

Prediction of rRNA modifications in single molecule resolution with hidden Markov models

Kübra Altinel¹, Miten Jain², Hugh E Olsen², Martin D. Jansson¹, Sophia J. Häfner¹, Mark Akesson² & Anders H. Lund¹

Co-supervisors: Albin Sandelin¹³ and Martin Miller⁴

¹ Biotech Research and Innovation Centre (BRIC), University of Copenhagen, Copenhagen, Denmark

² UC Santa Cruz Genomics Institute, University of California, Santa Cruz, California, United States of America

³ Bioinformatics Centre, Computational and RNA Biology, Department of Biology, University of Copenhagen, Copenhagen, Denmark

⁴ Cancer Research UK, Cambridge Institute, University of Cambridge, Cambridge, UK



Hypothesis

Different combinations of rRNA modifications may alter ribosome behaviour in translation

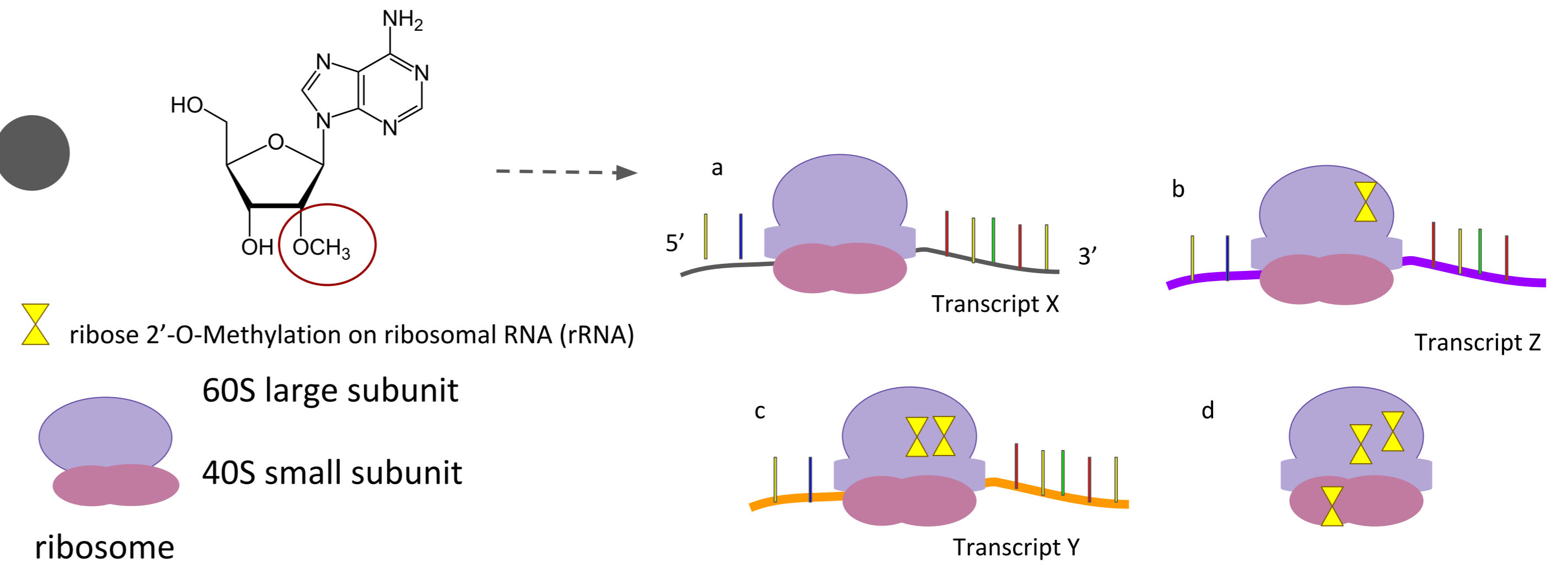


Fig.1: rRNAs are targets for modifications. One common modification type is 2'-O-methylation (2'-O-Me), which may regulate the ribosome and lead to changes in canonical translation (a) via preferring different mRNAs to translate (see b & c) or decreasing the amount of translation (d).

