

## ORIGINAL ARTICLE

# An epigenome-wide DNA methylation study of patients with COVID-19

Shengyu Zhou<sup>1,2,#</sup> | Jintao Zhang<sup>3,#</sup>  | Jiawei Xu<sup>3,#</sup> | Fayan Zhang<sup>4</sup> | Peng Li<sup>5</sup> | Yujie He<sup>6</sup> | Julong Wu<sup>6</sup> | Chunting Wang<sup>7</sup> | Ximing Wang<sup>8</sup> | Wei Zhang<sup>9</sup> | Kang Ning<sup>3</sup> | Yun Pan<sup>3</sup> | Tian Liu<sup>1</sup> | Jiping Zhao<sup>1</sup> | Lixia Yin<sup>10</sup> | Rumin Zhang<sup>11</sup> | Feng Gao<sup>12</sup> | Jintong Zhao<sup>11</sup> | Liang Dong<sup>3,13</sup>

<sup>1</sup> Department of Respiratory and Critical Care Medicine, Qilu Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China

<sup>2</sup> Clinical Nursing Department, School of Nursing, Cheeloo College of Medicine, Shandong University, Jinan, China

<sup>3</sup> Department of Respiratory, Shandong Qianfoshan Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China

<sup>4</sup> College of Traditional Chinese Medicine, Shandong University of Traditional Chinese Medicine, Jinan, China

<sup>5</sup> Blood Center of Shandong Province, Jinan, China

<sup>6</sup> Institute of Virology, Shandong Center for Diseases Prevention and Control, Jinan, China

<sup>7</sup> Department of Critical Care Medicine, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China

<sup>8</sup> Department of Radiology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Shandong University, Jinan, China

<sup>9</sup> Department of Lung Disease, Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan, China

<sup>10</sup> Department of Respiratory Medicine, Dezhou People's Hospital, Dezhou, China

<sup>11</sup> Department of Critical Care Medicine, Zibo Central Hospital, Zibo, China

<sup>12</sup> Department of Infectious Disease, Linyi People's Hospital, Linyi, China

<sup>13</sup> Department of Respiratory, Shandong Provincial Qianfoshan Hospital, Shandong University, The First Affiliated Hospital of Shandong First Medical University, Shandong Institute of Respiratory Diseases, Jinan, China

## Correspondence

Liang Dong, Department of Respiratory, Shandong Provincial Qianfoshan Hospital, Shandong University, The First Affiliated Hospital of Shandong First Medical University, Shandong Institute of Respiratory Diseases, Jinan, China.

Email: dl5506@126.com

<sup>#</sup>The authors Shengyu Zhou, Jintao Zhang, and Jiawei Xu contributed equally to this work.

## Abstract

In the early 2000s, emerging SARS-CoV-2, which is highly pathogenic, posed a great threat to public health. During COVID-19, epigenetic regulation is deemed to be an important part of the pathophysiology and illness severity. Using the Illumina Infinium Methylation EPIC BeadChip (850 K), we investigated genome-wide differences in DNA methylation between healthy subjects and COVID-19 patients with different disease severities. We conducted a combined analysis and selected 35 “marker” genes that could indicate a SARS-CoV-2 infection, including 12 (*ATHL1*, *CHN2*, *CHST15*, *CPLX2*, *CRHR2*, *DCAKD*, *GNAI2*, *HECW1*, *HYALI*, *MIR510*, *PDE11A*, and *SMG6*) situated in the promoter region. The functions and pathways of differentially methylated genes were enriched in biological processes, signal transduction, and the immune system. In the “Severe versus Mild” group, differentially methylated genes, after eliminating duplicates, were used for PPI analyses. The four hub genes (*GNG7*, *GNAS*, *PRKCZ*, and *PRKAG2*) that had the highest degree of nodes were identified and among them, *GNG7* and *GNAS* genes expressions were also downregulated in the severe group in

sequencing results. Above all, the results suggest that GNG7 and GNAS may play a non-ignorable role in the progression of COVID-19. In conclusion, the identified key genes and related pathways in the current study can be used to study the molecular mechanisms of COVID-19 and may provide possibilities for specific treatments.

#### KEYWORDS

COVID-19, DNA methylation, epigenetic regulation, SARS-CoV-2

## 1 | INTRODUCTION

Among the varieties of viruses, coronaviruses with the largest RNA genome have drawn much attention in light of the recent emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Yang et al., 2020; Zhou et al., 2020). Several studies have described the clinical characteristics of COVID-19 (Chen, Zhou et al., 2020; Wang et al., 2020). Infected patients with SARS-CoV-2 have varied symptoms, from mild feverish symptoms to those with severe clinical outcomes (Yang et al., 2020). Additionally, it is worth noting that there were increased concentrations of proinflammatory cytokines (e.g., interferon  $\gamma$  [IFN $\gamma$ ] and interleukin 6 [IL-6]) in coronavirus disease 2019 (COVID-19) patients, and these were associated with extensive lung damage and disease severity (Huang et al., 2020). In recent studies, mild and severe COVID-19 patients were observed to have different inflammatory characteristics, and the disturbed immune responses were related to the severe manifestations of COVID-19 (Bost et al., 2020; Huang et al., 2020). Poorly understood interactions between SARS-CoV-2 and the host limit the development of efficient treatment strategies. A thorough understanding of the interactions between the host and virus during infections will likely save lives.

DNA methylation is a potent epigenetic mechanism that regulates gene expression without altering DNA sequences. It has been shown that the production rate of the ACE2 receptor by its gene, known as the most critical virus receptor for SARS-CoV-2, is under epigenetic control (Zill et al., 2012). Age correlates with hypomethylation of the ACE2 gene in lung tissue, which could, to some extent, explain why aging is a risk factor for SARS-CoV-2 fatality (Corley & Ndhlovu, 2020). Existing studies have revealed that DNA methylation is implicated in the pathophysiology of many different diseases including, virus infections (Silmon de Monerri & Kim, 2014; Chlamydas et al., 2020). For many years, it has been known that viral infections usually use epigenetic

mechanisms to find ways to induce syncytium development (Pruimboom, 2020).

Moreover, virus-driven dysregulation of gene methylation levels could also help induce aberrant host immune responses and influence disease development (Kuss-Duerkop et al., 2018). As an individual risk factor for COVID-19, some tumors showed hypo-DNA methylation levels of ACE2, which may cause ACE2 to be highly expressed and increase COVID-19 susceptibility and severity (Chai et al., 2020). In patients with other comorbidities, epigenetic dysregulation, especially DNA methylation, may lead to an abnormal increase in the expression levels of genes and elicit more robust immune response to SARS-CoV-2 (Sawalha et al., 2020). The focus of this study was on the epigenetic events in the pathogenesis of SARS-CoV-2 infection. We aimed to shed light on the DNA methylation changes in healthy subjects and patients with different severities of COVID-19 to explore novel targets and effective methods for treating COVID-19.

## 2 | MATERIALS AND METHODS

### 2.1 | Clinical data

This study was approved by the ethics committee of Qilu Hospital of Shandong University (ethical approval no: KYLL-2020-GXB-001). After receiving written informed consent, 21 blood samples were obtained from seven healthy volunteers and 14 COVID-19 patients just before the discharge. These 21 subjects (12 male, nine female) were from Shandong Province Chest Hospital (Jinan, China). The patients admitted to hospitals with confirmed SARS-CoV-2 infection were diagnosed with mild or severe COVID-19 between January 27, 2020, and May 18, 2020. All the subjects were biologically unrelated to each other. The severity of the disease was classified based on the "Guideline on the management of COVID-19" published by the National Health Commission of the People's Republic of China (version 8.0). Mild patients had mild symptoms and

no imaging observation of pneumonia. According to the diagnostic criteria for COVID-19 in adults “the New Coronavirus Pneumonia Prevention and Control Program” published by the National Health Commission of China, severe patients had one of the following characteristics: dyspnea with a respiratory rate  $\geq 30$  breaths per min, oxygen saturation  $\leq 93\%$ , or arterial blood oxygen partial pressure ( $\text{PaO}_2$ )/oxygen concentration ( $\text{FiO}_2$ )  $\leq 40$  kPa.

