

RESEARCH

Open Access



Methylation patterns of the nasal epigenome of hospitalized SARS-CoV-2 positive patients reveal insights into molecular mechanisms of COVID-19

Benjamin L Spector^{1,2*}, Boryana Koseva², Rebecca McLennan², Dithi Banerjee³, Kamani Lankachandra⁴, Todd Bradley², Rangaraj Selvarangan³ and Elin Grundberg^{2*}

Abstract

Background Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has varied presentations from asymptomatic to death. Efforts to identify factors responsible for differential COVID-19 severity include but are not limited to genome wide association studies (GWAS) and transcriptomic analysis. More recently, variability in host epigenomic profiles have garnered attention, providing links to disease severity. However, whole epigenome analysis of the respiratory tract, the target tissue of SARS-CoV-2, remains ill-defined.

Results We interrogated the nasal methylome to identify pathophysiologic drivers in COVID-19 severity through whole genome bisulfite sequencing (WGBS) of nasal samples from COVID-19 positive individuals with severe and mild presentation of disease. We noted differential DNA methylation in intergenic regions and low methylated regions (LMRs), demonstrating the importance of distal regulatory elements in gene regulation in COVID-19 illness. Additionally, we demonstrated differential methylation of pathways implicated in immune cell recruitment and function, and the inflammatory response. We found significant hypermethylation of the *FUT4* promoter implicating impaired neutrophil adhesion in severe disease. We also identified hypermethylation of *ELF5* binding sites suggesting downregulation of *ELF5* targets in the nasal cavity as a factor in COVID-19 phenotypic variability.

Conclusions This study demonstrated DNA methylation as a marker of the immune response to SARS-CoV-2 infection, with enhancer-like elements playing significant roles. It is difficult to discern whether this differential methylation is a predisposing factor to severe COVID-19, or if methylation differences occur in response to disease severity. These differences in the nasal methylome may contribute to disease severity, or conversely, the nasal immune system may respond to severe infection through differential immune cell recruitment and immune function, and through differential regulation of the inflammatory response.

Keywords COVID-19, SARS-CoV-2, Methylation, Epigenome

*Correspondence:

Benjamin L Spector

bspector2@wisc.edu

Elin Grundberg

egrundberg@cmh.edu

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Since its emergence in December 2019, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in nearly 775 million cases of coronavirus disease 2019 (COVID-19) and over 7 million deaths worldwide [1]. This placed COVID-19 as the third leading cause of death in the United States in 2021 [2]. The symptomatology of this disease is extremely variable, ranging from asymptomatic infection to mild respiratory complaints, to cardiopulmonary failure and death [3]. Mortality associated with COVID-19 has most commonly been attributed to septic shock and multiorgan failure, often secondary to suppurative pneumonia [4]. These wide-reaching and catastrophic outcomes associated with COVID-19 led to global efforts to define the pathophysiology of this disease in hopes of identifying therapeutic targets and pharmacologic interventions to reduce morbidity and mortality. However, many of the causative mechanisms that determine COVID-19 severity remain elusive.

Proposed mechanisms and associations of severe SARS-CoV-2 infections, evaluated mainly through transcriptomic and proteomic analyses, include increased levels of proinflammatory cytokines [5–7], modulations of immune cells including leukocyte exhaustion with depletion of T lymphocytes in particular [5, 8–10], and increased binding affinity of the SARS-CoV-2 spike protein to the host angiotensin 2 (ACE 2) receptor [11–13]. Large genome-wide association studies (GWAS) of common and rare variants have provided additional insight into the biological underpinnings of infection severity, identifying single nucleotide variants (SNV) in or near genes involved in the innate immune response to viral SARS-CoV-2 infections, type I interferon (IFN) immunity, blood group phenotype, and viral entry [14–16].

Recently, interindividual differences in epigenetic footprints have been postulated as key drivers in some of the proposed pathways and determinants of differential clinical outcomes between patients [17–22]. DNA methylation, the most studied epigenetic mark, is cell-specific and occurs on cytosine residues in the context of cytosine-guanine dinucleotides (CpG). Generally, methylation (i.e., hypermethylation) of a gene promoter induces a closed chromatin configuration, such that methylation serves as a silencer of gene expression [23]. Conversely, lack of methylation (i.e., hypomethylation) is commonly associated with activation of gene transcription [24]. Though DNA methylation can be dynamic in response to environmental stimuli [25, 26], methylation patterns are typically propagated across cell divisions such that changes in methylation state can result in long-lasting effects on gene expression [23, 27, 28]. In fact, it was recently demonstrated that individuals previously

hospitalized with COVID-19 exhibit changes in their methylomes that persist for at least one year after hospital discharge [29].

In the examination of DNA methylation during COVID-19 pathogenesis, the most highly considered contribution to disease pathogenesis has been related to methylation status of the *ACE2* promoter region, with relative hypomethylation (i.e., increased gene expression) noted in individuals with severe disease as compared to uninfected controls [30]. Other associations suggesting methylation as a causative factor in COVID-19 severity include differentially methylated regions (DMRs) of interferon-related genes and interferon-effector genes in severe COVID-19 cases which correlate with observations of decreased transcriptional products of antiviral *IFN* genes [30, 31]. Further, hypomethylation of other inflammatory regulators leading to elevated cytokine/chemokine gene expression has been described [30]. Taken together, these findings indicate that differential methylation patterns impacting host-viral interactions may predispose certain individuals to more severe infection. Additionally, like other RNA viruses, SARS-CoV-2 may induce innate immune dysfunction thereby leading to impairment in host immune defenses.

In this study, we apply whole-genome bisulfite sequencing (WGBS) to define the global epigenomic landscape of the nasal mucosa as it relates to the host response in severe versus mild cases of SARS-CoV-2 infection utilizing biospecimens collected early in the COVID-19 pandemic, representing primary infections prior to the advent of the vaccination initiative. We demonstrate supporting data of differential antiviral responses and immune cell populations between disease severities. We highlight the importance and interplay between multiple inflammatory mechanisms, including the phosphoinositide 3-kinase/serine-threonine kinase (PI3K/Akt) pathway, Notch, and nuclear factor kappa B (NF- κ B) signaling. We identify differential methylation of *FUT4* – an immature neutrophil marker and adhesion molecule— as putative factors in severe COVID-19 pathogenesis. Finally, we expand the understanding of the roles of *ELF4* and *ELF5* in controlling transcriptional regulation as it relates to COVID-19 severity.

Results

Characterization of the regulatory genomic landscape in nasal mucosa by WGBS

Nasal samples from 61 subjects positive for the alpha or beta variants of SARS-CoV-2 presenting to a single center at the time of symptomatic presentation concerning for COVID-19 (4 severe and 57 mild) from April 8, 2020, through June 8, 2020, were included in the study. Demographic features of hospitalized (defined as severe;

Table 1 Demographic features of hospitalized versus non-hospitalized COVID-19 positive subjects

Item	Hospitalized (n=4)	Non-hospitalized (n=57)
^a Age (years)	60 (50, 66.75)	37 (27, 51)
^b Gender		
Female	1 (25)	25 (44)
Male	3 (75)	32 (56)
^b Race		
Black	1 (25)	27 (47)
White	0 (0)	13 (23)
Other	3 (75)	17 (30)

^a Presented as median (IQR)^b Presented as n (%)

$n=4$) versus non-hospitalized (defined as mild; $n=57$) patients are summarized in Table 1. Of note, all hospitalized subjects required intensive care unit admission. Three of the four hospitalized subjects required supplemental oxygen support with two of the four hospitalized subjects requiring intubation and mechanical ventilation.

WGBS data was generated at high depth, identifying on average 13.2 million CpGs per sample each at $>10X$ coverage. Hierarchical clustering was performed on the top 25th percentile most variably methylated regions showing no clustering structure for confounders such as self-reported race and gender (Supplemental Fig. 1). Using the combined WGBS datasets of severe and mild COVID-19 cases, we characterized active regulatory regions in nasal mucosa. Specifically, we performed methylation segmentations to extract unmethylated regions (UMR) and low methylated regions (LMR) which are known to correlate with promoter- and enhancer-like elements, respectively [25]. We identified 19,187 UMRs (average 2,366 bp) with methylation across the regions being $<5\%$ and containing on average 117 CpGs per region. LMRs were, as expected, more abundant identifying 43,924 regions with an intermediate methylation status (5–50%) and more CpG-sparse (8 CpGs per region, average 642 bp). We annotated these regions based on publicly available reference maps of regulatory DNA based on DNase I hypersensitive sites (DHSs) across 16 different cell types [32] and found 99% and 98% of UMRs and LMRs, respectively, overlapped a DHS. Of these annotated regions, the vast majority (98%) of the UMRs overlapped with a DHS detected in multiple cell types. In contrast, LMRs were shown to represent to a larger extent cell-specific regulatory DNA with 27% of LMRs ($N=11,976$) overlapping a DHS unique to a specific cell type. Of these 11,976 cell-specific LMRs from our aggregated COVID-19 positive

samples (severe and mild disease), we noted 11% and 20% being lymphoid and myeloid regulatory elements, respectively (Supplemental Table 1).

To better characterize the cell types in which differential methylation may be associated with COVID-19 severity, we examined the UMRs and LMRs of severe and mild WGBS datasets separately, as well as WGBS data from nasal samples derived from non-infected individuals. While there were not substantial differences in the cell-type proportions of identified DHSs between severe and mild cases of COVID-19 (Supplemental Table 2), there were notable differences between COVID-19 positive (combined severe and mild cases) and COVID-19 negative individuals, particularly in the case of LMRs (Fig. 1, Supplemental Table 1). Most notably, when comparing COVID-19 positive versus negative individuals, we found that 46% compared to 25% of LMRs overlapped with immune cell regulatory elements ($X^2=5189.6$, $df=1$, $p<2.2\times 10^{-16}$). More specifically, in COVID-19 positive compared to COVID-19 negative individuals, 24% versus 15% of LMRs overlapped with lymphoid cell regulatory elements ($X^2=1264.0$, $df=1$, $p<2.2\times 10^{-16}$), and 28% versus 12% of LMRs overlapped with myeloid cell regulatory elements ($X^2=4557.1$, $df=1$, $p<2.2\times 10^{-16}$). While the same trend was seen for UMRs, the magnitude of effect was strikingly stronger when contrasting LMRs in COVID-19 positive versus COVID-19 negative individuals (Fig. 1).

