ID: 362

Prediction of Liver Cancer Using *Tp53* Polymorphisms: A Comparison of Advanced Machine Learning Models

Saba Rasheed[1], Haris Khurram[2], Aqsa Parvaiz[1], Rehan Sadiq Shaikh[3], Saba Zafar[1*]

[1] Department of Biochemistry and Biotechnology, The Women University, Multan, Pakistan
[2]Department of Sciences and Humanities, National University of Computer and Emerging Sciences, Chiniot-Faisalabad Campus, Pakistan.
[3]Center for Applied Molecular Biology, University of the Punjab, Lahore, Pakistan

Abstract

Globally, liver cancer ranks second in terms of cancer-related deaths and is the sixth most common type of cancer diagnosed. Most initial liver malignancies (70%–90%) are hepatocellular carcinomas. The early detection of HCC can be challenging, as it tends to develop without obvious symptoms. Liver cancer has several risk factors like other cancer forms, both preventable and unavoidable. The risk factors for liver cancer may be demographic, environmental, clinical, and genetic.

We investigated single nucleotide polymorphism (SNP) in the *Tp53* gene with liver cancer in the southern Punjab population of Pakistan. Two SNPs of Tp53, rs9895829 and rs9891744 were selected to investigate their impact on liver cancer development. We additionally investigated several clinical and demographic factors, such as age, diabetes, BMI, cancer family history, smoking, HCV, HBsAg, and haplotypes affected by liver cancer. We recruited 100 individuals for this study including cases and controls. Tetra-Arms PCR polymerase chain reaction technique was used for genotyping. The impact of genetic and non-genetic factors on the development of liver cancer was predicted by using artificial neural networks (ANNs) machine learning models, the multilayer perceptron model, and the radial basis function neural network model. Further, factor importance is also predicted by using the classification regression tree (CRT) model. Our results indicated that age, smoking, BMI, and SNP of *Tp53* rs9891744 (C/T) may be the most contributing risk factors for causing liver cancer in the population of Southern Punjab, Pakistan. This information may help the medical practitioner to predict the development of liver cancer and help for the generation of personalized medicine.

Key Words: Liver cancer, Single nucleotide Polymorphism, Artificial neural network, Classification regression method, TP53