“Common genes findings and key pathways detections of shared genes between Severe Acute Respiratory Syndrome Coronavirus 2 and Cystic Fibrosis: A bioinformatic approach.”

A Thesis

Submitted for the completion for the Bachelor Degree in the Department of Software Engineering, Daffodil International University

By

Md. Tanvir Hasan

Academic ID: 171-35-209

Daffodil International University

Supervision of

Mr. Bikash Kumar Paul

Senior Lecturer

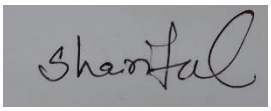
Department of Software Engineering

Daffodil International University

Approval

Thesis titled as "Common genes findings and key pathways detections of shared genes between Severe Acute Respiratory Syndrome Coronavirus 2 and Cystic Fibrosis: A bioinformatic approach", by Md. Tanvir Hasan. This thesis has been acknowledged as meeting the requirements for graduation with a B.Sc in Software Engineering and has been approved in terms of style and substance.

Board of Examiner



……..

Md. Shariful Islam

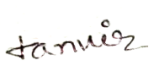
Declaration

I, Md. Tanvir Hasan, hereby swear that the material provided below is my original work, and that it's not been published or submitted somewhere for a degree program prerequisite. Any references to literature or work done by others that are mentioned in this thesis have been acknowledged and listed in the reference section.

Md. Tanvir Hasan

Daffodil International University

Submitted by



……...

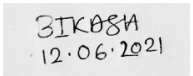
Md. Tanvir Hasan

Id: 171-35-209

Department of Software Engineering

Daffodil International University

Supervised by



….……

Bikash Kumar Paul

Senior Lecturer

Department of Software Engineering

Daffodil International University

Abstract

A comparative transcriptomic analysis has performed to establish essential profiles of the gene expression of SARS-CoV-2 linked to Cystic Fibrosis (CF). Transcriptomic studies have been carried out in relation to SARS-CoV-2 since a number of people have been diagnosed with CF. The recognition of differentially expressed genes demonstrated 8 concordant genes shared between the SARS-CoV-2 and CF. Extensive gene ontology analysis and the discovery of pathways enrichment demonstrated SARS-CoV-2 response to the CF. The gene ontological terms and pathways enrichment mechanisms derived from this research may affect the production of successful drugs, especially for the people with the following disorder. Identification of TF-miRNA association network reveals the interconnection between TF genes and miRNAs, which may effective to reveals the other influenced disease that occurs for SARS-CoV-2 to CF. The enrichment of pathways reveals SARS-CoV-2 associate CF mostly engaged with the type of Innate Immune System, Toll-like receptor signaling pathway, Pantothenate and CoA biosynthesis, Allograft rejection, Graft-versus-host disease, Intestinal immune network for IgA production, Mineral absorption, Autoimmune thyroid disease, Legionellosis, Viral myocarditis, Inflammatory bowel disease (IBD), etc. The drug compounds identification demonstrates the drug targets of IMIQUIMOD and raloxifene are the most significant with the significant hub DEGs.

Acknowledgement

I'd like to thank Dr. Md. Imran Mahmud, Head of the Department of Software Engineering, for enabling me to participate in this research.

I’m also grateful to my Associated Professor and Associate Head S A M Matiur Rahman, for continuous support and care.

I am grateful to my supervisor, Bikash Kumar Paul, Department of Software Engineering, for his continual advice and invertible suggestion during the research.

I am also glad to my teacher Md. Shariful Islam, Department of Software Engineering, for his special guidelines and cares about my work.

List of Abbreviation

SARS-CoV-2: Severe Acute Respiratory Syndrome Coronavirus 2

CF: Cystic Fibrosis

COVID-19: Coronavirus Disease 2019

ECFS: European Cystic Fibrosis Society

NCBI: National Center for Biotechnology Information

GO: Gene Ontology

BP: Biological Process

CC: Cellular Component

MF: Molecular Function

KEGG: Kyoto Encyclopedia Genes and Genome

PPI: Protein-protein Interaction Network

DEG: Differential Expressed Gene

MCC: Maximal Clique Centrality

TF: Transcription Factor

List of Table

List of Figure

indicates connections between the nodes. Yellow colored 4 nodes including TLR2, PROS1, CD86, and TRPM6 are seed nodes.

Figure 7. The hub genes identified from the protein interaction network. Highlighted octagon nodes considered as hub genes including TLR2, PROS1, CD86, and TRPM6.

Figure 8. The TF-miRNA network build up with TF genes and miRNA. The network conducts 25 TF genes and 80 miRNA. Blue colored rectangle nodes indicates miRNA, diamond shape green colored nodes indicates TF genes, and red colored octagon nodes indicates seed nodes. TF genes and miRNA interconnected with 6 seed nodes.

Table of Content

Chapter 1: Introduction

1.1. Background

Coronavirus disease 19 (COVID-19) is a result of a novel coronavirus known as the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1]. Coronavirus are a wide range of viruses that cause disease, from the usual cold to serious conditions such as Middle East Respiratory Syndrome (MERS) and SARS. In Wuhan, China, this virus was first found in people exposed to a seafood or wet market [2]. Chinese public health, healthcare communities and scientists have been rapidly responding, helping us to detect clinical disease and understand the infection epidemiology. First news suggested that transmission between people has been restricted or not existing, but we know that this is happening, even though it remains unclear to what degree [3, 4]. As of 25/03/2021 around 125,608,151 cases of coronavirus disease 2019 (COVID-19) and 2,759,429 deaths have been reported [5]. The virus is initially spread between peoples during close contact, most often via small droplets produced by coughing, sneezing, and talking [6]. People of all ages could be affected by COVID-19 [7, 8]. The contaminated droplets will scatter 1-2 m over the surfaces and deposit. The virus can remain viable on surfaces for several days, but common disinfectants such as sodium hypochlorite, hydrogen peroxide etc., are killed in less than one minute [9]. The clinical characteristics of COVID-19 vary from asymptomatic to acute respiratory distress and multi organ disorder. Popular health characteristics include fever, cough, sore throat, anxiety, tiredness, headache, myalgia, and respiratory failure. Conjoins are also identified. They are also inseparable from other breathing pathogens [10]. Most of peoples who died with COVID-19 have diseases that still exist, including high blood pressure, diabetes mellitus and cardiovascular disease [11]. The most common comorbidities are hypertension (66% of deaths), type 2 diabetes (29.8% of deaths), Ischemic Heart Disease (27.6% of deaths), atrial fibrillation (23.1% of deaths) and chronic renal failure (20.2% of deaths) [12]. According to the Centers for Disease Control and prevention (CDC), the most critical respiratory comorbidities are moderate or severe asthma, pre-existing COPD, pulmonary fibrosis, cystic fibrosis. In 13 June 2020, Elliot McClenaghan, Rebecca Cosgriff, et al. reported 181 people with CF (32 post-transplant) from 19 countries diagnosed with SARS-CoV-2. SARS-CoV-2 infections tend to have a similar range of effects as in the general population, 11 patients are admitted to intensive care with 7 deaths [13]. On the other hand in February 2021, Sara Manti, Giuseppe Fabio Parisi, et al. submitted a report on the patients with CF diagnosed with SARS-CoV-2. This study was based on 58 patients with CF where 12 patients diagnosed with SARS-CoV-2 [14]. Considering all these aspects, this study aimed to explore the common core mechanisms including key genes, common pathways, drug target, disease pathways, etc.

In this study, have established a comprehensive and quantitative network-based approach to explore the gene expression influence of the SARS-CoV-2 and CF and how these effects

could be representative of how they promote other disorder occurrence and development across pathways and pathway genes, which are also impaired in these disorders. Thus, equating the impact of SARS-CoV-2 exposure on gene expression with the changed pattern of gene expression observed in CF. The first step was to examine various gene profiles and then filtering these genes with networks of gene disorder connections, signaling, and ontological pathways, and networks of interactions between protein-protein, TF-gene interactions. All the processes shortly demonstrated in Figure 1.

