Journal of Infection and Public Health 15 (2022) 437–447

Contents lists available at ScienceDirect

Journal of Infection and Public Health

journal homepage: www.elsevier.com/locate/jiph

Original article

Bioinformatics analysis of potential therapeutic targets for COVID-19 infection in patients with carotid atherosclerosis

Liang Yanchao[a,1], Zhang Sibin[b,1], Ilgiz Gareev[g,1], Xiang Huan[a], Zhao Junfei[c], Liu Chunyang[a], Ozal Beylerli[d], Albert Sufianov[d,e], Yuan Chao[a], Gai Yuyan[c], Xu Xun[a], Aamir Ahmad[f], Liang Peng[b,*], Yang Guang[a,**]

[a] Department of Neurosurgery, First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang Province, China
[b] The Third Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang Province, China
[c] Harbin Institute of Information Technology, No.9 University Town, Binxi Economic and Technological Development Zone, Harbin, Heilongjiang, China
[d] Federal Center of Neurosurgery, Tyumen, Russia
[e] Department of Neurosurgery, Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russia
[f] Interim Translational Research Institute, Academic Health System, Hamad Medical Corporation, Doha, Qatar
[g] Bashkir State Medical University, Ufa, Republic of Bashkortostan 450008, Russia

ARTICLE INFO

Article history:
Received 31 December 2021
Received in revised form 19 February 2022
Accepted 9 March 2022

Keywords:
SARS-CoV-2
COVID-19 infection
Carotid atherosclerosis
Differentially expressed genes
Gene ontology
Protein–protein interactions
Hub gene
Drug

ABSTRACT

Background: COVID-19 is a new coronavirus that constitutes a great challenge to human health. At this stage, there are still cases of COVID-19 infection in some countries and regions, in which ischemic stroke (IS) is a risk factor for new coronavirus pneumonia, and patients with COVID-19 infection have a dramatically elevated risk of stroke. At the same time, patients with long-term IS are vulnerable to COVID-19 infection and have more severe disease, and carotid atherosclerosis is an early lesion in IS.
Methods: This study used human induced pluripotent stem cell (hiPSC)-derived monolayer brain cell dataset and human carotid atherosclerosis genome-wide dataset to analyze COVID-19 infection and carotid atherosclerosis patients to determine the synergistic effect of new coronavirus infection on carotid atherosclerosis patients, to clarify the common genes of both, and to identify common pathways and potential drugs for carotid atherosclerosis in patients with COVID-19 infection
Results: Using several advanced bioinformatics tools, we present the causes of COVID-19 infection leading to increased mortality in carotid atherosclerosis patients and the susceptibility of carotid atherosclerosis patients to COVID-19. Potential therapeutic agents for COVID-19 -infected patients with carotid atherosclerosis are also proposed.
Conclusions: With COVID-19 being a relatively new disease, associations have been proposed for its connections with several ailments and conditions, including IS and carotid atherosclerosis. More patient-based data-sets and studies are needed to fully explore and understand the relationship.

Introduction

In late 2019, COVID-19 begins to spread widely around the world, and in late 2019, the World Health Organization declared COVID-19 a serious epidemic of the 21st century [1]. As of July 20, 2021 according to WHO (https://covid19.who.int/), a total of 190,671,330 positive cases of COVID-19 have been confirmed, including 4,098,758 deaths. Of these, the top five countries with the highest number of confirmed cases were the United States, India, Brazil, Russia, and France. At this stage, the overall novel coronavirus pneumonia epidemic worldwide is stabilizing, with almost all countries and regions

* Correspondence to: The Third Affiliated Hospital of Harbin Medical University, No.150 Ha ping Road, Nan gang District, Harbin, Heilongjiang, China.
** Correspondence to: The First Affiliated Hospital of Harbin Medical University, 23 Youzheng St, Harbin 150001, Heilongjiang Province, China.
E-mail addresses: liangpeng@hrbmu.edu.cn (L. Peng), yangguang1227@163.com (Y. Guang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jiph.2022.03.004
1876-0341/

worldwide experiencing difficult periods of outbreaks of coronavirus disease (COVID-19). At this phase, there are still regions with localized outbreaks of new severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2).

The virus which causes COVID-19. SARS-CoV-2 is a family of coronaviruses [2], Patients with SARS-CoV-2 infection have predominantly respiratory symptoms such as fever, dry cough and dyspnea [3], However, the manifestations of gastrointestinal, cardiovascular, and neurological involvement have been increasingly appreciated and widely reported. A growing body of evidence suggests that SARS-CoV-2 infection can cause neurological symptoms and complications. Several medical centers in various countries, including China, have identified a large proportion of patients with SARS-CoV-2 infection who have a combination of ischemic stroke. Also, patients with ischemic stroke have a higher susceptibility to SARS-CoV-2 and have a higher mortality rate [4–6]. Therefore, it is of great importance to understand the underlying mechanisms between COVID-19 and ischemic stroke.

Stroke has been the number one cause of death and disability in the 50–74 and 75 years and older age groups worldwide [7]. Epidemiological data on stroke show that ischemic stroke accounts for 84.4% of the total number of strokes [8] where patients with carotid atherosclerotic disease have an elevated risk of embolic stroke and other major adverse cardiovascular events[9,10]. The current study shows that SARS-CoV-2 is bound to the angiotensin-converting enzyme 2 (ACE2) receptor to infect the human nervous system [11]. Therefore, cells which expressed ACE2 may be susceptible targets for SARS-CoV-2 infection, and SARSCoV-2 contains S proteins with high interaction with angiotensin-converting enzyme 2 [12]. In parallel, glial cells and neurons also express ACE2 receptors, which made them potential targets for SARS-CoV-2 [13–15] , This feedback suggests that SARS-CoV-2 may be associated with neuroaffinity15–17. Meanwhile, ACE2 receptors located in vascular endothelial cells allow viral attachment and penetration into cells [16] , Local pro-inflammatory mediators (cytokines and chemokines), coagulation cascade factors, growth factors and nitric oxide affect the reduction of the barrier integrity [18]. In addition, it was demonstrated that tumor necrosis factor(TNF-α) and interleukin 1β(IL-1β) released from endothelial cells during viral infection can trigger endothelial cells through the NF-κB pathway, ultimately inducing the expression of new adhesion molecules associated with the inflammatory response [19]. Consequently, viral infection triggers a series of cascade reactions leading to the complications of thromboembolism, and therefore, these investigations have sparked interest and research into the potential relationship between ischemic stroke and COVID-19.

In the field of biomedical research, high-throughput methods have assumed a pivotal role, and microarray data profiling is at the forefront of high-throughput methods for large-scale analysis of gene expression [20]. Microarray studies also assist the investigator in gene expression studies [21]. Relevant studies have demonstrated the outstanding performance of high-throughput sequencing analysis of SARS-CoV in assessing its data quality and gene representation [22]. There are no GeneChip data analysis for SARS-CoV-2 and carotid atherosclerosis.

This study attempted to identify the biological pathways of SARS-CoV-2 and carotid atherosclerosis and their interrelationships (Fig. 1). Two datasets were selected for analysis in this study; GSE157852 for human SARS-CoV-2 infection and GSE43292 for carotid atherosclerosis gene expression analysis. both datasets were collected from the Gene Expression Omnibus (GEO) database. First, we identified differentially expressed genes (DEGs) for GSE157852 and GSE43292, and the common differentially expressed genes were the basis of the whole study as well as the raw data. Further analysis based on common differential genes, including gene set enrichment analysis and pathway analysis, was performed to understand the biological process of genome-based expression studies. Subsequently, protein interaction networks are formed to identify HUB genes from DEGs and search for potential therapeutic agents.

