JIDC | THE JOURNAL OF INFECTION IN DEVELOPING COUNTRIES

Coronavirus Pandemic

Detection of the DNA methylation of seven genes contributes to monitoring recovery from COVID-19

Xiao Xiao[1], Ren Xu[1], Chaoxiang Du[2,3], Jun Yin[2], Beibei Xin[4], Zhonghe Ke[4], Xiyan Li[4], Hao Zhang[5], Xinyu Chen[1]

[1] *School of Physics, Changchun University of Science and Technology, Changchun 130022, China*
[2] *Department of Thoracic Surgery, Cancer Center, Zhongshan Hospital of Fudan University, Shanghai, 200030, China*
[3] *Zhongshan Hospital (Xiamen), Fudan University, Xiamen, 361015, China*
[4] *Shanghai Rightongene Biotechnology Co Ltd, Shanghai 201403, China*
[5] *School of Life Science and Technology, Changchun University of Science and Technology, Changchun 130022, China*

Abstract

Introduction: DNA methylation might influence the expression of genes that regulate coronavirus disease 2019 (COVID-19) progression. This work explored the significance of DNA methylation of 7 genes (*TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, *RASSF1A*, and *SHOX2*) in blood circulating free DNA (cfDNA) in differentiating COVID-19 infections and recoveries. The correlation with changes in the proportion of immune cell populations in the recovery period was analyzed.

Methodology: 18 COVID-19-infected, 65 COVID-19-recovered, and 11 uninfected individuals were included. DNA methylation expression was determined by quantitative multiplex methylation-specific PCR (qMSP). The immune function of the recovered group versus uninfected group was evaluated by full-spectrum flow cytometry.

Results: The infected population showed a higher methylation positivity rate for 7 genes compared to the uninfected/recovered population. A model was constructed to distinguish the infected patients from uninfected/recovered individuals using the methylation status of 7 genes, with a sensitivity, specificity and area under curve (AUC) of 0.889, 0.842 and 0.931, respectively. The results of flow cytometry revealed that CD8+ T cells and CD38+ CD8+ T cells were significantly upregulated in recovered individuals compared to those in uninfected individuals. DNA methylation was correlated with immune cell changes, with a significant increase in the percentage of T cells, PD-1+ function CD4+ T cells, TCRγδ+ cells, and CD38+ NKT cells upon an increase in 7 gene methylation positivity.

Conclusions: This work revealed the significance of 7 gene methylation in the diagnosis of COVID-19 recovery, and demonstrated that these genes were significant in evaluating the immune function during the recovery period.

Key words: COVID-19; DNA methylation; flow cytometry; recovery.

J Infect Dev Ctries 2026; 20(7):944-953. doi:10.3855/jidc.21864

(Received 20 May 2025 – Accepted 17 August 2025)

Introduction

The severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) emerged in Wuhan China at the end of 2019, and triggered a rare outbreak of viral pneumonia [1,2]. This novel coronavirus disease, also known as coronavirus disease 2019 (COVID-19), was highly contagious and spread rapidly around the world [3,4]. According to the World Health Organization (WHO), as of November 2023, there were more than 772 million confirmed cases of COVID-19, and the pandemic claimed nearly 7 million lives [5].

SARS-CoV-2 detection technologies mainly target either specific viral nucleic acids (molecular testing), proteins (antigen testing), or anti-SARS-CoV-2 antibodies (serological testing) [6]. Real-time reverse transcription polymerase chain reaction (RT-PCR) is the most common method, relying on detecting the presence of specific viral RNA. However, viral mutations can potentially lead to unpredictable test performances and false-negatives [7]. Additionally, antigen testing is also commonly used to detect SARS-CoV-2, yet its sensitivity lags behind molecular testing [8]. Notably, the success rate of these tests depends largely on factors such as viral load. Specifically, these tests can reliably detect SARS-CoV-2 in the clinical specimens with high viral load (i.e., typically from 1–3 days before, to 5–7 days after symptom onset), whereas the likelihood of SARS-CoV-2 detection declines in the samples with a low viral load [9]. Thus, these methods may have limited ability to distinguish patients with low SARS-CoV-2 load and those without it (uninfected healthy individuals and recovered patients).

Interestingly, accumulating evidence suggests that epigenetics can reveal the underlying biology and can improve early diagnosis and precision medicine in infectious diseases [10,11]. Specifically, epigenetics is the science of gene expression without nucleotide sequence alteration [12]. Notably, the epigenetic modifications have been found in SARS-CoV-2 and

other viral infections [13,14], and, these mechanisms appear to be an important component of COVID-19 pathophysiology and disease severity [13,15]. Importantly, many processes affect epigenetic expression, including gene ubiquitination, histone acetylation, and especially DNA methylation [16]. DNA methylation is the earliest recognized, as well as the most deeply studied and the most important epigenetic modification mode in mammals [17]. It selectively promotes or silences the expression of certain genes, leading to phenotypic changes [17]. Recent studies have discovered that DNA methylation in COVID-19 can elucidate pathogenic mechanisms, modulate disease progression, predict clinical outcomes, and identify potential therapeutic targets [18,19]. For example, Zhou *et al.* suggested that DNA methylation might influence the expression of genes that regulate COVID-19 progression [24]. Arnold *et al.* discovered that differences in methylation patterns may distinguish asymptomatic from symptomatic infection [26]. Balnis *et al*. revealed that the prevalent hyper-methylated status in blood samples was associated with worse outcomes in COVID-19-positive patients [23]. Notably, previous studies have indicated that the methylations of *TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, *RASSF1A*, and *SHOX2* serve as excellent molecular diagnostic and prognostic biomarkers in multiple diseases, especially cancers [20,21]. For example, the detection of *RASSF1A* and *SHOX2* methylation has higher sensitivity and specificity than other commonly used methods for diagnosing lung cancer, especially in the early stage [22]. It was speculated that the methylations of these genes may have the potential as biomarkers for distinguishing SARS-CoV-2 infection status. However, to date, there is no clear report on the clinical utility of DNA methylation of these genes in blood for detecting the COVID-19 infection status.

In addition, increasing evidence suggests that the immune system is intimately involved in the process of COVID-19 infection [23]. One of the main pathogenetic mechanisms responsible for COVID-19 attack is the potent suppression of innate immunity [24]. For example, after infection, activated B, CD4+ T, and CD8+ T lymphocytes respond to viral entry and then generate protective immunity [25]. During the inflammatory response, the amounts of NK and T lymphocytes decrease, whereas proinflammatory cytokine levels, as well as the amounts of T cell activation and exhaustion are increased following COVID-19 infection [26]. A study has shown that alterations in lymphocyte subsets correlate with markers of COVID-19 severity. The researchers observed an increase in certain lymphocytes as the disease progressed, leading to decreased levels of C-reactive protein, while potentially affecting COVID-19 recovery and immune response homeostasis [27]. Importantly, it has been recently revealed that DNA methylation plays an important role in the regulation of immune system in COVID-19 infections [28,29].

In this study, the performance of DNA methylation of 7 genes, including *TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, *RASSF1A*, and *SHOX2*, was investigated in distinguishing COVID-19-infected versus recovered and uninfected populations, to reveal the relationship between immune cell subpopulations and the DNA methylation that occurs during the recovery phase of COVID-19.

Methodology

Sample collection

Blood samples from 94 individuals were collected from July 2022 to December 2022 for inclusion in this study. The inclusion criteria are as follows: a) age ≥ 18 years; b) undergoing SARS-CoV-2 nucleic acid testing by RT-PCR; c) not suffering from immunological diseases or receiving immunosuppressive medications. Among the study participants, 11 were not infected with COVID-19 (defined as those who tested persistently negative for SARS-CoV-2 by RT-PCR before and after enrollment), 65 individuals had recovered from COVID-19 infection within 30 days (defined as those who tested positive for SARS-CoV-2 by RT-PCR for no more than 30 days, and subsequently converted to negative prior to enrollment), and 18 individuals were in the COVID-19 infection stage (defined as those who tested persistently positive for SARS-CoV-2 by RT-PCR before and after enrollment).

