5th International Conference on Electrical Information and Communication Technology (EICT), 17-19 December 2021, Khulna, Bangladesh

Genetic Effects of Covid 19 on the Development of Neurodegenerative Diseases

Nitun Kumar Podder
Dept. of Computer Science & Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
E-mail: nituncse@gmail.com

Pintu Chandra Shill
Dept. of Computer Science & Engineering
Khulna University of Engineering & Technology
Khulna, Bangladesh
E-mail: pintu@cse.kuet.ac.bd

Humayan Kabir Rana
Dept. of Computer Science & Engineering
Green University of Bangladesh
Dhaka, Bangladesh
E-mail: humayan.pustcse@gmail.com

Shudeb Babu Sen Omit
Dept. of Computer Science & Engineering
Z. H. Sikder University of Science and Technology
Shariatpur, Bangladesh
E-mail: shudebbabu0111@student.nstu.edu.bd

Md. Mehedee Hasan Al Shahriar
Dept. of Computer Science & Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
E-mail: mhsnayon@gmail.com

Md. Shafiul Azam
Dept. of Computer Science & Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
E-mail: shahincseru@gmail.com

2021 5th International Conference on Electrical Information and Communication Technology (EICT) | 978-1-6654-0906-3/21/$31.00 ©2021 IEEE | DOI: 10.1109/EICT54103.2021.9733698

***Abstract*— The coronavirus disease (COVID-19) caused by SARS-COV-2, a highly infectious pathogen, genetically similar to SARS-COV is an unprecedented worldwide health crisis. Rapidly accumulating clinical research revealed that COVID-19 is manifested by various neurological symptoms and also affects the brain in many ways including direct infection to systemic inflammation. Which indicates it may substantially increase the incidence of developing neurodegenerative diseases (NDGDs). To discourse this issue we studied the computable frameworks to address the gene expression association of COVID-19 and NDGDs to identify the link among them. We analyzed GEO microarray datasets from COVID-19 and NDGDs including Epilepsy, Stroke, Multiple Sclerosis, Alzheimer's disease, and Parkinson's disease. We constructed disease-gene relationship networks and identified dysregulated pathways, ontological pathways, protein-protein interaction (PPI) network, and protein-drug interaction (PDI) network. We observed that COVID-19 associated genes share 19, 26, 20, 19, 22 differentially expressed genes with Epilepsy, Stroke, Multiple Sclerosis, Alzheimer's disease, and Parkinson's disease respectively. Gene expression dysregulation, PPI and PDI relationship networks, different pathways suggest that COVID-19 may have a significant link to the development of these NDGDs. This analysis may help to develop therapeutic strategies and raise awareness about the influence of COVID-19 on the progression of NDGDs.**

Keywords—Covid-19, Comorbidities, Neurodegenerative diseases (NDGDs), Gene enrichment, Diseasome network, PPIs Network, PDIs Network.

I. INTRODUCTION

The COVID-19 is disrupting the health systems and threatening the lives, health of people, and present the world with a medical challenge. Though COVID-19 normally a respiratory disease that triggers both the upper (sinus, nose, and throat) and lower (wide pipe, lungs) respiratory tract, emerging evidence indicates that it has a various impact on several organs and systems of the human body. Intriguing evidence shows that it plays a significant role in the progression of neurodegenerative diseases like epilepsy, stroke, multiple sclerosis, Alzheimer's disease, Parkinson's disease. Our study aims to find the relation of COVID-19 in the progression of neurodegenerative diseases. COVID-19 affects the brain in many ways including direct infection of neural cells with SARS-COV-2, damaging nervous cells by severe systemic inflammation floods with pro-inflammatory agents, initiate stroke, brain ischemia by respiratory failure. COVID-19 has a tendency for the blood to thicken or sticky which blocks blood vessels supply to the brain and makes clotting which generates stroke. In various studies, it shows that Stroke has been a manifestation of COVID-19, and around 0.9% to 23% of COVID-19 patients developed stroke. Epilepsy is a family of CNS disorders in which the brain activity becomes abnormal in its actions and causes seizures. COVID-19 affects CNS systems in a way that leads epilepsy to generate. Clinical evidence shows the effect of COVID-19 on nervous cells and my line seath become one of a reason to develop Multiple Sclerosis. Extensive CNS inflammation in response to SARS-COV-2 infection leads aggressive inflammatory response called cytokine storm which imbalance the nervous system in consequence generates many neurological complications. The actions and reactions of the body towards COVID-19 initiate or enlarge the evidence of Alzheimer's and Parkinson's disease as well as its complications [10].