Materials and methods

Data collection

The dataset (GSE157852) investigates the susceptibility of human induced pluripotent stem cell (HiPSC)-derived monolayer brain cells

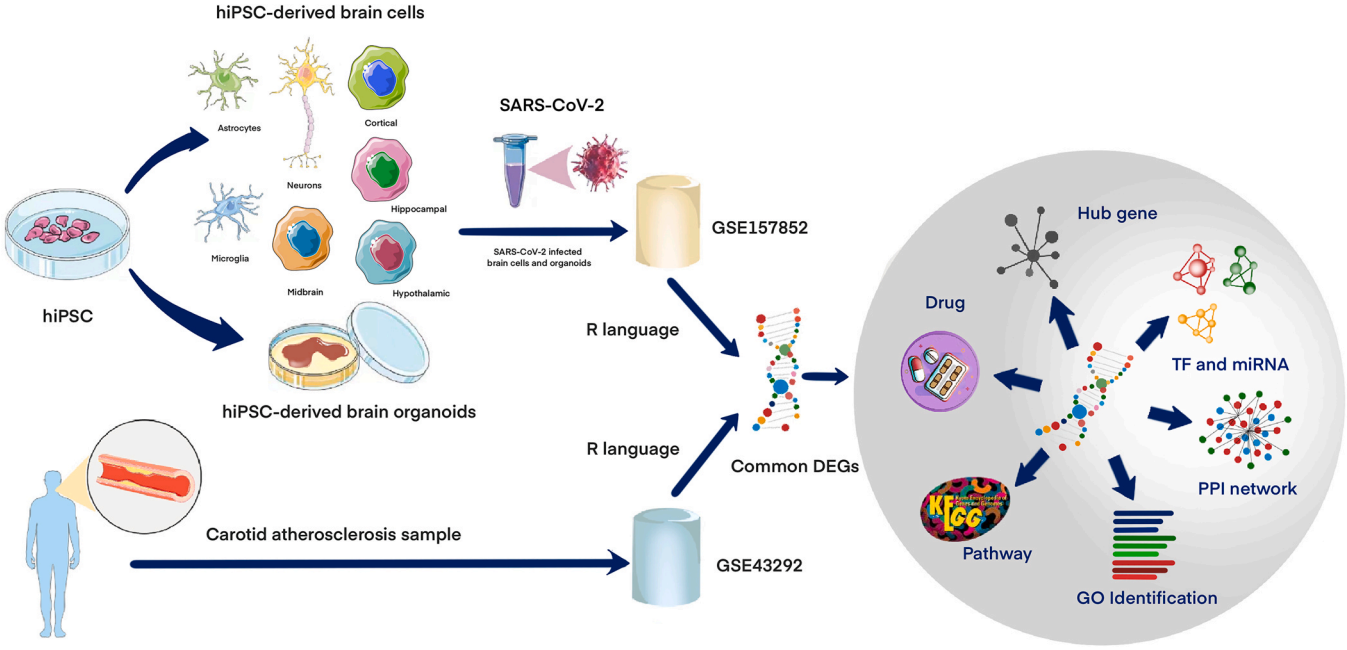

Fig. 1. Methodological workflow of the present research. Two categories of samples (carotid atherosclerosis samples, SARS-CoV-2 infected cells and organoids) were collected from SARS-CoV-2 infected hiPSC-derived brain cells and organoids both are included in the GSE157852 dataset and GSE 43292. Identifying common DEGs from two datasets using the R programming language. From the common DEGs, GO identification, KEGG pathway, PPIs network, TF and miRNA analysis, hub gene identification and module analysis was designed and based on those analysis drug molecule identification was performed.

and region-specific brain organs to SARS-CoV-2 infection, and the dataset GSE43292 investigates gene expression in carotid atherosclerosis, both from GEO database.The GSE157852 dataset has RNA sequence extraction was performed using the GPL21697 NextSeq 550 platform, and the GSE43292 dataset was performed using the GPL6244 (Affymetrix Human Gene 1.0 ST Array) platform. the GSE157852 dataset was provided by Jacob et al. [23]. Carotid atherosclerosis dataset (GSE43292) was proposed by Ayari et al. [24]. The COVID19 database (GSE157852) provides samples of human induced pluripotent stem cell (HiPSC)-derived SARS-CoV-2 infection. The carotid atherosclerosis dataset (GSE43292) contains 64 samples.

Characterization of differential genes (DEGs) and co-dominant genes in COVID-19 and carotid atherosclerosis

The DEG identification of GSE157852 and GSE43292 datasets was the principal task of this study. We analyzed the data using Linux system and Rstudio, library (DESeq2) to collate the GSE157852 dataset, and the DEG of GSE43292 dataset by dividing into two categories, Macroscopically intact tissue, Atheroma plaque. Rstudio, library(limma) was used to collate and analyze the data, and a web tool (http://bioinfogp.cnb.csic.es/tools/venny/index.html) was used to achieve universal gene identification between the DEGs of the GSE157852 and GSE43292 datasets.

Gene ontology and pathway identification in genome enrichment analysis

Gene set enrichment analysis is the functional analysis of gene sets with general biological functions and chromosomal locations [25]. The analysis results contain three main categories: biological processes, molecular functions and cellular components [26]. The main principle of gene enrichment analysis is to further understand the molecular activity, cellular role and location in the cell where the gene performs its function. kegg pathway is usually applied to understand metabolic pathways and contains more essential applications than gene annotation [27]. In the purpose of significant pathway analysis WikiPathways [28], Reactome [29] databases were also used alongside the KEGG pathway. GO terms and all the pathways were obtained through web-based platform Enrichr (https://maayanlab.cloud/Enrichr/)for the common genes that were identified in the previous step.

Analysis of PPI networks

PPI activity is considered a primary target for cell biology research and functions as a prerequisite for systems biology [30]. Proteins operate intracellularly through interactions with another protein, and information generated from the PPI network has boosted the insight into protein function [31]. The discrepant genes were imported into STRING11.0 for protein interaction network analysis, and the data were imported into Cytoscape software, and the top 10 Hub gene were identified by using the MCC algorithm in the cytohubba plugin in Cytoscape. The data were analyzed by MCODE plugin to obtain MCODE1 and MCODE2.

TF-gene interactions

Interaction of TF genes with identified common DEGs as a result of TF on functional pathways and gene expression levels was evaluated [32]. Identification of TF gene interactions with common genes via https://www.networkanalyst.ca/ platform, NetworkAnalyst is a web-based platform for performing gene comparisons, quantification, and differential gene expression analysis for numerous species [33]. The TF-gene interaction network was obtained from the ENCODE (https://www.encodeproject.org/)) database in the NetworkAnalyst platform.

TF-miRNA coregulatory network

TF-miRNA coregulatory network was obtained through the analysis of differential genes on https://www.networkanalyst.ca/ website. Identification of candidate drugs is a key component of ongoing research. The differential genes were screened and analyzed by https://maayanlab.cloud/Enrichr/ website.

Statistical analysis

The majority of statistical analyses were performed using the bioinformatics tools mentioned above, and the rest were performed using the default parameters of the R software (V4.11), such as DESeq2, clusterProfiler, ReactomePA, limma, biomaRt, etc., The identification of differentially expressed genes was evaluated by the DESeq2 package, applying the Benjamini-Hochberg FDR method to adjust the P-value.Results were considered to be statistically significant when P-value was < 0.05.

Results

Identification of DEGs and common gene between COVID-19 and Carotid atherosclerosis

The GSE157852 dataset was applied for the identification of DEG in COVID-19, taking padj < 0.05, log2FoldChange > 1 as highly expressed genes and log2FoldChange < −1 as low expressed genes. A total of 983 differentially expressed genes were obtained, of which 586 genes were up-regulated and 397 genes were down-regulated in expression. Using the GSE 43292 carotid atherosclerosis dataset, a total of 92 differentially expressed genes were identified, of which 53 genes were upregulated and 49 genes were downregulated. A total of 983 collected COVID-19 genes and 92 carotid atherosclerosis genes were compared using R language, and a total of 9 common genes were identified for expression (TM4SF18, TDO2, MMP7, MYOM1, CNN1, TTLL7, PLD5, DPP4, NPR3). The common DEGSs between the two datasets were visually compared by the Venn diagram in Fig. 2. The results of the Venn diagram showed that the common DEGSs accounted for 0.8% of the total 1075 DEGSs.

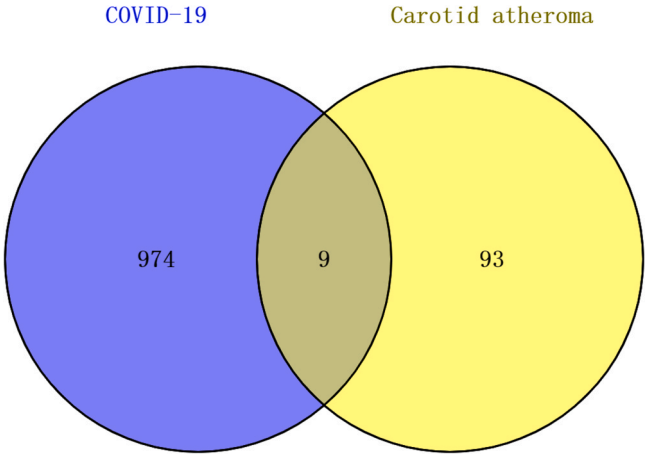

Fig. 2. Commonly differentially expressed genes were represented by Venn diagrams. 9 common genes were identified among 974 SARS-CoV-2 infection differentially expressed genes and 93 differentially expressed genes in patients with carotid atherosclerosis. Common differentially expressed genes accounted for 0.8% of the 1075 differentially expressed genes.

