INDRAPRASTHA INSTITUTE *of* INFORMATION TECHNOLOGY **DELHI**

10 YEARS *of* EXCELLENCE

Invitation to M.Tech. Thesis Defense of Neha Mishra: May 27, 2020 (Wednesday): 12.00-13.00 IST

In Partial Fulfillment of the Requirements for the Degree of

M.Tech. CB

Neha Mishra (MT18239)

Will defend her thesis

Title: "**Assessment of drug efficacy in Indian population: A pharma-genomic study**"

IIIT-D Faculty and Students are invited

Date: May 21, 2020 (Wednesday)
Time: 12.00-13.00 IST
Place: Online (Google Meet)

Examiner:	**Internal:**	**Gaurav Ahuja**
	External/Internal:	**Sridhar Sivasubbu, CSIR-IGIB**
	Advisor:	**Arjun Ray**

****

Abstract

India with a population of more than 1.2 billion has always been a great interest in genetic studies. As most of the Indian subpopulations are still following the tradition of endogamy and social stratification; it ensures that this diversity contains a treasure of clinically relevant rare mutations which may have evolved differently in different subpopulations. Previous studies have categorized the Indian population into four dominant subgroups based on ancestry dominance based on the samples taken from mainland India. These subgroups are Ancestral North-Indian (ANI), Ancestral South-Indian (ASI), Ancestral Tibeto-Burman (ATB) and Ancestral Austro Asiatic (AAA). The 1000 Genome data phase-III on the other hand contains around 500 samples of Indian ethnicity belonging to five linguistically defined Indian subcontinent populations i.e. Punjabi, Gujrati, Bengali, Telugu and Tamil. Some of these samples even include recent migrants of India to the USA or UK. Keeping the role of epigenetic in causing genetic variants in mind, these datasets are biased toward European populations and failed to address the Indian populations in these databases. Most of the research outcomes are based on these databases only and are most likely to be biased towards US and UK based populations. India is also the second largest market for US and UK based drugs. The human trials of these drugs are mostly done on European or American populations. The current healthcare system also works in a trial and error prescription model, where each patient is prescribed with the same drug therapy regardless of the genetic diversity Indian population contains. These prescriptions end up causing adverse drug reactions in patients which cost a lot of money and health hazards.

We did a systematic qualitative and comparative study on kinase coding genes as it is the second most targeted group of drug targets. We used the amino acid sequence data to analyze the population specific variants in these genes. Using homology modelling we studied the three-dimensional structure of these genes to study the effect of these variants in protein function, stability, protein-ligand binding and disease. Our study shows a number structural and functional changes caused by single nucleotide variants at the structure level including change in structure stability, binding and pathway enrichment analysis. Whereas, the sequence level analysis is mainly focused on finding the similarities and differences among different populations based on the SNVs and amino acid exchange frequencies. The comparison between native and mutant models provide an insight into atomic-level structural differences caused by single amino acid exchange. In-silico screening of the known drugs to these proteins reveal critical differences imparted in the strength of binding due to the variations present in the Indian population.