A powerful method for the estimation of cancer-driver genes using a weighted iterative zero-truncated negative-binomial regression

Miaoxin Li

Sun Yat-sen University
E-mail: limiaoxin@mail.sysu.edu.cn

Abstract: Genomic identification of driver mutations and genes in cancer cells are critical for precision medicine. Due to difficulty in modeling distribution of background mutations, existing statistical methods are often underpowered to discriminate driver genes from passenger genes. Here we propose a novel statistical approach, weighted iterative zero-truncated negative-binomial regression (WITER), to detect cancer-driver genes showing an excess of somatic mutations. By solving the problem of inaccurately modeling background mutations, this approach works even in small or moderate samples. Compared to alternative methods, it detected more significant and cancer-consensus genes in all tested cancers. Applying this approach, we estimated 178 driver genes in 26 different cancers types. In silico validation confirmed 90.5% of predicted genes as likely known drivers and 7 genes unique for individual cancers as very likely new drivers. The technical advances of WITER enable the detection of driver genes in TCGA datasets as small as 30 subjects, rescuing more genes missed by alternative tools. The tool is available at http://grass.cgs.hku.hk/limx/witer/ and http://grass.cgs.hku.hk/limx/kggseq/.