Detecting mRNA modifications using single molecule nanopore sequencing

RACHAEL WORKMAN

TIMP LAB, JOHNS HOPKINS UNIVERSITY, DEPARTMENT OF BIOMEDICAL ENGINEERING

NANOPORE HUMAN RNA CONSORTIUM

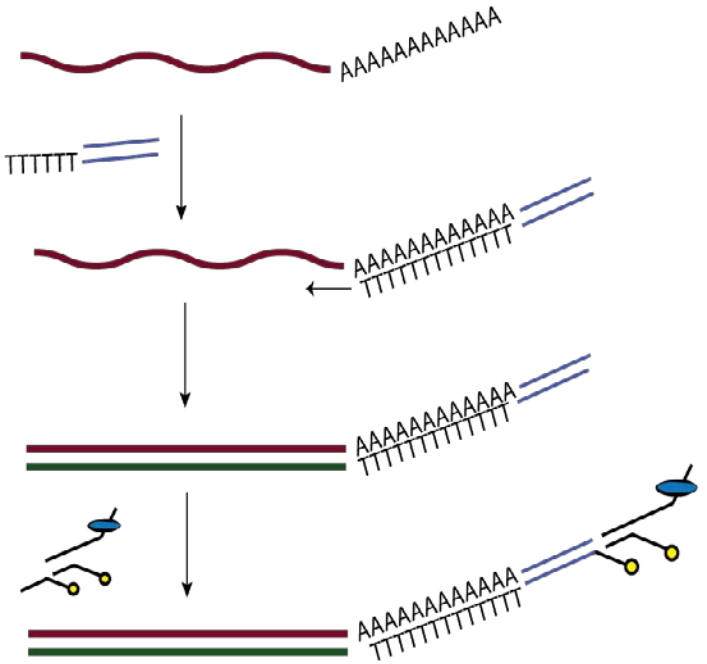
Native RNA sequencing opens new frontiers for long read RNA-seq analysis

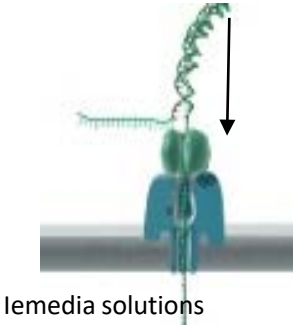
PolyA+ RNA captured

Splint poly-T adapter ligation

Reverse transcription (optional)