We developed a systemic and quantitative network-based structure to deal with COVID-19 and its influence factors on developing neurodegenerative diseases (NDGDs). We have analyzed commonly DEGs of COVID-19 with Stroke, Alzheimer's disease, Epilepsy, Parkinson's disease and Multiple sclerosis, from rns-Seq, and microarray datasets. We developed two distinct diseasome network for commonly upregulated and downregulated genes. We performed signaling and ontological pathway analysis, protein-protein interaction (PPI) analysis, and protein-drug interaction (PDI) analysis. We validate our investigation through two gold benchmark charts dbGaP- the database of genotypes and phenotypes, and OMIM- online Mendelian

inheritance in man. Our study provides a shred of evidence for the absorbency of COVID-19 on disease NDGDs progression [2].

II. METHODOLOGY

A. Working Principle and Datasets

We have analyzed Gene Expression Omnibus (GEO) datasets to categorize the relationship between COVID-19 and corresponding neurodegenerative diseases. These datasets are collected from National Center for Biotechnology Information (NCBI) [1]. Datasets information are given below table1.

S. No.	Disease Name	Accession Number	Sample (Case and Control)
1	COVID-19	GSE166552	06 (03 + 03)
2	Epilepsy	GSE22779	16
3	Stroke	GSE58294	92
4	Multiple Sclerosis	GSE38010	07
5	Alzheimer's Disease	GSE28146	29
6	Parkinson's Disease	GSE19587	22

Table 1: Datasets information used for expression profiling.

We represented an analytical standard approach to achieve the genetic profiling of COVID-19, Epilepsy, Alzheimer's disease, Stroke, Parkinson's disease and Multiple sclerosis at the molecular level. The pictorial view of this quantifiable approach is shown in figure 1 as a flowchart.

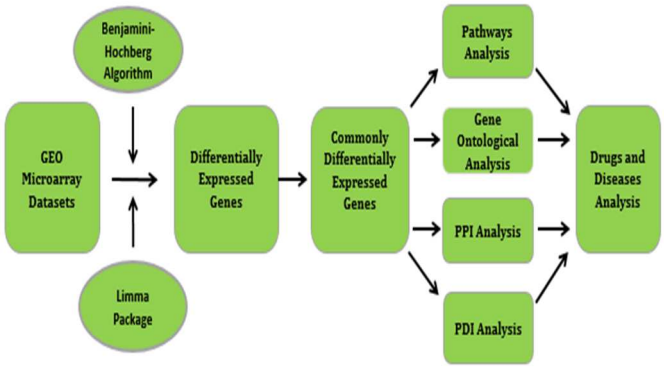

Fig. 1: Analytical approach of this investigation.

This approach used an R package named Limma to identify differentially expressed genes (DEGs) and find common, differentially expressed genes. Those were used to construct the gene-disease networks. We performed pathway analysis, ontological analysis, protein-protein interaction analysis, and protein-diseases interaction analysis. Lastly, validation was achieved by two gold benchmark databases named dbGaP, OMIM.