The research was conducted in accordance with the ethical standards set out in the Declaration of Helsinki. This study was approved by the Ethical Committee of Zhongshan Hospital Affiliated to Fudan University (No. B2021-715). All participants provided their informed consent to participate.

DNA isolation and quantitative multiplex methylation-specific PCR (qMSP)

The cfDNA was obtained from blood samples using the Plasma Free DNA Extraction Kit (Yuanqi Biomedical, Shanghai, China) according to the manufacturer's instructions. The methylation positive levels of the 7 genes were characterized using a DNA methylation kit (Yuanqi Biomedical, Shanghai, China). Briefly, 10 μL of DNA was bisulfite-converted using a

bisulfite solution according to the manufacturer's instructions. Subsequently, 2 μL of pre-amplified primer mix and 13 μL of PCR master mix from the above kit were mixed with 10 μL of bisulfite-converted DNA to perform the pre-amplification procedure. Next, 160 μL of water was added to the pre-amplified product for dilution, and 5 μL of it was mixed with 7 μL of fluorescently labeled PCR Mix (from the kit) and 13 μL of PCR master mix. The DNA methylation of *TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, *RASSF1A*, and *SHOX2* genes was detected by real-time PCR with *β-actin* (*ACTB*) as an internal reference. Samples with Ct > 35 for *ACTB* were considered as poor-quality samples and were not included. Based on Ct < 35 for *ACTB*, the gene was defined as methylated when the Ct value was < 45. The Ct value was defined as 50 for the negative methylated genes in samples, as the number of cycles of the PCR assay was set as 50.

Full spectrum flow cytometry

The immunophenotypic analysis for lymphocytes in peripheral blood samples was performed using full-spectral flow cytometry. The mixed brilliant violet antibodies and brilliant staining buffer (Becton-Dickinson and Company, New York, USA) were added to peripheral blood samples and incubated in dark at room temperature for 15 minutes. Subsequently, the samples were diluted with buffer and centrifuged at 600 g for 10 minutes to remove the supernatant. The precipitates were washed by centrifugation with fixation/permeability buffer (Thermo Fisher Scientific, Massachusetts, USA). Thereafter, they were incubated with cytoplasmic antibodies for 30 minutes in the dark. Finally, the samples were washed with phosphate buffer, centrifuged to remove the supernatant, and the processed samples were analyzed using a full spectrum flow cytometer (Shanghai Xiatai Biotechnology Co. Ltd., Shanghai, China). The data analysis was performed using FlowJo™ v10.7 software (Becton-Dickinson and Company, New York, USA).

Statistical analysis

Differences in methylation levels were assessed using Chi-square tests or Fisher's exact tests. The model was visualized using receiver operating characteristic (ROC) curves and assessed by the area under the curve (AUC). The methylation statuses of the 7 genes were included in a multivariate predictive logistic regression analysis model, which was performed by R open-source software version 4.0.2. ROC analysis was performed by the pROC package. Other statistical analyses were performed in IBM SPSS Statistics v24.0 software (IBM Corporation, New York, USA). The reported *p* values were 2-sided. $p < 0.05$ was considered to be significantly different.

Results

Individual characteristics

A total of 94 individuals were included in this study, including 40 males and 54 females, with a median age of 33 (range from 21 to 66) years. According to the infection status, 18 individuals were in the stage of COVID-19 infection (infected group), 65 individuals had recovered within 30 days after infection with COVID-19 (recovered group), and 11 individuals had not been infected with COVID-19 at the time of this study (uninfected group). The ratio of male to female in the infected group, recovered group and uninfected group was 1:0.8 (10/8), 1:1.8 (23/42), and 1:0.6 (7/4), respectively ($p = 0.945$). In addition, there was no significant difference in the age distribution of the three groups ($p = 0.920$; Table 1).

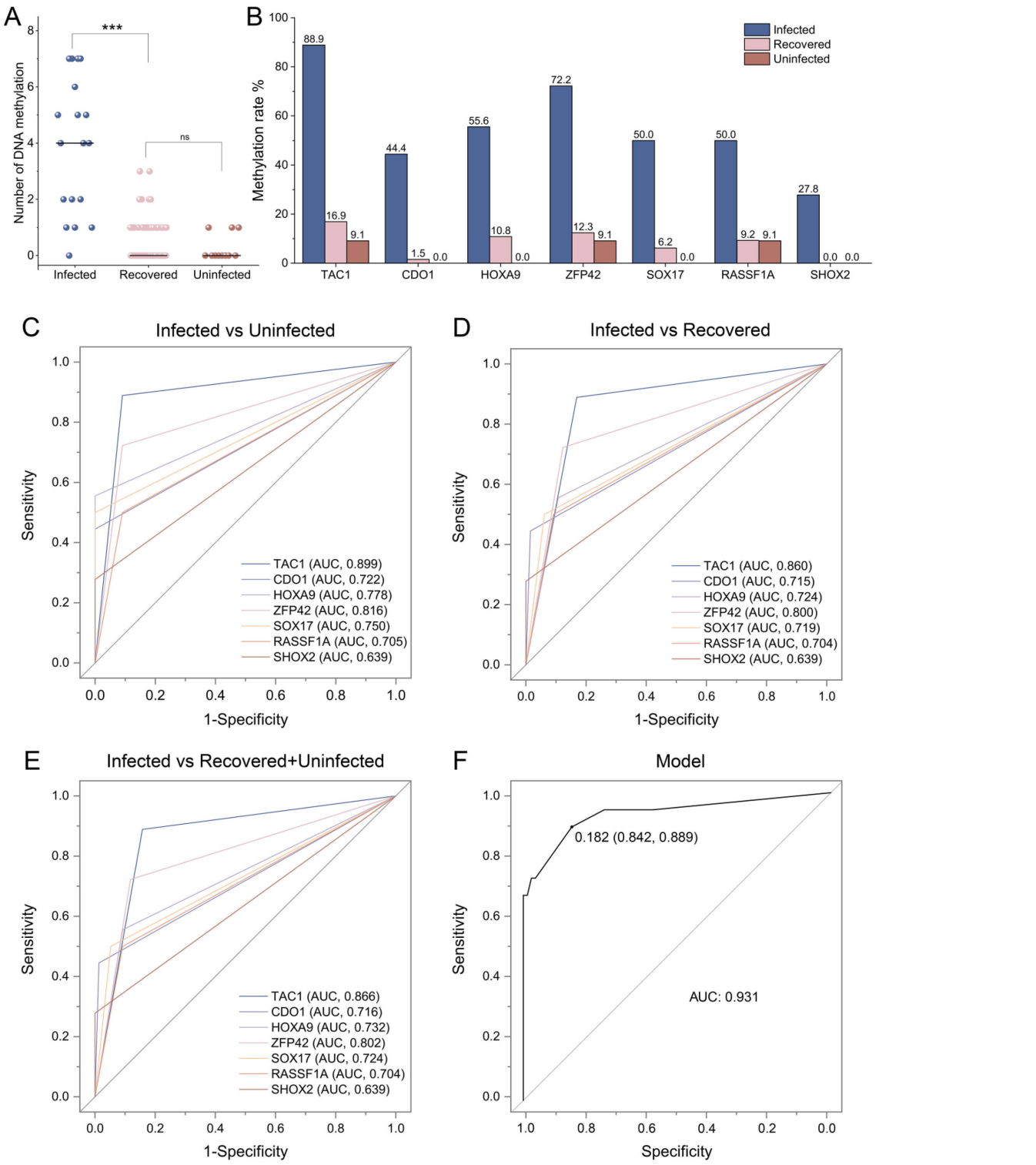
Significance of DNA methylation in the diagnosis of COVID-19 infection

In order to investigate the correlation between DNA methylation levels and COVID-19 infection, the study detected the rates of positive methylation of 7 genes in the 3 groups of individuals: infected, recovered, and uninfected, as mentioned above. The results showed that the median number of positively methylated genes in infected, recovered and uninfected groups were 4, 0, and 0, respectively. In addition, the positive rate of DNA methylation in infected patients was significantly higher than that in the recovered and uninfected groups (infected vs. recovered, $p < 0.001$; infected vs. uninfected, $p < 0.001$), but there was no significant difference in the positive rate of DNA methylation between the recovered and uninfected groups (recovered vs. uninfected, $p = 0.738$) (Figure 1A).