## 2.2 | DNA extraction

Genomic DNA was extracted from 200  $\mu\text{l}$  of whole blood using a QIAamp DNA Blood Mini Kit (Qiagen) following the manufacturer’s instructions. A total of 200  $\mu\text{l}$  of the blood sample and 20  $\mu\text{l}$  of proteinase K were mixed in a 1.5-ml microcentrifuge tube. A total of 200  $\mu\text{l}$  of buffer AL was added to the sample. The solution was mixed with pulse-vortexing for 15 s and then incubated at 56°C for 10 min. A total of 200  $\mu\text{l}$  absolute ethanol was added and mixed again with pulse-vortexing for 15 s. After mixing, the 1.5-ml microcentrifuge tube was centrifuged to remove drops from the inside of the lid. Buffer AW1 (500  $\mu\text{l}$ ) was added and centrifuged at 8000 rpm for 1 min. Then the collection tube containing the filtrate was discarded. Buffer AW2 (500  $\mu\text{l}$ ) was added and centrifuged at 14,000 rpm for 3 min. The spin column was placed in a new collection tube. The tube was centrifuged at full speed for 1 min. The column was placed in a clean microcentrifuge tube, and the collection tube containing the filtrate was discarded. Buffer AE (200  $\mu\text{l}$ ) was then added. The solution was incubated at room temperature for 1 min, and then centrifuged at 8000 rpm for 1 min. All of the above steps were completed according to the guidelines of the Institute for Prevention and Control of Viral Infectious Diseases of Shandong Center for Disease Control and Prevention.

## 2.3 | Detection of DNA methylation in the peripheral blood by microarray

In this study, the methylation-specific chip used was the Illumina Infinium Methylation EPIC 850 K BeadChip (Illumina). The applicable protocol for the 850 K BeadChip followed the manufacturer’s guidelines from OE biotech Co., Ltd (Shanghai, China). After extracting DNA from the peripheral blood, gDNA was obtained by bisulfite treatment. After hybridization with each probe, the specific capture probe on the chip was used to combine with the complementary gene fragments. The original data were obtained by hybridization overnight followed by

cleaning, extension, staining, and scanning. The degree of methylation was determined by the fluorescence intensity of the probes.

## 2.4 | Screening differential genes based on bioinformatics analysis

The data expression profiles of COVID-19 patients were searched in the GEO database, and one microarray analysis of the whole genome transcriptome to peripheral blood sample taken from severe and mild COVID-19 patients, GSE164805, was selected. Differential mRNAs were screened and adjusted  $P$  value  $< 0.05$  is considered as statistically significant.

## 2.5 | Statistical analysis

The signal value was extracted from the original data using a genome studio (genomic software 2011.1; Illumina), and the resulting data were normalized by performing background correction using the methylation module. The diff-score value ( $\text{diffscore} = 10 * \text{sgn}(\beta_{\text{ref}} - \beta_{\text{cond}}) * \log_{10}(p)$ ) of the experimental group was less than  $-13$  or  $> 13$ ; and  $\Delta\beta > 0.17$  or  $< -0.17$ , which indicated it was the differential methylation site. (Note: The calculation of the  $\Delta\beta$  value is based on the control group and the experimental group. Results for the AVG  $\beta$  difference; that is, the degree of methylation difference at each site between the experimental group and the normal control group; a negative value indicated that it is a low methylation site compared with the control group, otherwise it was a high methylation site; For a  $P$ -value of 0.05,  $\text{diffscore} = \pm 13$ .) Both Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed to clarify the biological and functional pathways of differentially expressed methylation loci. GO terms and KEGG with  $P$ -value  $< 0.05$  were considered significantly enriched from our data. Cytoscape v3.0 software was used for constructing an interaction network, which was based on differentially methylated genes (DMGs). The matrix of gene expression values was visualized graphically. The gene-interaction relationship was represented by nodes and edges in the graph. In the signaling network, the color of the cycle was considered to be the frequency of the gene interaction. The most prominent central genes in the network indicated the genes with the highest frequency. The top genes which had the highest degree of nodes and hit more than five cytohubba algorithms among the DMGs were defined as hub genes.

**TABLE 1** Demographics and baseline characteristics

	Control ( <i>n</i> = 7)	Mild ( <i>n</i> = 7)	Severe ( <i>n</i> = 7)	All ( <i>n</i> = 21)	<i>P</i> value
Age	48.1(43.5–52.0)	42.4(36.5–47.0)	55(48.0–57.5)	48.5(42.0–54.0)	0.17
Sex					
Male	4(57%)	4(57%)	4(57%)	12(57%)	..
Female	3(43%)	3(43%)	3(43%)	9(43%)	..
Days from illness onset to admission	..	5.0(0.5–9.5)	3.4(2.0–4.5)	..	0.75 <sup>a</sup>
Days from admission to discharge	..	15.1(11.5–18.0)	36.9(34.0–41.0)	..	<0.01 <sup>a</sup>
White blood cell count, × 10 <sup>9</sup> per L	..	6.1(4.5–7.2)	7.8(5.3–7.2)	..	0.28 <sup>a</sup>
Neutrophil count, × 10 <sup>9</sup> per L	..	4.1(2.9–5.1)	5.8(3.6–6.0)	..	0.11 <sup>a</sup>
Lymphocyte count, × 10 <sup>9</sup> per L	..	1.4(0.7–1.9)	0.9(0.7–1.0)	..	0.37 <sup>a</sup>
Current smoker	0	0	0	0	..
Cardiovascular diseases	0	0	0	0	..
Obesity	0	0	0	0	..

Data are shown as the average (interquartile range) or *n* (*n*/*N*%) where *N* is the total number of subjects with available data. *P* values comparing the ages of healthy controls, mild patients, and severely ill patients are from a one-way ANOVA.

<sup>a</sup>*P* values comparing basic information and laboratory parameters at discharge of severe and mild patients are from a Mann–Whitney U test.

### 3 | RESULTS

#### 3.1 | Baseline data of patients and healthy volunteers

Data were collected from designated hospitals for COVID-19 treatment in Shandong Province, China. Eight patients (57%) were male. Among the three study groups, the sex proportion was similar. The average age for all subjects was 48.5 years (interquartile range, 42.0–54.0), ranging from 34 years to 88 years. Though severely ill patients had a relatively older median age [55.0 (48.0–57.5) years] than healthy volunteers [48.1 (43.5–52.0) years] and mild patients [42.4 (36.5–47.0) years], there was no significant statistical difference in the age among the three groups (Table 1). The average duration from the onset of symptoms to admission was 5.0 (0.5–9.5) days for mild patients, while that from admission to discharge was 15.1 (11.5–18.0) days. The time from admission to discharge was 36.9 (34.0–41.0) days for severely ill patients. Some laboratory indexes from COVID-19 patients on admission were also analyzed statistically.

#### 3.2 | Identification of differential methylation

Following our screening method of differential methylation loci (the sample diffscore value was less than –13 or greater than 13, and delta beta was greater than 0.17 or less than –0.17), 214, 245, and 228 differentially methylated loci were identified in each group (Severe vs Control, Severe vs Mild and Mild vs Control, respectively). Among them, there were 114, 126, and 113 hypermethylation sites and 100, 119, and 115 low methylation sites. All

delta beta values, including maximum and minimum values, are shown in Figure 1D as violin plots. We then compared the methylation status of different substructures, such as CpG islands, shores (the 2 kb flanking the islands), shelves (the 2 kb flanking the shores), and sea (regions outside the previous three categories) in our datasets (Figure 1). The results showed robust alteration in DNA methylation of open sea probes in all samples. Moreover, hypomethylated sites were preferentially distributed in shelves and shores, and hypermethylated sites were preferentially located in shores rather than islands and shelves (Figure 1).