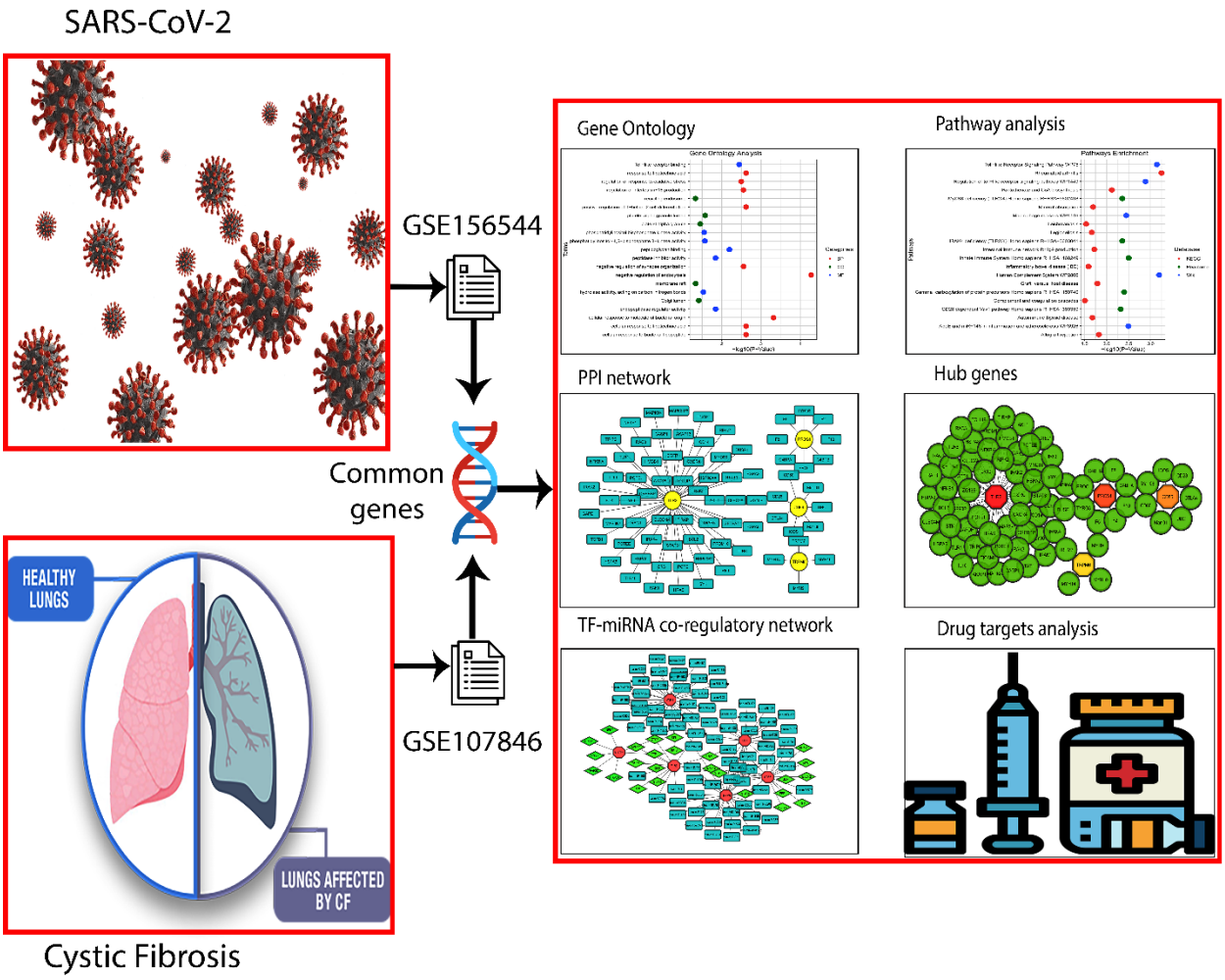


Figure 1. Image workflow of this study methods. A transcriptomic comparative analysis has performed between SARS-CoV-2 and CF. Two datasets have been collected for SARS-CoV-2 and CF and differentially expressed genes (DEGs) were identified using GEO2R and both datasets filtered to normalize datasets and identify shared DEGs. Using shared datasets have performed a transcriptomic analysis.

1.2. Objectives and limitations

The broad objective was to determine the bio-marker for the Severe Acute Respiratory Syndrome Coronavirus 2.

The specific objective of this study:

- Determine the common genes between the SARS-CoV-2 and CF
- Construct the protein interaction network
- Explore the gene ontological data
- Identification of pathways enrichment
- TF-miRNA co-regulatory network construction
- Hub genes identification
- Drug target analysis

Limitations of this study

- Poor number of datasets
- Very few number of samples

1.3. Motivation

I. In the future study may help to develop a biomarker for the SARS-CoV-2 associated CF.

II. This study explores the common disease connectivity of SARS-CoV-2 and CF.

III. Also, this research exposes the core common gene ontological data between SARS-CoV-2 and CF.

1.4. Problem Statement

Nowadays Coronavirus disease is the biggest headache in the world. Total of 3,557,413 people's death as of 31.05.2021. On the other hand, people with Cystic Fibrosis gradually affected by Coronavirus disease, reported by UCFS. But till now there are no bio-marker identified that will cure the patients of SARS-CoV-2 linked Cystic Fibrosis.

1.5. Thesis Organization

The entire paper has been divided into some chapters. Where the section number 2 introduced related work of this research that has already exist. In the section 3, marked the whole procedure to completion this study. Section number 4 indicates what actually we have got in this study. And the last section 5 mentioned the brief of this study.

Chapter II: Literature Review

There are huge of study have published, where the authors worked on statistical analysis or worked on a single of disease. A study, named "Systemic review: Cystic Fibrosis in the SARS-CoV-2/Covid-19 pandemic" Methew, H.R et al. have been introduced this paper, where they are demonstrated the pandemic situation for the CF associated SARS-CoV-2. This kind of huge work has been published but there are no study introduced where worked on the core common mechanism between the SARS-CoV-2 and CF.

Chapter III: Methodology

3.1. Statistical significance of patients with CF diagnosed with SARS-CoV-2

European Cystic Fibrosis Society (ECFS) started a project named COVID-CF project in Europe (https://www.ecfs.eu/covid-cf-project-europe). They have collected data throughout Europe about people with CF who become infected with SARS-Cov-2, causing the illness COVID-19. ECFS invited 38 countries to contribute the information and 38 countries provide the information. As of 8 March 2021 27 countries reported 1126 known case and 11 countries reported zero known case. Most of the patients are in the group of age 18-29 (303) and 30-49 (267) years [Figure 2].

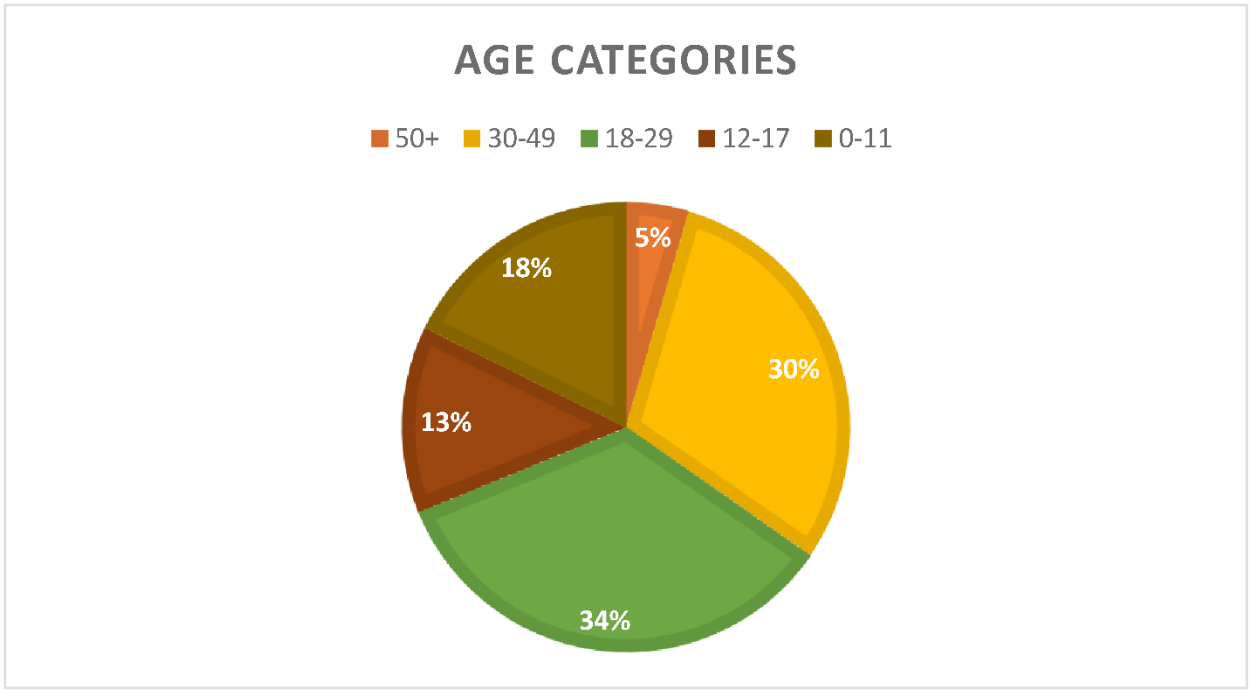


Figure 2. Pie chart of age categories in people with CF affected COVID-19. People with ages of 18 to 29 category is the most significantly affected by COVID-19 with CF.

3.2. Dataset consideration

We considered the gene expression microarray data sets from the National Center for Biotechnology Information (NCBI) in order to define a gene expression dysregulation common to SARS-CoV-2 and CF is examined. For this finding, considered two differentially expressed datasets with accession numbers GSE156544 and GSE107846 [15, 16]. GSE156544 dataset considered as Epithelial response to IFN-γ promotes SARS-CoV-2 infection. The dataset stands on GPL21272 microarray platform (Agilent-048908 8x60K whole genome incl V1-V2 linc BioVacSafe final 048908) and consist total of 8 samples: 6 normal samples and 2 infected samples. GSE107846 considered as Secondhand smoke alters arachidonic acid metabolism in infants and children with cystic fibrosis. GSE107846 also stands on GPL13369 microarray platform (Illumina Human Whole-Genome DASL HT) and consist total of 40 samples: 12 healthy samples and 28 cystic fibrosis infected samples.

3.3. Datasets filtration and retrieve DEGs to identify common DEGs

Microarray datasets GSE156544 and GSE147507 both of considered for Epithelial response to IFN-γ promotes SARS-CoV-2 infection and Secondhand smoke alters arachidonic acid metabolism in infants and children with cystic fibrosis respectively for this study. For both of datasets, GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/) assist to identify the expected DEGs. GEO2R is a web-based microarray data analysis tool. Comparing contaminated samples against managed samples is carried out in a comparatively manner and comparison is carried out using the limma and GEOquery [17] packages from Bioconductor [18] project in the R programming language framework. To determine significant DEGs as cutoff criteria 0.75 log2 fold change and P-value<0.05 has been considered for both of datasets. A comparison method named map-reduce algorithm has been build using Hadoop Distributed File System (HDFS) to determine the common DEGs between SARS-CoV-2 and CF. The software library Apache Hadoop is a big data framework platform that enables the distributed processing of large data sets through computer groups using basic programming models. It is planned to reach thousands of machines from single servers and provide local calculation and storage.