Sequencing adapter ligation



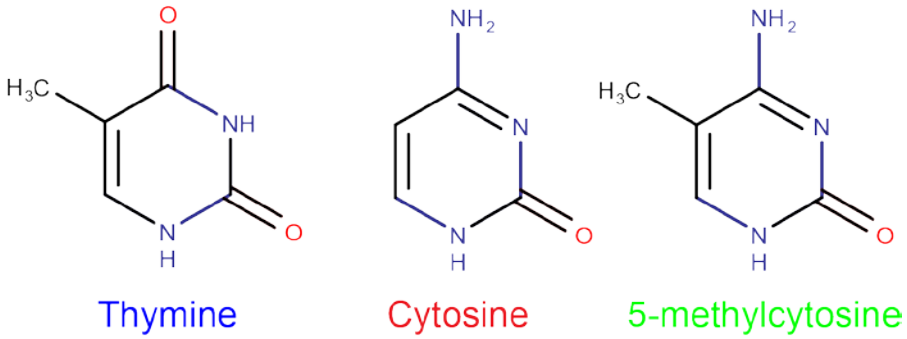


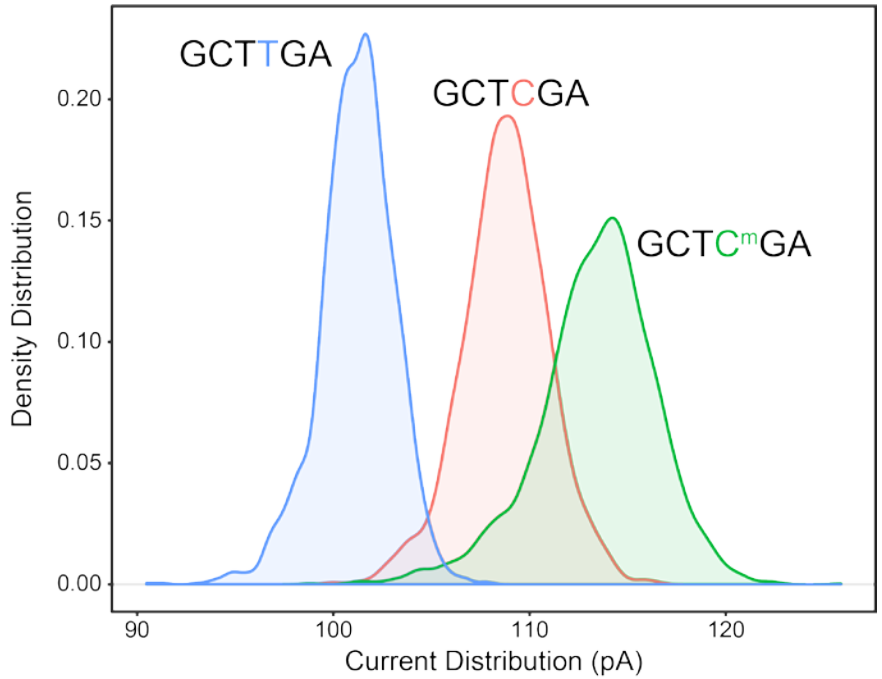
Iemedia solutions

Raw Current Signal

Current Signal K-mers Sequence

Nanopore current is affected by modifications

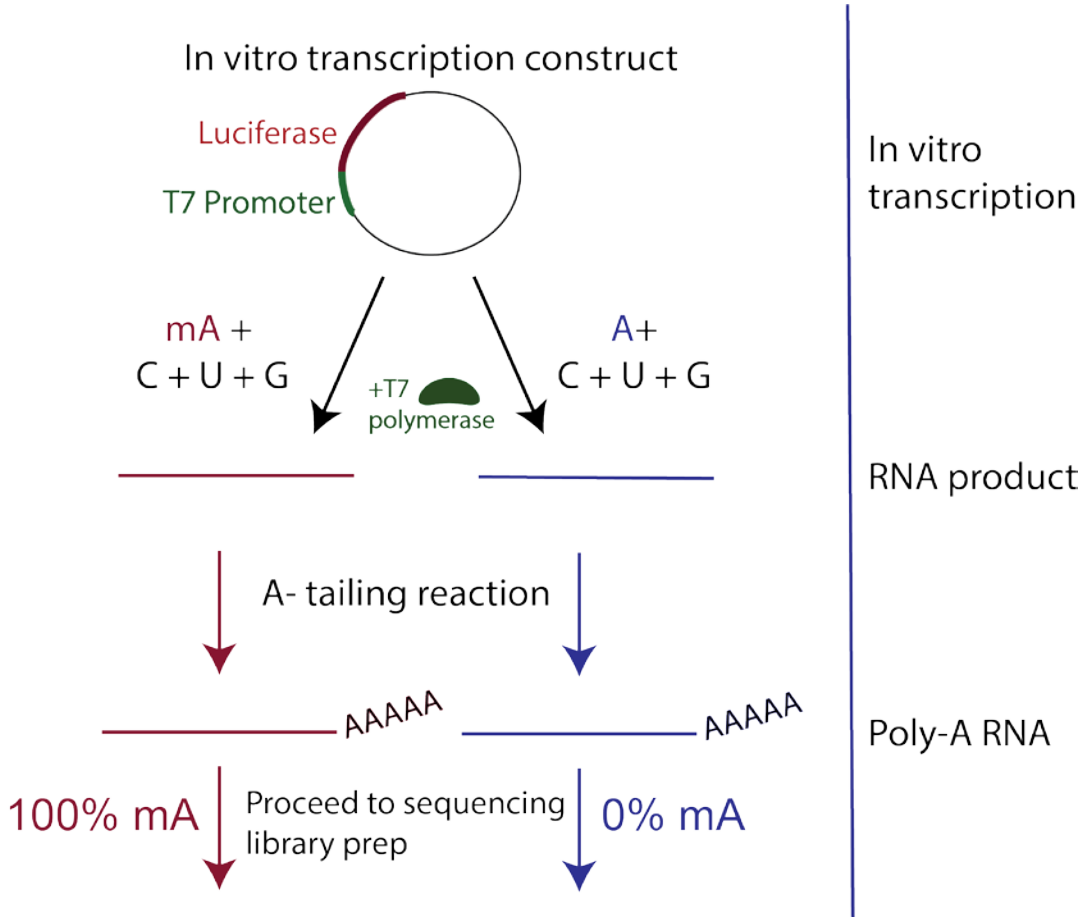




- 5mC detected and modeled in DNA
