Can Nanopore be used to detect RNA-protein interaction?

Anh Thu Thi Phan

Dr Chun Shen Lim
Assoc Prof Chris Brown
Dr Augustine Chen

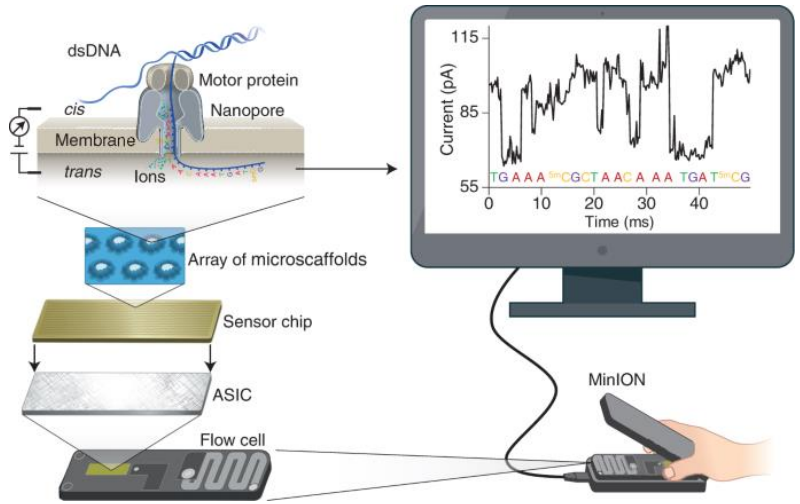
University of Otago
Te Whare Wānanga o Otāgo
New Zealand