Table 1
GO terms, GO pathways and their corresponding P-values and genes for common differentially expressed genes.

Category	GO ID	GO pathways	P-values	Genes
Biological Process	GO:0051000	positive regulation of nitric-oxide synthase activity	0.007178	NPR3
	GO:0006084	acetyl-CoA metabolic process	0.006284	TDO2
	GO:0010715	regulation of extracellular matrix disassembly	0.006284	DPP4
	GO:0071688	striated muscle myosin thick filament assembly	0.005836	MYOM1
	GO:1904706	negative regulation of vascular smooth muscle cell proliferation	0.005388	CNN1
	GO:0031034	myosin filament assembly	0.004940	MYOM1
	GO:0010737	protein kinase A signaling	0.004492	MYOM1
	GO:0018095	protein polyglutamylation	0.004043	TTLL7
	GO:0070189	kynurenine metabolic process	0.003595	TDO2
	GO:0033632	regulation of cell-cell adhesion mediated by integrin	0.003146	DPP4
Cellular Component	GO:0030139	endocytic vesicle	0.047572	DPP4
	GO:0005925	focal adhesion	0.010471	DPP4;CNN1
	GO:0005865	striated muscle thin filament	0.008519	MYOM1
	GO:0071437	invadopodium	0.005388	DPP4
	GO:0032982	myosin filament	0.004043	MYOM1
Molecular Function	GO:0004252	serine-type endopeptidase activity	0.003288	DPP4;MMP7
	GO:0070011	peptidase activity, acting on L-amino acid peptides	0.002936	DPP4;MMP7
	GO:0004177	aminopeptidase activity	0.015199	DPP4
	GO:0051393	alpha-actinin binding	0.012977	MYOM1
	GO:0048487	beta-tubulin binding	0.012087	TTLL7
	GO:0043014	alpha-tubulin binding	0.011642	TTLL7
	GO:0016702	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen	0.009412	TDO2
	GO:0051371	muscle alpha-actinin binding	0.007178	MYOM1
	GO:0008239	dipeptidyl-peptidase activity	0.003595	DPP4
	GO:0070739	protein-glutamic acid ligase activity	0.003146	TTLL7

GO and pathway findings in gene set enrichment analysis

Enrichr web tool was employed for gene set enrichment analysis. In this research, we analyzed the GOterm and KEGGpathway of 9 common DEGs (TM4SF18, TDO2, MMP7, MYOM1, CNN1, TTLL7, PLD5, DPP4, NPR3). go term includes biological process, molecular functions We list the top 10 GO terms for each component. As the data in Table1 shows, among the biological processes, positive regulation of nitric-oxide synthase activity is at the top of Biological Process, MYOM1 gene is involved in striated muscle MYOM1 gene is involved in striated muscle myosin thick filament assembly, myosin filament assembly, and protein kinase A signaling. Molecular functional data suggest that it is mainly involved in DPP4 and MYOM1 genes, including endocytic vesicle and focal adhesion. Cellular component studies have shown that peptidase activity, acting on L-amino acid peptides, aminopeptidase activity and alpha-actinin bindin have

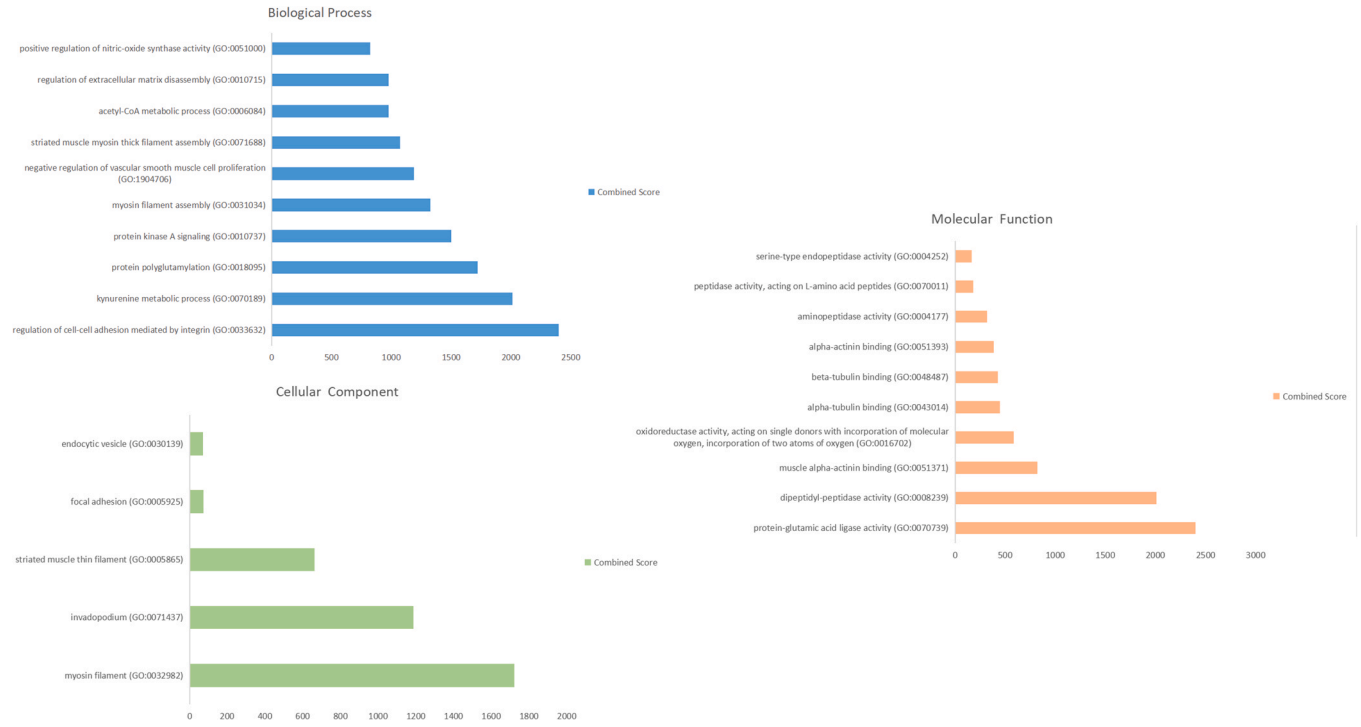

Fig. 3. (A) The results of the pathway terms were identified through the combined score. The results are identified according to the combination of scored biological processes, molecular functions and cellular components related to GO terms. The higher the enrichment score, the higher number of genes are involved in a certain ontology. (B) Identification of pathway analysis results by KEGG, WikiPathways, Reactome. The results of the pathway terms were determined by the combined score.

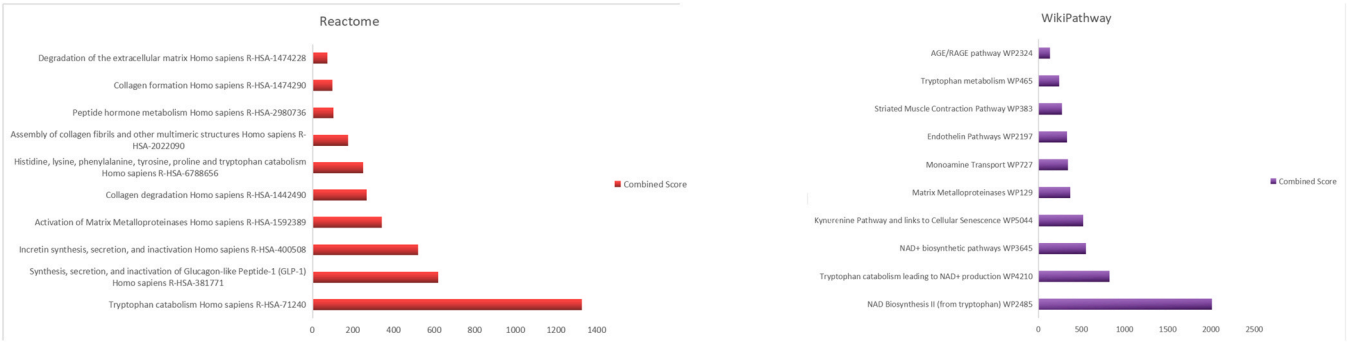

Fig. 3. (*continued*)

important roles in the disease. KEGG pathways, WikiPathways, and Reactome pathways were analyzed as shown in Fig. 3. The information obtained from Fig. 3 shows that the protein digestion and absorption pathway tryptophan metabolism pathway interacts with the most number of genes according to the KEGG pathway database.

