RESEARCH ARTICLE

Identifying potential drug targets and candidate drugs for COVID-19: biological networks and structural modeling approaches [version 1; peer review: 1 approved with reservations]

Gurudeeban Selvaraj[1,2], Satyavani Kaliamurthi[1,2], Gilles H. Peslherbe[1], Dong-Qing Wei[2-4]

[1]Centre for Research in Molecular Modeling, Concordia University, Montreal, Quebec, H4B 1R6, Canada
[2]Centre of Interdisciplinary Science-Computational Life Sciences, College of Chemistry and Chemical Engineering,, Henan University of Technology, Zhengzhou, Henan, 450001, China
[3]The State Key Laboratory of Microbial Metabolism, College of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, Shanghai, 200240, China
[4]IASIA (International Association of Scientists in the Interdisciplinary Areas), 125 Boul. de Bromont, Quebec, J2L 2K7, Canada

v1 **First published:** 18 Feb 2021, **10**:127
https://doi.org/10.12688/f1000research.50850.1
Latest published: 06 Apr 2021, **10**:127
https://doi.org/10.12688/f1000research.50850.2

Abstract

Background: Coronavirus (CoV) is an emerging human pathogen causing severe acute respiratory syndrome (SARS) around the world. Earlier identification of biomarkers for SARS can facilitate detection and reduce the mortality rate of the disease. Thus, by integrated network analysis and structural modeling approach, we aimed to explore the potential drug targets and the candidate drugs for coronavirus medicated SARS.
Methods: Differentially expression (DE) analysis of CoV infected host genes (HGs) expression profiles was conducted by using the Limma. Highly integrated DE-CoV-HGs were selected to construct the protein-protein interaction (PPI) network.
Results: Using the Walktrap algorithm highly interconnected modules include module 1 (202 nodes); module 2 (126 nodes) and module 3 (121 nodes) modules were retrieved from the PPI network. MYC, HDAC9, NCOA3, CEBPB, VEGFA, BCL3, SMAD3, SMURF1, KLHL12, CBL, ERBB4, and CRKL were identified as potential drug targets (PDTs), which are highly expressed in the human respiratory system after CoV infection. Functional terms growth factor receptor binding, c-type lectin receptor signaling, interleukin-1 mediated signaling, TAP dependent antigen processing and presentation of peptide antigen via MHC class I, stimulatory T cell receptor signaling, and innate immune response signaling pathways, signal transduction and cytokine immune signaling pathways were enriched in the modules.

Open Peer Review

Reviewer Status ✔ ? ?

	Invited Reviewers		
	1	**2**	**3**
version 2 (revision) 06 Apr 2021	✔ report	? report	? report
version 1 18 Feb 2021	? report		

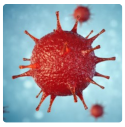
1. **Sugunadevi Sakkiah**, National Center for Toxicological Research, Jefferson, USA
2. **Dhurvas C. Dinesh**, Institute of Organic Chemistry and Biochemistry, Prague, Czech Republic
3. **Turanli Beste**, Marmara University, Istanbul, Turkey

Any reports and responses or comments on the

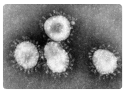
Protein-protein docking results demonstrated the strong binding affinity (-314.57 kcal/mol) of the ERBB4-3cLpro complex which was selected as a drug target. In addition, molecular dynamics simulations indicated the structural stability and flexibility of the ERBB4-3cLpro complex. Further, Wortmannin was proposed as a candidate drug to ERBB4 to control SARS-CoV-2 pathogenesis through inhibit receptor tyrosine kinase-dependent macropinocytosis, MAPK signaling, and NF-kb singling pathways that regulate host cell entry, replication, and modulation of the host immune system.
Conclusion: We conclude that CoV drug target "ERBB4" and candidate drug "Wortmannin" provide insights on the possible personalized therapeutics for emerging COVID-19.

Keywords
Biological network analysis, COVID-19, drug targets, ERBB4, growth factor receptor binding, Limma, protein-protein docking, SARS-CoV-2, signal transduction pathways, Walktrap algorithm, wortmannin

article can be found at the end of the article.



This article is included in the Disease Outbreaks gateway.



This article is included in the Coronavirus collection.

Corresponding authors: Gurudeeban Selvaraj (gurudeeb99@haut.edu.cn), Gilles H. Peslherbe (gilles.peslherbe@concordia.ca), Dong-Qing Wei (dqwei@sjtu.edu.cn)

Author roles: Selvaraj G: Conceptualization, Formal Analysis, Methodology, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Kaliamurthi S**: Data Curation, Formal Analysis, Methodology, Validation, Visualization, Writing – Original Draft Preparation; **Peslherbe GH**: Conceptualization, Funding Acquisition, Investigation, Resources, Supervision, Writing – Review & Editing; **Wei DQ**: Conceptualization, Funding Acquisition, Resources, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: Partial support of this work was provided by the National Natural Science Foundation of China (DQ.W. -Grant no. 61832019, 61503244), the Ministry of Science and Technology of China (DQ.W.-Grant no.: 2016YFA0501703); Natural Sciences and Engineering Research Council (NSERC) of Canada (G.H.P. - Grant no. 216940); G.S. and S.K. are grateful to the MITACS Global Research award, Concordia University for bridge funding from the Faculty of Arts and Science, the Office of the Vice-President Research and Graduate Studies, and Concordia International. Computational resources were provided by Compute Canada and the Centre for Research in Molecular Modeling.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Selvaraj G, Kaliamurthi S, Peslherbe GH and Wei DQ. **Identifying potential drug targets and candidate drugs for COVID-19: biological networks and structural modeling approaches [version 1; peer review: 1 approved with reservations]** F1000Research 2021, **10**:127 https://doi.org/10.12688/f1000research.50850.1

First published: 18 Feb 2021, **10**:127 https://doi.org/10.12688/f1000research.50850.1

Introduction

Coronavirus (CoV) are the largest group of enveloped and single-stranded-ribonucleic acid (ssRNA) pathogens[1]. The genome of CoV is typically larger (27-33Kb) than all other RNA viruses. Briefly, the genome contains an envelope protein, membrane protein, nucleoprotein, and spike protein which are responsible for capsid formation, assembly of virus particles and entry of the virus particle into the host[2]. The CoVs are zoonotic in nature, i.e. transmitted between animals and humans. There were three unforgettable zoonotic infectious diseases outbreak caused by CoVs recorded in the past two decades. In the 20th century (2002–2003), the foremost epidemic outbreak was originated in China by the severe acute respiratory syndrome (SARS)[3]. Later, in the 21st century (2012), the next epidemic outbreak happened in Saudi Arabia and the gulf countries with the Middle East respiratory syndrome (MERS)[4]. In December 2019, the third epidemic outbreak was recorded in China, especially from Wuhan city through a novel severe acute respiratory syndrome CoV-2 (SARS-CoV-2) (COVID-19)[5]. Paraskevis *et al.* (2020)[2] reported that the SARS-CoV-2 full-genome belongs to betacoronavirus, but it differs from the epidemic causing SARS and MERS[6]. Moreover, the SARS-CoV-2 genome exhibits 96.3% similarity with the Bat-SARS viruses like CoVs. Among the various human CoVs (229E, NL63, HKU1, OC43, SARS, and MERS), SARS and MERS caused severe respiratory-related mortality rates of 10% and 37% respectively[7].

Recently, Huang *et al.*[8] reported that the most widespread symptoms of COVID-19 patients at the beginning of the disease's condition were respectively fever (98%), cough (76%), fatigue (44%) followed by sputum production (28%), headache (8%), blood-stained mucus (5%) and diarrhea (3%). Moreover, nearly one-fifth (19.2%) of individuals with COVID-19 were asymptomatic[9]. COVID-19 positive cases showing the most common symptoms were cough hyposmia, sputum, and fever respectively[9]. In this current pandemic situation, identification of novel biomarkers, which plays an effective role in prognosis and monitoring the status of the SARS-COV-2 disease condition, is necessary. A high-throughput oligonucleotide microarray profiling has been widely employed in measuring the expression of genes (>1000) level significantly. In addition, microarray facilitates the identification of biomarkers (diagnosis/therapeutic), classification of the disease condition, treatment options, and mechanism of action of the gene involved in pathogenesis[10]. Biomarkers are identified through the determination of DEGs between the control and diseased/infected subjects, conversely finding the proteins that each biomarker is in connection with them, assisting to discover the key pathways related to the mechanism of the disease[11–13]. Many genes were involved in the pathogenesis, which make-up very complicated network systems. Coronavirus and influenza are pathogens causing severe respiratory illness to humans[14]. Recent reports demonstrated that SP110, HERC5, SAMD9L, RTP4, ESPT11 genes were identified to control and improve the immune system and defense mechanism against influenza by network analysis[15]. It has the supremacy to act as biomarkers and targets for pediatric influenza mediated therapy. A recent systematic review by 16 and his colleagues showed the increased level of biomarkers such as C-reactive protein, serum amyloid A, interleukin-6, lactate dehydrogenase, D-dimer, cardiac troponin and renal biomarkers (urea and creatinine) were higher, and low level of lymphocytes and platelet count were recorded in severe complicated COVID-19 patients than non-severe COVID-19 patients' plasma and infected lung tissues.

Thus, the present study aimed to identify modules consists of the highly interconnected genes, which are involved in the pathogenesis of coronavirus medicated respiratory syndrome in humans. Microarray analysis of gene expression profiles is a standard and well-known method to identify potential targets and pathways[17,18]. Initially, we collected the gene expression profiles from the Gene Expression Omnibus (GEO) database. Then, we performed differential expression of CoV infected host genes (DE-CoV-HGs) using the Limma algorithm. The protein-protein interaction (PPI) network was constructed with DE-CoV-HGs and extracted different modules from the network using the Walktrap algorithm. Besides, the potential targets were predicted from the selected modules based on the degree and betweenness centrality measures, and determined their functional and pathway enrichment terms. The possible validation of potential targets interaction with proteins of SARS-CoV-2 was performed using protein-protein docking (PPD) and molecular dynamics (MD) simulations. Further, possible candidate drug for defined potential drug target (PDT) "ERBB4" using Drug Gene Budger. An insight into the study provides possible personalized therapeutic target and candidate drug for coronavirus medicated respiratory syndrome. Figure 1 illustrates the work flow of the study.

Methods

Data collection

The coronavirus infected host gene expression profiles were extracted from the GEO database[19]. Agilent-014850 whole human genome microarray 4x44K G4112F platform was employed to GSE100509, GSE86529, GSE47962, GSE47961, GSE47960, GSE45042, GSE37827 and GSE33267 studies and one study GSE56677 which was performed based on Agilent-039494 SurePrint G3 human GE v2 8x60K microarray 039381 platform was included[20–24]. All the selected studies contain control and CoV infected human lung epithelial cells samples. The samples infected with other viruses were excluded. The sample details of different datasets are indicated in Table 1.

Differential expression of CoV infected host genes

Using NetworkAnalyst (NA) web interface, the tab-delimited text (.txt) files of selected datasets were given as an input for preprocessing, normalization, and probe identification[25]. Variance stabilizing normalization (VSN) and quantile normalization was applied to reduce false-positive errors and equal distributions of datasets for statistical analysis[26,27]. R package "LIMMA" (Linear models for microarray analysis) of from the Bioconductor project were used to perform differential expression of CoV infected host genes[28]. Log2 transformation, Benjamini and Hochberg and t-test were used to perform normalization and calculate false discovery rate (FDR; $p<0.05$) of samples[29].