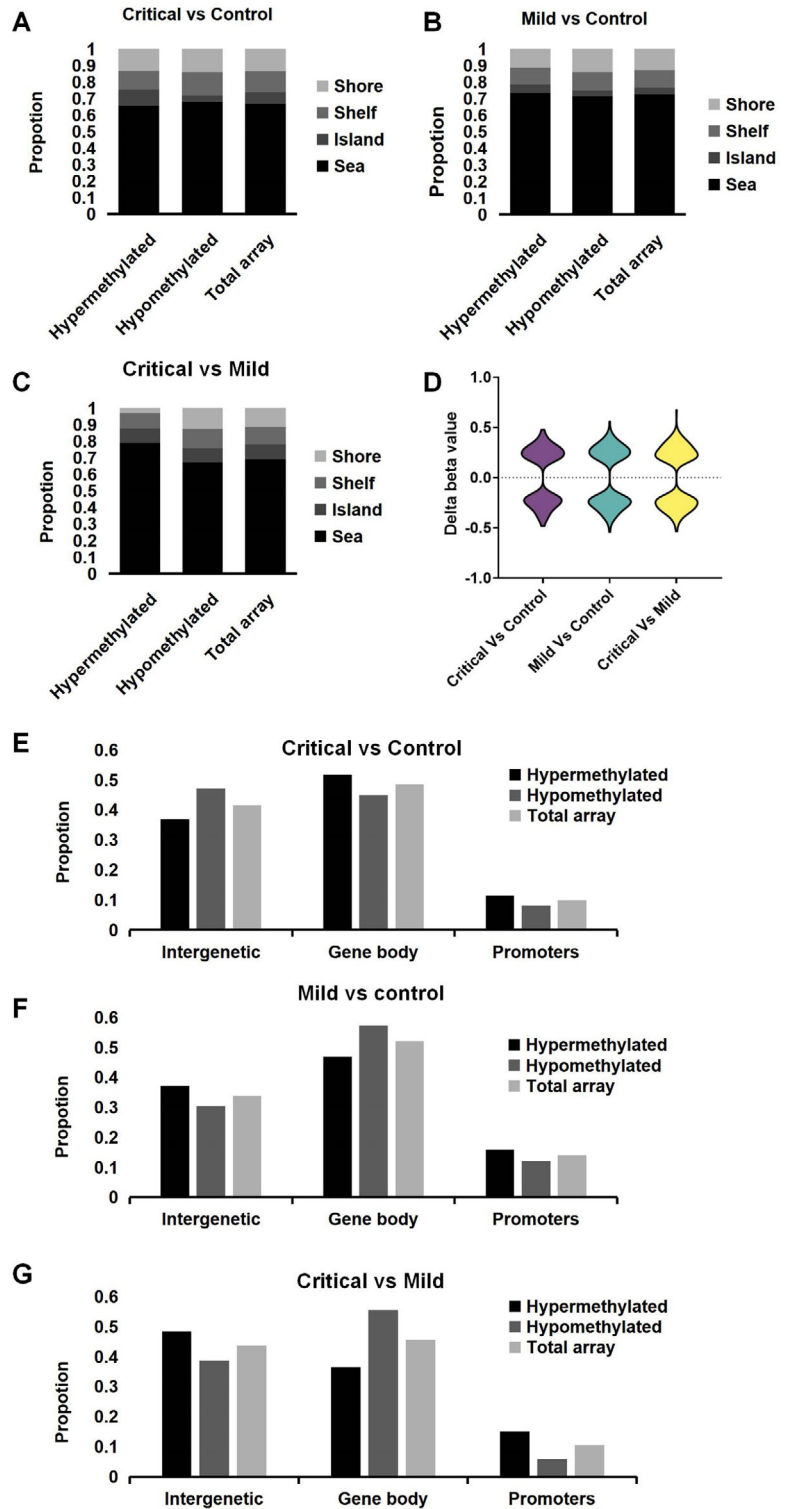
#### 3.3 | Cluster analysis

Unsupervised cluster analysis clustered the samples into two groups (Figure 2). These results showed statistically significant differences in DNA methylation in the healthy controls and COVID-19 patients with different disease severities.

#### 3.4 | Intersection of two groups

To find the key “marker” genes in the process of SARS-CoV-2 infection, we carried out Venn gathering analysis. The distribution of the target genes in two groups (“Mild vs Control” and “Severe vs Control”) is shown in Figure 3. As shown, a total of 35 target genes were detected simultaneously in the two groups. In this context, 35 genes were identified as “markers” of SARS-CoV-2 infection based on gathering analysis between the “Severe versus Control” group and the “Mild versus Control” group (common genes are listed in Table 2).

**FIGURE 1** Distribution of differentially methylated sites. (A–C) Distribution of sites in CpG islands, shores, shelves, and sea. (D) Distribution of delta beta values in each group. (E–F) Sites relative to UCSC\_RefSeq gene promoters, gene bodies, and intergenic regions. All the differentially methylated sites are referred to simply as “Total array”

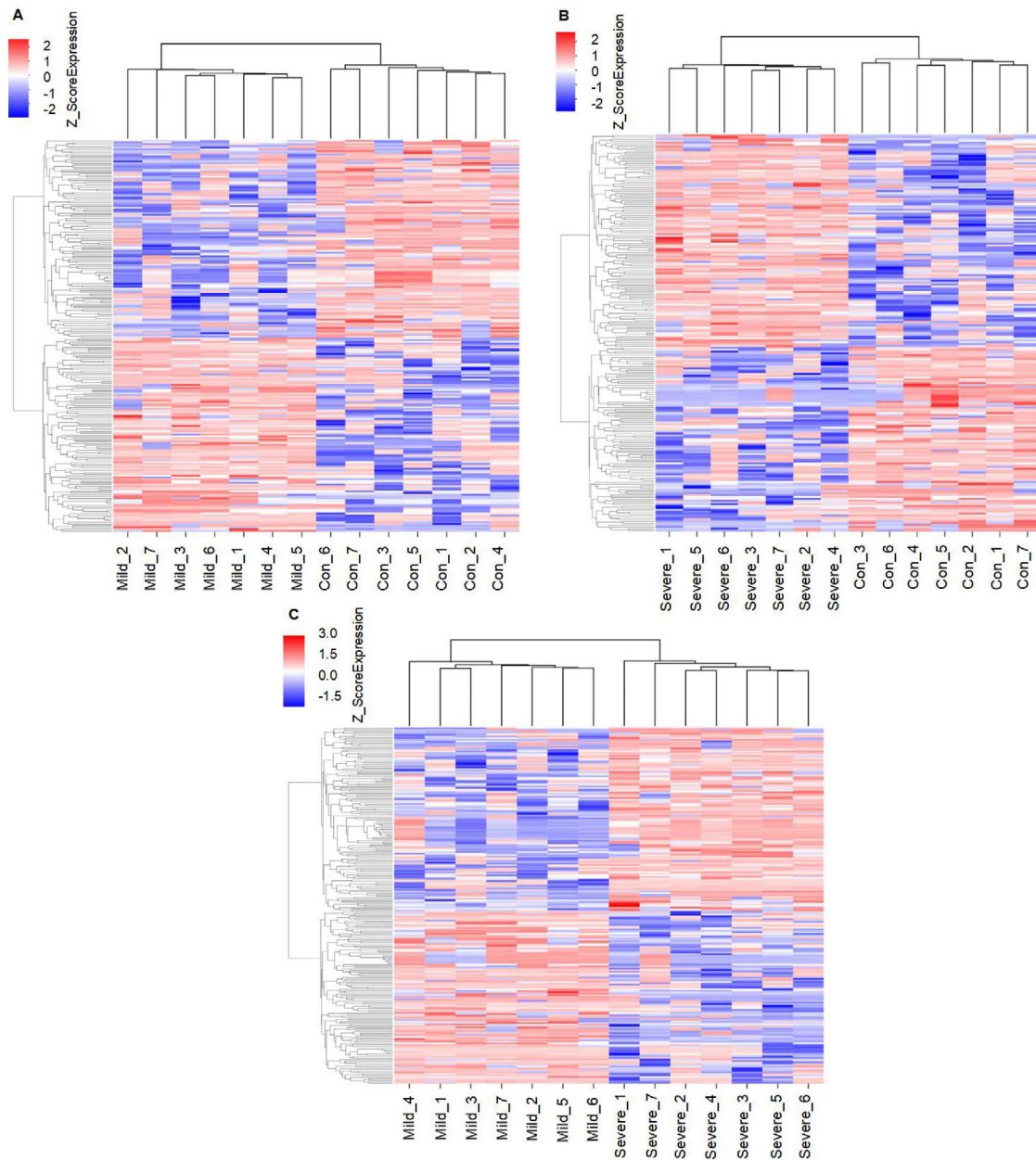


### 3.5 | GO functional enrichment analysis of differential methylated genes

To address the functional significance of the widespread changes in DNA methylation, we performed a GO analysis of all the datasets. The top 20 biologically

enriched results of the significantly enriched GO terms were highlighted (Supporting Information Table 1). Enrichment was seen in GO (Severe vs Control): peptide antigen binding ( $P = 1.5 \times 10^{-5}$ ), neuromuscular process controlling balance ( $P = 1.06 \times 10^{-4}$ ), integral component of the luminal side of endoplasmic