B. Methods of Analysis

GEO analysis is a unique way for analyzing different types of biological datasets. These are used to determine human cases with respect to control sets. For gene expression analysis we used the Benjamini-Hochberg algorithm [8]. Which is:
Benjamini-Hochberg = (i/m) Q

Here,
i = individual p value's rank
m = total number of tests
Q = false discovering rate

We have used neighborhood-based benchmarking with multilayer topological methods to find the association between the diseases and common DEGs. We have implemented the diseasome network (GDN) by using gene-disease association. In GDN every node of the network represented either gene or disease. We used an enrichment analysis tool (Enrichr) to find the gene ontology and pathways of several disorders. We used the Markov clustering algorithm to analyze Protein-Protein interaction (PPI) and protein-diseases interaction (PDA). We formed a PPI network, and a PDI network through the STRING database. We have used NetworkAnalyst, and Cytoscape for these constructions. Finally, we validated the results with gold benchmark databases (dbGaP and OMIM) [1].

III. RESULT ANALYSIS

A. Genetic relationship analysis

For identifying genetic relationship among diseases, we have analyzed DEGs for COVID-19, Epilepsy, Alzheimer's disease, Stroke, Parkinson's disease and Multiple sclerosis with statistical p-value <= 0.01 to reject the null hypothesis of our investigation and | logFC | is greater than or equal to 1 for upregulated identification and | logFC | is less than or equal to -1 for down-regulated identification [3].

We have identified an adequate number of DEGs for each disease as well as identified common DEGs between COVID-19 and each neurodegenerative disease too. We have got 19 (6 up-regulated and 13 down-regulated), 26 (8 up-regulated and 18 down-regulated), 20 (6 up-regulated and 14 down-regulated), 19 (14 up-regulated and 5 down-regulated), and 22 (5 up-regulated and 17 down-regulated) common dysregulated genes for Epilepsy, Stroke, Multiple Sclerosis, Alzheimer's disease, and Parkinson's diseases respectively with COVID-19.

We have formed two different diseasome networks for shared up and down regulated genes between COVID-19 and neurodegenerative diseases. Here, diseases nodes are the source node and common DEGs are the target node. There must be a link between the source node and target node (Jaccard Coefficient procedure) [2]. These diseasome networks are presented in figure 2 and 3.

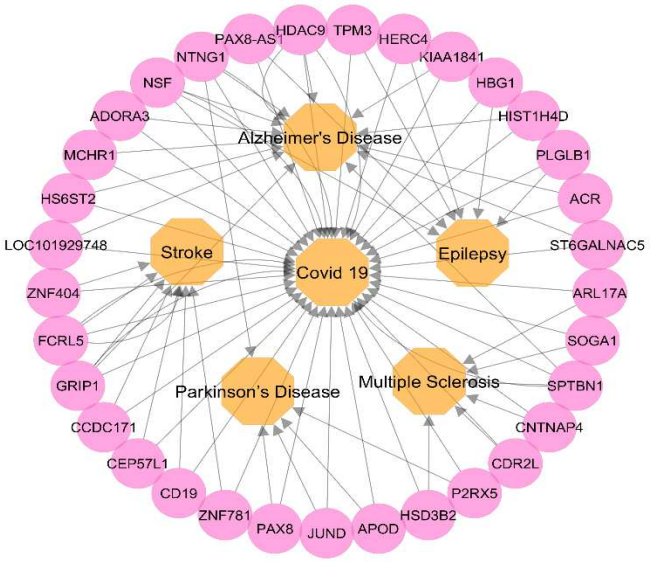

Fig. 2: Up-regulated diseasome network of common DEGs. Octagons represent the diseases and Circles represent the common DEGs. The arrow signs are represented from target nodes to source nodes.