When narrowing our focus to only those UMRs and LMRs appreciated for a single cell type, we noted similar findings. Our comparisons of cell-type specific UMRs and LMRs were relatively similar between those with severe and mild disease (Supplemental Table 2). However, when comparing the cell-specific UMRs and LMRs in SARS-CoV-2 positive as compared to negative individuals, substantial differences were appreciated in LMR distribution. Specifically, when examining cell-specific UMRs in COVID-19 positive versus negative individuals, 8% and 4% of cell-specific UMRs overlapped with immune regulatory elements ($p=0.064$) and 0.6% and 3% overlapped with epithelial regulatory elements ($p=0.032$) (Supplemental Table 1). Regarding cell-specific LMRs in COVID-19 positive as compared to negative individuals, 31% and 8% of cell-specific LMRs overlapped immune regulatory elements ($p=0.0037$), 11% and 4% overlapped with lymphoid regulatory elements ($p=5.6\times 10^{-16}$), 20% and 3% overlapped with myeloid regulatory elements ($p<2.2\times 10^{-16}$), 2% and 3% overlapped with pulmonary elements ($p<2.2\times 10^{-16}$), and 3% and 11% overlapped with epithelial regulatory elements ($p<2.2\times 10^{-16}$) (Supplemental Table 1).

In all, these results point towards activation of immune cells in nasal mucosa after SARS-CoV-2 infection and indicate that WGBS can capture methylation

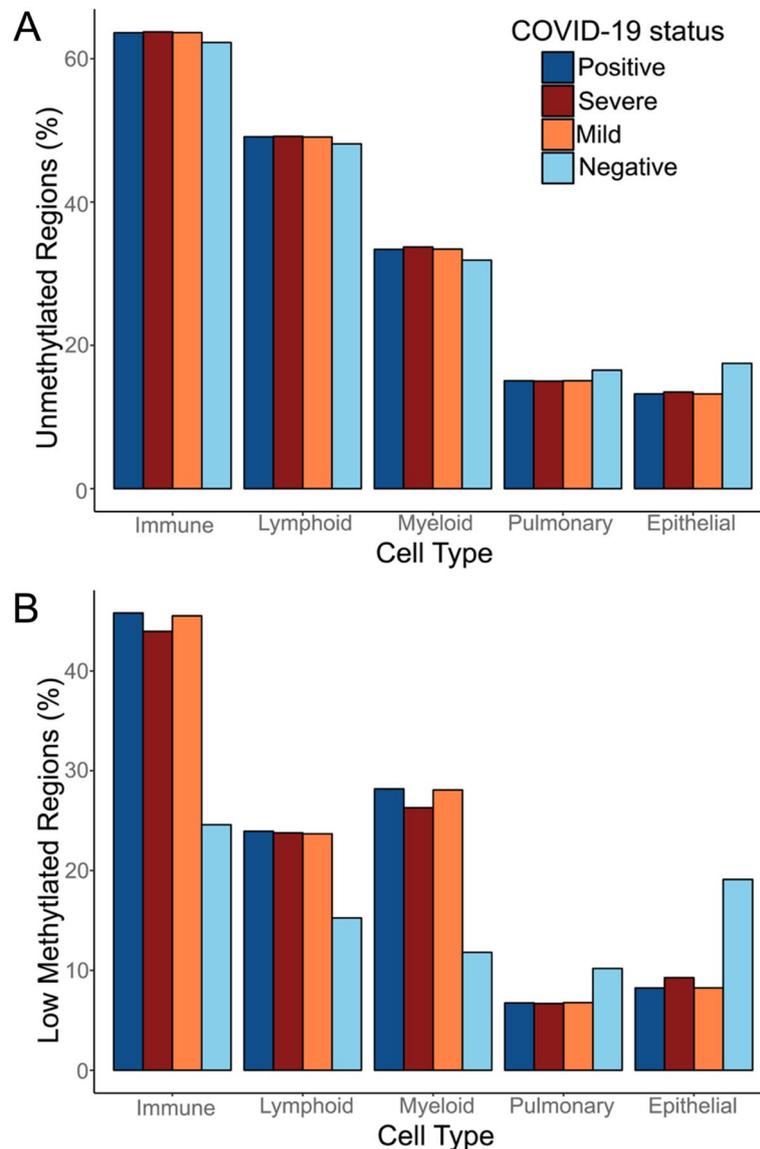


Fig. 1 Characterization of the regulatory landscape of the nasal methylome in COVID-19 positive and negative individuals. **A** The percentage (%) of un-methylated regions (UMR) overlapping cell-type specific regulatory elements as determined by DHS are depicted across COVID-19 positive individuals (severe + mild) (dark blue), individuals with severe disease (red), mild disease (orange), or COVID-19 negative (light blue). **B** The percentage (%) of low-methylated regions (LMR) overlapping cell-type specific regulatory elements as determined by DHS are depicted across COVID-19 positive individuals (severe + mild) (dark blue), individuals with severe disease (red), mild disease (orange), or COVID-19 negative (light blue)

signatures specific to these cell types. This suggests that our WGBS analysis of the nasal methylome can be used to infer differential regulation of immune-mediated pathways in the setting of severe as compared to mild SARS-CoV-2 infection.

Preponderance of hypomethylated regions in severe COVID-19 patients

To identify DMRs between individuals suffering from severe COVID-19 (i.e. requiring inpatient admission)

versus individuals experiencing mild COVID-19 (i.e. remaining outpatient) we performed tiling window analysis using the WGBS data sets and logistic regression models with the self-reported measures of age, race, and gender included as covariates. To evaluate meaningful methylation differences, the top 10,000 DMRs as ranked by q-value ($q\text{-value} < 1.20 \times 10^{-15}$) were selected for further analysis. These DMRs demonstrated a preponderance of hypomethylated regions ($n=7,256$) as compared to hypermethylated regions ($n=2,744$) in hospitalized

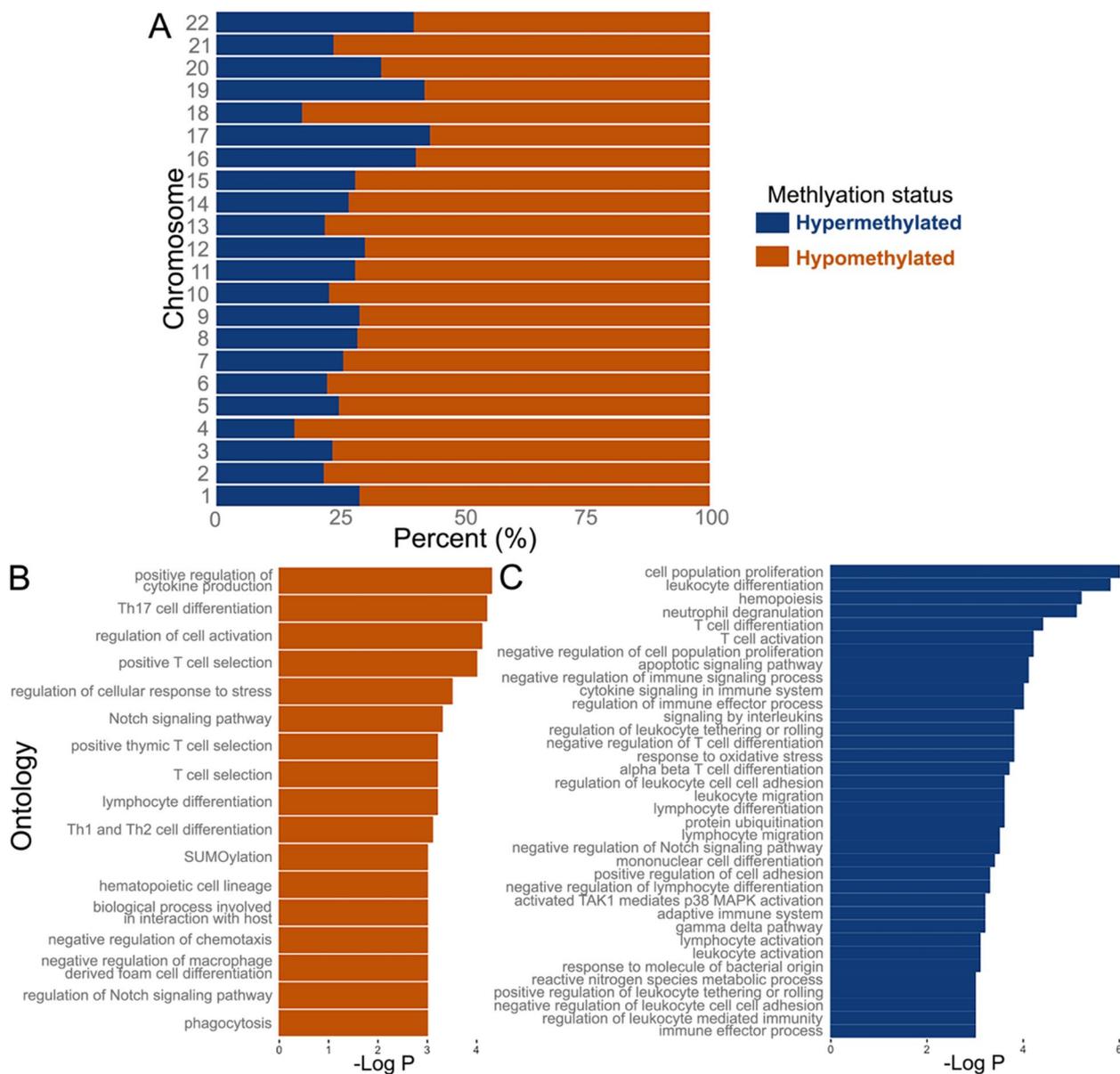


Fig. 2 Distribution of hypomethylated and hypermethylated differentially methylated regions (DMRs) by chromosome in severe versus mild COVID-19 and GO Biological Processes, KEGG Pathways, and Reactome Gene Sets associated with differentially methylated genes (DMGs) in severe versus mild COVID-19. **A** The top 10,000 DMRs according to q-value were subdivided by chromosomal location (y-axis). The percentage (x-axis) of relatively hypermethylated (blue) and hypomethylated (orange) DMRs in severe as compared to mild COVID-19 cases is shown on a per chromosome basis. **B-C** The most significant ($p < 1 \times 10^{-3}$) GO Biological Processes, KEGG Pathways, and Reactome Gene Sets related to the immune response in relatively hypomethylated (**B**) and hypermethylated (**C**) DMGs in individuals with severe as compared to mild COVID-19. -Log P values are displayed along the x-axis

