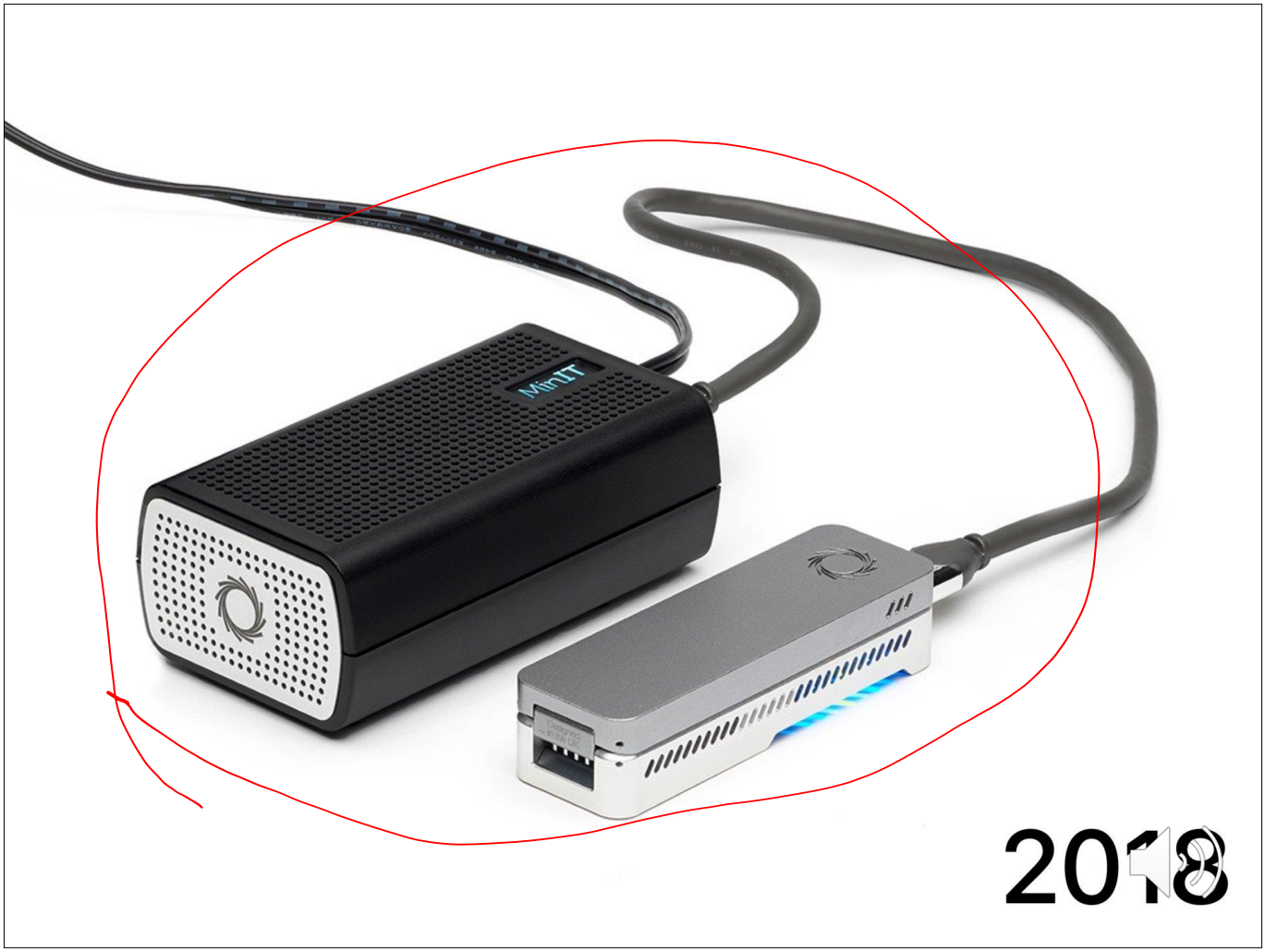
Helicase

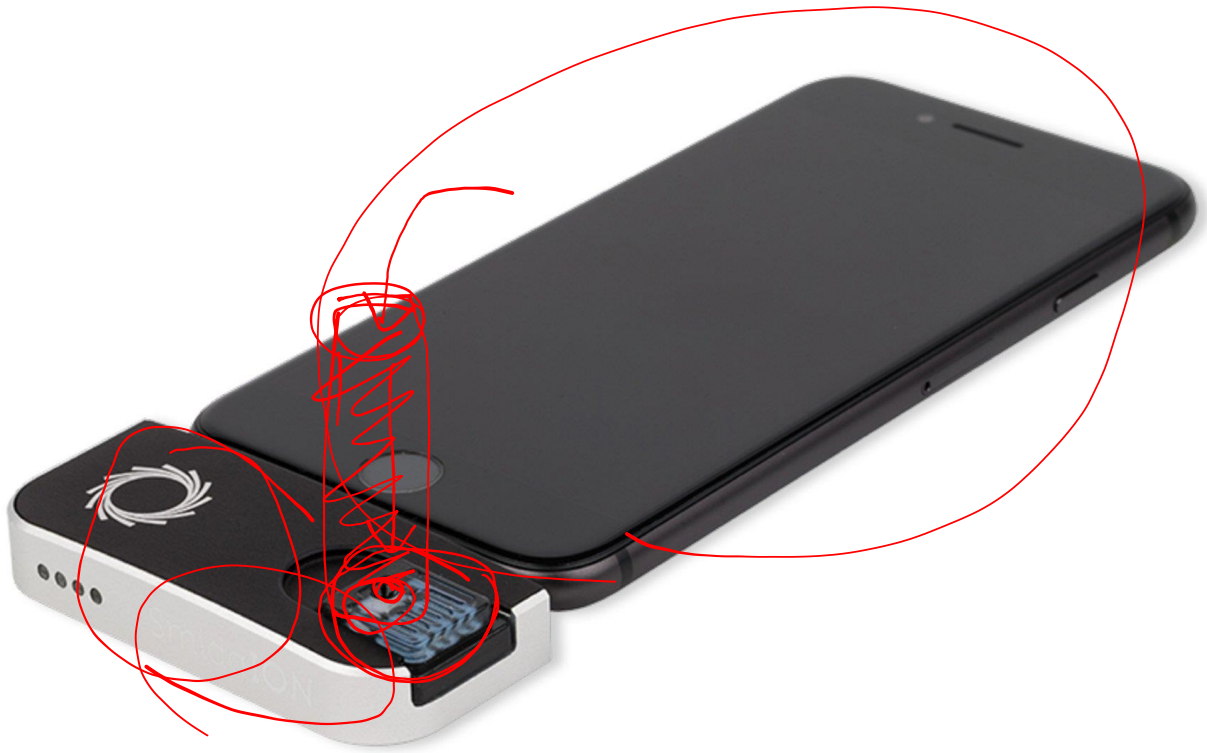


Polymerase

Oxford Nanopore Technologies, KeyGene Seminar (2016)