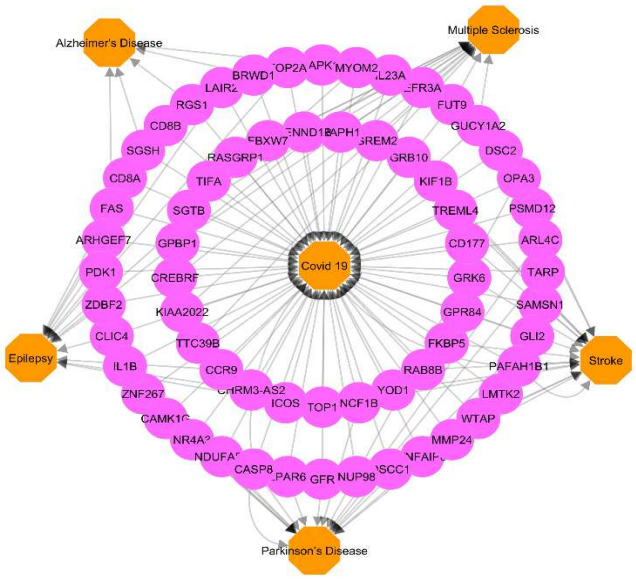

Fig. 3: Down-regulated diseasome network of common DEGs. Octagons represent the diseases and circles represent the common DEGs. The arrow signs are represented from target nodes to source nodes.

B. Functional association analysis (Pathways identifiaction)

Pathway is the interactions among molecules to conduct frequent changes in a cell [3]. It is a biology term used for representing a well-characterized segment of molecular physiological machinery. Pathway analysis identifies the specific protein functions, biological pathways profiling, physical interactions are enriched in a particular group. It helps researchers gain mechanistic insight into gene lists generated from genome-scale, making it a widely used tool in biological research. Here we have used an enrichment analysis tool (Enrichr) to analyze the pathway of dysregulated genes. WikiPathways, BioCarta, Reactome, KEGG are the four databases we have used for pathway analysis [4]. We have used common DEGs between COVID-19 and neurodegenerative diseases for pathway identification. We have identified a large number of disease pathways and performed some statistical and bioinformatics operation to identify the most significant 24 pathways between COVID-19 and neurodegenerative diseases as epilepsy, stroke, multiple sclerosis, Alzheimer's disease, Parkinson's disease to show the association between those diseases [10]. Significant pathways, gene symbols, and adjusted p values are enlisted in table 2.

Significant Pathways	Gene Symbols	Adjusted P-value
Purinergic signaling WP4900	P2RX5;LPAR6; ADORA3	0.044175991
Nucleotide GPCRs WP80	LPAR6;ADORA3	0.053224139
MAPK Signaling Pathway WP382	IL1B;FAS;RASGRP1; MAPK12;FGFR1	0.075740025
Apoptosis Modulation by HSP70 WP384	CASP8;FAS	0.071720967
Unfolded protein response WP4925	CASP8;IL1B	0.075740025
Regulation of toll-like receptor signaling pathway WP1449	CASP8;IL1B;TIFA;MAPK12	0.071720967
FAS signaling pathway (CD95)	CASP8;FAS	0.09796921
Signal Transduction Homo sapiens R-HSA-162582	PSMD12;FBXW7;GPR84;RASGRP1;HDAC9;MCHR1;MAPK12;GLI2;CASP8;RGS1	0.041257503
CLEC7A/inflammasome pathway Homo sapiens R-HSA-5660668	CASP8;IL1B	0.041257503
Adaptive Immune System R-HSA-1280218	HERC4;PSMD12;CD8B;FBXW7;CD8A;TREML4;CD19;LAIR2;ICOS;FGFR1	0.044708277
CASP8 activity is inhibited R-HSA-5218900	PSMD12;FBXW7;GPR84;RASGRP1;HDAC9;MAPK12;GLI2;GREM2;CASP8;RGS1	0.044708277
Regulation by c-FLIP R-HSA-3371378	CASP8;IL1B	0.044708277
Glutamate Binding, Activation of AMPA Receptors and Synaptic Plasticity R-HSA-399721	HERC4;PSMD12;CD8B;FBXW7;CD8A;TREML4;CD19;LAIR2;ICOS;RASGRP1; FGFR1	0.127144407
Dimerization of procaspase-8 R-HSA-69416	CASP8;FAS	0.044708277
RIPK1-mediated regulated necrosis R-HSA-5213460	CASP8;FAS	0.067865221