**FIGURE 2** Unsupervised hierarchical clustering. (A) Two different groups are represented: mild and control groups. (B) Two different groups are represented: severe and control groups. (C) Two different groups are represented: severe and mild groups. Each column represents an individual patient and each row represents an individual CpG. Z-scores of the  $\beta$ -values are shown in the heatmap

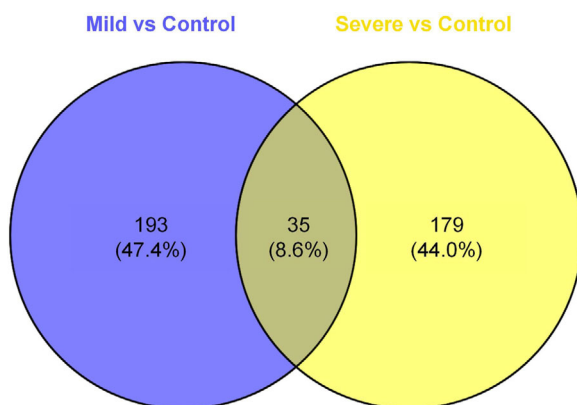
reticulum membrane ( $P = 6.09 \times 10^{-4}$ ); GO (Severe vs Mild): neuropeptide signaling pathway ( $P = 4.28 \times 10^{-4}$ ), positive regulation of interleukin-13 secretion ( $P = 9.98 \times 10^{-4}$ ), positive regulation of T-helper 2 cell differentiation ( $P = 1.592 \times 10^{-3}$ ); GO (Mild vs Control): PDZ domain binding ( $P = 2.52 \times 10^{-4}$ ), cytosol ( $P = 3.56 \times 10^{-4}$ ), high-density lipoprotein particle ( $P = 4.47 \times 10^{-4}$ ).

### 3.6 | KEGG functional enrichment analysis of differential methylated genes

The results from KEGG functional enrichment analysis showed that there were 121 (Severe vs Control), 145 (Severe vs Mild), and 183 (Mild vs Control) signal pathways involved in the three groups of differential methylation genes (DMGs); 23, 9, and 21 of which met the screening

**TABLE 2** “Markers” of SARS-CoV-2 infection

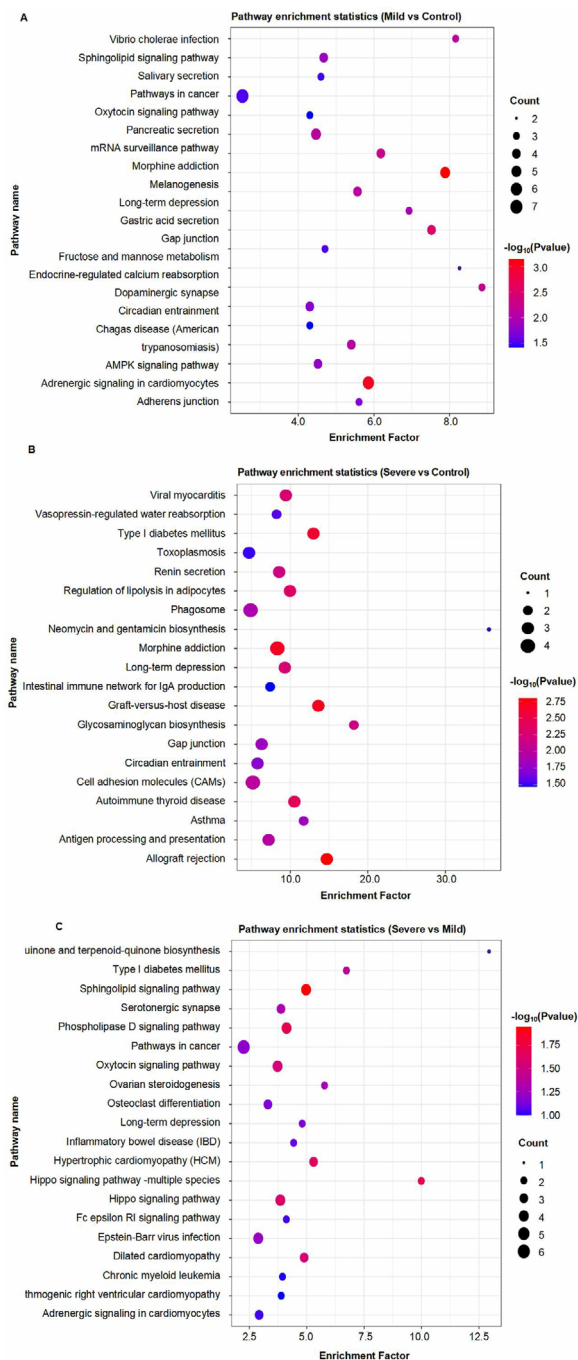
Target ID	Gene	Methylation status	UCSC ref gene group
cg19080354	ATHL1	Hypomethylated	TSS1500
cg22451923	BTNL2	Hypomethylated	Body
cg18698799	C6orf10	Hypermethylated	Body
cg12469381	CHN2	Hypomethylated	TSS200; Body
cg01825287	CHST15	Hypermethylated	5'UTR
cg264440059	COL23A1	Hypermethylated	Body
cg11128983	CPLX2	Hypomethylated	TSS1500; 5'UTR
cg03165426	CRHR2	Hypermethylated	Body;5'UTR
cg00597445	CRMP1	Hypermethylated	Body
cg25538415	DCAKD	Hypomethylated	TSS1500;5'UTR
cg09119656	EPAS1	Hypomethylated	Body
cg23836570	FBRSL1	Hypermethylated	Body
cg07733779	GNAI2	Hypermethylated	TSS1500; 5'UTR; Body
cg06096382	HECW1	Hypermethylated	TSS1500
cg27287527	HYAL1	Hypermethylated	TSS1500; 5'UTR
cg20107632	JPH3	Hypermethylated	Body
cg10644916	KLHL17	Hypermethylated	Body
cg12033072	KLRG2	Hypomethylated	Body
cg20963263	MGMT	Hypomethylated	Body
cg06875704	MIR510	Hypermethylated	TSS1500
cg02745784	PDE11A	Hypermethylated	5'UTR
cg18467790	RADIL	Hypomethylated	Body
cg03086067	SH2D4B	Hypermethylated	Body
cg07404046	SIRPB1	Hypomethylated	Body
cg03959986	SMG6	Hypermethylated	TSS1500; TSS200
cg07615678	SPON1	Hypermethylated	Body

**FIGURE 3** Venn diagram of two groups. The blue circle indicates datasets of Mild vs Control and yellow circle indicates datasets of Severe vs Control

criteria ( $P < 0.05$ ) involving 14, 16, and 25 DMGs, respectively (Figure 4).

### 3.7 | Analysis of differential methylated sites in COVID-19 groups with different severities

We compared the methylation  $\beta$ -values of CpG sites between severe and mild samples using a  $t$ -test model. Figure 5 presents a general view of the differentially methylated loci in a single chromosome. The PPI network of DMGs in the Severe vs Mild group was constructed with Cytoscape v3.0 software (Figure 6). Hub genes (including *GNG7*, *GNAZ*, *PRKCZ*, and *PRKAG2*) had the highest degree of nodes and hit more than five cytoHubba algorithms among the DMGs (Supporting Information Table 2). In addition, DMGs involved in the progression of tumors, such as *NMU* (Chen et al., 2018), were found. Moreover, *IL4R* and *IL17RA*, reported to be associated with the pathogenesis of several inflammatory disorders, were also shown to be significant in the PPI network (O'Connor et al., 2009; Massoud et al., 2016). The genes that have up-regulated or down-regulated methylation levels had



**FIGURE 4** Top 20 KEGG pathways of differential methylated genes in the three groups. (A–C) represents “Mild vs Control,” “Severe vs Control,” and “Severe vs Mild,” respectively. Top 20 enriched KEGG pathways were filtered with enrichment factors. The size and color of the dots indicate the number of differential methylated genes hit and the  $p$ -value of the enriched KEGG pathway, respectively

complex interactions with each other in the signaling network (Figure 6).