MinION
2014
AGBT


MinIT
2018

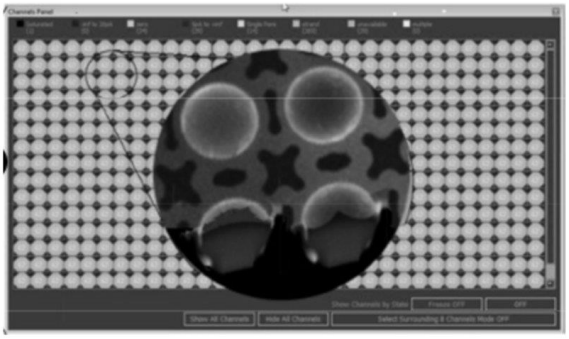


2019

Oxford NANOPORE Technologies

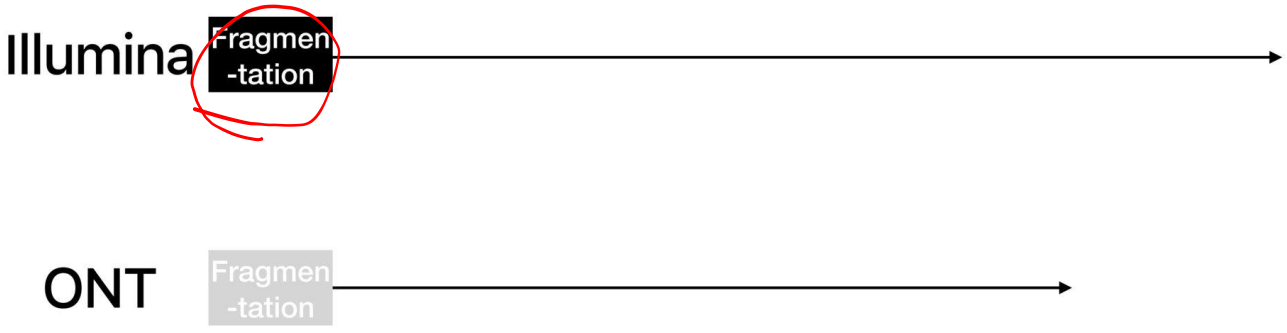
Nanopore DNA Sequencing

2048

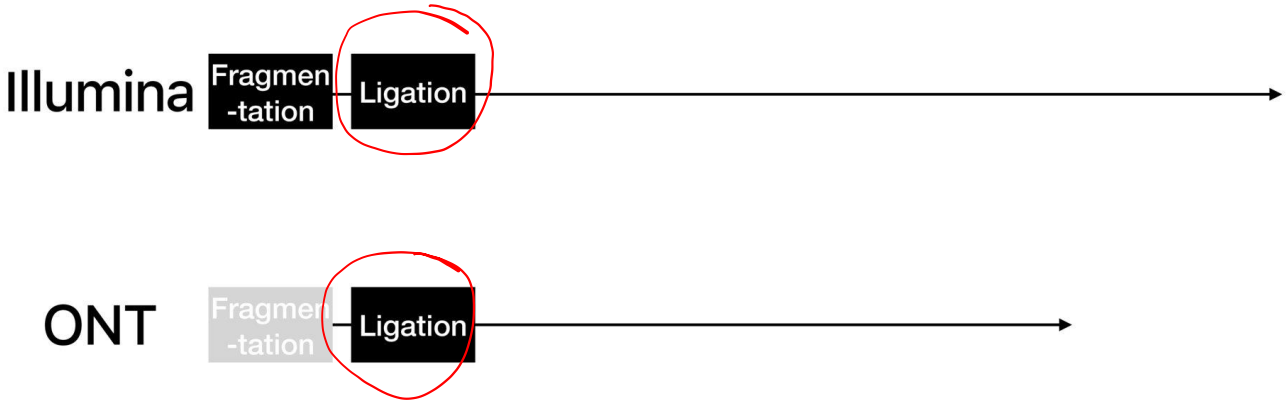


What's unique in nanopore sequencing

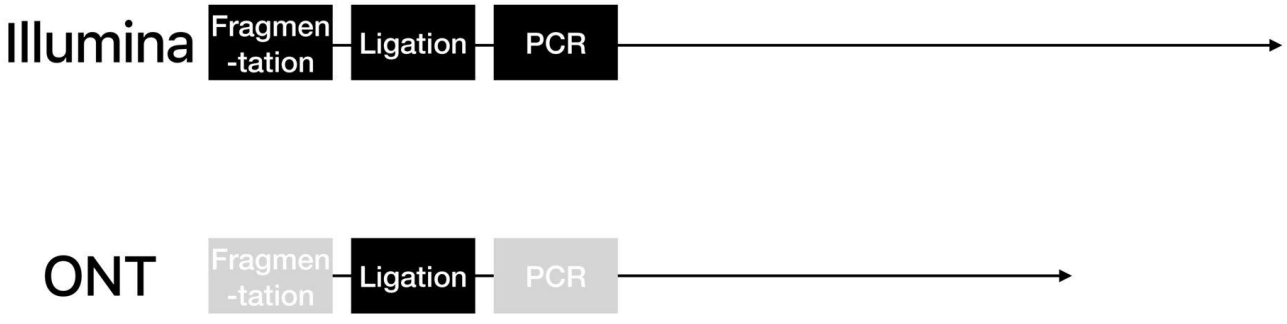
ONT vs Illumina – Workflow



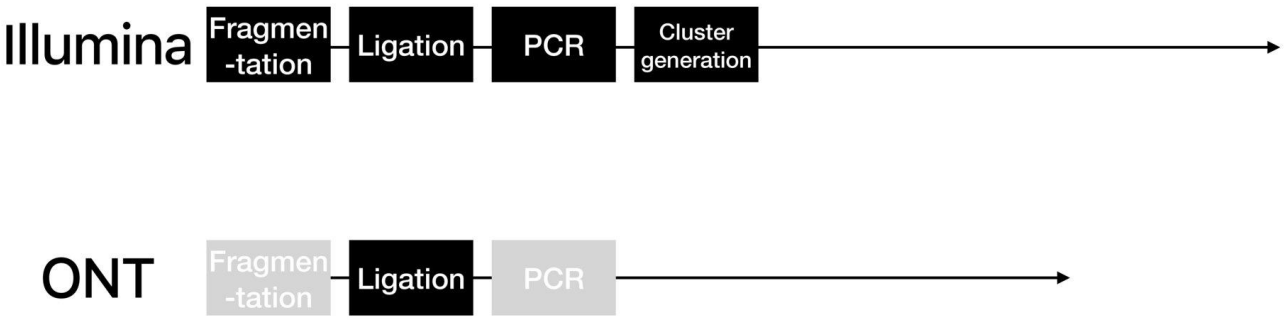
ONT vs Illumina – Workflow



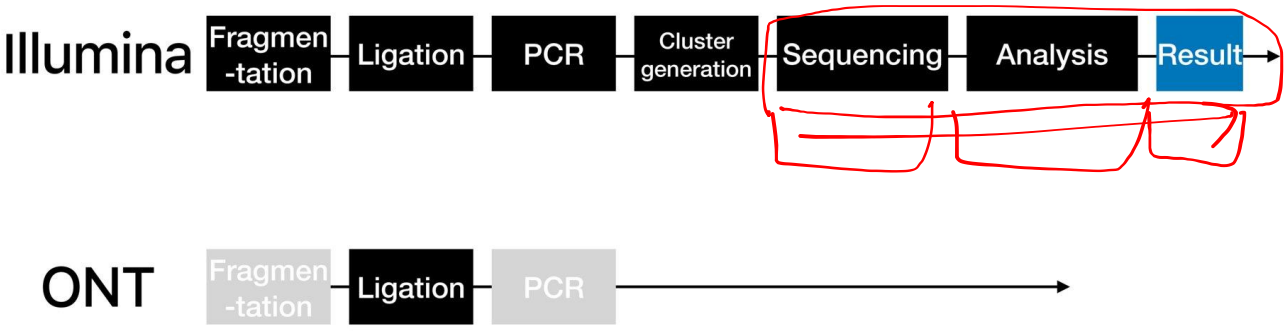
ONT vs Illumina – Workflow



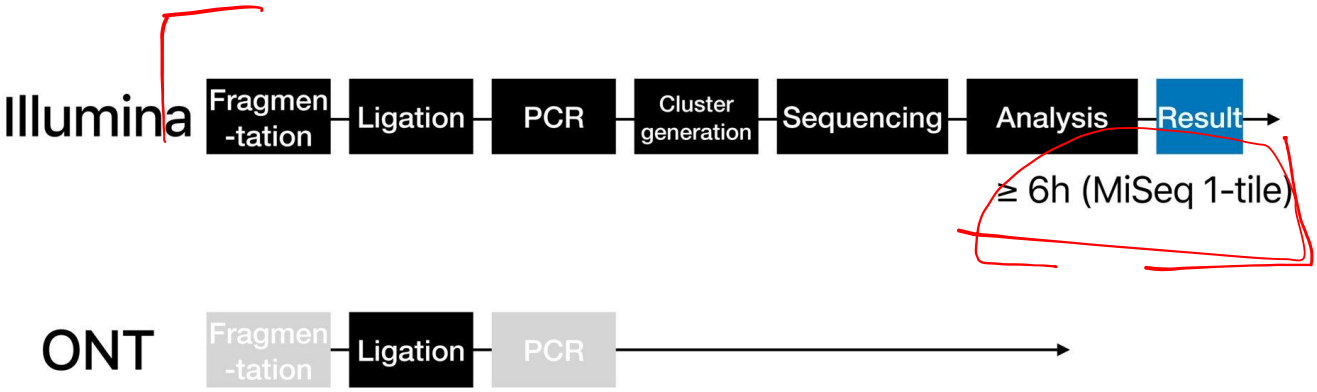
ONT vs Illumina – Workflow



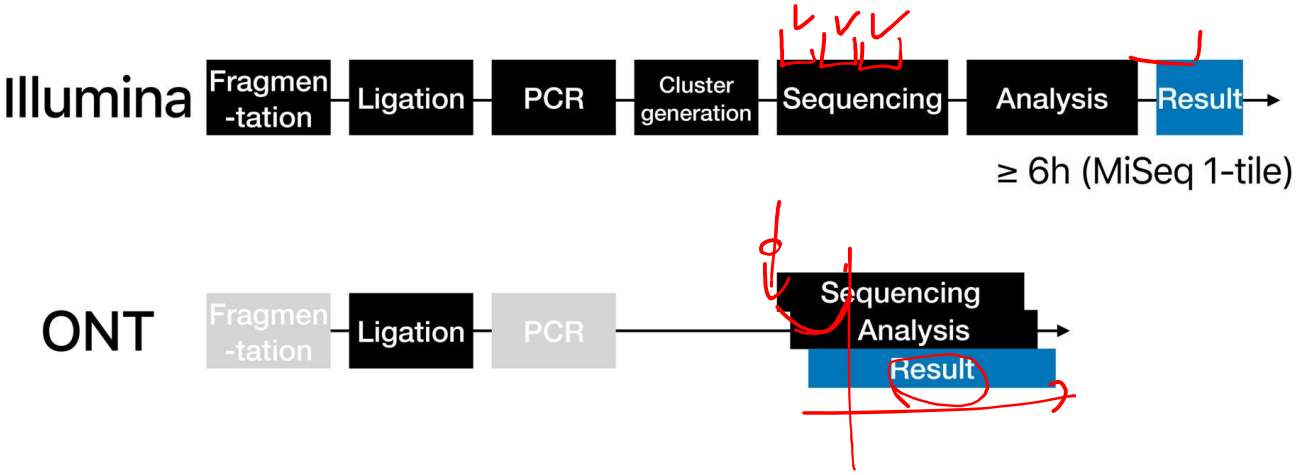
ONT vs Illumina – Workflow



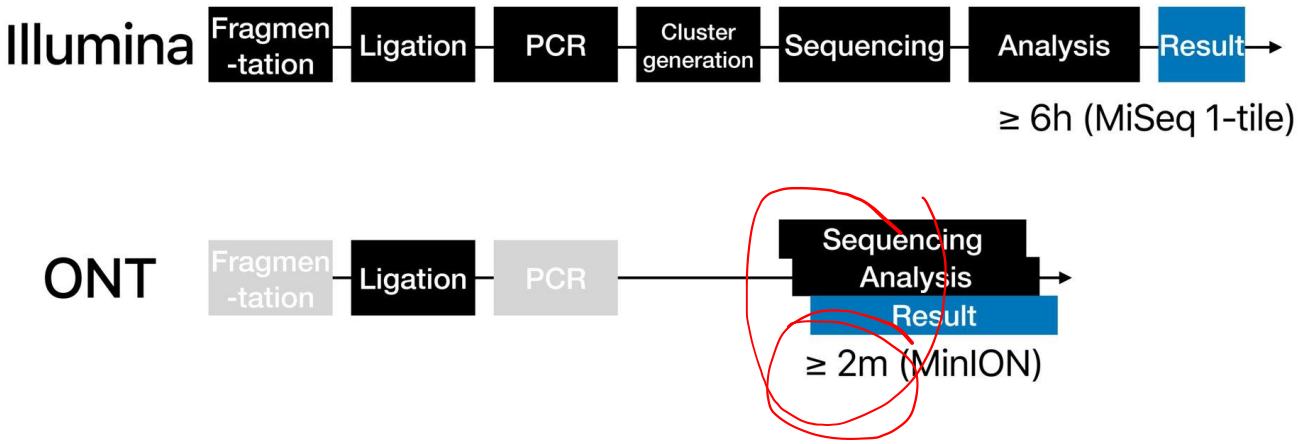
ONT vs Illumina – Workflow



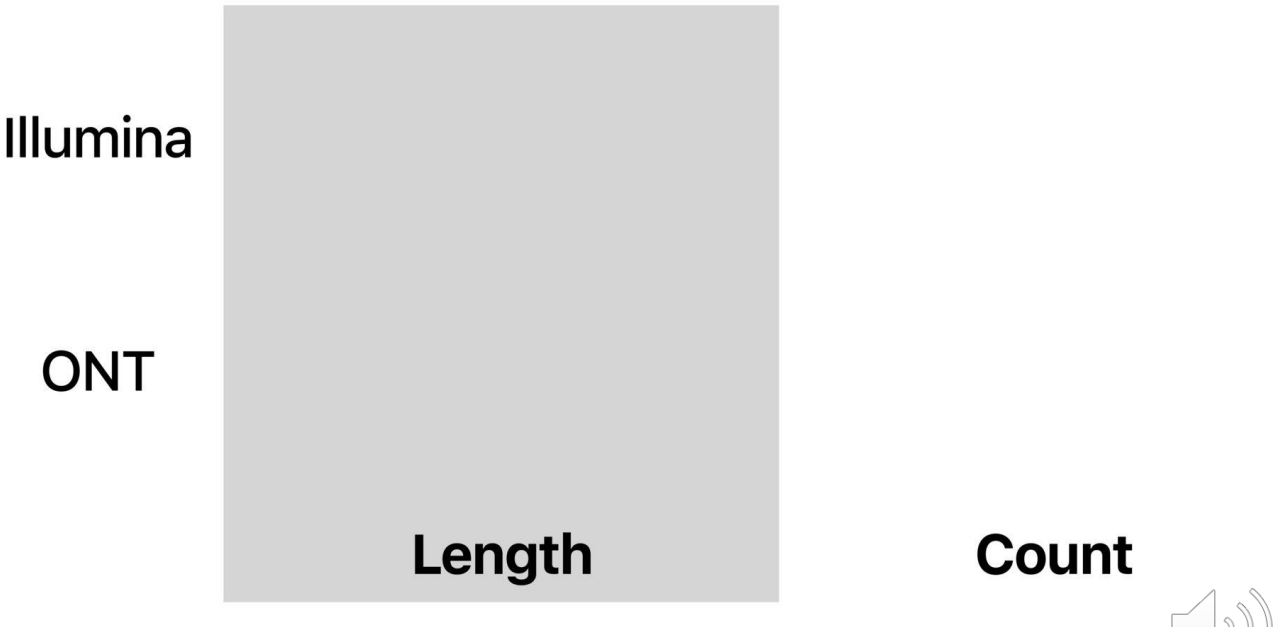
ONT vs Illumina – Workflow



ONT vs Illumina – Workflow



ONT vs Illumina – Read length



ONT vs Illumina – Read length

Illumina

Fixed, 50–800nt

ONT

Length

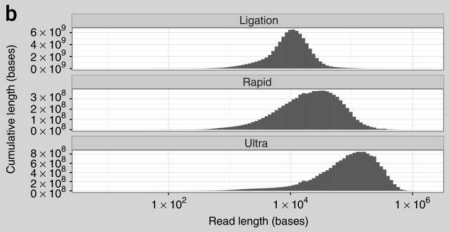
Count

ONT vs Illumina – Read length

Illumina

Fixed, 50–800nt

ONT



Jain *et al.* (2018) *Nat Biotechnol*

Length

Count

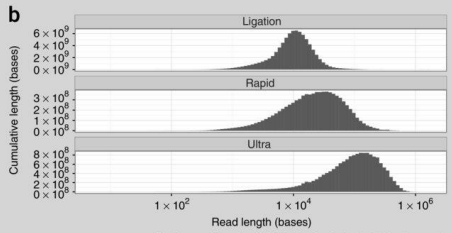
ONT vs Illumina – Read length

Illumina

Fixed, 50–800nt

1–2 (SE)
2–4 (PE)

ONT



Jain *et al.* (2018) *Nat Biotechnol*

Length

Count

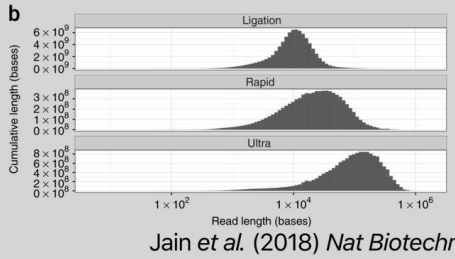
ONT vs Illumina – Read length

Illumina

Fixed, 50–800nt

1–2 (SE)
2–4 (PE)

