

4<sup>th</sup> Conference on Next Generation Sequencing for Adventitious Virus Detection in  
Biologics for Humans and Animals

Comparison of Long- and Short-Read Sequencing  
Methods for Recombinant Adeno-Associated Virus  
Sequence and Adventitious Agent Identification in a  
GMP Environment

Megan Gura, Ph.D.

Sr QC Virology Scientist

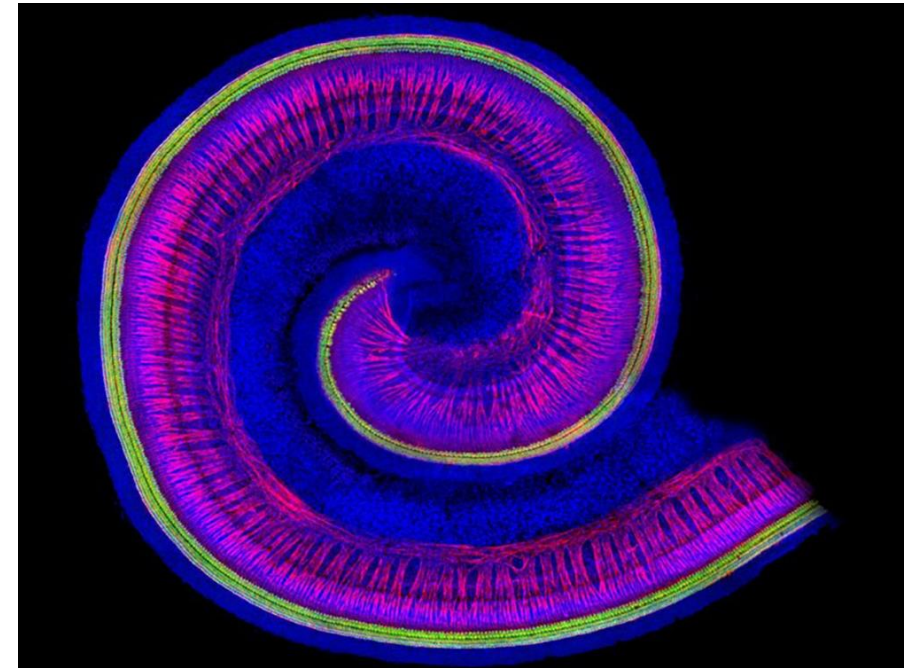
05Dec2024

**REGENERON**<sup>®</sup>

# Recombinant adeno-adenovirus (rAAV) is an emerging Cell & Gene Therapy treatment

rAAV therapies have the potential to bring medicine to previously untreatable illnesses

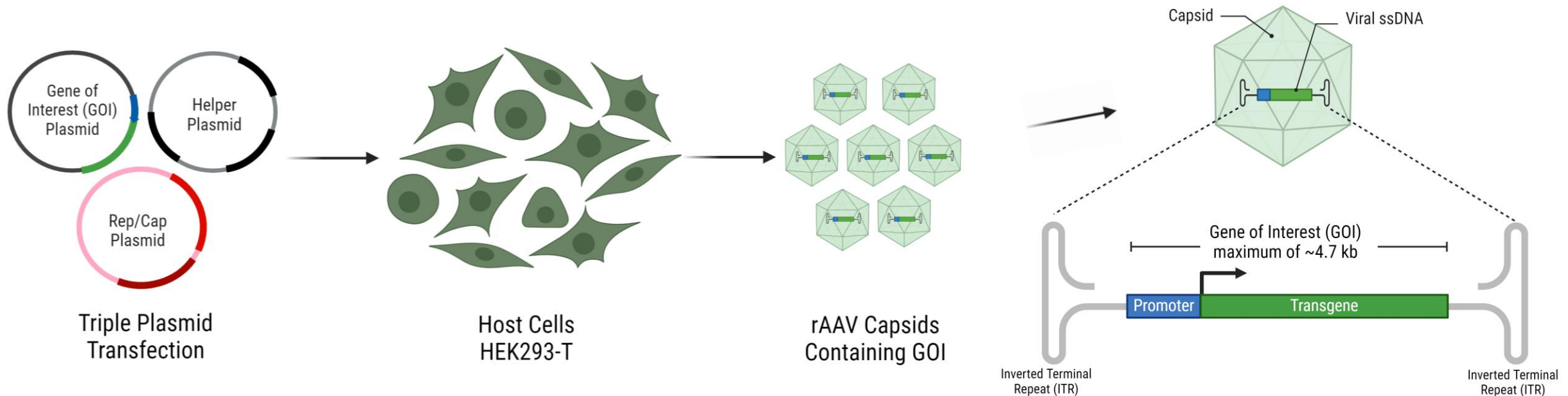
- In May 2024, the BBC reported on Opal, a UK girl born deaf, whose hearing was restored
  - Regeneron's first auditory program treats single-gene defect in otoferlin gene
  - Clinical trials ongoing



# Recombinant adeno-adenovirus (rAAV) is an emerging Cell & Gene Therapy treatment

Regeneron rAAVs are made through triple-transfection of HEK293-T cells

- rAAV is an encapsidated ssDNA virus that contains a gene of interest (GOI) between two inverted terminal repeats (ITRs)
- Critical Quality Attributes (CQAs) for rAAV products must be characterized including:
  - Sequence identity
  - Surveillance for contaminants
    - Non-target rAAV GOI sequences
    - Adventitious agents

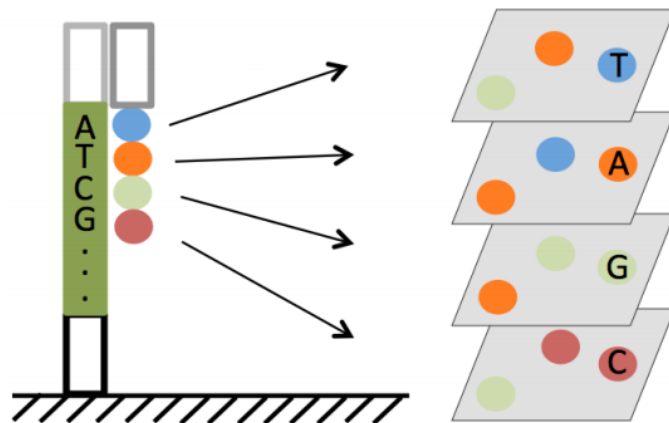


# Two sequencing approaches available in the Regeneron QC Virology laboratories

Illumina uses sequencing by synthesis, Nanopore pulls DNA strands through a channel

- Both NextSeq 2000 and PromethION use dsDNA as input

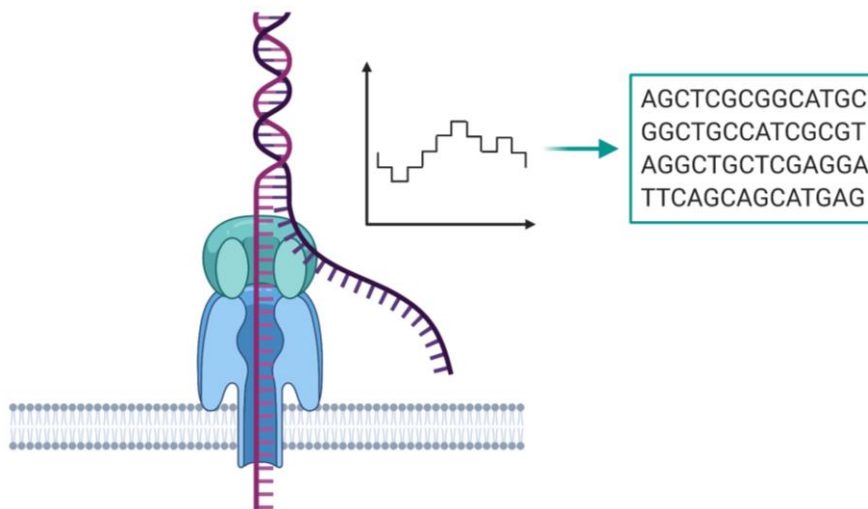
## Illumina NextSeq 2000



*Lu et al., 2016*

- Short-read (300 bp)
- Hundreds of million of reads
- 50-500 ng input DNA
- Extremely accurate (99%)