versus non-hospitalized subjects (Fig. 2A). Among these regions, 3,599 (49.6%) hypomethylated DMRs fell in intergenic regions as compared to hypermethylated DMRs where only 378 (13.8%) mapped to similar regions ($X^2=1,064.4$, $df=1$, $p < 2.2 \times 10^{-16}$). We then queried genes associated with hypo- and hypermethylated DMRs using Genomic Regions Enrichment of Annotations Tool

(GREAT) algorithm [33, 34]. This yielded associations with 487 genes in hypomethylated regions and 503 genes in hypermethylated regions (Supplemental Tables 3 and 4). This similarity in gene counts associated with hypo-versus hypermethylated regions despite the substantial difference in DMR suggests interdependency of regulatory elements involved in gene activation associated with

SARS-CoV-2 infection. Of note, this state of global hypomethylation has been previously appreciated in response to other viral infections [35].

To discern the relevant biologic pathways associated with these differentially methylated genes (DMGs) we performed pathway enrichment analysis. Enrichment of hypomethylated and hypermethylated DMGs in severe versus mild COVID-19 patients were carried out separately. Pathway enrichment analysis using Coronascape [36] was performed on 487 relatively hypomethylated genes with a p -value threshold of ≤ 0.01 ($\text{Log}_{10}P \leq -2$) resulting in 353 Gene Ontology (GO) processes [37, 38], 28 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways [39], 21 Reactome Gene Sets [40], 10 canonical pathways [41], and 2 CORUM complexes [42] (total 414 enriched pathways) previously associated with SARS-CoV-2 infection (Supplemental Table 5). Of these, associations with $p \leq 10^{-3}$ were manually evaluated ($n=124$) and results associated with immune response were extracted ($n=17$) (Fig. 2B, C). Themes emerging from these relatively hypomethylated immune-related pathways in severe compared to mild COVID-19 cases included processes related to Th1, Th2 and Th17 cell differentiation, regulation of T lymphocytes and hematopoietic cell lines, regulation of the Notch signaling pathway, cytokine production and chemotaxis, and regulation of macrophages and phagocytosis (Fig. 2B). In the similar analysis of relatively hypermethylated DMGs in severe versus mild COVID-19 cases, 423 GO Biological processes [37, 38], 40 KEGG pathways [39], 57 Reactome Gene Sets [40], and 12 canonical pathways [41] (total 532 enriched pathways) met the p -value threshold of ≤ 0.01 ($\text{Log}P \leq -2$), while 181 had an associated p -value of $\leq 10^{-3}$ (Supplemental Table 6). Of these 181 enriched pathways, 36 were identified as being associated with the immune response (Fig. 2C). Emerging themes of these enriched immune-related pathways associated with hypermethylated genes in severe versus mild COVID-19 patients included processes associated with cell activation, proliferation, and differentiation including hematopoietic cells from myeloid (e.g., neutrophils) and lymphoid (e.g., T cells) populations, leukocyte migration and adhesion, neutrophil degranulation, adaptive and leukocyte mediated immunity, cytokine signaling, and host stress responses (Fig. 2C). These cumulative findings support previous reports of aberrancies in immune response in the face of mild/moderate as opposed to severe COVID-19 [43–45].

Differential methylation of genes within the PI3K/Akt pathway and COVID-19 severity

Upon closer manual evaluation of genes implicated in various enriched GO terms, KEGG pathways, and

Reactome gene sets, we noted differential methylation between severe and mild COVID-19 individuals in many genes involved in the PI3K/Akt pathway (Fig. 3). Specifically, in the severe cohort compared to those with mild COVID-19, we noted relative hypomethylation of promoters whose genes have interplay with the PI3K/Akt pathway and SARS-CoV-2 infection, including genes within the NF- κ B (e.g., *NFKBIA*) and Notch (e.g., *NOTCH1*) signaling pathways (Fig. 3A). Genes mapping to the Notch signaling pathway were similarly appreciated as a significantly hypomethylated GO term in our enrichment analysis of severe as compared to mild COVID-19 patients (Fig. 2B).

Comparing severe versus mild SARS-CoV-2 infection, DMGs included relative hypomethylation of the *AKT1* promoter (chr14: 104,796,001 – 104,796,500, methylation difference = -18.38% , $q = 2.22 \times 10^{-22}$) and of its downstream target, the type I interferon signaling molecule, *ISG15* (chr1: 1,013,751 – 1,014,250, methylation difference = -32.27% , $q = 2.49 \times 10^{-35}$) (Fig. 3B, C). Further support implicating the PI3K/Akt pathway in the severe COVID-19 phenotype are the relative hypomethylation of *ZEB2* (chr2: 144,524,251 – 144,524,750, methylation difference = -24.39% , $q = 2.52 \times 10^{-32}$) and *SNAI1* (chr20: 49,983,251 – 49,983,750, methylation difference = -28.98% , $q = 6.11 \times 10^{-16}$) promoters in severe versus mild disease. Coordinates and methylation difference of DMRs in genes associated with the PI3k/Akt pathway from our dataset are summarized in Supplemental Table 7.

Hypermethylation of immune cell surface markers in severe cases of COVID-19

Differential expression of various immune cell surface makers have been noted in the setting of severe SARS-CoV-2 infections, including CD15 (the gene product of *FUT4*) and CD8 [8, 43, 45–52]. In examination of the most significant differentially methylated bins in severe versus mild COVID-19 cases, we found the *FUT4* promoter to be relatively hypermethylated over a long region in hospitalized as compared to non-hospitalized subjects (chr11: 94,545,001–94,552,500, $q = 2.90 \times 10^{-112}$) (Fig. 4A). To further evaluate this finding, we visualized this region using the UCSC Genome Browser [53] comparing severe to mildly infected individuals, in addition to pooled samples from healthy controls ($n=7$) and found that this relatively hypermethylated state of the *FUT4* promoter among severe COVID-19 patients persisted across comparison groups (Fig. 4A). CD15, the gene product of *FUT4*, is predominately expressed in myeloid cells, as is shown in Fig. 4B, generated using the dataset of Monaco et al. [54] and the Human Protein Atlas (proteintlas.org) [55]. We additionally noted significant hypermethylation