Biochemistry
te tari matū koiora

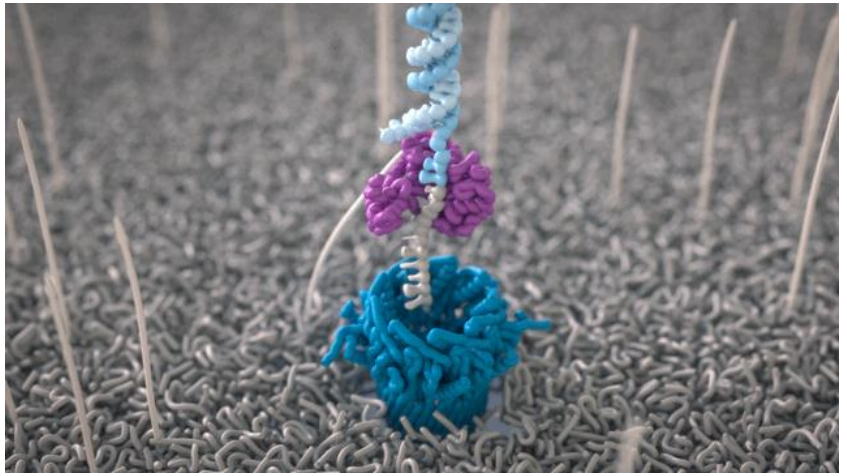
Direct RNA sequencing

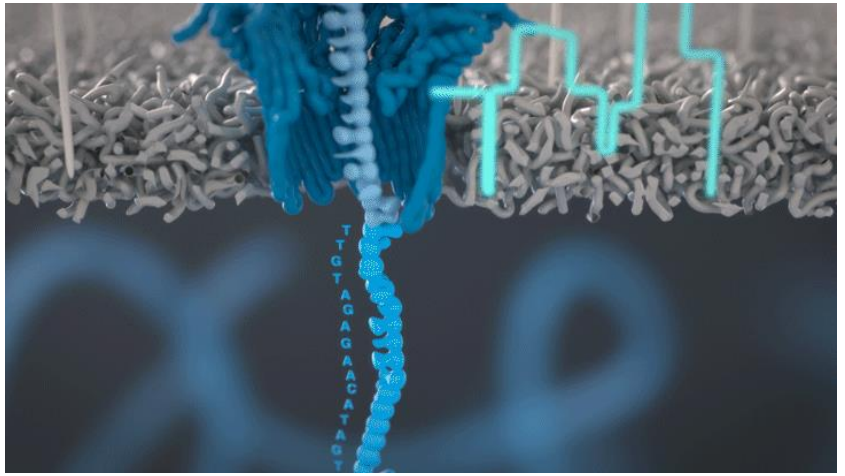
Each base creates a signature signal

- Magnitude of current blockage
- Event duration



Wang, 2021, Nat Biotechnol 39, 1348-1365





Oxford Nanopore Technologies, 2023
How Nanopore sequencing works

Objective

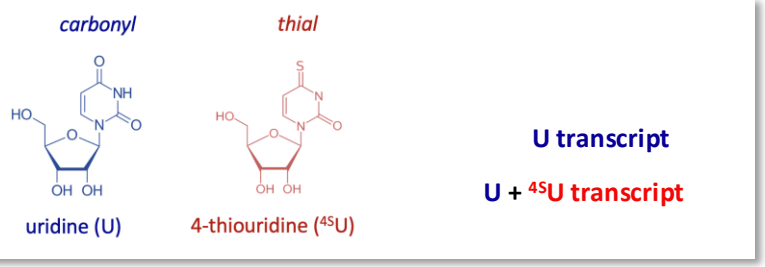
Nanopore direct RNA sequencing

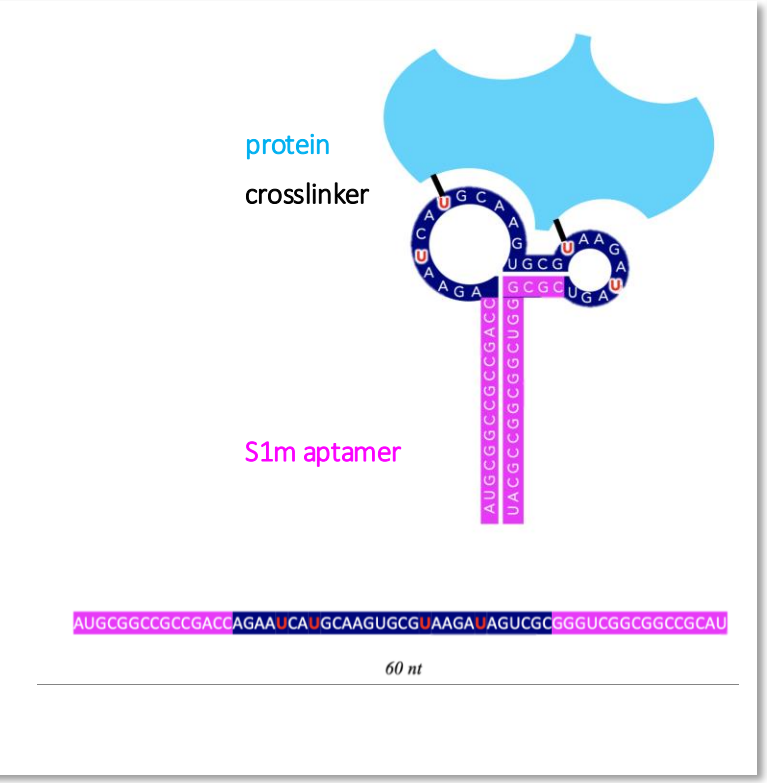
- *No PCR and RT bias*
- *Detect modified bases*
- *Identify protein binding sites?*

To detect protein-nucleic acid interaction sites on native RNA

- To create an *in vitro* model for protein-RNA interaction
  - To design and make an *in vitro* transcribed mRNA
  - To modify the RNA at the site of protein binding
  - To directly sequence the RNA, detecting the modifications

