Name: Aftab Alam
Supervisor: Dr. Romana Ishrat
Centre: Centre for Interdisciplinary Research in Basic Sciences

Title: **Network Meta-Analysis of Human Gene Expression Profiles to Identify Potential Therapeutic Targets for Tuberculosis Infection**

Tuberculosis (TB) has been one of the ancient human diseases for centuries and still is one of the major causes of mortality in many developing countries. The TB disease is mainly caused by Mycobacterium tuberculosis, a highly elusive bacillus. The disease tuberculosis is categorized as active or latent; active TB, sometimes referred to as TB disease, has been developed, causes symptoms, and is contagious. However, TB's latent form is very common, not contagious; people are infected with M. tuberculosis but do not have TB-like symptoms. About one-fourth of the world's population has latent TB, and individuals with latent TB infection have a 5–15% lifetime risk of becoming active TB patients. M. Tuberculosis does not possess the classic bacterial virulence factors. However, this intracellular pathogen efficiently dodges the immune response by complex and devious mechanisms, facilitating the survival of the pathogen in the host for as long as decades. In our first objective, we have analyzed gene expression data (microarray data) and compared the gene expression profiles of patients with different datasets of healthy control, latent infection, and active TB. We observed the transition of genes from normal condition to various stages of the TB and identified and annotated those genes/pathways/processes that have important roles in TB disease during its cyclic interventions in the human body. We identified 488 genes that were differentially expressed at various stages of TB and allocated to pathways and gene set enrichment analysis. These pathways as well as GSEA's importance were evaluated according to the number of DEGs presents in both. In addition, we studied the gene interaction networks that may help to further understand the molecular mechanism of immune response against the TB infection and provide us a new angle for future biomarker and therapeutic targets. In this study, we identified 31 key regulator genes, of which 14 genes were up-regulated and 17 genes were down-regulated. The proposed approach is based on gene-expression profiling, and network analysis approaches predict some unknown TB associated genes, which can be considered (or can be tested) as reliable candidates for further (in vivo/in vitro) studies. In our second objective, we again analyzed microarray data sets and compared the transcriptome profiling of the healthy individual with latent infection (LTBI) and active TB (TB) patients and identified the differentially expressed genes (DEGs) but using a different approach. Here, we have built a

gene interaction network using differentially expressed genes which obtained from latent and active TB patients. Our goal was to identify key genes and pathways associated with tuberculosis infection by network centrality analysis which measure the most influencing gene (potential key regulators) in the gene interaction network. We have identified seven prominent hub genes namely IL6, IL1B, TNF, NFKB1, STAT1, JAK2 and MAPK8 with higher centrality value which are involved in various biological processes like cell cycle control, apoptosis, pro-inflammatory responses, complement signalling enhanced cytokine & chemokine signalling and host immune responses. Additionally, we also identified 22 inferred genes that are mainly engaged in the induction of innate immune response, cellular response to interleukin-6, inflammatory response, apoptotic process, I-kappa B phosphorylation, JAK-STAT signalling pathway, macrophage activation, cell growth, and cell signalling. The proper attention of these inferred genes may open up a new horizon to understand the defensive mechanisms of TB disease. The transcriptome profiling and network approach can enhance the understanding of the molecular pathogenesis of tuberculosis infection and have implications for the plan and execution of mRNA expression tools to support early diagnostics and treatment of Mycobacterium tuberculosis (M. TB). Our third objective focuses on network medicine approach, as we know that Tuberculosis (TB) is the leading cause of death from a single infectious agent. The estimated total global TB deaths in 2019 were 1.4 million. The decline in TB incidence rate is very slow, while the burden of noncommunicable diseases (NCDs) is exponentially increasing throughout the world. Emerging empirical evidence justify the convergence of TB with NCDs. The current study was proposed to build a disease-gene network based on overlapping TB with NCDs, viz Parkinson's disease, Cardiovascular disease, Diabetes mellitus, Rheumatoid arthritis, and Lung cancer. We compared the TB-associated genes with genes of its overlapping NCDs to determine the gene-disease relationship. Next, we constructed the protein-protein interaction network of disease-genes by integrating curated and experimentally validated interactions in humans and the find the 13 most important modules which contains A total of 86 key regulatory genes that are commonly associated with TB, and its overlapping NCDs were identified, and further their involvement in diverse biological processes and pathways was endorsed by gene ontology (GO) enrichment analysis. Additionally, the identified genes and their respective drugs were exploited for building a bipartite network that assists in deciphering the drug-target interaction that highlights influential roles for such drugs on apparently unrelated targets and pathways via direct and indirect interactions. The current findings insight into the overall molecular picture of TB and its overlapping NCDs and to consider the fact that these diseases are somewhere associated

with each other at the gene level. This study provides the basis for an alternative approach to investigate new targets and multidrug treatment in TB and other overlapping NCDs. Since, it is important for a network-based approach to disease is to find disease module for the patho phenotype of interest, that can guide further investigational work towards finding the disease mechanism. Therefore, this study provides a holistic understanding of TB-associated comorbidities for the identification of potential drug candidates, that can be synergistically used with current therapies and additional support for the patients with immunity related conditions. Finally, we acknowledge several potential limitations. First, although we integrated data from multiple sources to build the interactome and the drug–target network, that relied on host gene/protein/disease datasets, and their quality and literature bias may influence the performance and the results of network analysis. Second, our method can only be applied to diseases with well-characterized genetic information and may not be applicable for diseases that lack such information, such as rare diseases (i.e., cerebral palsy or mental conditions). Potential literature bias of disease associated genes and the human interactome may also influence our findings.