Direct RNA sequencing with Oxford Nanopore Technologies

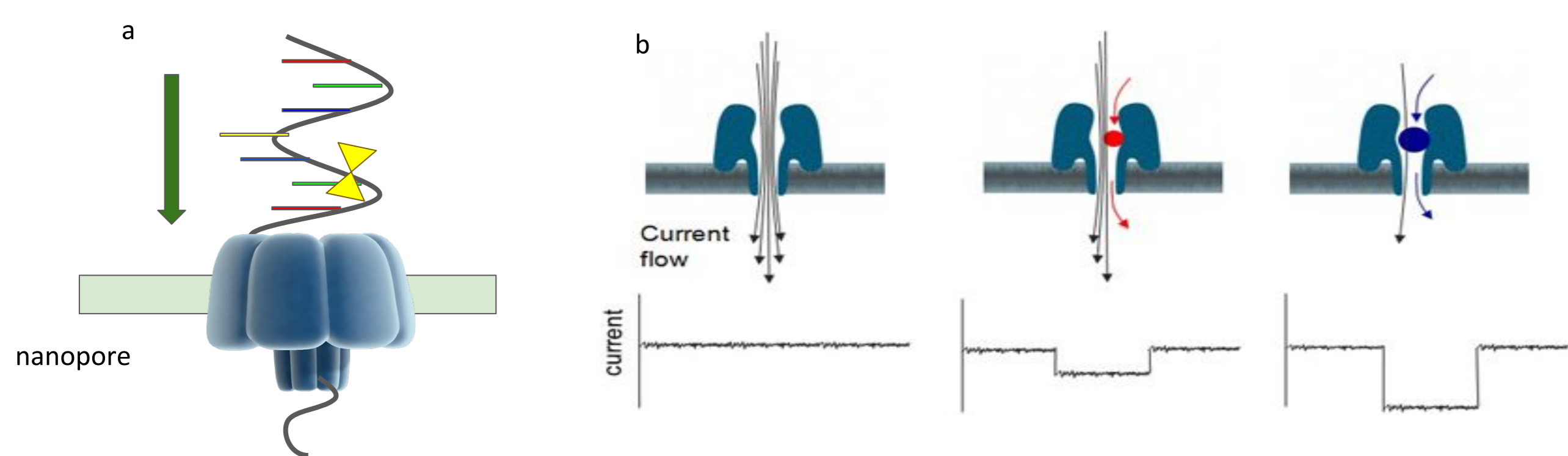


Fig.2: Schematic representation of direct RNA sequencing with Oxford Nanopore Technologies (ONT). a) A single stranded RNA with modified nucleotide is passing through an ONT nanopore embedded in a lipid membrane. b) Plot of ionic current change when a molecule is passing through the nanopore (images are from <https://nanoporetech.com>)

AIM

Classification of individual ribosomes based on their specific modification patterns and their possible impacts on translation