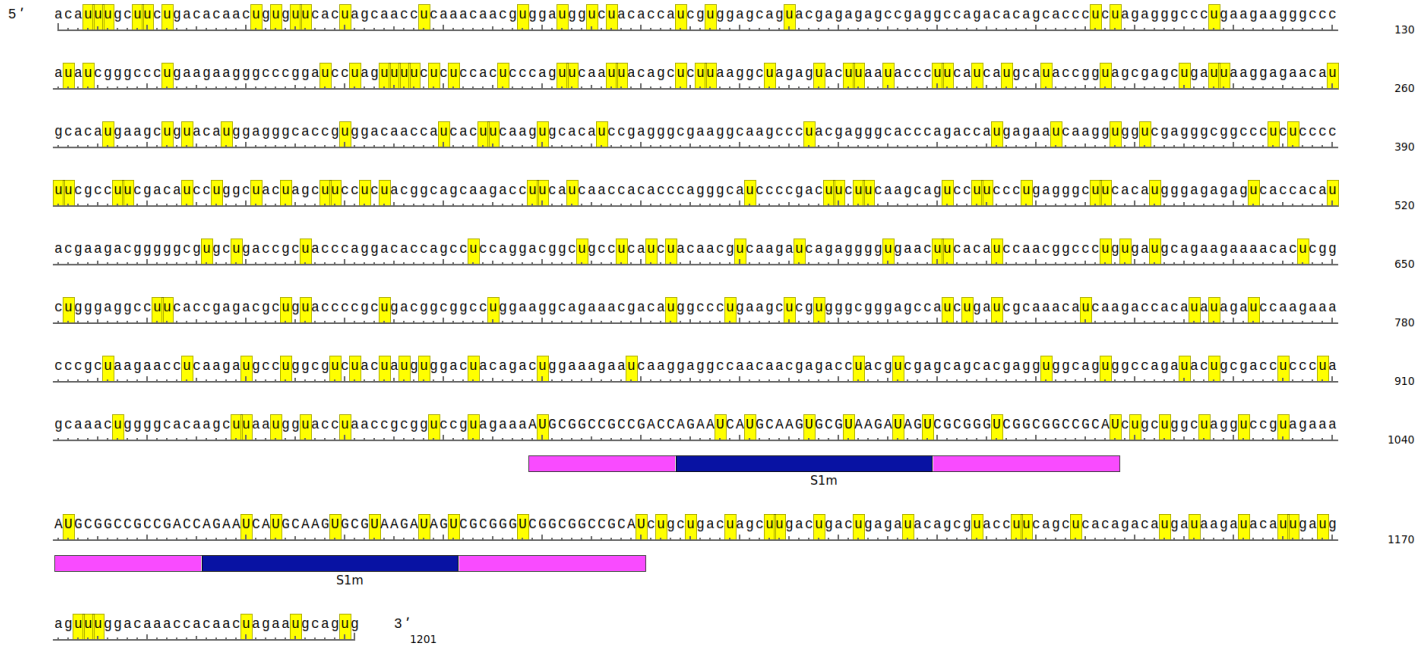
RNA and protein interaction



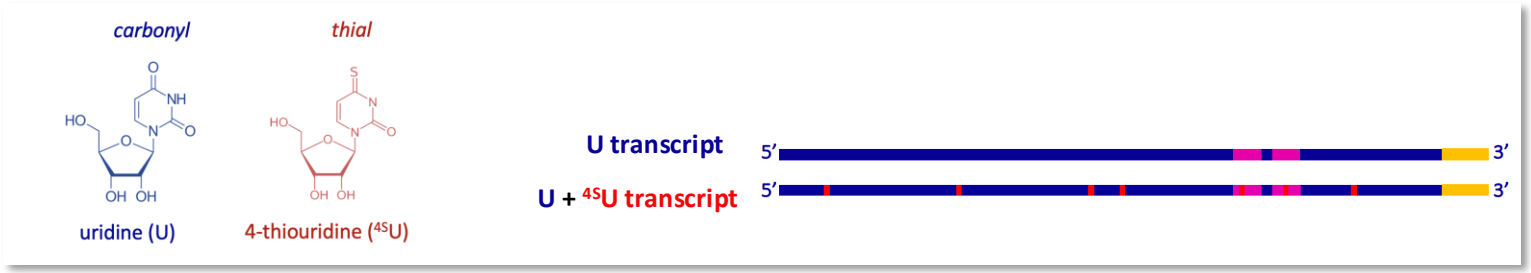


RNA 1201 nt

A (324) =27.0% U (213) =17.7% C (347) =28.9% G (317) =26.4%

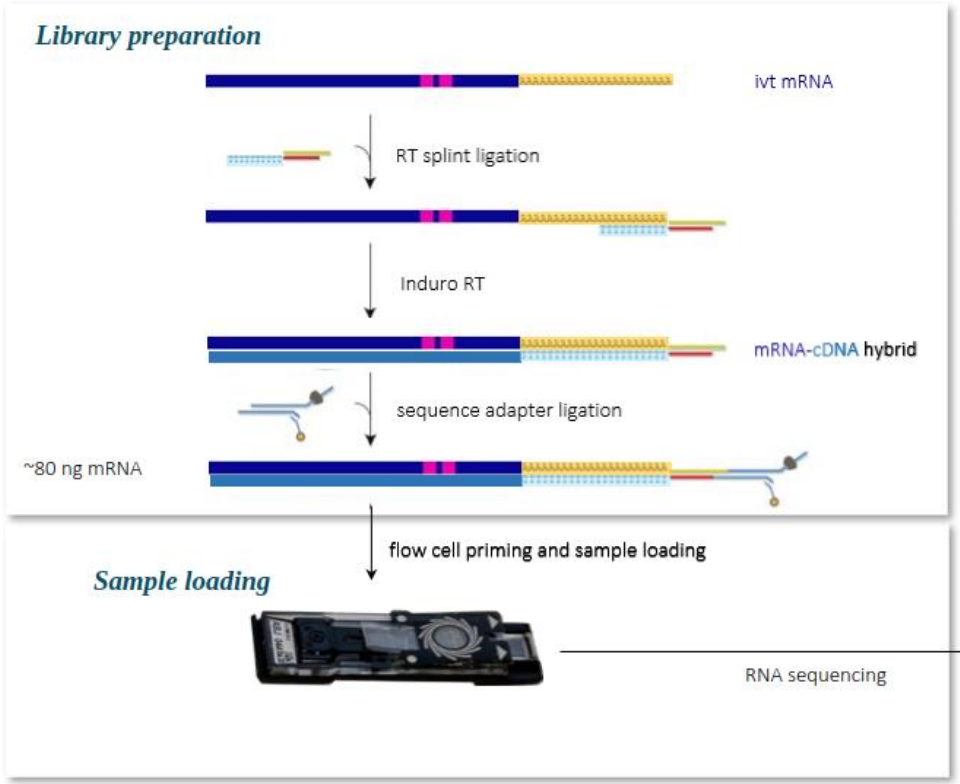


ONT direct RNA sequencing

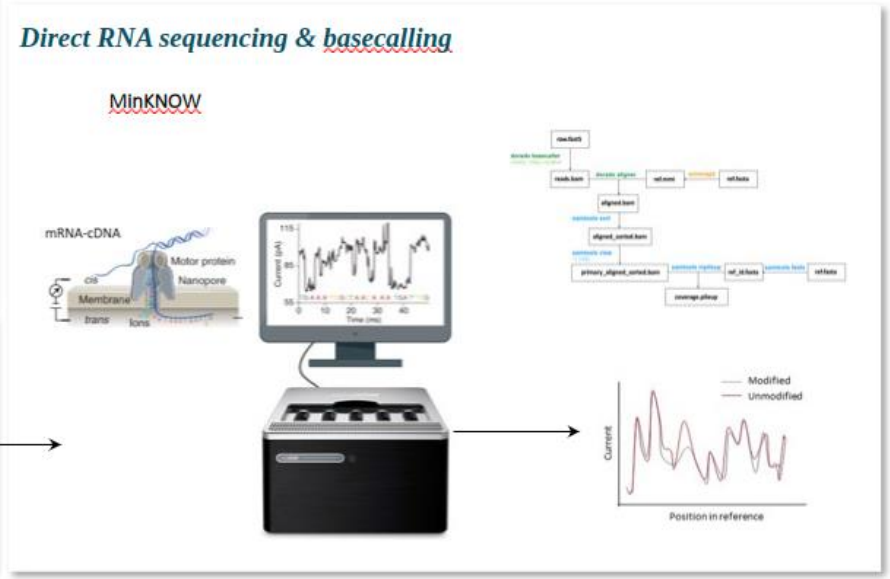


Direct RNA sequencing kit (SQK-RNA002)