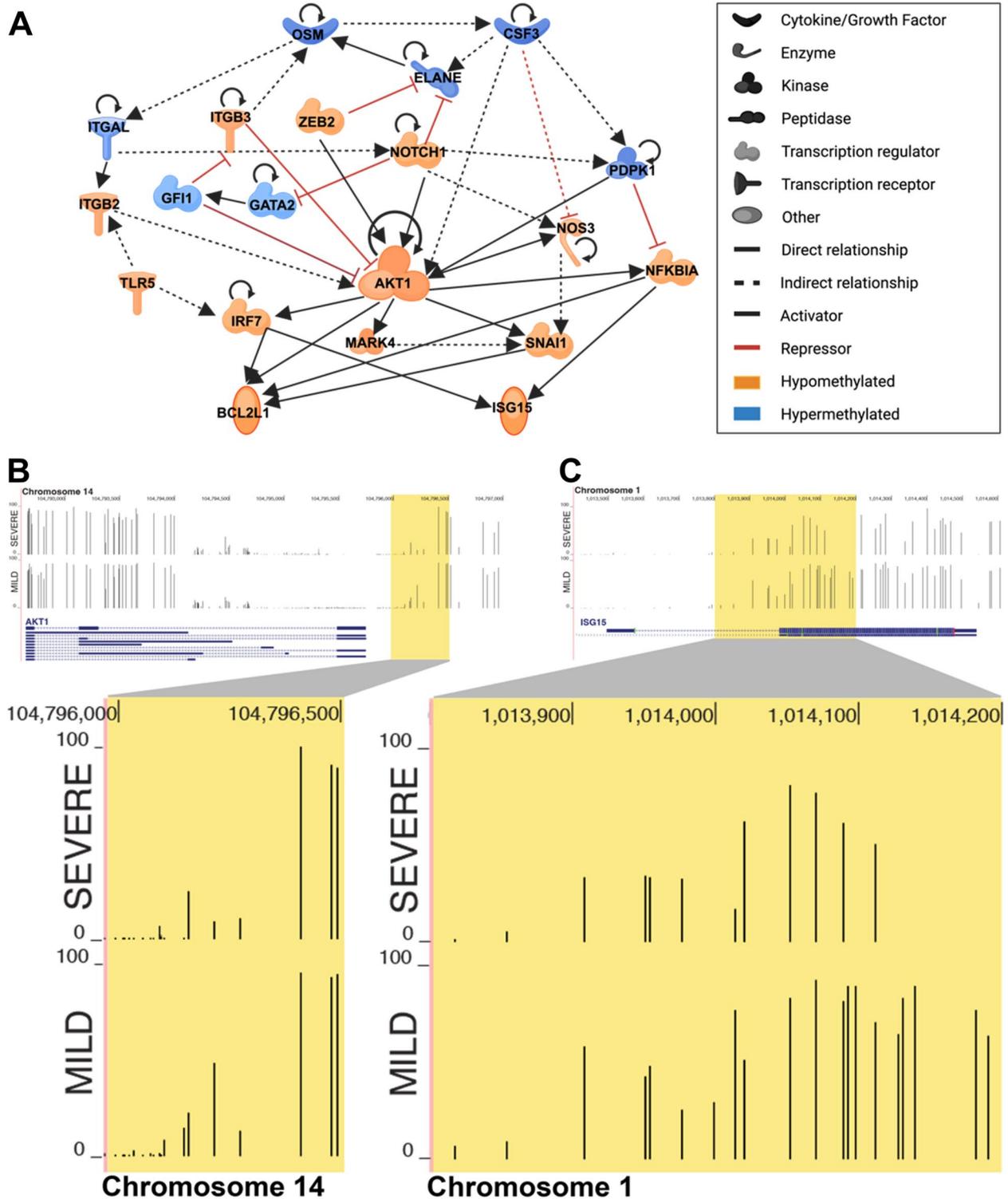


Fig. 3 Differentially methylated genes in the PI3K/Akt pathway in severe versus mild COVID-19 patients. **A** Ingenuity Pathway Analysis (QIAGEN Ingenuity Pathway Analysis (IPA) version 01–21-03, Venlo, Netherlands) was performed demonstrating the interplay of differentially methylated genes (DMGs) with *AKT1*. Hypomethylated DMGs in severe versus mild COVID-19 within our dataset are depicted in orange; hypermethylated DMGs in severe versus mild COVID-19 are depicted in blue. Known relationships between genes as activators (black) and repressors (red) are shown, with direct interactions displayed as solid lines and indirect as dashed lines. Created with BioRender.com. **B, C** WGBS methylation analysis demonstrating hypomethylation of the *AKT1* (**B**) and *ISG15* (**C**) promoters (yellow rectangle) in individuals with severe COVID-19 as compared to mild COVID-19. Y-axis demonstrates percent methylation at a given CpG site (0–100%). Figure generated using the UCSC Genome Browser

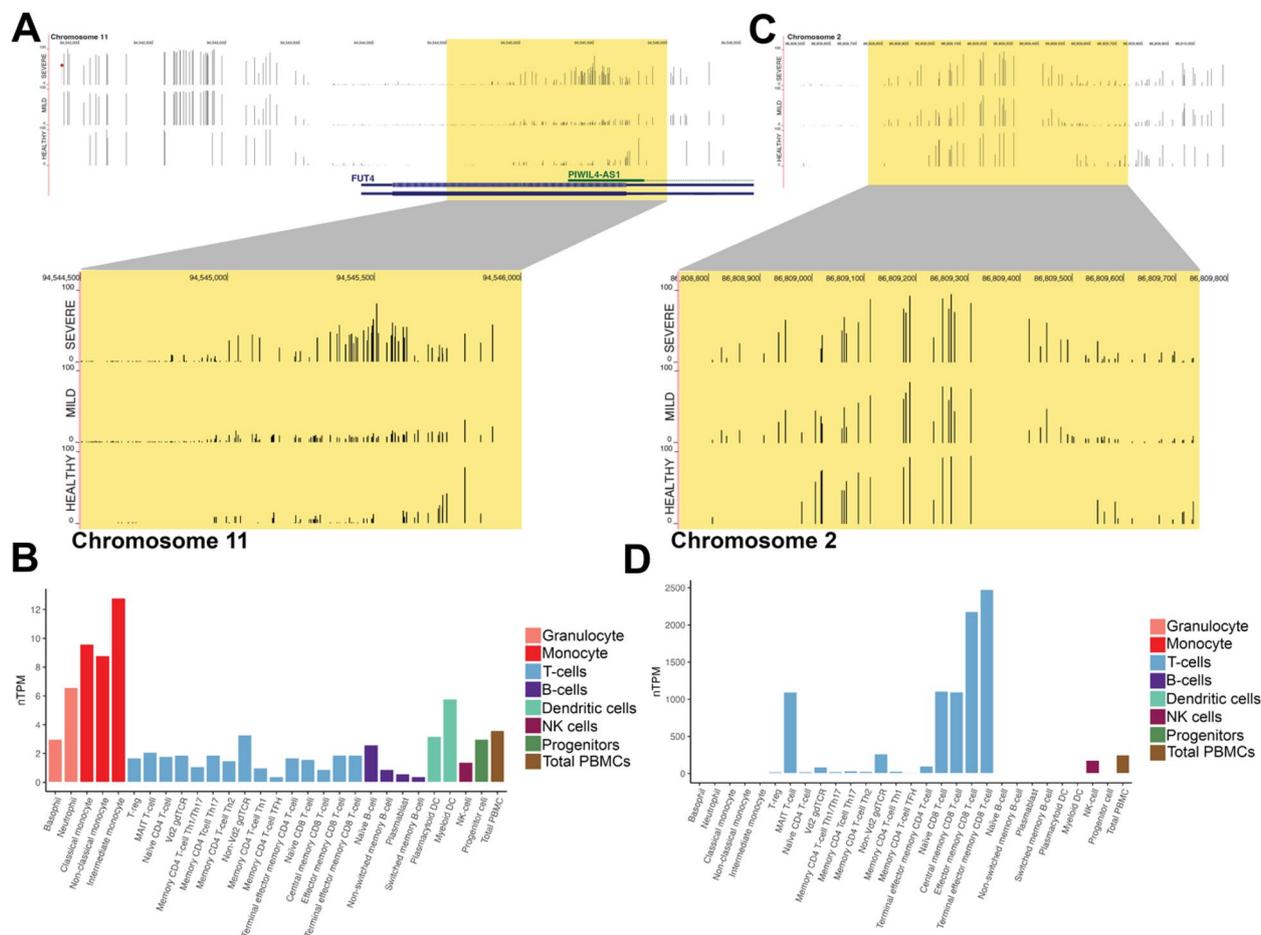


Fig. 4 Hypermethylation of the *FUT4* and *CD8A* promoters in severe COVID-19 patients with predominate expression in myeloid cells. **A** WGBS methylation analysis demonstrating hypermethylation of the *FUT4* promoter (yellow rectangle) in individuals with severe COVID-19 (top panel) as compared to mild COVID-19 (middle panel) and negative controls (bottom panel). Y-axis demonstrates percent methylation at a given CpG site (0–100%). Red lollipop (top panel) represents locus of SNP rs117126460, shown by the COVID-19 Host Genetics Initiative (HGI) to confer increased risk for COVID-19. Figure generated using the UCSC Genome Browser. **B** *FUT4* is preferentially expressed in immune cells derived from the myeloid cell lineage (normalized transcripts per million, y-axis). Granulocytes = pink, Monocytes = red, T-cells = blue, B-cells = purple, Dendritic cells = teal, NK cells = magenta, Progenitors = olive, Total PBMC = brown. Image credit: Human Protein Atlas. Image reproduced from: v23.proteinatlas.org/ENSG00000196371-FUT4/immune+cell#top. **C** WGBS methylation analysis demonstrating hypermethylation of the *CD8A* promoter. **D** *CD8A* is strongly expressed in T-cells. Image credit: Human Protein Atlas. Image reproduced from: v23.proteinatlas.org/ENSG00000153563-CD8A/immune+cell

of the *CD8A* promoter (chr2: 86,809,001–86,809,750, $q=1.50 \times 10^{-21}$), which is strongly expressed in T cells (Fig. 4C, D). Cumulatively, these data provide evidence that myeloid cell dysfunction is associated with COVID-19 severity, and that differential regulation of cell surface genes (e.g., *FUT4*, *CD8A*) may be either a cause or effect of disease severity.

Enrichment of *ELF4* and *ELF5* transcriptional motifs in hypermethylated regions

To better understand the regulatory pathways involved in COVID-19 pathogenesis and severity, we examined

transcription factor binding sites among DMRs in severe versus mild COVID-19 samples. The top 25,000 hypo- and hypermethylated bins as determined by q-value between severe and mildly infected individuals were evaluated separately using HOMER motif analysis (Tables 2 and 3) [56]. Evaluation of hypermethylated regions revealed targets of *ELF4* ($p=1 \times 10^{-59}$, target sequences with motif=11.34%, background sequences with motif=8.33%) and *ELF5* ($p=1 \times 10^{-54}$, target sequences with motif=8.84%, background sequences with motif=6.31%) as among the most significantly enriched motifs (Table 3).