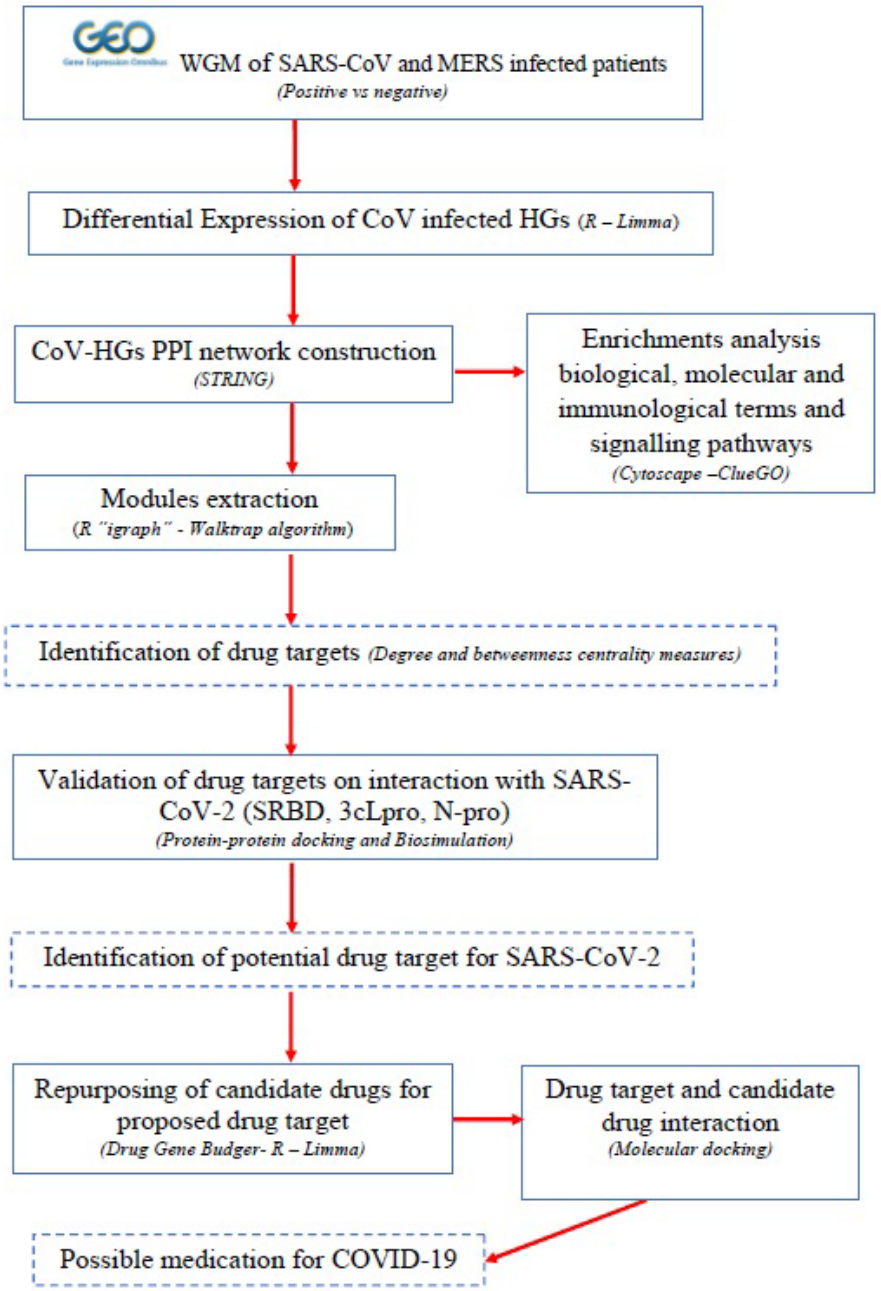


Figure 1. Schematic illustration of the study. WGM-whole genome microarray; SARS-CoV- ; MERS-; HGs-host genes; PPI-protein-protein interaction; SRBD-spike receptor binding domain; 3cLpro-3C-like protease; N pro-nucleocapsid protein;

Construction of DE-CoV-HGs PPI network, modules and PDTs identification

We constructed the PPI network from identified DE-CoV-HGs using Search tool for retrieval of interacting genes/proteins (STRING) interactome[30]. The highest confidence interaction score was set to 0.9, which reduces false positive interactions[31]. The random walks (R package "*igraph*") was used to extract modules based on the Walktrap algorithm from the DE-CoV-HGs interaction network[32]. It runs several short random walks within a group of nodes that are highly connected to detect small modules. From the modules, the hub genes (nodes) were identified using two different centrality measures "degree" and "betweenness"[32]. The degree of the gene is the many connections it has to other genes. Genes with a high degree act as hubs within the network. The betweenness of a gene is the number of paths that pass through it when considering

Table 1. List of microarray datasets used in this study.

Profile	Platform	Sample size (C vs I)	Pathogen	Reference
GSE47962	GPL6480	60 (33 vs 27)	SARS-CoV	20
GSE47961	GPL6480	60 (33 vs 27)	SARS-CoV	20
GSE47960	GPL6480	70 (38 vs 32)	SARS-CoV	20
GSE45042	GPL6480	32 (15 vs 17)	SARS-CoV	21,24
GSE37827	GPL6480	60 (30 vs 30)	SARS-CoV	22
GSE33267	GPL6480	66 (33 vs 33)	SARS-CoV	22
GSE100509	GPL13497	50 (25vs 25)	MERS-CoV	24
GSE86529	GPL13497	50 (25vs 25)	MERS-CoV	24
GSE56677	GPL17077	33 (15 vs 18)	MERS-CoV	23

Note: C- Control; I- Infected; GPL13497 -Agilent-026652 Whole Human Genome Microarray 4x44K v2; GPL17077-Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray 039381; GPL6480 -Agilent-014850 Whole Human Genome Microarray 4x44K G4112F

the pair-wise shortest paths between all genes in the network. A node that occurs between two dense clusters will have a high betweenness. In the present study, the PDTs of DE-CoV-HGs were identified based on the degree centrality and betweenness centrality measures.

Pathway enrichment analysis

We have used ClueGO v2.5.3, which is a Cytoscape plugin for function and pathway enrichment analysis of DE-CoV-HGs[33]. A list of PDTs were provided as input into ClueGO with select specific parameters like species such as-Homo sapiens; ID type-Entrez gene ID; enrichment functions -KEGG pathways for the analysis. Each enrichment was calculated based on the Bonferroni method (kappa score 0.96; $p > 0.005$). ImageGP was employed to visualize the results of functional enrichment analysis.

Structural modeling approach to study host-SARS-CoV-2 proteins interaction and drug-target interaction

Preparation of host and SARS-CoV-2 viral receptor. The FASTA sequence of selected proteins namely 3C-like protease (3cLpro), angiotensin I converting enzyme 2 (ACE2), B-cell lymphoma 3 (BCL3), casitas b-lineage lymphoma (CBL), CCAAT enhancer binding protein beta (CEBPB), CRK like proto-oncogene (CRKL), histone deacetylase 9 (HDAC9), Kelch-like protein 12 (KLHL12), v-myc myelocytomatosis viral oncogene (MYC), mothers against decapentaplegic homolog 3 (SMAD3), nucleocapsid protein (Npro), nuclear receptor coactivator 3 (NCOA3), receptor tyrosine-protein kinase erbB-4 (ERBB4), spike protein-receptor binding domain (SRBD), SMAD-specific E3 ubiquitin protein ligase 1 (SMURF1), ubiquitin C (UBC), and vascular endothelial growth factor A (VEGFA) were obtained from Uniprot for protein-protein docking studies (Table 3)[34].

Protein-protein docking. PPD was performed in the HDOCK web-interface using *ab initio* docking with a fast Fourier transform (FFT) based hierarchical algorithm[35,36]. The FASTA sequence of the selected protein (protein 1 and protein 2) was given as input. The server obtained and chosen one of the highest sequence coverage and similarity and with the highest resolution template automatically from PDB and then, the homology modeling was performed using MODELLER. Then, the corresponding structures constructed by superimposing the modeled protein structure on to the template, where 5,000steps of AMBER MD minimization was also performed to remove severe atomic clashes at the interfaces. The putative binding modes between the protein-1 and protein-2 are systematically samples based on the FFT algorithm. In total 4,392 binding modes are ranked depending on their cluster and binding energy scores. The final complex was selected with RMSD of two binding modes was below a cut off 5Å. RMSD was calculated based on the backbone atoms of the protein 2[36].

$$u_{xy}^{a+1}(z) = u_{xy}^{a}(z) + \Delta u_{xy}^{a}(z),$$

$$\Delta u_{xy}^{a}(z) = \frac{1}{2} a_B T \left[h_{xy}^{a}(z) - h_{xy}^{obs}(z) \right]$$

Whereas, "*a*" indicates as iterative step. "*x*" and "*y*" indicates the types of a pair of atoms in the protein 1 and protein 2. $h_{xy}^{obs}(z)$ is the pair distribution function for atom pair xy calculated from the ensemble of experimentally measured protein 1- protein 2 complex structures. $h_{xy}^{a}(z)$ is the pair distribution function calculated from the ensemble of the binding modes of the protein 1 – protein 2 complexes predicted with the trial potentials $\{h_{xy}^{a}(z)\}$ at the *a*-th step. $u_{xy}^{a+1}(z)$ are the improved potentials from $u_{xy}^{a}(z)$. $a_B T$. is the generality value set as 1[37].

Molecular dynamics simulations. The results of the PPD complex with the highest binding affinity were used to prepare input files for MD simulations. Solution builder in the CHARMM-GUI web interface was employed to solvate the complex with TIP3P and neutralized by potassium chloride (KCl) ions at the 0.15 mol-1 concentration under CHARMM36 force field[38,39]. The initial configuration of KCl ions is then estimated using short Monte Carlo simulations (2000 steps) through Coulombic and the van der Waals (vdW) interactions. The solvated complex was simulated by using Nanoscale molecular dynamics (NAMD)[40]. Full system long-range Coulombic interactions were determined by the particle mesh Ewald (PME) method[41]. The integrator parameters time was set to 2 fs/step. The simulations were executed in the constant pressure and temperature (NpT ensemble) with pressure bar (1) and temperature (303 K) through the Langevin dynamics with a damping coefficient of 1/ps. After 9×10^7 steps of minimization, water molecules and ions were equilibrated for 2 ns around proteins, which were restrained using harmonic forces. MD simulation of selected drug target (ERBB4-3cLpro) was started from the last frame of restrained equilibration. The production was carried out for 150 ns.

Identification of possible candidate drug for PDT and molecular docking

Drug Gene Budger a web-interface was employed to identify possible candidate drugs or small molecules in LINCS L1000 data that significantly regulated selected drug target (ERBB4) expression based on LIMMA algorithm[42]. While entering the gene symbol of drug target, we could retrieve the following information such as log-transformed fold change, *p*-value and *q*-value for each potential small molecule. Small molecules/drugs that are liable for considerable overexpression and under expression of downregulated and upregulated drug target can increase the therapeutic action against COVID-19. The threshold *q* value was derived from the following equation

$$Q = X/Y = X/(X+Z)$$

Where X is the number of false discoveries and Z is the number of true discoveries. Y=X+Z is the number of rejected null hypotheses. The false discovery rate (FDR) is as follows

$$FDR = Qe = E\,[Q]$$

Where E [Q] is the expected value of Q. The drugs/small molecule have the highest log2FC value was selected for further protein-docking studies. The drug target ERBB4 sequence was retrieved from the Uniprot database (ID: Q15303) for three structure modeling. The protein structure modeling was performed in a Swiss modeling server. The structure of the candidate drug "wortmannin" was retrieved from the ZINC database (ID: ZINC1619592). Then, protein-ligand docking was performed to calculate the binding energy of the docked complex in the Swiss dock web-interface, which was developed based on the EADock dihedral space sampling (DSS) algorithm. The energy functions were calculated using the CHARMM force field. Each docking cluster was the output of 250 different sequential runs. The binding models (BM) having the most favorable energies were calculated by fast analytical continuum treatment of solvation (FACTS) and clustered. Binding modes were estimated through Full fitness and clustered. Then, the Full fitness of the clusters was estimated by averaging 30% of the most favorable effective energy of their elements.

The effective energy calculated by the following formula

$$G_{eff.eng} = E_{intra}^{Ligand} + E_{intra}^{receptor} + E_{inter} + \Delta G_{elec.solv} + \sigma \times SASA$$

Whereas, E_{intra}^{Ligand}, $E_{intra}^{receptor}$ are the internal energies of ligand and receptor; E_{inter} is interaction energy between the receptor and ligand; σ value (0.0072 kcal/mol); SASA is the solvent accessible surface area. The results were visualized in Maestro Visualizer. All the supporting data associated with this study freely available (http://doi.org/10.5281/zenodo.4458252).

Results

Differentially expressed -CoV-HGs

In total, 1,968 DE-CoV-HGs in human lung epithelial cells were obtained from the initial analysis. To obtain better results of biomarker identification, we have selected 315 overexpressed and 112 under expressed DE-CoV-HGs based on the log-fold changes -1>log2FC>1. Table 2 illustrates the top 10 DE-CoV-HGs, their description, and their fold change and *p* value.

Modules and potential drug targets identification

PPI network was constructed using 427 DE-CoV-HGs (Figure 1a). The major subnetwork contains 531 edges with an average local clustering coefficient 0.353 and the PPI enrichment *p*-value was 2.36 ×10-3. Then, using the Walktrap algorithm, 47 modules were retrieved from the subnetwork. Among

Table 2. Top 10 differentially expressed CoV-HGs.

Gene symbol	Name	Log2FC	*p* value
Overexpressed genes			
CALCA	Calcitonin related polypeptide alpha	3.149	2.78×10^{-12}
GALNTL2	UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-like 2	2.581	5.46×10^{-14}
SPIN4	Spindlin family, member 4	2.362	7.50×10^{-16}
ARL4A	ADP-ribosylation factor-like 4A	2.336	1.84×10^{-19}
CCL14	Chemokine (C-C motif) ligand 14	2.255	4.08×10^{-14}
PGM5	Phosphoglucomutase 5	2.229	7.28×10^{-14}
OTUD3	OTU deubiquitinase 3	2.213	6.63×10^{-16}
GIMAP7	GTPase, IMAP family member 7	2.202	2.52×10^{-15}
YAP1	Yes associated protein 1	2.178	1.22×10^{-04}
CLU	Clusterin	2.155	6.87×10^{-15}

Gene symbol	Name	Log2FC	*p* value
Under expressed genes			
SSX2	Synovial sarcoma, X breakpoint 2	-8.803	4.39×10^{-20}
FBXL8	F-box and leucine-rich repeat protein 8	-5.253	5.18×10^{-06}
NEURL3	Neutralized homolog 3 (Drosophila) pseudo gene	-5.084	5.86×10^{-15}
TSPAN18	Tetraspanin 18	-4.866	2.07×10^{-04}
FXYD3	FXYD domain containing ion transport regulator 3	-3.949	9.70×10^{-13}
CELF4	CUGBP, Elav-like family member 4	-3.616	2.95×10^{-04}
SLC12A1	Solute carrier family 12 member 1	-3.443	2.29×10^{-04}
BNC2	Basonuclin 2	-3.412	2.51×10^{-04}
KLF10	Kruppel-like factor 10	-3.386	5.91×10^{-22}
TNFAIP3	Tumor necrosis factor, alpha-induced protein 3	-3.314	1.68×10^{-16}

Table 3. List of drug targets and PPD interaction with SARS-CoV-2 proteins.