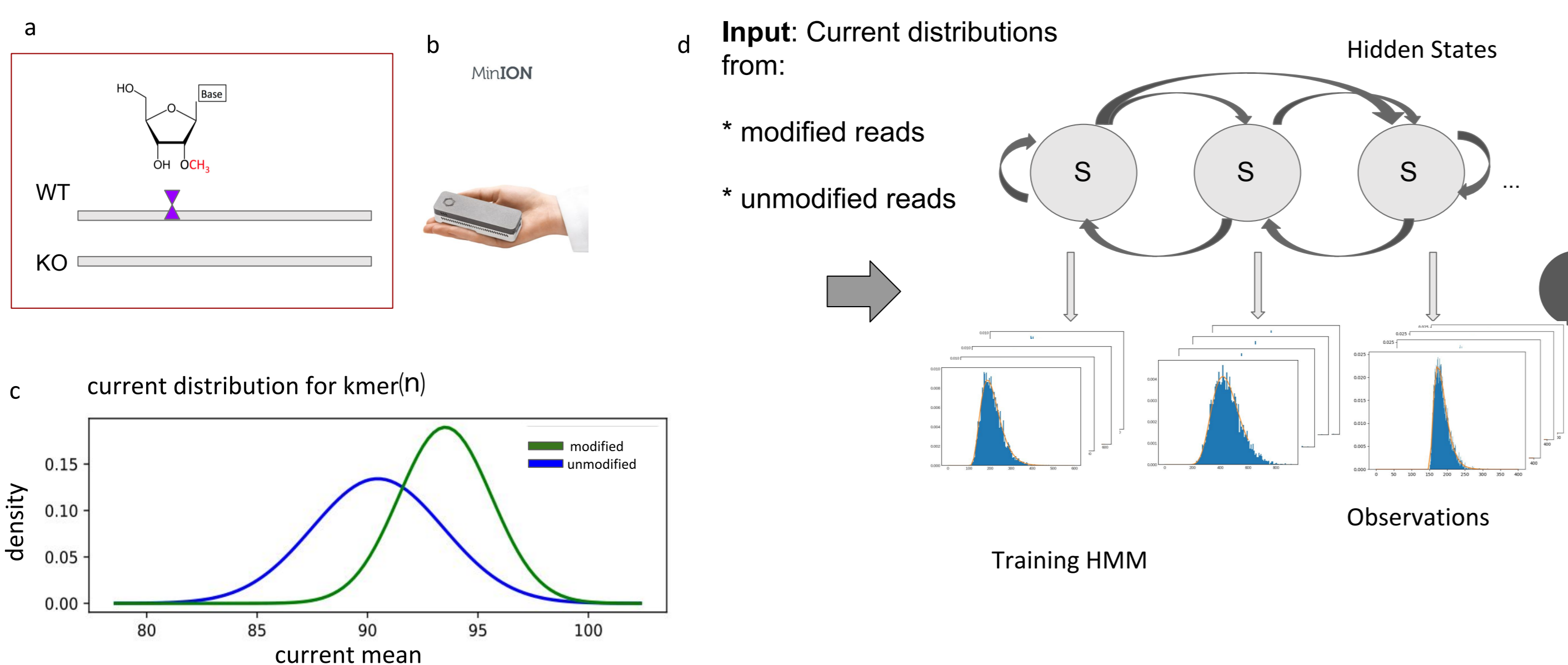


Fig.3: Schematic representation of the current workflow. a) Our libraries include a cell line carrying a methylation (WT) and a modified cell line which does not carry a specific methylation (KO) b) for direct RNA Sequencing with Oxford Nanopore Technologies MinION sequencer. c) After basecalling with Guppy (<https://nanoporetech.com>), and aligning reads to reference genome by Nanopolish (Simpson, Jared T., et al. "Detecting DNA cytosine methylation using nanopore sequencing." *nature methods* 14.4 (2017): 407-410), current distributions for the genomic regions of interest were extracted in k-mer resolution. d) A hidden Markov model (HMM) was implemented to predict the modification statuses of the given reads. Hidden states are consisting of kmers with their current distributions from modified and unmodified reads as their observables (Rabiner L R. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE.* 1989;77:257-286).