## Oxford Nanopore Technologies PromethION



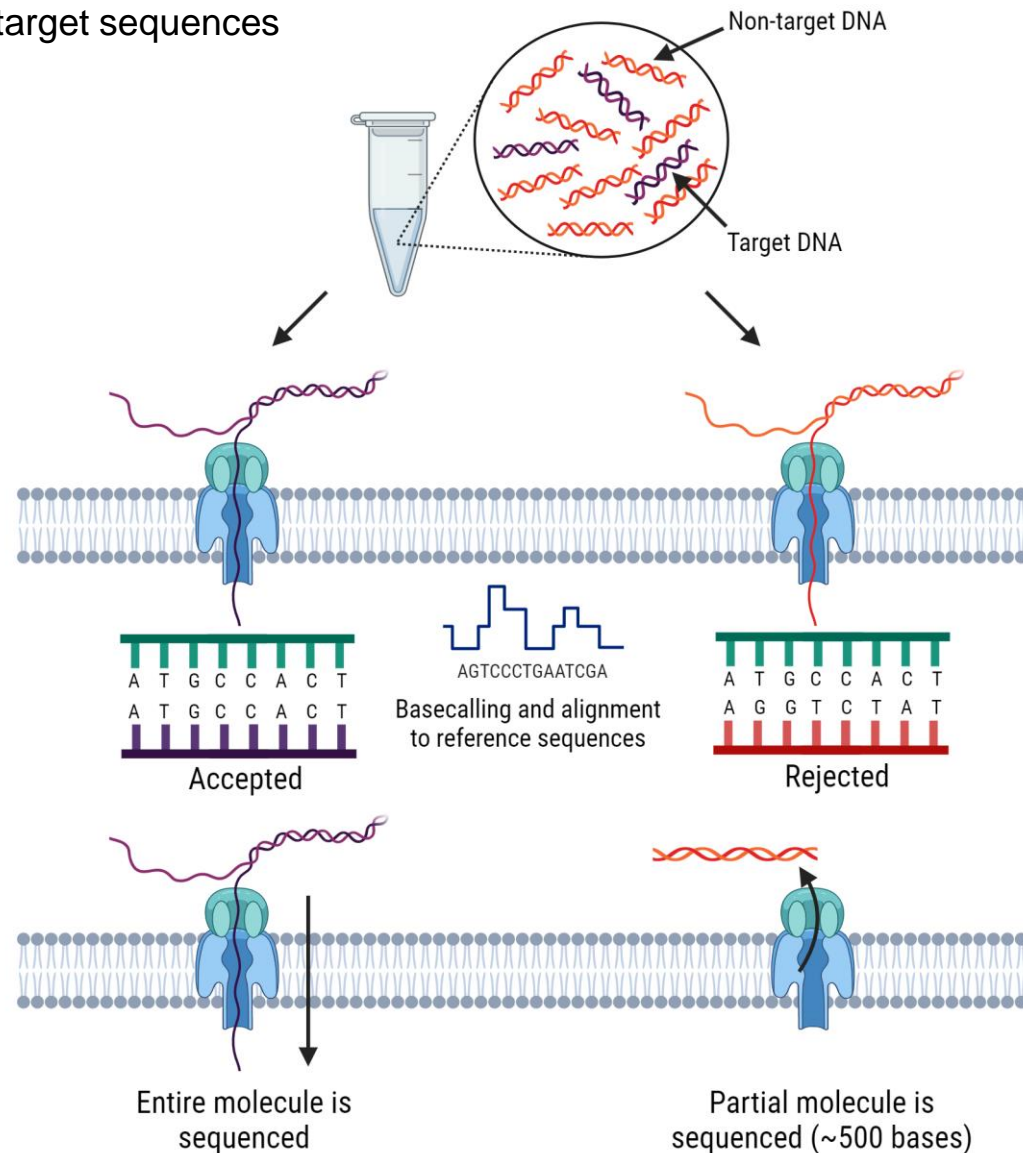
*BC Cancer Research Institute*

- Long-read (>1000 bp)
- Million of reads
- 1 µg input DNA
- Very accurate (92-97%)

# Adaptive sampling is a unique feature available on the PromethION

Illumina uses sequencing by synthesis, Nanopore pulls DNA strands through a channel

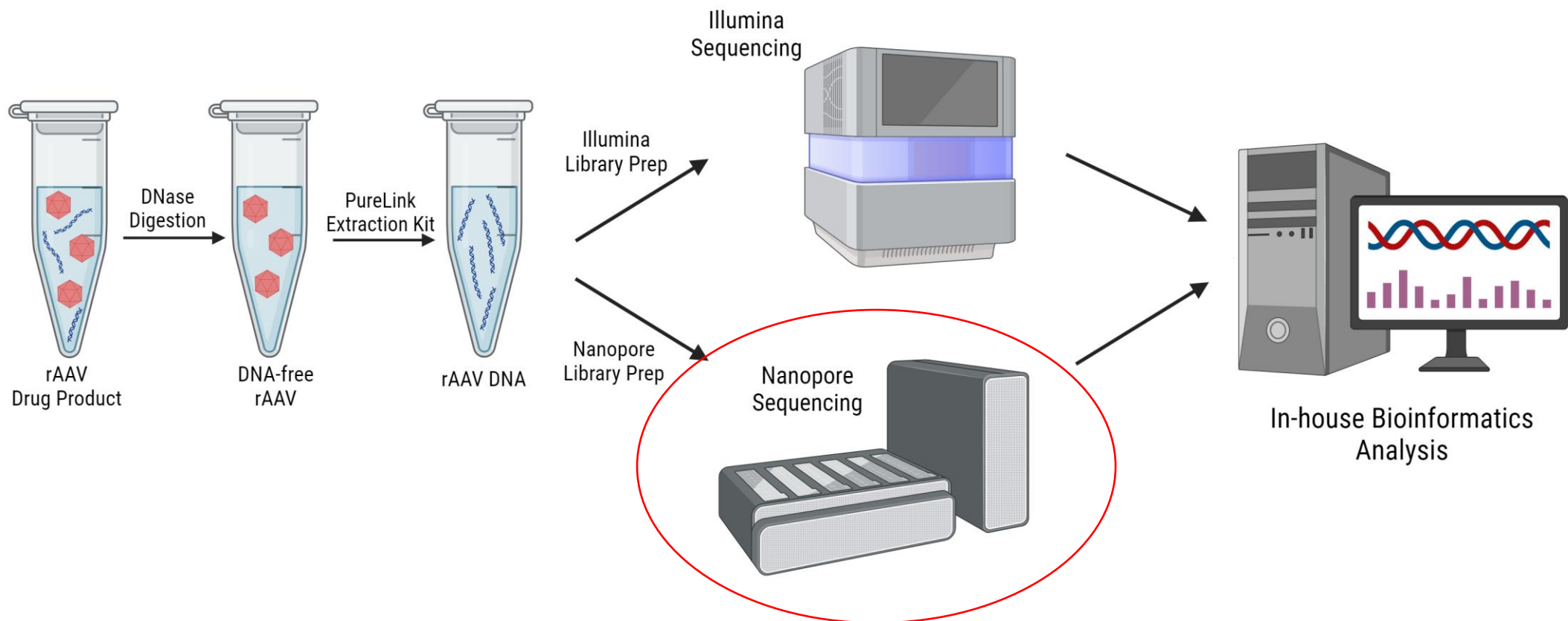
- Allows enrichment or depletion of target sequences



# rAAV experiments were performed on both NextSeq 2000 and PromethION

How do these sequencing platforms compare?

- Using the same extracted rAAV material, created sequencing libraries for both Illumina and Nanopore instruments
- How does normal sequencing on the PromethION compare to adaptive sampling of rAAV?



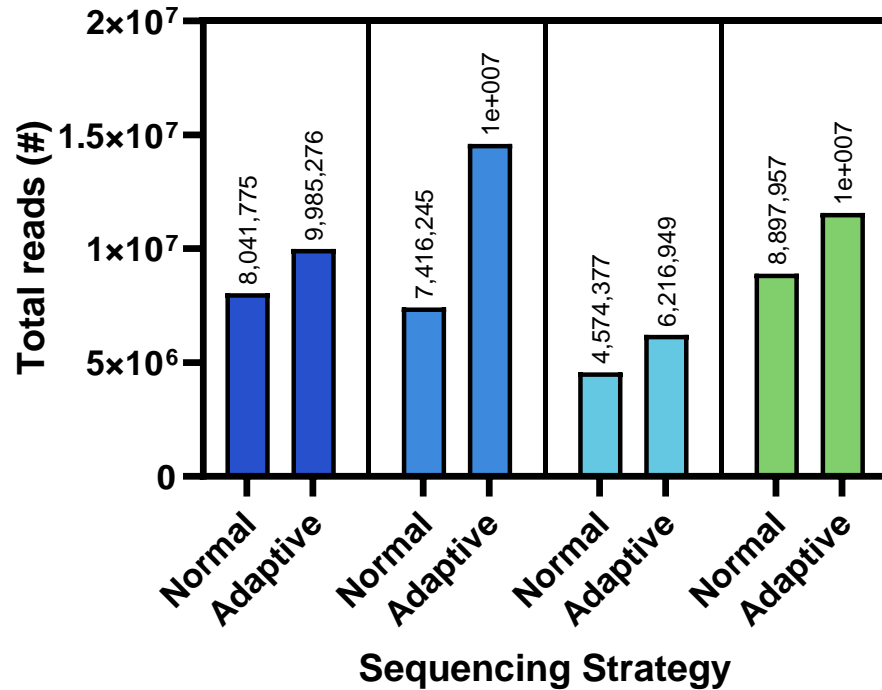
# Adaptive sampling yields more reads across multiple rAAV samples