ONT

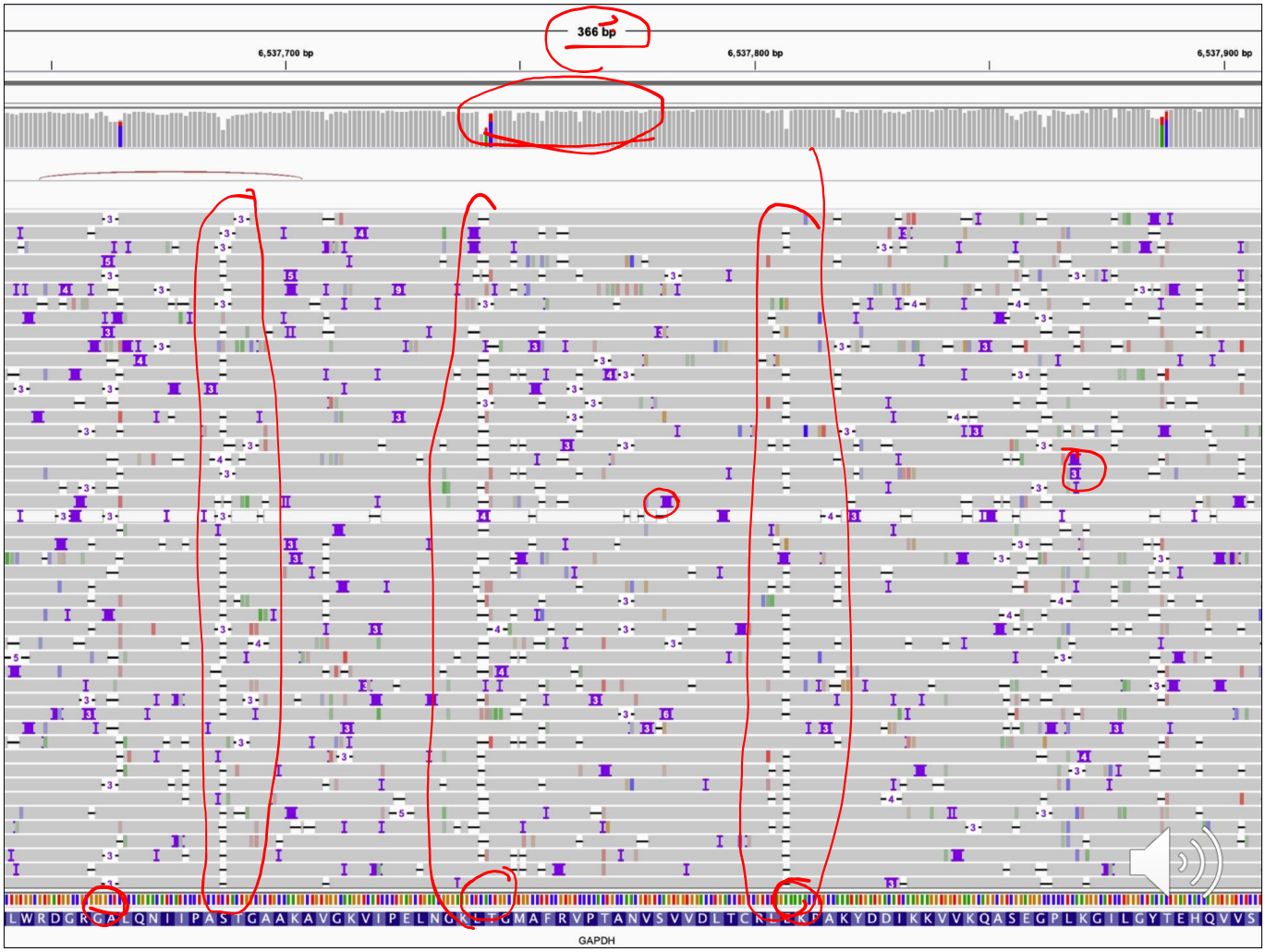


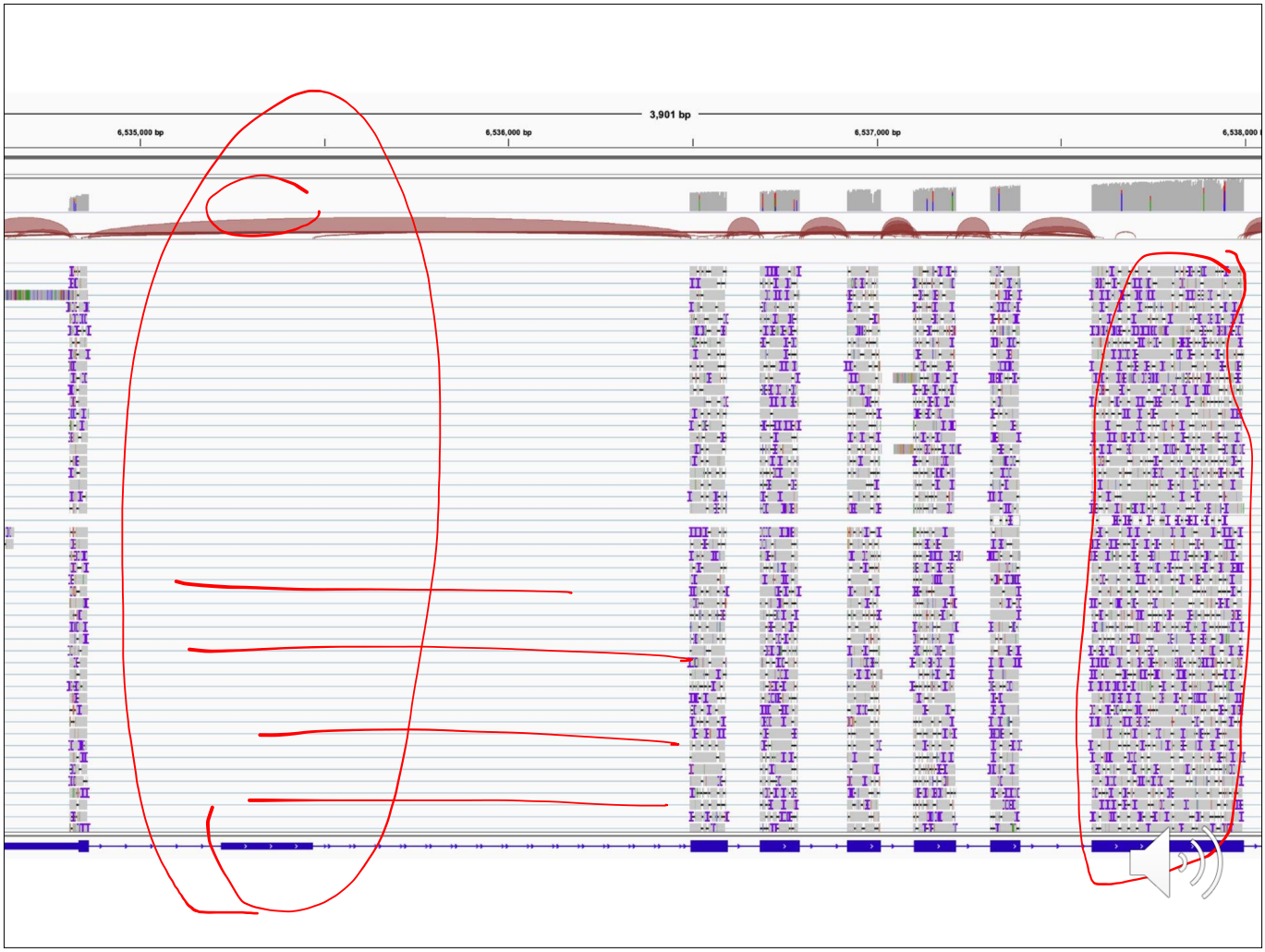
Jain *et al.* (2018) *Nat Biotechnol*

1 (1D)
1–2 (1D²)
1–20 (LCS+1D)