- Current shifts occur when a non-canonical nucleotide traverses the pore, as we know from DNA studies
- These shifts can be modeled and used to call modifications

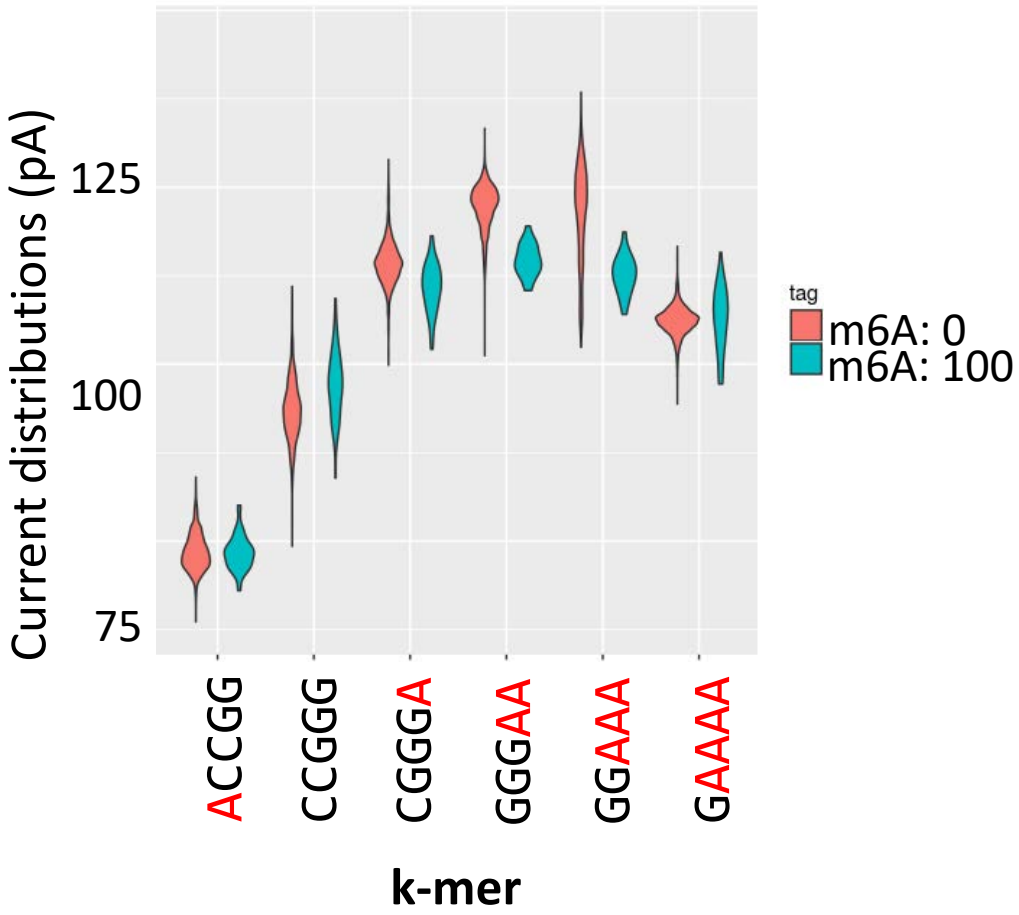
Simpson et. al *Nature* Methods 2017

Training RNA basecaller to recognize modified sites requires truth sets: **modIVT**