Identification and modular analysis of hub genes

The common DEGs were provided as an input in STRING and the file generated from the analysis is reintroduced into Cytoscape software for visual representation. The PPIs network is created for further analysis of this study including hub gene detection for identifying drug molecules for COVID-19 and Carotid atherosclerosis (Fig. 4). Ultimately, the results of the PPIs network connect for suggesting drug compounds that establish the PPIs analysis as centroid of this study. The high density module was designed from PPI network using MCODE, which is also a plug-in for Cytoscape software. Then MCODE plug-in was used for analysis to obtain MCODE1, characteristic: number of nodes 37, number of edges 147. MCODE2: characteristic number of nodes 17, number of edges 74.

Identification of hub genes and module analysis for proposing therapeutic solutions

To identify the hub genes in the PPI network, Cytoshubba, a plug-in for Cytoscape software, was to be used. HUB genes were ranked according to a score, which indicates the number of gene interactions in the PPI network. Table 2 shows 10 HUB genes identified

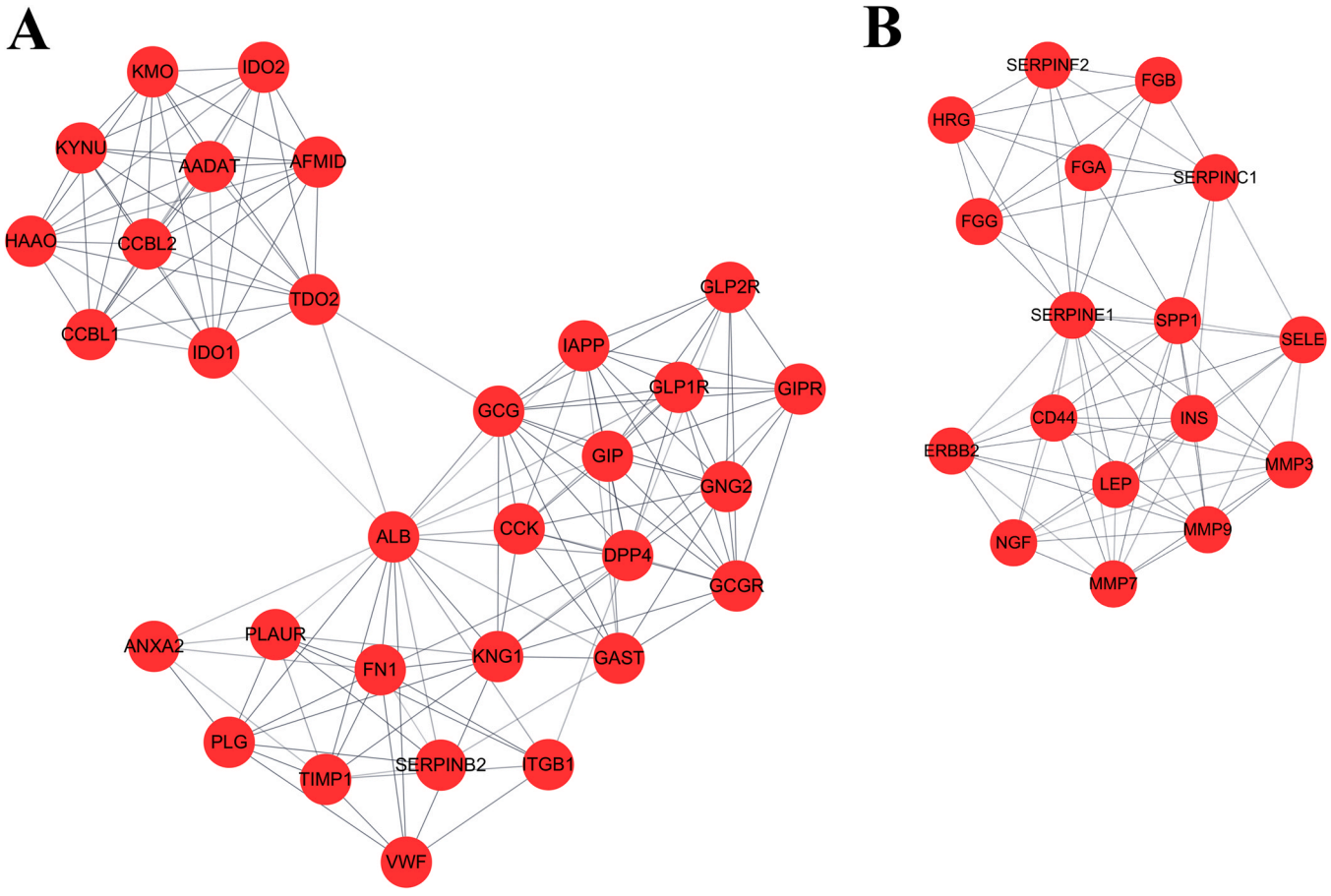

Fig. 4. Protein-protein interaction (PPI) network for the identification of differentially expressed genes common to both diseases (COVID-19 and carotid atherosclerosis). The analyzed network A holds 37 nodes and 147 edges. The analyzed network B holds 17 nodes and 74 edges.

Table 2
10 HUB genes and their scores.

Hub genes	Score
ALB	206316306.00
FN1	206259864.00
TIMP1	206256960.00
PLG	205148160.00
VWF	164757600.00
SERPINE1	164508504.00
MMP9	121326024.00
INS	113746838.00
SPP1	107825760.00
MMP3	97977600.00

were ALB, FN1, TIMP1, PLG, VWF, SERPINE1, MMP9, INS, SPP1, MMP3. Fig. 5 shows the interactions of HUB proteins with other proteins in the PPI network. number of nodes 43, number of edges 328. Network diameter 3, Network radius 2,Network density 0.363, where ALB, FN1, and TIMP1 are the three genes highlighted in the modular network, as these three genes are also common between the two datasets.

TF gene interactions

TF-gene interactions were collected using NetworkAnalyst. Differential genes (TM4SF18, TDO2, MMP7, MYOM1, CNN1, TTLL7, PLD5, DPP4, NPR3) were screened for TF gene identification(Fig. 6). dPP4 is regulated by 13 TF genes, NPR3 by 8 TF genes, MYOM1 by 16 TF genes, TDO2 by 2 TF genes, and CNN1 is regulated by 15 TF genes, and MMP7 is regulated by 3 TF genes, and these TF genes regulate over one common differential gene in the network, indicating a highdegree of interaction between TF genes and differential genes.

TF-miRNA coregulatory network

TF-miRNA co-regulatory network was generated using NetworkAnalyst. The analysis of the TF-miRNA co-regulatory network provided the interaction of miRNAs and TF with the common DEG. This interaction may be responsible for the regulation of DEGS expression. the network created by TF-miRNA co-regulatory network consists of 170 Nodes,182 Edges, consisting of 47 TF genes and 115 miRNAs interacting with differential genes. Fig. 7 shows the TF-miRNA co-regulatory network.

Identification of candidate drugs

10 common drug molecules were identified using the Enrichr platform. The results of drug candidates were generated based on P-values (Table 3). The analysis showed that ascorbic acid CTD 00005445, progesterone CTD 00006624 and cyclosporin A CTD 00007121 were the three drug molecules that interacted with most genes. Therefore these drugs represent the common drugs for neocoronary pneumonia and carotid atherosclerosis. Table 3 indicates the common drug candidates.

Discussion

Carotid atherosclerosis is recognized as a risk factor for COVID-19. When the Nervous system of a person gets disrupted and that is the time when the functionality of the system cannot adjust properly to its task. People who have carotid atherosclerosis are at a higher risk of developing COVID-19. This study assists in the narrative of the bioinformatics course used for the susceptibility of human induced pluripotent stem cell (HiPSC)-derived monolayer brain cells and region-specific brain organs to SARS-CoV-2 infection and the genome-wide dataset of human carotid atherosclerosis. Bioinformatics-related methods were used to investigate the identification of 983 and 92 DEGs from GSE 157852 and GSE 43292, respectively. To establish relationships and identify common DEGs between the GSE 157852 and GSE 43292 datasets based on COVID-19 and carotid atherosclerosis assay drug candidate. After identification, 9 common DEGs (TM4SF18, TDO2, MMP7, MYOM1, CNN1, TTLL7, PLD5, DPP4, NPR3) were identified. next studies continued with GO analysis, KEGG pathway analysis, PPI, TF-gene interactions, TF-miRNA coregulatory network, and drug candidate assays.