### 3.8 | Combined methylation and transcriptome analysis to validate hub genes

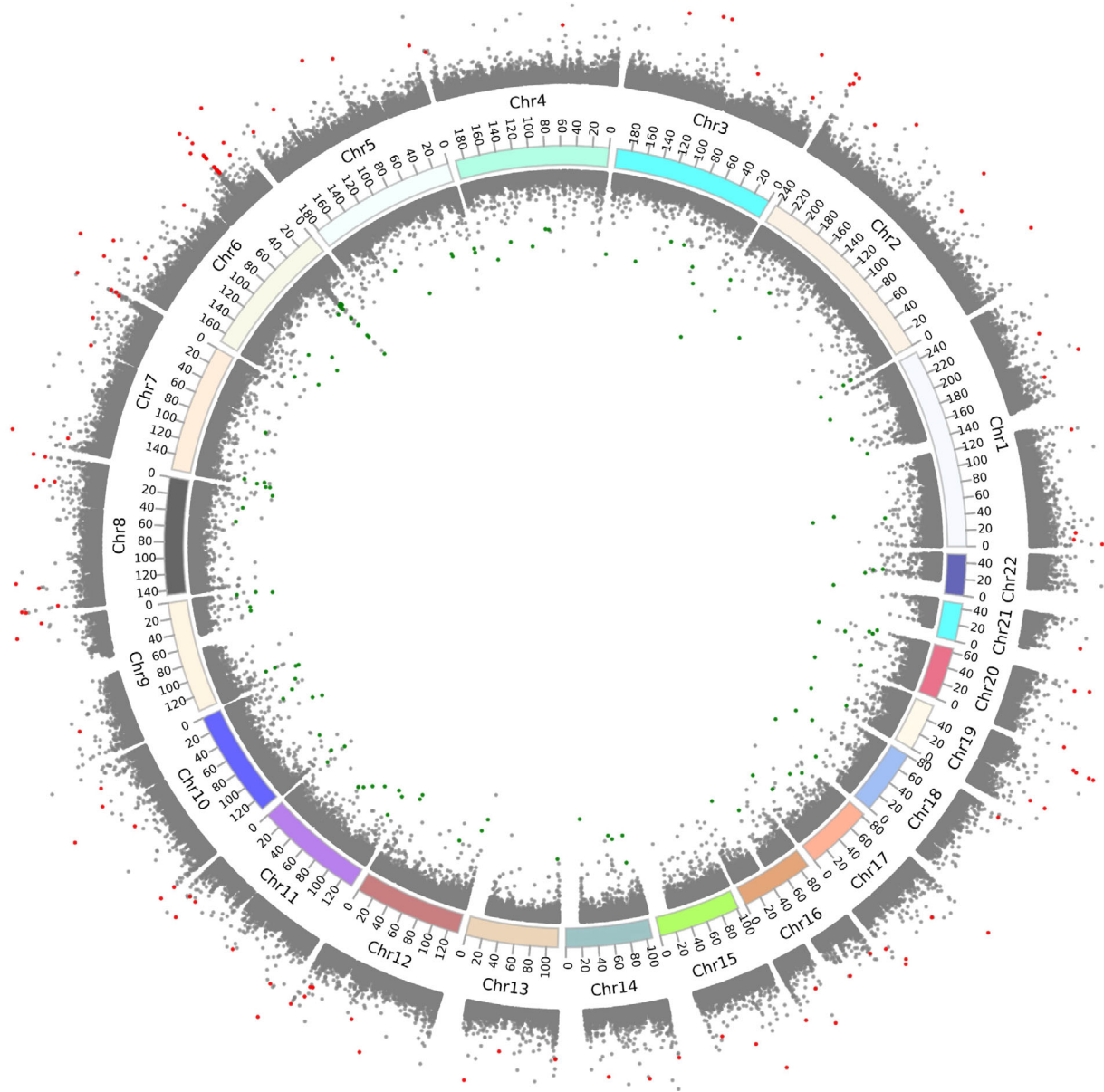
Next, we analyzed a public database GSE689128 including ten COVID-19 patients, whose COVID-19 diagnostic and severity criteria was identical to us.  $|FC| > 1.5$  and adjusted  $P$  value  $< 0.05$  were considered as criteria to screen the differential expression (DE)-mRNAs between severe group and mild group. As shown in Supporting Information Figure S1, 14132 DE-mRNAs, including 8056 downregulated and 6076 upregulated mRNA, were identified in the GSE164805 profile. For four hub genes, we found that the expression of *GNAS* gene (adjusted  $P$  value  $< 0.01$ ) and *GNG7* gene (adjusted  $P$  value = 0.026) were both downregulated in the severe group, and the mRNA expression of the other two hub genes had no significant difference between the two groups (adjusted  $P$  value  $> 0.05$ ) (Table 3).

## 4 | DISCUSSION

At present, there is a lack of research that provides evidence of the COVID-19 pathogenesis pathway by epigenetic regulation. Epigenetic disorders might favor viral infection and provide useful biomarkers to stratify disease severity in patients more prone to disseminating SARS-CoV-2 infection (Crimi et al., 2020). In this study, the Illumina 850k methylation chip was used to detect the changes of DNA methylation expression in whole blood samples from a cohort of COVID-19 patients and subjects with no clinically detectable SARS-CoV-2 infection. This study laid a foundation for finding novel SARS-CoV-2 infection “markers” and therapeutic targets. The hub genes and signal pathways identified also point to the need to study the COVID-19 regulation mechanism.

Through the Venn diagram analysis, we combined the DMGs of the “Severe versus Control” and “Mild versus Control” groups and found 35 “marker” genes that acted as candidate genes that may play an important role in indicating SARS-CoV-2 infection. Previous studies have shown that the methylation status in the promoter region (including 5’UTR, TSS200, and TSS1500) of genes can regulate gene expression, and studies based on DNA methylation have also focused on these regions (Bayarsaihan, 2011; Bierne et al., 2012). In these 35 candidate genes, differential methylation sites of 12 genes