Four different rAAV test articles sequenced

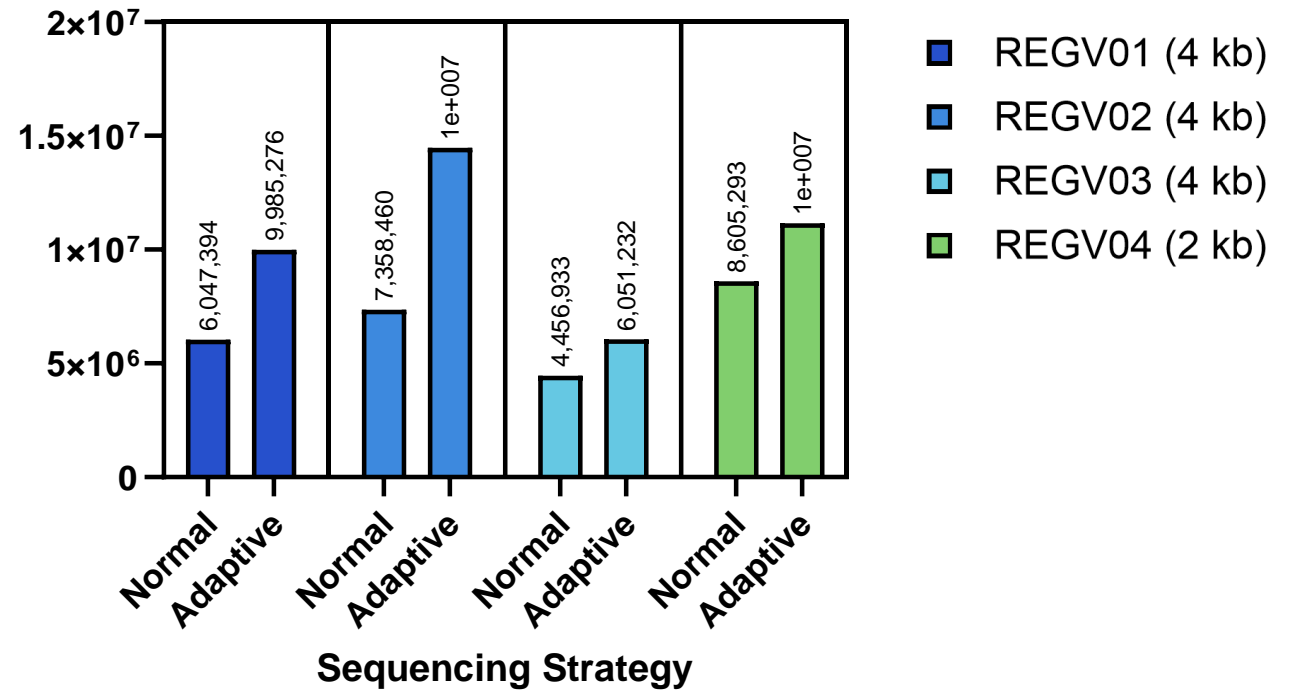
- Adaptive sampling increased total number of reads and rAAV reads



### Reads after filtering



### rAAV Reads



- REGV01 (4 kb)
- REGV02 (4 kb)
- REGV03 (4 kb)
- REGV04 (2 kb)

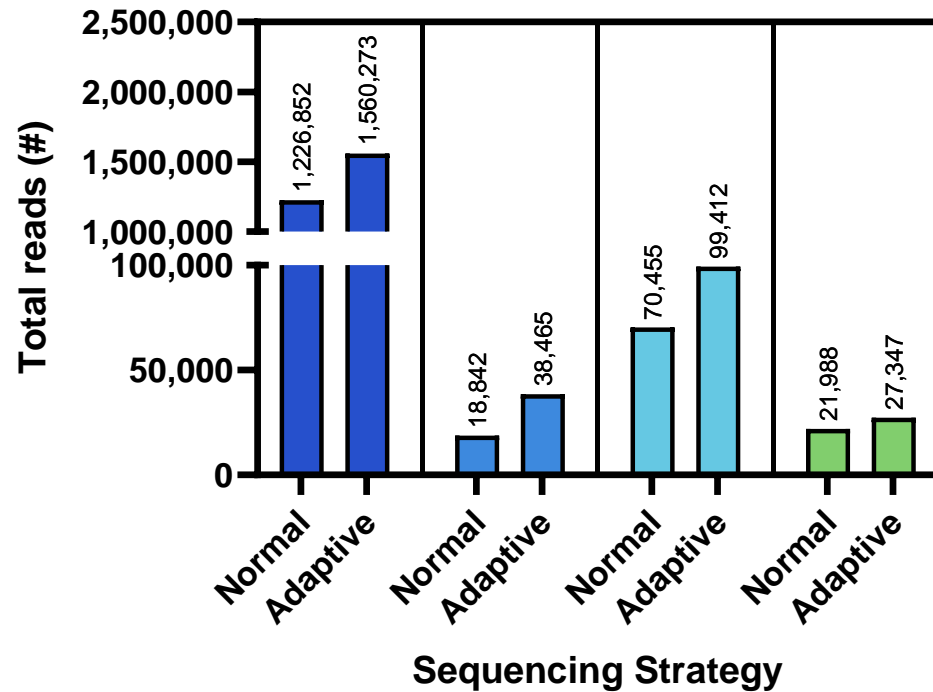
# Adaptive sampling yields more reads across multiple rAAV samples

Four different rAAV test articles sequenced

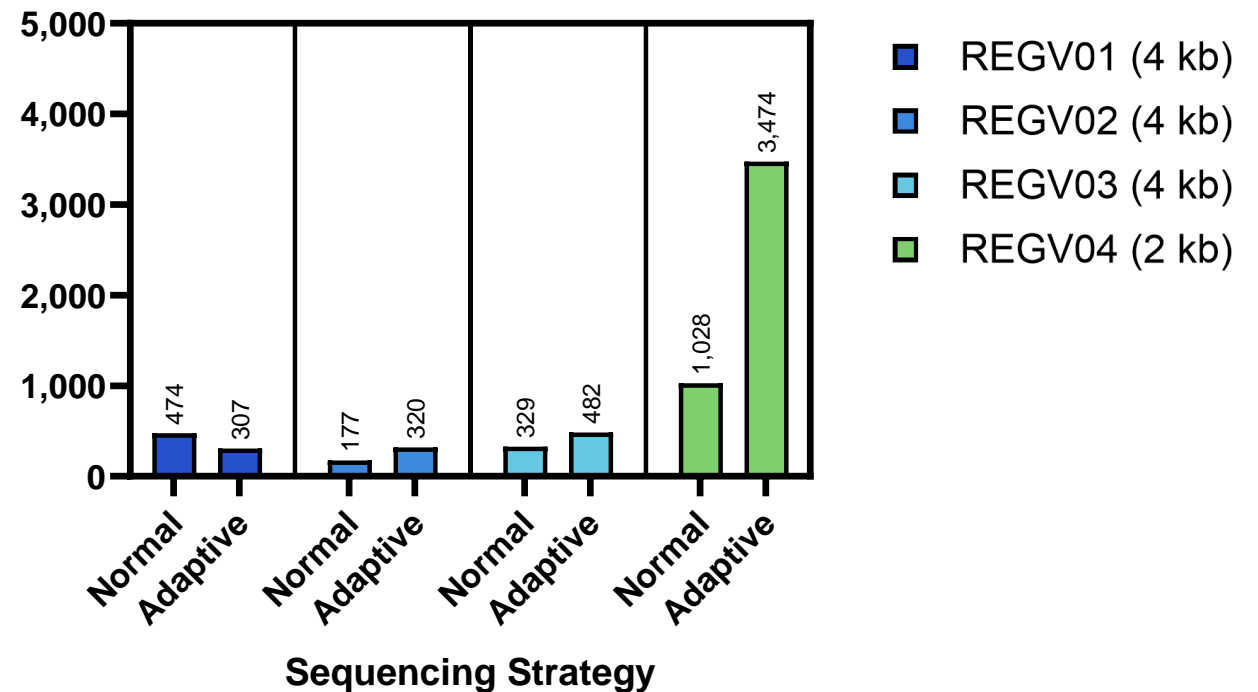
- Adaptive sampling also increased the number of non-rAAV reads detected
- Number of pores across flow cells was relatively equal between the normal and adaptive sampling samples