The 10 common digenomes identified were used to detect GO terms. GO terms were selected based on P values. Related studies confirmed that the ERK1/2-RSK-nNOS signaling pathway may play an important role in Ang II-mediated regulation of central blood pressure, particularly in terms of impaired vasoconstriction and dilation.[34,35]. Considering that SARS-CoV-2 infects the nervous system by binding to angiotensin-converting enzyme 2 (ACE2) receptors, and that most ischemic stroke patients have long-term combined hypertension and poor vasoconstriction and diastolic function, SARS-CoV-2 invades the nervous system through ACE2 by regulating nitric-oxide synthase activity and thus the adverse effects of vasoconstriction and diastolic function are exacerbated by SARS-CoV-2 invasion of the nervous system through ACE2. Whether this is a possible mechanism for the exacerbation of SARS-CoV-2 infection in patients with ischemic stroke needs to be further explored. Studies have shown that DPP4 is widely expressed in the vascular system, including endothelial cells, cardiomyocytes, smooth muscle cells, macrophages and many other types of cells, which means it may contribute to cardiovascular disease [36,37]. CoV virus entry into host epithelial cells is mediated by interactions between viral envelope S protein homologs and cell surface receptors. hydrolytic cleavage of the CoV S protein results in recognition of membrane receptors [angiotensin-converting enzyme 2 (ACE2) for SARS-Cov and SARS-Cov-2 and dipeptidyl peptidase 4 (DPP4) for MERS-Cov] by the S1 outer structural domain, while the S2 C -terminal structural domain is involved in cell fusion and viral entry [38–40].

The KEGG pathway was identified for 10 common DEGs. the KEGG pathway includes the protein digestion and uptake pathway, the tryptophan metabolism pathway.Increasingly, tryptophan and its metabolites are considered to be important aspects of the pathophysiology of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) that drove the COVID-19 pandemic [41], Targeting the various aspects of tryptophan metabolism may have major preventive and therapeutic implications for the administration of SARS-CoV-2 infection, especially since the emergence of new variants would suggest that this virus may affect for multiple years [42]. Systemic low-grade immune-mediated inflammation is associated with atherosclerosis, in which pro-inflammatory cytokines such as interferon-gamma (IFN-γ) play an important role. upregulation of indoleamine 2,3-dioxygenase (IDO) by IFN-γ decreases serum levels of tryptophan increasing levels of kynurenine metabolites. Increased IDO expression and activity can accelerate the atherosclerotic process [43]. Simultaneously, from the results of WikiPathways, the most interacted gene pathway are NAD Biosynthesis II(from tryptophan)WP2485, Tryptophan catabolism leading to NAD+ production WP4210,NAD+ biosynthetic pathways wp3645. Results from the Reactome pathway produce Tryptophan catabolism Homo sapiens R-HAS-71240 and Synthesis,secretion,and inactivation of Glucagon-like Peptide-1(GLP-1) Homo sapiens R-HAS-381771 pathway.

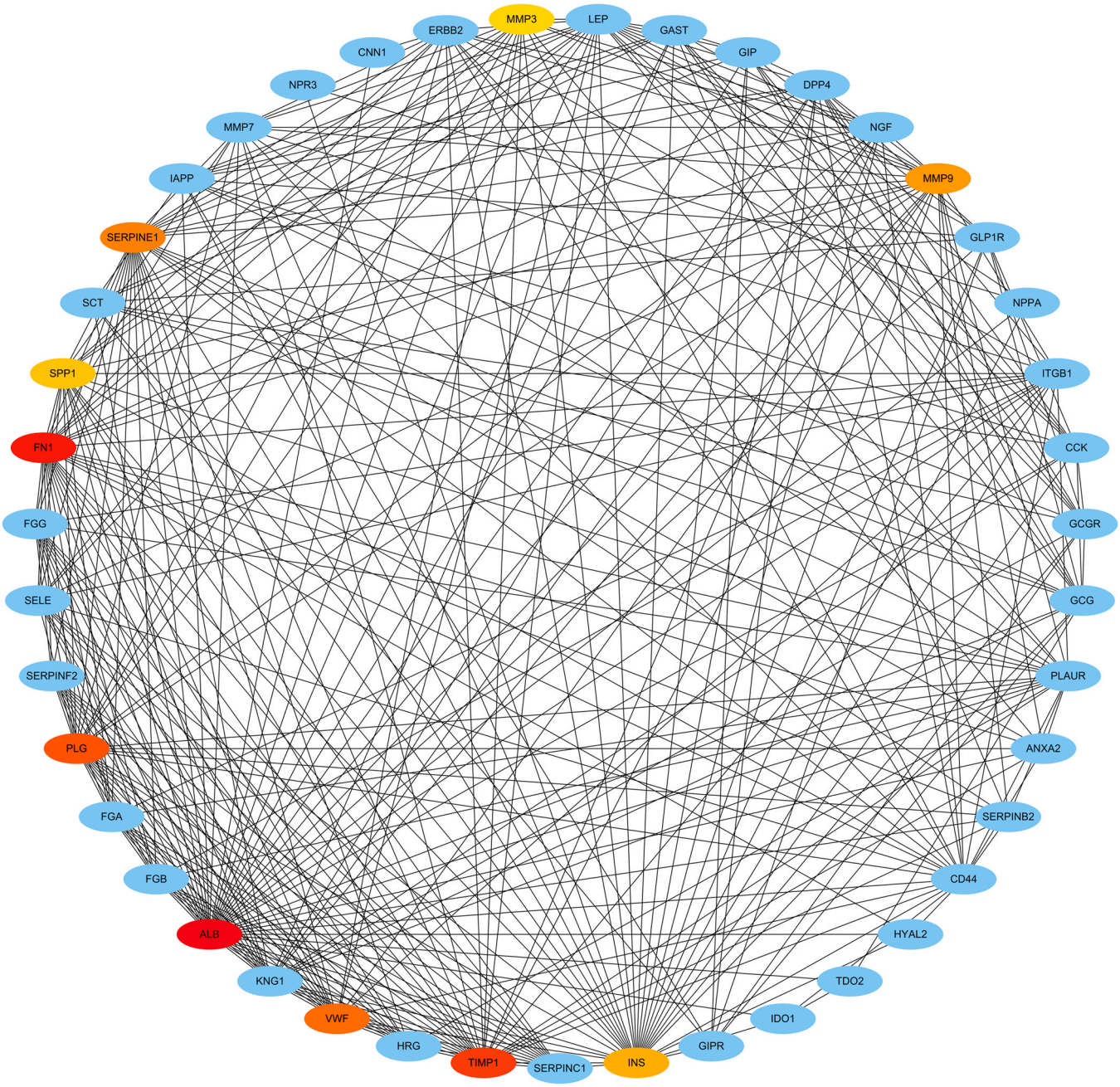

Fig. 5. HUB genes were assayed from a PPI network of co-differentially expressed genes. The three genes highlighted are ALB, FN1, and TIMP1. these three genes are considered as hub genes based on their degree values. According to the analysis number of nodes 43, number of edges 328, Network diameter 3, Network radius 2, Network density 0.363.

PPI network analysis is the most prominent part of the investigation because the assay of pivotal genes, analysis of modules and identification of drugs were totally dependent on PPIs network. the analysis of PPIs also arose from TM4SF18, TDO2, MMP7, MYOM1, CNN1, TTLL7, PLD5, DPP4, NPR3 genes as these genes are common DEGs. According to the PPI network, ALB, FN1, TIMP1, PLG, VWF, SERPINE1, MMP9, INS, SPP1, MMP3 genes were declared as hub genes due to their high interaction rate or degree values. A retrospective analysis study found a significant decrease in ALB levels in the blood of patients with Covid-19 infection [44], The levels of hemoglobin (Hb) and albumin (Alb) predicts incremental risk for severe respiratory failure in Covid-19 patients with pneumonia [45]. To focus on important regions of the PPI network, a modular analysis of hub genes was implemented. The reason for concentrating on highly focused regions was more effective drug compound recommendations.

