ANN based Method for Prediction of Multidrug and Toxin Extrusion (MATE) Proteins

Tamanna[1], Jayashree Ramana[2]

[1] Tamanna, Research Scholar, Department of Biotechnology and Bioinformatics, Jaypee University of Information and Technology, Waknaghat, Solan (Himachal Pradesh)

[2] Dr. Jayashree Ramana, Assistant Professor (Grade-II) Department of Biotechnology and Bioinformatics, Jaypee University of Information and Technology, Waknaghat, Solan (Himachal Pradesh)

[1]tanu_tanu14@yahoo.com, [2]jayashree.ramana13@gmail.com

ABSTRACT

Multidrug And Toxin Extrusion (MATE) proteins form a class of proteins that function as drug and proton antiporters. It plays a vital role in the secretion of cationic drugs across the cell membrane. MATE transporters confer resistance to bacterial pathogens and cancer cells, thus causing critical reductions in the curative efficacies of drugs. Due to the increasing problem of drug resistance, control of infectious diseases is becoming more unmanageable. Hence, it is crucial to understand the underlying causes of drug resistance and devise novel solutions to address this problem. In this work, we propose Artificial Neural Network (ANN) based method for prediction of MATE proteins using SNNS package. The data set employed for training consists of 189 non-redundant protein sequences, that are further classified as positive (63 sequences) set comprising of sequences from MATE family, and negative (126 sequences) set having protein sequences from the other transporter families. First, we used different features (Amino Acid Composition [AAC], Dipeptide Composition [DC] and Charge Composition [CC]) for generating networks. However, it was observed that the performance of these network were poor when compared to AAC network. The fully-connected network was derived using amino acid composition, with 20 nodes in the input layer, 12 in the hidden layer and 1 in the output layer. The training was carried out for 7000 cycles. Random weights were used for initializing the network and Standard Backpropagation algorithm was used to minimize the differences between the computed output and the target value. This yielded an overall accuracy of 84.45%, with sensitivity and specificity of 64.90% and 99.22% respectively at a threshold value of 0.4 in three fold cross validation. We anticipate that this work will aid rapid and rational identification of MATE proteins, and would be a useful tool for the research community.

Keywords: Antibiotics, drug resistance, MATE, ANN.