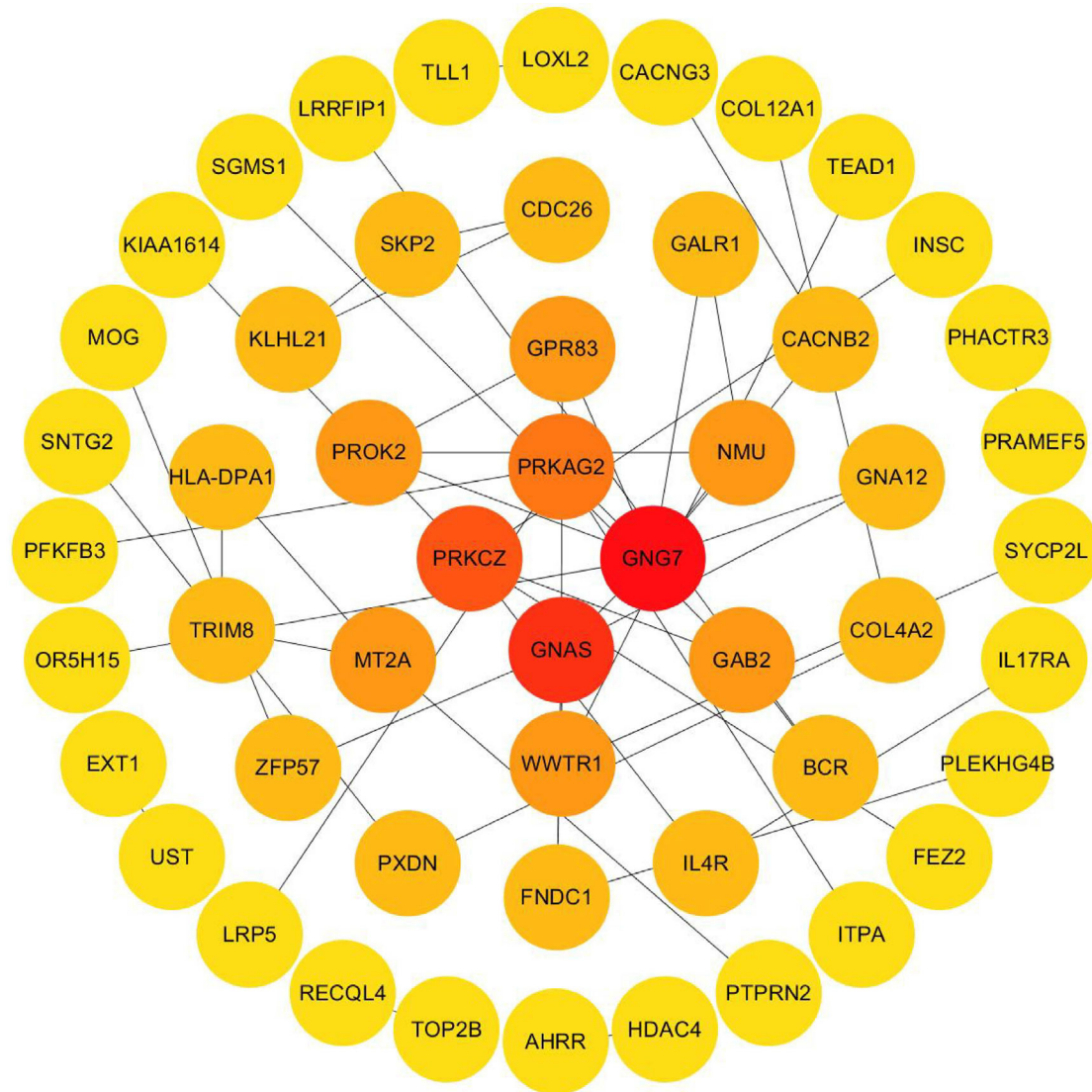




**FIGURE 5** Circos plot of genome-wide DNA methylation changes between the severe and mild group. Red dots denote hypermethylated sites ( $\beta > 0.17$ ), green dots denote hypomethylated sites ( $\beta < -0.17$ ), and gray dots denote sites  $|\beta| < 0.17$  or  $|\text{diffScore}| < 13$

**TABLE 3** mRNA expression analysis of four hub genes (severe vs mild)

Gene symbol	Delta Beta (methylation difference)	DiffScore (methylation difference)	Fold change (mRNA expression)	P value (mRNA expression)	adjusted P value (mRNA expression)
GNAS	-0.23643	-14.44971	-2.6818	0.00033	0.00585
GNG7	-0.36106	-15.92767	-2.02017	0.00446	0.02627
PRKAG2	-0.26721	-15.82896	1.61695	0.02680	0.08674
PRKCZ	0.28300	15.68705	1.33840	0.36300	0.49675



**FIGURE 6** PPI analysis. Protein interactions between the differential methylated genes. Darker colors indicate higher scores that represent more network involvement and significance

(*ATHL1*, *CHN2*, *CHST15*, *CPLX2*, *CRHR2*, *DCAKD*, *GNAI2*, *HECW1*, *HYAL1*, *MIR510*, *PDE11A*, and *SMG6*) were situated in the promoter region. Among them, the molecular functions of *ATHL1*, *CHN2*, *DCAKD*, *MIR510*, and *HECW1* are largely unexplored. Emerging evidence suggests that *CHST15* involves multiple tumor progression and metastasis (Liu et al., 2019). *CPLX2* regulates the formation of synaptic vesicles and acts as an inhibitor of secretion of some autoimmune antibodies, thus it may participate in the pathogenesis of some related autoimmune diseases (Tsuru et al., 2019). Both *CRHR2* and *PDE11A* play roles in signal transduction by regulating the intracellular concentration of cyclic nucleotide cAMP and are reported to be associated with many human cancers (Gu et al., 2010; Koch, 2011). *GNAI2* can be upregulated by NF- $\kappa$ B activation and is associated with sponta-

neous colitis and colitis-associated cancer (Li et al., 2019). *HYAL1* catalyzes the hydrolyzation of 20 kDa HA into low-molecular-weight HA fragments with pathophysiological effects including angiogenic, inflammatory, and immunosuppressive effects. In a recent study, researchers revealed overexpression of *HYAL1* could prevent fibroblast fibroproliferation. They also showed significant downregulation of *HYAL1* in idiopathic pulmonary fibrosis lung tissues (Leng et al., 2020). As an RNA surveillance protein, *SMG6* could affect HIV-1 gene expression via the influence of viral reactivation at a posttranscriptional level (Rao et al., 2018). It has been suggested that elucidation of the role of *SMG6* may aid in explaining the maintenance of SARS-CoV-2 persistence. In addition to these 12 genes, some other genes whose differential methylation sites were identified in the gene body may also play crucial roles,

such as *BTNL2*, which is involved in multiple inflammatory and autoimmune diseases (Zhao et al., 2020). Whether these DMGs, can be used as new biomarkers of SARS-CoV-2 infection need further study. The function of these DMGs and their role in SARS-CoV-2 infection need to be explored. Further study may provide new direction and strategies for specific treatments and early diagnosis of COVID-19.

Function enrichment and pathway enrichment analyses were conducted in three groups. Among these, the results regarding the “Severe versus Mild” group may be of great value. GO pathway analysis showed that, in the “Severe versus Mild” group, the functions of the DMGs were mainly focused on neuropeptides, interleukin-13, chemokine secretion, T-helper 2 cell differentiation, and mast cell degranulation, indicating that the activation of inflammatory and allergic reactions are closely related to the progression of COVID-19. In the KEGG analysis, DMGs were involved in the sphingolipid signaling pathway, hippo signaling pathway, phospholipase D signaling pathway, hypertrophic cardiomyopathy (HCM), oxytocin signaling pathway, dilated cardiomyopathy, type I diabetes mellitus, and serotonergic synapses. It has been suggested that the occurrence and development of COVID-19 is caused by the joint action of multiple molecular biological signal pathways. Sphingosine has been confirmed to prevent the binding of SARS-CoV-2 spikes to its cellular receptor ACE2 (Edwards et al., 2020). Its derivatives also act as a selective and potent candidate for COVID-19 treatment (Dave et al., 2020). In addition, sphingosine 1-phosphate mediated signaling was suggested to be a potential target for ameliorating neurological symptoms induced by SARS-CoV-2 (Meacci et al., 2020). In a recent study, oxytocin was shown to have anti-inflammatory and proimmune functions, and they act as a potential defense against COVID-19 (Thakur et al., 2020). As for HCM, Bos et al. (2020) extracted RNA from cardiac tissue of patients with HCM and healthy donor hearts. Their results showed a 5-fold increase in ACE2 protein in patients with HCM compared to healthy donors, which may lead to an increased risk for COVID-19 outcomes. In line with the study by Bos et al., a study by Chen, Li et al. yielded similar results. They found patients with basic heart failure disease (including HCM and dilated cardiomyopathy) exhibited increased expression of ACE2, and might have a high possibility of heart attack and progression to a severe condition after SARS-CoV-2 infection (Chen, Li et al., 2020). However, the mechanisms of hippo signaling, phospholipase D, and serotonergic synapse signaling pathway in SARS-COV-2 infection are still not clear, which may be new directions for COVID-19 research.

In PPI network analysis, we identified hub genes, which consisted of the four DMGs with the high-

est degrees of interaction. Combined methylation and transcriptome analysis further highlights the significance of *GNG7* and *GNAS2* during SARS-CoV-2 infection. Among these four hub genes, guanine nucleotide-binding protein  $\gamma$ -7 (*GNG7*) is a subunit of heterotrimeric G protein, is ubiquitously expressed in multiple tissues, and is downregulated in various cancers (Liu et al., 2016). It has also been confirmed to be highly methylated in tumors and unmethylated in normal tissue, and it stands as a highly specific promoter methylated gene associated with some cancers (Demokan et al., 2013). Through a methylated DNA immunoprecipitation chip, researchers showed that protein kinase C zeta (*PRKCZ*) is hypermethylated in type 2 diabetes mellitus and the hypermethylation *PRKCZ* gene may be involved in the pathogenesis of type 2 diabetes mellitus (Zou et al., 2013). The methylation level of *PRKAG2* also correlated with insulin resistance, arterial blood pressure, and respiratory diseases (Ramos-Lopez et al., 2018; Lee et al., 2019; Olloquequi, 2020). Noteworthy, patients with the aforementioned conditions, such as elevated glucose levels, cancer, and respiratory diseases, likely favor SARS-CoV-2 infection and increase COVID-19 severity (Dai et al., 2020; Zhu et al., 2020). Until now, no previous study assessed whether a relationship exists between changes in DNA methylation levels in *GNAS2* and some cellular processes. The specific function of *GNAS2* remains to be determined.

