PolyA-miner assesses the effect of alternative polyadenylation on gene expression

July 23 2020

Researchers with an interest in unraveling gene regulation in human health and disease are expanding their horizons by closely looking at alternative polyadenylation (APA), an under-charted mechanism that regulates gene expression.

"APA is about modifying one of the ends, called the 3-prime end (3'end), of RNA strands that are transcribed from DNA. The modification consists of changing the length of a tail of adenosines, one of the RNA building blocks, at the 3'end before RNA is translated into proteins," said first author Dr. Hari Krishna Yalamanchili, a postdoctoral associate in the lab of Dr. Zhandong Liu at Baylor College of Medicine. "This adenosine chain helps to determine how long the messenger RNA lasts in the cell, influencing how much protein is produced from it."

The interest in APA has resulted in the development of several 3' sequencing (3'Seq) techniques that allow for precise identification on APA sites on RNA strands. But what researchers are missing is a robust computational tool that is specifically designed to analyze the wealth of 3'Seq data that has been generated.

Meet PolyA-miner

"Until now, researchers have been using traditional RNA sequencing computational tools to analyze the 3'Seq datasets. Although this approach produces results, it does not maximize the potential amount of

information that can be extracted from that data," Yalamanchili said. "Here we developed a computational tool that precisely analyzes 3'Seq data. We call it PolyA-miner."

Yalamanchili and his colleagues used their new computational tool to analyze existing 3'Seq datasets. PolyA-miner not only reproduced the analyses achieved with traditional computational tools, but also identified novel APA sites that were not detected with the other analytical approaches.

"We were surprised when the PolyA-miner analysis of a glioblastoma cell line dataset identified more than twice the number of genes with APA changes than were initially reported," Yalamanchili said.

"I think that the most exciting part of this new tool is that it enables us to precisely reflect gene-level 3' changes and to identify many more APA events than before. With other analytical approaches, we underestimate the effect and number of poly-adenylation events," said Liu, associate professor of pediatrics and neurology at Baylor and the Jan and Dan Duncan Neurological Research Institute at Texas Children's Hospital.

Immediate applications

This development has tremendous implications for basic research and for the potential translation of scientific findings into the clinic. APA is considered a major mechanism for RNA regulation that has strong relevance both in cancer and neurological diseases. PolyA-miner can assist scientists looking to identify the genetic causes of these diseases by determining whether there are differences in APA between diseased and normal cells. With this new analysis, scientists can take a fresh look at existing genomic datasets that may provide an answer to the cause of human conditions, as well as studying newly developed datasets.

"Previously, people knew about APA changes, but did not consider them to be major contributors to gene regulation mainly because we lacked the computational tools to determine APA's overall influence on gene expression," Yalamanchili said. "PolyA-miner has shown that APA seems to play a larger role in gene regulation than we had previously thought."

More information: Hari Krishna Yalamanchili et al, PolyA-miner: accurate assessment of differential alternative poly-adenylation from 3′Seq data using vector projections and non-negative matrix factorization, *Nucleic Acids Research* (2020). DOI: 10.1093/nar/gkaa398

Provided by Baylor College of Medicine

Citation: PolyA-miner assesses the effect of alternative polyadenylation on gene expression (2020, July 23) retrieved 25 April 2024 from https://phys.org/news/2020-07-polya-miner-effect-alternative-polyadenylation-gene.html