

JIANHUA LUO

UNIVERSITY OF PITTSBURGH SCHOOL OF MEDICINE,
USA



Serum fusion transcripts based machine learning model to predict the occurrence of hepatocellular carcinoma

Abstract:

Hepatocellular carcinoma (HCC) is one of the most lethal malignancies for human. Early diagnosis of HCC is crucial in reducing the mortality of the disease. In this study, a panel of 9 fusion transcripts in the serum samples from 136 individuals using TaqMan qRT-PCR was analyzed. Seven fusion genes were frequently detected in the serum samples of HCC patients, including MAN2A1-FER (100%), SLC45A2-AMACR (62.3%), ZMPSTE24-ZMYM4 (62.3%), PTEN-NOLC1 (57.4%), CCNH-C5orf30 (55.7%), STAMBPL1-FAS (26.2%) and PCMTD1-SNTG1 (16.4%). Machine learning models were constructed based on serum fusion gene levels to predict HCC occurrence in the training cohort using leave-one-out-cross-validation approach. One of the machine learning models called 4-fusion genes logistic regression model (MAN2A1-FER<40, CCNH-C5orf30<38, SLC45A2-AMACR<41, PTEN-NOLC1<40) produced a 91.5% accuracy in the training cohort. The same model generated an accuracy of 83.3% in the testing cohort. When serum α -fetal protein (AFP) level was incorporated into the machine learning model, a 2-fusion gene+AFP logistic regression model (MAN2A1-FER<40, CCNH-C5orf30<38, AFP) was found to generate an accuracy of 94.8% in the training cohort. The same model generated 95% accuracy in both the testing cohort and the combined cohorts. Cancer treatment reduced most of the serum fusion transcript levels. Serum fusion gene machine learning models may serve as important tools in screening HCC and monitoring the impact of HCC treatment.

Biography

Dr. Luo has been studying molecular mechanisms of human malignancies in the last 36 years. Currently, he is a Professor of Pathology and Director of High Throughput Genome Center at University of Pittsburgh. In the last 30 years, Dr. Luo has been largely focusing on the genetic and molecular mechanism of human cancers such as prostate cancer. He is one of the pioneers in utilizing high throughput gene expression and genome analyses to analyze field effects in prostate cancer and liver cancer. He is also the first in using methylation array and whole genome methylation sequencing to analyze prostate cancer. He and his colleague helped to develop an ultra-low error synthetic long-read sequencing technology called LOOPSeq that can be utilized to quantify mRNA isoforms and mutation isoform distributions in single cell level. His group has discovered 21 novel fusion genes in prostate, liver and colon cancers. Subsequently, his group discovered that many of these fusion genes are recurrent in many other types of human cancers. His group also developed a genome intervention strategy targeting at the chromosomal breakpoint of fusion gene to treat cancers. Recently, his group has been focusing on machine learning prediction of prostate and liver cancers. Overall, these findings advance our understanding of how cancer develops and behaves, and lay down the foundation for better future diagnosis and treatment for human malignancies.