There are several limitations in this study. First, given that differences between “severe patients” and “mild patients” may be subtle, the sample size of our study is limited. For EPIC array DNA methylation studies according to the latest research (Mansell et al., 2019), using a threshold of *P* value less than  $9 \times 10^{-8}$  to screen differential methylation site may be more appropriate. As such, we could have overestimated the impact of some methylation sites, and further exploration on this issue and researches with large sample size were needed. Another limitation was that the methylation levels of COVID-19 patients might be dynamic. Thus, longitudinal studies will be needed to examine this hypothesis more rigorously. Additionally, differences of methylation level between asymptomatic and symptomatic patients should also be compared. In the context of a global epidemic, more in-depth studies are urgently needed to clarify the complete mechanism of SARS-CoV-2 infection.

In conclusion, this study used an Illumina 850k methylation chip to screen genes with different methylation levels in peripheral blood, further studying the relationship between epigenetic regulation and the pathogenesis of COVID-19. Given that there are no effective therapeutic agents available for the treatment of COVID-19, this study provided a new idea for the diagnosis and therapy of COVID-19.



## ACKNOWLEDGEMENTS

We appreciate Lin Liu (Division of Medical Administration, Health Commission of Shandong Province) for providing assistance.

## CONFLICT OF INTERESTS

The authors declare that they have no competing interests.

## FUNDING

This work was supported by grants from the National Natural Science Foundation of China (81770029) and Key Research Project of Shandong Province (2020SFXGFY03-1, 2020SFXGFY03-3, 2020SFXGFY04). The funders had no role in the study design, data collection, data analysis or preparation of the manuscript.

## DATA AVAILABILITY STATEMENT

The data used and analyzed during this research are available from the corresponding author on reasonable request.

## ETHICS APPROVAL

This study was approved by the ethics committee of Qilu Hospital of Shandong University (ethical approval no: KYLL-2020-GXB-001).

## CONSENT FOR PUBLICATION

Written informed consent for publication was obtained from all participants.

## AUTHORS CONTRIBUTIONS

LD and SZ conceived and designed the study, had full access to all of the data in the study, and take responsibility for the integrity of the data and the accuracy of the data analysis. JZ, JX, FZ, and LP drafted the paper and performed the analysis. YH, JW, CW, XW, WZ, YL, KN, YP, TL, JZ, LY, RZ, FG, JZ, and XZ collected the key data and critically reviewed the article before submitting it. All authors made substantial contributions to this manuscript and agree to be accountable for all aspects of the work.

## ORCID

Jintao Zhang  <https://orcid.org/0000-0002-1309-6836>

## REFERENCES

- Bayarsaihan, D. (2011). Epigenetic mechanisms in inflammation. *Journal of Dental Research*, 90(1), 9–17. <https://doi.org/10.1177/0022034510378683>
- Bierne, H., Hamon, M., & Cossart, P. (2012). Epigenetics and bacterial infections. *Cold Spring Harbor Perspectives in Medicine*, 2(12), a010272. <https://doi.org/10.1101/cshperspect.a010272>
- Bos, J., Hebl, V., Oberg, A., Sun, Z., Herman, D., Teekakirikul, P., Bos, J. M., Hebl, V. B., Oberg, A. L., Sun, Z., Herman, D. S., Teekakirikul, P., Seidman, J. G., Seidman, C. E., Dos Remedios, C. G., Maleszewski, J. J., Schaff, H. V., Dearani, J. A., Noseworthy, P. A., Friedman, P. A., Ommen, S. R., Brozovich, F. V., & Ackerman, M. (2020). Marked up-regulation of ACE2 in hearts of patients with obstructive hypertrophic cardiomyopathy: Implications for SARS-CoV-2-Mediated COVID-19. *Mayo Clinic Proceedings*, 95(7), 1354–1368. <https://doi.org/10.1016/j.mayocp.2020.04.028>
- Bost, P., Giladi, A., Liu, Y., Bendjelal, Y., Xu, G., David, E., Blecher-Gonen, R., Cohen, M., Medaglia, C., Li, H., Deczkowska, A., Zhang, S., Schwikowski, B., Zhang Z., & Amit, I. (2020). Host-viral infection maps reveal signatures of severe COVID-19 Patients. *Cell*, 181(7), 1475–1488.e1412. <https://doi.org/10.1016/j.cell.2020.05.006>
- Chai, P., Yu, J., Ge, S., Jia, R., & Fan, X. (2020). Genetic alteration, RNA expression, and DNA methylation profiling of coronavirus disease 2019 (COVID-19) receptor ACE2 in malignancies: a pan-cancer analysis. *Journal of Hematology & Oncology*, 13(1), 43. <https://doi.org/10.1186/s13045-020-00883-5>
- Chen, L., Li, X., Chen, M., Feng, Y., & Xiong, C. (2020). The ACE2 expression in human heart indicates new potential mechanism of heart injury among patients infected with SARS-CoV-2. *Cardiovascular Research*, 116(6), 1097–1100. <https://doi.org/10.1093/cvr/cvaa078>
- Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., . . . , Zhang, L. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet (London, England)*, 395(10223), 507–513. [https://doi.org/10.1016/s0140-6736\(20\)30211-7](https://doi.org/10.1016/s0140-6736(20)30211-7)
- Chen, X., Ma, L., & Xu, J. (2018). Abnormal DNA methylation may contribute to the progression of osteosarcoma. *Molecular Medicine Reports*, 17(1), 193–199. <https://doi.org/10.3892/mmr.2017.7869>
- Chlamydas, S., Papavassiliou, A., & Piperi, C. (2020). Epigenetic mechanisms regulating COVID-19 infection. *Epigenetics*, 16(3). <https://doi.org/10.1080/15592294.2020.1796896>
- Corley, M. J., & Ndhlovu, L. C. (2020). DNA methylation analysis of the COVID-19 host cell receptor, angiotensin I converting enzyme 2 gene (ACE2) in the respiratory system reveal age and gender differences. *Preprints*. <https://doi.org/10.20944/preprints202003.0295.v1>
- Crimi, E., Benincasa, G., Figueroa-Marrero, N., Galdiero, M., & Napoli, C. (2020). Epigenetic susceptibility to severe respiratory viral infections: pathogenic and therapeutic implications: a narrative review. *British Journal of Anaesthesia*, <https://doi.org/10.1016/j.bja.2020.06.060>
- Dai, M., Liu, D., Liu, M., Zhou, F., Li, G., Chen, Z., . . . , & Cai, H. (2020). Patients with cancer appear more vulnerable to SARS-CoV-2: A multicenter study during the COVID-19 outbreak. *Cancer Discovery*, 10(6), 783–791. <https://doi.org/10.1158/2159-8290.cd-20-0422>
- Dave, G., Rakholiya, K., Kaneria, M., Galvadiya, B., Vyas, S., Kanbi, V., & Patel, M. (2020). High affinity interaction of *Solanum tuberosum* and *Brassica juncea* residue smoke water compounds with proteins involved in coronavirus infection. *Phytotherapy Research: PTR*, 34(12), 3400–3410. <https://doi.org/10.1002/ptr.6796>
- Demokan, S., Chuang, A., Chang, X., Khan, T., Smith, I., Pattani, K., . . . , & Califano, J. (2013). Identification of guanine nucleotide-binding protein  $\gamma$ -7 as an epigenetically silenced gene in head and neck cancer by gene expression profiling. *International journal of oncology*, 42(4), 1427–1436. <https://doi.org/10.3892/ijo.2013.1808>