Signalling by NGF R-HSA-166520	PSMD12;CD19;ARHGEF7;RASGRP1;SPTBN1;MAPK12	0.104749017
Signaling by EGFR R-HSA-177929	PSMD12;CD19;ARHGEF7;RASGRP1;SPTBN1;FGFR1	0.118677876
NOD1/2 Signaling Pathway R-HSA-168638	CASP8;MAPK12	0.127144407
Trafficking of AMPA receptors R-HSA-399719	NSF;GRIP1	0.127144407
Signaling by Interleukins R-HSA-449147	PSMD12;IL23A;IL1B;RASGRP1;SPTBN1;FGFR1	0.127144407
TNF signaling pathway	CASP8;IL1B;FAS;MAPK12	0.038463182
IL-17 signaling pathway	CASP8;IL1B;MAPK12	0.096553553

Table 2: Functional associative significant pathways of neurodegenerative diseases (Stroke, Alzheimer's disease, Epilepsy, Parkinson's disease and Multiple sclerosis) with COVID-19 of shared DEGs.

C. Gene ontological analysis (GO)

The GO is a theoretical model that describes a gene or gene product in detail as its molecular functions, biological process in which it participates, and cellular location [5]. Gene Ontology is the largest knowledge base and most comprehensive platform for the information of genes and its functions [1]. Here we have used Enrichr, the universal gene enrichment analysis tool to analyze the GO of commonly DEGs. We have used two databases of Enrichr named GO Biological process (GO term) and Human Phenotype ontology (HP term) for analysis on commonly dysregulated genes [5]. We did some statistical operations on analyzed data then pointed out some significant gene ontology terms (20) and human phenotype term (4) for COVID-19 with NDGDs which are given below in table 3.

Term	Name	Adjusted P-value	Genes
GO:0021952	central nervous system projection neuron axonogenesis	0.098910314	PAFAH1B1;GLI2
GO:0045833	negative regulation of lipid metabolic process	0.127355778	IL1B;APOD
GO:0042327	positive regulation of phosphorylation	0.126781557	IL1B;LMTK2;FGFR1
GO:0000165	MAPK cascade	0.167559528	PSMD12;IL1B;SPTBN1;FGFR1
GO:0006265	DNA topological change	0.098910314	TOP2A;TOP1
GO:0030201	heparan sulfate proteoglycan metabolic process	0.127355778	HS6ST2;SGSH
GO:0010648	negative regulation of cell communication	0.145934924	IL1B;LMTK2;FGFR1
GO:0032456	endocytic recycling	0.082315162	ARL4C;DENND1B;LMTK2
GO:0071637	regulation of monocyte chemotactic protein-1 production	0.098910314	IL1B;APOD
GO:0009968	negative regulation of signal transduction	0.114022036	FBXW7;IL1B;GRB10;LMTK2;APOD;FGFR1
GO:0032645	regulation of granulocyte macrophage colony-stimulating factor production	0.012649396	IL23A;IL1B;RASGRP1
GO:0007204	positive regulation of cytosolic calcium ion concentration	0.127355778	P2RX5;LPAR6;CCR9;MCHR1
GO:0032729	positive regulation of interferon-gamma production	0.098910314	IL23A;IL1B;RASGRP1
GO:0045582	positive regulation of T cell differentiation	0.167559528	IL23A;GLI2
GO:0071260	cellular response to mechanical stimulus	0.104756888	CASP8;IL1B;FAS
GO:0045862	positive regulation of proteolysis	0.114022036	CASP8;IL1B;MAPK12
GO:0018205	peptidyl-lysine modification	0.114022036	TOP2A;NUP98;TOP1;HDAC9
GO:0032743	positive regulation of interleukin-2 production	0.127355778	IL1B;SPTBN1
GO:0043001	Golgi to plasma membrane protein transport	0.145934924	NSF;SPTBN1
GO:0070372	regulation of ERK1 and ERK2 cascade	0.145934924	FBXW7;IL1B;LMTK2
HP:0002843	Abnormality of T cells	0.007964771	CASP8;CD8A;FAS;ICOS
HP:0012140	Abnormality of cells of the lymphoid lineage	0.074553644	CD19;FAS;ICOS
HP:0001744	Splenomegaly	0.074553644	CASP8;HBG1;CD19;FAS;ICOS;SGSH
HP:0000528	Anophthalmia	0.102325032	GRIP1;GLI2