Table 2 Most significantly enriched transcription factor binding motifs of hypomethylated regions in hospitalized versus non-hospitalized COVID-19 patients

Motif	Name	P-value	q-value (Benjamini)	# Target Sequences with Motif	% Targets Sequences with Motif	# Background Sequences with Motif	% Background Sequences with Motif	^a Fold change
	Mef2c	1e-12	0.0000	2905.0	11.62%	2543.9	10.19%	1.14
	Fosl2	1e-12	0.0000	709.0	2.84%	537.1	2.15%	1.32
	Fra2	1e-10	0.0000	1111.0	4.44%	905.8	3.63%	1.22
	Jun-AP1	1e-10	0.0000	520.0	2.08%	383.2	1.54%	1.35
	JunB	1e-10	0.0000	1346.0	5.38%	1123.3	4.50%	1.20
	AP-1	1e-10	0.0000	1797.0	7.19%	1545.5	6.19%	1.16
	Mef2a	1e-9	0.0000	2473.0	9.89%	2189.6	8.77%	1.13
	Fos	1e-8	0.0000	1466.0	5.86%	1253.4	5.02%	1.17
	Atf3	1e-8	0.0000	1601.0	6.40%	1380.9	5.53%	1.16
	EWS:ERG -fusion	1e-8	0.0000	1947.0	7.79%	1711.0	6.86%	1.14

^a Fold change represented as % targeted sequences/% background sequences

Among these relatively hypermethylated DMRs in hospitalized as compared to non-hospitalized individuals, 2,829 were found to be targets of *ELF4*. More specifically, when annotating these regions based on DHSs [32], we found that 11.2% ($n=317$) of these *ELF4* targets overlapped signatures of myeloid cells as compared to non-*ELF4* targets, of which 8.6% ($n=1,906$) overlapped with myeloid signatures ($X^2=20.75$, $df=1$, $p=5.23 \times 10^{-6}$). Regarding lymphoid signatures, we found that the percentage of *ELF4* and non-*ELF4* targets overlapping with lymphoid cell DHSs were roughly equivalent, 9.9% ($n=280$) and 9.1% ($n=2,024$), respectively ($X^2=1.68$, $df=1$, $p=0.19$). These findings are in keeping with the known preferential upregulation of *ELF4* within myeloid cell lines (Supplemental Fig. 2) as well as its known role in host antiviral response [57]. The distribution of *ELF4* transcription factor binding motifs across cell types is summarized in Supplemental Fig. 3).

We identified 2,199 relatively hypermethylated DMRs as targets of *ELF5* in severe as compared to mild COVID-19 individuals. Additionally, we noted relative

hypermethylation (i.e., downregulation) of the *ELF5* promoter in COVID-19 positive individuals (severe or mild) as compared to COVID-19 negative ($n=7$) patients (Fig. 5). Recently, Pietzner et al. [58] suggested several genes to be potentially regulated or co-expressed with *ELF5*, of which 23 were also identified within our dataset as being relatively hypermethylated targets of *ELF5* in individuals with severe as compared to mild COVID-19 (Supplemental Table 8). Notably, we found that *C1orf116* (chr1: 207,031,251–207,031,750, $q=7.85 \times 10^{-7}$), *PLAC8* (chr4: 83,128,251–83,128,750, $q=4.43 \times 10^{-10}$), and *IFRD1* (chr7: 112,421,501–112,422,000, $q=8.15 \times 10^{-9}$) were among these relatively hypermethylated *ELF5* targets in severe as compared to mild COVID-19, all of which have been implicated in the host response to SARS-CoV-2 infection [59–61].

Supportive evidence

To validate our results and thereby overcome the limited sample size of our severely affected cohort, we analyzed external publicly available data sources. These sources

Table 3 Most significantly enriched transcription factor binding motifs of hypermethylated regions in hospitalized versus non-hospitalized COVID-19 patients

Motif	Name	P-value	q-value (Benjamini)	# Target Sequences with Motif	% Targets Sequences with Motif	# Background Sequences with Motif	% Background Sequences with Motif	^a Fold change
	CEBP	1e-62	0.0000	2758.0	11.03%	2005.5	8.02%	1.38
	EHF	1e-60	0.0000	3794.0	15.18%	2924.5	11.69%	1.30
	ELF3	1e-60	0.0000	2272.0	9.09%	1597.9	6.39%	1.42
	ELF4	1e-59	0.0000	2835.0	11.34%	2085.4	8.33%	1.36
	HLF	1e-57	0.0000	3096.0	12.38%	2328.9	9.31%	1.33
	NFIL3	1e-54	0.0000	2381.0	9.52%	1722.0	6.88%	1.38
	ELF5	1e-54	0.0000	2209.0	8.84%	1578.5	6.31%	1.40
	ERG	1e-53	0.0000	4373.0	17.49%	3501.1	13.99%	1.25
	ETS1	1e-50	0.0000	2776.0	11.10%	2089.7	8.35%	1.33
	CEBP: AP1	1e-49	0.0000	2481.0	9.92%	1837.8	7.34%	1.35

^a Fold change represented as % Targeted Sequences/% Background sequences

corroborated many of our interesting findings including the biological processes associated with COVID-19 severity, candidate genes associated with SARS-CoV-2 infection susceptibility or severity, and association of the PI3K/Akt pathway with COVID-19 severity.

Differential regulation of processes related to cell cycle regulation, cell migration, the cytokine response, immune cell regulation, and infection response are associated with SARS-CoV-2 infection severity

We utilized the Gene Expression Omnibus (GEO) tool, GEO2R, to examine differential gene expression in severe compared to mild COVID-19 patients in the datasets of Gómez-Carballa et al. [62] and Rombauts et al. [63]. In the case of Gómez-Carballa et al., differentially expressed genes (DEGs) from the nasal mucosa were evaluated from individuals with severe ($n=14$) as compared to mild ($n=17$) COVID-19. This yielded 159 DEGs meeting a significance threshold of $p_{adj} < 0.05$. In the case of the dataset of Rombauts et al., differential

gene expression was evaluated in whole blood of hospitalized COVID-19 positive patients at the time of admission between those who developed acute respiratory distress syndrome (ARDS) ($n=19$) and those who did not ($n=31$). When considering differentially expressed autosomal gene-associated loci, this resulted in 152 DEGs. The DEG lists generated from the datasets of Gómez-Carballa et al. and Rombauts et al. were then cross-referenced to the DMGs derived from our dataset. From these combined datasets, 28 unique DEGs overlapped with the DMGs noted within our data (differential expression of *HLA-DPA1* was present in both datasets). From the dataset of Gómez-Carballa et al., 19 DEGs overlapped with DMGs from our dataset (11 hypomethylated, 8 hypermethylated). From the dataset of Rombauts et al., 10 DEGs overlapped with our identified DMGs (5 hypomethylated, 5 hypermethylated). Among these 28 overlapping genes, 16 genes (57%) demonstrated a direction of differential expression concordant with what is expected from our methylation data

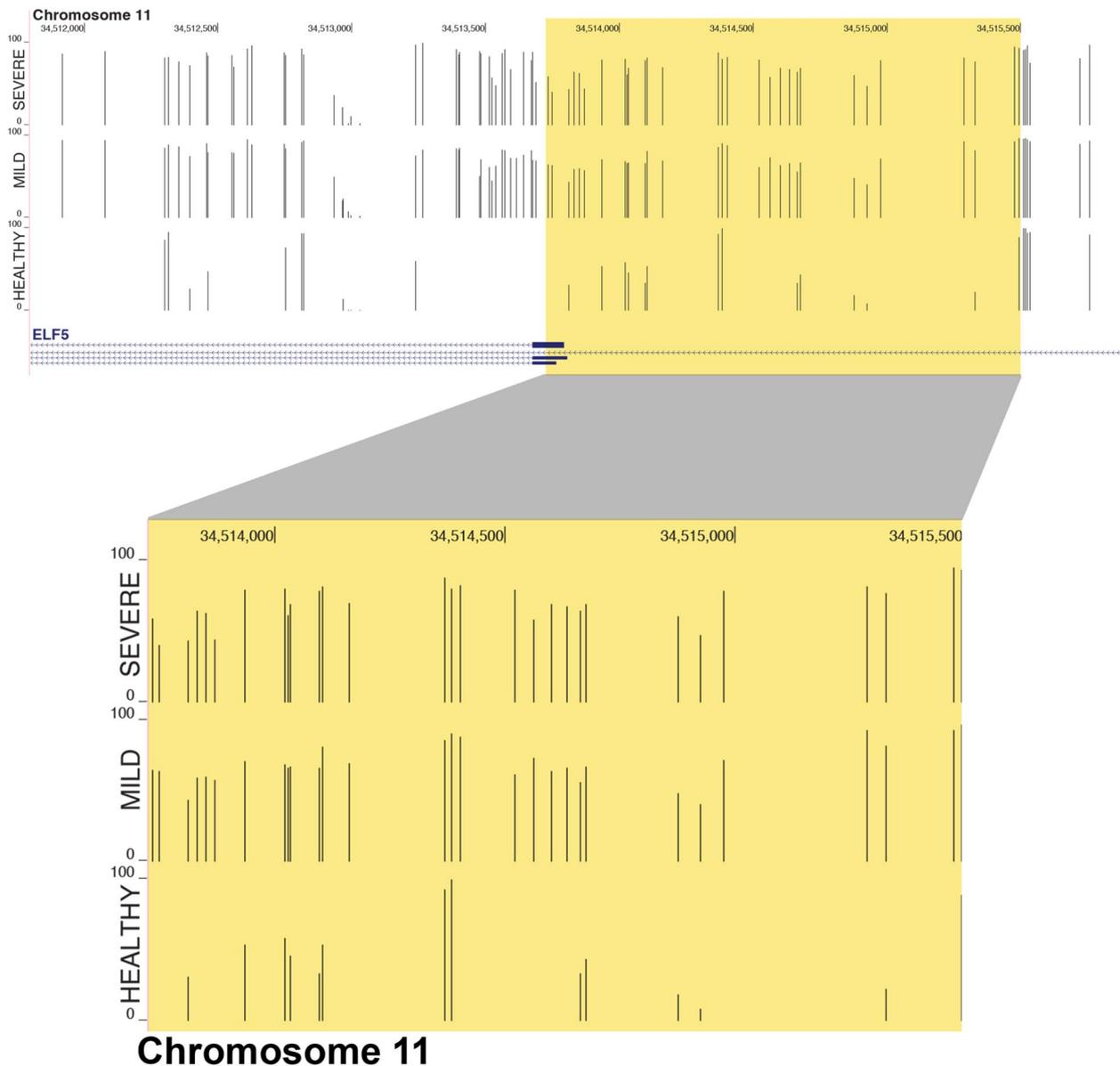


Fig. 5 Hypermethylation of the *ELF5* promoter in nasopharyngeal samples of COVID-19 positive as compared to COVID-19 negative individuals. WGBS methylation analysis demonstrating hypermethylation of the *ELF5* promoter (yellow rectangle) in individuals with severe COVID-19 (top panel) and mild COVID-19 (middle panel) as compared to negative controls (bottom panel). Y-axis demonstrates percent methylation at a given CpG site (0–100%). Figure generated using the UCSC Genome Browser

(i.e., increased expression of hypomethylated genes; decreased expression of hypermethylated genes). These data are displayed in Table 4.