TF-gene interaction was obtained with the common DEGs. From the network, MYOM1 shows a high interaction rate with other TF-genes. Among the regulators, ATF3 and SUZ12 have significant interaction. The degree value of ATF3 and SUZ12 are 2, respectively in the TF-gene interactions network. ATF3 (activating transcription factor 3), a member of the CREB family of basic leucine zipper

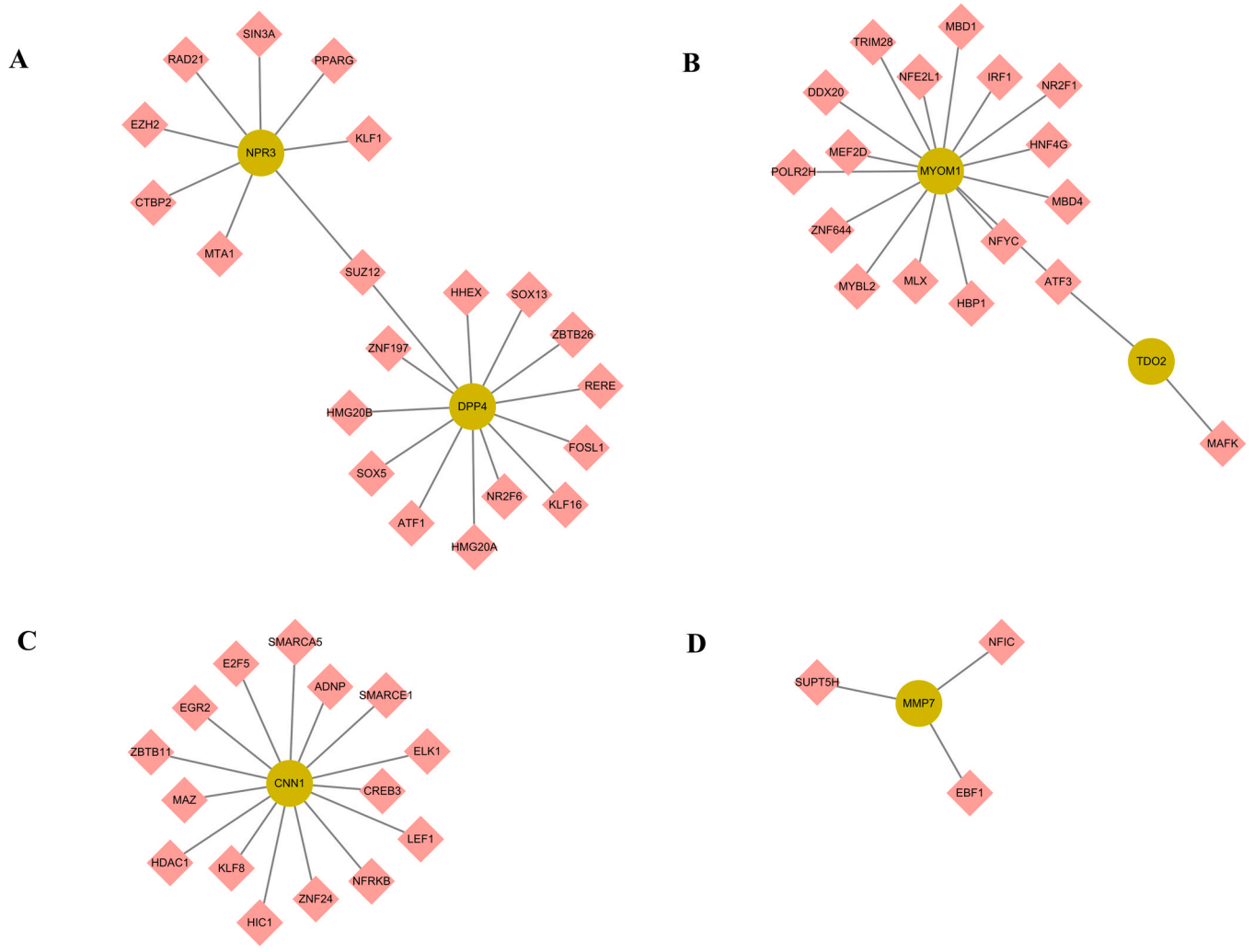

Fig. 6. Network of TF genes interacting with differentially expressed genes, A contains Nodes 22 Edges 21, DPP4 regulated by 13 TF genes and NPR3, regulated by 8 TF genes, B contains Nodes 19 Edges 18, MYOM1 is regulated by 16 TF genes and TDO2 is regulated by 2 TF genes, C contains Nodes 16 Edges 15, CNN1 is regulated by 15 TF genes, D contains Nodes 4, Edges3, and MMP7 regulated by 3 TF genes.

transcription factors (TFs), has been demonstrated to be involved in both inflammatory and metabolic pathways [46,47]. ATF3 is expressed at higher levels in human atherosclerotic vessels and ATF3 is strongly expressed in macrophages within atherosclerotic vessels [48,49]. SUZ12 (Suppressor of zeste 12 protein homolog) is a core component of the Polycomb PRC2-HMTase complex, which has been proved to be participating in stem cell maintenance and development [50–52]. SUZ12 is upregulated in many human cancers including gastric [53,54]、breast cancers [55] and bladder [51]. For illustration, the expression of SUZ12 was markedly higher in gastric tumor tissues as opposed to normal tissues [53]. SUZ12 expression was revealed to be correlated with pathological stage, metastatic distance and lower overall survival in GC patients.

Regulatory biomolecules serve as potential biomarkers in multiple complex illnesses. The activities of miRNAs and TF genes for the regulatory analysis of common DEGs were visualized in a TF-miRNA co-regulatory network. 115 miRNAs and 47 TF-genes are identified in the research. Among the most interacted TFs,CTCF has the higher degree value of 4. CTCF plays an important transcriptional regulatory role in lung cancer and atherosclerosis [56–58]. TF-genes are responders for the regulation, which is accomplished by binding to target genes and miRNAs, and are capable of regulating gene expression through mRNA degradation [59].

Pharmaceutical molecules were proposed by 10 common DEGs according to the DSigDB database. Among all drug candidates, the current research accentuates the top 10 significant drugs. ascorbic acid CTD 00005445 are the peak drug candidates for COVID-19 and Carotid atherosclerosis. There have been relevant randomized clinical trials using ascorbic acid in the treatment of ambulatory patients with SARS-CoV-2 infection [60]. At this stage, there are no studies that have used progesterone in the clinical treatment of SARS-CoV-2. In addition, the therapeutic effect of progesterone on atherosclerosis has been demonstrated for a long time [61,62]. These agents can be considered for further validation by chemical experiments. Since SARS-CoV-2 is a new virus, the studies done so far are relatively few. This is the reason why fewer samples were collected to analyze the results. In the future, with more samples, the current studies will be more effective in the case of a SARS-CoV-2 pandemic.