### Human Reads



### E. coli Reads



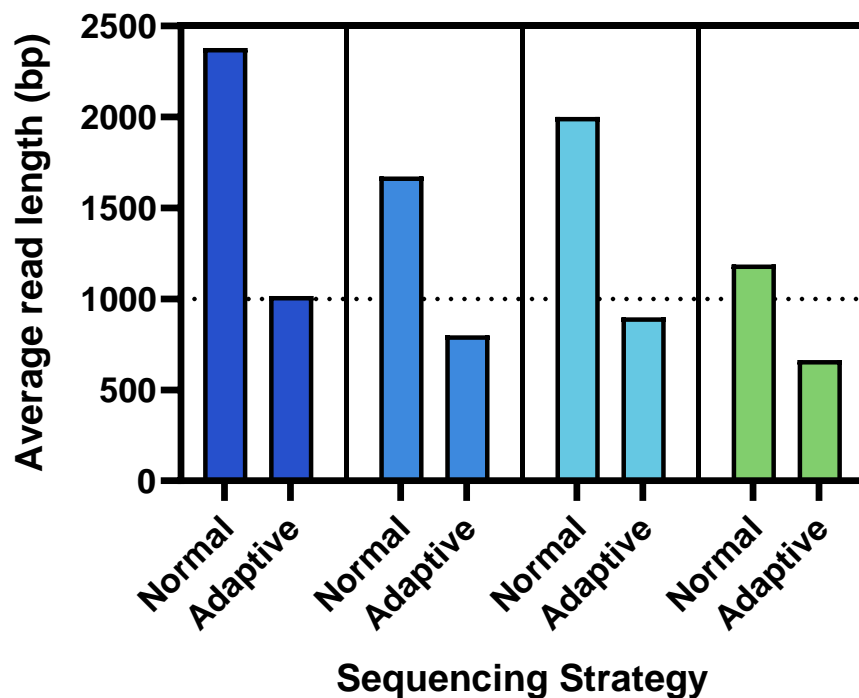
# Adaptive sampling yields more reads across multiple rAAV samples

Four different rAAV test articles sequenced

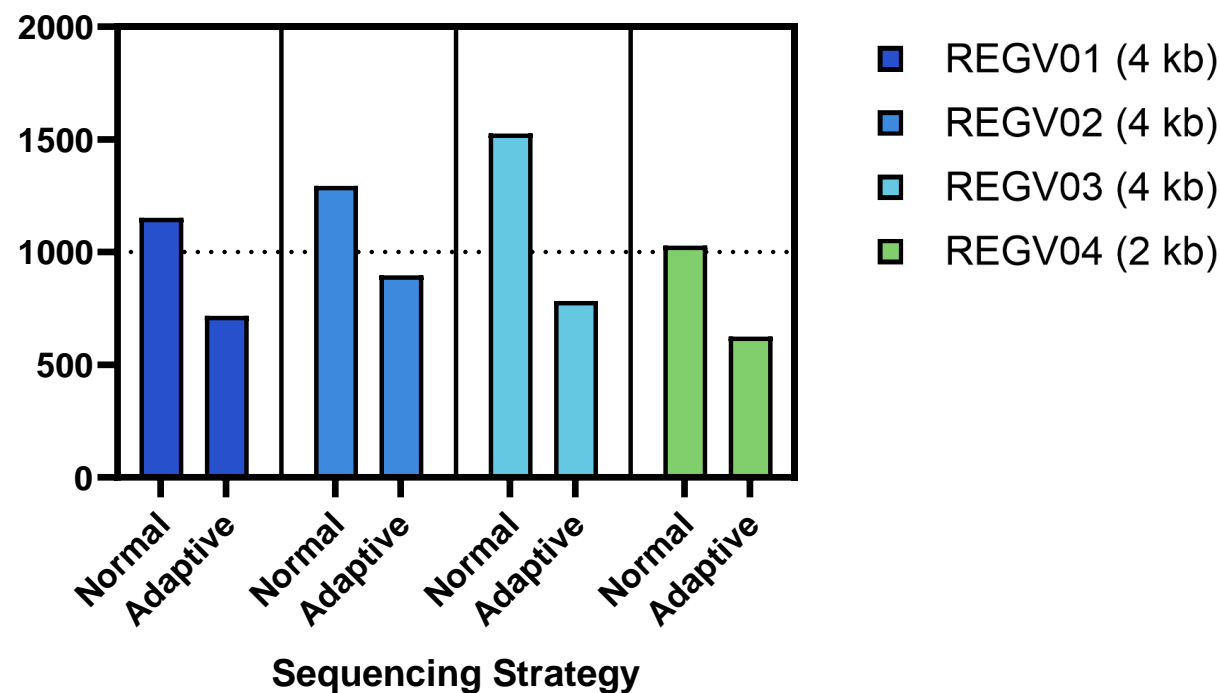
- Adaptive sampling also increased the number of non-rAAV reads detected
- Short non-rAAV reads predominantly sequenced



### Human Read Lengths



### *E. coli* Read Length



- REGV01 (4 kb)
- REGV02 (4 kb)
- REGV03 (4 kb)
- REGV04 (2 kb)

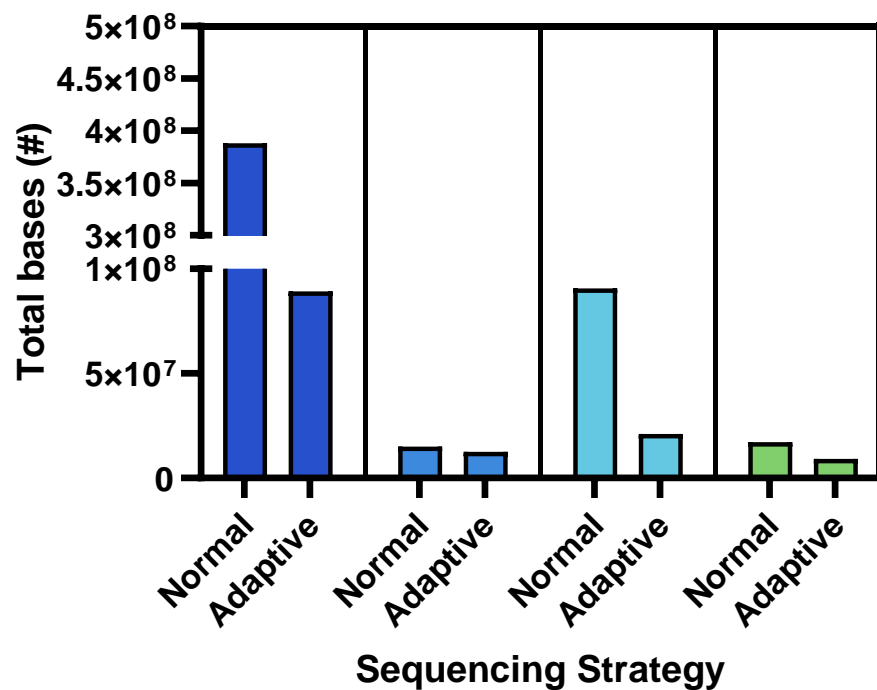
# Adaptive sampling yields more reads across multiple rAAV samples

Four different rAAV test articles sequenced

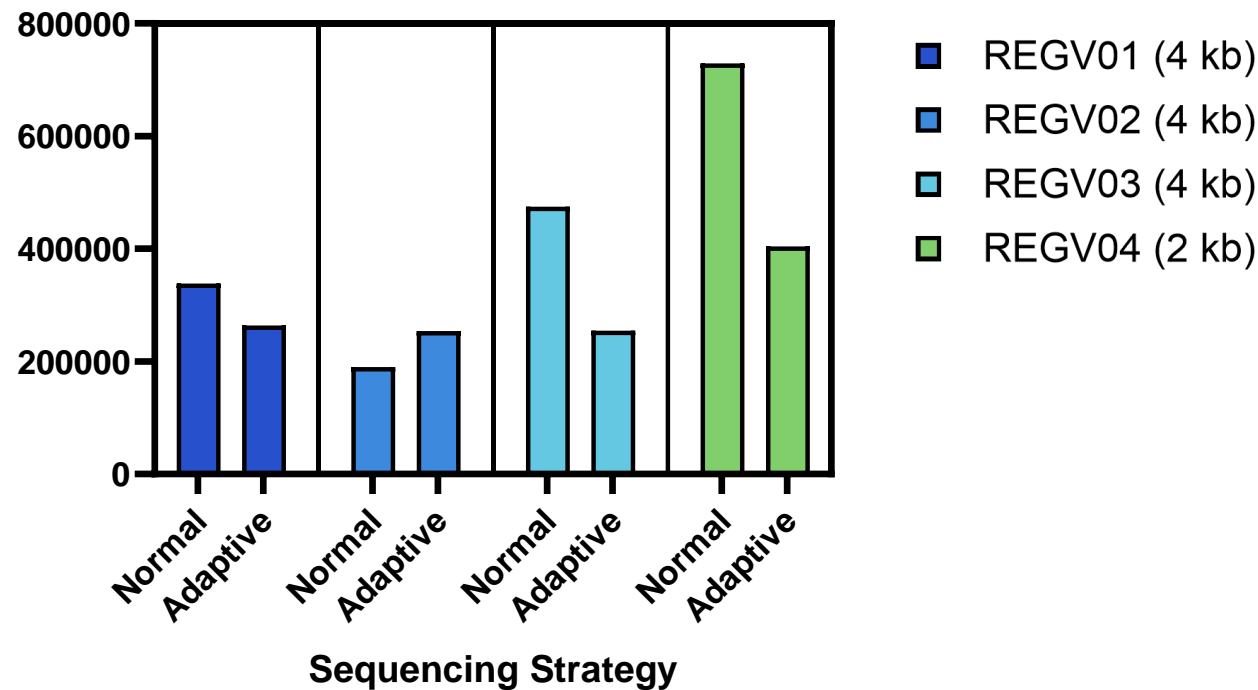
- Adaptive sampling also increased the number of non-rAAV reads detected
- Short non-rAAV reads were sequenced before adaptive sampling could reject the read
- Adaptive sampling reduces number of non-target bases



### Human Bases



### *E. coli* Bases



- REGV01 (4 kb)
- REGV02 (4 kb)
- REGV03 (4 kb)
- REGV04 (2 kb)

# Side-by-side comparison of Illumina and Nanopore sequencing approaches



(Adaptive Sampling!)