Notably, many of these DEGs/DMGs play crucial roles in the immune response. Downregulated and hypermethylated genes include: *CD96* which inhibits NK cell and T cell activation; *CD8A*, a T cell surface marker; *IRF4*, a regulator of B cell development, *LCK* which typically activates the T cell receptor and in which mutations

can cause severe combined immunodeficiency; *ITGAL* which mediates immune cell adhesion, and *PTPN22*, an immune marker. Overlapping genes that are hypomethylated and show increased expression in severe as compared to mild COVID-19 cases include: *TLR5*, involved in the response to bacterial pathogens; *IL1R2* which suppresses the immune response; *IKZF1*, a regulator of lymphocyte development; and *NOTCH1*, involved in T cell development.

Table 4 DMGs overlapping with DEGs in severe vs mild COVID-19 as identified in the datasets of Gómez-Carballa et al. and Rombauts et al. [62, 63]. Negative values represent relative hypomethylation in the case of Methylation difference (%) and decreased relative gene expression in the case of analyses of the Gómez-Carballa et al. and Rombauts et al. datasets. Methylation difference and q-value represent data derived from our original dataset. Log₂ Fold Change and P_{adj} represent data from comparative datasets. Rows in boldface indicate genes in which the direction of differential gene expression is concordant with what is expected from the methylation difference noted in our data

Gene	Methylation difference (%)	q-value	Overlapping dataset	Log ₂ Fold Change	P _{adj}
CD96	11.2	1.50E-17	Gómez-Carballa	-1.25	8.56E-04
IRF4	12.4	1.18E-20	Gómez-Carballa	-0.90	8.79E-04
LCK	15.1	5.13E-18	Gómez-Carballa	-0.96	1.50E-03
PTPN22	24.6	5.30E-17	Gómez-Carballa	-0.74	7.69E-03
ITGAL	16.6	2.10E-16	Gómez-Carballa	-0.62	7.69E-03
CD59	10.1	9.74E-23	Gómez-Carballa	0.64	7.69E-03
CD8A	17.3	1.50E-21	Gómez-Carballa	-0.98	9.25E-03
IRAK2	12.5	3.40E-27	Gómez-Carballa	0.82	3.87E-02
TLR5	-16.7	2.71E-18	Gómez-Carballa	0.72	4.63E-05
IL1R2	-15.9	2.49E-16	Gómez-Carballa	1.85	1.06E-04
IKZF1	-14.1	5.54E-16	Gómez-Carballa	0.58	2.45E-03
CD6	-14.6	8.45E-16	Gómez-Carballa	-1.17	7.96E-03
TBX21	-17.4	1.67E-22	Gómez-Carballa	-1.40	9.55E-03
NOTCH1	-20.2	2.23E-16	Gómez-Carballa	0.63	1.17E-02
CD3E	-14.7	4.09E-19	Gómez-Carballa	-1.79	1.24E-02
NFKB1A	-15.8	1.02E-16	Gómez-Carballa	-2.19	1.70E-02
ICOS	-21.7	4.62E-20	Gómez-Carballa	-2.69	2.51E-02
GP1BB	-23.7	5.95E-19	Gómez-Carballa	0.57	3.74E-02
HLA-DPA1	-20.1	1.10E-16	Gómez-Carballa	-1.24	2.00E-03
			Rombauts	-0.83	5.79E-03
CEBPE	27	3.50E-17	Rombauts	0.32	3.47E-02
DPM2	17.9	1.32E-20	Rombauts	-0.26	1.89E-02
LRG1	30.3	3.61E-23	Rombauts	0.50	4.59E-02
PSMG4	15.5	1.10E-21	Rombauts	-0.24	4.13E-02
SEPT1	22.5	2.67E-19	Rombauts	-0.58	1.11E-02
ALYREF	-12.3	7.45E-17	Rombauts	-0.33	4.79E-02
EEF2	-33.7	7.92E-24	Rombauts	-0.34	2.24E-02
HHEX	-12.6	4.84E-20	Rombauts	0.39	3.59E-02
KRT39	-10.1	7.67E-20	Rombauts	0.26	3.04E-02

These overlapping DEGs/DMGs were further evaluated for pathway enrichment using Coronascape [36]. As previously described, hypo- and hypermethylated genes were evaluated separately. Regarding hypomethylated genes, a total of 73 enriched pathways (58 GO processes, 10 KEGG pathways, and 5 Reactome Gene Sets) meeting a significance threshold of $p \leq 0.01$ were identified (Supplemental Table 9A). Regarding hypermethylated genes, a total of 42 enriched pathways (37 GO processes and 5 Reactome Gene Sets) meeting a significance threshold of $p \leq 0.01$ were identified (Supplemental Table 9B). Predominate themes emerging in these enriched pathways included regulation of immune cells and the immune

response, as well as response to infectious pathogens. The top 25 enriched hypo- and hypermethylated pathways as determined by p -value are shown in Fig. 6. In total, this comparative evaluation of differential gene expression and pathway enrichment lends further evidence to the presence of immune dysregulation and alteration in the setting of severe COVID-19.

Differentially methylated genes in severe COVID-19 overlap genetic loci

To add validity to our identified differentially methylated gene list in individuals with severe as compared to mild SARS-CoV-2 infection, we used data from a recent

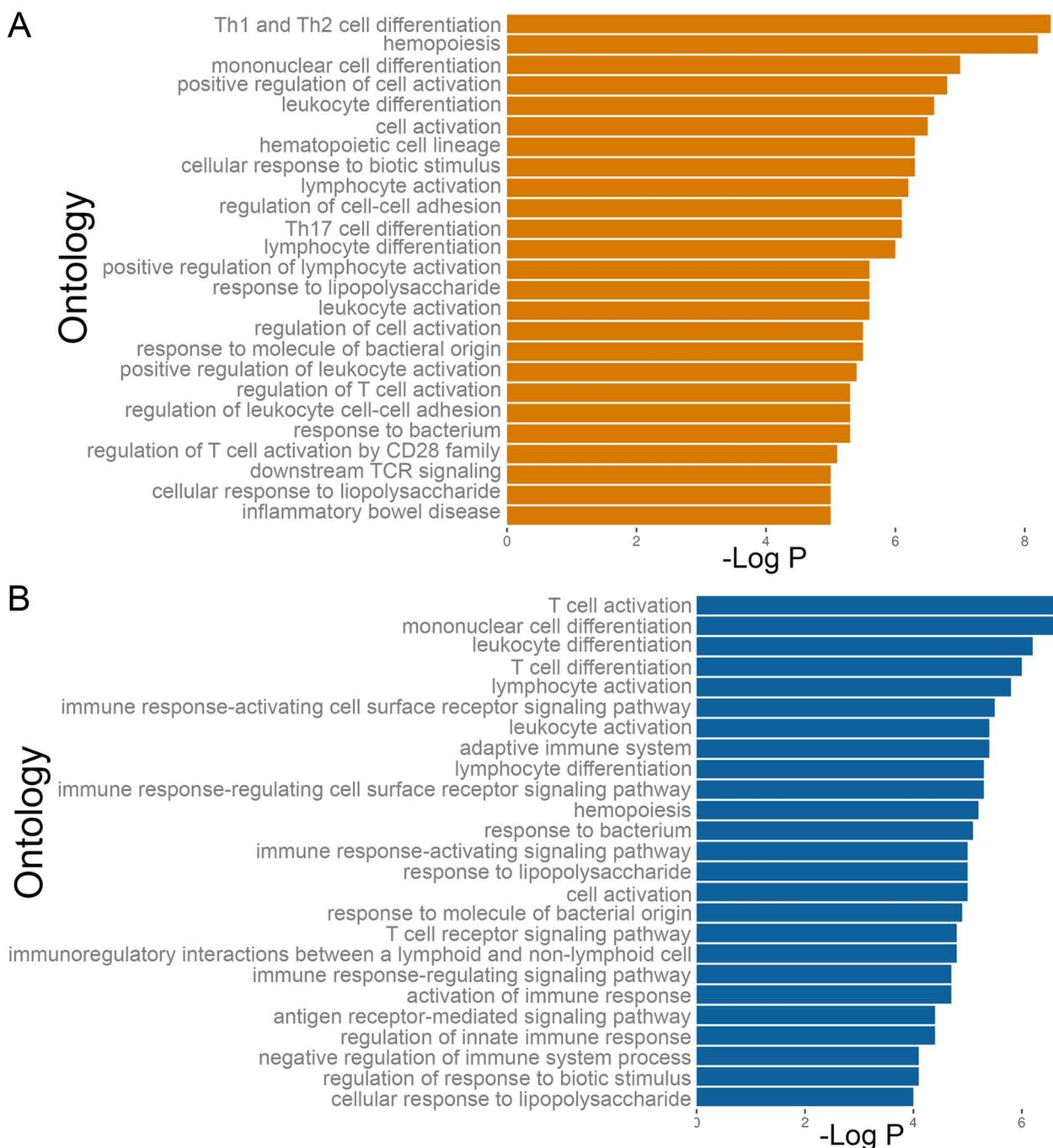


Fig. 6 Differential regulation of immune cells and the response, as well as differential response to foreign pathogens are present in severe vs mild COVID-19. Utilizing the overlapping genes identified in our original dataset and the datasets of Gómez-Carballa *et al.* [62] and Rombauts *et al.* [63], the top 25 most significantly enriched GO Biological Processes, KEGG Pathways, and Reactome Gene Sets are displayed from relatively (A) hypo- and (B) hypermethylated in severe vs mild COVID-19. Ontologies are listed along the y-axis. -Log P values are displayed along the x-axis

