Pseudogenes may provide clearer understanding of biomarkers

7 July 2014

Alas, the thankless pseudogene. Dysfunctional, unloved and seemingly of little use, these poor-cousin relatives of genes have lost their protein-coding abilities. They contain material not essential for an organism's survival and are the "last stop" for removal of genomic waste.

Not any more. The pseudogene's day may have arrived thanks to scientists at The University of Texas MD Anderson Cancer Center in Houston.

Han Liang, Ph.D., an assistant professor in the Department of Bioinformatics and Computational Biology at the Cancer Center is advancing knowledge of these largely overlooked but increasingly attractive genetic oddities. He and his team completed a study that generated pseudogene expression profiles in 2,808 patient samples representing seven cancer types. That meant analyzing 378 billion RNA sequences to measure the expression levels of close to 10,000 pseudogenes.

The results indicated that the science of pseudogene expression analysis may very well play a key role in explaining how cancer occurs by helping medical experts in the discovery of new biomarkers. The study's findings appear in today's issue of *Nature Communications*.

Understanding of biomarkers is important for developing therapies that targeted specific tumor sites and for gaining new insight into how patients will fare with various cancers and treatments. Biomarkers are molecules that can indicate the presence of a condition or disease, and are increasingly being used to measure how the body responds to therapies. The emerging field of personalized medicine is built on customizing treatment for patients based on biomarkers.

Liang's study is novel in that understanding of pseudogenes relies on analyzing large numbers of patient samples. Previous studies have been limited by the size of the patient sample groups.

Liang's team analyzed data made available from The Cancer Genome Atlas research program. The program is supported by the National Cancer Institute and National Human Genome Research Institute within the National Institutes of Health and is looking at genomic changes in more than 20 different types of cancer.

"The study surveyed seven cancer subtypes including those for breast, kidney, ovarian, colorectal, lung and uterine," said Liang. "Across the cancer types, the tumor subtypes revealed by pseudogene expression showed extensive and strong similarities with subtypes defined by other molecular data."

Liang believes that the study highlights the potential of pseudogene expression analysis as a new "gold standard" for investigating cancer mechanisms and discovering prognostic biomarkers. These biomarkers will allow medical experts to more accurately predict cancer survival rates.

"Pseudogene expression alone can accurately classify the major subtypes of endometrial cancer," said Liang. "Strikingly, in kidney cancer, the pseudogene expression subtypes not only significantly correlate with patient survival, but also help stratify patients in combination with clinical variables."

Provided by University of Texas M. D. Anderson Cancer Center