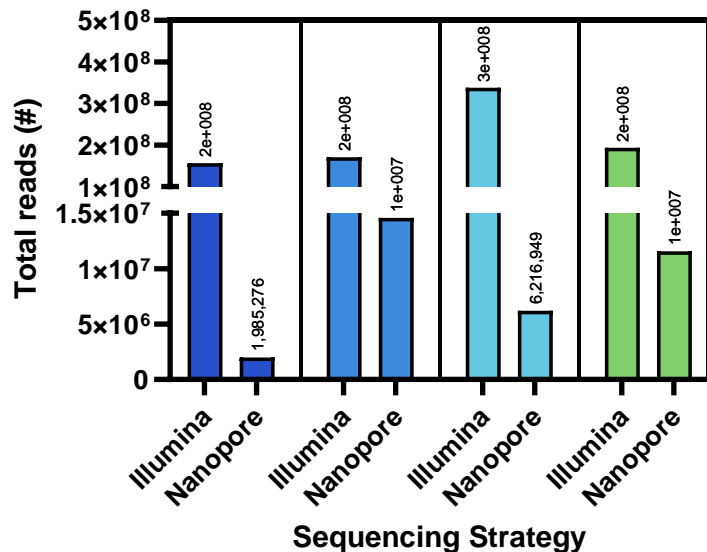
Quality	NextSeq 2000	PromethION
# Reads		
Read length		
Average read quality		
rAAV genome coverage		
FLIP/FLOP ratio		
rAAV cross-contaminant detection		
Variant/indel detection		
Cost		

# Illumina and Nanopore sequencing show their typical strengths

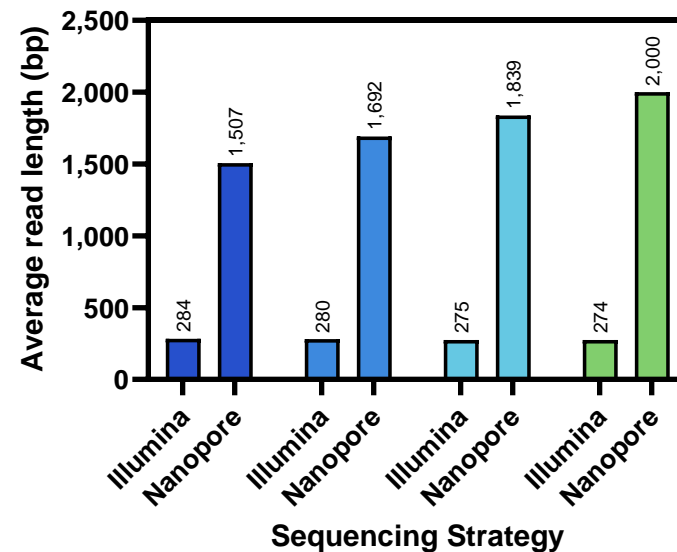
Interestingly, average read quality is not significantly better with Illumina sequencing

- Illumina produces far more reads than Nanopore
- Nanopore sequencing produces far longer reads than Illumina
- Average read quality approximately the same

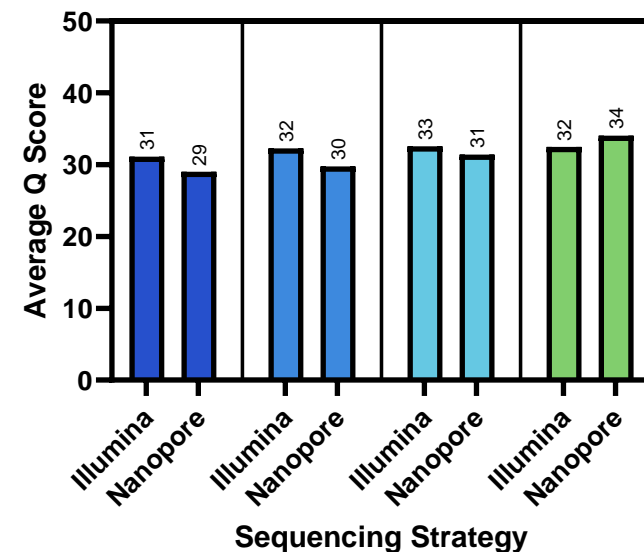
Reads after filtering



Comparing Aligned Read Length



Average Read Quality



- REGV01 (4 kb)
- REGV02 (4 kb)
- REGV03 (4 kb)
- REGV04 (2 kb)

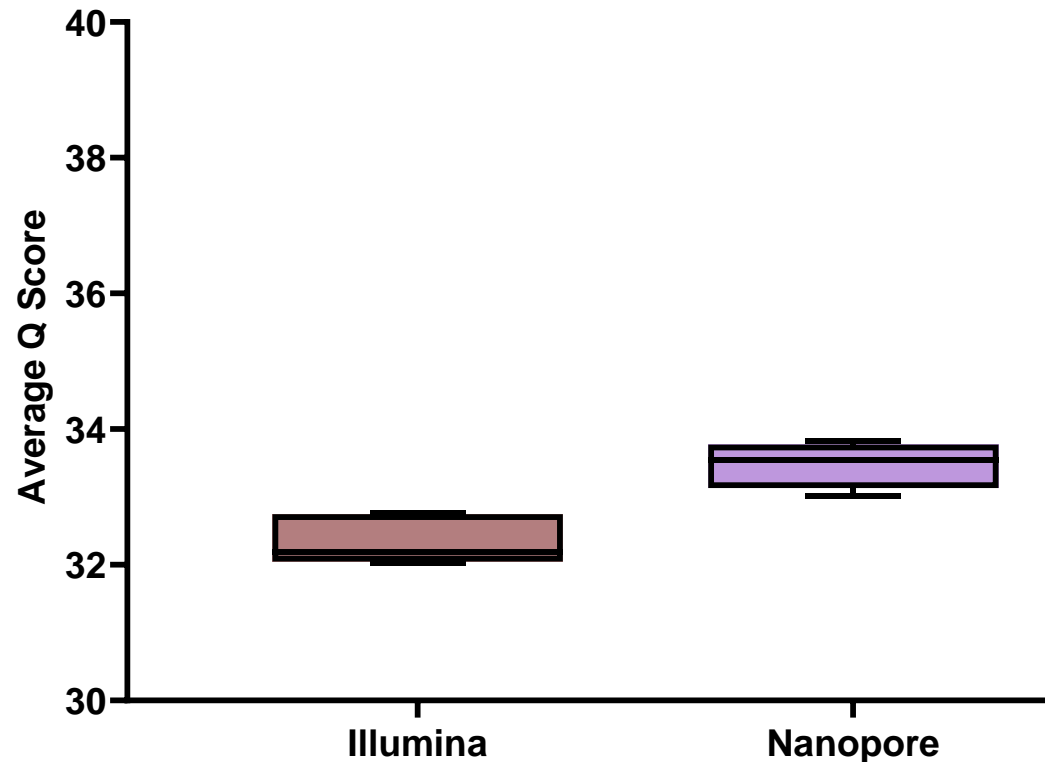
# New XLEAP-SBS chemistry pulls Illumina read quality back into the lead



Nanopore read quality outshined by the new Illumina chemistry

- In some experiments using NextSeq 2000 Standard SBS sequencing reagents, PromethION average read quality was slightly better

Old Chemistry versus Nanopore



# Side-by-side comparison of Illumina and Nanopore sequencing approaches



(Adaptive Sampling!)