Table 1. Characteristics of the infected, recovered, and uninfected groups.

Characteristic	Infected (n = 18)	Recovered (n = 65)	Uninfected (n = 11)	*p* value
Gender (n, %)				0.9450
Male	10 (55.6)	23 (35.4)	7 (63.6)	
Female	8 (44.4)	42 (64.6)	4 (36.4)	
Age (median, range) in years	32 (21–55)	33 (21–66)	33 (22–57)	0.9199

Figure 1. Overview of DNA methylation and evaluation of diagnostic pneumonia recovery models.



A: Number of DNA methylations in the different individuals, ***p < 0.001; **B:** The seven gene methylation rates in the infected individuals, recovered individuals, and uninfected individuals; receiver operating characteristic (ROC) curves reveal the sensitivity, specificity, and area under curve (AUC) of the 7 gene methylation statuses distinguishing **C:** infected from uninfected individuals, **D:** infected from recovered individuals, and **E:** infected from negative (recovered and uninfected) individuals; **F:** The diagnostic model of the rehabilitation was evaluated for sensitivity, specificity, and AUC using seven-gene methylation status.

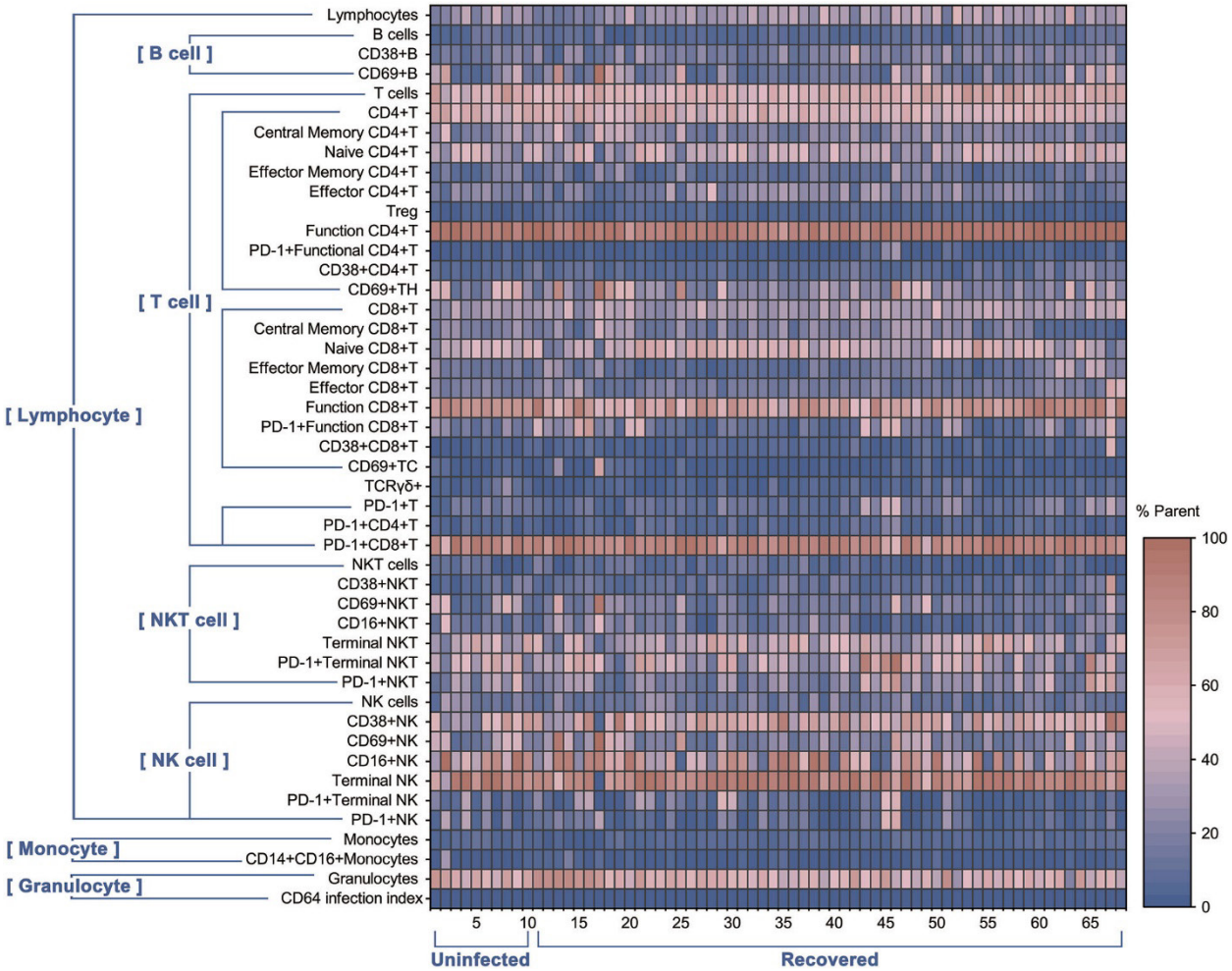
Table 2. Sensitivity and specificity of different genes in different groups.

Gene	Infected vs Uninfected			Infected vs Recovered			Infected vs Recovered + Uninfected		
	AUC	Specificity	Sensitivity	AUC	Specificity	Sensitivity	AUC	Specificity	Sensitivity
TAC1	0.899	0.909	0.889	0.860	0.831	0.889	0.866	0.842	0.889
CDO1	0.722	1.000	0.444	0.715	0.985	0.444	0.716	0.987	0.444
HOXA9	0.778	1.000	0.556	0.724	0.892	0.556	0.732	0.908	0.556
ZFP42	0.816	0.909	0.722	0.800	0.877	0.722	0.802	0.882	0.722
SOX17	0.750	1.000	0.500	0.719	0.939	0.500	0.724	0.947	0.500
RASSF1A	0.705	0.909	0.500	0.704	0.908	0.500	0.704	0.908	0.500
SHOX2	0.639	1.000	0.278	0.639	1.000	0.278	0.639	1.000	0.278

The analysis for methylation levels of different genes revealed that the positive methylation rates of the 7 genes in the infected group were *TAC1* 88.9%, *CDO1* 44.4%, *HOXA9* 55.6%, *ZFP42* 72.2%, *SOX17* 50.0%, *RASSF1A* 50.0% and *SHOX2* 27.8%. The positive percentage of DNA methylation was obviously higher in the infected period than in the recovered and uninfected periods (infected vs. recovered, $p < 0.001$). Meanwhile, in comparison, 7 genes were found to have markedly lower positive methylation levels in the uninfected periods, among which no positive methylation was detected in *COD1*, *HOXA9*, *SOX17*, and *SHOX2* (Figure 1B).