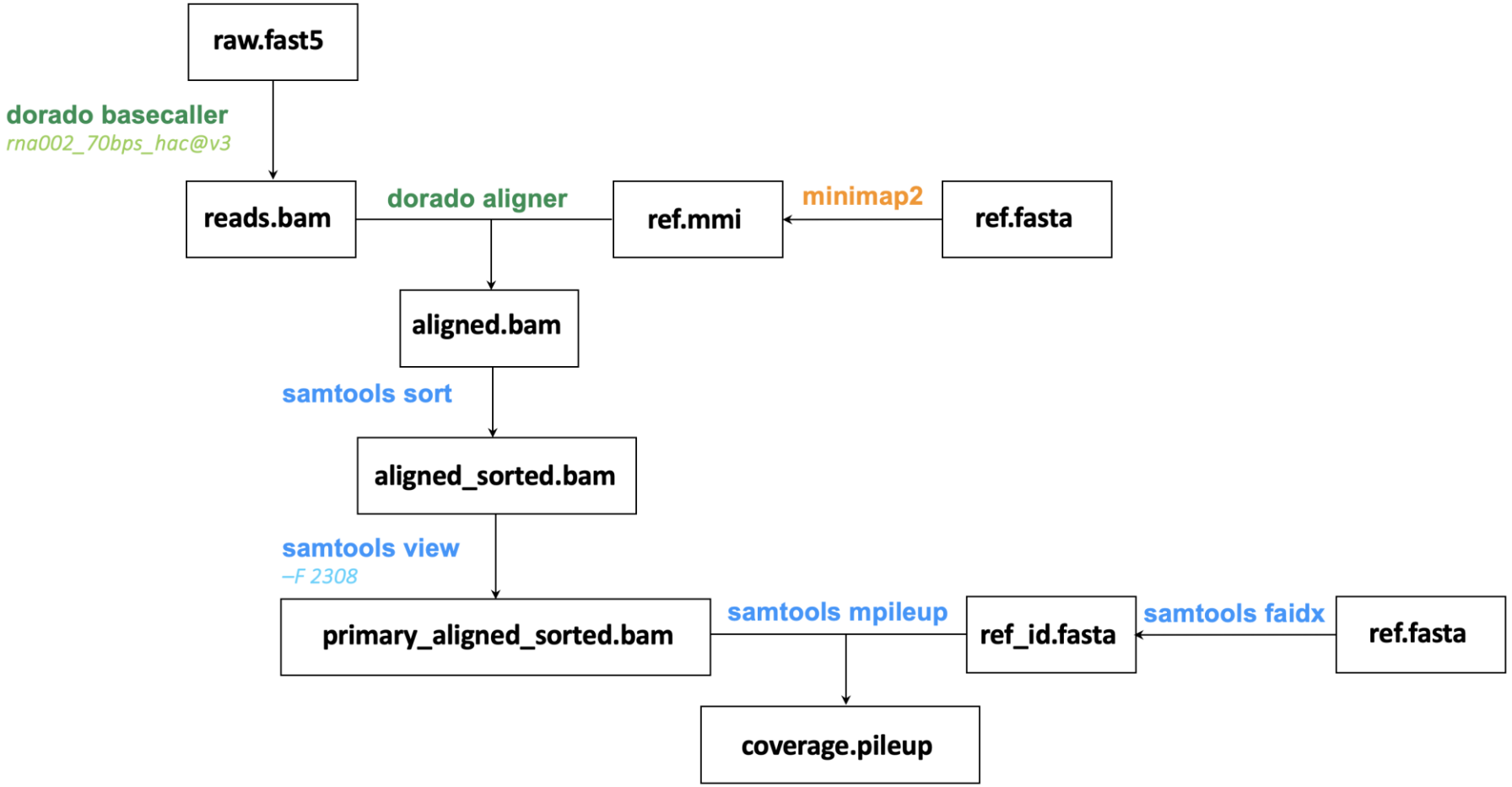
Flongle flow cell R9.4.1



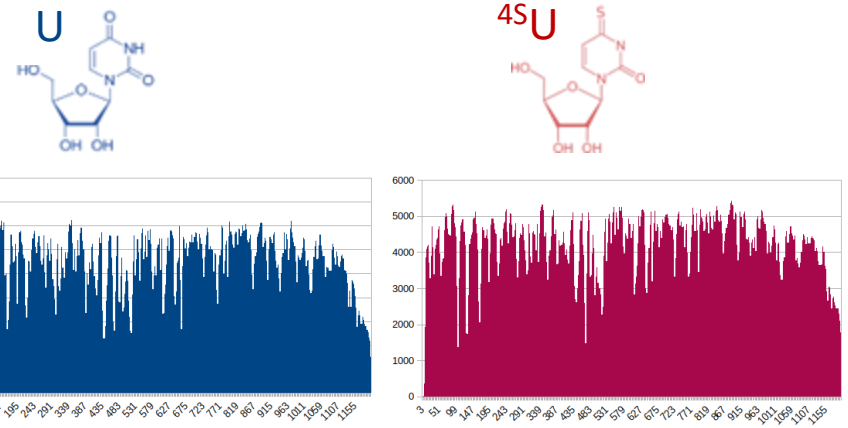
GridION sequencer



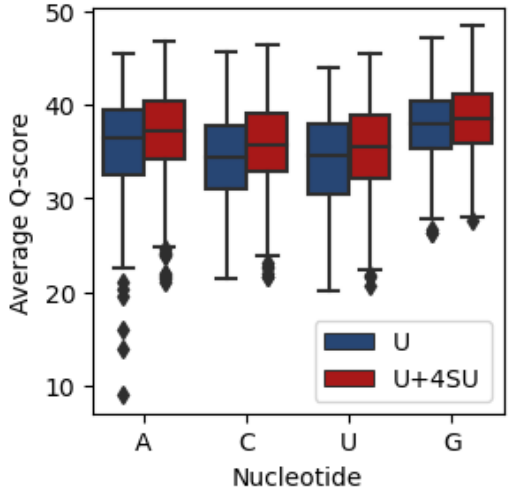
Data analysis -alignment workflow



Results comparison: U and U + 4sU



Coverage plot for transcripts with U and transcripts with U + ^{4s}U



Sequencing accuracy of direct RNA seq by base

1201 nt + ~100 As

	U	U + 4SU
Readouts (Q-score >7)		
Base-called reads	24.4 k	11.4 k
N50	1.18 kb	1.18 kb
Primary aligned reads		
Aligned reads	17.7 k (73.2%)	8.2 k (72.1%)
Mean coverage	12.2 k	5.3 k
Mean base Q-score (Dorado)	31.4	33.5
Alignment error rate (M/D/I) Consensus	8.3% (3.7/2.6/2.0) 0.25% (3 in 1.2 kb)	6.3% (2.7/2.1/1.5) 0.08% (1 in 1.2 kb)

Next steps

- To detect protein-nucleic acid interaction sites on native RNA
  - To create an *in vitro* model for protein-RNA