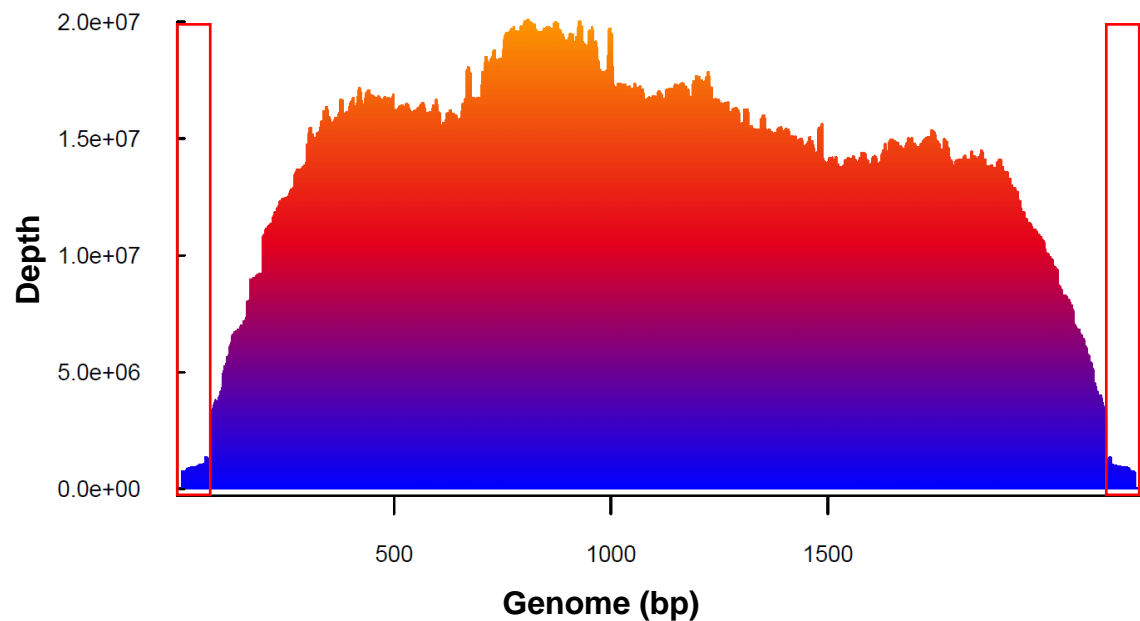


Quality	NextSeq 2000	PromethION
# Reads	✓	✗
Read length	✗	✓
Average read quality	✓	✗
rAAV genome coverage		
FLIP/FLOP ratio		
rAAV cross-contaminant detection		
Variant/indel detection		
Cost		

# Illumina sequencing struggles with ITRs, but still achieves sufficient coverage

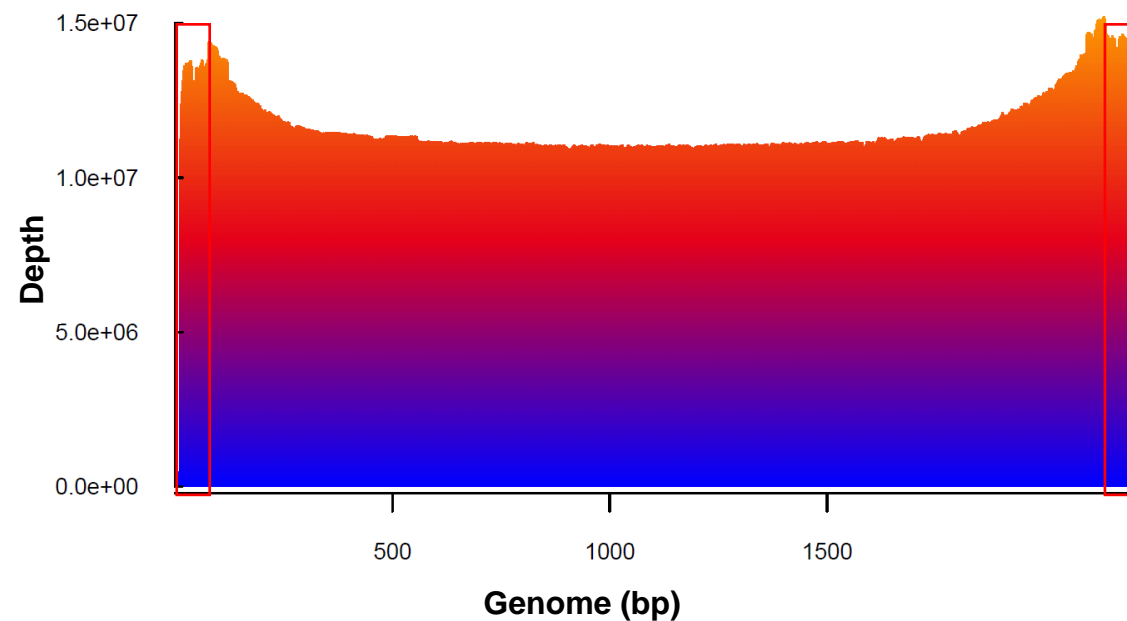
Minimum coverage criteria: 1000X average coverage, 10X per base coverage

## Illumina



- Reference Coverage: 100%
- Average Depth: ~14 M
- Minimum Depth: ~42 K (3'ITR)

## Nanopore

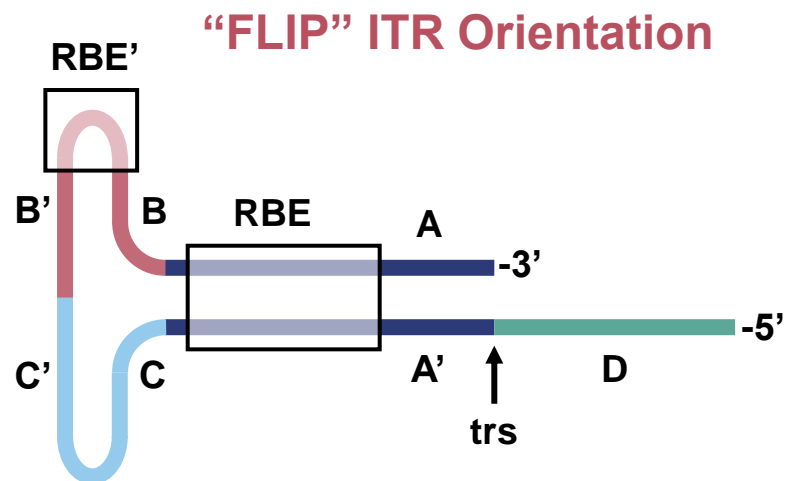


- Reference Coverage: 100%
- Average Depth: ~12 M
- Minimum Depth: ~360 K (3'ITR)

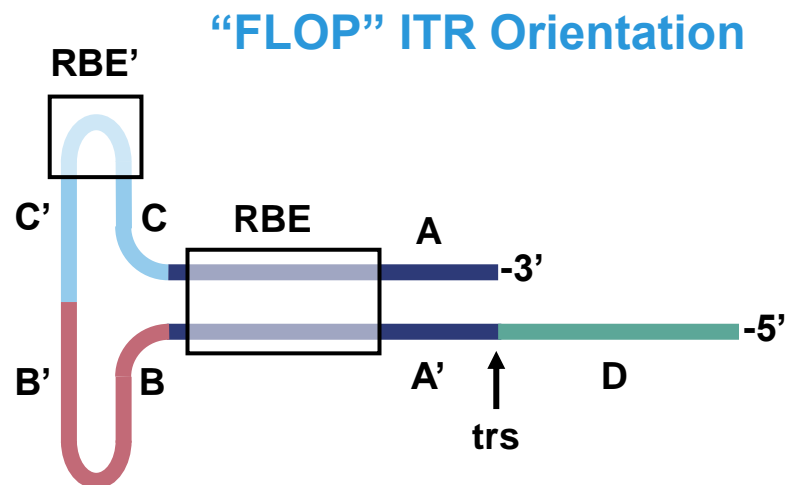
# rAAV ITRs come in two possible versions: FLIP & FLOP

Four rAAV genome orientations are expected in rAAV drug product

- Only remaining region of genome from WT, required for rAAV replication and genome insertion
- Methylation status of ITRs affects host immunity



25%	FLIP	-AAV	FLIP
25%	FLIP	-AAV	FLOP
25%	FLOP	-AAV	FLIP
25%	FLOP	-AAV	FLOP



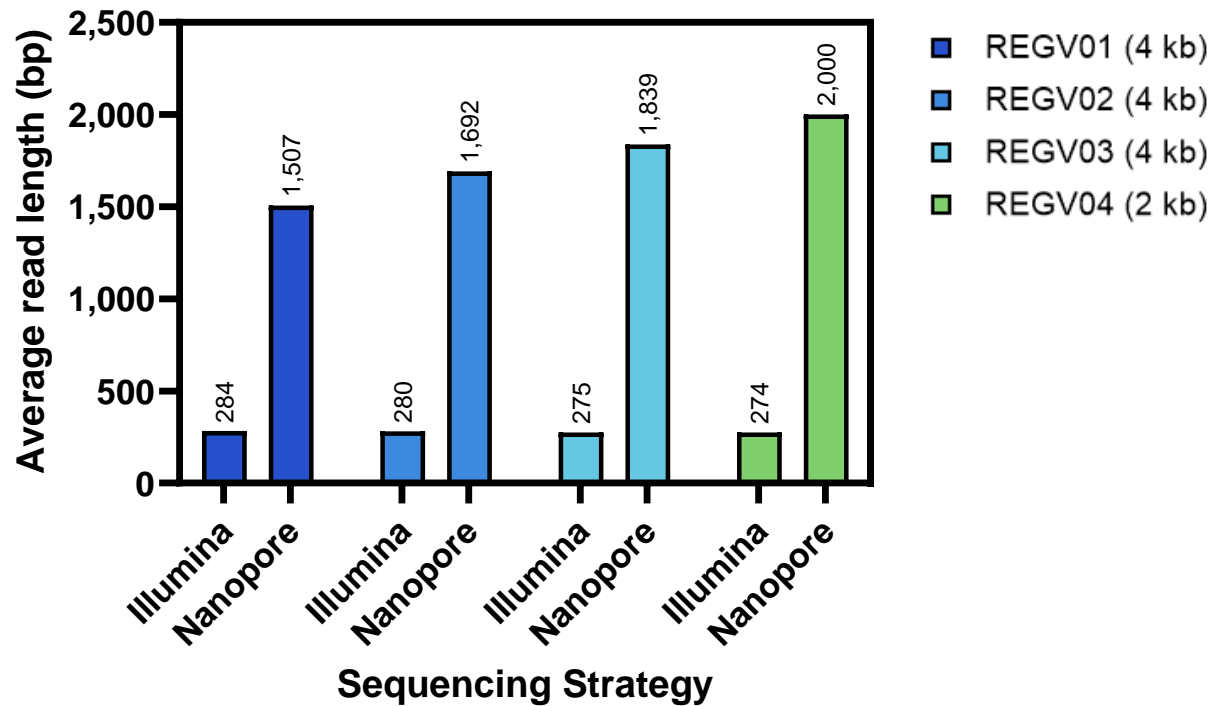
# FLIP/FLOP ratio quantification simpler and easier with Nanopore