GWAS meta-analysis published by the COVID-19 Host Genetics Initiative [64] reporting 23 genetic loci associated with SARS-CoV-2 infection susceptibility and/or COVID-19 disease severity. Within our dataset, we

identified differential methylation of eight genes that overlapped with those implicated in this previously published meta-analysis (Table 5). Within our dataset, we found that relative hypermethylation of *ABO* was present

Table 5 DMGs overlapping with known loci associated with COVID-19 risk

Gene	Chromosome	Start	Stop	Methylation difference (%)	q-value
<i>FDPS</i>	chr1	155,306,501	155,307,000	-12.73	3.32×10^{-19}
<i>CEP97</i>	chr3	101,727,251	101,727,750	-19.01	4.03×10^{-16}
<i>HLA-DPA1</i>	chr6	33,080,751	33,081,250	-20.08	1.07×10^{-16}
<i>OBP2B</i>	chr9	133,211,251	133,211,750	-15.64	8.28×10^{-16}
<i>ABO</i>	chr9	133,275,751	133,276,250	13.43	2.14×10^{-29}
<i>MUC5B</i>	chr11	1,225,751	1,226,250	-23.41	1.11×10^{-16}
<i>PPP1R15A</i>	chr19	48,871,251	48,871,750	-12.33	7.39×10^{-23}
<i>NAPSA</i>	chr19	50,369,501	50,370,000	-35.15	3.38×10^{-43}

in severely as compared to mildly affected individuals. In severe as compared to mild cases, we noted relative hypomethylation of *FDPS*, *CEP97*, *HLA-DPA1*, *OBP2B*, *MUC5B*, *PPP1R15A*, and *NAPSA*.

PI3K/Akt pathway is associated with COVID-19 severity

To support our findings indicating differential regulation of the PI3K/Akt pathway in severe as compared to mild COVID-19 cases, we utilized the gene expression dataset of Gómez-Carballa et al. [62]. We evaluated differential expression of the candidate genes identified in our methylation analysis in the nasal mucosa of individuals with severe ($n=14$) compared to mild ($n=17$) SARS-CoV-2 infection. Among the relatively hypomethylated candidate genes within the PI3K/Akt pathway identified in our dataset (Fig. 3A), Gómez-Carballa and colleagues found increased expression of *NFKBIA* (Log_2 -fold change=0.851, $p_{\text{adj}}=1.70 \times 10^{-2}$), *NOTCH1* (Log_2 -fold change=0.627, $p_{\text{adj}}=1.17 \times 10^{-2}$) and *TLR5* (Log_2 -fold change=0.717, $p_{\text{adj}}=4.63 \times 10^{-5}$). Among the relatively hypermethylated genes in the PI3K/Akt pathway from our dataset, Gómez-Carballa et al. found differential expression of *ITGAL* (Log_2 -fold change=-0.624, $p_{\text{adj}}=7.69 \times 10^{-3}$) and *GFI1* (Log_2 -fold change=-0.751, $p_{\text{adj}}=5.55 \times 10^{-3}$) in individuals with severe compared to mild disease, and in keeping with our findings, both of these genes were significantly downregulated. For further validation, we used publicly available single cell RNA sequencing (scRNA-seq) data from Chua et al. accessed through the UCSC Cell Browser (<https://covid-airways.cells.ucsc.edu>) [44, 65]. In the dataset of Chua et al. we were able to examine scRNA-seq data derived from nasopharyngeal samples of individuals with critical cases of COVID-19, moderate cases, and control subjects (Fig. 7A). These data similarly showed greater expression of *AKT1*, *TLR5*, and *MARK4* in secretory and ciliated cells of individuals with moderate and severe COVID-19 as compared to healthy controls. *ZEB2*, *NOTCH1*, and *ITGB2*, showed increased expression in neutrophils and

non-resident macrophage populations of individuals with severe COVID-19. *IRF7*, *NFKBIA*, and *ISG15* showed upregulation in secretory, ciliated, and squamous cells, as well as in cytotoxic T lymphocytes, T regulatory cells, non-resident macrophages, and neutrophils of severe/moderate cases compared to healthy controls (Fig. 7, Supplemental Fig. 4). These comparative data further strengthen our findings suggesting differential immune cell expression of PI3K/Akt-related genes in the nasopharynx of individuals with severe COVID-19 disease.

Discussion

In our exploration of the nasal epigenome, we highlight differential methylation status as a key correlate of COVID-19 severity. Importantly, our approach of WGBS provides a comprehensive and unbiased evaluation of the nasal methylome in SARS-CoV-2 infection at single base pair resolution. This is in contrast to previous epigenome wide association studies (EWAS) of COVID-19, many of which performed restricted analyses in peripheral blood at predefined loci using array-based methods. As a result, these previous studies were unable to provide a comprehensive whole genome approach to methylation analysis and failed to evaluate epigenetic modulations in the respiratory tract (i.e., the target tissue of SARS-CoV-2) [30, 31].

In this study, we demonstrate increased proportions of the LMR overlapping with immune cell regulatory elements in COVID-19 positive individuals, along with the preponderance hypomethylated DMRs in intergenic regions among individuals with severe disease. In so doing, we shed new light to the importance of enhancer-like regions and distal regulatory elements in immune system regulation as it relates to COVID-19 severity.

As we explore our findings individually, our pathway enrichment analysis suggests the presence of aberrancies in the immune system in the face of severe as opposed to mild COVID-19, which is in keeping with prior studies [43–45]. In the comparison of severe versus mild

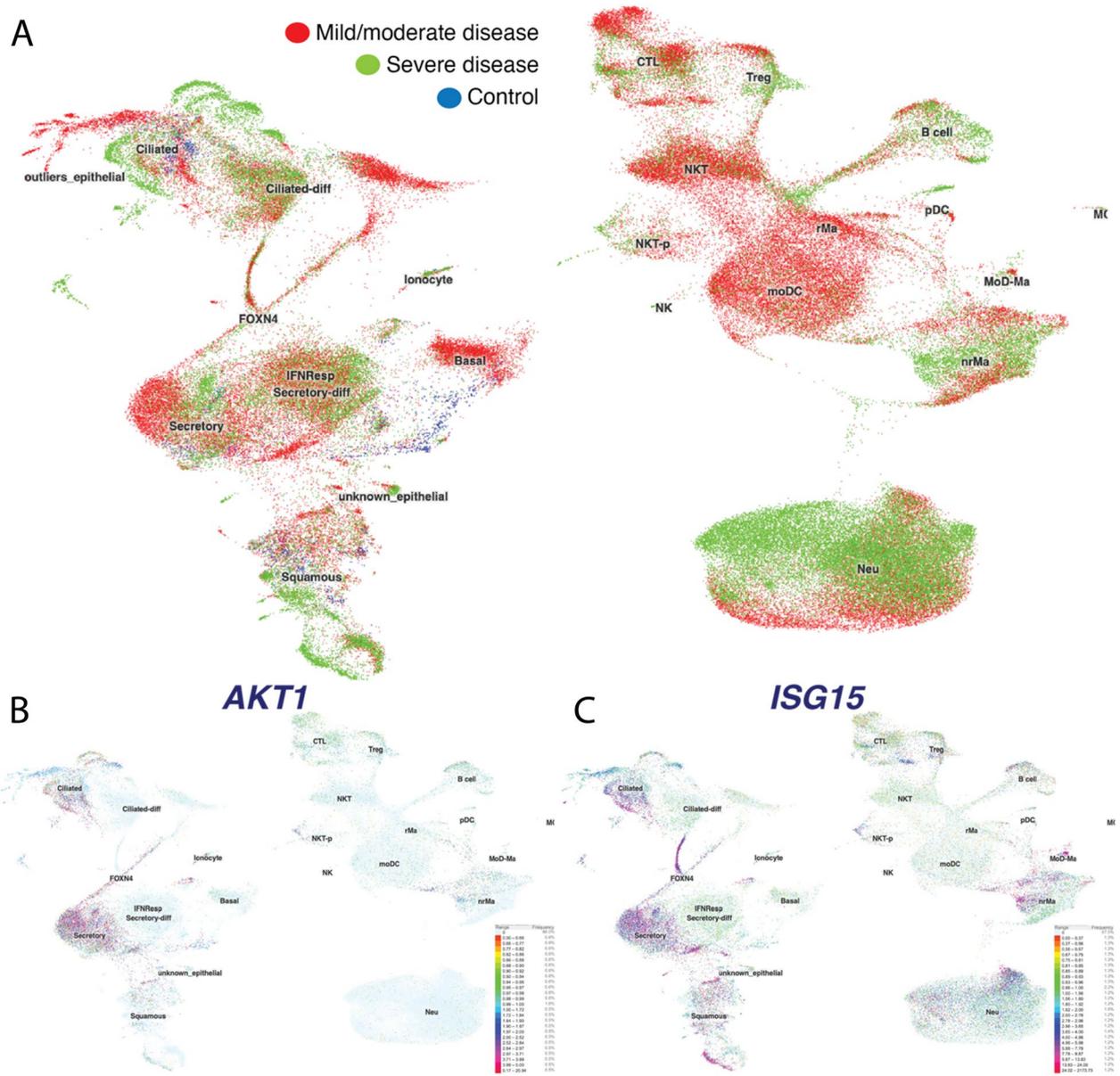


Fig. 7 Increased gene expression of *AKT1* and *ISG15* in the nasopharynx of moderate/severe COVID-19 patients as compared to COVID-19 negative individuals. CTL= Cytotoxic T lymphocytes, MC = Mast cell, moDC = Monocyte-derived dendritic cell, MoD-Ma = Monocyte-derived macrophage, Neu = Neutrophil, NK = Natural killer cell, NKT = NKT cell, NKT-p = Proliferating NKT cell, nrMA = non-resident macrophage, pDC = Plasmacytoid dendritic cell, rMA = Resident macrophage, Treg = Regulatory T cell. UMAP demonstrating differential gene expression by disease severity and cell-type in the nasopharynx. (A) Clustering by COVID-19 severity, Blue = Control, Red= Mild/moderate disease, Green = Critical disease. Gene expression as determined by scRNA-seq of (B) *AKT1*, (C) *ISG15*. Figure generated using the dataset of Chua et al. and the UCSC Cell Browser (<https://covid-airways.cells.ucsc.edu>) [44, 65]

COVID-19, we appreciate differential activation of Th17, Th1, and Th2 as well as T cell selection (hypomethylated), but also note downregulation (hypermethylation) in leukocyte activation, neutrophil degranulation and negative regulation of macrophage differentiation. These findings suggest differential activation of T cell subsets is occurring in severe as compared to mild cases of COVID-19

and indicate impairment or dysregulation in host innate immunity in those individuals who go on to develop more severe phenotypes. However, whether this is a causal relationship is unclear as it is possible that this immune dysregulation is a predisposing factor to development of severe disease but is also conceivable that immune dysfunction arises because of severe illness.