Potential target gene encoded host target proteins	Cellular components	Degree centrality	Betweenness centrality	Log2 FC	Uniprot ID	Reference PDB ID	PPD binding affinity (kcal/mol)		
							SRBD Uniprot ID: P0DTC2	3cLpro Uniprot ID: P0DTD1	N-protein Uniprot ID: P0DTC9
Bcl3	N, NP, CP	16	1364.72	-1.621	P20749	1K1A	-265.8	-288.55	-244.73
CBL	N, CP, C, PM	81	6405.31	1.046	P22681	5HKX	-285.74	-257.97	-244.71
CEBPB	N, NP, CP	28	2647.13	-1.869	P17676	3A5T	-235.81	-220.39	-227.41
CRKL	EN, C, EX	32	1158.21	1.081	P46108	2EYZ	-302.98	-273.78	-276.28
ERBB4*	**N, NP, M, C, PM**	**34**	**3037.11**	**-2.052**	Q15303	3U9U	**-289.18**	**-280.6**	**-281.11**
HDAC9	N, NP, CP	46	6027.63	1.027	Q9UKV0	5JI5	-286.06	-262.5	-261.77
KLHL12	GM, C	15	538.67	1.023	Q53G59	2VPJ	-290.49	-289.48	-279.35
MYC	N, NP, M, C	89	12884.19	-2.021	P01106	4ATI	-238.61	-244.15	-197.38
NCOA3	N, CP, EX	30	3580.11	1.059	Q9Y6Q9	4F3L	-265.44	-281.1	-239.99
SMAD3	N, CP, C, PM	69	5195.05	-1.353	P84022	1KHX	-288.66	-271.12	-254.11
SMURF1	N, CP, M, C, PM	48	3655.56	-1.163	Q9HCE7	5HPL	-269.9	-241.05	-233.53
VEGFA	C, PM	9	1123.93	0	P15692	2VWE	-234.15	-230.48	-243.59
UBC	N, NP, CP, C, PM	21	2832.51	-1.905	P0CG48	3RT3	-240.33	-218.27	-222.44
ACE	PM	-	-	-	Q9BYF1	1R42	-291.07	-	-

Note: * denotes ERBB4 selected as drug target; 3cLpro -3C-like protease; ACE2-angiotensin I converting enzyme 2; Bcl3-B-cell lymphoma 3; CBL- casitas b-lineage lymphoma; CEBPB- CCAAT enhancer binding protein beta; CRKL- CRK like proto-oncogene; C-cytosol; CP-cytoplasm; ERBB4- receptor tyrosine-protein kinase erbB-4; EN-Endosome; EX-exosome; FC -fold change; HDAC9- histone deacetylase 9; KLHL12- Kelch-like protein 12; M-mitochondria; MYC- v-myc myelocytomatosis viral oncogene; GM-golgi membrane; N protein- nucleocapsid protein; N-Nucleus; NP-nucleoplasm; PM-plasma membrane; PPD-protein-protein docking; SMAD3-mothers against decapentaplegic homolog 3; NCOA3- nuclear receptor coactivator 3; SMURF1-SMAD-specific E3 ubiquitin protein ligase 1; SRBD- spike protein-receptor binding domain; UBC-ubiquitin C; VEGFA-vascular endothelial growth factor A; FC-Fold change

them, highly interconnected modules includes module 1 (No. of nodes=202; $p=5.20\times10^{-32}$); module 2 (No. of nodes=126; $p=1.90\times10^{-28}$) and module 3 (No. of nodes=121; $p=3.13\times10^{-26}$) was employed to identify the PDTs (Figure 2b–d). Table 3 demonstrated that list of PDTs identified from different modules based on the degree and betweenness centrality measures. BCL3, CBL, CEBPB, CRKL, HDAC9, KLHL12, MYC, SMAD3, NCOA3, ERBB4, SMURF1, UBC, VEGFA were identified as PDTs which are highly expressed in human cells after CoV infection.

Functional and pathway enrichments of PPI network and PDTs

Many important gene ontology terms and pathways were identified from the PPI network namely, cellular response to growth factor stimulus, wound healing as a biological process term; growth factor receptor binding as a molecular function term; and TNF signaling pathway and transcriptional mis-regulation in cancer as a pathway enrichment terms (Figure 3). Moreover, the biological process term that is directly or indirectly related to the viral life cycle and human immune system includes RNA polymerase II promoter in response to hypoxia, peptidyl-tyrosine auto-phosphorylation, ERBB2 signaling pathway, anaphase-promoting complex dependent catabolic process, interleukin-1 mediated signaling pathway, TAP dependent antigen processing and presentation of peptide antigen via MHC class I, regulation of cellular amino acid and metabolic process, negative regulation of protein acetylation, negative regulation of gene silencing by miRNA, white fat cell differentiation and histone H4 deacetylation enriched in three different modules. Besides, stimulatory c-type lectin receptor signaling pathway (module 1), regulation of hematopoietic stem cell differentiation, T cell receptor signaling, innate immune response signaling and Fc-epsilon receptor signaling pathway (module 2), lymphocyte co-stimulation (module 3) were highly enriched immunological terms of three different modules (Figure 4). Further, pathways enrichment of modules demonstrated that numerous signal transduction pathways and cytokine immune signaling pathways were illustrated in Figure 5.

Molecular interaction of host drug targets with SARS-CoV-2 protein

Protein-protein docking interaction. The potential target genes encoding proteins are defined as drug targets to control

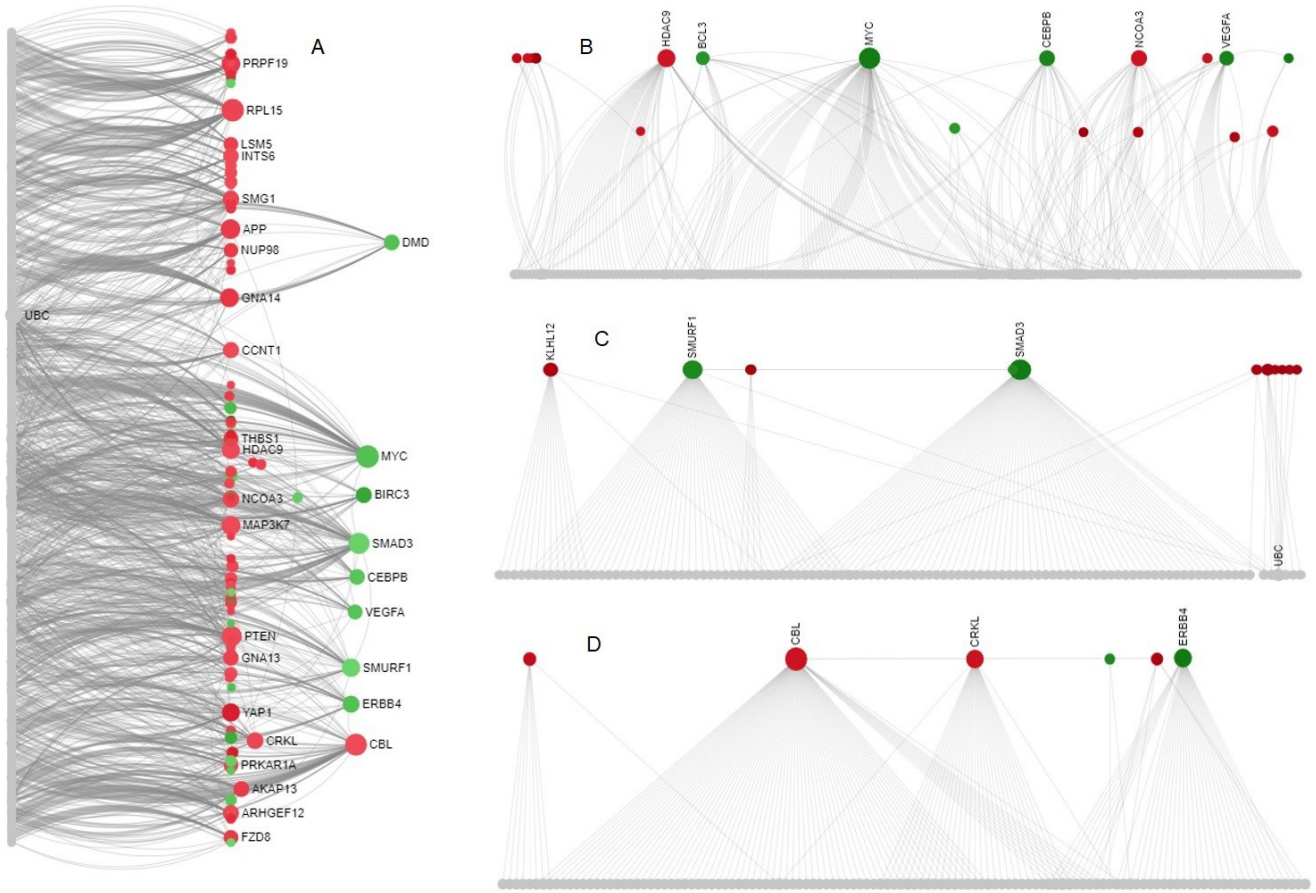


Figure 2. PPI network and Walk trap modules from DE-CoV-HGs. (**A**) PPI network with highly interconnected nodes; (**B**) Module-1; (**C**) Module-2; (**D**) Module-3. Overexpressed genes are represented in green color, under-expressed genes in red color, interconnected nodes in gray color.

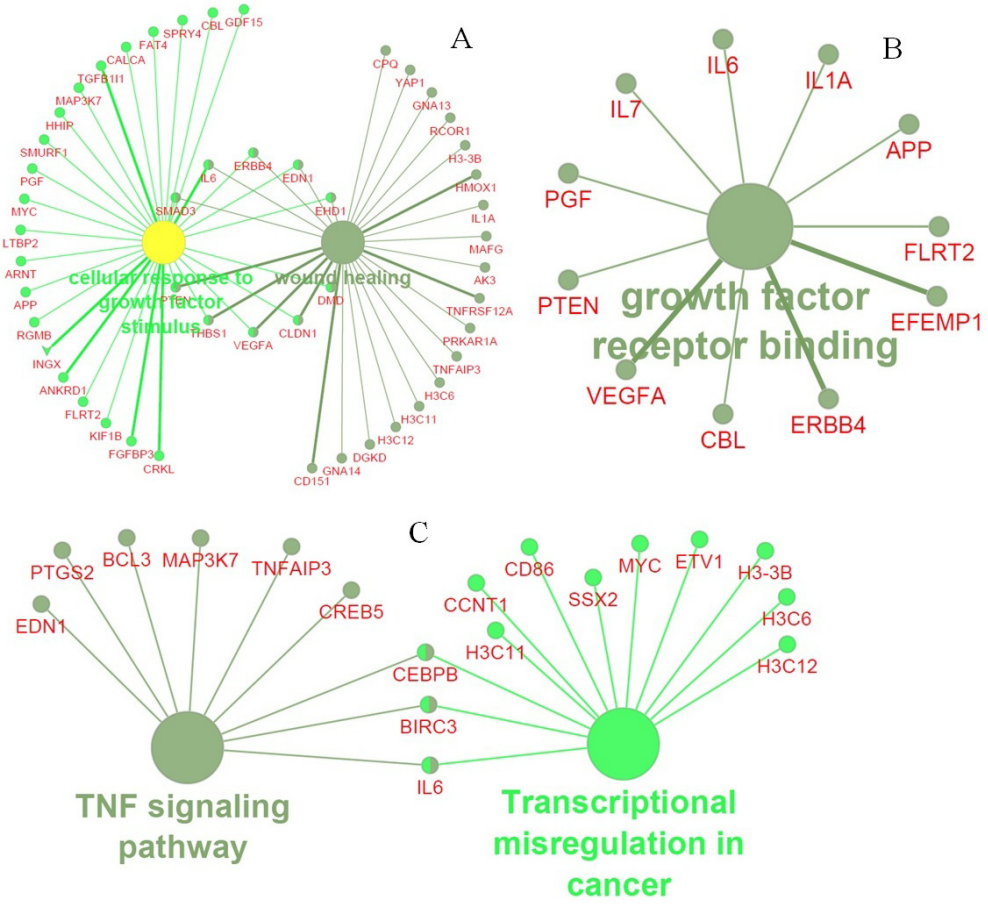


Figure 3. GO and Pathway enrichments of the PPI network. (**A**) Biological process term; (**B**) Molecular function term; (**C**) Pathway enrichment terms of PPI network

