Thesis Proposal

University of Rajshahi

Project Name: Identification of comorbidities for type-2 diabetic patients by analyzing gene expression data

Course Code : CSE4192

Course Title : Thesis/Project (Part1)

Submission date: 5th October, 2023

Submitted By:

Ashiq Uddin Pranto

Roll : 1911176136

Year : 4th Year Odd semester

Session:2018-19

Department of Computer

Science and Engineering,

University of Rajshahi

Supervised By:

Md. Humayun Kabir

Associate Professor

Department of Computer

Science and Engineering,

University of Rajshahi

_____

Signature of Supervisor

Introduction:

Diabetes is a medical condition characterized by high levels of glucose(sugar) in the blood. Among the two types of diabetes(type-1 and type-2), Type 2 diabetes(T2D) is the most severe form of diabetes. The majority of diabetes patients suffer from type 2 diabetes worldwide. In type 2 diabetes the body's cells do not respond effectively to insulin, which normally push the body cells to take glucose from the bloodstream. As the cell doesn't respond, the level of glucose increases in the bloodstream.

Diabetes often occurs in conjunction with other medical conditions, known as comorbidities. For example hypertension(High Blood Pressure), obesity, cardiovascular disease etc. These comorbidities can significantly impact the overall health and treatment of T2D patients. That's why understanding the specific comorbidities that commonly occur with T2D is essential for providing more effective treatment.

Literature Review:

In this study, gene expression differences were observed between tuberculosis (TB) patients and those with tuberculosis-diabetes comorbidity (TB-DM) at diagnosis[1]. While TB treatment influenced gene expression in all groups, TB-DM patients showed persistently distinct expression patterns throughout treatment. Specifically, genes associated with immune responses and inflammation were significantly upregulated in TB-DM patients, indicating heightened immune activity during treatment.

This study investigates differences among pancreatic β-cells in Type II Diabetes (T2D)[2]. Using advanced techniques to study individual cells, the research uncovers variations in β-cells related to T2D. They also find that a specific gene called HNF1A is important in driving these differences. Interestingly, people with T2D have lower HNF1A, which affects the flow of sodium in the cells. Another gene called FXYD2 plays a significant role. This research helps us understand T2D better by looking at individual cells and their diversity.

Research Objectives:

This research proposal aims to identify and characterize comorbidities associated with T2D through the analysis of gene expression data using RNA-seq. Additionally, we seek to understand the potential molecular mechanisms, conduct a detailed analysis of gene expression

patterns to discover genes responsible for diabetes, and determine the most prevalent comorbidities among patients diagnosed with T2D within the selected dataset.

Datasets:

The dataset for this research is collected from publicly available repositories, notably the NCBI Gene Expression Omnibus (GEO) repository[3]. It comprises RNA-seq data from patients diagnosed with Type 2 Diabetes (T2D) along with information regarding the presence or absence of comorbidities. The dataset is chosen for its diversity and represents a valuable resource for understanding the molecular mechanisms associated with T2D and its comorbidities. The dataset I collected till now GSE193978, GSE234754, GSE142025, GSE97655, GSE137317.

Methodology:

We will initiate our study by collecting RNA-seq data from patients with Type 2 Diabetes (T2D), both with and without comorbidities, sourced from publicly available datasets in the NCBI GEO repository. Upon obtaining the datasets, our first step will involve a quality check and preprocessing of the gene expression data. This rigorous process aims to ensure data quality and maintain consistency throughout the analysis.

Subsequently, we will conduct an extensive data analysis to identify genes that exhibit high expression levels and are responsible for specific diseases or comorbidities. This investigation will allow us to uncover the relationships between gene expression patterns and disease conditions.

Conclusions:

In conclusion, this research proposal tries to mention the critical issue of Type 2 Diabetes (T2D). T2D is a severe form of diabetes, and the co-occurrence of comorbidities significantly complicates the management and treatment of T2D patients. Identifying the genes responsible for this disease holds the potential to provide valuable insights for healthcare professionals. These insights can guide the development of more effective strategies to mitigate the risk of diabetes and improve the overall quality of patient care. This research represents a vital step toward addressing the challenges posed by T2D and its associated comorbidities, ultimately contributing to better health outcomes for individuals affected by this condition.

References:

1. Eckold C, van Doorn CLR, Ruslami R, et al. Impaired resolution of blood transcriptomes through tuberculosis treatment with diabetes comorbidity. *Clin Transl Med*. 2023;13(9):e1375. doi:10.1002/ctm2.1375

2. Weng C, Gu A, Zhang S, et al. Single cell multiomic analysis reveals diabetes-associated β-cell heterogeneity driven by HNF1A. *Nat Commun*. 2023;14(1):5400. Published 2023 Sep 5. doi:10.1038/s41467-023-41228-3

3. https://www.ncbi.nlm.nih.gov/geo/