Length

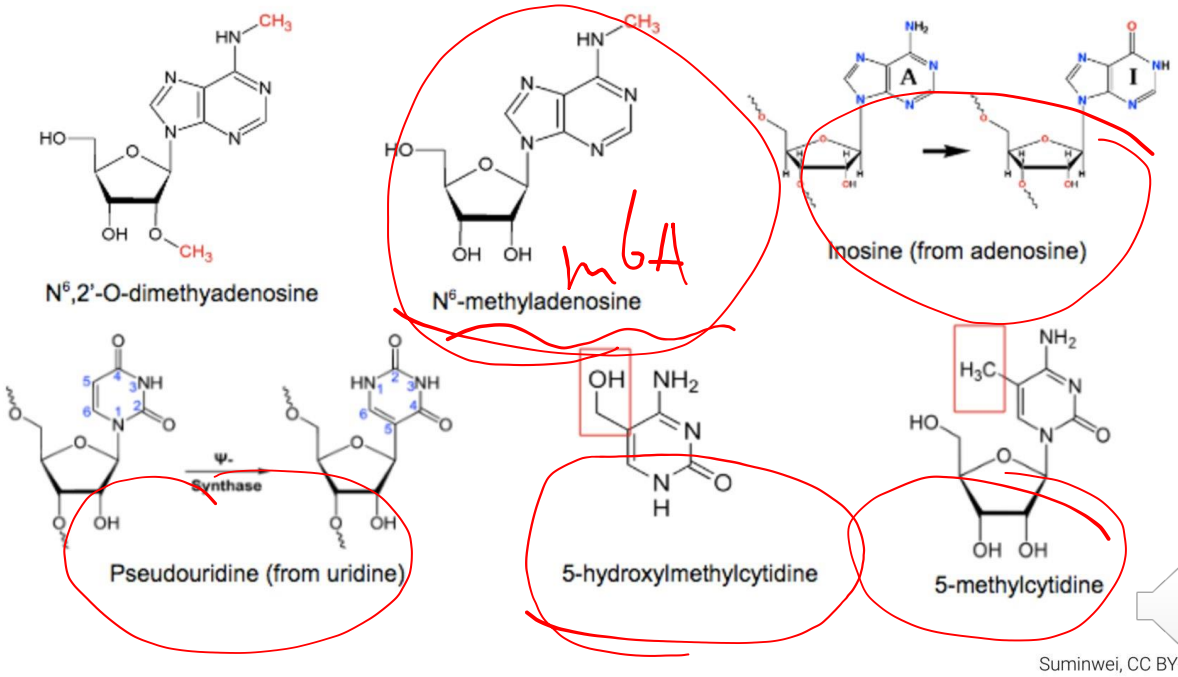
Count


366 bp
6,537,700 bp
6,537,800 bp
6,537,900 bp
GAPDH

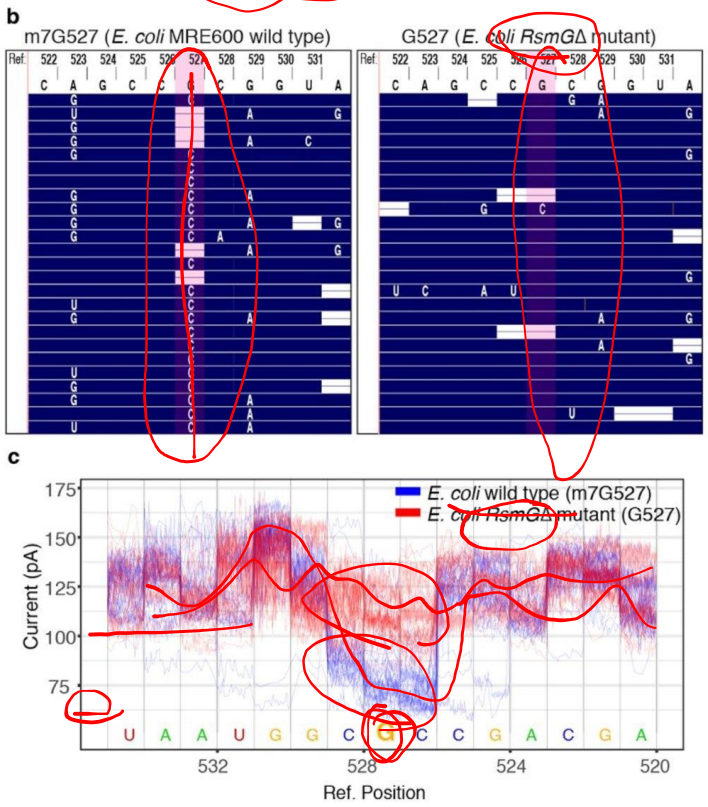

3,901 bp
6,535,000 bp
6,536,000 bp
6,537,000 bp

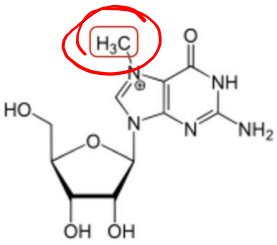
RNA sequencing with nanopore

Modified RNA bases can be detected



m7G in *E. coli* 16S rRNA





7-methylguanosine

Smith, Jain *et al.* (2017) *bioRxiv*, doi:10.1101/132274

Long reads


“Detects” the splicing isoforms

"Detects" the splicing isoforms

Spans the full length

"Detects" the splicing isoforms

Spans the full length

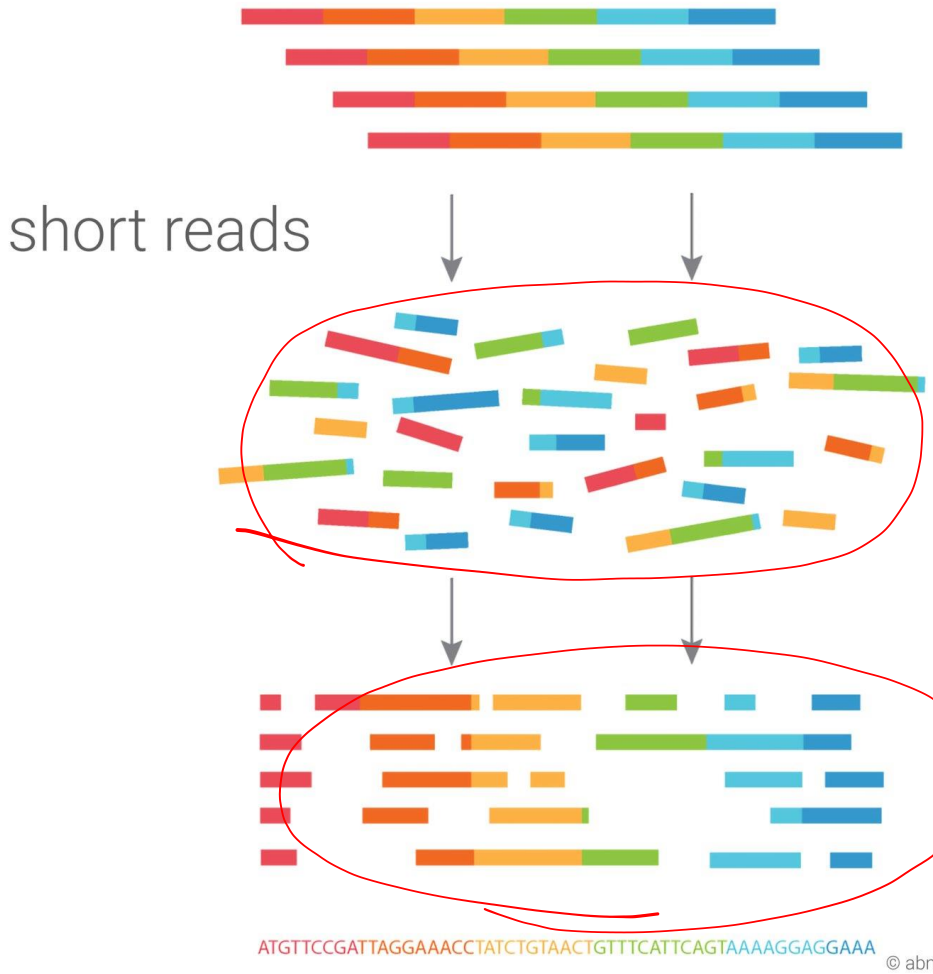
Read count ~ molecule count

"Detects" the splicing isoforms

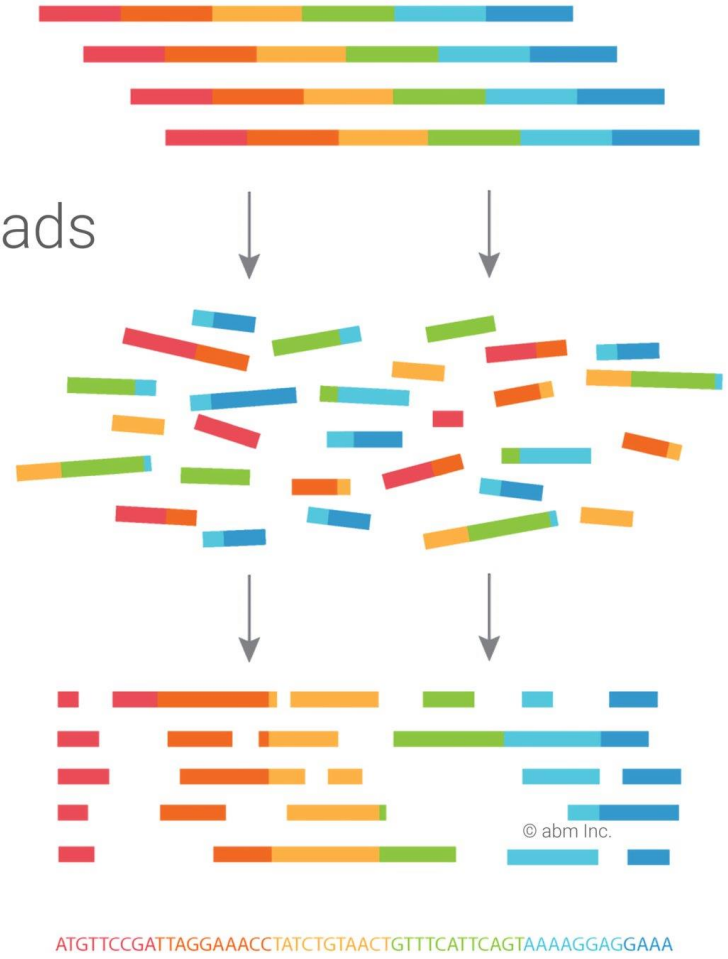
Spans the full length

Read count ~ molecule count

Can read long range associations



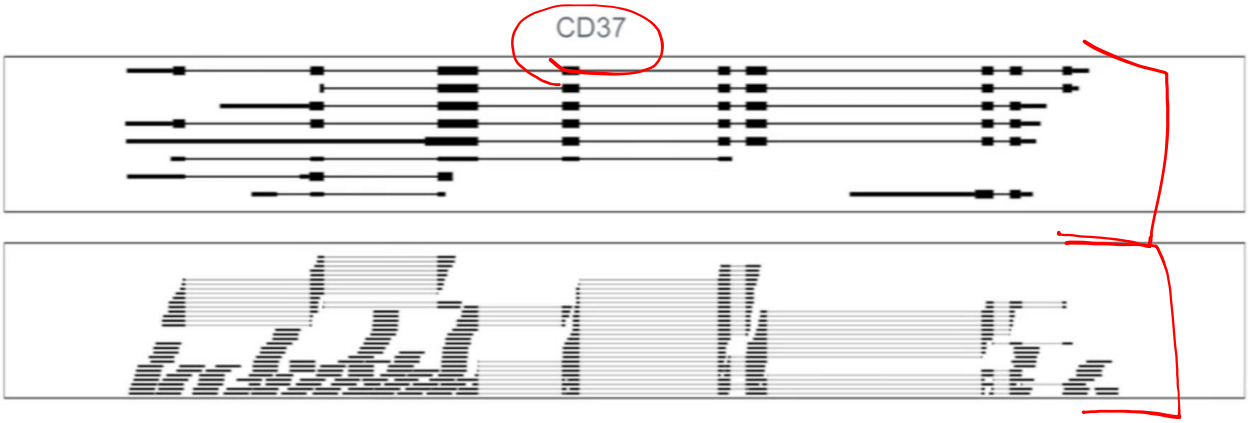
long reads



ATGTTCCGATTAGGAAACCTATCTGTAACTGTTTCATTCAGTAAAAGGAGGAAA

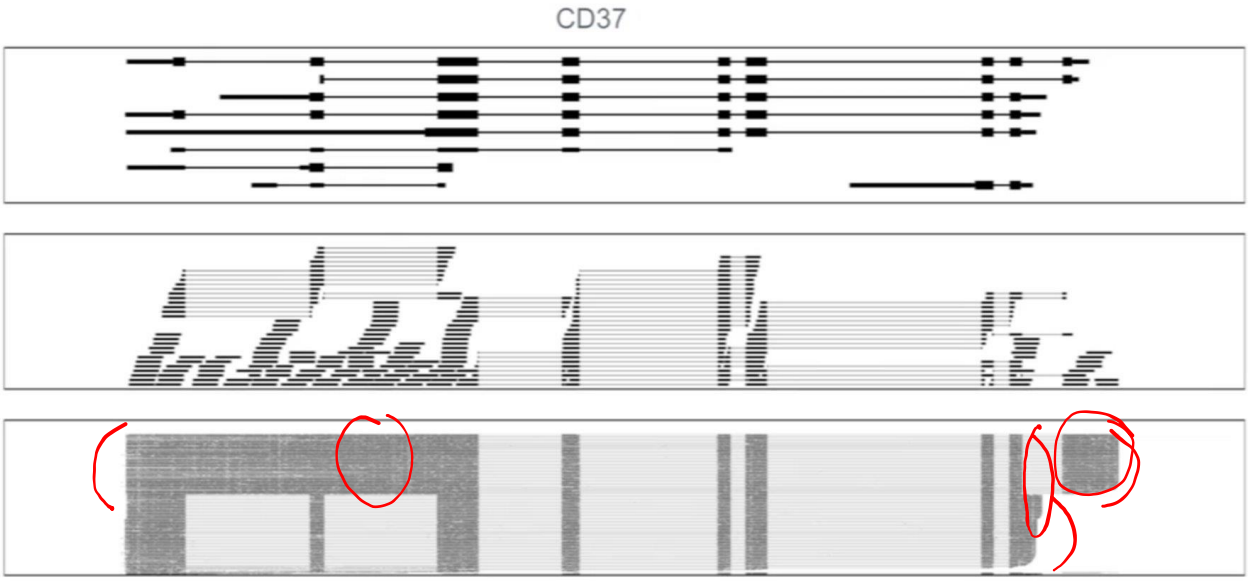
long reads

Long reads enable “reading” of isoforms



Christopher Vollmers (2017) *ONT Knowledge Exchange*

Long reads enable “reading” of isoforms



Christopher Vollmers (2017) *ONT Knowledge Exchange*

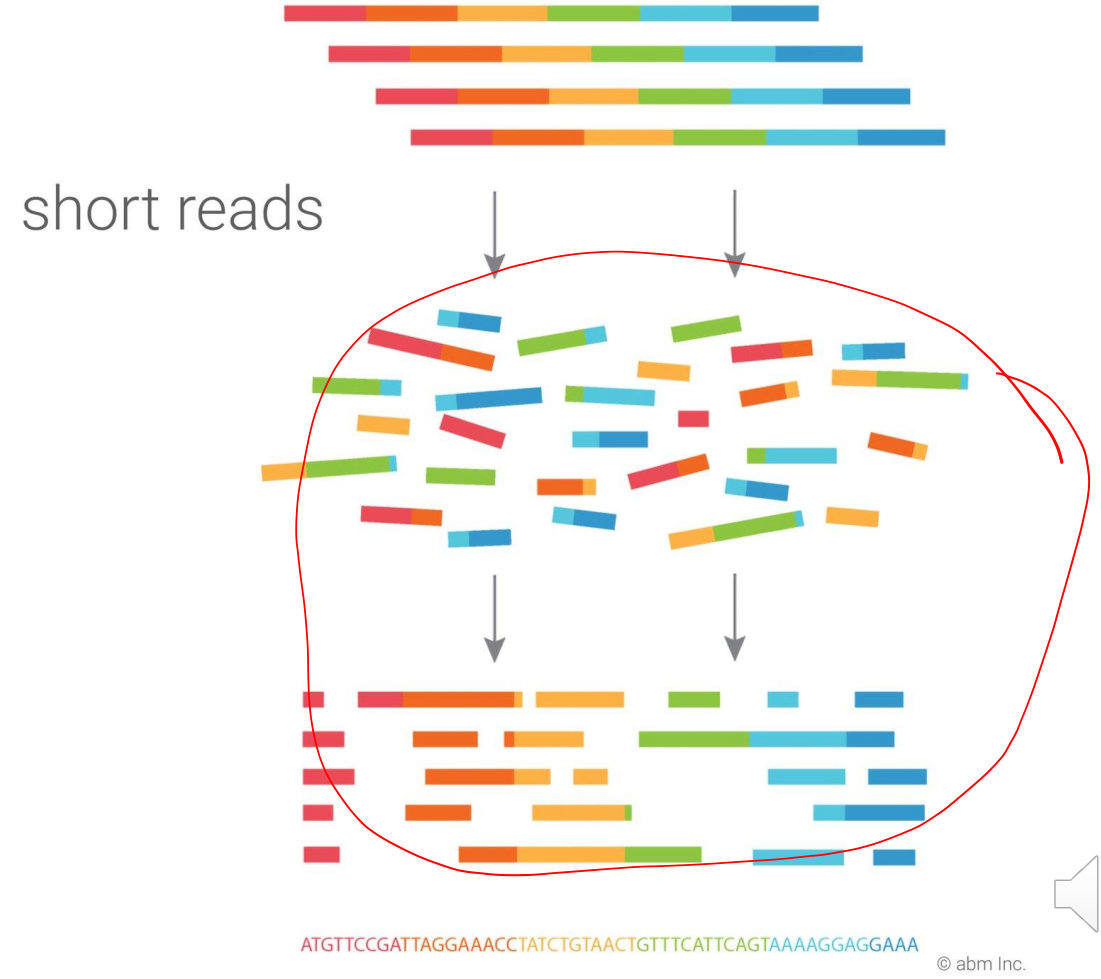
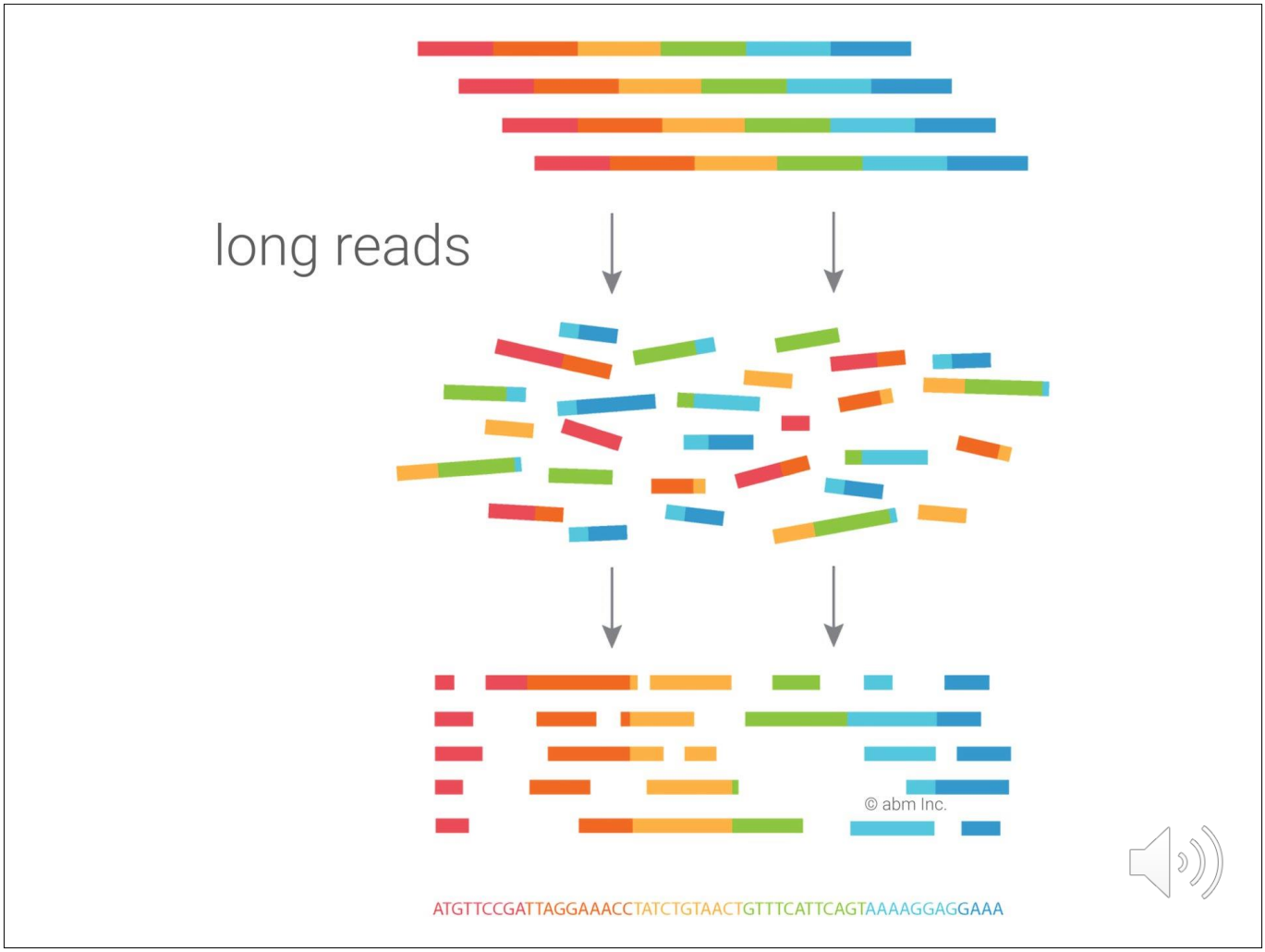
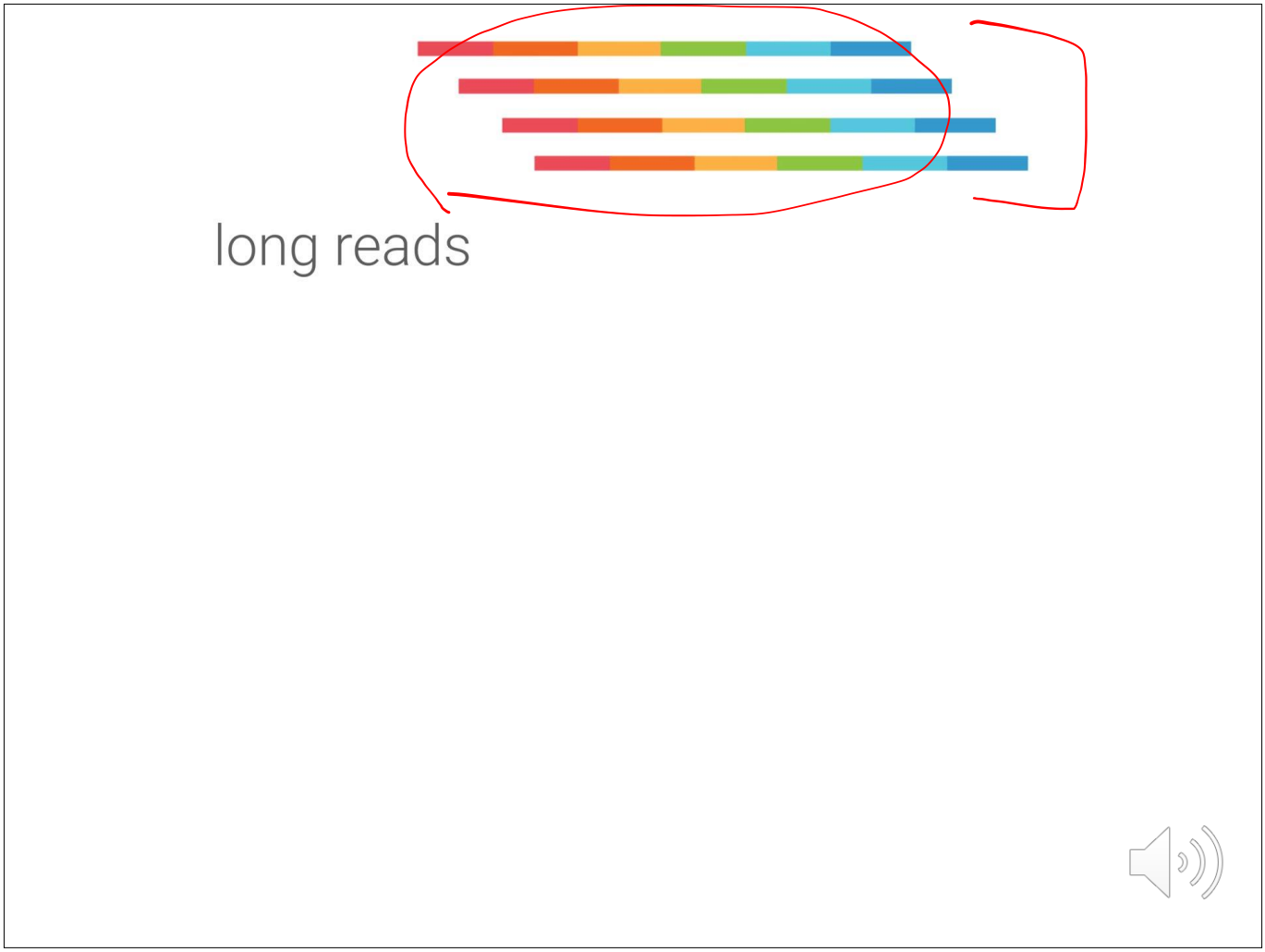


Table 11: Normalization methods for the comparison of gene read counts within the same sample.