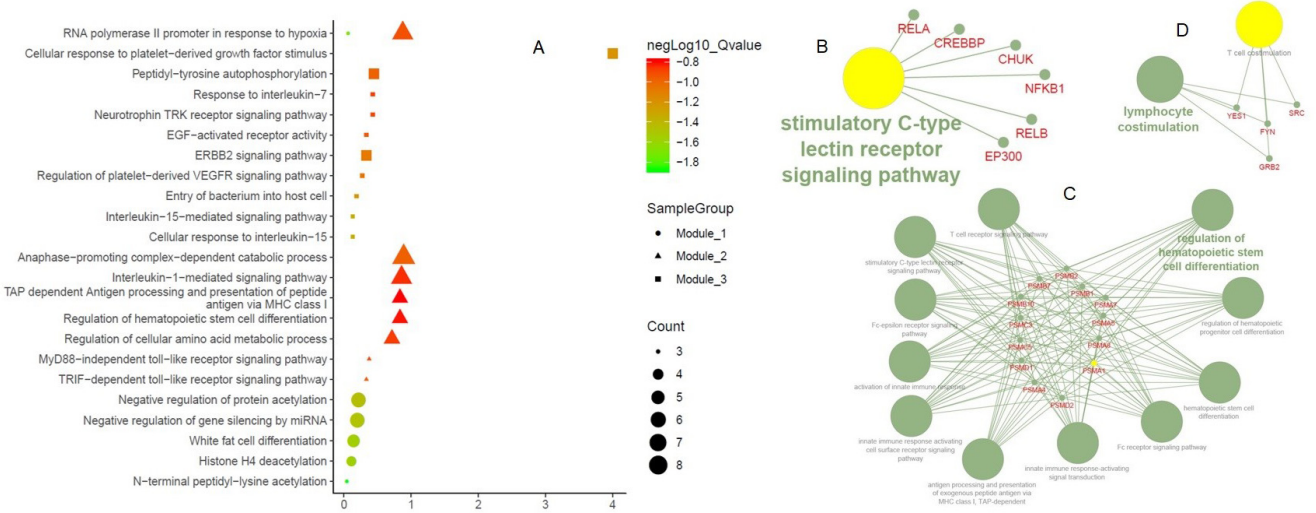


Figure 4. Functional enrichments of the Walk trap modules. (**A**) Biological process terms of three different modules; Immunological terms of (**B**) Module-1; (**C**) Module-2; (**D**) Module-3.

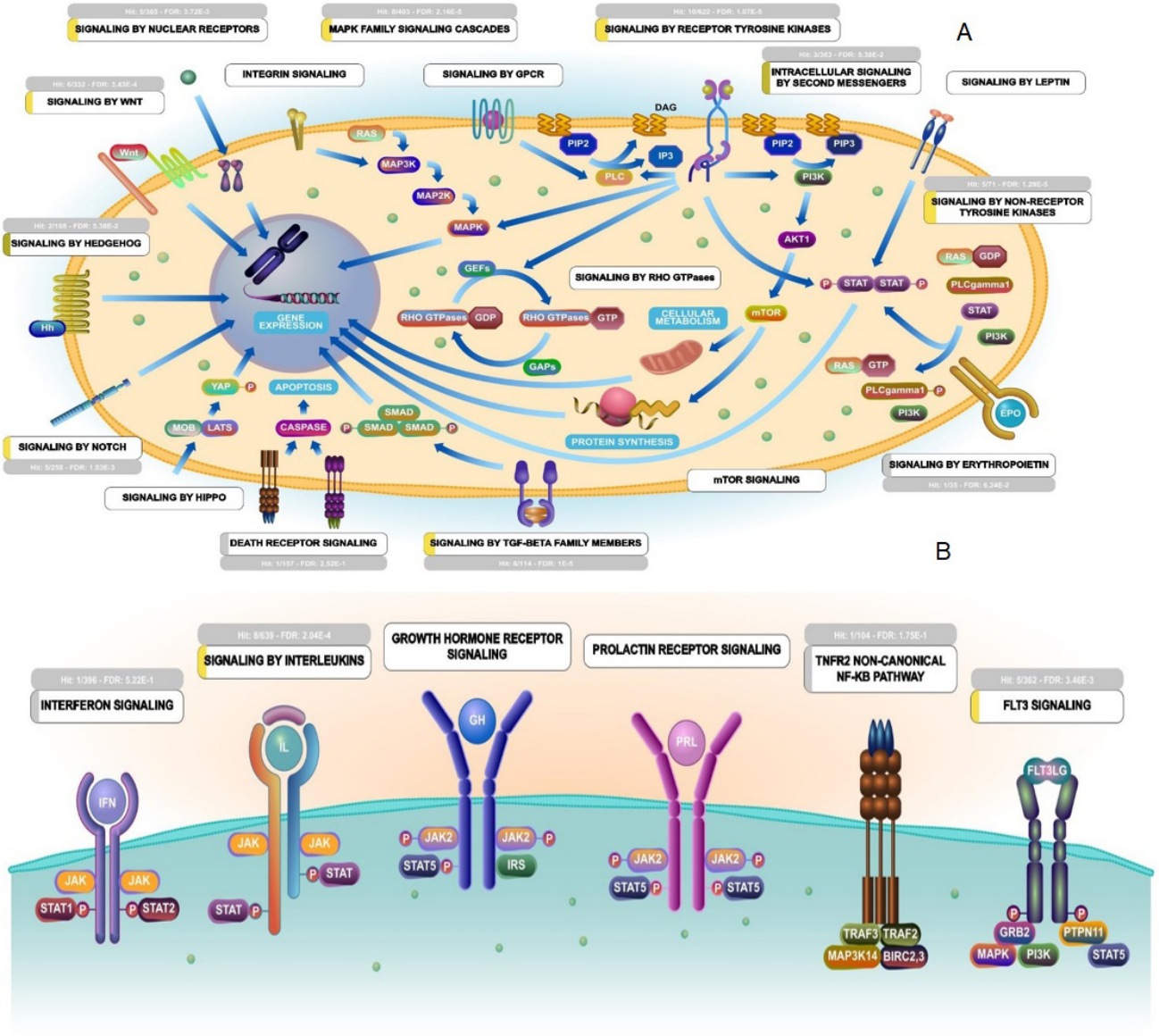


Figure 5. Pathways enrichment of Walktrap modules. (**A**) Signal transduction pathways; (**B**). Cytokine immune signaling pathways. The figures are retrieved from Reactome pathway analyzer using PTGs.

the pathogenesis of SARS-CoV-2. The drug targets regulated (overexpressed or under-expressed) by SARS-CoV-2 during pathogenesis in the host cytoplasm and different cellular components. Accordingly, PPD of SRBD, 3cLpro, and N-protein of SARS-CoV-2 with proposed drug targets happened by different biological signaling pathways. Thus, we hypothesized, proposed drug targets may control the pathogenesis of SARS-CoV-2 through the respective signaling pathways. The possible molecular interaction of host drug targets with SRBD, 3cLpro, and N-protein of SARS-CoV-2 was determined by PPD method. In the present study, we performed protein-protein rigid-body docking, which treats both proteins as rigid and discovers only six degrees of translational and rotational freedom. It excludes any kind of flexibility. Drug targets include HDAC9, SRBD, SMURF1, SRBD, CBL, ERBB4, BCL3, CEBPB, CRKL, KLHL12, MYC, NCOA3, SMAD3, UBC, and VEGFA were in the human respiratory system, which was possibly able to interact with the S-RBD, 3cLpro, and N - protein of SARS-CoV-2 and regulate the viral pathogenesis through replication and development. PPD binding affinities were illustrated in Table 3. Among the drug targets, ERBB4 demonstrated showed high binding affinity -260.46, -314.57 and -291.72 kcal/mol with SRBD, 3cLpro and N-protein of SARS-CoV-2 respectively. Thus “ERBB4” was selected as potential drug target and

used for candidate drug selection study. Hydrogen bond forming residues between the ERBB4 and SRBD, 3cLpro, and N-protein of SARS-CoV-2 complexes represented in Figure 6.

MD simulation of ERBB4-3cLpro complex. MD simulations demonstrated the stability of the ERBB4-3cLpro complex. The structure and dynamic properties of the complex were

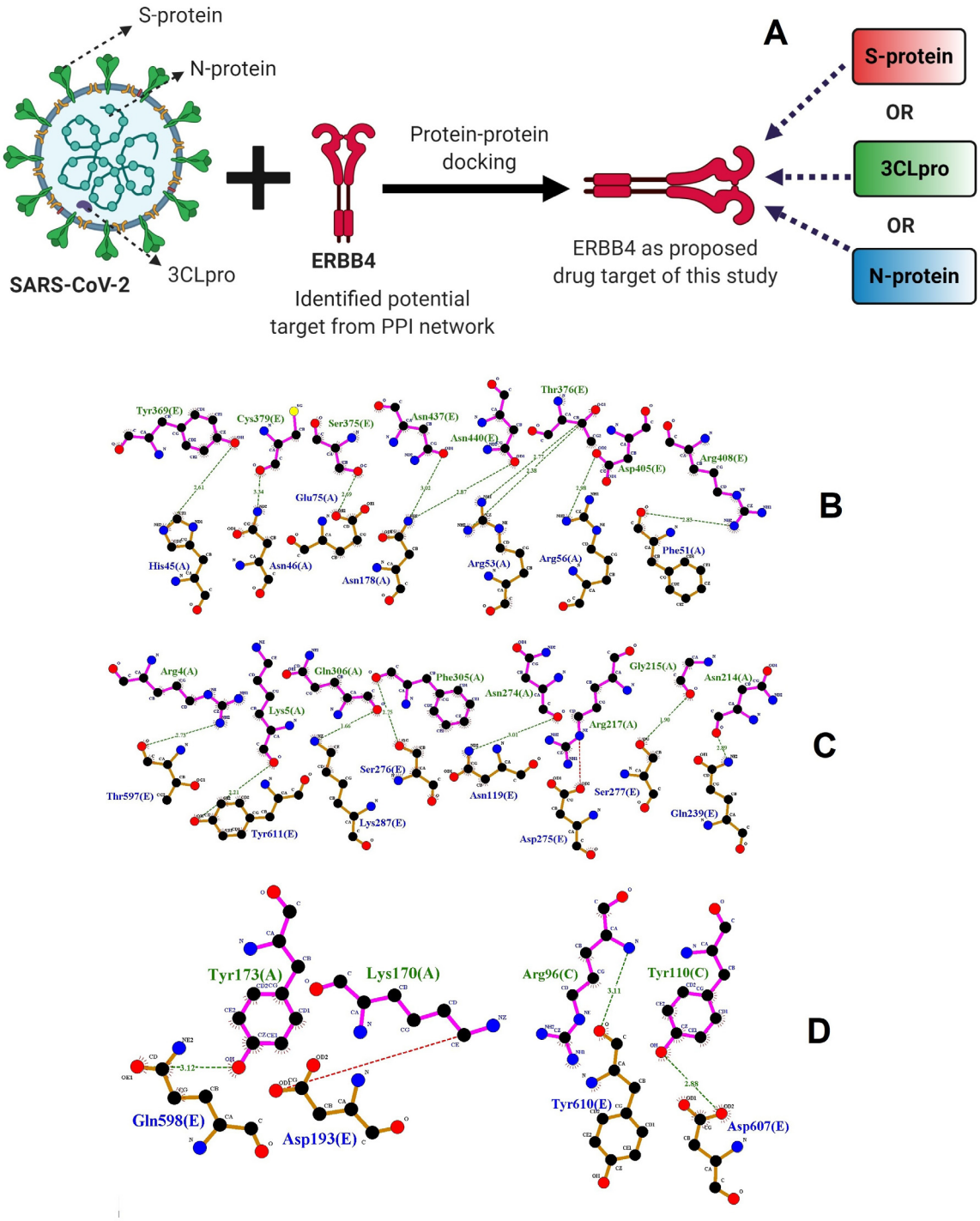


Figure 6. Possible action and PPD of drug target ERBB4 with SARS-CoV-2 proteins. (**A**) Possible pathological action of SRBD, 3cLpro, and N-protein of SARS-CoV-2 with "ERBB4". Results of PPD of (**B**) ERBB4-SRBD; (**C**) ERBB4-3cLpro; and (**D**) ERBB4-N-protein complex by using LigPlot. The interacting residues of ERBB4 (magenta) are labelled in green color and SARS-CoV-2 protein structures (brown) are labelled in blue color. The interacting hydrogen bonds indicated in green color.

followed by monitoring the backbone Cα root mean square deviation (RMSD) from the initial structure during the simulation period (150ns). The RMSD of the ERBB4-3cLpro complex gradually increased for the first 95 ns before steadily oscillating around a value of ~0.85 nm, indicative of the ERBB4-3cLpro complex stability over the simulation timescale (Figure 7a). The root mean square fluctuation (RMSF) reflects the fluctuations in the positions of the ERBB4-3cLpro complex residues. Briefly, notable fluctuations have been observed in the multiple furin-like cysteine-rich domains, a tyrosine kinase domain, and a phosphatidylinositol-3 kinase binding site of ERBB4. However, the other amino acid residues in the PDZ binding domain (amino acid range 400–600) were not deviating more (Figure 7b). This indicates the reason for the stable RMSD plot of the ERBB4-3cLpro complex. The strong hydrogen bonding interaction (average ~1.89) between ERBB4 and 3cLpro observed throughout the 150ns of MD simulation (Figure 7c).

