Title

A synergistic approach to high-throughput detection of RNA modifications using direct RNA sequencing and algorithmic analysis

Authors

Salvador Rodarte,[1] Luis Gracia Mazuca,[1] Isaac Weislow,[2] Davor Beltran,[2] Jonathon E. Mohl,[1,3] and Eda Koculi[2]

[1]Bioinformatics Program, [2]Department of Chemistry and Biochemistry, and [3]Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, TX, 79968

Abstract

RNA post-transcriptional modifications play critical roles in regulating various cellular processes, and when aberrant, can lead to severe health conditions such as cancer and neurodegenerative diseases. Of roughly 300 identified classes of RNA modifications, there remains a significant demand for methods that combine high throughput and single nucleotide resolution detection. In this research, the potential of the RNA direct-sequencing platform developed by Oxford Nanopore Technologies was harnessed, focusing on bacterial ribosomal RNA as a model molecule. While efforts led to the pinpointing of several classes of RNA post-transcriptional modifications, including pseudouridine (Ψ) and numerous nucleotide methylations, certain modifications like 5-hydroxycytosine (5HC) and others displaying consistent mutation rates between control and sample sets remain elusive. To enhance the method's robustness and accuracy, the capabilities of two pre-existing tools, Nanopore-PSU and EpiNano, were incorporated. Nanopore-PSU, developed for Ψ site prediction, exploits native content training and machine learning modeling to map Ψ modifications across varied transcriptomes. Simultaneously, EpiNano in its EpiNano-Error mode was utilized to derive crucial features from RNA sequencing reads—such as current intensity and mismatch frequency—that boost the reliability and accuracy of RNA modification detection. This approach, which integrates custom algorithms and external tools, not only augments comprehension of RNA modifications but also underscores the efficacy of employing diverse analytical methods to progress the domain of post-transcriptional RNA modification detection.