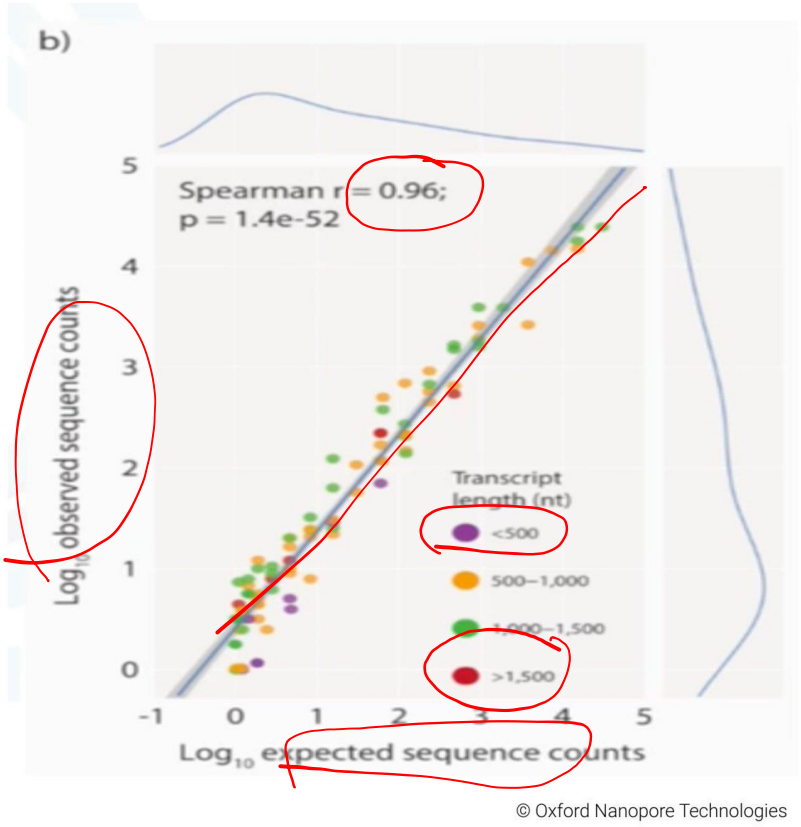
Name	Details	Comment
RPKM (reads per kilobase of exons per million mapped reads)	1. For each gene, count the number of reads mapping to it. 2. Divide that count by: the length of the gene in base pairs divided by 1,000 multiplied by the total number of mapped reads divided by 10^6. $$RPKM_i = \frac{read\,count\,of\,gene\,i}{(\frac{length\,of\,gene\,i}{10^3})(\frac{library\,size}{10^6})}$$	• introduces a bias in the per-gene variances, in particular for lowly expressed genes (Oshlack and Wakefield, 2009) • implemented in `edgeR`'s `rpkm()` function
FPKM (fragments per kilobase...)	1. Same as RPKM, but for paired-end reads: 2. The number of fragments (defined by two reads each) is used.	• implemented in `DESeq2`'s `fpkm()` function
TPM	Instead of normalizing to the total library size, TPM represents the abundance of an individual gene i in relation to the abundances of the other transcripts (e.g., j) in the sample. 1. For each gene, count the number of reads mapping to it and divide by its length in base pairs (= counts per base). 2. Multiply that value by 1 divided by the sum of all counts per base of every gene. 3. Multiply that number by 10^6. $$TPM_i = \frac{X_i}{l_i} * \frac{1}{\sum_j \frac{X_j}{l_k}}$$	• details in Wagner et al. (2012)

Dünbar *et al.* (2015)