Approach

Building a prediction model to distinguish modified nucleotides from the rest

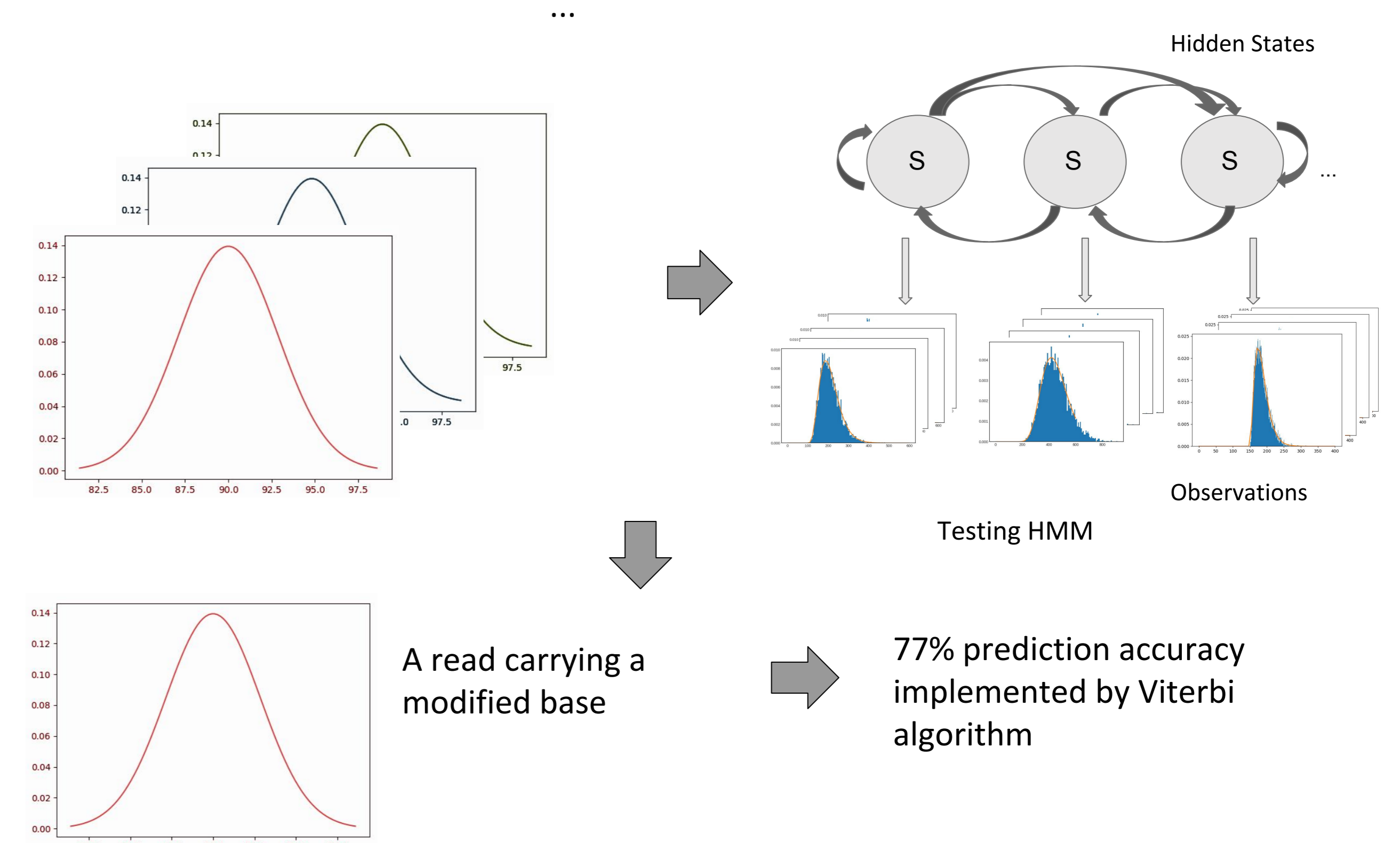
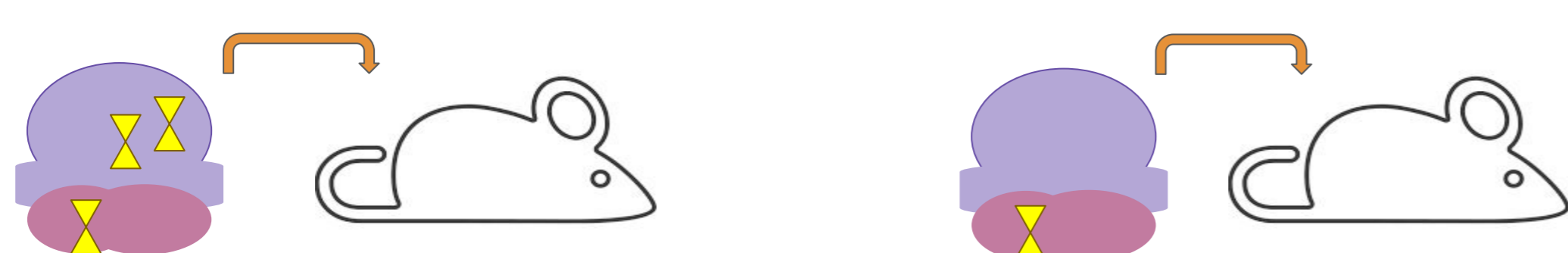


Fig.4: Schematic representation of the HMM predictions. Ionic current distributions from modified and unmodified reads were fed into the trained HMM for testing purposes. Viterbi algorithm implemented an HMM to find the most likely sequence of hidden states. The current prediction accuracy of the model is 77% for a single modification site (Churbanov, A., Winters-Hilt, S. Implementing EM and Viterbi algorithms for Hidden Markov Model in linear memory. *BMC Bioinformatics* 9, 224 (2008). <https://doi.org/10.1186/1471-2105-9-224>).

Results

Prediction of a single modification site with ~77% accuracy

- 1 Improving the model to increase its prediction accuracy
- 2 Extending the analyzed genomic region to cover all 2'-O-methylations on rRNAs
- 3 Assessing effects of 2'-O-methylations on translation via experimental validation



Future Steps & Experimental Validation

Importance & Conclusion

This work might provide unique and groundbreaking insight into how translation is regulated and coordinated in human cells through rRNA modifications. As the modification system is deeply conserved, we expect this potential new knowledge to be directly applicable across several kingdoms of life.