    - To create a synthetic RNA construct with:
    - To perform direct RNA sequencing on *in vitro* transcribed mRNA
    - To crosslink protein and *in vitro* transcribed mRNA
    - To perform direct RNA seq on crosslinked mRNA
      - ***New RNA flow cell and chemistry available this month?***
        - Direct RNA seq kit (SQK-RNA004) and RP4 pore.

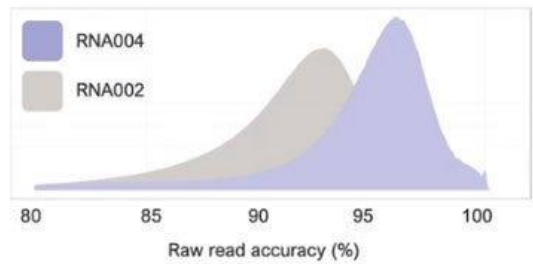
Direct RNA Sequencing

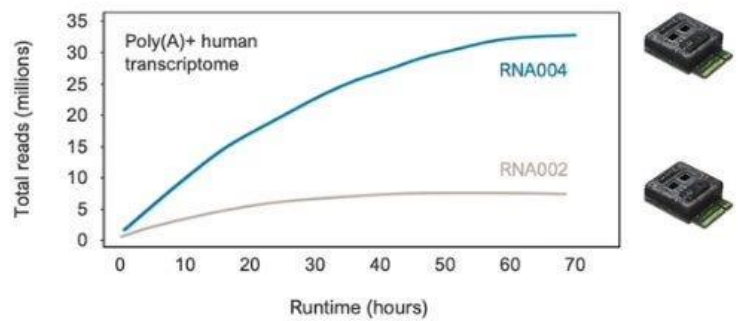
Upgrades

Improvements to Direct RNA

- "RP4" nanopore optimised specifically for RNA accuracy
  - Allows best signal for base calling without compromise
  - Requires dedicated flowcells for best RNA results
- RNA enzyme motor developed for better speed and output
  - Now averaging speed of 125 bps (~2x improvement)
  - Hitting outputs of 30 million reads from PromethION flowcell