Drug candidate for proposed drug target

The results of Drug Gene Budger demonstrated that Wortmannin (*q*- value = 3.31×10^{-21}; log2 Fold change = 2.997; Specificity = 1.09×10^{-04}) was identified as the candidate drug for the proposed target ERBB4. The ERBB4 was downregulated

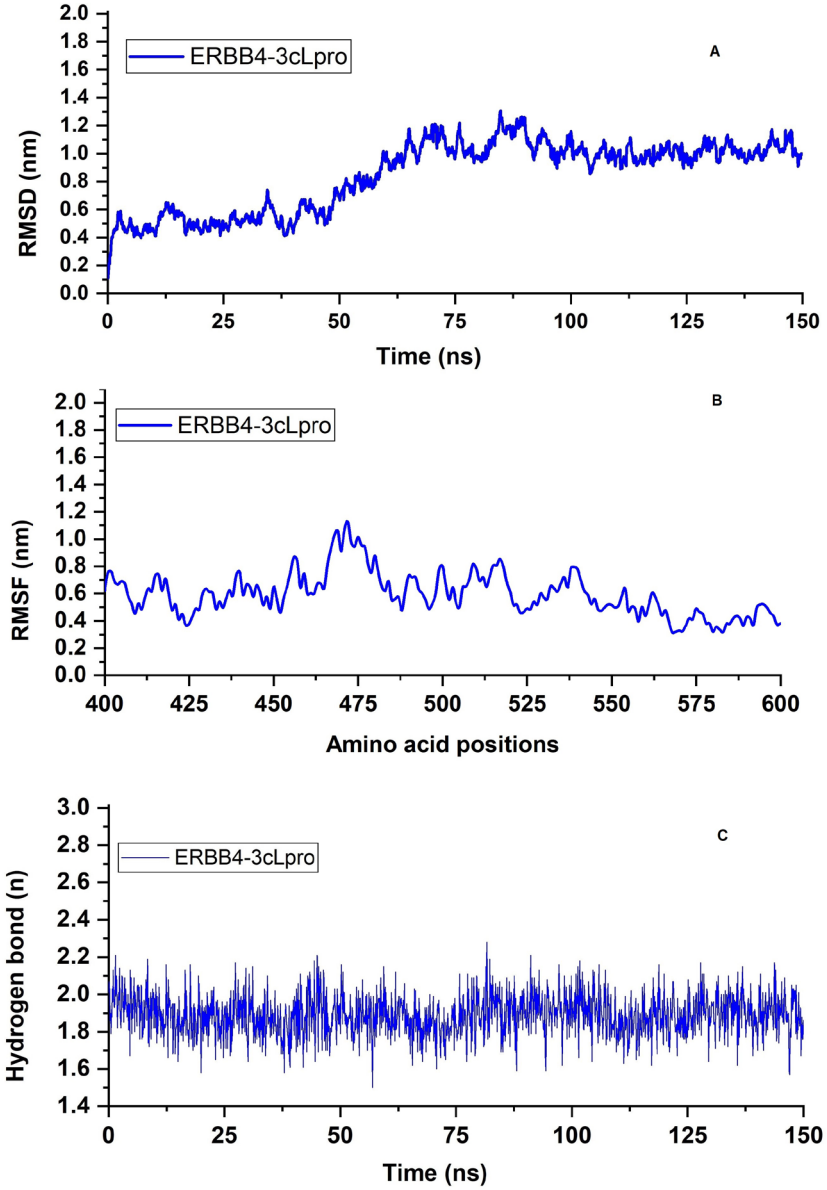


Figure 7. MD simulation of ERBB4-3cLpro complex. (**A**) NAMD trajectories of RMSD; (**B**) RMSF; (**C**) Total number of intermolecular hydrogen bond formation trajectories of ERBB4-3cLpro complex.

after the coronavirus infection that was then upregulated by the possible use of the Wortmannin as candidate drugs, which may inhibit the pathological cycle and development of SARS-CoV-2 in the human hosts. 30 different protein-ligand cluster of binding modes were obtained after molecular docking. Among them, the binding mode of cluster 1 indicated that the ERBB4-Wortmannin complex has the energy in term of full fitness -2563.6987 kcal/mol and Gibbs free energy ΔG was -7.93 kcal/mol (Figure 8).

Discussion

The discovery of biomarkers from gene-expression profiles as well as key functions in a pathogenic-related pandemic disease like COVID-19 is of consequence for diagnosis, drug development, and monitoring of disease progression. Network analysis provides new insights into the biological and cellular organization. The single gene can be involved in many biological functions and regulated different genes at different times[43]. Thus, biomarkers from gene-expression profiles of coronavirus infected subjects have the potential to be employed as drug-targets for COVID-19. Therefore, the present study aimed to identify drug targets from modules containing highly interconnected potential target genes which involve in the pandemic coronavirus pathogenesis. DEGs analysis demonstrated that 315 overexpressed and 112 under expressed genes between normal and CoV-infected subjects. Besides, the potential target genes namely MYC, HDAC9, NCOA3, CEBPB, VEGFA, BCL3, SMAD3, SMURF1, KLHL12, CBL, ERBB4, and CRKL were identified from Walktrap modules, which are significantly associated with the coronavirus infection. Then, the potential target genes encoding proteins employed as PDTs (receptor) for further analysis. Thus, we discussed the possible association of PDTs with pathogenesis of SARS-CoV-2.

MYC regulates numerous cellular functions including cell activation, differentiation, cell cycle progress, transformation, and apoptosis in virus (HIV, hepatitis C, influenza, and Epstein-Barr virus (EBV)) infected host cells[44–46]. In addition, mutation of MYC can disturb human B cells to proliferate indefinitely[47,48]. Further, influenza A virus infection dysregulates the expression of miRNA-22 and induces CD147 in asthmatics through the MYC transcription factor[49]. Moreover, overexpression of the MYC gene associated with PHACTR3 and E2F4 mutation in NSCLC is considered as a potential biomarker of NSCLC and its specific subtypes[50]. This information suggesting that targeting MYC for COVID-19 or CoV mediated respiratory disease has a potential role in the SARS-CoV-2 pathogenesis. SMAD3 is a transcriptional regulator and plays as a mediator of the cellular signals initiated by the transforming growth factor-beta (TGF-β) superfamily of cytokines, which control proliferation, differentiation, and apoptosis. It is a complex regulator in adipose physiology and the pathogenesis of chronic obstructive pulmonary disease, diabetes, inflammatory disease, and obesity[51,52]. In addition, TGF-β/SMAD3-induced repression of target genes, it required repression of the MYC gene. Moreover, SMAD3 deficiency induces inflammatory abnormal bulge on aortic in angiotensin II-infused animal models through the activation of nitric oxide synthases[53]. This information suggesting that under-expressed SMAD3 in CoV mediated respiratory disease has a prospective role in the pathogenesis of SARS-CoV-2. CBL is a proto-oncogene, encoding one of the ubiquitin ligase protein CBL involving in cell signaling, and causes a mutation in the lung, colon, endometrial adenocarcinoma, cutaneous, and breast cancer[54]. In addition, CBL is a potential gatekeeper of immune activation through its function as a non-redundant negative regulator of immune activation. Further[55], reported that targeting CBL and lipid rafts have the potential to block Kaposi's sarcoma-associated herpesvirus infection of endothelial cells. Moreover, CBL was identified as modulators of immature dendritic cell (DC) activation. The deficiency of CBL in DC upregulates toll-like receptors and inducing pro-inflammatory cytokines and chemokines[56]. Then, reduced expression of CBL in CD4+T cells is reported in several autoimmune diseases namely asthma, lupus erythematodes, multiple sclerosis, and type 1 diabetes[57]. This information is proposing that overexpressed CBL in CoV infected subjects have a prospective role in the host immunological system.

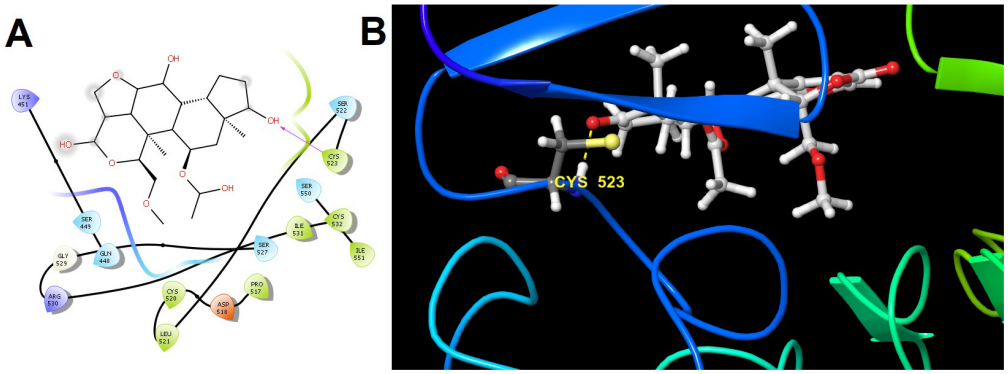


Figure 8. Molecular interactions of Wortmannin within the catalytic pocket residues of receptor ERBB4. (**A**) Two dimensional (hydrogen bond formation between Cys523 residue of ERBB4 and OH atom of Wortmannin showed in magenta line) and (**B**) three-dimensional interaction (hydrogen bond formation between Cys523 residue of protein-ligand complex showed in yellow color dotted line).

HDAC9 belongs to class II histone deacetylases, and plays a crucial role in regulating adipocyte and myocyte differentiation, and cardiac muscle development. It generally applies its function in the nucleus, and thus its activity is inhibited due to cytoplasmic retention[58]. In addition, HDAC9 transcription is observed significantly high in CD4+ T-cells from lupus subjects compared with healthy subjects[59]. NCOA3 is a transcriptional coactivator that directly binds different nuclear receptors and stimulates hormone-dependent transcriptional activities. In addition, it plays a central role in the remodeling of chromatin, histone acetyltransferase activity, and NF-kB pathway[60,61]. It is effectively used as a biomarker in breast and hepatocellular cancer[62,63]. VEGFA is a signaling protein produced by different types of cells includes macrophages, tumor cells, platelets, keratinocytes, and renal mesangial cells, which play an active role in endothelial cell functions including the formation of bones, and the angiogenesis and hematopoiesis processes. Moreover, the VEGF-A isoform enhanced the entry of the virus into the host cells through the activation of this Akt pathway[64]. Furthermore, SARS coronavirus can diffuse alveolar damage with changeable degrees of acute edema and hyaline membranes, organization, fibrosis, macrophagic infiltration, multinuclear giant cells, atypical reactive pneumocytes, and vascular injury[65,66]. This literature information suggested that accepting the relationship between coronavirus mediated respiratory disease and VEGF signaling biomarkers will cover the way to design targeted and effective therapeutic approaches for emerging COVID-19. BCL3 is a proto-oncogene that resides in the nucleus and plays a key role in the regulation of inflammatory response through the transcription of genes dependent on the NF-κB. In the context of respiratory syncytial virus (RSV) infection, the BCL-3 changes the status of histone acetylation and transcription factors on activated inflammatory (chemokines) promoters[67]. In another study, the BCL-2 gene was significantly up-regulated in PBMC samples of SARS patients (5 fold) than compared to healthy individuals[68]. BCL-3 prevents the inflammation in injured lungs through pinning-down the emergency granulopoiesis process[69]. The SARS-CoV open reading frame 7a (SARS-CoV-7a) protein participated in various pathogenesis processes including interaction with host protein and inhibiting their synthesis, promoting apoptosis, and activation of p38MAPK in the host system[70]. Further, the overexpression of BcL-xL protein in transected Jurkat T-cells significantly block the induction of apoptosis by SARS-CoV-7a protein[71].

For the biological process, we determined that cellular response to growth factor stimulus wound healing was the key function of the PPI network. The top five functional terms were RNA polymerase II promoter in response to hypoxia, peptidyl-tyrosine auto-phosphorylation, ERBB2 signaling pathway, anaphase-promoting complex dependent catabolic process, interleukin-1 mediated signaling pathway of the modules. For molecular function, growth factor receptor binding is the major term. Growth factor receptor (GFR) mostly belongs to tyrosine kinase receptors or serine-threonine kinases. The cytoplasmic domain of GFR acts as an enzyme or binds to another protein and forms a complex that acts as an enzyme. In addition, binding of GFR leads to phosphorylation of tyrosine residue and transmitting cell signals within the cell[72]. After, viral infection, the ligand (viral proteins) binds to the GFR through cell signaling, phosphorylation, and rearrangement and activates different downstream signaling that enhances cell survival, proliferation, angiogenesis, and endocytosis[73]. Moreover, a recent report demonstrated that GFR as a potential target for drug repurposing against emerging viral diseases such as COVID-19[74,75].