The diagnostic performance of single positive methylation of *TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, *RASSF1A*, and *SHOX2* was examined. The study was conducted to pairwise compare infected and uninfected groups, infected and recovered groups, and infected and negative (uninfected + recovered) groups. In the single methylation prediction model, the AUC ranged from 0.639 to 0.899, the specificity ranged from 0.831 to 1.000, and the sensitivity ranged from 0.278 to 0.889 (Figure 1C–1E and Table 2). In order to achieve a more accurate prediction for the recovery level of each

Figure 2. Overview of immune cell subpopulations in the uninfected and recovered groups.



individual, the study combined the DNA methylation status of *TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, *RASSF1A*, and *SHOX2* in infected and negative (uninfected + recovered) groups to construct a recovery-evaluation model (Supplementary Table 1). The individual overall DNA positive methylation level (M value) was calculated as follows: M = exp ((-4.596+3.356 × TAC1-1.105 × CDO1+0.095 × HOXA9+2.772 × ZFP42+1.287 × SOX17+1.006 × RASSF1A+15.664 × SHOX2)). Methylation positivity was scored as 1, and negativity as 0. In this model, the cutoff value was 0.182, the AUC was 0.931, and the sensitivity and specificity were 0.889 and 0.842, respectively (Figure 1F). Next, a robustness assessment was conducted using a 5-fold cross-validation, representing results with ROC curves (Supplementary Figure 1). More importantly, the model successfully distinguished the infected from the recovered or uninfected population by characterizing positive methylation.

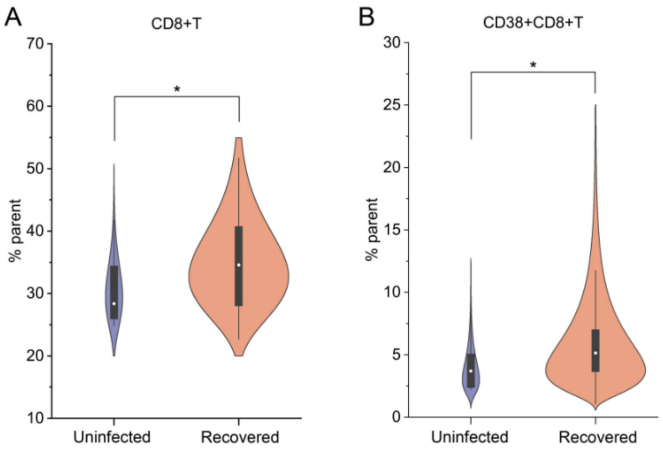
Different immune cell profiles in recovered and uninfected individuals

The study performed an assessment of immune function using full-spectrum flow cytometry in 58 recovered individuals, with 10 uninfected individuals serving as negative controls (Figure 2). The results demonstrated that the ratios of CD8+ T-cells and CD8+ CD38+ T-cells were significantly higher in the recovered individuals compared to the uninfected individuals (recovered vs. uninfected, $p < 0.05$) (Figure 3A–3B). In contrast, other immune cell populations did not exhibit significant differences (Supplementary Figures 2 and 3). Therefore, potential differences in immune function were found between recovered individuals and uninfected individuals.

Correlation of DNA methylation with immune cell profiles in recovered and uninfected individuals

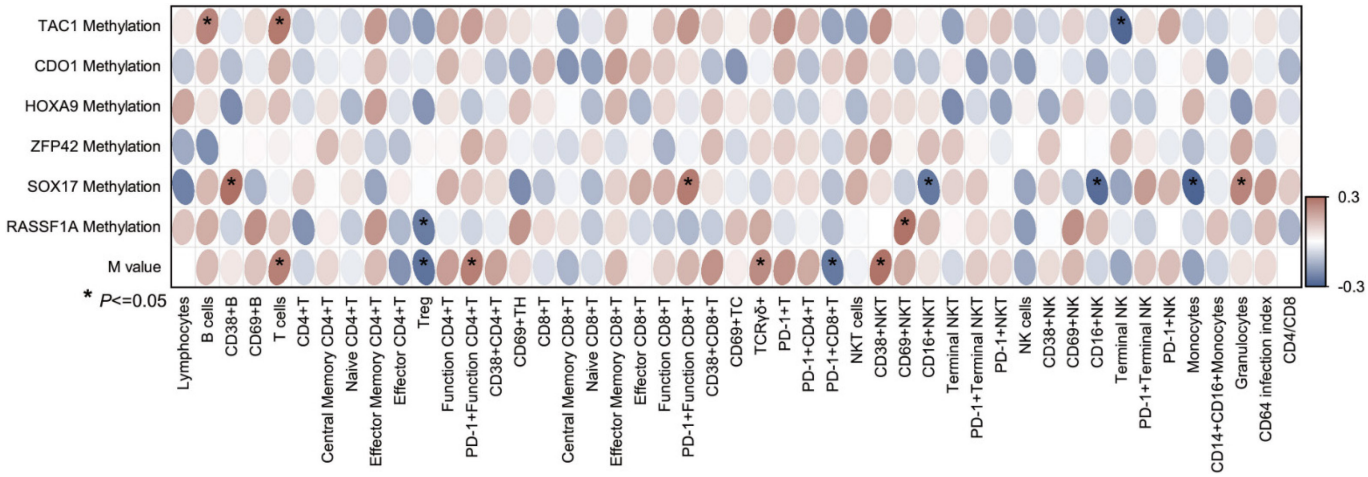
Taking together the results of DNA positive methylation and flow cytometry assays, the single positive methylation of *TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, and *RASSF1A*; as well as 7-gene methylation positive levels (M-values), in correlation with immune cell populations were further characterized. The results revealed that the *TAC1* methylation was significantly positively correlated with the proportion of B cells and T cells, and significantly negatively correlated with the proportion of terminal NK cells. *SOX17* methylation was significantly positively correlated with CD38+ B cells, PD-1+ function CD8+T cells, and granulocytes; and significantly negatively correlated with CD16+

Figure 3. Analysis of immune cell populations in the uninfected versus recovered group.



A: CD8+T cell ratio levels differed between uninfected group and recovered group; **B:** CD38+CD8+T cell ratio levels differed between uninfected group and recovered group. *p < 0.05.

Figure 4. Overview of the gene methylation correlation with immune cell populations. *p < 0.05.



NKT cells, CD16+ NK cells, and monocytes. *RASSF1A* methylation was significantly positively correlated with the proportion of CD69+ NKT cells and significantly negatively correlated with the proportion of Treg cells. In addition, the total M values calculated by the positive methylation levels of 7 genes demonstrated significant positive correlations with the proportions of T cells, PD-1+ function CD4+ T cells, TCRγδ+ cells, and CD38+ NKT cells; and significant negative correlations with the proportions of Treg cells and PD-1+ CD8+ T cells in the recovered population (Figure 4).

Discussion

With the growing prevalence of COVID-19, disease surveillance from infection to rehabilitation has received increased attention from a growing number of researchers [30]. In this study, a plasma DNA biomarker detection model (including *TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, *RASSF1A*, and *SHOX2*) was successfully constructed as a result of the analysis of the changes of cfDNA methylation in COVID-19-infected and recovered/uninfected populations. This model was used to monitor the positive amount of plasma DNA methylation in the recovery of COVID-19 patients to distinguish them from the infected population. Meanwhile, the correlation between 7 gene methylation and the proportion of immune cells was demonstrated in the study, which further characterized that this multigene methylation model could be equally meaningful for monitoring the immune function during the recovery period of COVID-19.

Methylation alterations can modify gene expression, which in turn can affect cellular pathways and phenotypes [31]. In fact, there have been numerous studies reporting associations between DNA methylation levels and human diseases, especially inflammatory diseases such as pneumonia [32,33]. The dynamic and reversible nature of DNA methylation has made it a suitable biomarker [34]. In addition, DNA methylation profiles can be generated in a single blood draw and measurement, providing a holistic view of whole-body metabolism and health status without analyzing multiple types of biomarkers, and can help to predict susceptibility to relevant comorbidities while improving and accelerating diagnosis [35,36].