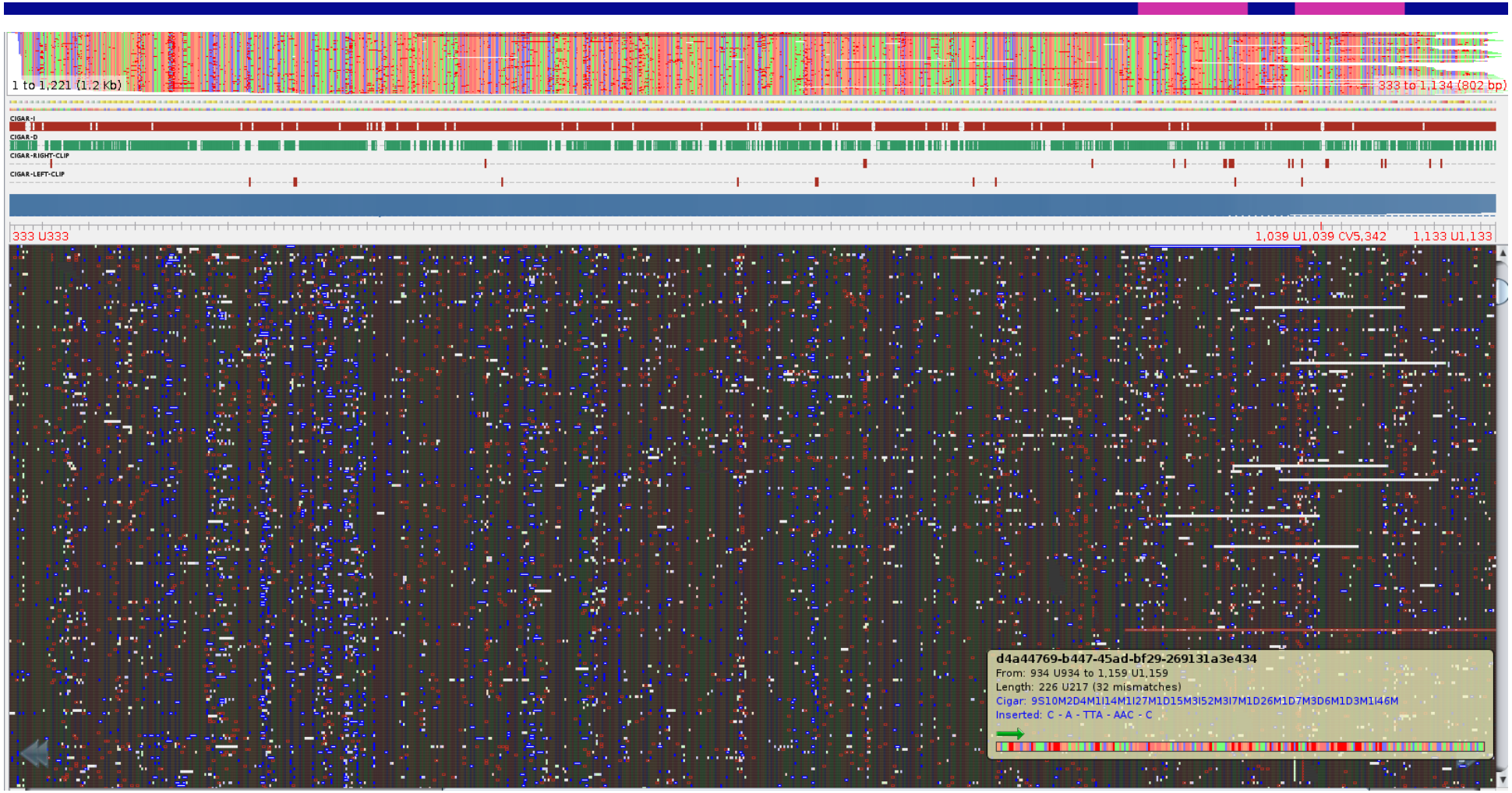
9 RNA02	RP4 RNA04
FLO-MIN106D FLO-PRO002	FLO-MIN004R FLO-PRO004R
SQK-RNA0002	SQK-RNA004