Moreover, we also expected to determine which immunological terms linked with coronavirus pathogenesis; namely stimulatory c-type lectin receptor (CLR) signaling pathway, regulation of hematopoietic stem cell differentiation, T cell receptor signaling, innate immune response signaling, and Fc-epsilon receptor signaling pathway, lymphocyte co-stimulation were highly enriched immunological terms of three different modules. The carbohydrate-recognition domains of CLRs are important pattern recognition receptors to identify viral pathogens and induce antiviral innate immune responses and T helper differentiation[76]. Further, reports suggested that the activation of CLRs and RIG-I-like receptors enhances pro-inflammatory response in MERS coronavirus-infected macrophages[77].

For KEGG pathways, we identified 12 important pathways that might have an essential role in the pathogenesis and human immune system. Eight different signal transduction pathways include nuclear receptor, MAPK, and WNT signaling pathway, receptor tyrosine kinase, intracellular signaling by second messenger, TGF-β, Notch and Hedgehog signaling pathway; four key cytokine immune signaling pathways namely signaling of interleukins, TNFR2-NF-kB pathway, FLT3, and interferon signaling pathways were identified using the hub-genes. Many studies reported that MAPK, WNT, receptor tyrosine kinase, intracellular signaling by the second messenger, TGF-β, Notch, and Hedgehog signaling pathway regulated in viral pathogenesis[78,79]. As we stated above, overexpression of HDAC9, NCOA3, KLHL12, CBL, and CRKL and under expression of MYC, CEBPB, VEGFA, BCL3, SMAD3, SMURF1, and ERBB4 may be involved in the regulation of pathways in signal transduction and cytokine immune signaling pathways. It might act as a novel drug target of coronavirus infected patients and plays a key role to improve host immune system.

Many biological functions, like cell signaling, cellular metabolism, enzyme inhibition and antibody-antigen recognition, have essential molecular interaction whether receptor-receptor or receptor-drug interaction. These molecular interactions frequently lead to form protein-protein or protein-drug complexes to carry out molecular functions[80]. Unfortunately, there are insufficient gene expression profiles of SARS-CoV-2 infected patients in the GEO database. Thus, in the present study, we have validated that drug targets identified from SARS-CoV-1 infected patients can have the possible interaction with SRBD, 3cLpro and N-protein of SARS-CoV-2 to define possible drug targets. Understanding the binding mode and affinity between PTs encoded host proteins and target proteins of SARS-CoV-2 requires tertiary structure of these proteins.

However, it is generally difficult and expensive to obtain complex structures by nuclear magnetic resonance spectroscopy and X-ray crystallography. Thus, protein-protein docking is an important method for understanding interactions of drug targets with SRBD, 3cLpro and N-protein of SARS-CoV-2. In the present study, the protein-protein docking results demonstrated that drug-targets with three different protein of SARS-CoV-2 showed the highest interaction energies Hydrogen bond forming residues between the protein complexes. Among them, the highly under-expressed gene ERBB4 (Table 3) encoding drug target "ERBB4" demonstrated high docking interaction energies -260.46, -314.57 and -291.72 kcal/mol with SRBD, 3cLpro and N-protein of SARS-CoV-2 respectively, which was selected for further candidate drug selection studies.

Precision based therapeutic countermeasures for COVID-19 patients

In the present study, the highly downregulated gene ERBB4 was selected as a drug target to control SARS-CoV-2 infection. Figure 9 illustrates the hypothesized mechanism action of downregulated ERBB4 in SARS-CoV-2 infected cells. A key receptor tyrosine kinase ERBB4 contains four domains includes multiple furin-like cysteine-rich domains, a tyrosine kinase domain, a phosphatidylinositol-3 kinase binding site, and a PDZ binding domain. Ligands are binding to the ERBB4 which induces different cellular response including alveolarization, alveolar epithelial type 2 cells (AEC2) differentiation, anchoring of cell surface receptor, and morphogenesis via extracellular-signal-regulated kinase pathway[81,82]. Downregulation of ERBB4 inhibits regular action namely alveolarization, morphogenesis, and AEC2 differentiation. Many proteolytic events allow for the release of ERBB4 cytoplasmic fragments. The downregulation of ERBB4 by the SARS-CoV-2 in ACE2 cells may induce receptor tyrosine kinase-mediated macropinocytosis for the host cell entry. Besides, activation of macropinocytosis can increase non-specific fluid uptake, the pH in the vesicle is not decreased and, in that scenario, SARS-CoV-2 may be recycled to the cell surface. Interestingly, fluid-phase uptake (macropinocytosis and cell-to-cell spread) was reported in murine coronavirus and SARS-CoV infection[83–85]. Both macropinocytosis and PI3K signaling are directly associated with the progression of the number of tumor cells including pancreatic and lung cancer[86,87]. Further, growing evidence suggested that PI3K signaling involves

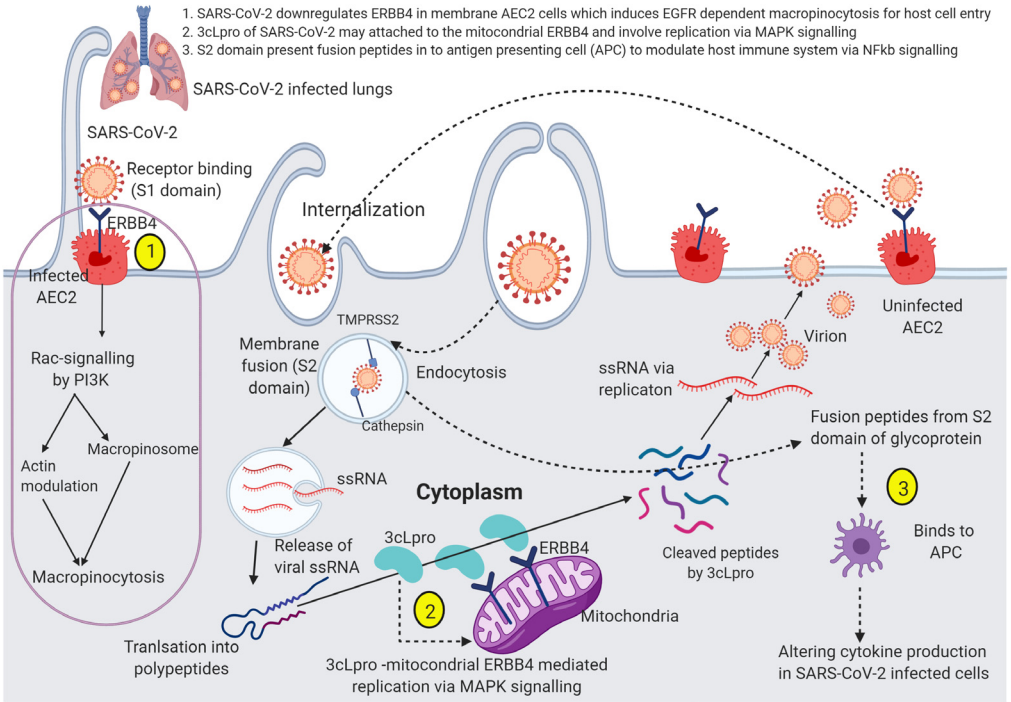


Figure 9. Hypothesized mechanism of action of downregulated ERBB4 in SARS-CoV-2 infected cells. There are three possible mechanisms (1) SARS-CoV-2 an enveloped single-strand RNA virus, virions are internalized probably by receptor tyrosine kinase (EGFR) family receptor ERBB4 dependent macropinocytosis like SARS-CoV-1 (the process of macropinocytosis indicated in magenta circle). The cellular entry of SARS-CoV-2 involves binding to ERBB4 expressing AEC2 in the plasma membrane, macropinocytosis, and cathepsin based cleavage of the spike glycoproteins. The host transmembrane serine protease 2 (TMPRSS2) for spike priming and cathepsin cleaves spike glycoprotein S1 domain, permits fusion with endosomal membranes and release virions genome into the host cytoplasm; (2) Single-strand RNA translated polypeptides cleaved by the 3cLpro of SARS-CoV-2 possibly attached with mitochondrial ERBB4 and involve viral replication through MAPK signaling pathway. Then, the cell to cell spread action, the replicated SARS-CoV-2 attached to the uninfected AEC2 receptor and then restart macropinocytosis; (3) The part of S2 domain presents peptides into antigen-presenting cells (APC) to modulate host immune system through NFkB- signaling.

the generation of the second messenger lipids and activates Rac-Pak signaling which plays a critical function namely macropinocytosis, mTOR signaling clathrin-mediated endocytosis, and cytoskeletal rearrangements[88,89]. A recent systematic review by 16 and his colleagues showed the increased level of biomarkers such as C-reactive protein, serum amyloid A, interleukin-6, lactate dehydrogenase, D-dimer, cardiac troponin and renal biomarkers were higher in severe complicated COVID-19 patients' plasma and infected lung tissues. Another study reported that low levels of type I and III interferons reduced innate antiviral defenses, which act as driving features of COVID-19[90]. In addition, proteome analysis of SARS-CoV-2 infected hosts demonstrated that proteins involved in translation, splicing, carbon metabolism, proteostasis and nucleic acid metabolism could act as a drug target for SARS-CoV-2[91]. Moreover, Gordon *et al*. reported 67 druggable targets and repurposed 69 FDA-approved drugs for COVID-19[92]. This information supports the technical concept of the present study which enhances drug repurposing strategy and advances on precision-based therapeutics to COVID-19.

In the present study, we repurposed Wortmannin as a candidate drug that can inhibit PI3K signaling and receptor tyrosine kinase (ERBB4) mediated macropinocytosis, to control virus-cell proliferation and may enhance ERBB4 functions namely alveolarization, morphogenesis, and AEC2 differentiation. Wortmannin is a covalent inhibitor of phosphatidylinositol-3-kinase (PI3K) and related enzymes include the mammalian target of rapamycin (mTOR), DNA-dependent protein kinase catalytic subunit (DNA-PKcs), and mitogen-activated protein kinase (MAPK)[93,94]. In humans, PI3K activates cell signaling pathways in the head and neck, urinary tract, cervical, ovarian, and lung cancer. Studies reported that the inhibition of PI3K signaling enhances tumor suppression and anti-tumor activity[95]. Besides, Wortmannin is an extensively used cell biology research includes to inhibit cell proliferation, DNA repair, and receptor-mediated endocytosis[96,97]. This report supported that the repurposed Wortmannin may have the potential to control SARS-CoV-2 infection in COVID-19 patients with downregulated expression of receptor tyrosine kinase ERBB4.

Conclusion

Biomarkers can be used as part of a personalized medicine paradigm to customize treatment to the specific disease characteristics of an individual patient. It can also be used to better understand disease mechanisms and to identify novel disease targets. In the present study, the highly downregulated gene ERBB4 was selected as a drug target and Wortmannin was repurposed as candidate drug to control SARS-CoV-2 infection. Further, this study provides insights on the possible personalized therapeutics for emerging COVID-19.

Data availability

Underlying data

Zenodo: Datasets for SARS-CoV-2 drug target and candidate drug identification, http://doi.org/10.5281/zenodo.4458252[98]

This project contains the following underlying data:

- List of selected differentially expressed coronavirus infected host genes for protein-protein interaction network construction
- List of protein-protein interaction subnetworks
- List of Walktrap modules and potential drug targets
- Potential Drug Targets and SARS-CoV-2 drug targets docking results
- List of predicted candidate drugs for proposed drug target

Acknowledgments

The authors are grateful to the Concordia International, GPU-accelerated Graham server of Compute Canada for high-performance computer facilities during this study.

References

1. Miłek, J, Blicharz-Domańska K: **Coronaviruses in Avian Species - Review with Focus on Epidemiology and Diagnosis in Wild Birds.** *J Vet Res.* 2018; **62**(3): 249–255.
PubMed Abstract | Publisher Full Text | Free Full Text

2. Paraskevis D, Kostaki EG, Magiorkinis G, *et al.*: **Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event.** *Infect Genet Evol.* 2020; **79**: 104212.
PubMed Abstract | Publisher Full Text | Free Full Text

3. Xu RH, He JF, Evans MR, *et al.*: **Epidemiologic clues to SARS origin in China.** *Emerg Infect Dis.* 2004; **10**(6): 1030–1037.
PubMed Abstract | Publisher Full Text | Free Full Text

4. Zaki AM, Van Boheemen S, Bestebroer TM, *et al.*: **Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia.** *N Engl J Med.* 2012; **367**(19): 1814–1820.
PubMed Abstract | Publisher Full Text

5. Hui DS, Azhar EI, Madani TA, *et al.*: **The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health — The latest 2019 novel coronavirus outbreak in Wuhan, China.** *Int J Infect Dis.* 2020; **91**: 264–266.
PubMed Abstract | Publisher Full Text | Free Full Text

6. Zhu N, Zhang D, Wang W, *et al.*: **A Novel Coronavirus from Patients with Pneumonia in China, 2019.** *N Engl J Med.* 2020; **382**(8): 727–733.
PubMed Abstract | Publisher Full Text | Free Full Text

7. Chen Y, Liu Q, Guo D: **Emerging coronaviruses: Genome structure, replication, and pathogenesis.** *J Med Virol.* 2020; **92**(4): 418–423.
PubMed Abstract | Publisher Full Text | Free Full Text

8. Huang C, Wang Y, Li X, *et al.*: **Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.** *Lancet.* 2020; **395**(10223): 497–506.
PubMed Abstract | Publisher Full Text | Free Full Text