Table 3: Ontological associative significant pathways of neurodegenerative diseases (Epilepsy, Stroke, Multiple sclerosis, Alzheimer's disease, Parkinson's disease) with COVID-19 of shared DEGs.

D . Protien protien interaction analysis (PPI)

PPI is the major objective of system biology, shows the prediction of protein functions of molecules' target protein and drug ability [6]. It handles a wide range of biological processes as forming of macromolecular structures and enzyme complexes. It is a mathematical representation of physical contacts among proteins in the molecules. A web resource tool and gold benchmark biological database (STRING) is used for investigating the PPIs of commonly dysregulated genes of COVID-19 and neurodegenerative diseases (epilepsy, stroke, multiple sclerosis, Alzheimer's disease, Parkinson's disease) [7]. We have used a comprehensive network visual analytics platform (NetworkAnalyst) for constructing PPI network, which is shown on figure 4.

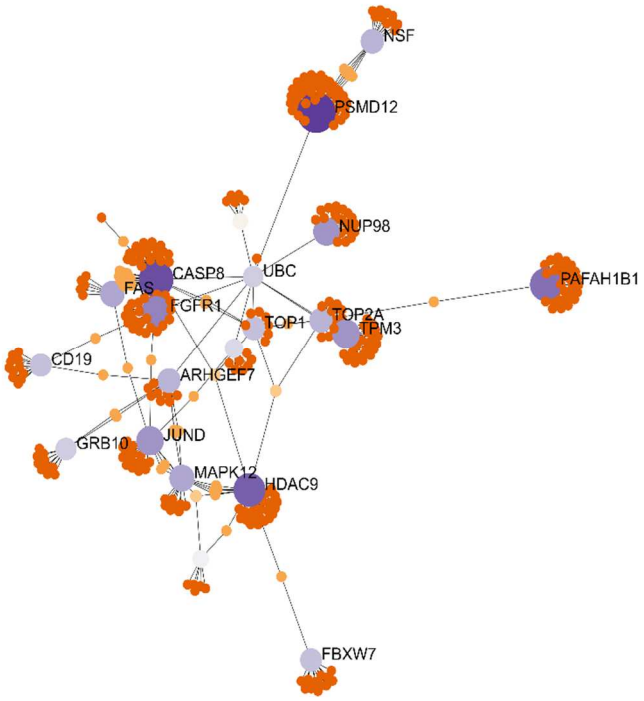

Fig. 4: PPIs network of NDGDs with Covid-19

We have used high STRING interactome confidence score (950) for identifying interactions at molecular level [8]. We have identified some hub proteins. The most critical hub proteins are PSMD12 (degree 49 and betweeness centrality 19001), PAFAH1B1 (degree 30 and betweeness centrality 9831), TPM3 (degree 22 and betweeness centrality 7203), FAS (degree 17 and betweeness centrality 2167.717), TOP2A (degree 15 and betweeness centrality 1511.62), CD19 (degree 13 and betweeness centrality 3859.95), SPTBN1 (degree 09 and betweeness centrality 4181.101), TP53 (degree 03 and betweeness centrality 5693.501), HDAC9, FBXW7, MAPK12, JUND, GRB10 etc. These hub proteins are responsible for developing maximum NDGDs [2]. These hub proteins are also represented in figure 4.