3.4. Gene ontology enrichment and gene associated pathways analysis

Gene set enrichment analysis is typically a quantitative and mathematical approach that determines whether a set of determined genes show statistical importance in various biological conditions [19]. Gene Ontology (GO) is a bioinformatics resource that provides information about the role of a gene product through ontology representing biological knowledge [20]. Gene ontology (GO) comprises three hierarchies describing the gene function based on biological process (BP), cellular component (CC), and molecular function [21]. For this study GO terms has been extracted from the Enrichr (https://maayanlab.cloud/Enrichr/) web based platform which is publicly open repository and contains in total, 180184 annotated gene sets from 102 gene set libraries [22]. Kyoto Encyclopedia Genes and Genome (KEGG) [23], Reactome [24], and WikiPathways [25] are used for cellular descriptive research pathways. This analysis also extracted from the Enrichr database.

3.5. Protein interaction network construction and analysis

Protein interaction network is the graphical representation of the connections between proteins where nodes and edges represents proteins and connections respectively [26]. The backbone of the transduction pathways and networks in a number of physiological processes are protein-protein interactions (PPIs). As the 'unwieldy' PPIs are now a possible new class of medicinal drugs, their essential functions in the relaying of cell growth signals in both normal and cancer cells are very significant [27]. Overlapped DEGs were applied to construct protein interaction network and extract protein-protein interaction network (PPIN) from NetworkAnalyst tool (https://www.networkanalyst.ca/) using STRING database [28, 29]. NetworkAnalyst is a robust web-based platform that enables bench researchers to conduct popular and nuanced meta-analyses of gene expression data using an intuitive web interface [30]. The extracted PPIN was visualized, analyzed, and re-designed using the Cytoscape software tool, which is one of the most common open-source software tools for visualizing biomedical networks composed of protein, gene, and other forms of interactions [31].

3.6. Cytoscape plugin algorithms analysis

Hub genes or key genes are defined as genes with high correlation in the large scale of PPIs network [32]. CytoHubba plugin (http://apps.cytoscape.org/apps/cytohubba) tool used to detect highly connected protein node from the PPI network. CytoHubba provides 11 topological analysis methods including Degree, Edge Percolated Component, Maximum Neighborhood Component, Density of Maximum Neighborhood Component, Maximal Clique Centrality (MCC) and six centralities (Bottleneck, EcCentricity, Closeness, Radiality, Betweenness, and Stress) based on shortest paths. Among them, MCC is a new proposed method, who has a better performance on the precision of predicting essential proteins from the yeast PPI network [33]. Maximal Clique Centrality (MCC) algorithm was applied to determine the hub genes for this study [34]. The PEWCC (http://apps.cytoscape.org/apps/pewcc) algorithm mete the reliability of the interaction data, then predicts protein complexes based on the concept of weighted clustering coefficient which is used to determine the complex network or cluster network [35].

3.7. miRNA-TF co-regulatory network development and analysis

The miRNA-Transcription factor (TF) co-regulatory networks was constructed using RegNetwork repository (http://www.regnetworkweb.org/), which provided new insights on the role of the proposed miRNA-TF co-regulation in the affected cell of SARS-CoV-2 diagnosed with CF and its cross talk with the surrounding microenvironment [36]. RegNetwork is a repository of human and mouse gene regulatory networks that captures and integrates known regulatory interactions between transcription factors (TFs), microRNAs (miRNAs) and target genes in 25 selected databases. RegNetwork comprises a detailed collection of transcriptional and post-transcriptional

regulatory interactions that are experimentally observed or predicted, and the database structure is flexible for possible expansions into gene regulatory networks for other species [37].

3.8. Drug candidate identification

Drug candidate has been extracted from the drug signature database (DSigDB) for the significant genes [38]. DSigDB is publicly open source online repository which is holds 22527 gene sets, consists of 17 389 unique compounds covering 19531 genes at present. DSigDB is a database of drug-related and small molecular gene samples focused on improvements in quantitative gene expression and / or drug-induced data. In the following ways, DSigDB varies from current tools: 1. DSigDB genes have been extracted from a range of databases and journals and collected from quantitative drug/compound inhibition evidence, 2. The DSigDB gene sets are obtained from both automated and manual computing processes, 3. DSigDB gene sets was specifically configured to provide the GSEA program with smooth integration, 4. DSigDB features the highest possible number of gene sets to date for drugs/compounds [39].

Chapter IV: Result and Discussion

4.1. Overlapped DEGs identification between SARS-CoV-2 and CF

For the dataset of SARS-CoV-2 a total of 322 DEGs were found after filtration, from where 167 DEGs showed up-regulation and the remaining 155 DEGs showed down-regulation. On the other hand the dataset of CF a total of 1537 filtered DEGs including 663 up-regulated and 874 down-regulated DEGs [Figure 3]. Mapreduce algorithm has been implemented using HDFS for two datasets of SARS-CoV-2 and CF to identify the overlapped DEGs for further analysis. Total of 8 overlapped DEGs showed in HDFS including ROPN1L, VNN2, TLR2, MCTP1, PROS1, LBH, TRPM6, and CD86.

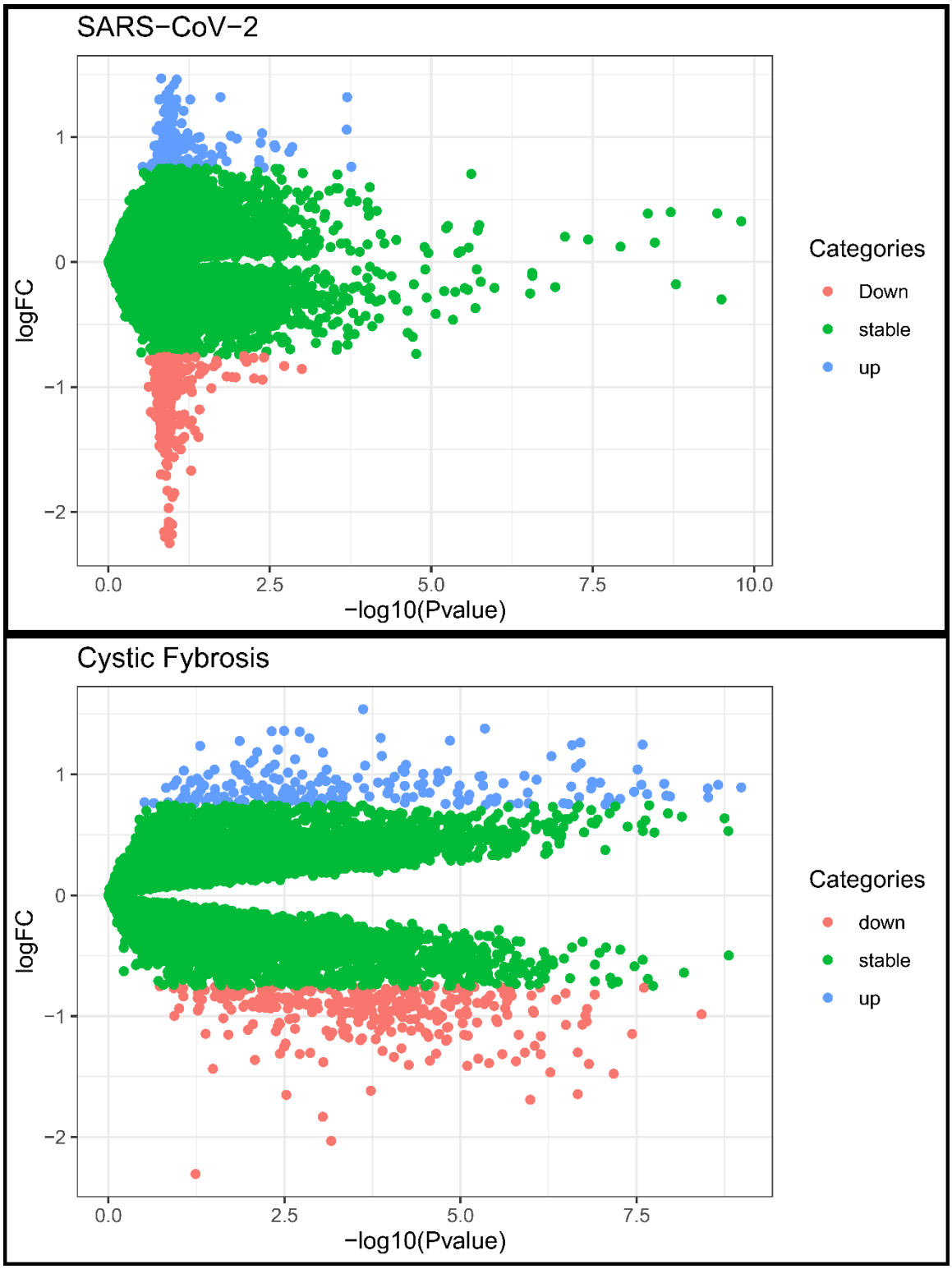


Figure 3. Skater plot shows the regulation of DEGs for SARS-CoV-2 and CF. Light blue colored nodes indicates the up regulated DEGs, pink colored nodes indicates the down regulated DEGs, Green colored nodes indicates stable DEGs.