Long Nanopore reads can uniquely map to the FLIP/FLOP variants

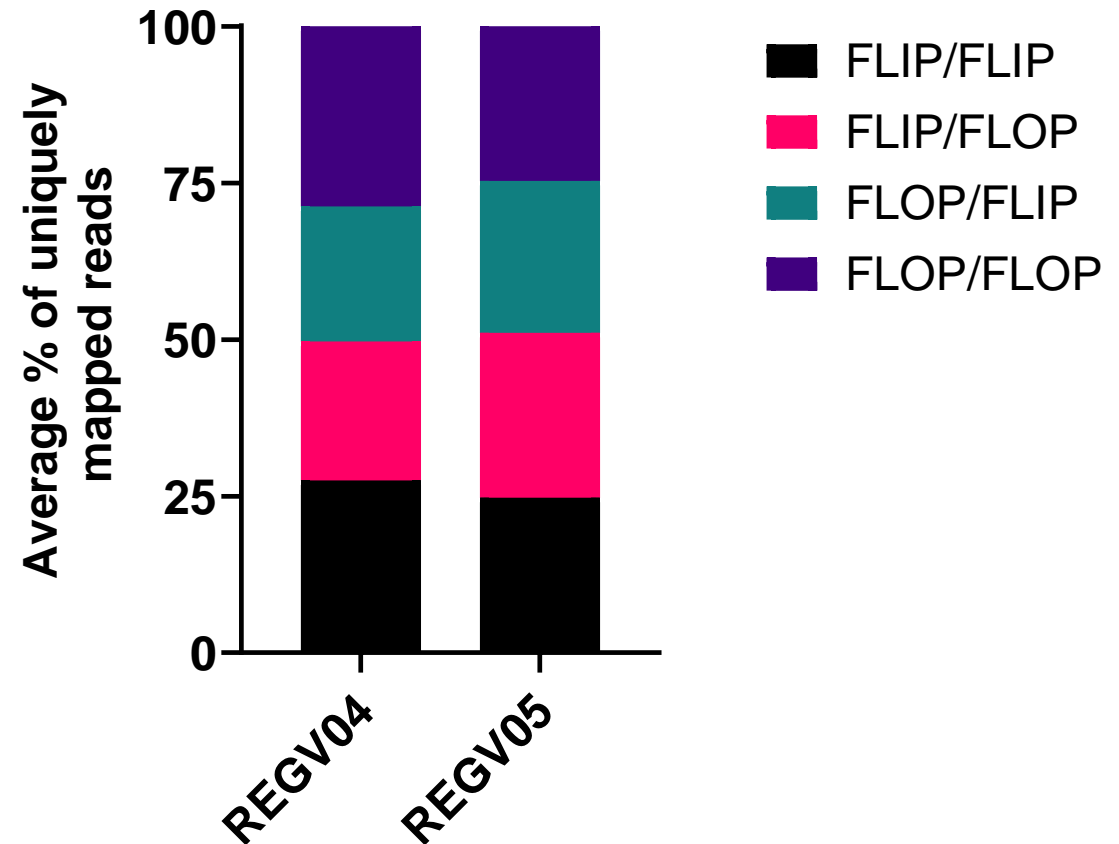
- Illumina reads are ~280 bp long, not possible to sequence rAAV ITR-to-ITR
  - May be able to be bioinformatically resolved

FLIP	-AAV	FLIP
FLIP	-AAV	FLOP
FLOP	-AAV	FLIP
FLOP	-AAV	FLOP

### Comparing Aligned Read Length



### FLIP/FLOP ratio from Nanopore



# Side-by-side comparison of Illumina and Nanopore sequencing approaches



(Adaptive Sampling!)

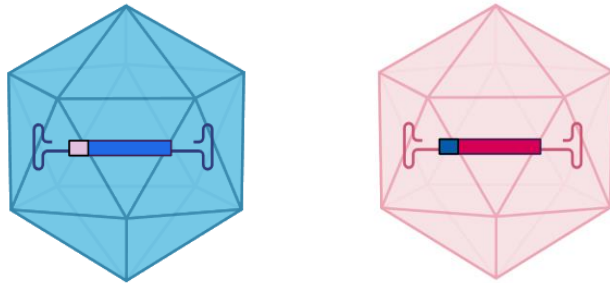


Quality	NextSeq 2000	PromethION
# Reads	✓	✗
Read length	✗	✓
Average read quality	✓	✗
rAAV genome coverage	✓	✓
FLIP/FLOP ratio	✗	✓
rAAV cross-contaminant detection		
Variant/indel detection		
Cost		

# If there was rAAV cross-contamination, both Illumina and Nanopore can detect it

Minimum coverage criteria: 1000X average coverage, 10X per base coverage

- Per 21CFR § 610.14, the contents of a final container of each filling must be specific for each product
- rAAVs are high-titer products, often over 1E13 vg/mL, miniscule cross-contamination could still be many rAAV copies
- Tested decreasing amounts of **REGV05** spiked into **REGV04**

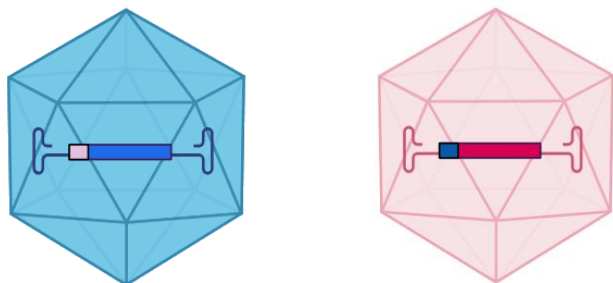


REGV04	REGV05
90%	10%
99%	1%
99.9%	0.1%
99.99%	0.01%
100%	0%

# If there was rAAV cross-contamination, both Illumina and Nanopore can detect it

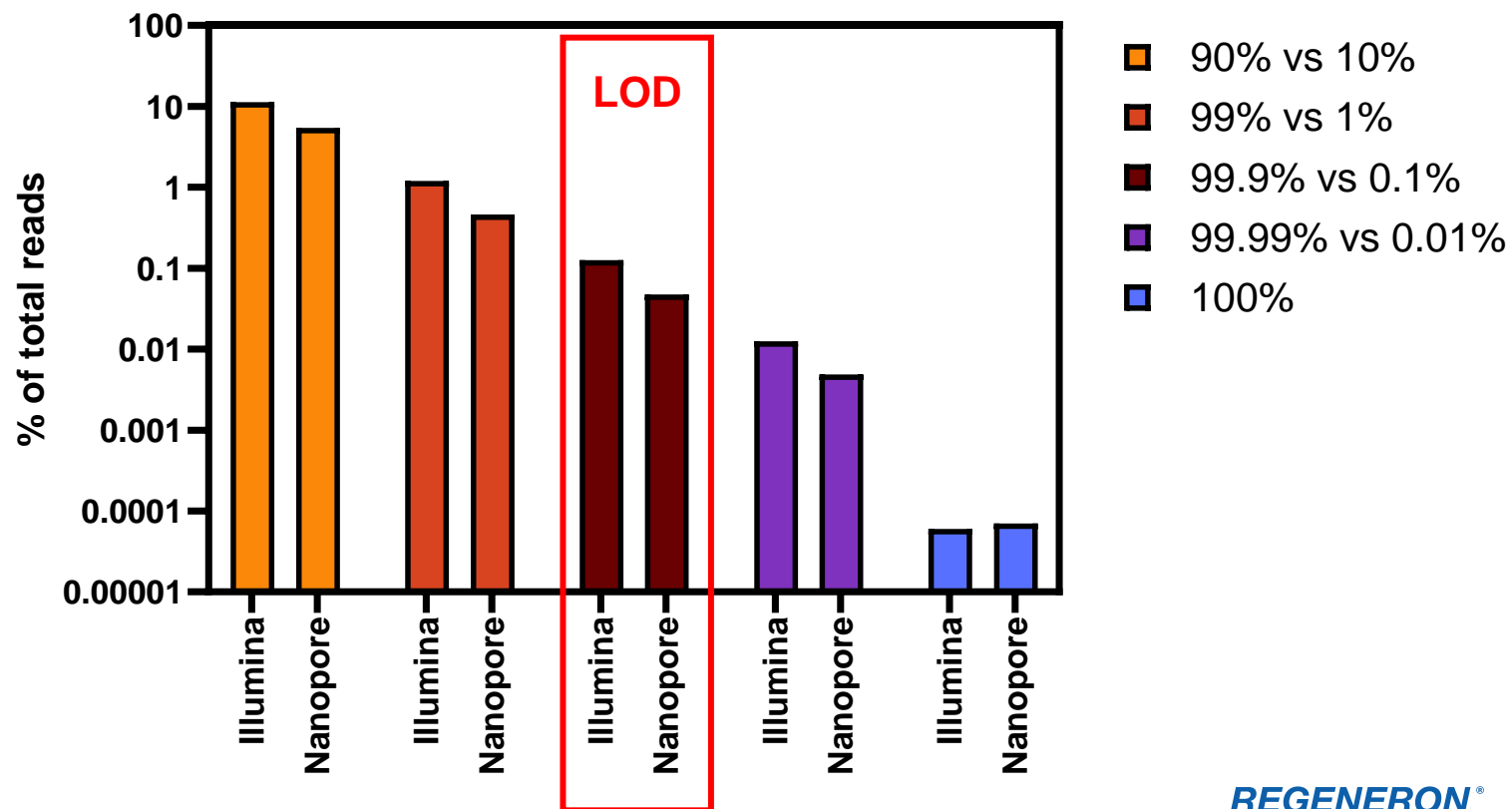
Minimum coverage criteria: 1000X average coverage, 10X per base coverage