long reads
© abm Inc.
ATGTTCCGATTAGGAAACCTATCTGTAACTGTTTCATTCAGTAAAAGGAGGAAA


long reads

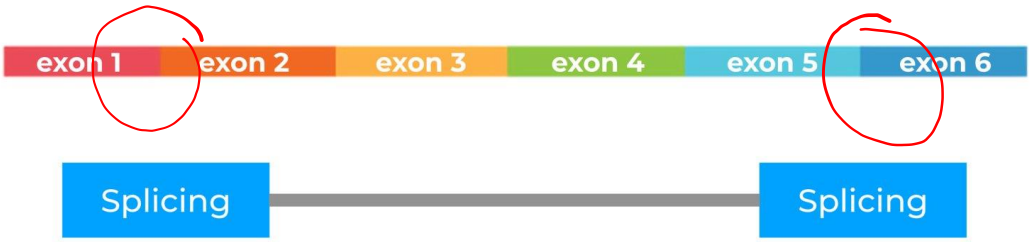
Read count ~ molecule count



ERCC

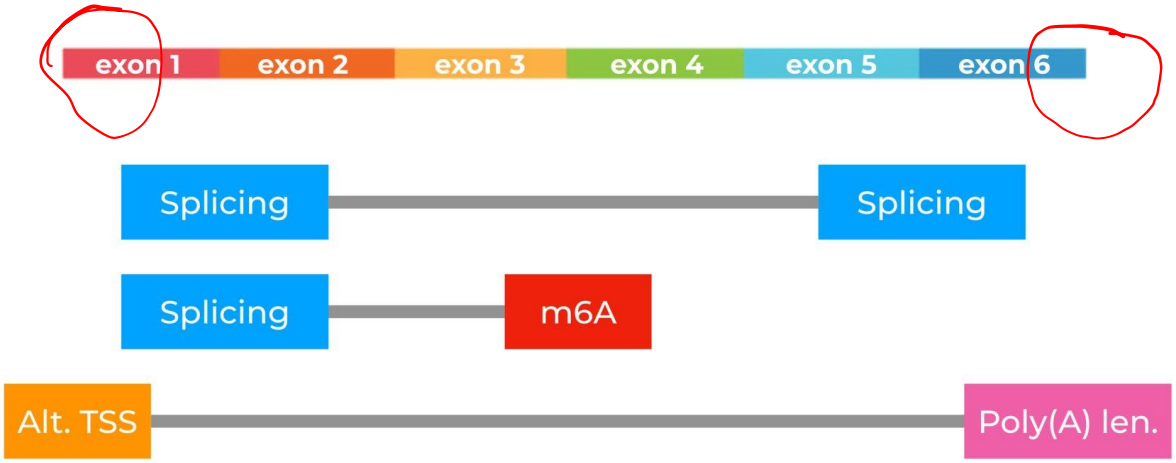
Capturing long-range associations

Capturing long-range associations

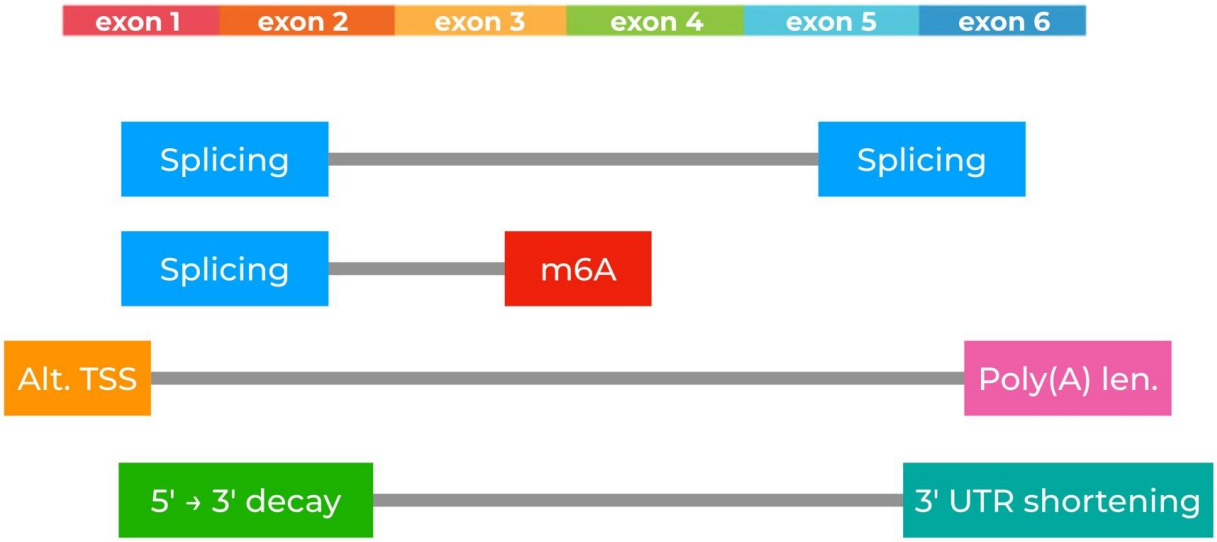


Capturing long-range associations

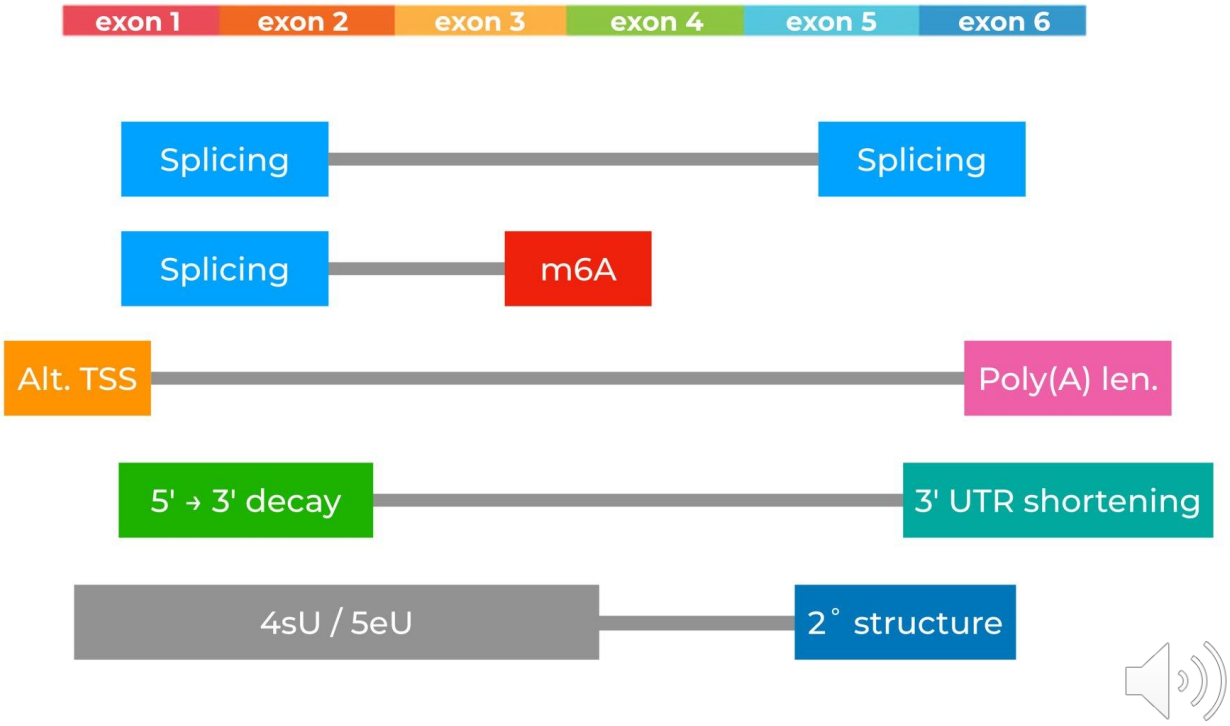
Capturing long-range associations



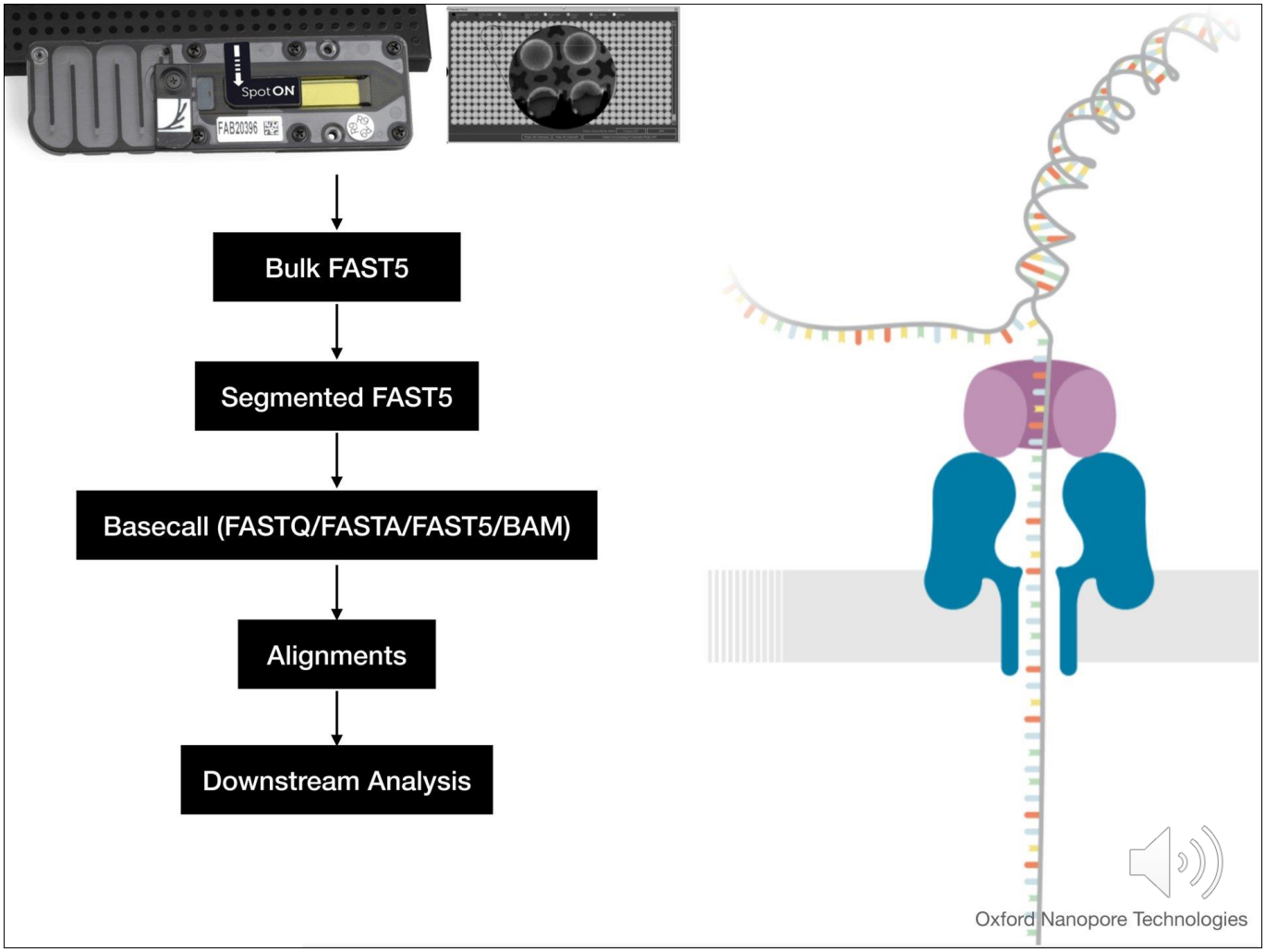
Capturing long-range associations

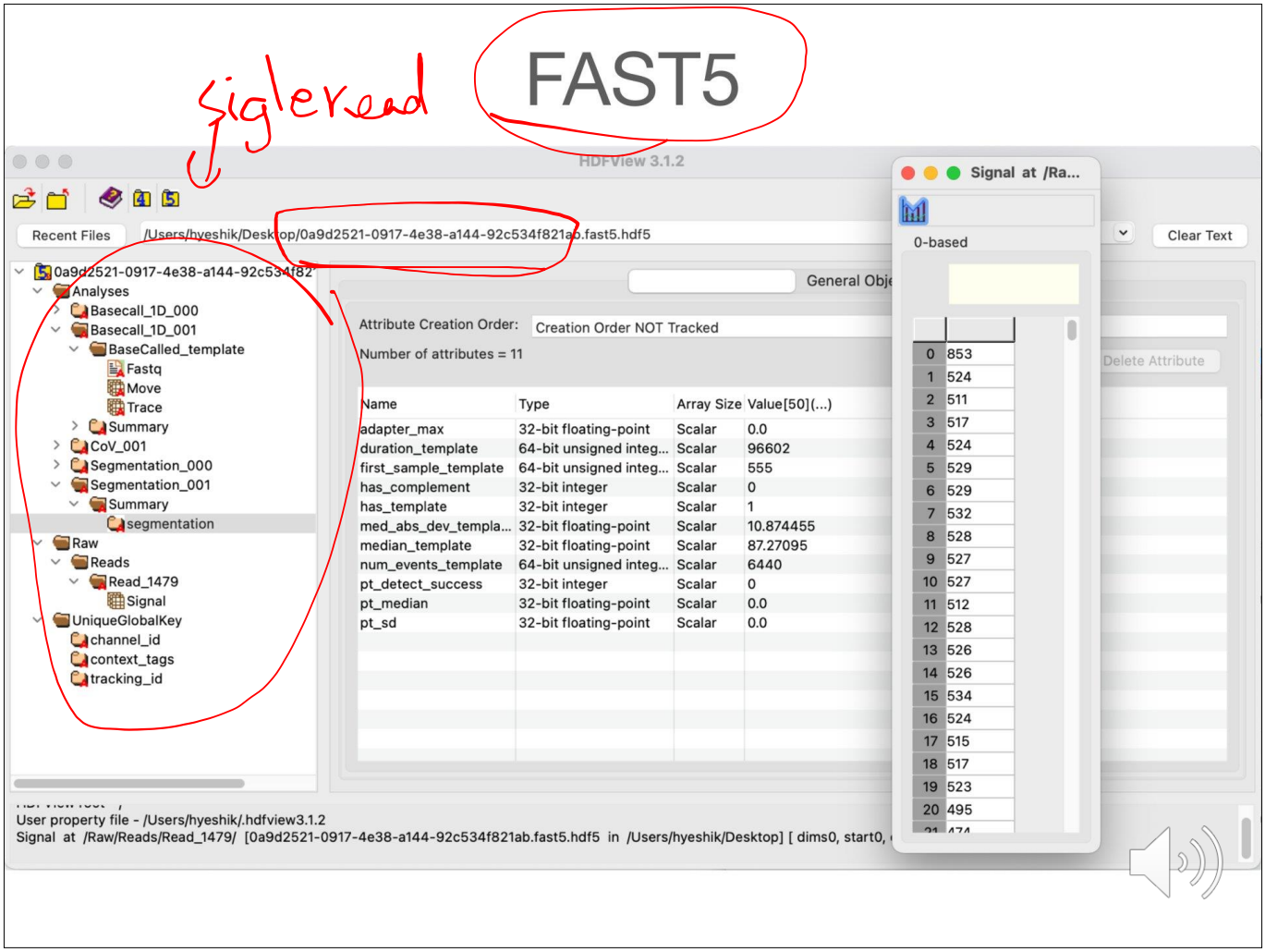


Capturing long-range associations



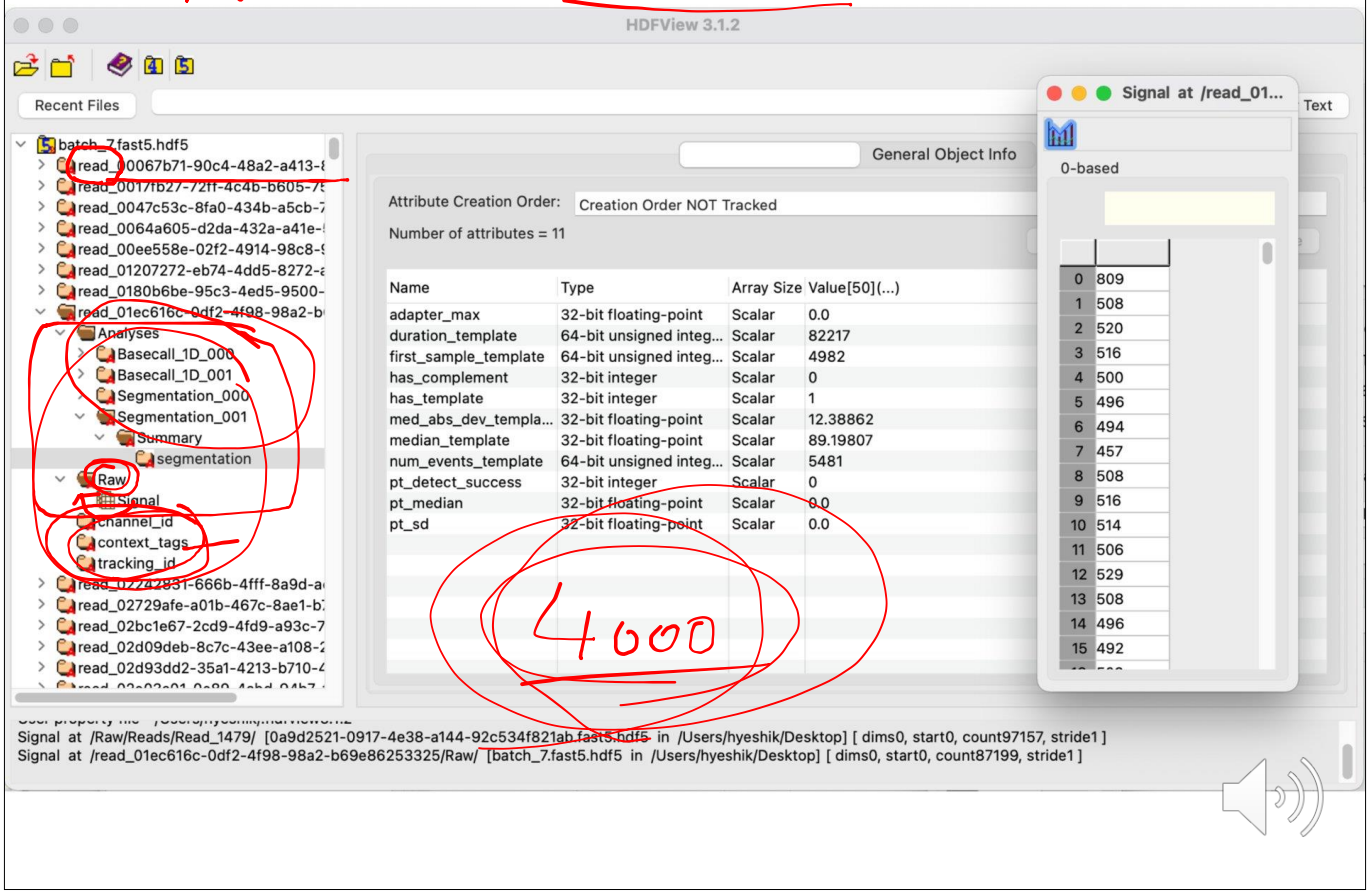
Data flow in nanopore sequencing


Spot ON
FAB20396
Bulk FAST5
Segmented FAST5
Basecall (FASTQ/FASTA/FAST5/BAM)
Alignments
Downstream Analysis
Oxford Nanopore Technologies

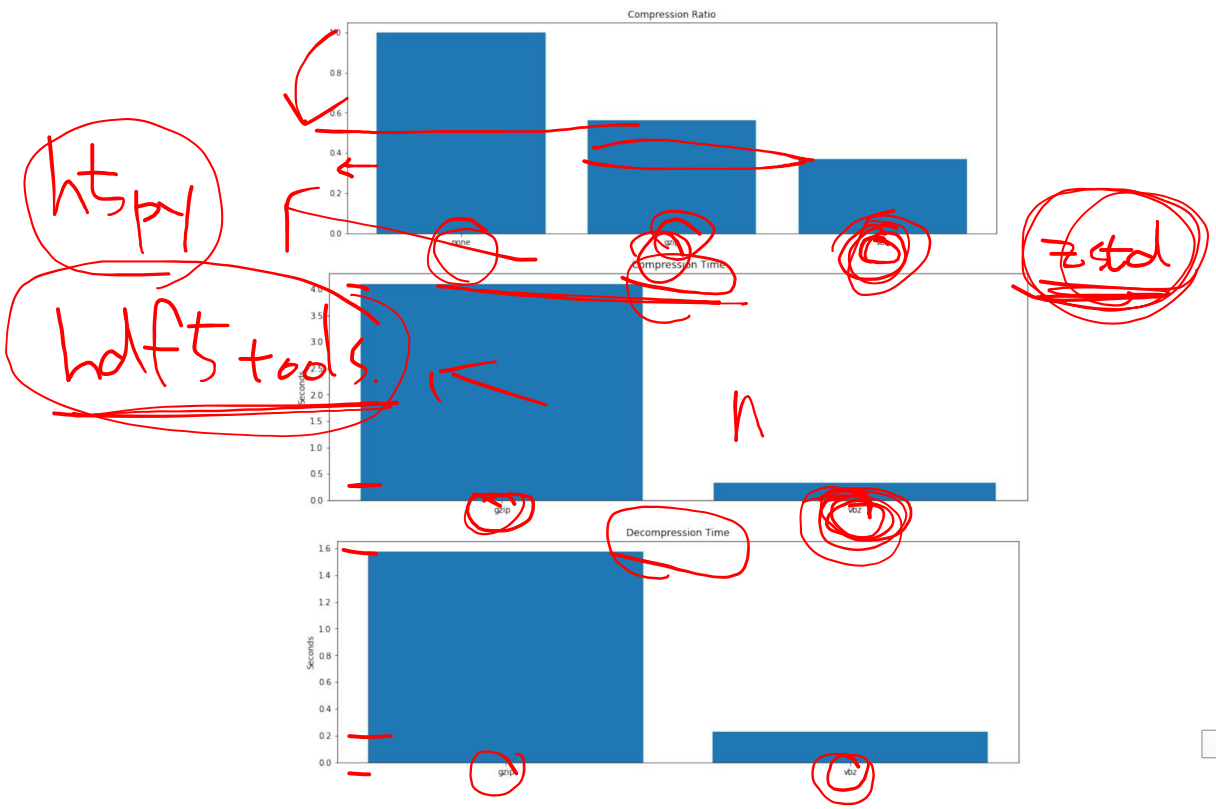

FAST5
Sigleread
HDFView 3.1.2
Recent Files
/Users/hyeshik/Desktop/0a9d2521-0917-4e38-a144-92c534f821ab.fast5.hdf5
Clear Text
0a9d2521-0917-4e38-a144-92c534f82
Analyses
Basecall_1D_000
Basecall_1D_001
BaseCalled_template
Fastq
Move
Trace
Summary
CoV_001
Segmentation_000
Segmentation_001
Summary
segmentation
Raw
Reads
Read_1479
Signal
UniqueGlobalKey
channel_id
context_tags
tracking_id
General Obj
Attribute Creation Order:
Creation Order NOT Tracked
Number of attributes = 11
Delete Attribute
Name
Type
Array Size
Value[50](...)
adapter_max
32-bit floating-point
Scalar
0.0
duration_template
64-bit unsigned integ...
Scalar
96602
first_sample_template
64-bit unsigned integ...
Scalar
555
has_complement
32-bit integer
Scalar
0
has_template
32-bit integer
Scalar
1
med_abs_dev_templa...
32-bit floating-point
Scalar
10.874455
median_template
32-bit floating-point
Scalar
87.27095
num_events_template
64-bit unsigned integ...
Scalar
6440
pt_detect_success
32-bit integer
Scalar
0
pt_median
32-bit floating-point
Scalar
0.0
pt_sd
32-bit floating-point
Scalar
0.0
Signal at /Ra...
0-based
0 853
1 524
2 511
3 517
4 524
5 529
6 529
7 532
8 528
9 527
10 527
11 512
12 528
13 526
14 526
15 534
16 524
17 515
18 517
19 523
20 495
User property file - /Users/hyeshik/.hdfview3.1.2
Signal at /Raw/Reads/Read_1479/ [0a9d2521-0917-4e38-a144-92c534f821ab.fast5.hdf5 in /Users/hyeshik/Desktop] [dims0, start0,