4.2. Gene ontology analysis and gene associated pathways enrichment

GO enrichment were analyzed using identified overlapped DEGs between SARS-CoV-2 and CF. The results of GO enrichment provides biological process highly associated with negative regulation of endocytosis, cellular response to molecule of bacterial origin, cellular response to lipoteichoic acid, positive regulation of T-helper 2 cell differentiation, cellular response to bacterial lipopeptide, response to lipoteichoic acid, negative regulation of synapse organization, etc. Toll-like receptor binding, peptidoglycan binding, peptidase inhibitor activity, endopeptidase regulator activity, phosphatidylinositol-4,5-bisphosphate 3-kinase activity, etc. associated with molecular function. Cellular component mostly engaged with platelet alpha granule lumen, platelet alpha granule, Golgi lumen, membrane raft, and recycling endosome [Figure 4, Table 1]. On the other hand pathways enrichment analysis of KEGG, Reactome, Wiki pathways has been extracted through Enrichr open source database [41]. Extracted results showed Rheumatoid arthritis, Toll-like receptor signaling pathway, Pantothenate and CoA biosynthesis, Allograft rejection, Graft-versus-host disease, Intestinal immune network for IgA production, Mineral absorption, Autoimmune thyroid disease, Legionellosis, Viral myocarditis, Inflammatory bowel disease (IBD), Innate Immune System Homo sapiens R-HSA-168249, Gamma-carboxylation of protein precursors Homo sapiens R-HSA-159740, Gamma carboxylation, transport, and amino-terminal cleavage of proteins Homo sapiens R-HSA-159854, MyD88 deficiency (TLR2/4) Homo sapiens R-HSA-5602498, Human Complement System WP2806, Toll-like Receptor Signaling Pathway WP75, Regulation of toll-like receptor signaling pathway WP1449, ApoE and miR-146 in inflammation and atherosclerosis WP3926, and Macrophage markers WP4146, etc. pathways of KEGG, Reactome, and Wiki pathways are highly connected with the common DEGs [Figure 5, Table 2].

Table 1. Gene ontological terms analysis for shared genes between SARS-CoV-2 and CF with their p-value and adjusted p-value.

Categories	**Term**	**GO ID**	**P-value**	**Adjusted P-value**
BP	negative regulation of endocytosis	(GO:0045806)	5.26E-05	0.007792
	cellular response to molecule of bacterial origin	(GO:0071219)	4.80E-04	0.031636
	cellular response to lipoteichoic acid	(GO:0071223)	0.002398	0.031636
	positive regulation of T-helper 2 cell differentiation	(GO:0045630)	0.002398	0.031636
	cellular response to bacterial lipopeptide	(GO:0071221)	0.002398	0.031636
	response to lipoteichoic acid	(GO:0070391)	0.002398	0.031636

	negative regulation of synapse organization	(GO:1905809)	0.002797	0.031636
	regulation of interleukin-18 production	(GO:0032661)	0.002797	0.031636
	regulation of response to oxidative stress	(GO:1902882)	0.003196	0.031636
	regulation of T-helper 2 cell differentiation	(GO:0045628)	0.003196	0.031636
	positive regulation of interleukin-2 biosynthetic process	(GO:0045086)	0.003196	0.031636
CC	platelet alpha granule lumen	(GO:0031093)	0.026492	0.083931
	platelet alpha granule	(GO:0031091)	0.035444	0.083931
	Golgi lumen	(GO:0005796)	0.038541	0.083931
	membrane raft	(GO:0045121)	0.046628	0.083931
	recycling endosome	(GO:0055037)	0.046628	0.083931
MF	Toll-like receptor binding	(GO:0035325)	0.003595	0.047874
	peptidoglycan binding	(GO:0042834)	0.006383	0.047874
	peptidase inhibitor activity	(GO:0030414)	0.014312	0.050655
	endopeptidase regulator activity	(GO:0061135)	0.014312	0.050655
	phosphatidylinositol-4,5-bisphosphate 3-kinase activity	(GO:0046934)	0.026883	0.050655
	phosphatidylinositol bisphosphate kinase activity	(GO:0052813)	0.028054	0.050655
	hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds, in linear amides	(GO:0016811)	0.029614	0.050655

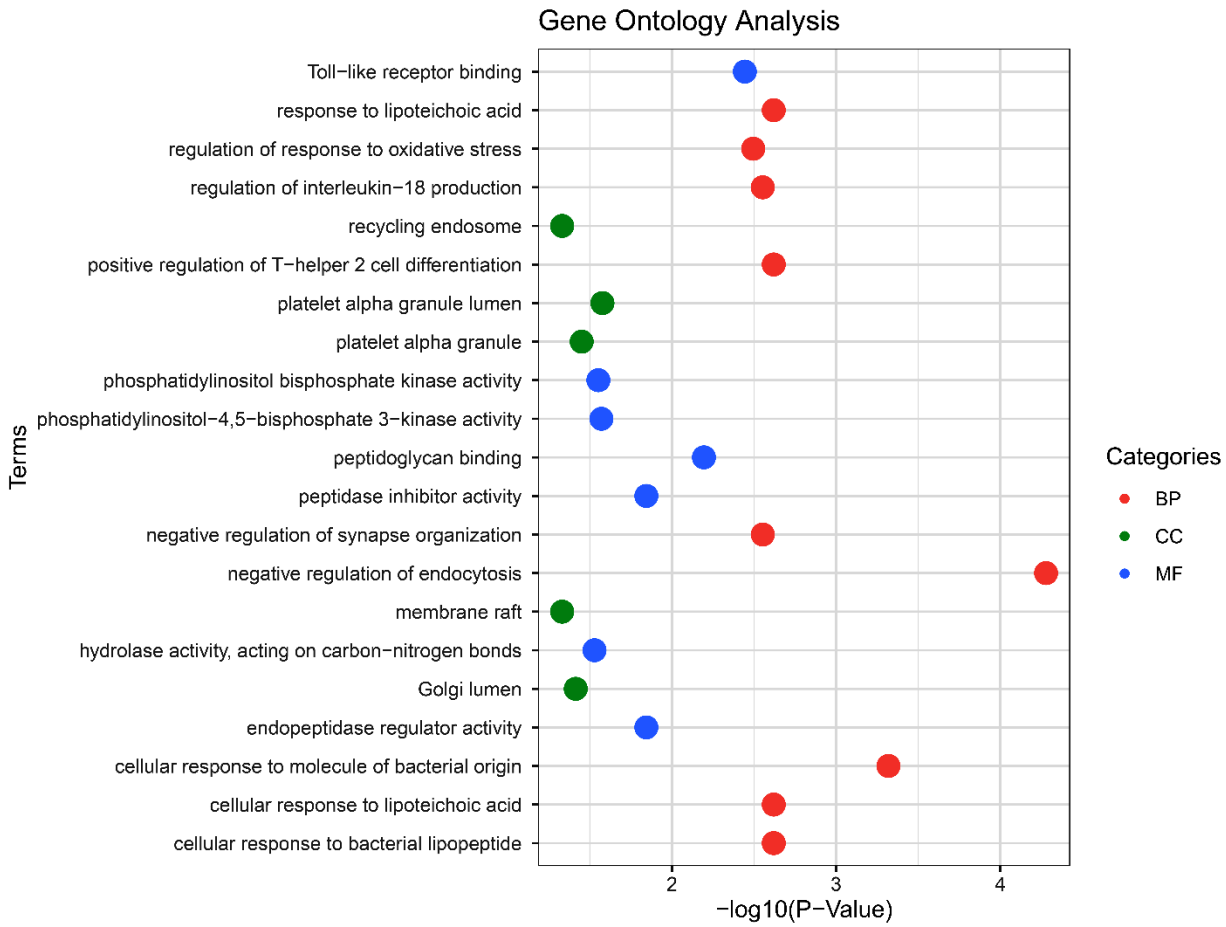


Figure 4. Gene ontological analysis for biological process (BP), cellular component (CC), molecular function (MF). Red colored terms indicates BP, green colored terms indicates CC, and blue colored terms indicates MF.

