GMI's director, Jong-Il Kim, looking at a specimen through a microscope.

PUTTING SOUTH KOREA ON THE GENOMIC MAP

A DATABASE OF GENOMIC VARIATIONS will greatly improve the search for disease-associated genetic mutations in Asian populations.

Since publishing the first whole-genome sequence of a South Korean individual in *Nature* in 2009, researchers at the Genomic Medicine Institute (GMI) in Seoul National University have been focusing on deciphering the genomic traits of the Asian population and on uncovering Asian-specific links to disease. With just a small fraction of the genomic variations recorded in the world's major databases being of Asian origin, GMI has taken the ambitious step of building its own database and research platform that targets Asian genomic variations and their associations with gene expression at the single-cell level.

"One of the biggest challenges in genome research for disease is accurately distinguishing true causal genomic variants from a sea of unfounded correlations," says Kim.

When the genomic variation is within the coding region of a specific gene, the research approach is usually straightforward. Scientists can use previously acquired biological knowledge, such as the functional role or structure of the protein coded by the mutant gene. They can determine how the variation may change the amino acid sequence and thus alter the protein's physical structure. This allows them to infer the functional change of the protein and relate it to the phenotype or the disease in question.

However, when the genomic variation is not on the coding region of a gene but affects the cell's biological processes, the variant will mostly likely affect the expression levels of certain genes near or far from it. Such an association is known as an expression quantitative trait locus (eQTL).

"Gene expression differs based not only on the organ or the tissue type, but also on the different cell types within a certain tissue," explains GMI's director, Jong-Il Kim. "For example, lung tissue consists of various cell types such as basal cells, secretory cells, ciliated cells, alveolar cells and immune cells, to name just a few. Using single-cell RNA sequencing, we can collect gene-expression data specific for each cell or cell type, which can then help us search for a statistically significant association between the presence of a genomic variation and the expression level of a certain gene." Well-established single-cell eQTLs database will enable researchers to search in a more efficient and targeted way for genetic mutations linked to disease.

Researchers at GMI are building a single-cell eQTL database by cultivating 'organoids' — cultures of stem cells collected from biopsy tissues — for lung and stomach tissues. Using organoids for single-cell RNA sequencing assures consistency and high quality, since all sequencing experiments are conducted under lab conditions. "Furthermore, this organoid bank with known genomic information will be a great resource for further experiments for data collection of other phenotypes such as response to viral infections," says Kim.

"A database of single-cell eQTLs will enable us to infer which non-coding variations alter the expression of which genes in which organs or cell types, and how strong those associations are," says Kim. "With all these tools and new knowledge, we will be better equipped to accurately find meaningful genomic variants with causal links to disease."

This project is funded by the Ministry of Education and the Ministry of Science and ICT, Republic of Korea. ■

서울대학교 유전체의학연구소
Genomic Medicine Institute, Seoul National University

gmi.snu.ac.kr