- IVT based RNA synthesis allows incorporation of labeled nucleotides
- All or none reaction right now, T7 has a strong preference for the unmodified nucleotides, making mixtures hard

Training RNA basecaller to recognize modified sites requires truth sets: **modIVT**

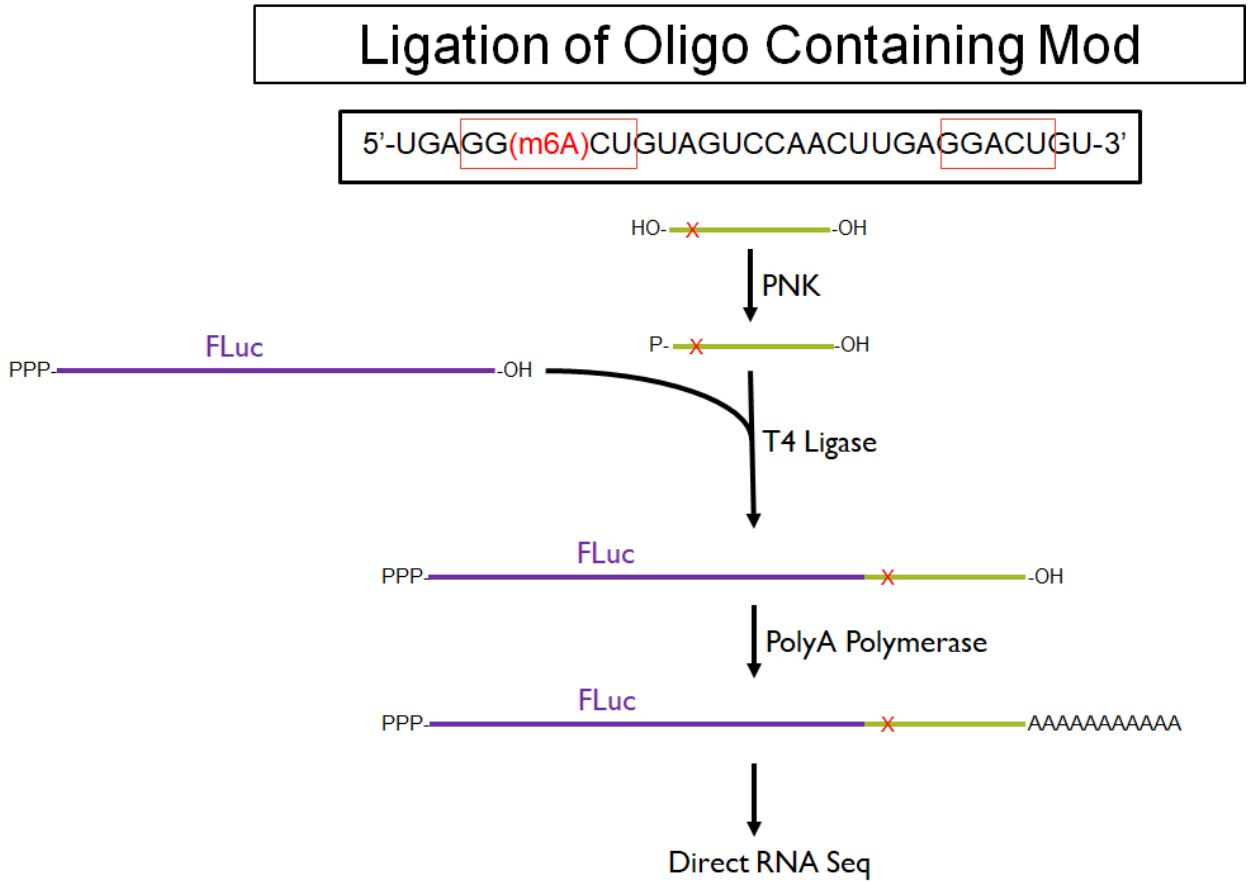


- From luciferase we can already see strong signal depending on context
- Using nanopolish eventalign, we can extract the distribution of current values along the RNA strand