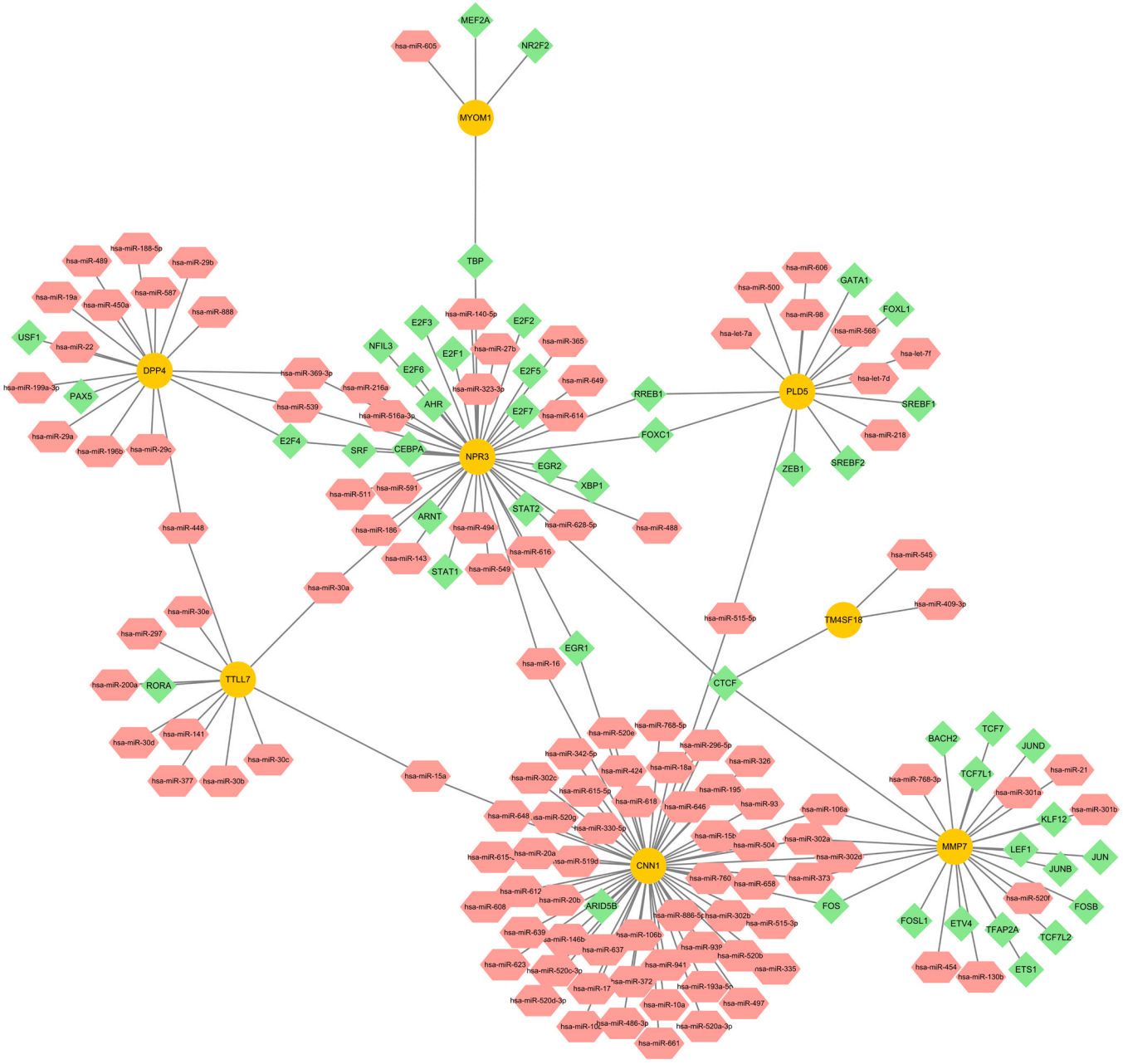

Fig. 7. The network of TF genes interacting with miRNAs. The highlighted yellow nodes indicate co-expressed genes, green nodes are TF genes, and other nodes indicate miRNAs. the network consists of 170 nodes and 182 edges.

Table 3
Suggested top drug compounds for the carotid atherosclerosis SARS-CoV-2 infections.

Candidate drugs	P-value	Combined score	Genes
ascorbic acid CTD 00005445	0.000174	331.0154894	DPP4;CNN1;MMP7
progesterone CTD 00006624	0.000725	85.54243324	DPP4;CNN1;MMP7;TDO2;NPR3
3,4-DICHLOROANILINE CTD 00000754	0.001578	275.2122078	DPP4;PLD5
diuron CTD 00005864	0.001768	254.9198622	DPP4;PLD5
MANGANESE CHLORIDE CTD 00001187	0.003288	166.7035725	DPP4;MYOM1
MANGANESE CTD 00006240	0.003869	148.7806594	DPP4;MYOM1
CGS-27023A TTD 00002801	0.005836	1070.484451	MMP7
benzyl isothiocyanate CTD 00001509	0.006284	973.8925198	MMP7
fucose TTD 00008125	0.008072	707.7995527	DPP4
cyclosporin A CTD 00007121	0.008287	30.18071977	TTLL7;DPP4;TM4SF18;MMP7;TDO2;NPR3

Conclusions

In terms of transcriptome analysis, there are no other studies on SARS-CoV-2 and carotid atherosclerosis to date. We have completed the analysis of DEGs between the two datasets and filtered the material by common gene identification in an attempt to find the infection response between SARS-CoV-2 and carotid atherosclerosis. The analysis on SARS-CoV-2 and carotid atherosclerosis predicts the way to detect infection in various diseases. The suggestions for drug targets are logical because they were derived by identifying the central gene, which may act as a positive precursor to already approved drugs. Because SARS-CoV-2 is a recent discovery, few studies have been conducted on its risk factors. Unique studies on SARS-CoV-2 will become ever more important with the availability of exceeding datasets.

Funding

This work was supported by National Natural Science Foundation of China, Grant/Award Number: 81971135 and 81572482; Natural Science Foundations of Heilongjiang, Grant/Award Number: YQ2020H014; “Chunhui Plan” of Ministry of Education, Grant/Award Number: HLJ2019009; Distinguished Young Foundations of the First Affiliated Hospital of Harbin Medical University (HYD2020JQ0014).

Declaration of Competing Interest

The authors declare no potential conflicts of interest.

Acknowledgments

We sincerely thank Harbin Institute of Information Technology for our support and help, and also grateful with all the participants of this study. We would like to give our sincere appreciation to the reviewers for their helpful comments on this article.

References

[1] Zou L, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. N Engl J Med 2020;382:1177–9. https://doi.org/10.1056/NEJMc2001737
[2] Alexander e Gorbalenya, et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol 2020;5:536–44. https://doi.org/10.1038/s41564-020-0695-z
[3] Xu X, et al. Imaging and clinical features of patients with 2019 novel coronavirus SARS-CoV-2. Eur J Nucl Med Mol Imaging 2020;47:1275–80. https://doi.org/10.1007/s00259-020-04735-9
[4] Li Y, et al. Acute cerebrovascular disease following COVID-19: a single center, retrospective, observational study. Stroke Vasc Neurol 2020;5:279–84. https://doi.org/10.1136/svn-2020-000431
[5] Szelenberger R, Saluk-Bijak J, Bijak M. Ischemic stroke among the symptoms caused by the COVID-19 infection. J Clin Med 2020;9. https://doi.org/10.3390/jcm9092688
[6] Tsivgoulis G, et al. COVID-19 and cerebrovascular diseases: a comprehensive overview. Ther Adv Neurol Disord 2020;13. https://doi.org/10.1177/1756286420978004
[7] GBD 2019 Diseases and Injuries Collaborators.Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet (London, England) 396, 1204–1222, doi:10.1016/s0140–6736(20)30925–9 (2020).
[8] GBD 2016 Neurology Collaborators.Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. Neurology 18, 459–480, doi:10.1016/s1474–4422(18)30499-x (2019).
[9] Naylor A, et al. Choice - Management of Atherosclerotic Carotid and Vertebral Artery Disease: 2017 Clinical Practice Guidelines of the European Society for Vascular Surgery (ESVS). Eur J Vasc Endovasc Surg: Off J Eur Soc Vasc Surg 2018;55:3–81. https://doi.org/10.1016/j.ejvs.2017.06.021
[10] Ntaios G, Hart R. Embolic stroke. Circulation 2017;136:2403–5. https://doi.org/10.1161/circulationaha.117.030509
[11] Walls A, et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 2020;183:1735. https://doi.org/10.1016/j.cell.2020.11.032
[12] Brake S, et al. Smoking upregulates angiotensin-converting enzyme-2 receptor: a potential adhesion site for novel coronavirus SARS-CoV-2 (Covid-19). J Clin Med 2020;9. https://doi.org/10.3390/jcm9030841
[13] Zou X, et al. Single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to 2019-nCoV infection. Front Med 2020;14:185–92. https://doi.org/10.1007/s11684-020-0754-0
[14] Alenina N, Bader M. ACE2 in brain physiology and pathophysiology: evidence from transgenic animal models. Neurochem Res 2019;44:1323–9. https://doi.org/10.1007/s11064-018-2679-4
[15] Baig A, Khaleeq A, Ali U, Syeda H. Evidence of the COVID-19 virus targeting the CNS: tissue distribution, host-virus interaction, and proposed neurotropic mechanisms. ACS Chem Neurosci 2020;11:995–8. https://doi.org/10.1021/acschemneuro.0c00122
[16] Li Y, Bai W, Hashikawa T. The neuroinvasive potential of SARS-CoV2 may play a role in the respiratory failure of COVID-19 patients. J Med Virol 2020;92:552–5. https://doi.org/10.1002/jmv.25728
[17] Wu Y, et al. Nervous system involvement after infection with COVID-19 and other coronaviruses. Brain Behav Immun 2020;87:18–22. https://doi.org/10.1016/j.bbi.2020.03.031
[18] Dalrymple N, Mackow E. Virus interactions with endothelial cell receptors: implications for viral pathogenesis. Curr Opin Virol 2014;7:134–40. https://doi.org/10.1016/j.coviro.2014.06.006
[19] van Hinsbergh V. Endothelium–role in regulation of coagulation and inflammation. Semin Immunopathol 2012;34:93–106. https://doi.org/10.1007/s00281-011-0285-5
[20] Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. Bioinformatics 2002;18:207–8. https://doi.org/10.1093/bioinformatics/18.1.207
[21] Lee M, Kuo F, Whitmore G, Sklar J. Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. Proc Natl Acad Sci USA 2000;97:9834–9. https://doi.org/10.1073/pnas.97.18.9834
[22] Irigoyen N, et al. High-resolution analysis of coronavirus gene expression by RNA sequencing and ribosome profiling. PLOS Pathog 2016;12:e1005473https://doi.org/10.1371/journal.ppat.1005473
[23] Jacob F, et al. Human pluripotent stem cell-derived neural cells and brain organoids reveal SARS-CoV-2 neurotropism predominates in choroid plexus epithelium. Cell Stem Cell 2020;27(937–950):e939https://doi.org/10.1016/j.stem.2020.09.016
[24] Ayari H, Bricca G. Identification of two genes potentially associated in iron-heme homeostasis in human carotid plaque using microarray analysis. J Biosci 2013;38:311–5. https://doi.org/10.1007/s12038-013-9310-2
[25] Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545–50. https://doi.org/10.1073/pnas.0506580102
[26] Doms A, Schroeder M. GoPubMed: exploring PubMed with the gene ontology. Nucleic Acids Res 2005;33:W783–6. https://doi.org/10.1093/nar/gki470
[27] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27–30. https://doi.org/10.1093/nar/28.1.27
[28] Slenter D, et al. WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. Nucleic Acids Res 2018;46:D661–7. https://doi.org/10.1093/nar/gkx1064
[29] Jassal B, et al. The reactome pathway knowledgebase. Nucleic Acids Res 2020;48:D498–503. https://doi.org/10.1093/nar/gkz1031
[30] Ewing R, et al. Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol Syst Biol 2007;3:89. https://doi.org/10.1038/msb4100134
[31] Ben-Hur A, Noble W. Kernel methods for predicting protein-protein interactions. Bioinformatics 2005:i38–46. https://doi.org/10.1093/bioinformatics/bti1016
[32] Ye Z, et al. Bioinformatic identification of candidate biomarkers and related transcription factors in nasopharyngeal carcinoma. World J Surg Oncol 2019;17:60. https://doi.org/10.1186/s12957-019-1605-9
[33] Zhou G, et al. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic Acids Res 2019;47:W234–41. https://doi.org/10.1093/nar/gkz240
[34] Li Z, Lindner D, Bishop N, Cipolla M. ACE (angiotensin-converting enzyme) inhibition reverses vasoconstriction and impaired dilation of pial collaterals in chronic hypertension. Hypertension 2020;76:226–35. https://doi.org/10.1161/hypertensionaha.119.14315
[35] Cheng W, et al. Angiotensin II inhibits neuronal nitric oxide synthase activation through the ERK1/2-RSK signaling pathway to modulate central control of blood pressure. Circ Res 2010;106:788–95. https://doi.org/10.1161/circresaha.109.208439
[36] Röhrborn D, Eckel J, Sell H. Shedding of dipeptidyl peptidase 4 is mediated by metalloproteases and up-regulated by hypoxia in human adipocytes and smooth muscle cells. FEBS Lett 2014;588:3870–7. https://doi.org/10.1016/j.febslet.2014.08.029
[37] Ishibashi Y, Matsui T, Maeda S, Higashimoto Y, Yamagishi S. Advanced glycation end products evoke endothelial cell damage by stimulating soluble dipeptidyl peptidase-4 production and its interaction with mannose 6-phosphate/insulin-like growth factor II receptor. Cardiovasc Diabetol 2013;12:125. https://doi.org/10.1186/1475-2840-12-125
[38] Millet J, Whittaker G. Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. Virus Res 2015;202:120–34. https://doi.org/10.1016/j.virusres.2014.11.021
[39] Tortorici M, Veesler D. Structural insights into coronavirus entry. Adv Virus Res 2019;105:93–116. https://doi.org/10.1016/bs.aivir.2019.08.002