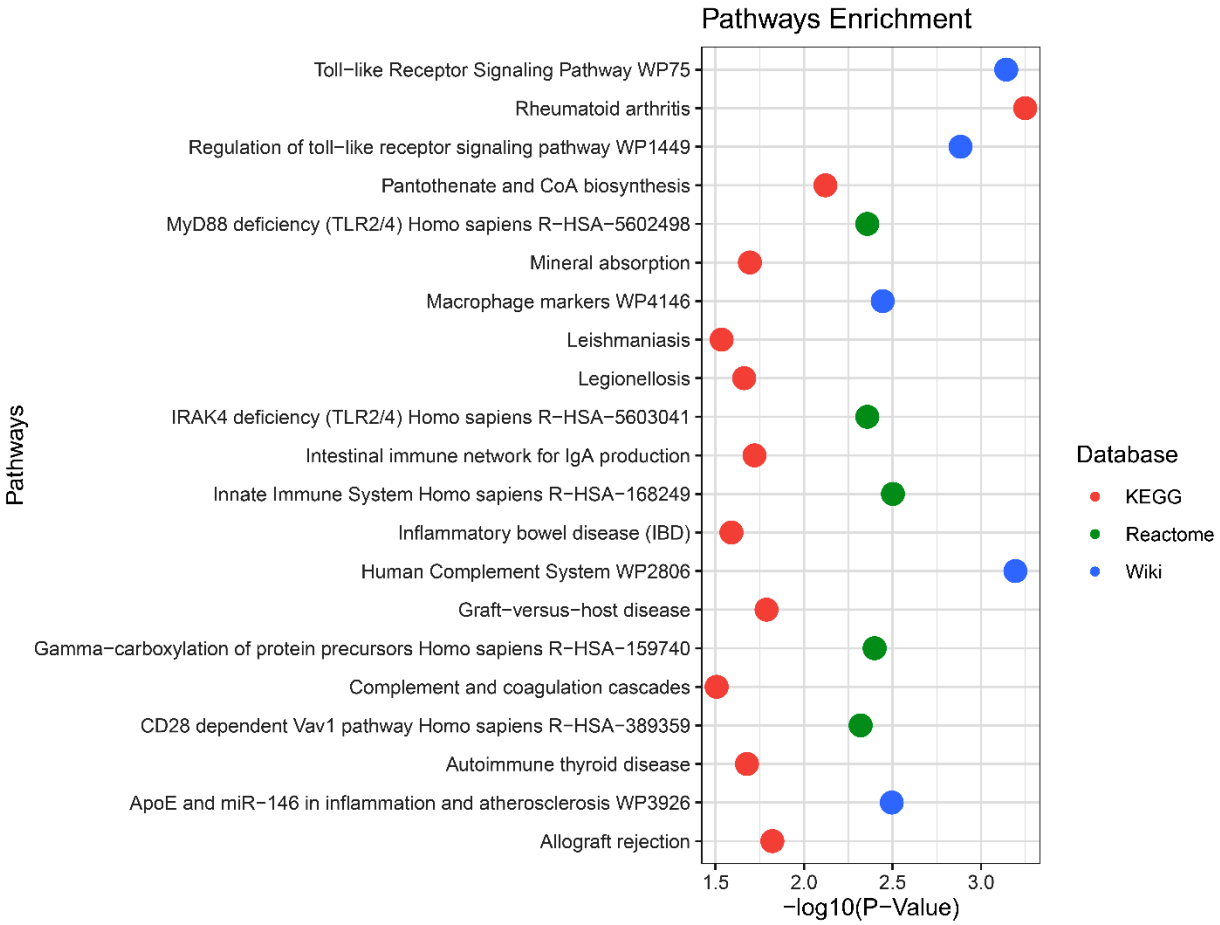


Figure 5. Pathways enrichment analysis through three databases. Red colored pathways for KEGG, green colored for Reactome, and blue colored for WikiPathways. In graph X axis for –log10(P-Value) and Y axis for pathways.

Table 2. Pathways enrichment analysis of shared genes between SARS-CoV-2 and CF through three databases including KEGG, Reactome, and WikiPathways.

Database	**Pathways**	**P-value**	**Adjusted P-value**	**Combined Score**
KEGG	Rheumatoid arthritis	5.63E-04	0.011388	557.7232
	Toll-like receptor signaling pathway	7.35E-04	0.011388	469.0428
	Pantothenate and CoA biosynthesis	0.007576	0.060351	774.035
	Allograft rejection	0.015102	0.060351	323.0508
	Graft-versus-host disease	0.016286	0.060351	293.3998

	Intestinal immune network for IgA production	0.019043	0.060351	240.1326
	Mineral absorption	0.020222	0.060351	222.2663
	Autoimmune thyroid disease	0.021008	0.060351	211.6082
	Legionellosis	0.021793	0.060351	201.815
	Viral myocarditis	0.023362	0.060351	184.4464
	Inflammatory bowel disease (IBD)	0.02571	0.06131	162.8428
	Leishmaniasis	0.029224	0.064422	137.7082
	Complement and coagulation cascades	0.031172	0.064422	126.4956
Reactome	Innate Immune System Homo sapiens R-HSA-168249	0.003147	0.046709	82.49672
	Transport of gamma-carboxylated protein precursors from the endoplasmic reticulum to the Golgi apparatus Homo sapiens R-HSA-159763	0.003595	0.046709	2008.476
	Removal of aminoterminal propeptides from gamma-carboxylated proteins Homo sapiens R-HSA-159782	0.003994	0.046709	1751.859
	Gamma-carboxylation of protein precursors Homo sapiens R-HSA-159740	0.003994	0.046709	1751.859
	Gamma-carboxylation, transport, and amino-terminal cleavage of proteins Homo sapiens R-HSA-159854	0.004392	0.046709	1549.437
	MyD88 deficiency (TLR2/4) Homo sapiens R-HSA-5602498	0.004392	0.046709	1549.437
	IRAK4 deficiency (TLR2/4) Homo sapiens R-HSA-5603041	0.004392	0.046709	1549.437
	CD28 dependent Vav1 pathway Homo sapiens R-HSA-389359	0.004791	0.046709	1385.975
Wiki	Human Complement System WP2806	6.40E-04	0.005766	513.4615
	Toll-like Receptor Signaling Pathway WP75	7.21E-04	0.005766	474.9722
	Regulation of toll-like receptor signaling pathway WP1449	0.001306	0.006968	320.7932
	ApoE and miR-146 in inflammation and atherosclerosis WP3926	0.003196	0.011504	2343.484
	Macrophage markers WP4146	0.003595	0.011504	2008.476

4.3. Protein interaction network construction and analysis

The NetworkAnalyst database used to construct the PPI network through the overlapped DEGs including ROPN1L, VNN2, TLR2, MCTP1, PROS1, LBH, TRPM6, and CD86. The extracted PPI network build with the confidence score 0.90 and network represents connectivity of TLR2, PROS1, CD86, and TRPM6 to other nodes [Figure 6]. The network build up with 80 nodes and 76 edges where TLR2, PROS1, CD86, and TRPM6 is the seed nodes, from them TLR2 is the highly connected seed node who consist 57 nodes. The PPI network is the base of the forward analysis including hub node identification, module analysis, drug compounds prediction, etc.

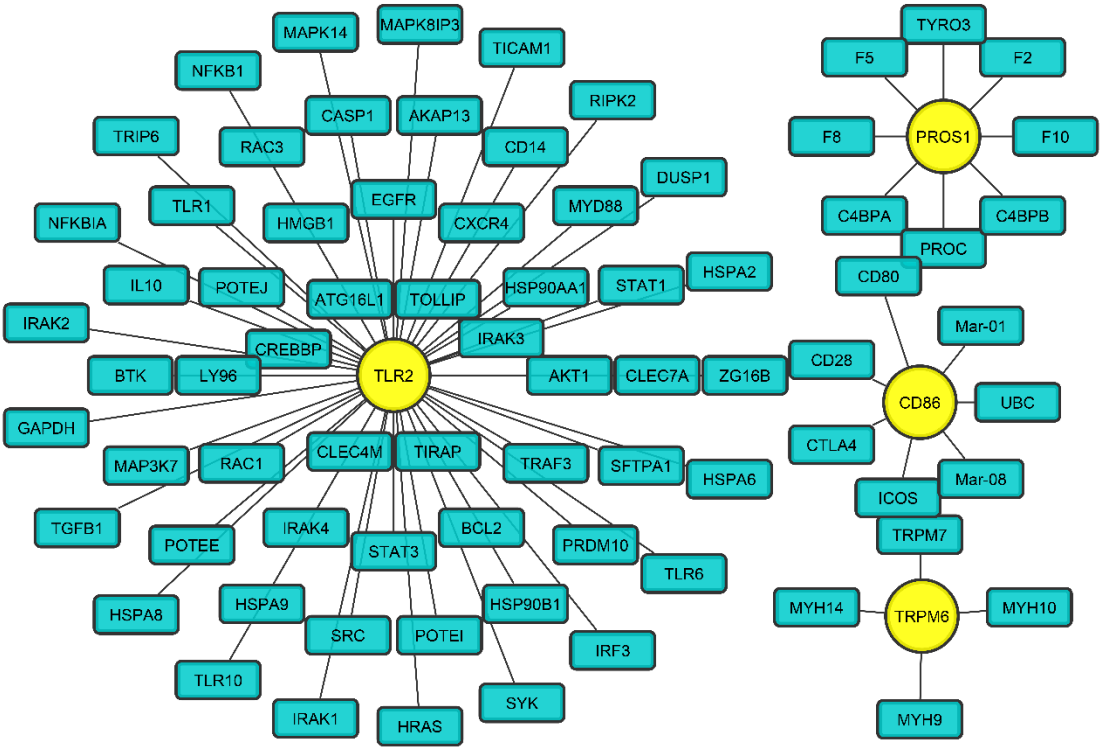


Figure 6. Protein interaction for the common differentially expressed genes between SARS-CoV-2 and CF. The network conducts 80 nodes and 76 edges where nodes indicates genes and edges indicates connections between the nodes. Yellow colored 4 nodes including TLR2, PROS1, CD86, and TRPM6 are seed nodes.

4.4. Cytoscape plugins algorithm analysis

The cytoHubba plugin algorithms used to predict the key nodes with the MCC method of eleven methods provide by cytoHubba. From the result of cytoHubba four nodes has taken as hub node including TLR2, PROS1, CD86, and TRPM6 [Figure 7]. Toll Like Receptor 2 known as TLR2 [42], who consist highest score on the Maximal Clique Centrality (MCC) algorithm. The PPI network were visualized to extract the protein complex network using MCODE algorithm. There was no protein complex network showed by MCODE algorithm.

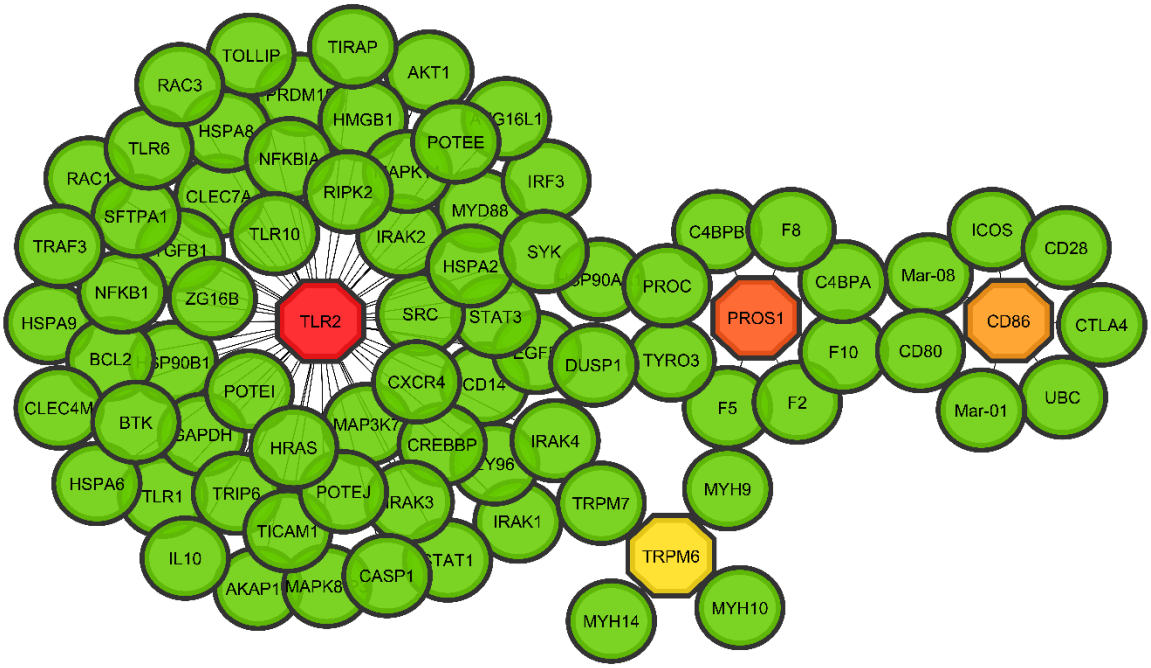


Figure 7. The hub genes identified from the protein interaction network. Highlighted octagon nodes considered as hub genes including TLR2, PROS1, CD86, and TRPM6.

