A NUANCED LOOK AT HOW GENES REGULATE HUMAN DISEASE

Researchers at Shenzhen Bay Laboratory are developing BIOINFORMATICS APPROACHES AND TOOLKITS to identify disease-relevant genes for precision medicine.

jxfzsy/iStock/Getty

▲ **Taking a deep dive into the genetic basis for human diseases is made easier through Shenzhen Bay Laboratory's strong investment in computational approaches.**

Responses to prescribed medication may vary due to genetic differences. Yet, surprisingly little is known about genetic regulatory mechanisms that underly complex traits and diseases in humans. A group of researchers at the Computational and Disease Genomics Lab of Shenzhen Bay Laboratory (SZBL) is developing bioinformatics methods to investigate genetic regulation, aiming to uncover the genetic causes of human diseases.

"The focus of my wet and dry lab research is to decipher the genetic basis of disease-associated transcriptional and post-transcriptional regulations," says Lei Li, a computational biologist and the lab's principal investigator.

"We use quantitative techniques on large-scale, high-throughput, multi-dimensional data to provide new directions for discovering driver genes and potential therapeutics," says Li. Driver genes are important as their mutations directly impact the biology of cancer cells.

More than 70% of human genes are associated with alternative polyadenylation (APA), whereby alternative cleavage and polyadenylation sites of a single gene can generate many mRNA isoforms that have different 3'untranslated regions (3'UTRs). Since these diverse 3'UTRs contain multiple cis-regulatory elements such as miRNA binding sites, many target RNAs possess different functions, stability and translation efficiency, says Li.

Li was among the first researchers to use computational approaches to delve into RNA sequencing (RNA-seq) data with an 'APA-aware' approach.

In 2019, Li was involved in a major breakthrough that identified a master APA regulator that induces 3'UTR shortening in many tumour types. The 3'UTR alterations between control and regulated cells were profiled using 2.5 × 10^8 reads of RNA-seq and bioinformatics algorithms, shedding light on the molecular mechanisms driving APA in cancer.

To build the capacity of multithreading and joint analysis of multiple samples without relying on control samples, Li further optimised the algorithms. He constructed a database by analysing more than 8,000 RNA-seq datasets, coupled with whole-genome sequencing genotype data derived from 46 tissues and isolated from 467 individuals.

These analytical tools were used by Li's team to demonstrate that 3'UTR APA quantitative trait loci (3'aQTLs) co-localized with about 16.1% of trait-associated variants and may play a role in the molecular mechanisms that cause Alzheimer's disease and rheumatoid arthritis.

Li's research results have attracted interest from pharmaceutical companies, resulting in projects that are jointly developing drugs targeting APA for clinical applications. "We continue to identify new genes that directly contribute to diseases with the goal of identifying new druggable targets for precision medicine," says Li.

"Next steps include developing new computational tools for analysing and integrating multi-level single-cell data. Potentially, this could reveal cell-to-cell heterogeneity during disease development, treatment, and drug response for individual patients." ■

深圳湾实验室
Shenzhen Bay Laboratory

Contact info
+86-755-86726648
hr_pi@szbl.ac.cn
www.szbl.ac.cn