[BIOINFO 2023]

Pseudo-temporal analysis of ribonucleoprotein complex assembly using proximity labeling and nanopore sequencing

Heeseung Yoo[1,2], Keewon Sung[1], and Hyeshik Chang[1,2,*]

[1] *Center for RNA Research, Institute for Basic Science (IBS)*
[2] *Department of Bioinformatics, Seoul National University*
**Corresponding author: hyeshik@snu.ac.kr*

RNA binding proteins (RBPs) play a crucial role throughout the entire RNA life cycle. By binding to specific regions of RNA and modulating their stability, transport, and function, RBPs underscore the importance of identifying these proteins and their binding sites for a deeper understanding of RNA biology. Here, we developed a novel method that enables us to capture RNA-protein interactions in RNA-centric and long-range using Direct RNA sequencing. We utilized proximity labeling to attach chemical probes to RNA molecules as RBP footprints that are small enough to pass through nanopores, thereby eliciting significant differences in nanopore signals. The difference appears in signal drops and elongation on dwell time, which allows us to build a detection model. Therefore we developed a computational tool named DOMI, based on the Hidden semi-Markov Model (HsMM), that could detect RBP binding sites at the single-molecule level. We build HsMM with a control set and apply the Viterbi algorithm to both the control and treated set for calculating the maximum log-likelihood. RBP binding sites could be captured by their low log-likelihood caused by abnormal signals. By using DOMI, we explored cooperative RBP interactions and the biogenesis pathway of the Telomerase RNA Component (TERC). The trajectory analysis showed that the proteins bind to Telomerase RNA in the order from 3' to 5' hairpin in the H/ACA lobe. We expect that DOMI will be able to track the interactions between RBPs and RNAs along the RNA life cycle at the molecular level.