4.5. Hub nodes related data

TLR2 is one of the most difficult receptors in the immune system and plays a significant role in immune system. TLR2 is a membrane protein, a receptor that is expressed on the surface of certain cells and identified by external substances and transmits adequate signals to the immune system cells [43]. TLR2 identifies a large number of bacterial, fungal, fungus, and certain endogenous compounds as a membrane surface receptor. A research found that the TLR2/6 agonist INNA-051's prophylactic intra-nasal administration in a ferret infection model SARS-CoV-2 effectively

lowers viral RNA levels in the nose and neck. The findings from their research help the clinical development of prophylactic TLR2/6-based therapy for innate URT immune activating, reducing transmission of SARS-CoV-2 and providing defense against COVID-19 [44]. The protein S is (known as S-Protein and Spike Protein) a vitamin K dependent plasma glycoprotein synthesized in the hepatic. Protein S appears in two different forms: a free shape and a complex shape bound by protein C4b-bound protein (C4BP). The PROS1 gene is coded in humans with protein S [45, 46]. The coronavirus spike protein S is involved in a variety of biological processes, including virus neutralization via antibody, cell attachment, and cell fusion. In a research, a S determinant typical in murine coronaviruses has elucidated the amino acid series [47]. Cluster Differentiation 86 is a protein that is constitutively expressed on dendritic cells, Langerhan cells, macrophages, B cells and on other antigenic cells. Cluster Differentiation 86 is also referred to as CD86 and B7-2 [48]. A study in 2020 have demonstrated of a distinct signaling event induced by CD80 and CD86 molecules in B cell lymphoma [49]. Another study showed insertion of host-derived costimulatory molecules CD80 and CD86 into human immunodeficiency virus type 1 affects the virus life cycle [50].

4.6. miRNA-TF co-regulatory network development and analysis

MicroRNAs (miRNAs) and transcription factors (TFs) are significant regulators of gene expression [51]. The miRNAs and TFs might play combinatory regulatory roles in SARS-CoV-2 and CF. After collecting SARS-CoV-2 and CF associated overlaped candidate genes, have constructed a comprehensive specific TF-miRNA co-regulatory network by integrating predicted and experimentally validated TF and miRNA targets. RegNetwork repository has been used to construct TF-miRNA co-regulatory network using overlapped genes of SARS-CoV-2 and CF. TF-miRNA co-regulatory network build up on 111 nodes, 134 edges, where 25 TF candidates, 6 seed nodes, and 80 miRNA candidates [Figure 8].

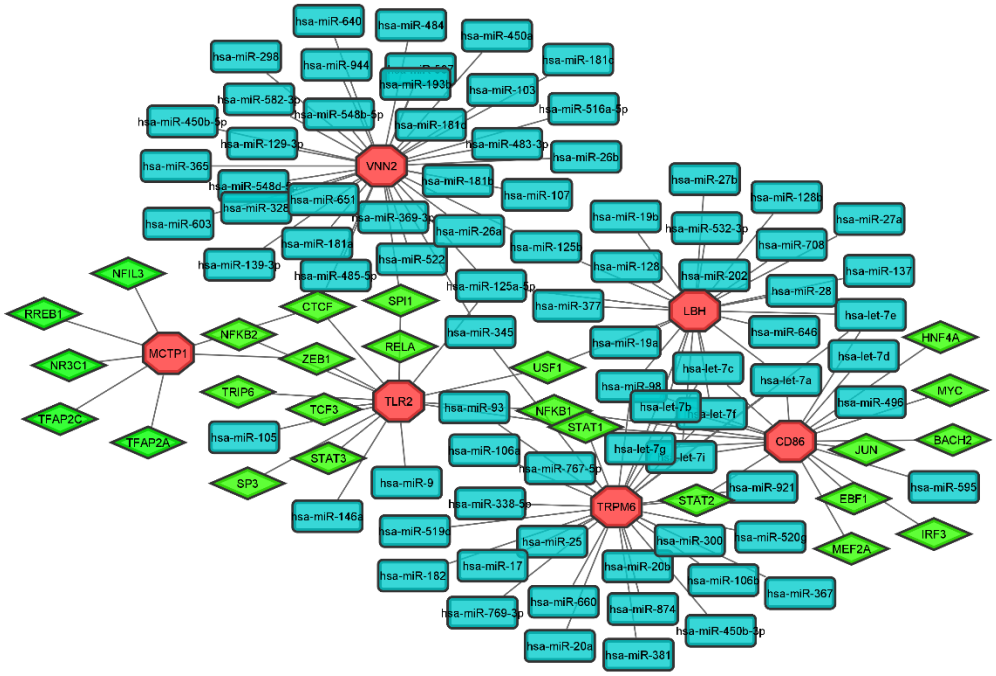


Figure 8. The TF-miRNA network build up with TF genes and miRNA. The network conducts 25 TF genes and 80 miRNA. Blue colored rectangle nodes indicates miRNA, diamond shape green colored nodes indicates TF genes, and red colored octagon nodes indicates seed nodes. TF genes and miRNA interconnected with 6 seed nodes.

4.7. Drug candidate identification

The predictive drug signature has been extracted from the DSigDB database using the overlapped DEGs of SARS-CoV-2 and CF including ROPN1L, VNN2, TLR2, MCTP1, PROS1, LBH, TRPM6, and CD86. The result has been taken based on adjusted P-Value and overlapped DEGs. IMIQUIMOD and raloxifene are mostly significant based on their adjusted P-Value and overlapped genes that shown in Table 3.

Table 3. Predicted drug target analysis using shared genes between SARS-CoV-2 and CF.

Drug targets	Adjsted P-Value	P-Value	Overlapped Genes
IMIQUIMOD BOSS	0.00708	5.65E-05	CD86;TLR2
POLY I-C CTD 00006579	0.00708	7.80E-05	CD86;TLR2
Ammonium hexachloroplatinate(IV) CTD 00000945	0.00708	1.03E-04	CD86;TLR2
Sodium dodecyl sulfate CTD 00006753	0.019408	3.53E-04	CD86;TLR2

hydrocortisone CTD 00006117	0.02336	6.40E-04	CD86;TLR2
chloroquine BOSS	0.02336	6.53E-04	CD86;TLR2
ACMC-20mvek CTD 00002629	0.02336	6.80E-04	CD86;TLR2
glutathione CTD 00006035	0.032208	0.001251	CD86;TLR2
diphenylpyraline BOSS	0.032208	0.0019	CD86;TLR2
raloxifene CTD 00007367	0.032208	0.001977	MCTP1;LBH;PROS1

4.8. Discussion

This research explored common connections between SARS-CoV-2 and CF and the role of SARS-CoV-2 on CF. As of 8 March 2021 according to the European Cystic Fibrosis Society total of 1126 cases reported CF diagnosed with SARS-CoV-2. This means there is significant connections between SARS-CoV-2 and CF. Considering all these perspectives this study aimed to demonstrate the association between SARS-CoV-2 and CF including GO analysis, pathways enrichment, PPI network construction, hub gene finding, TF-miRNA co-regulatory network, and drug candidate identification.

Two microarray datasets for SARS-CoV-2 and CF have been considered for this study. The selected dataset has been filtered to find common DEGs using HDFS of the HADOOP big data framework. Through common DEGs applied for the GO analysis and pathways enrichment. The GO analysis revealed the terms of negative regulation of endocytosis, cellular response to molecule of bacterial origin, cellular response to lipoteichoic acid, cellular response to bacterial lipopeptide, Toll-like receptor binding, peptidoglycan binding, peptidase inhibitor activity, endopeptidase regulator activity, phosphatidylinositol-4,5-bisphosphate 3-kinase activity, platelet alpha granule lumen, platelet alpha granule, Golgi lumen, membrane raft, and recycling endosome are significantly associated with the overlapped DEGs. Pathways enrichment has been collected from three databases (KEGG, Reactome, WikiPathways) through the Enrichr repository. The results of pathways enrichment showed Rheumatoid arthritis, Toll-like receptor signaling pathway, Pantothenate, CoA biosynthesis, Allograft rejection, Graft-versus-host disease, Intestinal immune network for IgA production, Mineral absorption, Autoimmune thyroid disease, Legionellosis, Viral myocarditis, Inflammatory bowel disease (IBD), Innate Immune System Homo sapiens R-HSA-168249, Gamma-carboxylation of protein precursors Homo sapiens R-HSA-159740, Gamma carboxylation, etc. are associated with the common DEGs.