9. Kim D, Lee JY, Yang JS, *et al.*: **The Architecture of SARS-CoV-2 Transcriptome.** *Cell.* 2020; **181**(4): 914–921.e10.
PubMed Abstract | Publisher Full Text | Free Full Text

10. Tarca AL, Romero R, Draghici S: **Analysis of microarray experiments of gene expression profiling.** *Am J Obstet Gynecol.* 2006; **195**(2): 373–388.
PubMed Abstract | Publisher Full Text | Free Full Text

11. Selvaraj G, Selvaraj C, Wei DQ: **Computational Advances in Chronic Diseases Diagnostics and Therapy - II.** *Curr Drug Targets.* 2020; **21**(2): 103–104.
PubMed Abstract | Publisher Full Text

12. Selvaraj G, Kaliamurthi S, Lin S, *et al.*: **Prognostic Impact of Tissue Inhibitor of Metalloproteinase-1 in Non- Small Cell Lung Cancer: Systematic Review and Meta-Analysis.** *Curr Med Chem.* 2020; **26**(42): 7694–7713.
PubMed Abstract | Publisher Full Text

13. Kaliamurthi S, Selvaraj G, Junaid M, *et al.*: **Cancer Immunoinformatics: A Promising Era in the Development of Peptide Vaccines for Human Papillomavirus-induced Cervical Cancer.** *Curr Pharm Des.* 2019; **24**(32): 3791–3817.
PubMed Abstract | Publisher Full Text

14. Taubenberger JK, Morens DM: **The pathology of influenza virus infections.** *Annu Rev Pathol.* 2008; **3**: 499–522.
PubMed Abstract | Publisher Full Text | Free Full Text

15. Ghobadi MZ, Mozhgani SH, Farzanehpour M, *et al.*: **Identifying novel biomarkers of the pediatric influenza infection by weighted co-expression network analysis.** *Virol J.* 2019; **16**(1): 124.
PubMed Abstract | Publisher Full Text | Free Full Text

16. Kermali M, Khalsa RK, Pillai K, *et al.*: **The role of biomarkers in diagnosis of COVID-19 - A systematic review.** *Life Sci.* 2020; **254**: 117788.
PubMed Abstract | Publisher Full Text | Free Full Text

17. Wei DQ, Selvaraj G, Kaushik AC: **Computational Perspective on the Current State of the Methods and New Challenges in Cancer Drug Discovery.** *Curr Pharm Des.* 2019; **24**(32): 3725–3726.
PubMed Abstract | Publisher Full Text

18. Selvaraj G, Kaliamurthi S, Kaushik AC, *et al.*: **Identification of target gene and prognostic evaluation for lung adenocarcinoma using gene expression meta-analysis, network analysis and neural network algorithms.** *J Biomed Inform.* 2018; **86**: 120–134.
PubMed Abstract | Publisher Full Text

19. Edgar R, Domrachev M, Lash AE: **Gene Expression Omnibus: NCBI gene expression and hybridization array data repository.** *Nucleic Acids Res.* 2002; **30**(1): 207–10.
PubMed Abstract | Publisher Full Text | Free Full Text

20. Mitchell HD, Eisfeld AJ, Sims AC, *et al.*: **A network integration approach to predict conserved regulators related to pathogenicity of influenza and SARS-CoV respiratory viruses.** *PLoS One.* 2013; **8**(7): e69374.
PubMed Abstract | Publisher Full Text | Free Full Text

21. Josset L, Menachery VD, Gralinski LE, *et al.*: **Cell host response to infection with novel human coronavirus EMC predicts potential antivirals and important differences with SARS coronavirus.** *mBio.* 2013; **4**(3): e00165–13.
PubMed Abstract | Publisher Full Text | Free Full Text

22. Sims AC, Tilton SC, Menachery VD, *et al.*: **Release of severe acute respiratory syndrome coronavirus nuclear import block enhances host transcription in human lung cells.** *J Virol.* 2013; **87**(7): 3885–3902.
PubMed Abstract | Publisher Full Text | Free Full Text

23. Selinger C, Tisoncik-Go J, Menachery VD, *et al.*: **Cytokine systems approach demonstrates differences in innate and pro-inflammatory host responses between genetically distinct MERS-CoV isolates.** *BMC Genomics.* 2014; **15**(1): 1161.
PubMed Abstract | Publisher Full Text | Free Full Text

24. Menachery VD, Mitchell HD, Cockrell AS, *et al.*: **MERS-CoV Accessory ORFs Play Key Role for Infection and Pathogenesis.** *mBio.* 2017; **8**(4): e00665–17.
PubMed Abstract | Publisher Full Text | Free Full Text

25. Xia J, Gill EE, Hancock REW: **NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data.** *Nat Protoc.* 2015; **10**(6): 823–844.
PubMed Abstract | Publisher Full Text

26. Konishi S: **Normalizing and variance stabilizing transformations for intraclass correlations.** *Ann Inst Stat Math.* 1985; **37**: 87–94.
Publisher Full Text

27. Hansen KD, Irizarry RA, Wu Z: **Removing technical variability in RNA-seq data using conditional quantile normalization.** *Biostatistics.* 2012; **13**(2): 204–216.
PubMed Abstract | Publisher Full Text | Free Full Text

28. Smyth GK: **limma: Linear Models for Microarray Data.** In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor.* Springer-Verlag, 2005; 397–420.
Publisher Full Text

29. Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.** *J R Stat Soc Ser B.* 1995; **57**: 289–300.
Publisher Full Text

30. Szklarczyk D, Gable AL, Lyon D, *et al.*: **STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.** *Nucleic Acids Res.* 2019; **47**(D1): D607–D613.
PubMed Abstract | Publisher Full Text | Free Full Text

31. Bozhilova LV, Whitmore AV, Wray J, *et al.*: **Measuring rank robustness in scored protein interaction networks.** *BMC Bioinformatics.* 2019; **20**(1): 446.
PubMed Abstract | Publisher Full Text | Free Full Text

32. Pons P, Latapy M: **Computing communities in large networks using random walks.** In *Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer, Berlin, Heidelberg, 2005; **3733 LNCS**: 284–293.
Publisher Full Text

33. Bindea G, Mlecnik B, Hackl H, *et al.*: **ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks.** *Bioinformatics.* 2009; **25**(8): 1091–1093.
PubMed Abstract | Publisher Full Text | Free Full Text

34. UniProt Consortium: **UniProt: a worldwide hub of protein knowledge.** *Nucleic Acids Res.* 2019; **47**(D1): D506–D515.
PubMed Abstract | Publisher Full Text | Free Full Text

35. Yan Y, Wen Z, Wang X, *et al.*: **Addressing recent docking challenges: A hybrid strategy to integrate template-based and free protein-protein docking.** *Proteins.* 2017; **85**(3): 497–512.
PubMed Abstract | Publisher Full Text

36. Yan Y, Tao H, He J, *et al.*: **The HDOCK server for integrated protein-protein docking.** *Nat Protoc.* 2020; **15**(5): 1829–1852.
PubMed Abstract | Publisher Full Text

37. Huang SY, Zou X: **An iterative knowledge-based scoring function for protein-protein recognition.** *Proteins.* 2008; **72**(2): 557–579.
PubMed Abstract | Publisher Full Text

38. Jo S, Kim T, Iyer VG, *et al.*: **CHARMM-GUI: A web-based graphical user interface for CHARMM.** *J Comput Chem.* 2008; **29**(11): 1859–1865.
PubMed Abstract | Publisher Full Text

39. Lee J, Cheng X, Swails JM, *et al.*: **CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field.** *J Chem Theory Comput.* 2016; **12**(1): 405–413.
PubMed Abstract | Publisher Full Text | Free Full Text

40. Phillips JC, Braun R, Wang W, *et al.*: **Scalable molecular dynamics with NAMD.** *J Comput Chem.* 2005; **26**(16): 1781–1802.
PubMed Abstract | Publisher Full Text | Free Full Text

41. Darden T, York D, Pedersen L: **Particle mesh Ewald: An *N*·log(*N*) method for Ewald sums in large systems.** *J Chem Phys.* 1993; **98**(12): 10089–10092.
Publisher Full Text

42. Wang Z, He E, Sani K, *et al.*: **Drug Gene Budger (DGB): An application for ranking drugs to modulate a specific gene based on transcriptomic signatures.** *Bioinformatics.* 2019; **35**(7): 1247–1248.
PubMed Abstract | Publisher Full Text | Free Full Text

43. Sun Y, Clark EA: **Expression of the *c-myc* proto-oncogene is essential for HIV-1 infection in activated T cells.** *J Exp Med.* 1999; **189**(9): 1391–1397.
PubMed Abstract | Publisher Full Text | Free Full Text

44. Higgs MR, Lerat H, Pawlotsky JM: **Hepatitis C virus-induced activation of β-catenin promotes c-Myc expression and a cascade of pro-carcinogenetic events.** *Oncogene.* 2013; **32**(39): 4683–4693.
PubMed Abstract | Publisher Full Text

45. Wang HF, He HX: **Regulation of Yamanaka factors during H5N1 virus infection in A549 cells and HEK293T cells.** *Biotechnol Biotechnol Equip.* 2018; **32**(6): 1–10.
Publisher Full Text

46. Price AM, Messinger JE, Luftig MA: **c-Myc Represses Transcription of Epstein-Barr Virus Latent Membrane Protein 1 Early after Primary B Cell Infection.** *J Virol.* 2018; **92**(2): e01178-17.
PubMed Abstract | Publisher Full Text | Free Full Text

47. Javier RT, Butel JS: **The history of tumor virology.** *Cancer Res.* 2008; **68**(19): 7693–7706.
PubMed Abstract | Publisher Full Text | Free Full Text

48. Rolls AES, Giovannoni G, Constantinescu CS, *et al.*: **Multiple Sclerosis, Lymphoma and Nasopharyngeal Carcinoma: The Central Role of Epstein-Barr Virus?** *Eur Neurol.* 2010; **63**(1): 29–35.
PubMed Abstract | Publisher Full Text

49. Moheimani F, Koops J, Williams T, *et al.*: **Influenza A virus infection dysregulates the expression of microRNA-22 and its targets; CD147 and HDAC4, in epithelium of asthmatics.** *Respir Res.* 2018; **19**(1): 145.
PubMed Abstract | Publisher Full Text | Free Full Text

50. Dragoj M, Bankovic J, Podolski-Renic A, *et al.*: **Association of overexpressed *MYC* gene with altered *PHACTR3* and *E2F4* genes contributes to non-small cell lung carcinoma pathogenesis.** *J Med Biochem.* 2019; **38**(2): 188–195.
PubMed Abstract | Publisher Full Text | Free Full Text

51. Frederick JP, Liberati NT, Waddell DS, *et al.*: **Transforming Growth Factor beta-Mediated Transcriptional Repression of c-*myc* Is Dependent on Direct Binding of Smad3 to a Novel Repressive Smad Binding Element.** *Mol Cell Biol.* 2004; **24**(6): 2546–2559.
PubMed Abstract | Publisher Full Text | Free Full Text

52. Yang T, Ying B, Song X, *et al.*: **Single-nucleotide polymorphisms in SMAD3 are associated with chronic obstructive pulmonary disease.** *Exp Biol Med (Maywood).* 2010; **235**(5): 599–605.
PubMed Abstract | Publisher Full Text

53. Tan CK, Tan EH, Luo B, *et al.*: **SMAD3 deficiency promotes inflammatory aortic aneurysms in angiotensin II-infused mice via activation of iNOS.**

J Am Heart Assoc. 2013; **2**(3): e000269.
PubMed Abstract | Publisher Full Text | Free Full Text

54. AACR Project GENIE Consortium: **AACR project genie: Powering precision medicine through an international consortium.** *Cancer Discov.* 2017; **7**(8): 818–831.
PubMed Abstract | Publisher Full Text | Free Full Text

55. Chakraborty S, Veettil MV, Chandran B: **Kaposi's sarcoma associated herpesvirus entry into target cells.** *Front Microbiol.* 2012; **3**: 6.
PubMed Abstract | Publisher Full Text | Free Full Text

56. Chiou SH, Shahi P, Wagner RT, *et al.*: **The E3 ligase c-Cbl regulates dendritic cell activation.** *EMBO Rep.* 2011; **12**(9): 971–979.
PubMed Abstract | Publisher Full Text | Free Full Text

57. Lutz-Nicoladoni C, Wolf D, Sopper S: **Modulation of immune cell functions by the E3 ligase CBL-b.** *Front Oncol.* 2015; **5**: 58.
PubMed Abstract | Publisher Full Text | Free Full Text

58. McKinsey TA, Zhang CL, Olson EN: **Identification of a Signal-Responsive Nuclear Export Sequence in Class II Histone Deacetylases.** *Mol Cell Biol.* 2001; **21**(18): 6312–6321.
PubMed Abstract | Publisher Full Text | Free Full Text

59. Yan K, Cao Q, Reilly CM, *et al.*: **Histone deacetylase 9 deficiency protects against effector T cell-mediated systemic autoimmunity.** *J Biol Chem.* 2011; **286**(33): 28833–28843.
PubMed Abstract | Publisher Full Text | Free Full Text

60. Xu J, Wu RC, O'Malley BW: **Normal and cancer-related functions of the p160 steroid receptor co-activator (SRC) family.** *Nat Rev Cancer.* 2009; **9**(9): 615–630.
PubMed Abstract | Publisher Full Text | Free Full Text