Notably, previous studies have shown that the methylation status of genes including *TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, *RASSF1A*, and *SHOX2* has broad application value [20,37]. Mechanistically, *TAC1*, *CDO1*, and *RASSF1A* are tumor suppressor genes (TSGs) that regulate normal cell division, repair DNA damage, and promote apoptosis or programmed cell death of damaged cells [38]. Hypermethylation of these genes can lead to their functional inactivation, potentially triggering unregulated cell proliferation and contributing to tumorigenesis [39,40]. Additionally, *HOXA9* is a member of the *HOX* cluster family of transcription factors that are critical for endothelial cell differentiation, and its methylation is involved in carcinogenesis of various cancers [41]. *ZFP42* methylation is associated with differentiation and malignant transformation of stem cells, and is frequently observed in various types of tumors [42]. *SOX17* has been recognized as an important antagonist and inhibitor of the canonical *Wnt* signaling pathway, and hypermethylation of *SOX17* occurs frequently in lung cancer [43,44]. *SHOX2* was found to modulate cell proliferation and differentiation and induce epithelial-mesenchymal transition, making it an important growth regulator in the body [45]. Notably, previous studies revealed that parallels exist between COVID-19 and cancer at the pathophysiological level [46,47], suggesting these gene methylations may play an equivalent role in SARS-CoV-2 infection and clinical management strategies.

Based on the biological implications of these genes, mounting evidence suggests that the methylations of these genes can serve as diagnostic biomarkers for various diseases [20,37]. For example, prior studies have found a higher prevalence of methylations of *CDO1*, *TAC1*, *HOXA9*, *SOX17*, and *ZFP42* in cancer patients' plasma compared with healthy individuals [21,48]; and, the methylation of these genes was significantly associated with the diagnosis of non-small cell lung cancer that was independent from age, race, and smoking pack-years [21]. Similarly, Tan *et al*. have reported that the hypermethylation status of the *RASSF1A* gene was significantly associated with the risk of lung cancer, which increased with elevated methylation status [49]. Konecny *et al.* proposed that methylation of the *SHOX2* gene was a reliable marker of lung malignancy, and that plasma, as an alternative material for the detection of *SHOX2* methylation, might offer the possibility of screening for patients who are unable to undergo bronchoscopy [50]. Although these genes appeared to signify lung cancer association, it is exciting to note that in this study, there was a significant increase in the methylation positive levels of these 7 genes (*TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, *RASSF1A*, and *SHOX2*) in plasma from individuals infected with COVID-19 as compared to uninfected/recovered individuals. This was probably due to the fact that lung cancer represented the most frequent cancer histology in COVID-19 patients [51].

On the other hand, epigenetic modifications (DNA methylation) have been reported to play an important role in innate immune responses, regulation of B-cell function, and control of T-cell differentiation and memory responses [52,53]. Memory T and B cell differentiation involves epigenetic changes, including DNA methylation, that are required to initiate and maintain hereditary gene expression programs [54,55]. Interestingly, regulatory T cells (Tregs) are CD4+T cell subsets characterized by increased CD25 expression and decreased CD127 expression [56]. A decreasing trend in amount of Tregs in critically ill COVID-19 patients has been observed [57]. The reduction in the number of Tregs in COVID-19 patients might affect COVID-19 patients in several aspects, such as attenuating inflammatory suppression, causing an imbalance in the Treg/Th17 ratio, and increasing the risk of respiratory failure [58,59]. In addition, it was found that changes in Tregs were significantly associated with recovery in critically ill COVID-19 patients, meaning that Tregs were reduced in COVID-19 patients with sequelae compared to completely recovered individuals [60]. In this study, *RASSF1A* methylation was found to be negatively correlated with the proportion of Tregs, and the 7 gene methylation scores (M-values) also showed that there was a lower proportion of Tregs with higher M-values.

Several limitations need to be disclosed. First, this study included a limited number of samples. The robustness and generalizability of the findings require validation in larger cohorts. Second, the study lacked the investigation into the mechanism of action of the methylation of these biomarkers in COVID-19 infection, which could be a direction for future research.

Conclusions

This study revealed that the methylation of 7 genes (*TAC1*, *CDO1*, *HOXA9*, *ZFP42*, *SOX17*, *RASSF1A*, and *SHOX2*) was crucial in the diagnosis of COVID-19 infection, and was used to construct a prediction model which had sensitivity, specificity, and AUC of 0.889, 0.842, and 0.931, respectively. Further, it was found that the occurrence of DNA methylation of the 7 genes in the rehabilitation population correlated with changes in the proportion of immune cell subpopulations. Collectively, this study not only contributes to the assessment of recovery from COVID-19 infection, but also aids in the evaluation of immune function during the recovery period.

Acknowledgements

We thank Shanghai Rightongene Biotechnology Co. Ltd. (Shanghai, China) for performance and analysis of the full-spectrum flow cytometry.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' contributions

XX, BX, ZK, JY, HZ, XC: study conception and design; CD, RX, XL: clinical data; XX, BZ, ZK, RX: data interpretation; XC, BX, XX: manuscript preparation; All authors critically reviewed and modified the manuscript. The final version of the manuscript was read and approved by all the authors.

Corresponding authors

Xinyu Chen. MD.
School of Physics, Changchun University of Science and Technology,
No. 7186, Weixing Road, Chaoyang District, Changchun 130022, China.
Tel: +8613944862172
Fax: 0431-85582189
Email: xinyu_chen0110@163.com

Hao Zhang, MD.
School of Life Science and Technology, Changchun University of Science and Technology,
No. 7186, Weixing Road, Chaoyang District, Changchun 130022, China.
Tel: +8618043162127
Fax: 0431-85582189
Email: zhanghao@cust.edu.cn

Conflict of interest

Beibei Xin, Xiyan Li, and Zhonghe Ke were employed by Shanghai Rightongene Biotechnology Co. Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Wang MY, Zhao R, Gao LJ, Gao XF, Wang DP, Cao JM (2020) SARS-CoV-2: structure, biology, and structure-based therapeutics development. Front Cell Infect Microbiol 10: 587269. doi: 10.3389/fcimb.2020.587269.
2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W (2020) A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 382: 727–733. doi: 10.1056/NEJMoa2001017.
3. Hu B, Guo H, Zhou P, Shi ZL (2021) Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 19: 141–154. doi: 10.1038/s41579-020-00459-7.
4. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (2020) The species severe acute respiratory syndrome-related coronavirus: classifying 2019-