E.. Protein drug interaction analysis (PDI)

One of the essential targets of this line of inquiring about centers on pinpointing the potential drug particles [9]. With the shared DEGs of Covid 19 with NDGDs, we determined the protein-drug interactions (PDIs) network using NetworkAnalyst online tools, prepared with DrugBank. The organized information was downloaded and customized with Cytoscape. PDI network of the whole 45 hubs, circles speak to the shared DEGs, which are TOP1 and TOP2A, whereas squares show the interacting drug molecules [6]. These representations are shown in figure 5.

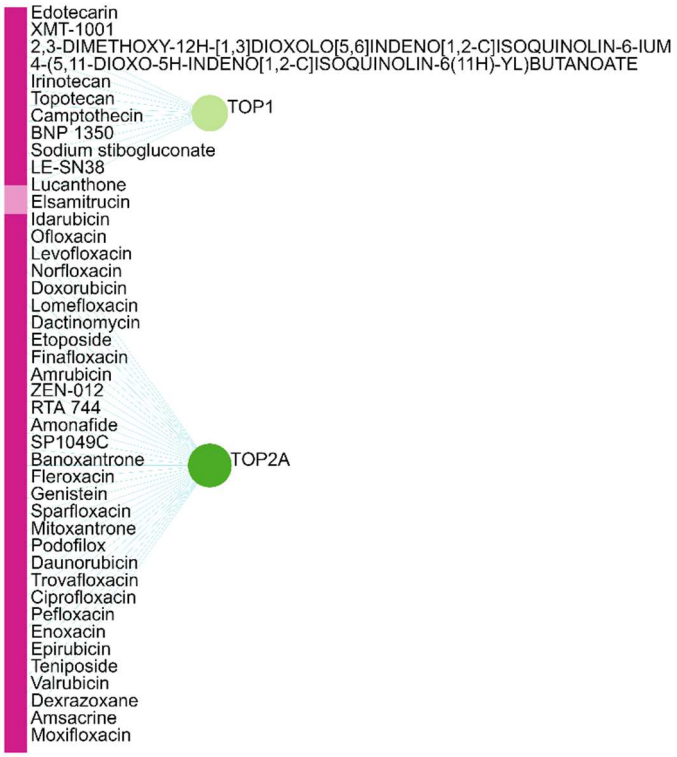

Fig 5: PDIs network of NDGDs with Covid-19

F.. Drugs and diseases analysis (Validation)

We have utilized two gold benchmark databases dbGaP and OMIM from Enrichr, a gene enrichment tool to check the validity of our investigation. We have utilized the DEGs of Covid 19 for these purposes [3]. From that gold benchmark databases, we have got some drug and disease lists with some additional information, which genes are common DEGs of Covid 19. After applying some statistical operations, and reviewing the literature, we have selected ten diseases from the diseases list and found that our selected five NDGDs (comorbidities of Covid 19) still existed on that ten diseases, which carries some evidence of validation on our investigation. So this evidence proves that Stroke, Alzheimer's disease, Epilepsy, Parkinson's disease and Multiple sclerosis are actually the comorbidities of Covid 19. The rest five diseases are also different types of NDGDs, and these five would be the comorbidities of Stroke, Alzheimer's disease, Epilepsy, Parkinson's disease

and Multiple sclerosis. We have constructed a drug diseases (validation) network as shown in figure 6 [9].