The protein interaction network also constructed with the gene association of SARS-CoV-2 and CF for further analysis. The PPI network conducts 80 nodes and 76 edges, from where hub nodes identified including TLR2, PROS1, CD86, and TRPM6. The regulatory networks could play a significant role as a potential biomarker to the various complex types of disease. Common DEGs have been performed to clarify the TF-miRNA coregulatory network which is the interaction of TF genes and miRNA. The regulatory network build-up with 25 TF genes and 80 miRNA candidates.

Finally, using hub nodes applied for the drug signature. From the identified drug signature Muramyl Dipeptide CTD 00005307 and raloxifene CTD 00007367 is highly significant. The Imiquimod (IMQ) is a heterocyclic amine that does not have nucleoside and belongs to the 1H-imidazole-[4, 5-c] family [52-55]. In general, imiquimod acts indirectly as an immune reaction modification inducing immune reactions and the section of several cytokines which stimulate T cells in turn [56]. The innate immune system works by recognizing toxins in the body and activating a variety of cell types to destroy them. Imiquimod works in the innate immune system by binding it to cell surface receptors including peak receptors (TLRs). In the early stages of the disease where activation of innate immunity by a TLR-7 agonist is crucial, Imiquimod's function as an active agent of the disease indicates its ability to treat viral infections such as SARS-CoV-2 [57]. Raloxifene is a modulatory selection of estrogen receptors for the treatment of postmenopausal osteoporosis and cancer that was approved for treatment and prevention by the FDA in 1997. Recently raloxifene also demonstrated effectiveness in treating Ebola, influenza A, and hepatitis C virus viral infections and exhibited possible drug repurposing to combat SARS-CoV-2 infections [58]. The outcomes of this study have been validated from the gold benchmark databases OMIM and dbGaP. In the future, this research may help to expose the other disease complexity, pathways, and transcription factor for the SARS-CoV-2 associated CF.

Chapter V: Conclusion

In this study, the identified biologic regions, regulatory components were briefly addressed, which may speed up clinical activity against SARS-CoV-2 associated with CF. The strength of our study should be taken into consideration, as it is the greatest transcriptomic study on SARS-CoV-2 associated CF. Transcriptomic analysis of this research has produced the common ontological entity, pathways, disease connectivity, transferrin genes. In CF associated SARS-CoV-2 have identified the corresponding genes between SARS-CoV-2 and CF that generate more molecular findings and demonstrate the interaction of DEGs. In the future, this study may help to develop drug targets and treatment.

Reference

1. P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C.L. Huang, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature, 579 (2020), pp. 270-273, https://doi.org/10.1038/s41586-020-2012-7.

2. He, F, Deng, Y, Li, W. Coronavirus disease 2019: What we know? J Med Virol. 2020; 92: 719– 725. https://doi.org/10.1002/jmv.25766.

3. Jinzhong Li, Xiaobing Gong, Zhigang Wang, Renzhou Chen, Taoyuan Li, Dongyu Zeng, Minran Li. Clinical features of familial clustering in patients infected with 2019 novel coronavirus in Wuhan, China, Virus Research, Volume 286, 2020, 198043, ISSN 0168-1702, https://doi.org/10.1016/j.virusres.2020.198043.

4. Stanley Perlman, M.D., Ph.D. Another Decade, Another Coronavirus. February 20, 2020, N Engl J Med 2020; 382:760-762 https://doi.org/10.1056/NEJMe2001126.

5. COVID-19 Coronavirus Pandemic. Worldometers, https://www.worldometers.info/coronavirus/.

6. Vignesh P., Dr. Brundha M. P. (2021). Active Monitoring of Person Exposed to Patients with Confirmed COVID - 19. Annals of the Romanian Society for Cell Biology, 738 - 749. Retrieved from http://annalsofrscb.ro/index.php/journal/article/view/1544.

7. Fakhroo, A.D.; Al Thani, A.A.; Yassine, H.M. Markers Associated with COVID-19 Susceptibility, Resistance, and Severity. Viruses 2021, 13, 45. https://doi.org/10.3390/v13010045.

8. Rothe C, Schunk M, Sothmann P, et al. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. N Engl J Med. 2020. https://doi.org/10.1056/NEJMc2001468.

9. Kampf G, Todt D, Pfaender S, Steinmann E. Persistence of coronaviruses on inanimate surfaces and its inactivation with biocidal agents. J Hosp Infect. 2020 Feb 6. pii: S0195–6701(20)30046–3.

10. Singhal, T. A Review of Coronavirus Disease-2019 (COVID-19). Indian J Pediatr 87, 281–286 (2020). https://doi.org/10.1007/s12098-020-03263-6.

11. Garg S, Kim L, Whitaker M, O'Halloran A, Cummings C, Holstein R, et al. (April 2020). "Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 – COVID-NET, 14 States, 1–30 March 2020". MMWR. Morbidity and Mortality Weekly Report. 69 (15): 458–464. https://doi.org/10.15585/mmwr.mm6915e3. PMC 7755063. PMID 32298251.

12. Palmieri L, Andrianou X, Barbariol P, Bella A, Bellino S, Benelli E, et al. (22 July 2020). Characteristics of SARS-CoV-2 patients dying in Italy Report based on available data on July 22nd, 2020 (PDF) (Report). Istituto Superiore di Sanità. Retrieved 4 October 2020.

13. Elliot McClenaghan, Rebecca Cosgriff, Keith Brownlee, Susannah Ahern, Pierre-Régis Burgel, et al. The global impact of SARS-CoV-2 in 181 people with cystic fibrosis, Journal

of Cystic Fibrosis, Volume 19, Issue 6, 2020, Pages 868-871, ISSN 1569-1993, https://doi.org/10.1016/j.jcf.2020.10.003.

14. Sara Manti, Giuseppe Fabio Parisi, Maria Papale, Enza Mulè, Donatella Aloisio, Novella Rotolo, Salvatore Leonardi, Looking beyond pulmonary disease in COVID-19: A lesson from patients with cystic fibrosis, Medical Hypotheses, Volume 147, 2021, 110481, ISSN 0306-9877, https://doi.org/10.1016/j.mehy.2020.110481.

15. Epithelial response to IFN-γ promotes SARS-CoV-2 infection. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE156544 (Public on Jan 01, 2021).

16. Secondhand smoke alters arachidonic acid metabolism in infants and children with cystic fibrosis. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE107846 (Public on Dec 01, 2020).

17. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007 Jul 15;23(14):1846-7. doi: 10.1093/bioinformatics/btm254. Epub 2007 May 12. PMID: 17496320.

18. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004;5(10):R80. doi: 10.1186/gb-2004-5-10-r80. Epub 2004 Sep 15. PMID: 15461798; PMCID: PMC545600.

19. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P: a desktop application for Gene Set Enrichment Analysis. Bioinformatics. 2007 Dec 1;23(23):3251-3. doi: 10.1093/bioinformatics/btm369. Epub 2007 Jul 20. PMID: 17644558.

20. Gene Ontology Consortium. Gene Ontology Consortium: going forward. Nucleic Acids Res. 2015 Jan;43(Database issue):D1049-56. doi: 10.1093/nar/gku1179. Epub 2014 Nov 26. PMID: 25428369; PMCID: PMC4383973.

21. Dejongh M, Van Dort P, Ramsay B. Linking molecular function and biological process terms in the ontology for gene expression data analysis. Conf Proc IEEE Eng Med Biol Soc. 2004;2004:2984-6. doi: 10.1109/IEMBS.2004.1403846. PMID: 17270905.

22. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016

update. Nucleic Acids Res. 2016 Jul 8;44(W1):W90-7. doi: 10.1093/nar/gkw377. Epub 2016 May 3. PMID: 27141961; PMCID: PMC4987924.

23. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000 Jan 1;28(1):27-30. doi: 10.1093/nar/28.1.27. PMID: 10592173; PMCID: PMC102409.

24. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, Haw R, Jassal B, Korninger F, May B, Milacic M, Roca CD, Rothfels K, Sevilla C, Shamovsky V, Shorser S, Varusai T, Viteri G, Weiser J, Wu G, Stein L, Hermjakob H, D'Eustachio P. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 2018 Jan 4;46(D1):D649-D655. doi: 10.1093/nar/gkx1132. PMID: 29145629; PMCID: PMC5753187.

25. Slenter DN, Kutmon M, Hanspers K, Riutta A, Windsor J, Nunes N, Mélius J, Cirillo E, Coort SL, Digles D, Ehrhart F, Giesbertz P, Kalafati M, Martens M, Miller R, Nishida K, Rieswijk L, Waagmeester A, Eijssen LMT, Evelo CT, Pico AR, Willighagen EL. WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. Nucleic Acids Res. 2018 Jan 4;46(D1):D661-D667. doi: 10.1093/nar/gkx1064. PMID: 29136241; PMCID: PMC5753270.

26. Md. Tanvir Hasan, Mehedi Hassan, Kawsar Ahmed, Md. Rakibul Islam, Kabirul Islam, Touhid Bhuyian, Muhammad Shahin Uddin, Bikash Kumar Paul. Network based study to explore genetic linkage between diabetes mellitus and myocardial ischemia: Bioinformatics approach, Gene Reports, Volume 21, 2020, 100809, ISSN 2452-0144, https://doi.org/10.1016/j.genrep.2020.100809.

27. Li, Z., Ivanov, A., Su, R. et al. The OncoPPi network of cancer-focused protein–protein interactions to inform biological insights and therapeutic strategies. Nat Commun 8, 14356 (2017). https://doi.org/10.1038/ncomms14356.

28. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic Acids Res. 2019 Jul 2;47(W1):W234-W241. doi: 10.1093/nar/gkz240. PMID: 30931480; PMCID: PMC6602507.