ncov and naming it SARS-CoV-2. Nat Microbiol 5: 536-544. doi: 10.1038/s41564-020-0695-z.
5. World Health Organization (2023) Coronavirus (COVID-19). Available: https://data.who.int/dashboards/covid19/more-resources. Accessed: 1 March 2024.
6. Fernandes Q, Inchakalody VP, Merhi M, Mestiri S, Taib N, Moustafa Abo El-Ella D, Bedhiafi T, Raza A, Al-Zaidan L, Mohsen MO, Yousuf Al-Nesf MA, Hssain AA, Yassine HM, Bachmann MF, Uddin S, Dermime S (2022) Emerging COVID-19 variants and their impact on SARS-CoV-2 diagnosis, therapeutics and vaccines. Ann Med 54: 524–540. doi: 10.1080/07853890.2022.2031274.
7. Tahan S, Parikh BA, Droit L, Wallace MA, Burnham CD, Wang D (2021) SARS-CoV-2 e gene variant alters analytical sensitivity characteristics of viral detection using a commercial reverse transcription-pcr assay. J Clin Microbiol 59: e0007521. doi: 10.1128/JCM.00075-21.
8. Safiabadi Tali SH, LeBlanc JJ, Sadiq Z, Oyewunmi OD, Camargo C, Nikpour B, Armanfard N, Sagan SM, Jahanshahi-Anbuhi S (2021) Tools and techniques for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)/COVID-19 detection. Clin Microbiol Rev 34: e00228–20. doi: 10.1128/CMR.00228-20.
9. Yüce M, Filiztekin E, Özkaya KG (2021) COVID-19 diagnosis -a review of current methods. Biosens Bioelectron 172: 112752. doi: 10.1016/j.bios.2020.112752.
10. Berdasco M, Esteller M (2019) Clinical epigenetics: seizing opportunities for translation. Nat Rev Genet 20: 109–127. doi: 10.1038/s41576-018-0074-2.
11. Cole J, Morris P, Dickman MJ, Dockrell DH (2016) The therapeutic potential of epigenetic manipulation during infectious diseases. Pharmacol Ther 167: 85–99. doi: 10.1016/j.pharmthera.2016.07.013.
12. Gayon J (2016) From Mendel to epigenetics: history of genetics. C R Biol 339: 225–230. doi: 10.1016/j.crvi.2016.05.009.
13. Pruimboom L (2020) Methylation pathways and SARS-CoV-2 lung infiltration and cell membrane-virus fusion are both subject to epigenetics. Front Cell Infect Microbiol 10: 290. doi: 10.3389/fcimb.2020.00290.
14. Nehme Z, Pasquereau S, Herbein G (2019) Control of viral infections by epigenetic-targeted therapy. Clin Epigenetics 11: 55. doi: 10.1186/s13148-019-0654-9.
15. Bhat S, Rishi P, Chadha VD (2022) Understanding the epigenetic mechanisms in SARS-CoV-2 infection and potential therapeutic approaches. Virus Res 318: 198853. doi: 10.1016/j.virusres.2022.198853.
16. Zhang L, Lu Q, Chang C (2020) Epigenetics in health and disease. Adv Exp Med Biol 1253: 3–55. doi: 10.1007/978-981-15-3449-2_1.
17. Kulis M, Esteller M (2010) DNA methylation and cancer. Adv Genet 70: 27–56. doi: 10.1016/B978-0-12-380866-0.60002-2.
18. Li YY, Yuan MM, Li YY, Li S, Wang JD, Wang YF, Li Q, Li J, Chen RR, Peng JM, Du B (2024) Cell-free DNA methylation reveals cell-specific tissue injury and correlates with disease severity and patient outcomes in COVID-19. Clin Epigenetics 16: 37. doi: 10.1186/s13148-024-01645-7.
19. Konigsberg IR, Barnes B, Campbell M, Davidson E, Zhen Y, Pallisard O, Boorgula MP, Cox C, Nandy D, Seal S, Crooks K, Sticca E, Harrison GF, Hopkinson A, Vest A, Arnold CG, Kahn MG, Kao DP, Peterson BR, Wicks SJ, Ghosh D, Horvath S, Zhou W, Mathias RA, Norman PJ, Porecha R, Yang IV, Gignoux CR, Monte AA, Taye A, Barnes KC (2021) Host methylation predicts SARS-CoV-2 infection and clinical outcome. Commun Med (Lond) 1: 42. doi: 10.1038/s43856-021-00042-y.
20. Zhong Q, Wang Y, Liang C, Wei F, She B (2023) Combined methylation of *shox2* and *rassf1a* genes in diagnosing malignant pleural effusion. Discov Med 35: 845–852. doi: 10.24976/Discov.Med.202335178.79.
21. Liu B, Ricarte Filho J, Mallisetty A, Villani C, Kottorou A, Rodgers K, Chen C, Ito T, Holmes K, Gastala N, Valyi-Nagy K, David O, Gaba RC, Ascoli C, Pasquinelli M, Feldman LE, Massad MG, Wang TH, Jusue-Torres I, Benedetti E, Winn RA, Brock MV, Herman JG, Hulbert A (2020) Detection of promoter DNA methylation in urine and plasma aids the detection of non-small cell lung cancer. Clin Cancer Res 26: 4339–4348. doi: 10.1158/1078-0432.CCR-19-2896.
22. Li N, Zeng Y, Huang J (2020) Signaling pathways and clinical application of *rassf1a* and *shox2* in lung cancer. J Cancer Res Clin Oncol 146: 1379–1393. doi: 10.1007/s00432-020-03188-9.
23. Sette A, Crotty S (2021) Adaptive immunity to SARS-CoV-2 and COVID-19. Cell 184: 861–880. doi: 10.1016/j.cell.2021.01.007.
24. Gusev E, Sarapultsev A, Solomatina L, Chereshnev V (2022) SARS-CoV-2-specific immune response and the pathogenesis of COVID-19. Int J Mol Sci 23: 1716. doi: 10.3390/ijms23031716.
25. Sette A, Crotty S (2022) Immunological memory to SARS-CoV-2 infection and COVID-19 vaccines. Immunol Rev 310: 27–46. doi: 10.1111/imr.13089.
26. Björkström NK, Ponzetta A (2021) Natural killer cells and unconventional t cells in COVID-19. Curr Opin Virol 49: 176–182. doi: 10.1016/j.coviro.2021.06.005.
27. Liontos A, Asimakopoulos AG, Markopoulos GS, Biros D, Athanasiou L, Tsourlos S, Dova L, Rapti IC, Tsiakas I, Ntzani E, Evangelou E, Tzoulaki I, Tsilidis K, Vartholomatos G, Dounousi E, Milionis H, Christaki E (2023) Correlation of lymphocyte subpopulations, clinical features and inflammatory markers during severe COVID-19 onset. Pathogens 12: 414. doi: 10.3390/pathogens12030414.
28. Qin W, Scicluna BP, van der Poll T (2021) The role of host cell DNA methylation in the immune response to bacterial infection. Front Immunol 12: 696280. doi: 10.3389/fimmu.2021.696280.
29. Barturen G, Carnero-Montoro E, Martínez-Bueno M, Rojo-Rello S, Sobrino B, Porras-Perales Ó, Alcántara-Domínguez C, Bernardo D, Alarcón-Riquelme ME (2022) Whole blood DNA methylation analysis reveals respiratory environmental traits involved in COVID-19 severity following SARS-CoV-2 infection. Nat Commun 13: 4597. doi: 10.1038/s41467-022-32357-2.
30. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, Cook JR, Nordvig AS, Shalev D, Sehrawat TS, Ahluwalia N, Bikdeli B, Dietz D, Der-Nigoghossian C, Liyanage-Don N, Rosner GF, Bernstein EJ, Mohan S, Beckley AA, Seres DS, Choueiri TK, Uriel N, Ausiello JC, Accili D, Freedberg DE, Baldwin M, Schwartz A, Brodie D, Garcia CK, Elkind MSV, Connors JM, Bilezikian JP, Landry DW, Wan EY (2021) Post-acute COVID-19 syndrome. Nat Med 27: 601–615. doi: 10.1038/s41591-021-01283-z.
31. Kappeler L, Meaney MJ (2010) Epigenetics and parental effects. BioEssays 32: 818–827. doi: 10.1002/bies.201000015.