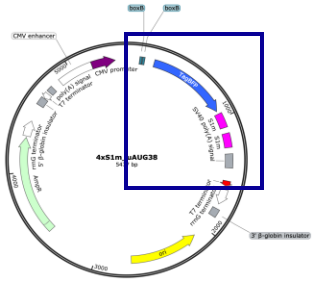
#NanoporeConf

S1m coverage



Tablet showing direct RNA reads against reference seq

Synthetic RNA

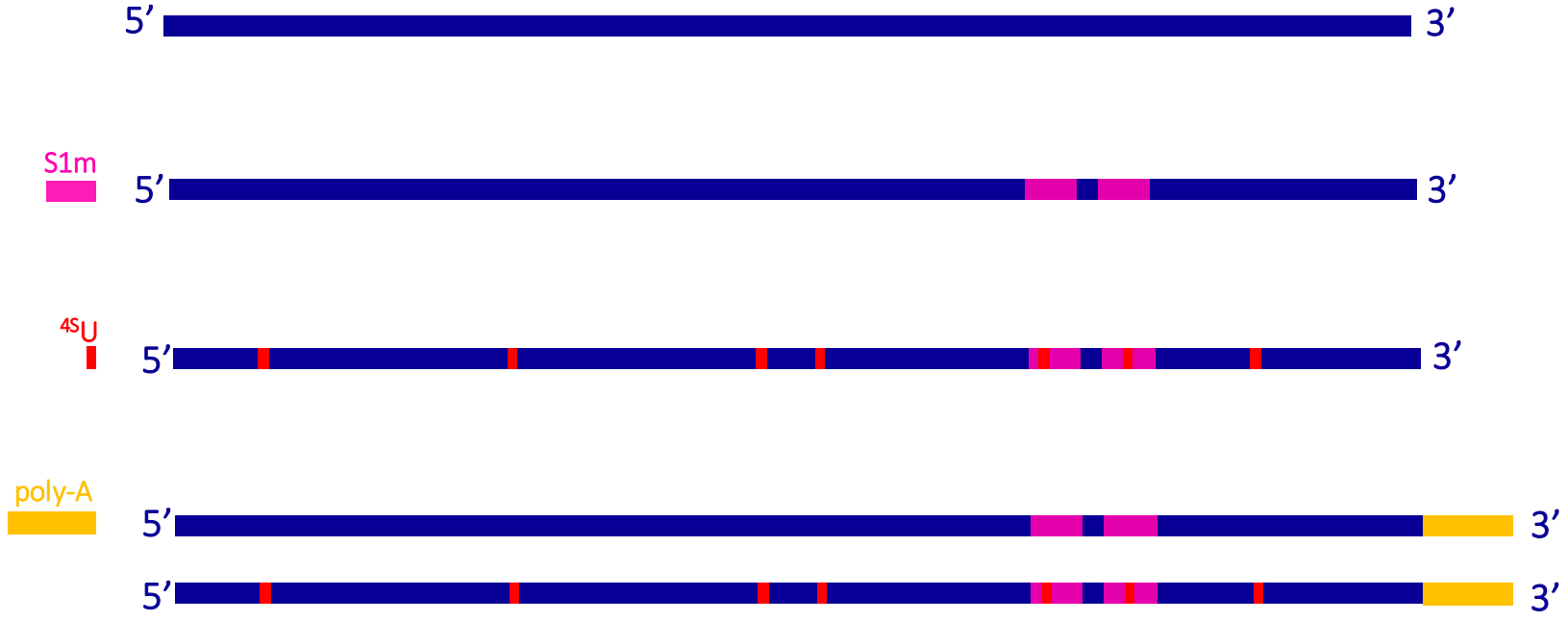


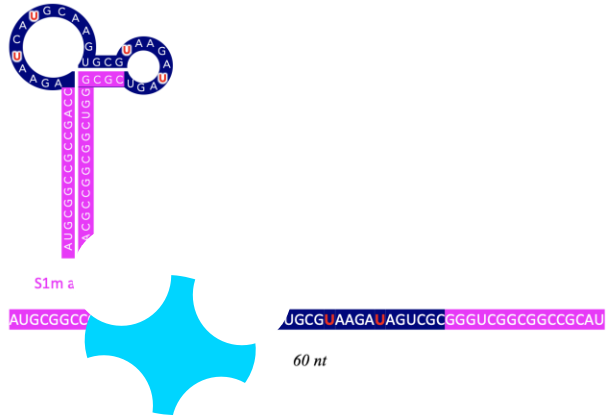
Known sequence
- *IVT from plasmid*

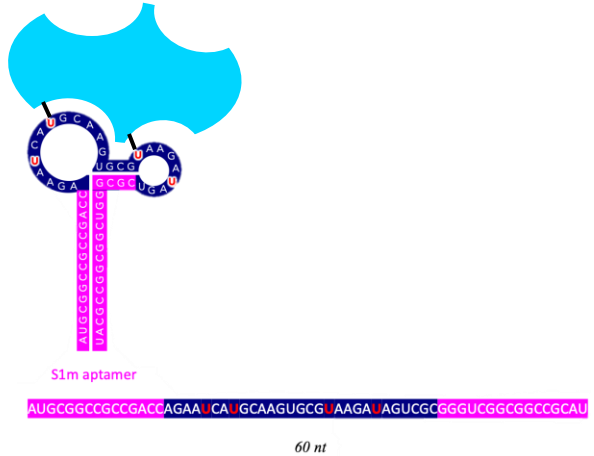
Protein interaction site
- *S1m introduction*

Modified base for crosslinking
- *^{4S}U incorporation*

Polyadenylated
- *Poly-A tail addition*







1201 bases (RNA)

5′ acauuugcuucugacacaacuguguucacuagcaaccucaaacaacguggaugucuacaccaucguggagcaguacgagagagccgaggccagacacagcacccucuagagggcccugaagaagggccc 130

auaucgggcccugaagaagggcccggauccuaguuuucucuccacucccaguucaauuacagcucuuaaggcuagaguacuuaauacccuucaucaugcauaccgguagcgagcugauuaaggagaacau 260

gcacaugaagcuguacauggagggcaccguggacaaccaucacuucaagugcacauccgagggcgaaggcaagcccuacgagggcacccagaccaugagaaucaagguggucgagggcggcccucucccc 390

uucgccuucgacauccuggcuacuagcuuccucuacggcagcaagaccuucaucaaccacacccagggcaucccgacuucuucaagcaguccuucccugagggcuucacaugggagagagucaccacau 520

acgaagacgggggcgugcugaccgcuacccaggacaccagccuccaggacggcugccucaucuacaacgucaagaucagagggugaacuucacauccaacggcccugugaugcagaagaaaacacucgg 650

cugggaggccuucaccgagacgcuguaccccgcugacggcggccuggaaggcagaaacgacauggcccugaagcucgugggcgggagccaucugaucgcaaacaucaagaccacauauagauccaagaaa 780

cccgcuaagaaccucaagaugccuggcgucuacuauguggacuacagacuggaaagaaucaaggaggccaacaacgagaccuacgucgagcagcacgagguggcaguggccagauacugcgaccucccua 910

gcaaacuggggcacaagcuuaaugguaccuaaccgcgguccguagaaaAUGCGGCCGCCGACCAGAAUCAUGCAAGUGCGUAAGAUAGUCGCGGGUCGGCGGCCGCAUcugcuggcuaggucguagaaa 1040