61. Chen H, Lin RJ, Xie W, *et al.*: **Regulation of hormone-induced histone hyperacetylation and gene activation via acetylation of an acetylase.** *Cell.* 1999; **98**(5): 675–686.
PubMed Abstract | Publisher Full Text

62. de Amicis F, Chiodo C, Morelli C, *et al.*: **AIB1 sequestration by androgen receptor inhibits estrogen-dependent cyclin D1 expression in breast cancer cells.** *BMC Cancer.* 2019; **19**(1): 1038.
PubMed Abstract | Publisher Full Text | Free Full Text

63. Ma L, Liu W, Xu A, *et al.*: **Activator of thyroid and retinoid receptor increases sorafenib resistance in hepatocellular carcinoma by facilitating the Warburg effect.** *Cancer Sci.* 2020; **111**(6): 2028–2040.
PubMed Abstract | Publisher Full Text | Free Full Text

64. Hiley CT, Chard LS, Gangeswaran R, *et al.*: **Vascular Endothelial Growth Factor A Promotes Vaccinia Virus Entry into Host Cells via Activation of the Akt Pathway.** *J Virol.* 2013; **87**(5): 2781–2790.
PubMed Abstract | Publisher Full Text | Free Full Text

65. Gu J, Korteweg C: **Pathology and pathogenesis of severe acute respiratory syndrome.** *Am J Pathol.* 2007; **170**(4): 1136–1147.
PubMed Abstract | Publisher Full Text | Free Full Text

66. Alkharsah KR: **VEGF upregulation in viral infections and its possible therapeutic implications.** *Int J Mol Sci.* 2018; **19**(6): 1642.
PubMed Abstract | Publisher Full Text | Free Full Text

67. Jamaluddin M, Choudhary S, Wang S, *et al.*: **Respiratory Syncytial Virus-Inducible BCL-3 Expression Antagonizes the STAT/IRF and NF-kappaB Signaling Pathways by Inducing Histone Deacetylase 1 Recruitment to the Interleukin-8 Promoter.** *J Virol.* 2005; **79**(24): 15302–15313.
PubMed Abstract | Publisher Full Text | Free Full Text

68. Reghunathan R, Jayapal M, Hsu LY, *et al.*: **Expression profile of immune response genes in patients with severe acute respiratory syndrome.** *BMC Immunol.* 2005; **6**: 2.
PubMed Abstract | Publisher Full Text | Free Full Text

69. Kreisel D, Sugimoto S, Tietjens J, *et al.*: **Bcl3 prevents acute inflammatory lung injury in mice by restraining emergency granulopoiesis.** *J Clin Invest.* 2011; **121**(1): 265–276.
PubMed Abstract | Publisher Full Text | Free Full Text

70. Vasilenko N, Moshynskyy I, Zakhartchouk A: **SARS coronavirus protein 7a interacts with human Ap4A-hydrolase.** *Virol J.* 2010; **7**: 31.
PubMed Abstract | Publisher Full Text | Free Full Text

71. Yang Y, Xiong Z, Zhang S, *et al.*: **Bcl-xL inhibits T-cell apoptosis induced by expression of SARS coronavirus E protein in the absence of growth factors.** *Biochem J.* 2005; **392**(Pt 1): 135–143.
PubMed Abstract | Publisher Full Text | Free Full Text

72. **Cell Biology.** 3rd Edition. (accessed on Sep 23, 2020).
Reference Source

73. Zheng K, Kitazato K, Wang Y: **Viruses exploit the function of epidermal growth factor receptor.** *Rev Med Virol.* 2014; **24**(4): 274–286.
PubMed Abstract | Publisher Full Text

74. Hondermarck H, Bartlett NW, Nurcombe V: **The role of growth factor receptors in viral infections: An opportunity for drug repurposing against emerging viral diseases such as COVID-19?** *FASEB Bioadv.* 2020; **2**(5): 296–303.
PubMed Abstract | Publisher Full Text | Free Full Text

75. Klann K, Bojkova D, Tascher G, *et al.*: **Growth Factor Receptor Signaling Inhibition Prevents SARS-CoV-2 Replication.** *Mol Cell.* 2020; **80**(1): 164–174.
PubMed Abstract | Publisher Full Text | Free Full Text

76. Bermejo-Jambrina M, Eder J, Helgers LC, *et al.*: **C-type lectin receptors in antiviral immunity and viral escape.** *Front Immunol.* 2018; **9**: 590.
PubMed Abstract | Publisher Full Text | Free Full Text

77. Zhao X, Chu H, Wong BHY, *et al.*: **Activation of C-Type Lectin Receptor and (RIG)-I-Like Receptors Contributes to Proinflammatory Response in Middle East Respiratory Syndrome Coronavirus-Infected Macrophages.** *J Infect Dis.* 2020; **221**(4): 647–659.
PubMed Abstract | Publisher Full Text | Free Full Text

78. Alto NM, Orth K: **Subversion of cell signaling by pathogens.** *Cold Spring Harb Perspect Biol.* 2012; **4**(9): a006114.
PubMed Abstract | Publisher Full Text | Free Full Text

79. Hu,L, Lin X, Lu H, *et al.*: **An overview of hedgehog signaling in fibrosis.** *Mol Pharmacol.* 2015; **87**(2): 174–182.
PubMed Abstract | Publisher Full Text

80. Vakser IA: **Protein-protein docking: From interaction to interactome.** *Biophy J.* 2014; **107**(8): 1785–1793.
PubMed Abstract | Publisher Full Text | Free Full Text

81. Sibilia M, Kroismayr R, Lichtenberger BM, *et al.*: **The epidermal growth factor receptor: from development to tumorigenesis.** *Differentiation.* 2007; **75**(9): 770–787.
PubMed Abstract | Publisher Full Text

82. Miettinen PJ, Berger JE, Meneses J, *et al.*: **Epithelial immaturity and multiorgan failure in mice lacking epidermal growth factor receptor.** *Nature.* 1995; **376**(6538): 337–341.
PubMed Abstract | Publisher Full Text

83. Freeman MC, Peek CT, Becker MM, *et al.*: **Coronaviruses induce entry-independent, continuous macropinocytosis.** *mBio.* 2014; **5**(4): e01340–01314.
PubMed Abstract | Publisher Full Text | Free Full Text

84. Kerr MC, Teasdale RD: **Defining macropinocytosis.** *Traffic.* 2009; **10**(4): 364–371.
PubMed Abstract | Publisher Full Text

85. Mercer J, Helenius A: **Virus entry by macropinocytosis.** *Nature cell biol.* 2009; **11**(5): 510–520.
PubMed Abstract | Publisher Full Text

86. Baer R, Cintas C, Dufresne M, *et al.*: **Pancreatic cell plasticity and cancer initiation induced by oncogenic Kras is completely dependent on wild-type PI 3-kinase p110α.** *Genes Dev.* 2014; **28**(23): 2621–2635.
PubMed Abstract | Publisher Full Text | Free Full Text

87. Hodakoski C, Hopkins BD, Zhang G, *et al.*: **Rac-mediated macropinocytosis of extracellular protein promotes glucose independence in non-small cell lung cancer.** *Cancers (Basel).* 2019; **11**(1): 37.
PubMed Abstract | Publisher Full Text | Free Full Text

88. Hawkins PT, Stephens LR: **Emerging evidence of signalling roles for PI(3,4)P_2 in class I and II PI3K-regulated pathways.** *Biochem Soc Trans.* 2016; **44**(1): 307–314.
PubMed Abstract | Publisher Full Text

89. Campa CC, Ciraolo E, Ghigo A, *et al.*: **Crossroads of PI3K and Rac pathways.** *Small GTPases.* 2015; **6**(2): 71–80.
PubMed Abstract | Publisher Full Text | Free Full Text

90. Blanco-Melo D, Nilsson-Payant BE, Liu WC, *et al.*: **Imbalanced host response to SARS-CoV-2 drives development of COVID-19.** *Cell.* 2020; **181**(5): 1036–1045.e9.
PubMed Abstract | Publisher Full Text | Free Full Text

91. Bojkova D, Klann K, Koch B, *et al.*: **Proteomics of SARS-CoV-2-infected host cells reveals therapy targets.** *Nature.* 2020; **583**(7816): 469–472.
PubMed Abstract | Publisher Full Text

92. Gordon DE, Jang GM, Bouhaddou M, *et al.*: **A SARS-CoV-2 protein interaction map reveals targets for drug repurposing.** *Nature.* 2020; **583**(7816): 459–468.
PubMed Abstract | Publisher Full Text | Free Full Text

93. Abliz A, Deng W, Sun R, *et al.*: **Wortmannin, PI3K/Akt signaling pathway inhibitor, attenuates thyroid injury associated with severe acute pancreatitis in rats.** *Int J Clin Exp Pathol.* 2015; **8**(11): 13821–33.
PubMed Abstract | Free Full Text

94. Gomes AM, Pinto TS, da Costa Fernandes CJ, *et al.*: **Wortmannin targeting phosphatidylinositol 3-kinase suppresses angiogenic factors in shear-stressed endothelial cells.** *J Cell Physiol.* 2020; **235**(6): 5256–5269.
PubMed Abstract | Publisher Full Text

95. Ihle NT, Williams R, Chow S, *et al.*: **Molecular pharmacology and antitumor activity of PX-866, a novel inhibitor of phosphoinositide-3-kinase signaling.** *Mol Cancer Ther.* 2004; **3**(7): 763–772.
PubMed Abstract

96. Liu Y, Shreder KR, Gai W, *et al.*: **Wortmannin, a widely used phosphoinositide 3-kinase inhibitor, also potently inhibits mammalian polo-like kinase.** *Chem Biol.* 2005; **12**(1): 99–107.
PubMed Abstract | Publisher Full Text

97. Kim SH, Jang YW, Hwang P, *et al.*: **The reno-protective effect of a phosphoinositide 3-kinase inhibitor wortmannin on streptozotocin-induced proteinuric renal disease rats.** *Exp Mol Med.* 2012; **44**(1): 45–51.
PubMed Abstract | Publisher Full Text | Free Full Text

98. Selvaraj G, Kaliamurthi S, Wei D, *et al.*: **Datasets for SARS-CoV-2 drug target and candidate drug identification [Data set].** *Zenodo.* 2021.
http://www.doi.org/10.5281/zenodo.4458252

Open Peer Review

Current Peer Review Status: ?

Version 1

Reviewer Report 22 March 2021

https://doi.org/10.5256/f1000research.53939.r79888

?

Sugunadevi Sakkiah
Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, Jefferson, AR, USA

The authors put a great effect to identify the potent target and candidate drugs for COVID-19 applying biological networks and structural modeling approaches. The authors should address the below comments before the indexing of this paper:

Minor revision:

1. In the introduction, Ref. 16, the authors should mention the authors' name.

2. In methods, data collection, the authors should be clear whether they selected all CoV infected samples or only CoV2 infected samples?

3. In Figure 1, SARS-CoV means 1 and 2 or only 2?

4. Any particular reason for the authors setting the 303K in molecular dynamics simulations?

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Cheminformatics, Endocrine disrupting chemicals

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 22 Mar 2021

Gurudeeban Selvaraj, Concordia University, Montreal, Canada

Comment 1: In the introduction, Ref. 16, the authors should mention the authors' name.
Author Response 1: Thanks for your suggestion. The author name is included in the text as follows----A recent systematic review by 16 Kermali and his colleagues showed the increased level of biomarkers such as C-reactive protein, serum amyloid A, interleukin-6, lactate dehydrogenase, D-dimer, cardiac troponin, and renal biomarkers (urea and creatinine) were higher, and low level of lymphocytes and platelet count were recorded in severe complicated COVID-19 patients than non-severe COVID-19 patients' plasma and infected lung tissues.

Comment 2: In methods, data collection, the authors should be clear whether they selected all CoV infected samples or only CoV2 infected samples?
Author Response 2: Thanks for the question. All the selected studies contain control and CoV-infected human lung epithelial cells samples. In earlier days the inadequate information of the CoV-2 genomic datasets in Public databases (i.e. NCBI), the authors have selected CoV samples (i.e. CoV-1 genomic information is more than 96% similar to CoV-2).

Comment 3: In Figure 1, SARS-CoV means 1 and 2 or only 2?
Author Response 3: Thanks for the key question. The initial drug targets were identified from the SARS-CoV-1 whole-genome datasets. Then, the targets verified and validated with RBD-spike receptor binding domain; 3cLpro-3C-like protease; N pro-nucleocapsid protein of SARS-CoV-2 using protein-protein docking. The protein-protein docking results confirm molecular interaction of the identified drug targets with proteins of SARS-CoV-2. Thus, the first box of Figure 1, represents the SARS-CoV or SARS-CoV-1.

Comment 4: Any particular reason for the authors setting the 303K in molecular dynamics simulations?
Author Response 4: Thanks for the key question. Yes, most computational and experimental studies have focused on room temperature (298K to 303K). The MD simulation input file preparation using the Charmm-gui web-interface was set at 303K as a room temperature (Jo et al., 2008 and Lee et al., 2016). Thus, the authors followed the same.

Competing Interests: No