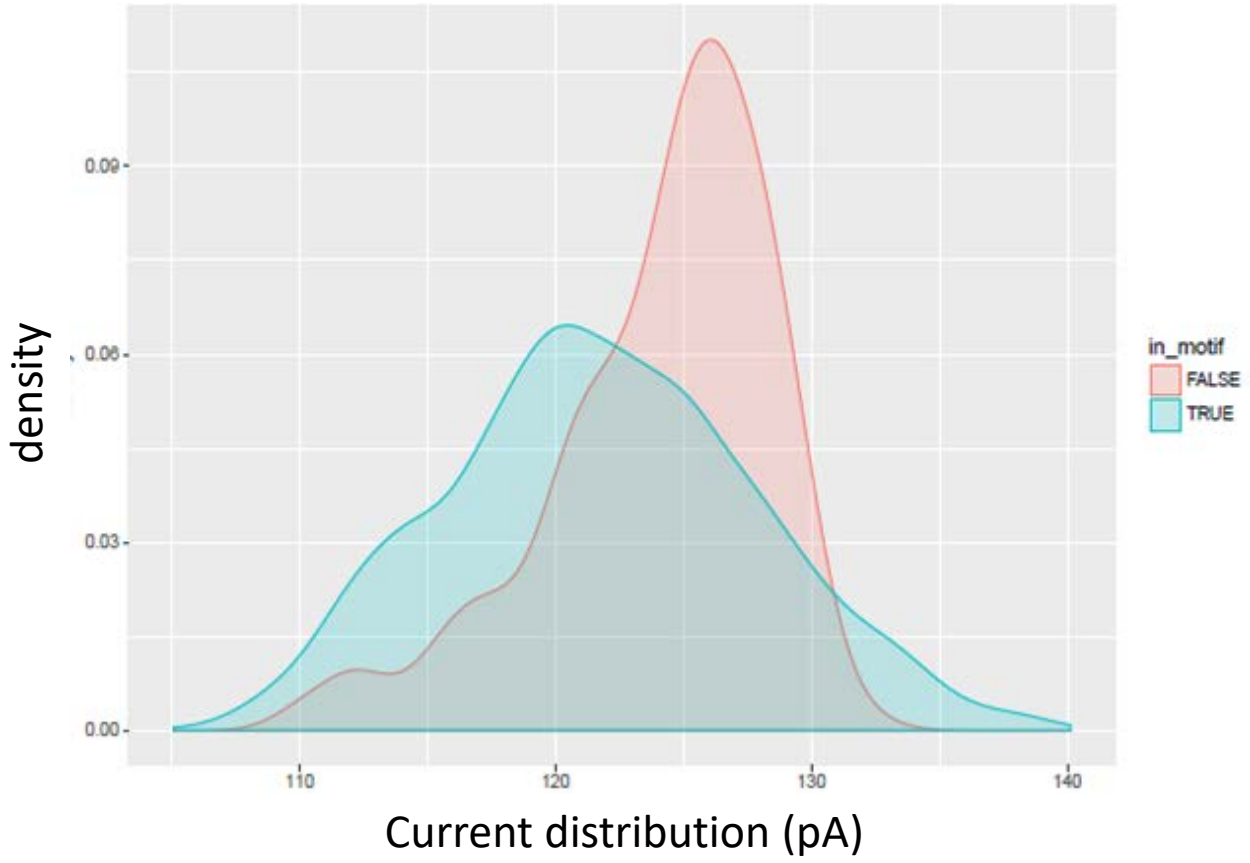
Nanopolish eventalign: https://github.com/jts/nanopolish

Training RNA basecaller to recognize modified sites requires truth sets: **oligo ligation**

- Short oligo (easier/cheaper to synthesize) ligated to handle (to give needed length for sequencing)
- Oligo containing modified and unmodified METTL3 motif



Training RNA basecaller to recognize modified sites requires truth sets: **oligo ligation**