- Edwards, M., Becker, K., Gripp, B., Hoffmann, M., Keitsch, S., Wilker, B., ..., & Gulbins, E. (2020). Sphingosine prevents binding of SARS-CoV-2 spike to its cellular receptor ACE2. *The Journal of Biological Chemistry*, 295(45). <https://doi.org/10.1074/jbc.RA120.015249>
- Gu, X., Qi, P., Zhou, F., Ji, Q., Wang, H., Dou, T., ..., & Gao, C. (2010). An intronic polymorphism in the corticotropin-releasing hormone receptor 2 gene increases susceptibility to HBV-related hepatocellular carcinoma in Chinese population. *Human genetics*, 127(1), 75–81. <https://doi.org/10.1007/s00439-009-0750-6>
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., ..., & Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, 395(10223), 497–506. [https://doi.org/10.1016/s0140-6736\(20\)30183-5](https://doi.org/10.1016/s0140-6736(20)30183-5)
- Koch, L. (2011). Cancer: PDE11a—a phenotype modifier in carney complex. *Nature Reviews. Endocrinology*, 7(2), 64. <https://doi.org/10.1038/nrendo.2010.225>
- Kuss-Duerkop, S. K., Westrich, J. A., & Peyer, D. (2018). DNA tumor virus regulation of host DNA methylation and its implications for immune evasion and oncogenesis. *Viruses*, 10(2). <https://doi.org/10.3390/v10020082>
- Lee, S., Hwang, H., Hsu, P., Chuang, T., Liu, C., & Wu, L. (2019). Whole-genome methylation profiling from PBMCs in acute-exacerbation COPD patients with good and poor responses to corticosteroid treatment. *Genomics*, 113(6), 1381–1386. <https://doi.org/10.1016/j.ygeno.2018.09.010>
- Leng, D., Huang, X., Yi, J., Zhao, H., & Zhang, Y. (2020). HYAL1 is downregulated in idiopathic pulmonary fibrosis and inhibits HFL-1 fibroblast proliferation when upregulated. *BioMed research international*, 2020, 3659451. <https://doi.org/10.1155/2020/3659451>
- Li, Z. W., Sun, B., Gong, T., Guo, S., Zhang, J., Wang, J., ..., & Chu, W. M. (2019). GNAI1 and GNAI3 reduce colitis-associated tumorigenesis in mice by blocking IL6 signaling and down-regulating expression of GNAI2. *Gastroenterology*, 156(8), 2297–2312. <https://doi.org/10.1053/j.gastro.2019.02.040>
- Liu, J., Ji, X., Li, Z., Yang, X., Wang, W., & Zhang, X. (2016). G protein  $\gamma$  subunit 7 induces autophagy and inhibits cell division. *Oncotarget*, 7(17), 24832–24847. <https://doi.org/10.18632/oncotarget.8559>
- Liu, L., Wang, Y., Lin, P., Zhang, X., Cheng, W., Liu, S., ..., & Wang, S. (2019). Long noncoding RNA HOTAIR promotes invasion of breast cancer cells through chondroitin sulfotransferase CHST15. *International Journal of Cancer*, 145(9), 2478–2487. <https://doi.org/10.1002/ijc.32319>
- Mansell, G., Gorrie-Stone, T., Bao, Y., Kumari, M., Schalkwyk, L., Mill, J., & Hannon, E. (2019). Guidance for DNA methylation studies: statistical insights from the Illumina EPIC array. *BMC Genomics [Electronic Resource]*, 20(1), 366. <https://doi.org/10.1186/s12864-019-5761-7>
- Massoud, A., Charbonnier, L., Lopez, D., Pellegrini, M., Phipatanakul, W., & Chatila, T. (2016). An asthma-associated IL4R variant exacerbates airway inflammation by promoting conversion of regulatory T cells to TH17-like cells. *Nature Medicine*, 22(9), 1013–1022. <https://doi.org/10.1038/nm.4147>
- Meacci, E., Garcia-Gil, M., & Pierucci, F. (2020). SARS-CoV-2 infection: A role for S1P/S1P receptor signaling in the nervous system? *International Journal of Molecular Sciences*, 21(18). <https://doi.org/10.3390/ijms21186773>
- O'Connor, W., Kamanaka, M., Booth, C., Town, T., Nakae, S., Iwakura, Y., ..., & Flavell, R. (2009). A protective function for interleukin 17A in T cell-mediated intestinal inflammation. *Nature immunology*, 10(6), 603–609. <https://doi.org/10.1038/ni.1736>
- Olloquequi, J. (2020). COVID-19 Susceptibility in chronic obstructive pulmonary disease. *European Journal of Clinical Investigation*, 50(10), e13382. <https://doi.org/10.1111/eci.13382>
- Pruimboom, L. (2020). Methylation pathways and SARS-CoV-2 lung infiltration and cell membrane-virus fusion are both subject to epigenetics. *Frontiers in cellular and infection microbiology*, 10, 290. <https://doi.org/10.3389/fcimb.2020.00290>
- Ramos-Lopez, O., Samblas, M., Milagro, F., Riezu-Boj, J., Crujeiras, A., Martinez, J., & Project, M. (2018). Circadian gene methylation profiles are associated with obesity, metabolic disturbances and carbohydrate intake. *Chronobiology International*, 35(7), 969–981. <https://doi.org/10.1080/07420528.2018.1446021>
- Rao, S., Amorim, R., Niu, M., Temzi, A., & Moulard, A. J. (2018). The RNA surveillance proteins UPF1, UPF2 and SMG6 affect HIV-1 reactivation at a post-transcriptional level. *Retrovirology*, 15(1), 42. <https://doi.org/10.1186/s12977-018-0425-2>
- Sawalha, A., Zhao, M., Coit, P., & Lu, Q. (2020). Epigenetic dysregulation of ACE2 and interferon-regulated genes might suggest increased COVID-19 susceptibility and severity in lupus patients. *Clinical immunology (Orlando, Fla.)*, 215, 108410. <https://doi.org/10.1016/j.clim.2020.108410>
- Silmon de Monerri, N. C., & Kim, K. (2014). Pathogens hijack the epigenome: a new twist on host-pathogen interactions. *American Journal of Pathology*, 184(4), 897–911. <https://doi.org/10.1016/j.ajpath.2013.12.022>
- Thakur, P., Shrivastava, R., & Shrivastava, V. (2020). Oxytocin as a potential adjuvant against COVID-19 infection. *Endocrine, Metabolic & Immune Disorders Drug Targets*. 20. <https://doi.org/10.2174/1871530320666200910114259>
- Tsuru, E., Oryu, K., Sawada, K., Nishihara, M., & Tsuda, M. (2019). Complexin 2 regulates secretion of immunoglobulin in antibody-secreting cells. *Immun Inflamm Dis*, 7(4), 318–325. <https://doi.org/10.1002/iid3.276>
- Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., ..., & Peng, Z. (2020). Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *Jama*, 323(11), 1061–1069. <https://doi.org/10.1001/jama.2020.1585>
- Yang, X., Yu, Y., Xu, J., Shu, H., Xia, J., Liu, H., ..., & Shang, Y. (2020). Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *The Lancet. Respiratory medicine*, 8(5), 475–481. [https://doi.org/10.1016/s2213-2600\(20\)30079-5](https://doi.org/10.1016/s2213-2600(20)30079-5)
- Zhao, Y., Zheng, Q., & Jin, L. (2020). The role of B7 family molecules in maternal-fetal immunity. *Frontiers in immunology*, 11, 458. <https://doi.org/10.3389/fimmu.2020.00458>
- Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., ..., & Cao, B. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet (London, England)*, 395(10229), 1054–1062. [https://doi.org/10.1016/s0140-6736\(20\)30566-3](https://doi.org/10.1016/s0140-6736(20)30566-3)
- Zhu, L., She, Z., Cheng, X., Qin, J., Zhang, X., Cai, J., ..., & Li, H. (2020). Association of blood glucose control and outcomes



in patients with COVID-19 and pre-existing type 2 diabetes. *Cell metabolism*, 31(6), 1068–1077.e1063. <https://doi.org/10.1016/j.cmet.2020.04.021>

Zill, P., Baghai, T. C., Schüle, C., Born, C., Frustuck, C., Butner, A., ..., & Bondy, B. (2012). DNA methylation analysis of the angiotensin converting enzyme (ACE) gene in major depression. *Plos One*, 7(7). <https://doi.org/10.1371/journal.pone.0040479>

Zou, L., Yan, S., Guan, X., Pan, Y., & Qu, X. (2013). Hypermethylation of the PRKCZ gene in Type 2 diabetes mellitus. *Journal of diabetes research*, 2013, 721493. <https://doi.org/10.1155/2013/721493>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Zhou S, Zhang J, Xu J, et al. An epigenome-wide DNA methylation study of patients with COVID-19. *Ann Hum Genet.* 2021;85:221–234. <https://doi.org/10.1111/ahg.12440>