29. Ye X, Zeng T, Kong W, Chen LL. Integrative Analyses of Genes Associated with Fulminant Type 1 Diabetes. J Immunol Res. 2020 Oct 6;2020:1025857. doi: 10.1155/2020/1025857. PMID: 33083497; PMCID: PMC7559223.

30. Xia J, Gill EE, Hancock RE. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. Nat Protoc. 2015 Jun;10(6):823-44. doi: 10.1038/nprot.2015.052. Epub 2015 May 7. PMID: 25950236.

31. Su G, Morris JH, Demchak B, Bader GD. Biological network exploration with Cytoscape 3. Curr Protoc Bioinformatics. 2014 Sep 8;47:8.13.1-24. doi: 10.1002/0471250953.bi0813s47. PMID: 25199793; PMCID: PMC4174321.

32. Chen P, Wang F, Feng J, Zhou R, Chang Y, Liu J, Zhao Q. Co-expression network analysis identified six hub genes in association with metastasis risk and prognosis in hepatocellular carcinoma. Oncotarget. 2017 Jul 25;8(30):48948-48958. doi: 10.18632/oncotarget.16896. PMID: 28430663; PMCID: PMC5564739.
33. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014;8 Suppl 4(Suppl 4):S11. doi: 10.1186/1752-0509-8-S4-S11. Epub 2014 Dec 8. PMID: 25521941; PMCID: PMC4290687.

34. Li C, Xu J. Feature selection with the Fisher score followed by the Maximal Clique Centrality algorithm can accurately identify the hub genes of hepatocellular carcinoma. Sci Rep. 2019 Nov 21;9(1):17283. doi: 10.1038/s41598-019-53471-0. PMID: 31754223; PMCID: PMC6872594.

35. Zaki N, Efimov D, Berengueres J. Protein complex detection using interaction reliability assessment and weighted clustering coefficient. BMC Bioinformatics. 2013 May 20;14:163. doi: 10.1186/1471-2105-14-163. PMID: 23688127; PMCID: PMC3680028.

36. Mohamed RH, Abu-Shahba N, Mahmoud M, Abdelfattah AMH, Zakaria W, ElHefnawi M. Co-regulatory Network of Oncosuppressor miRNAs and Transcription Factors for Pathology of Human Hepatic Cancer Stem Cells (HCSC). Sci Rep. 2019 Apr 3;9(1):5564. doi: 10.1038/s41598-019-41978-5. PMID: 30944375; PMCID: PMC6447552.

37. Liu ZP, Wu C, Miao H, Wu H. RegNetwork: an integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse. Database (Oxford). 2015 Sep 30;2015:bav095. doi: 10.1093/database/bav095. PMID: 26424082; PMCID: PMC4589691.

38. Yoo M, Shin J, Kim J, Ryall KA, Lee K, Lee S, Jeon M, Kang J, Tan AC. DSigDB: drug signatures database for gene set analysis. Bioinformatics. 2015 Sep 15;31(18):3069-71. doi: 10.1093/bioinformatics/btv313. Epub 2015 May 19. PMID: 25990557; PMCID: PMC4668778.

39. Yoo M, Shin J, Kim J, Ryall KA, Lee K, Lee S, Jeon M, Kang J, Tan AC. DSigDB: drug signatures database for gene set analysis. Bioinformatics. 2015 Sep 15;31(18):3069-71. doi: 10.1093/bioinformatics/btv313. Epub 2015 May 19. PMID: 25990557; PMCID: PMC4668778.

40. Konstantin Shvachko, Hairong Kuang, Sanjay Radia, Robert Chansler, "The Hadoop Distributed File System" Yahoo! Sunnyvale, California USA, IEEE 2010. Electronic ISBN:978-1-4244-7153-9, DOI: 10.1109/MSST.2010.5496972.

41. Pujar MK, Vastrad B, Vastrad C. Integrative Analyses of Genes Associated with Subcutaneous Insulin Resistance. Biomolecules. 2019 Jan 22;9(2):37. doi: 10.3390/biom9020037. PMID: 30678306; PMCID: PMC6406848.

42. Rock FL, Hardiman G, Timans JC, Kastelein RA, Bazan JF (January 1998). "A family of human receptors structurally related to Drosophila Toll". Proceedings of the National Academy of Sciences of the United States of America. 95 (2): 588–93. doi:10.1073/pnas.95.2.588. PMC 18464. PMID 9435236.

43. Sharma BK, Kakker NK, Bhadouriya S, Chhabra R. Effect of TLR agonist on infections bronchitis virus replication and cytokine expression in embryonated chicken eggs. Mol Immunol. 2020 Apr;120:52-60. doi: 10.1016/j.molimm.2020.02.001. Epub 2020 Feb 14. PMID: 32065987; PMCID: PMC7112572.

44. Proud PC, Tsitoura D, Watson RJ, Chua BY, Aram MJ, Bewley KR, Cavell BE, Cobb R, Dowall S, Fotheringham SA, Ho CMK, Lucas V, Ngabo D, Rayner E, Ryan KA, Slack GS, Thomas S, Wand NI, Yeates P, Demaison C, Zeng W, Holmes I, Jackson DC, Bartlett NW, Mercuri F, Carroll MW. Prophylactic intranasal administration of a TLR2/6 agonist reduces upper respiratory tract viral shedding in a SARS-CoV-2 challenge ferret model. EBioMedicine. 2021 Jan;63:103153. doi: 10.1016/j.ebiom.2020.103153. Epub 2020 Dec 3. PMID: 33279857; PMCID: PMC7711201.

45. Lundwall A, Dackowski W, Cohen E, Shaffer M, Mahr A, Dahlbäck B, Stenflo J, Wydro R. "Isolation and sequence of the cDNA for human protein S, a regulator of blood coagulation". Proc. Natl. Acad. Sci. U.S.A. 83 (18): 6716–20. doi:10.1073/pnas.83.18.6716. PMC 386580. PMID 2944113.
46. Long GL, Marshall A, Gardner JC, Naylor SL. "Genes for human vitamin K-dependent plasma proteins C and S are located on chromosomes 2 and 3, respectively". Somat. Cell Mol. Genet. 14 (1): 93–8. doi:10.1007/BF01535052. PMID 2829367. S2CID 31236887.

47. Koolen MJ, Borst MA, Horzinek MC, Spaan WJ. Immunogenic peptide comprising a mouse hepatitis virus A59 B-cell epitope and an influenza virus T-cell epitope protects against lethal infection. J Virol. 1990 Dec;64(12):6270-3. doi: 10.1128/JVI.64.12.6270-6273.1990. PMID: 1700835; PMCID: PMC248803.

48. Lenschow DJ, Su GH, Zuckerman LA, Nabavi N, Jellis CL, Gray GS, et al. (December 1993). "Expression and functional significance of an additional ligand for CTLA-4". Proceedings of the National Academy of Sciences of the United States of America. 90 (23): 11054–8. doi:10.1073/pnas.90.23.11054. PMC 47920. PMID 7504292.

49. Suvas S, Singh V, Sahdev S, Vohra H, Agrewala JN. Distinct role of CD80 and CD86 in the regulation of the activation of B cell and B cell lymphoma. J Biol Chem. 2002 Mar 8;277(10):7766-75. doi: 10.1074/jbc.M105902200. Epub 2001 Nov 28. PMID: 11726649.

50. Giguère JF, Bounou S, Paquette JS, Madrenas J, Tremblay MJ. Insertion of host-derived costimulatory molecules CD80 (B7.1) and CD86 (B7.2) into human immunodeficiency virus type 1 affects the virus life cycle. J Virol. 2004 Jun;78(12):6222-32. doi: 10.1128/JVI.78.12.6222-6232.2004. PMID: 15163715; PMCID: PMC416533.

51. Lin Y, Sibanda VL, Zhang HM, Hu H, Liu H, Guo AY. MiRNA and TF co-regulatory network analysis for the pathology and recurrence of myocardial infarction. Sci Rep. 2015 Apr 13;5:9653. doi: 10.1038/srep09653. PMID: 25867756; PMCID: PMC4394890.

52. Garland SM. Imiquimod. Curr Opin Infect Dis. 2003 Apr;16(2):85-9. doi: 10.1097/00001432-200304000-00004. PMID: 12734440.

53. Gupta AK, Browne M, Bluhm R. Imiquimod: a review. J Cutan Med Surg. 2002 Nov-Dec;6(6):554-60. doi: 10.1007/s10227-001-0134-6. Epub 2002 Oct 9. PMID: 12362256.

54. Bilu D, Sauder DN. Imiquimod: modes of action. Br J Dermatol. 2003 Nov;149 Suppl 66:5-8. doi: 10.1046/j.0366-077x.2003.05628.x. Erratum in: Br J Dermatol. 2006 Jun;154(6):1220. Bilu, Donna [added]. PMID: 14616337.

55. Suzuki H, Wang B, Shivji GM, Toto P, Amerio P, Tomai MA, Miller RL, Sauder DN. Imiquimod, a topical immune response modifier, induces migration of Langerhans cells. J Invest Dermatol. 2000 Jan;114(1):135-41. doi: 10.1046/j.1523-1747.2000.00833.x. PMID: 10620129.

56. Dahl MV. Imiquimod: a cytokine inducer. J Am Acad Dermatol. 2002 Oct;47(4 Suppl):S205-8. doi: 10.1067/mjd.2002.126586. PMID: 12271278.

57. Zhao S, Gao N, Qi H, Chi H, Liu B, He B, Wang J, Jin Z, He X, Zheng H, Wang Z, Wang X, Jin G. Suppressive effects of sunitinib on a TLR activation-induced cytokine storm. Eur J Pharmacol. 2019 Jul 5;854:347-353. doi: 10.1016/j.ejphar.2019.04.045. Epub 2019 Apr 27. PMID: 31039345.