- Tested decreasing amounts of **REGV05** spiked into **REGV04**
  - All spike levels detected
  - REGV009 0.1% was the lowest spike level that met minimum coverage criteria on both platforms



REGV04	REGV05
90%	10%
99%	1%
99.9%	0.1%
99.99%	0.01%
100%	0%

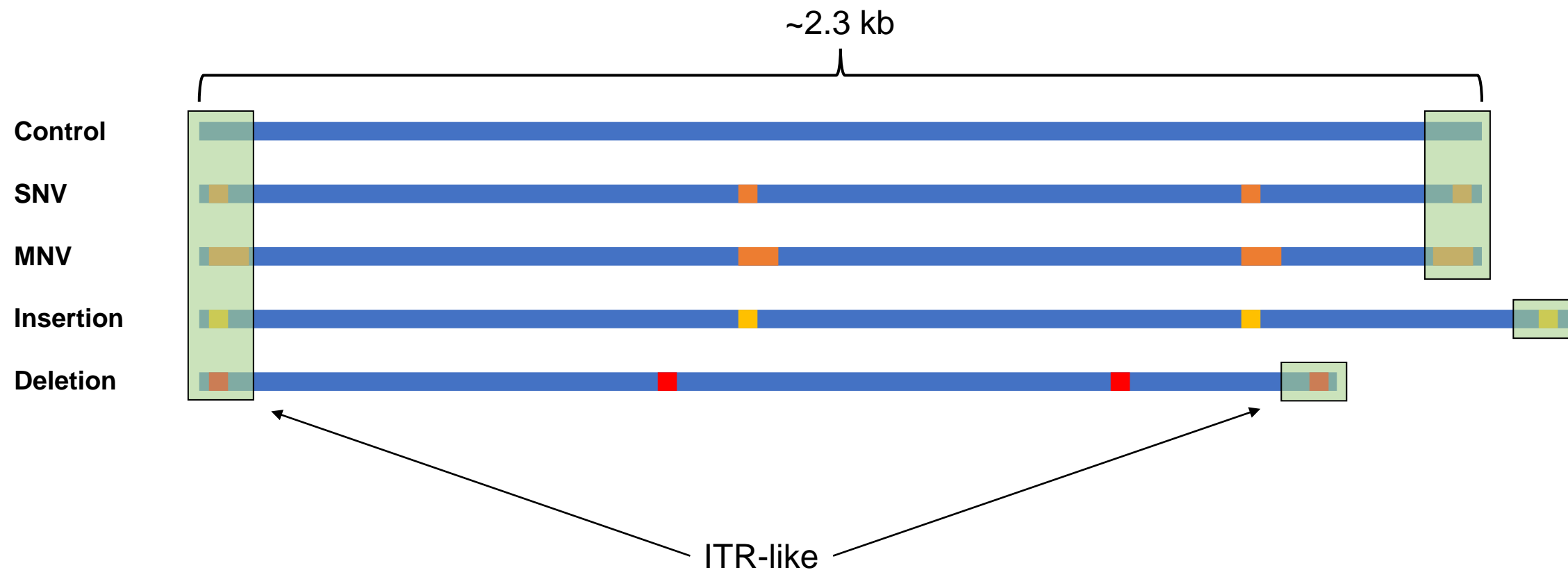
### REGV05 Uniquely Mapped Reads



# Can Illumina and Nanopore sequencing detect small rAAV differences?

Detection of rAAV variants is important for drug product potency specifications

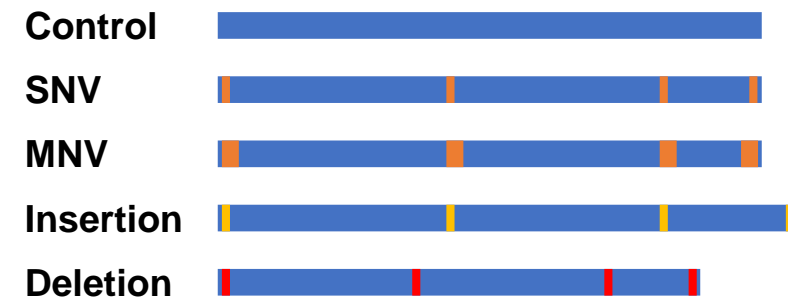
- 'Control' construct based originally on REGV02 sequence
- Mutations, insertions, or deletions added to a dsDNA control sequence



# Illumina sequencing better detected rAAV sequence variants

Combinations of dsDNA constructs to be tested in the future

- Illumina sequencing successful at detecting all variants
- Nanopore sequencing detected most variants



	Illumina				Nanopore			
	SNV	MNV	Insertion	Deletion	SNV	MNV	Insertion	Deletion
<b>5' End Variant</b>	Detected	Detected	Detected	Detected	Detected	Detected	Detected	Detected
<b>5' Middle Variant</b>	Detected	Detected	Detected	Detected	Detected	Detected	Detected	Detected
<b>3' Middle Variant</b>	Detected	Detected	Detected	Detected	Detected	Detected	Detected	Detected
<b>3' End Variant</b>	Detected	Detected	Detected	Detected	Detected	Not Detected	Detected	Detected

# Illumina sequencing is more cost-effective

Auxiliary materials like tubes and magnets not included in calculations

- Caveats: Multiplexing up to 8 samples on the NextSeq 2000 and no multiplexing on the PromethION
- If I was multiplexing up to the recommended maximum number of samples:
  - Illumina multiplexing
    - ≤ 30 samples
    - The cost of reagents for a 30-sample run would be **\$150 per sample**
  - Nanopore multiplexing
    - ≤ 24 samples
    - The cost of reagents for a 24-sample run would be **\$65 per sample**

Illumina Sequencing		Nanopore Sequencing	
Item	Cost	Item	Cost
Illumina DNA Prep	\$712	Ligation Kit	\$599
Dual Indexes Plate	\$480	PromethION Flow Cell	\$820
PhiX Control	\$165		
P2 Reagent Kit (600 cycle)	\$2,963		
NextSeq 2000	\$328,410	PromethION	\$195,455
<b>Cost per base (\$)</b>	0.01-0.05	<b>Cost per base (\$)</b>	0.10-0.20

# Side-by-side comparison of Illumina and Nanopore sequencing approaches

## Development of rAAV identity assay to proceed with short-read approach

- Adaptive sampling method 'enrichment' approach yields more on-target reads in Nanopore sequencing
  - Non-target reads < 1,000 bp will slip through
- Both sequencing approaches work and have their strengths
  - Cost and scalability: Illumina sequencing
  - ITRs: Nanopore sequencing



Quality	NextSeq 2000	PromethION
# Reads	✓	✗
Read length	✗	✓
Average read quality	✓	✗
rAAV genome coverage	✓	✓
FLIP/FLOP ratio	✗	✓
rAAV cross-contaminant detection	✓	✓
Variant/indel detection	✓	✓
Cost	✓	✗

# Acknowledgements

Questions?

## QC Virology Development Team

Seamus O'Connor, Ph.D.

David VanHoute, Ph.D.

## Wet-lab Team

Patrick Blatt, Ph.D.

Sriramana Kaginakudru, Ph.D.

Marissa Martinek, M.S.

Morgan Hewes

## Bioinformatics Team

Jared Richardson, Ph.D.

Christopher Remillard

Tyler Dion, M.S.

Devika Torvi

Ping Li, M.S.



- Some figures created with BioRender.com