INDRAPRASTHA INSTITUTE of INFORMATION TECHNOLOGY DELHI

10 YEARS of EXCELLENCE

Invitation to M.Tech. Thesis Defense of Dikscha Sapra: June 05, 2020 (Friday): 10.00-11.00 IST

In Partial Fulfillment of the Requirements for the Degree of

M.Tech. CB

Dikscha Sapra (MT18245)

Will defend her thesis

Title: "**Compilation and Prediction of Prostate Cancer Biomarkers**"

IIIT-D Faculty and Students are invited

Date: June 05, 2020 (Friday)
Time: 10.00-11.00 IST
Place: Online (Google Meet)

Examiner:	**Internal:**	**Gaurav Ahuja**
	External/Internal:	**Dinesh Gupta (ICGEB)**
	Advisor:	**G P S Raghava**

****

Abstract

"Prostate Cancer is the second leading cause of cancer related death in men worldwide. There are several biomarkers present in literature that help in the diagnosis and prognosis of the disease. However there does not exist a resource that can consolidate the information contained in all the published literature. We present- ProCanBio a manually curated database containing information from 861 published articles, and having a total of 1890 entries from 1302 biomarkers related to genomics, epigenomics, proteomics and metabolomics of the disease. It contains detailed information regarding each biomarker from the chosen literature. To analyse the information in the database, a user friendly web server with searching and browsing capabilities was developed which is freely available and is compatible with most web browsers and devices. ProCanBio can be highly useful for the research community in finding most commonly used biomarkers in prostate adenocarcinoma.

ProCanBio also consists of Stage prediction tools, which can help users to classify a Prostate Cancer sample into early and late stage based on a miRNA signature consisting of 20 unique miRNAs. This model was trained on TCGA cohort consisting of 487 samples containing a total of 1046 miRNAs. We split this data into 80:20 such that the same ratio of early and late stage samples is maintained in training and testing. Using SelectKBest for feature selection, and SVM for training our data, we achieved 74% accuracy and 0.80 ROC-AUC on our training data (5 fold cross validated), and 73% accuracy and 0.73 ROC-AUC for validation data. Users of ProCanBio can predict the stage of a Prostate Cancer patient(s) by entering the miRNA based signature in a csv file.

ProCanBio is available at: https://webs.iiitd.edu.in/raghava/procanbio/"