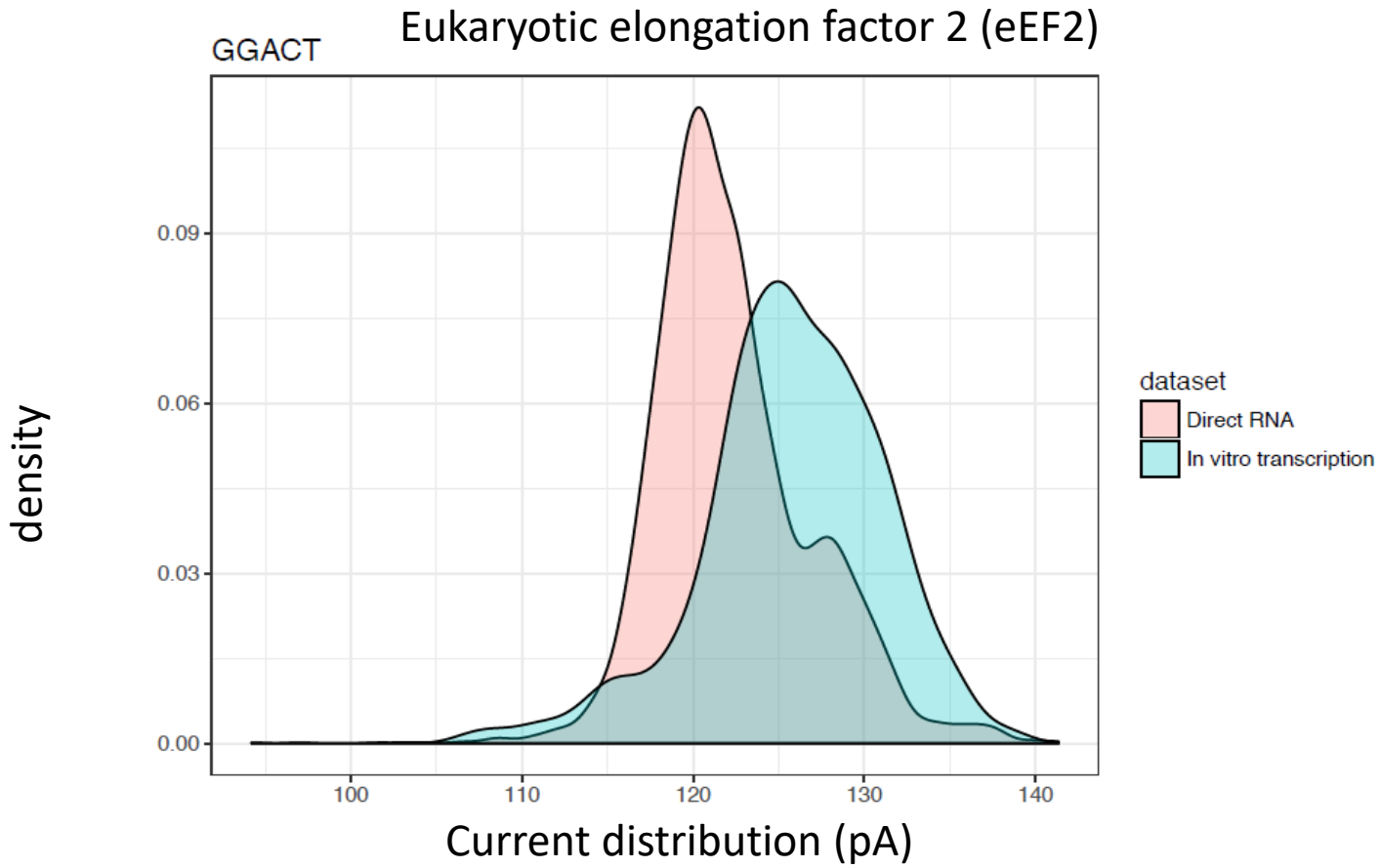


- Similar current distribution shift observed in modification–containing kmers

Nanopolish eventalign: https://github.com/jts/nanopolish

Independently identified m6A peaks in GM12878 reveal hints of mod signature

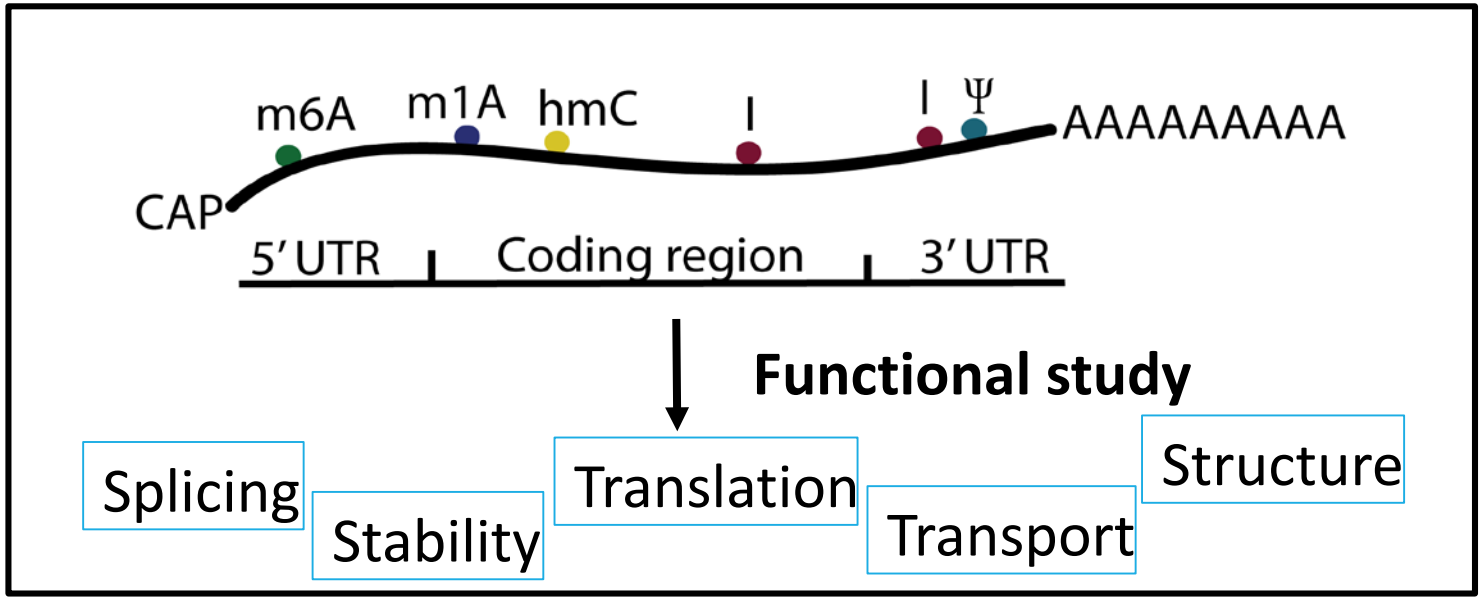
- Direct RNA data shown is publically available! https://github.com/nanopore-wgs-consortium/NA12878/blob/master/RNA.md
- Shift in current distribution at peak site in Eukaryotic elongation factor 2 (eEF2)



Target gene found using m6A immunoprecipitation peak datasets from: Molinie et al 2016. Nat Methods. Roost et al 2015. JACS.

Future directions for RNA modification detection

- Global m6A calling
- Simultaneous detection of multiple mods
- Investigation of relationship between modifications, splice variation and poly-A tail lengths



Acknowledgements

Birmingham
Nick Loman
Josh Quick
Andrew Beggs
Jaqueline Goes de Jesus

Nottingham
Matt Loose
Nadine Holmes
Matthew Carlile

Santa Cruz
Mark Akeson
Hugh Olsen
Benedict Paten
Angela Brooks
Miten Jain

British Columbia
Terry Snutch
John Tyson

Toronto (OICR)
Jared Simpson
Phil Zuzarte
Paul Tang

Johns Hopkins
Winston Timp
Timothy Gilpatrick
Norah Sadowski

ONT
Daniel Garalde
Daniel Jachimowicz
Andy Heron
Rosemary Dokos