FAST5

multi read



FAST5 Compression

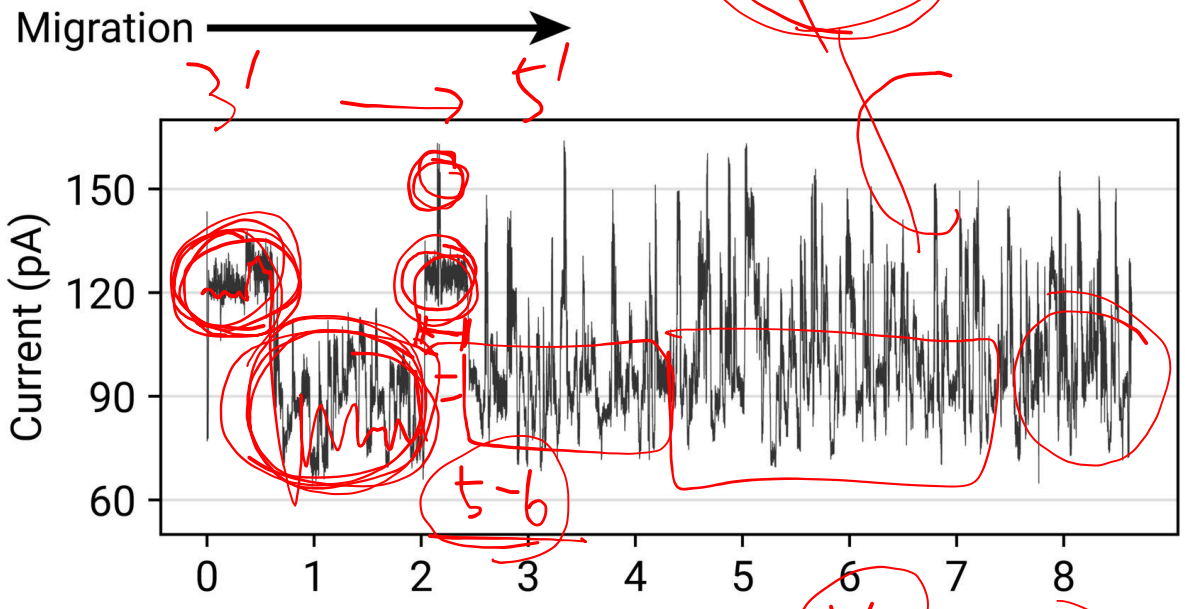


Practical workflows and tools for expression analysis using nanopore sequencing

Expression profiling using nanopore sequencing

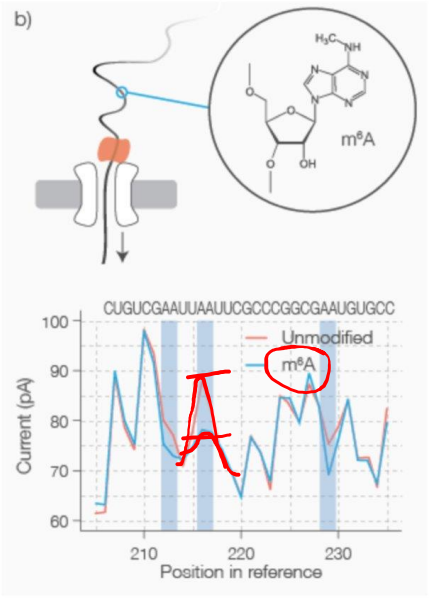
https://github.com/nanoporetech/pipeline-transcriptome-de/

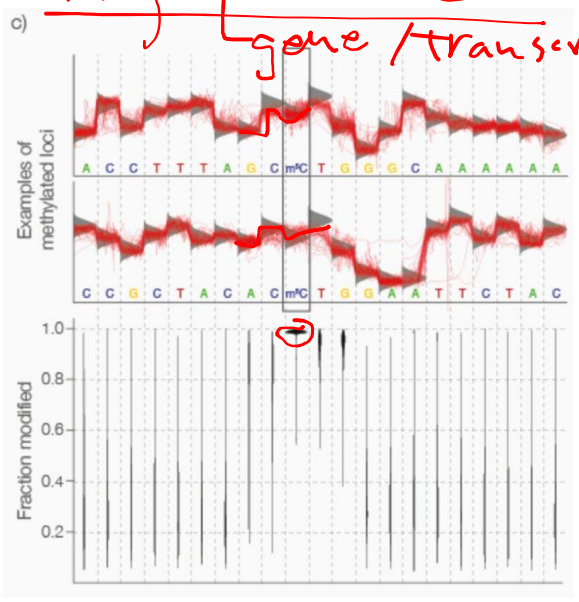
Poly(A) length measurement



tailfindr, nanopolish, poreplex

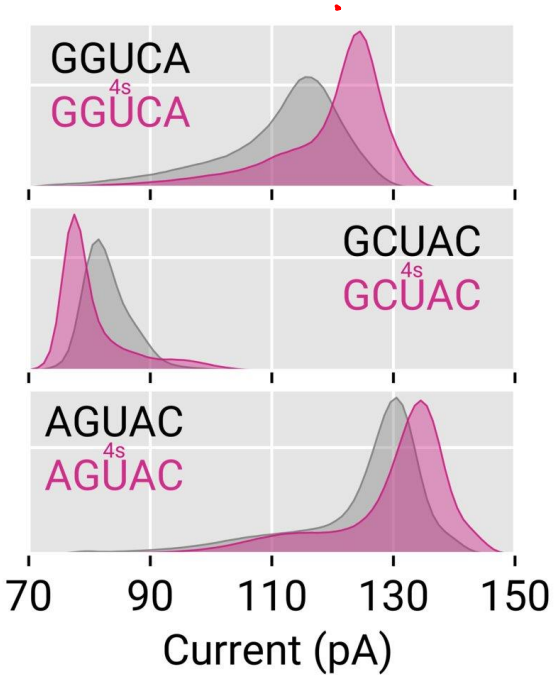
Calling RNA modifications



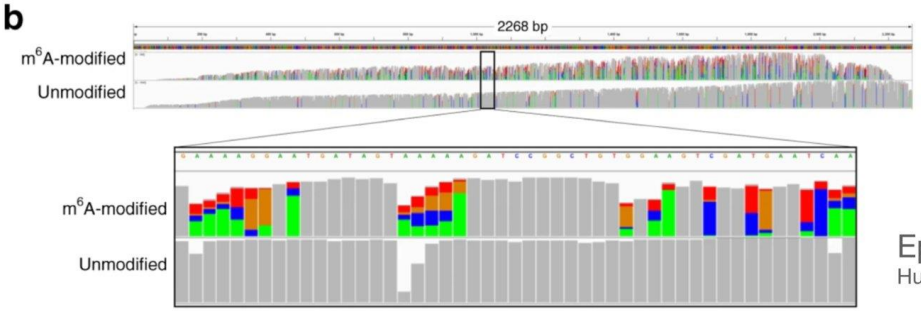


Oxford Nanopore Technologies

Calling RNA modifications

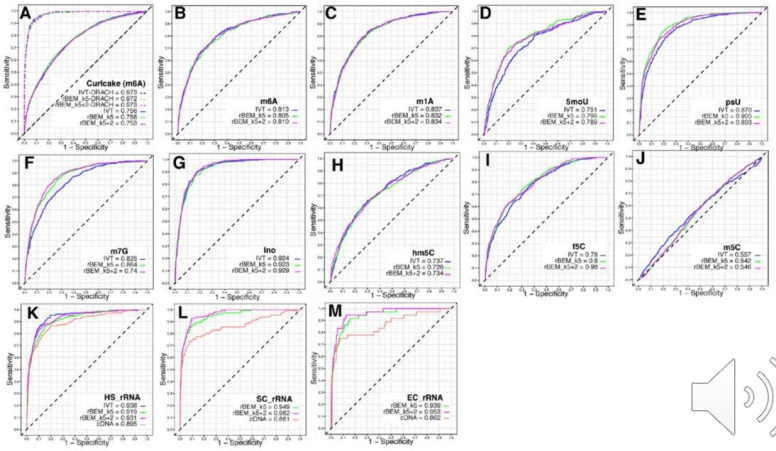


Calling RNA modifications



EpiNano
Huanle Liu et al. (2019) Nat Comms

ELIGOS2
Piroon Jenjaroenpun al. (2021) NAR

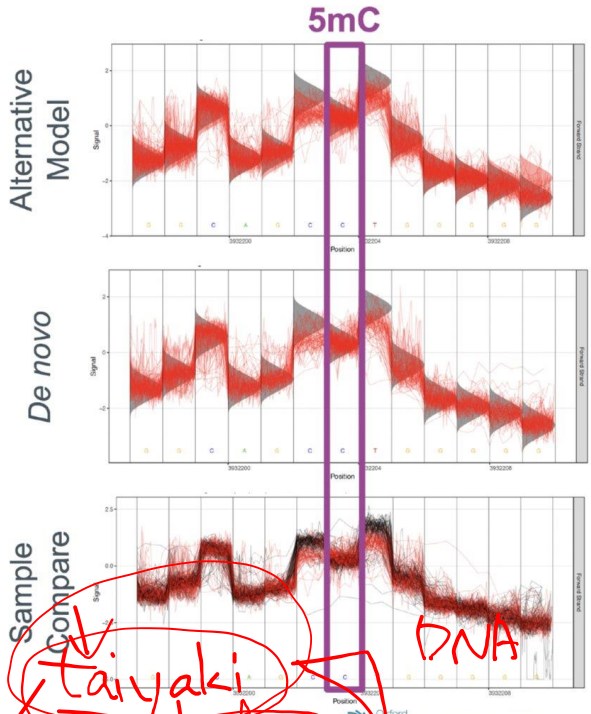


Calling RNA modifications

MODIFIED BASE DETECTION

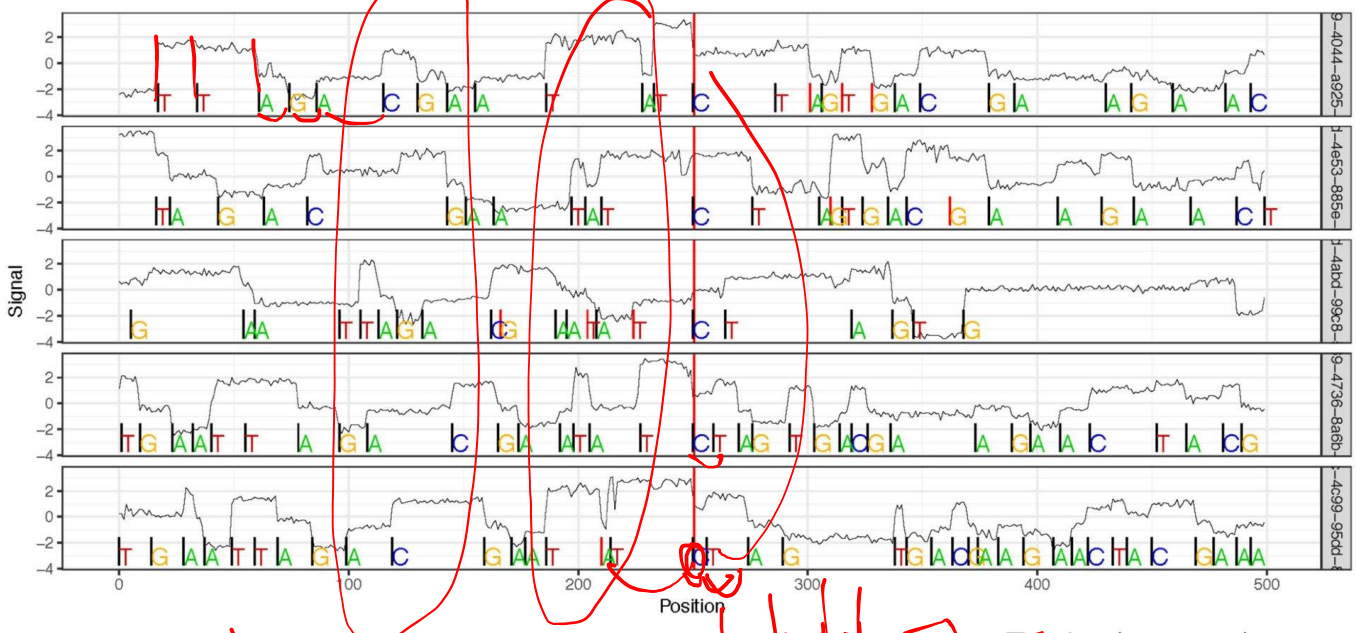
Tombo provides three methods for the identification of non-standard bases

	Advantages	Disadvantages
Alternative Model	• Known alt. base • Exact alt. loc. • Good accuracy	• Requires alt. model estimation
De novo	• Apply to any sample	• High error rate • Inexact location • Alt. base unknown
Sample Compare	• Best AUC • Most robust	• Inexact location • Alternative base unknown