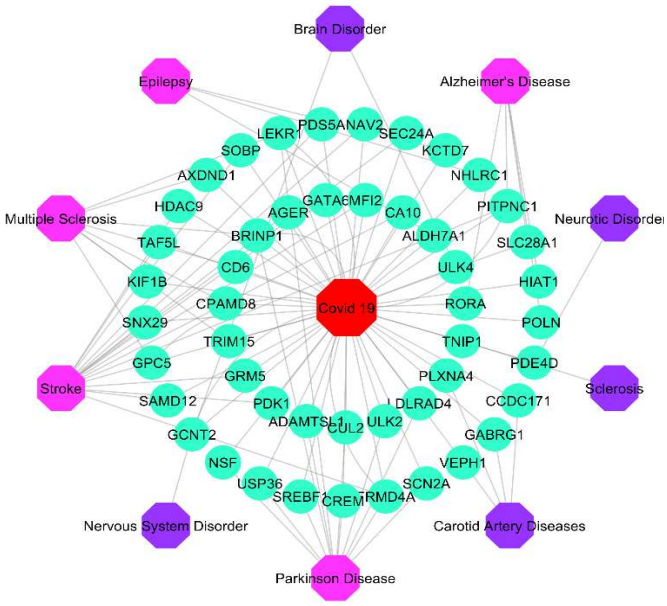

Fig. 6: Drug diseases analysis network of Covid 19. Purpled colored shape represents our selected diseases, and violet colored shape represents the comorbidities of our selected diseases. Green colored shape represents the shared genes.

IV. Discussion and Conclusion

We have investigated the gene expression relationship of COVID-19 and NDGDs (Epilepsy, Stroke, Multiple sclerosis, Alzheimer's disease, and Parkinson's disease) based on the association of common DEGs, pathways, and ontological analysis, PPIs, and PDIs. We found a significant number of commonly dysregulated genes among COVID-19 and NDGDs which indicates the genetic effects and risk factors of COVID-19 on the progression of NDGDs. Our analyzed significant pathways, ontological GO and HP terms, PPIs networks, and PDIs network also suggested the same outcome of significant relation of COVID-19 and NDGDs at the molecular and cellular level. Here we have analyzed and investigated the genetic effects of COVID-19 on the progression of NDGDs. We also validate our findings through gold benchmark databases. Our findings indicate that COVID-19 has a strong association with NDGDs on the genetic level. This type of study will be useful for making genomic evidence-based recommendations about accurate disease prediction, identification, therapeutic treatments. It will also useful for raising awareness about the threatening consequences of COVID-19 among people.

V. Acknowledgement

This research work has been financially supported in part by Green University of Bangladesh, and Khulna University of Engineering and Technology.

VI. References

[1] Podder, Nitun Kumar, et al. "A system biological approach to investigate the genetic profiling and comorbidities of type 2 diabetes." *Gene Reports* 21 (2020): 100830.

[2] Rana, Humayan Kabir, et al. "Genetic effects of welding fumes on the progression of neurodegenerative diseases." *Neurotoxicology* 71 (2019): 93-101.

[3] Rana, Humayan Kabir, et al. "Genetic effects of welding fumes on the development of respiratory system diseases." *Computers in biology and medicine* 108 (2019): 142-149.

[4] Rana, Humayan Kabir, et al. "Machine learning and bioinformatics models to identify pathways that mediate influences of welding fumes on cancer progression." *Scientific reports* 10.1 (2020): 1-15.

[5] Nain, Zulkar, et al. "Pathogenetic profiling of COVID-19 and SARS-like viruses." *Briefings in bioinformatics* 22.2 (2021): 1175-1196.

[6] Moni, Mohammad Ali, and Pietro Liò. "Network- Nain, Zulkar, et al. "Transcriptomic studies revealed pathophysiological impact of COVID-19 to predominant health conditions." *Briefings in Bioinformatics* (2021).

[7] Ferini-Strambi, Luigi, and Maria Salsone. "COVID-19 and neurological disorders: are neurodegenerative or neuroimmunological diseases more vulnerable?" *Journal of neurology* (2020): 1-11.

[8] Datta, Ratri, et al. "Bioinformatics approach to analyze gene expression profile and comorbidities of gastric cancer." *2020 23rd International Conference on Computer and Information Technology (ICCIT)*. IEEE, 2020.

[9] Moni, Mohammad Ali, et al. "COVID-19 patient transcriptomic and genomic profiling reveals comorbidity interactions with psychiatric disorders." *Translational psychiatry 11.1 (2021): 1-13*

[10] Rahman, Md Habibur, et al. "Bioinformatics and machine learning methodologies to identify the effects of central nervous system disorders on glioblastoma progression." *Briefings in Bioinformatics (2021).*