32. Younesian S, Yousefi AM, Momeny M, Ghaffari SH, Bashash D (2022) The DNA methylation in neurological diseases. Cells 11: 3439. doi: 10.3390/cells11213439.
33. Morselli M, Farrell C, Montoya D, Gören T, Sabırlı R, Türkçüer İ, Kurt Ö, Köseler A, Pellegrini M (2022) DNA methylation profiles in pneumonia patients reflect changes in cell types and pneumonia severity. Epigenetics 17: 1646–1660. doi: 10.1080/15592294.2022.2051862.
34. Heikkinen A, Bollepalli S, Ollikainen M (2022) The potential of DNA methylation as a biomarker for obesity and smoking. J Intern Med 292: 390–408. doi: 10.1111/joim.13496.
35. Meng H, Cao Y, Qin J, Song X, Zhang Q, Shi Y, Cao L (2015) DNA methylation, its mediators and genome integrity. Int J Biol Sci 11: 604–617. doi: 10.7150/ijbs.11218.
36. Mittelstraß K, Waldenberger M (2018) DNA methylation in human lipid metabolism and related diseases. Curr Opin Lipidol 29: 116–124. doi: 10.1097/MOL.0000000000000491.
37. Maekawa H, Ito T, Orita H, Kushida T, Sakurada M, Sato K, Hulbert A, Brock MV (2020) Analysis of the methylation of cpg islands in the *cdo1*, *tac1* and *chfr* genes in pancreatic ductal cancer. Oncol Lett 19: 2197–2204. doi: 10.3892/ol.2020.11340.
38. Zhu K, Liu Q, Zhou Y, Tao C, Zhao Z, Sun J, Xu H (2015) Oncogenes and tumor suppressor genes: comparative genomics and network perspectives. BMC Genomics 16 Suppl 7: S8. doi: 10.1186/1471-2164-16-S7-S8.
39. Chen X, Poetsch A (2024) The role of *cdo1* in ferroptosis and apoptosis in cancer. Biomedicines 12: 918. doi: 10.3390/biomedicines12040918.
40. Rameshwar P, Gascón P (1996) Induction of negative hematopoietic regulators by neurokinin-a in bone marrow stroma. Blood 88: 98–106. doi: 10.1182/blood.V88.1.98.98.
41. Shenoy US, Basavarajappa DS, Kabekkodu SP, Radhakrishnan R (2024) Pan-cancer exploration of oncogenic and clinical impacts revealed that *hoxa9* is a diagnostic indicator of tumorigenesis. Clin Exp Med 24: 134. doi: 10.1007/s10238-024-01389-x.
42. Dansranjavin T, Krehl S, Mueller T, Mueller LP, Schmoll HJ, Dammann RH (2009) The role of promoter *cpg* methylation in the epigenetic control of stem cell related genes during differentiation. Cell Cycle 8: 916–924. doi: 10.4161/cc.8.6.7934.
43. Sinner D, Rankin S, Lee M, Zorn AM (2004) *Sox17* and beta-catenin cooperate to regulate the transcription of endodermal genes. Development 131: 3069–3080. doi: 10.1242/dev.01176.
44. Yin D, Jia Y, Yu Y, Brock MV, Herman JG, Han C, Su X, Liu Y, Guo M (2012) *Sox17* methylation inhibits its antagonism of *wnt* signaling pathway in lung cancer. Discov Med 14: 33–40.
45. Yi J, Jin L, Chen J, Feng B, He Z, Chen L, Song H (2017) Mir-375 suppresses invasion and metastasis by direct targeting of *shox2* in esophageal squamous cell carcinoma. Acta Biochim Biophys Sin (Shanghai) 49: 159–169. doi: 10.1093/abbs/gmw131.
46. Tyagi S, Tyagi N, Singh A, Gautam A, Singh A, Jindal S, Singh RP, Chaturvedi R, Kushwaha HR (2025) Linking COVID-19 and cancer: underlying mechanism. Biochim Biophys Acta Mol Basis Dis 1871: 167563. doi: 10.1016/j.bbadis.2024.167563.
47. Tian Y, Qiu X, Wang C, Zhao J, Jiang X, Niu W, Huang J, Zhang F (2021) Cancer associates with risk and severe events of COVID-19: a systematic review and meta-analysis. Int J Cancer 148: 363–374. doi: 10.1002/ijc.33213.
48. Chen C, Huang X, Yin W, Peng M, Wu F, Wu X, Tang J, Chen M, Wang X, Hulbert A, Brock MV, Liu W, Herman JG, Yu F (2020) Ultrasensitive DNA hypermethylation detection using plasma for early detection of *nsclc*: a study in Chinese patients with very small nodules. Clin Epigenetics 12: 39. doi: 10.1186/s13148-020-00828-2.
49. Tan S, Sun C, Wei X, Li Y, Wu Y, Yan Z, Feng F, Wang J, Wu Y (2013) Quantitative assessment of lung cancer associated with genes methylation in the peripheral blood. Exp Lung Res 39: 182–190. doi: 10.3109/01902148.2013.790096.
50. Konecny M, Markus J, Waczulikova I, Dolesova L, Kozlova R, Repiska V, Novosadova H, Majer I (2016) The value of *shox2* methylation test in peripheral blood samples used for the differential diagnosis of lung cancer and other lung disorders. Neoplasma 63: 246–253. doi: 10.4149/210_150419N208.
51. Guarnera A, Santini E, Podda P (2022) COVID-19 pneumonia and lung cancer: a challenge for the radiologist review of the main radiological features, differential diagnosis and overlapping pathologies. Tomography 8: 513–528. doi: 10.3390/tomography8010041.
52. Cheong JG, Ravishankar A, Sharma S, Parkhurst CN, Grassmann SA, Wingert CK, Laurent P, Ma S, Paddock L, Miranda IC, Karakaslar EO, Nehar-Belaid D, Thibodeau A, Bale MJ, Kartha VK, Yee JK, Mays MY, Jiang C, Daman AW, Martinez de Paz A, Ahimovic D, Ramos V, Lercher A, Nielsen E, Alvarez-Mulett S, Zheng L, Earl A, Yallowitz A, Robbins L, LaFond E, Weidman KL, Racine-Brzostek S, Yang HS, Price DR, Leyre L, Rendeiro AF, Ravichandran H, Kim J, Borczuk AC, Rice CM, Jones RB, Schenck EJ, Kaner RJ, Chadburn A, Zhao Z, Pascual V, Elemento O, Schwartz RE, Buenrostro JD, Niec RE, Barrat FJ, Lief L, Sun JC, Ucar D, Josefowicz SZ (2023) Epigenetic memory of coronavirus infection in innate immune cells and their progenitors. Cell 186: 3882–3902.e3824. doi: 10.1016/j.cell.2023.07.019.
53. Reiner SL (2005) Epigenetic control in the immune response. Hum Mol Genet 14: R41–46. doi: 10.1093/hmg/ddi115.
54. Montacchiesi G, Pace L (2022) Epigenetics and cd8(+) t cell memory. Immunol Rev 305: 77–89. doi: 10.1111/imr.13057.
55. Wu H, Deng Y, Feng Y, Long D, Ma K, Wang X, Zhao M, Lu L, Lu Q (2018) Epigenetic regulation in b-cell maturation and its dysregulation in autoimmunity. Cell Mol Immunol 15: 676–684. doi: 10.1038/cmi.2017.133.
56. Staats J (2019) Immunophenotyping of human regulatory t cells. Methods Mol Biol 2032: 141–177. doi: 10.1007/978-1-4939-9650-6_9.
57. Rahimzadeh M, Naderi N (2021) Toward an understanding of regulatory t cells in COVID-19: a systematic review. J Med Virol 93: 4167–4181. doi: 10.1002/jmv.26891.
58. Wang H, Wang Z, Cao W, Wu Q, Yuan Y, Zhang X (2021) Regulatory t cells in COVID-19. Aging Dis 12: 1545–1553. doi: 10.14336/AD.2021.0709.
59. Muyayalo KP, Huang DH, Zhao SJ, Xie T, Mor G, Liao AH (2020) COVID-19 and treg/th17 imbalance: potential relationship to pregnancy outcomes. Am J Reprod Immunol 84: e13304. doi: 10.1111/aji.13304.
60. Hoffmann AD, Weinberg SE, Swaminathan S, Chaudhuri S, Almubarak HF, Schipma MJ, Mao C, Wang X, El-Shennawy L, Dashzeveg NK, Wei J, Mehl PJ, Shihadah LJ, Wai CM, Ostiguin C, Jia Y, D'Amico P, Wang NR, Luo Y, Demonbreun AR, Ison MG, Liu H, Fang D (2023) Unique molecular signatures sustained in circulating monocytes and regulatory t cells in convalescent COVID-19 patients. Clin Immunol 252: 109634. doi: 10.1016/j.clim.2023.109634.