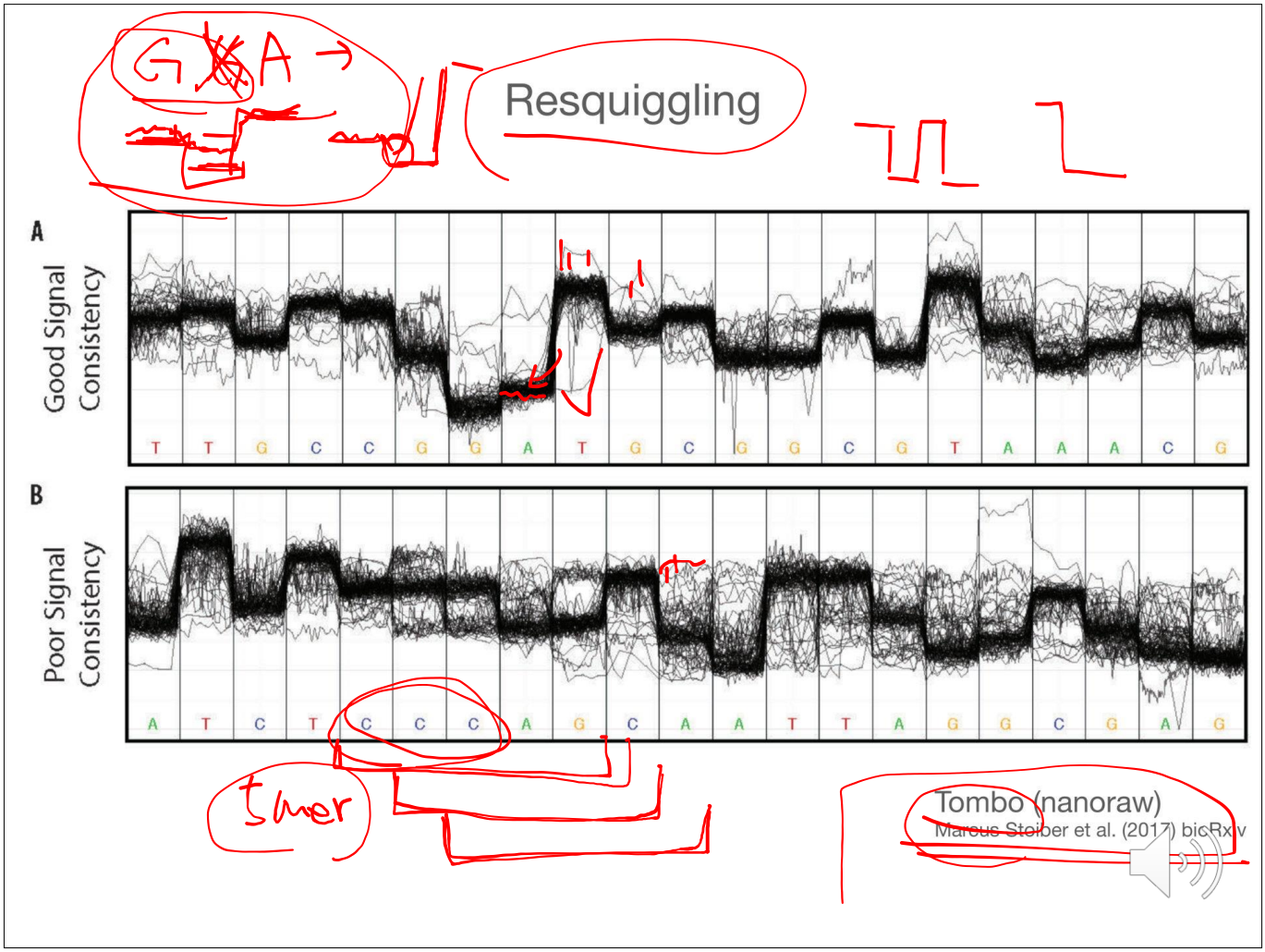


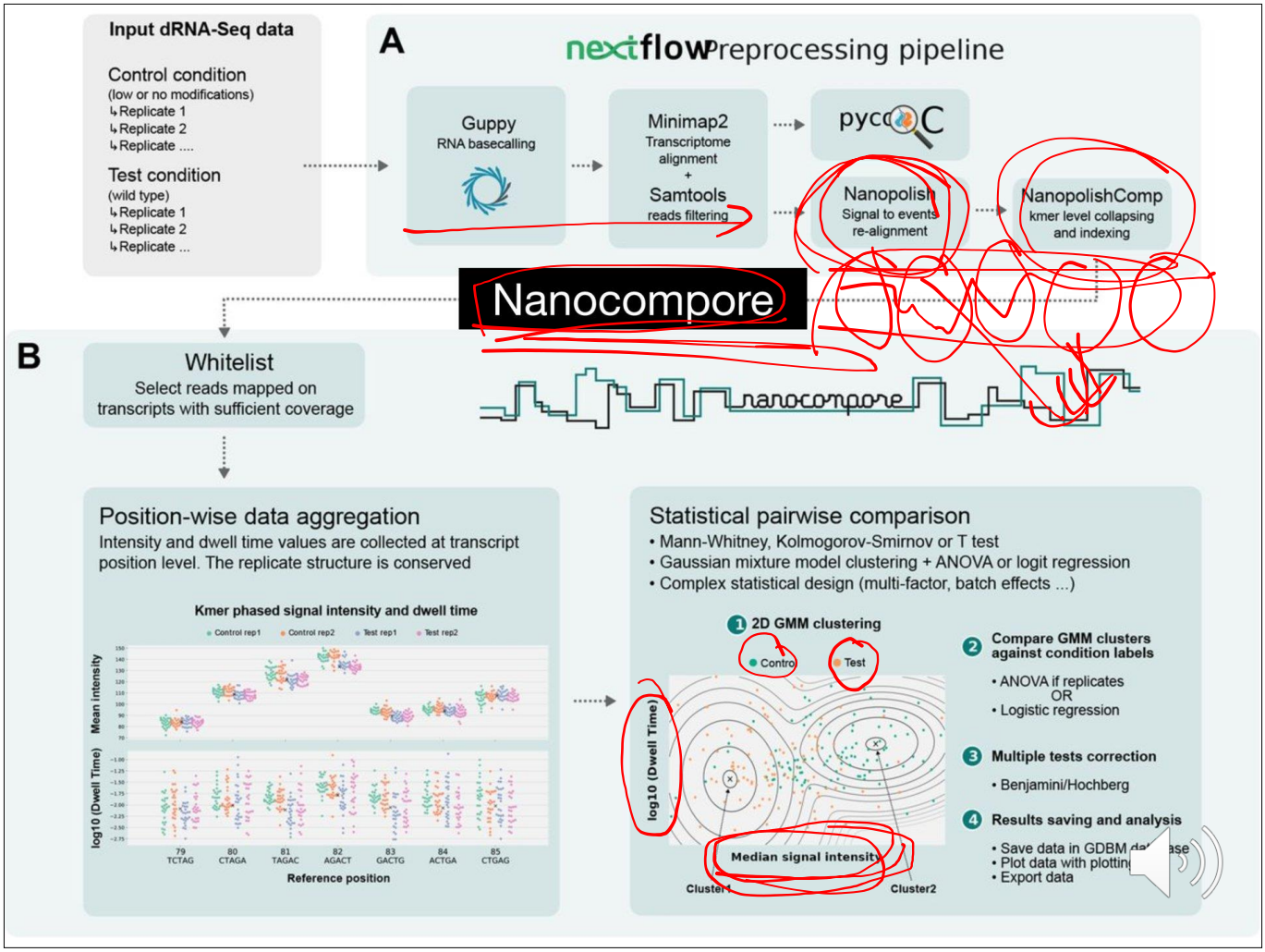
Tombo
Marcus Stoiber et al. (2017) bioRxiv

Resquiggling



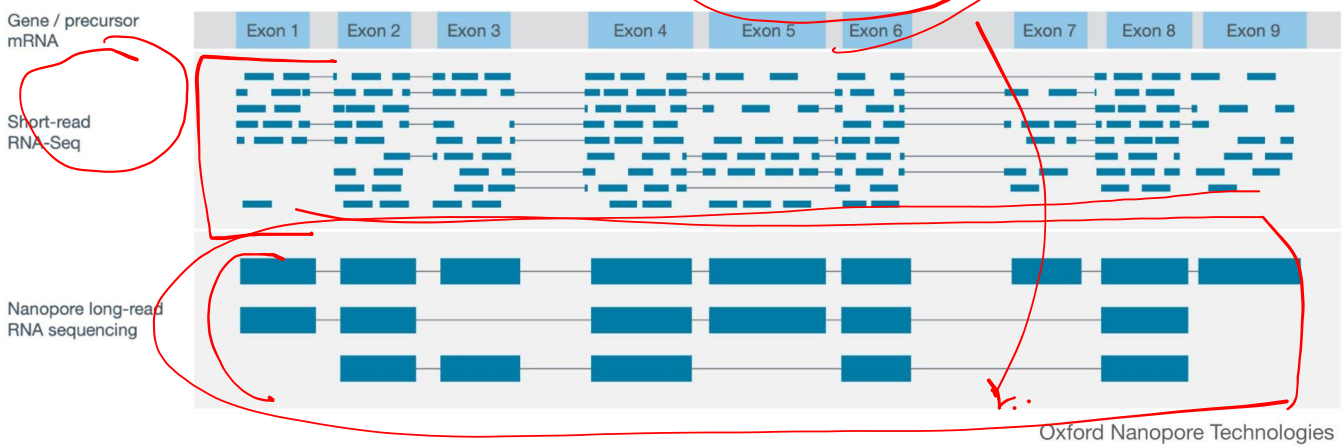
Tombo (nanoraw)
Marcus Stoiber et al. (2017) bioRxiv


Resquiggling
A
Good Signal Consistency
T T G C C G G A T G C G G C G T A A A C G
B
Poor Signal Consistency
A T C T C C C A G C A A T T A G G C G A G
Tombo (nanoraw)
Marcus Stoiber et al. (2017) bioRxiv

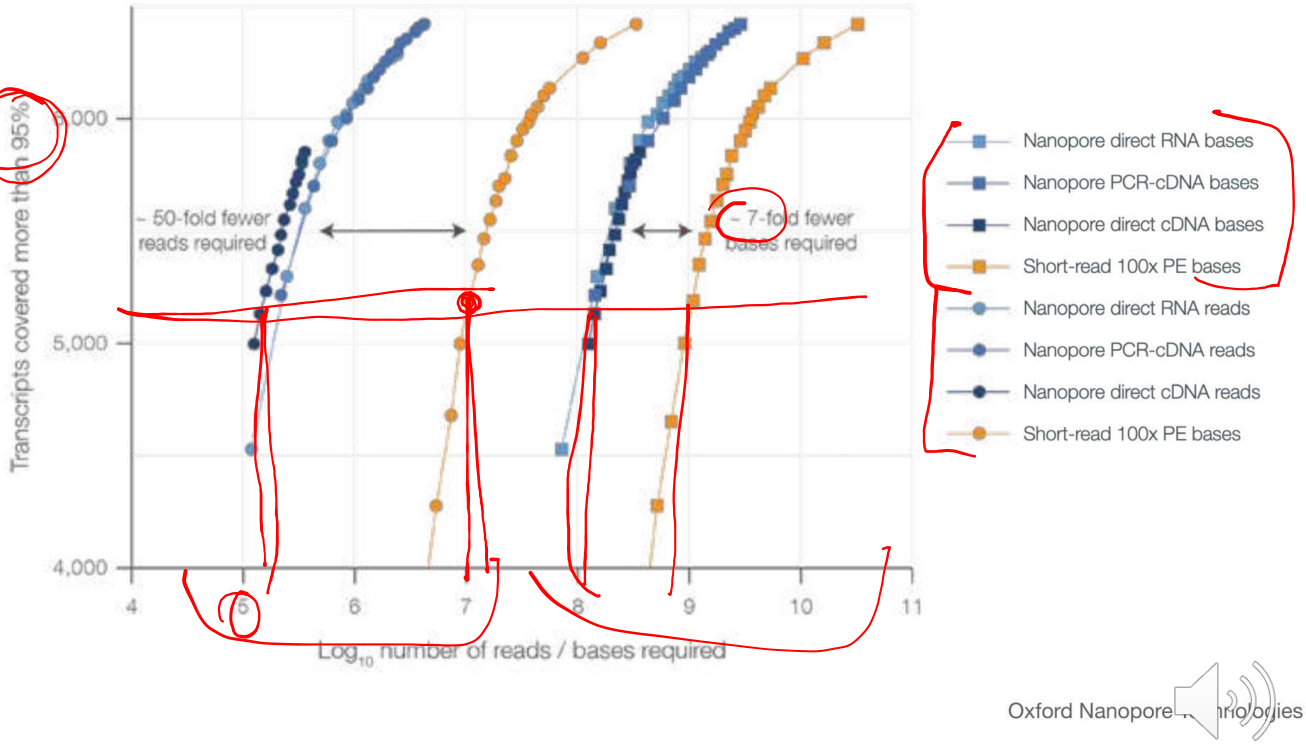

Input dRNA-Seq data
Control condition
(low or no modifications)
↳Replicate 1
↳Replicate 2
↳Replicate ....
Test condition
(wild type)
↳Replicate 1
↳Replicate 2
↳Replicate ...
A
nextflow Preprocessing pipeline
Guppy
RNA basecalling
Minimap2
Transcriptome alignment
+
Samtools
reads filtering
pycoQC
Nanopolish
Signal to events re-alignment
NanopolishComp
kmer level collapsing and indexing
Nanocompore
B
Whitelist
Select reads mapped on transcripts with sufficient coverage
nanocompore
Position-wise data aggregation
Intensity and dwell time values are collected at transcript position level. The replicate structure is conserved
Kmer phased signal intensity and dwell time
Control rep1
Control rep2
Test rep1
Test rep2
Mean intensity
log10 (Dwell Time)
79 TCTAG
80 CTAGA
81 TAGAC
82 AGACT
83 GACTG
84 ACTGA
85 CTGAG
Reference position
Statistical pairwise comparison
• Mann-Whitney, Kolmogorov-Smirnov or T test
• Gaussian mixture model clustering + ANOVA or logit regression
• Complex statistical design (multi-factor, batch effects ...)
1 2D GMM clustering
Control
Test
log10 (Dwell Time)
Median signal intensity
Cluster1
Cluster2
2 Compare GMM clusters against condition labels
• ANOVA if replicates
OR
• Logistic regression
3 Multiple tests correction
• Benjamini/Hochberg
4 Results saving and analysis
• Save data in GDBM database
• Plot data with plotting
• Export data

Which one to use?

Short-read vs ONT



Short-read vs ONT



Oxford Nanopore Technologies

DO NOT USE ONT direct RNA sequencing

DO NOT USE ONT direct RNA sequencing

(yet) for:

DO NOT USE ONT direct RNA sequencing

(yet) for:

- small RNA (miRNA, snoRNA, etc.)

DO NOT USE ONT direct RNA sequencing

(yet) for:

- small RNA (miRNA, snoRNA, etc.)
- Poly(A)⁻ RNAs with variable 3' ends

DO NOT USE ONT direct RNA sequencing

(yet) for:

- small RNA (miRNA, snoRNA, etc.)
- Poly(A)⁻ RNAs with variable 3' ends

- <100 ng input RNA

DO NOT USE ONT direct RNA sequencing

(yet) for:

- small RNA (miRNA, snoRNA, etc.)
- Poly(A)⁻ RNAs with variable 3' ends
- <100 ng input RNA
- Study objective requires very high base-level quality

Use ONT direct RNA sequencing for:

Use ONT direct RNA sequencing for:

- Quantification of poly(A)$^+$ RNAs

Use ONT direct RNA sequencing for:

- Quantification of poly(A)$^+$ RNAs
- Isoform-level analyses

Use ONT direct RNA sequencing for:

- Quantification of poly(A)$^+$ RNAs
- Isoform-level analyses
- When without a high quality reference

Use ONT direct RNA sequencing for:

- Quantification of poly(A)$^+$ RNAs
- Isoform-level analyses
- When without a high quality reference
- Detecting modified bases

Use ONT direct RNA sequencing for:

- Quantification of poly(A)$^+$ RNAs
- Isoform-level analyses
- When without a high quality reference
- Detecting modified bases
- Discovery of long-range associations

Use ONT direct RNA sequencing for:

- Quantification of poly(A)$^+$ RNAs
- Isoform-level analyses
- When without a high quality reference
- Detecting modified bases
- Discovery of long-range associations
- Measuring poly(A) lengths

Use ONT direct RNA sequencing for:

- Quantification of poly(A)$^+$ RNAs
- Isoform-level analyses
- When without a high quality reference
- Detecting modified bases
- Discovery of long-range associations
- Measuring poly(A) lengths
- When a fast feedback is desirable