

MSIpred: A Python 2 Package for the Classification of Tumor Microsatellite Instability from Tumor Mutation Annotation Data Using a Support Vector Machine

Authors : Chen Wang, Chun Liang

Abstract : Microsatellite instability (MSI) is characterized by high degree of polymorphism in microsatellite (MS) length due to a deficiency in mismatch repair (MMR) system. MSI is associated with several tumor types and its status can be considered as an important indicator for tumor prognostic. Conventional clinical diagnosis of MSI examines PCR products of a panel of MS markers using electrophoresis (MSI-PCR) which is laborious, time consuming, and less reliable. MSIpred, a python 2 package for automatic classification of MSI was released by this study. It computes important somatic mutation features from files in mutation annotation format (MAF) generated from paired tumor-normal exome sequencing data, subsequently using these to predict tumor MSI status with a support vector machine (SVM) classifier trained by MAF files of 1074 tumors belonging to four types. Evaluation of MSIpred on an independent 358-tumor test set achieved overall accuracy of over 98% and area under receiver operating characteristic (ROC) curve of 0.967. These results indicated that MSIpred is a robust pan-cancer MSI classification tool and can serve as a complementary diagnostic to MSI-PCR in MSI diagnosis.

Keywords : microsatellite instability, pan-cancer classification, somatic mutation, support vector machine

Conference Title : ICBMB 2018 : International Conference on Bioinformatics and Molecular Biology

Conference Location : New York, United States

Conference Dates : June 03-04, 2018