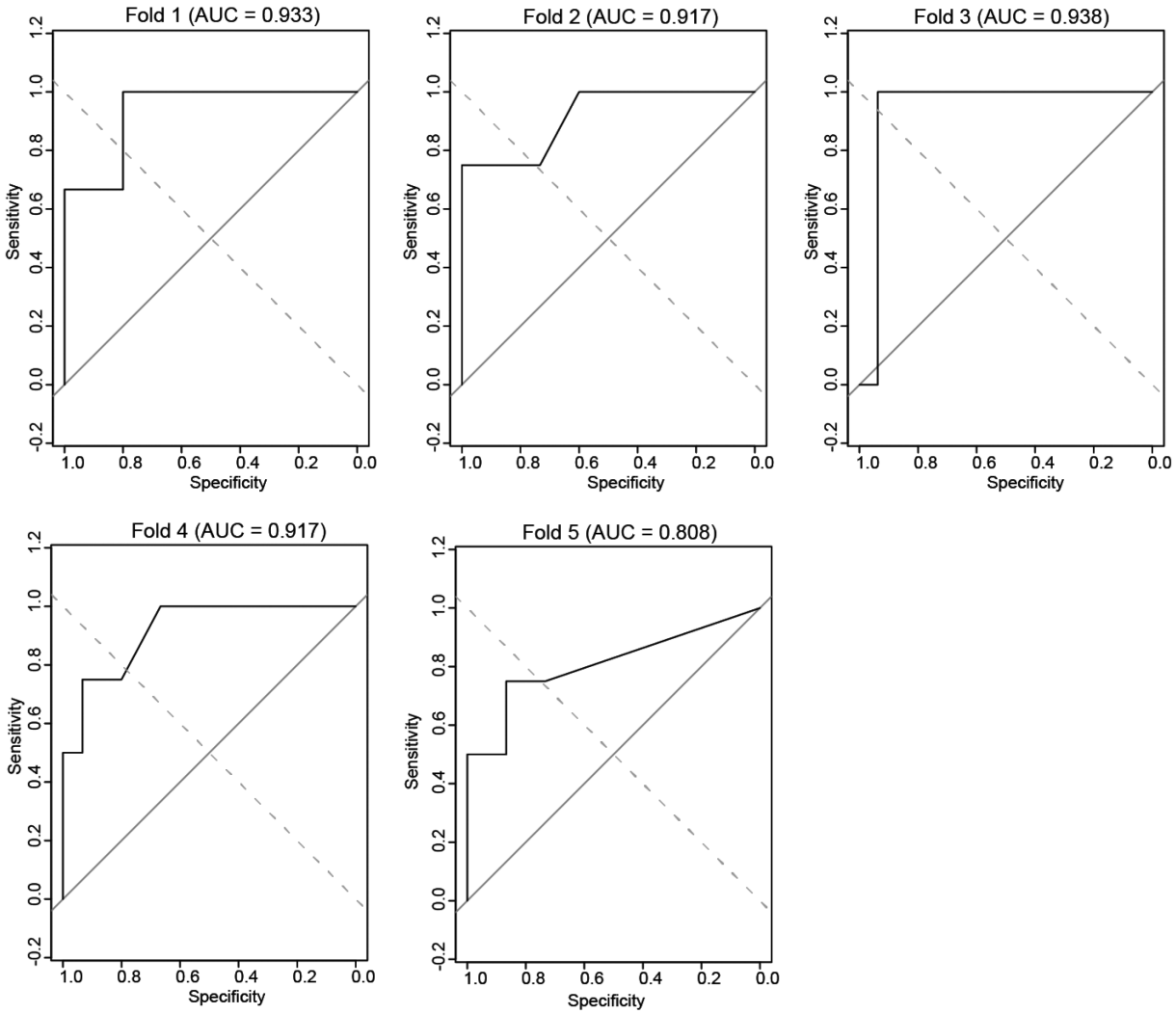
Annex – Supplementary Items

Supplementary Table 1. Multivariate logistic regression analysis of the variables.

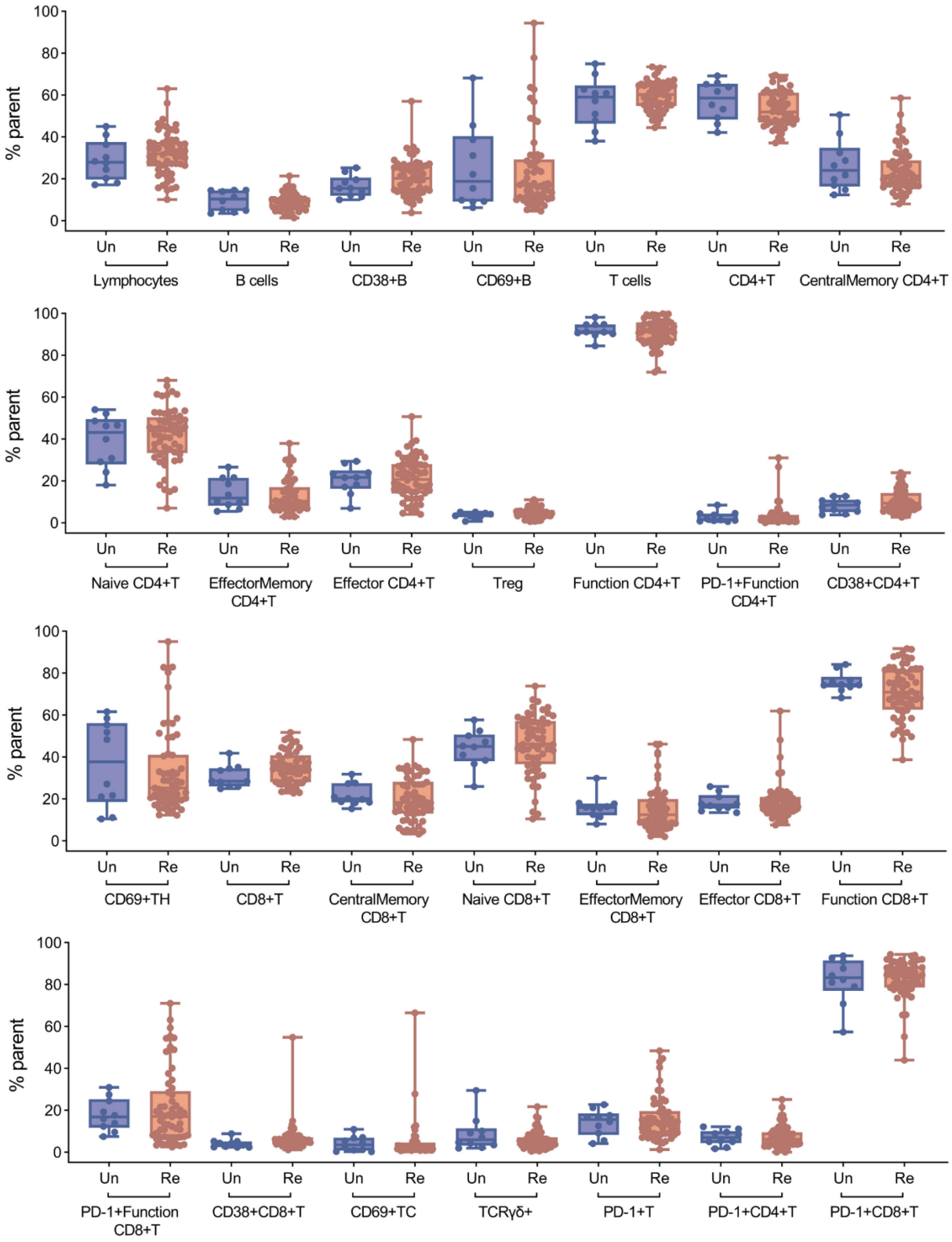
Variables	OR	95% CI	*p* value
TAC1	28.68	3.57–641.67	0.00579
CDO1	0.33	0–19.31	0.5825
HOXA9	1.1	0.09–10.22	0.93409
ZFP42	15.99	1.62–373.34	0.02892
SOX17	3.62	0.28–52.72	0.3215
RASSF1A	2.74	0.12–47.66	0.49397
SHOX2	6347825	0–NA	0.99557

OR: odds ratio; CI: confidence interval.

Supplementary Figure 1. The receiver operating curve (ROC) curve results were confirmed by a 5-fold cross-validation.



Supplementary Figure 2. Overview showing the percentage of immune cell populations in the uninfected (un) and recovered (re) groups.



Supplementary Figure 3. Overview showing the percentage of immune cell populations in the uninfected (un) and recovered (re) groups.