[40] Coutard B, et al. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. Antivir Res 2020;176:104742https://doi.org/10.1016/j.antiviral.2020.104742
[41] Anderson G, Carbone A, Mazzoccoli G. Tryptophan metabolites and aryl hydrocarbon receptor in severe acute respiratory syndrome, coronavirus-2 (SARS-CoV-2) pathophysiology. Int J Mol Sci 2021;22. https://doi.org/10.3390/ijms22041597
[42] Pan H, et al. Repurposed antiviral drugs for covid-19 – interim WHO solidarity trial results. N Engl J Med 2021;384:497–511. https://doi.org/10.1056/NEJMoa2023184
[43] Sudar-Milovanovic E, Gluvic Z, Obradovic M, Zaric B, Isenovic E. Tryptophan metabolism in atherosclerosis and diabetes. Curr Med Chem 2021:99–113. https://doi.org/10.2174/0929867328666210714153649
[44] Bi X, et al. Prediction of severe illness due to COVID-19 based on an analysis of initial fibrinogen to albumin ratio and platelet count. Platelets 2020;31:674–9. https://doi.org/10.1080/09537104.2020.1760230
[45] Anai M, et al. Decrease in hemoglobin level predicts increased risk for severe respiratory failure in COVID-19 patients with pneumonia. Respir Investig 2021;59:187–93. https://doi.org/10.1016/j.resinv.2020.10.009
[46] Shi Q, et al. ATF3 inhibits arsenic-induced malignant transformation of human bronchial epithelial cells by attenuating inflammation. Toxicology 2021;152890. https://doi.org/10.1016/j.tox.2021.152890
[47] Yan F, et al. ATF3 is positively involved in particulate matter-induced airway inflammation in vitro and in vivo. Toxicol Lett 2018;287:113–21. https://doi.org/10.1016/j.toxlet.2018.01.022
[48] Nawa T, et al. Expression of transcriptional repressor ATF3/LRF1 in human atherosclerosis: colocalization and possible involvement in cell death of vascular endothelial cells. Atherosclerosis 2002;161:281–91. https://doi.org/10.1016/s0021-9150(01)00639-6
[49] Gold E, et al. ATF3 protects against atherosclerosis by suppressing 25-hydroxycholesterol-induced lipid body formation. J Exp Med 2012;209:807–17. https://doi.org/10.1084/jem.20111202
[50] Vizán P, Beringer M, Ballaré C, Di Croce L. Role of PRC2-associated factors in stem cells and disease. FEBS J 2015;282:1723–35. https://doi.org/10.1111/febs.13083
[51] Lee S, et al. Activation of EZH2 and SUZ12 regulated by E2F1 predicts the disease progression and aggressive characteristics of bladder cancer. Clin Cancer Res: J Am Assoc Cancer Res 2015;21:5391–403. https://doi.org/10.1158/1078-0432.Ccr-14-2680
[52] Lee S, et al. Polycomb repressive complex 2 component Suz12 is required for hematopoietic stem cell function and lymphopoiesis. Blood 2015;126:167–75. https://doi.org/10.1182/blood-2014-12-615898
[53] Cui Y, Chen J, He Z, Xiao Y. SUZ12 depletion suppresses the proliferation of gastric cancer cells. Cell Physiol Biochem: Int J Exp Cell Physiol, Biochem, Pharmacol 2013;31:778–84. https://doi.org/10.1159/000350095
[54] Wu X, et al. Ubiquitin-specific protease 3 promotes cell migration and invasion by interacting with and deubiquitinating SUZ12 in gastric cancer. J Exp Clin Cancer Res 2019;38:277. https://doi.org/10.1186/s13046-019-1270-4
[55] Hwang-Verslues W, et al. Loss of corepressor PER2 under hypoxia up-regulates OCT1-mediated EMT gene expression and enhances tumor malignancy. Proc Natl Acad Sci USA 2013;110:12331–6. https://doi.org/10.1073/pnas.1222684110
[56] Yang Y, et al. Silencing of long non-coding RNA H19 downregulates CTCF to protect against atherosclerosis by upregulating PKD1 expression in ApoE knockout mice. Aging 2019;11:10016–30. https://doi.org/10.18632/aging.102388
[57] Dai J, et al. Systematical analyses of variants in CTCF-binding sites identified a novel lung cancer susceptibility locus among Chinese population. Sci Rep 2015;5:7833. https://doi.org/10.1038/srep07833
[58] Eldholm V, Haugen A, Zienolddiny S. CTCF mediates the TERT enhancer-promoter interactions in lung cancer cells: identification of a novel enhancer region involved in the regulation of TERT gene. Int J Cancer 2014;134:2305–13. https://doi.org/10.1002/ijc.28570
[59] Zhang H, et al. Transcription factor and microRNA co-regulatory loops: important regulatory motifs in biological processes and diseases. Brief Bioinforma 2015;16:45–58. https://doi.org/10.1093/bib/bbt085
[60] Thomas S, et al. Effect of high-dose zinc and ascorbic acid supplementation vs usual care on symptom length and reduction among ambulatory patients with SARS-CoV-2 infection: The COVID A to Z randomized clinical trial. JAMA Netw Open 2021;4:e210369https://doi.org/10.1001/jamanetworkopen.2021.0369
[61] Hanke H, et al. Different effects of estrogen and progesterone on experimental atherosclerosis in female versus male rabbits. Quantif Cell Prolif Bromodeoxyuridine Circ 1996;94:175–81. https://doi.org/10.1161/01.cir.94.2.175
[62] Hanke H, et al. Inhibition of the protective effect of estrogen by progesterone in experimental atherosclerosis. Atherosclerosis 1996;121:129–38. https://doi.org/10.1016/0021-9150(95)05710-2