S1m

AUGCGGCCGCCGACCAGAAUCAUGCAAGUGCGUAAGAUAGUCGCGGGUCGGCGGCCGCAUcugcugacuagcuugacugacugagauacagcguaccuucagcucacagacaugauaagauacauugaug 1170

S1m

aguuuggacaaaccacaacuagaaugcagug 3′ 1201

Sequencing kit and flow cell

Sequencing kit

- Direct RNA Sequencing Kit (SQK-RNA002)
  - Kit chemistry: kit 9

Flow cell

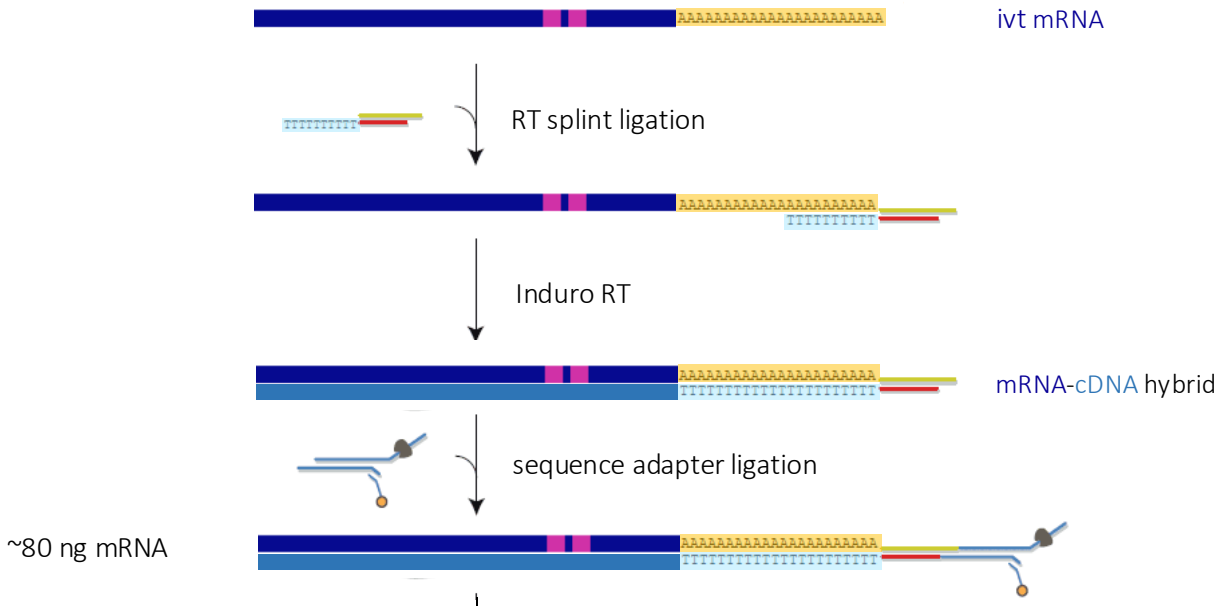
- Flongle Flow Cell R9.4.1 (FLO-FLG001)
  - Up to 126 channels

Sequencer

- GridION

ONT direct RNA sequencing

Library preparation

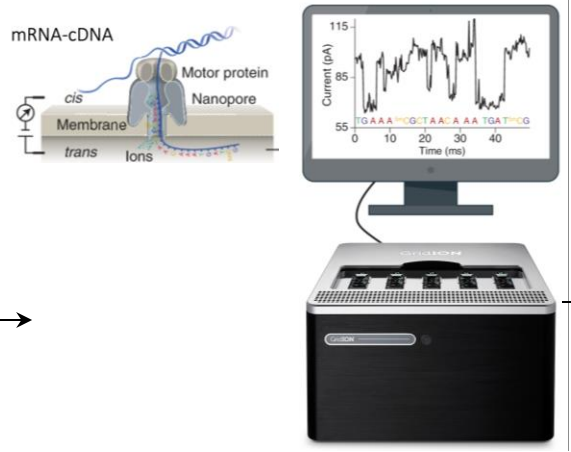


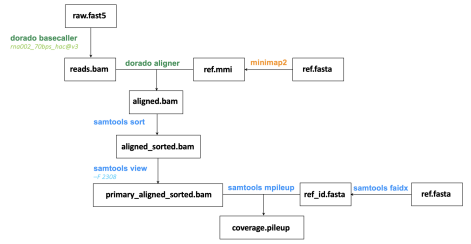
Sample loading

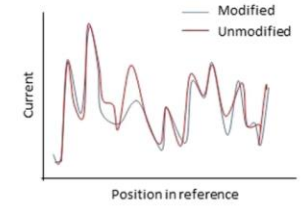
RNA sequencing

Direct RNA sequencing & basecalling

MinKNOW







Accuracy of the consensus sequence -tablet