Looking at the specific immune cell lineages and functions that may be dysregulated in severe versus mild disease, our findings show differential methylation of genes impacting innate and adaptive immune cell functions between disease severities. We note the substantial hypermethylation of the *FUT4* promoter in individuals with severe as compared to mild SARS-CoV-2 infection. *FUT4* has previously been highlighted as a marker of immature neutrophil or “proneutrophil” populations in the examination of peripheral blood of patients with severe COVID-19 [43, 45, 49]. In contrast to our finding of striking hypermethylation of the *FUT4* promoter, suggesting decreased gene expression in those with severe COVID-19, prior studies of SARS-CoV-2 infection demonstrated circulating *FUT4*⁺ neutrophils to be present in greater abundance in individuals with severe COVID-19 when compared to healthy controls and patients with mild COVID-19 cases [43, 45, 51, 52]. However, more recently and in keeping with our data, Karawajczyk et al. showed decreased neutrophilic expression of the cell surface marker and gene product of *FUT4*, CD15, in severe COVID-19 patients [47]. They focused on the functional role of CD15 as an adhesion molecule, noting the simultaneous lack of upregulation of other adhesion molecules in the presence of severe infection. The relative hypermethylation of the *FUT4* promoter in severe disease in our dataset, taken together with the significantly enriched GO terms involved in leukocyte migration, adhesion, and tethering among hypermethylated genes implicate impaired localization of leukocytes as a potential pathophysiologic player or respondent in severe COVID-19. Building on this further, our data indicate that in a broader sense, neutrophil dysregulation may be associated with COVID-19 severity. This is supported by our finding of relative hypermethylation of genes implicated in neutrophil degranulation (e.g., *ELANE*, *AZU1*, *CD59*, *GSDMD*, *SERPINB1*) in severe COVID-19.

We also identified relative hypermethylation of *CD8A* in individuals with severe as compared to mild COVID-19. Throughout the COVID-19 pandemic, data has accumulated demonstrating the importance of CD8⁺ T cells in the antiviral response to SARS-CoV-2 [8, 46, 48, 50, 66, 67]. Robust populations of SARS-CoV-2-specific CD8⁺ T cells are associated with mild COVID-19 presentations [50], whereas CD8⁺ T cell depletion and exhaustion have been shown to correlate with worsened disease severity and increased mortality [8, 46, 48]. It is possible that hypermethylation of genes responsible for the CD8 antigen (e.g., *CD8A*) contribute to the decreased expression of CD8⁺ T cells and is among the reasons behind the CD8⁺ cytopenia observed in severe COVID-19 cases.

In addition to differential immune cell line regulation between COVID-19 severities, we also appreciated

differential methylation of genes related to the inflammatory and cytokine responses. For example, relative hypomethylation of *PPP1R15A* in individuals with severe COVID-19. *PPP1R15A* is a stress-response gene. Prior studies indicate increased levels of *PPP1R15A* expression is present within immune cells of severely affected COVID-19 patients [18, 68]. More specifically, *PPP1R15A* has been reported to have highest expression in immune cells containing the highest levels of viral RNA [68]. It is thought that this increased expression of *PPP1R15A* may contribute to COVID-19 severity through the induction of proinflammatory cytokines and by enhancing the survival and multiplication of infected cells [18].

Among the most important differentially methylated pathways between severe and mild disease in our dataset was the differential methylation within the PI3K/Akt signaling pathway. The PI3K/Akt pathway is critical to regulation of IFN signaling and IFN effector genes as part of the host’s antiviral response. Increasingly, studies have demonstrated delayed or decreased type I IFN (IFN-I) and type III IFN (IFN-III) responses in severe COVID-19 cases [15, 69–72]. We appreciated hypomethylation of the interferon-stimulated gene, *ISG15*, in individuals with severe as compared to mild COVID-19, which is consistent with findings of the single cell profiling of airway cells in the COVID-19 Cell Atlas demonstrating *ISG15* as among the top three most differentially expressed genes between severe COVID-19 disease and non-severe cases (https://www.covid19cellatlas.org/meyer21_airway/) [73]. *ISG15* was suggested by Munnur et al. to be involved in COVID-19 pathogenesis at multiple levels: SARS-CoV-2 stimulates the release of intracellular interferon-stimulated gene 15 (ISG15) from infected macrophages, and extracellular ISG15 acts to exaggerate the cytokine/chemokine inflammatory response [74]. Interestingly, we found within our data that some IFN-I genes were downregulated in severe as compared to mild COVID-19 cases (i.e., *IRF1*), however other positive regulators of the IFN-I responses (e.g., *IRF3*, *IRF7*, *IRF8*) were hypomethylated indicative of increased expression. We also noted relative hypermethylation of *ELF4* targets among individuals with severe disease. *ELF4* has a crucial role in the host antiviral response including contributing to NK cell development and function, regulating cell cycle arrest in naïve CD8⁺ cells in the face of viral infection, and driving IFN-I responses [57, 75, 76]. Cumulatively, these findings emphasize that disordered regulation of the IFN response may be associated with a more severe COVID-19 phenotype, either as a cause of or in response to severe disease.

Even prior to the SARS-CoV-2 pandemic, the PI3K/Akt/mTOR pathway was recognized as crucial to the pathogenesis of other coronaviruses, namely the related Middle East respiratory syndrome coronavirus

(MERS-CoV) in which in vitro studies demonstrated inhibition of this pathway could block viral proliferation [77, 78]. Building on this historical role of PI3K/Akt in RNA coronaviruses, PI3K/Akt signaling has been implicated in SARS-CoV-2 pathogenesis in multiple organ systems and studies have demonstrated inhibition of SARS-CoV-2 replication in response to PI3K/Akt/mTOR blockade [77, 79, 80].

A possible explanation for the differential immune cell recruitment and inflammatory pathway activation seen in our dataset could be related to the concept of immune tolerance. Immune tolerance typically refers to an immune cell's inability to activate gene transcription and perform its function in response to restimulation by a previously encountered antigen [81]. However, it has also been demonstrated that exposure to one pathogen can induce tolerance of the immune response to an unrelated pathogen (i.e., "heterologous immune tolerance") [82]. Regardless of the primary exposure, immune tolerance leads to a less effective response to secondary stimuli.

Though not specifically evaluating SARS-CoV-2, Habibi et al. recently investigated the concept of differential mucosal immunity in the contraction of symptomatic respiratory syncytial virus (RSV) [83]. They acknowledged that despite all adults having exposure to RSV, that even healthy individuals experience repeated RSV reinfection. In this elegant experiment, they administered RSV to healthy volunteers and evaluated differential gene expression in the nasal mucosa of those individuals who developed symptomatic RSV as compared to those who remained RSV PCR negative despite inoculation. Similar to our findings, they found differential activation of immune cells, namely prior activation of neutrophils, seemed to predispose to symptomatic viral infection. Based on this human subjects research and correlating mouse models, they postulated that preexisting neutrophilic inflammation alters the tissue environment so that the recruitment of CD8+ T cells to the lung is increased later in the disease course thereby leading to a more severe phenotype. Though it remains fully possible that the differential methylation of the nasal epigenome in our study could be the result disease severity, rather than a predisposing factor to development of severe disease, these findings by Habibi et al. strengthen the argument that baseline differences in the existing nasal immune cell landscape, perhaps due to prior exposures, could play an important role in the severity of respiratory illnesses. The subjects of our current study were evaluated prior to the advent of the SARS-CoV-2 vaccination initiative; however it is possible that remote exposures to viruses with homologous antigens to SARS-CoV-2, including the seasonal human coronaviruses (HCoVs) that are most typically associated with mild respiratory disease, may have

induced this tolerance to the SARS-CoV-2 virus. The concept of immune tolerance may be particularly important in the context of methylation given that methylation changes can have long-term impacts on gene expression that propagate across cell divisions. In fact, emerging evidence indicates that DNA methylation changes in blood associated with lymphocyte activation and the immune response persist on a longitudinal basis, and factors regulating chromatin accessibility may be particularly important in the response to RNA-viruses [29, 84]. Our identification of multiple biologic processes associated with response to a variety of infectious stimuli within our dataset and supportive datasets [62, 63] lends further evidence to the postulation that immune tolerance may be associated with COVID-19 severity. It is possible that methylation status of the immune regulatory genes could have been altered in response to remote exposures and modulated the host's COVID-19 severity risk.

We found a relative degree of hypermethylation of *ELF5* binding sites among individuals with severe as compared to mild COVID-19. Pietzner et al. recently demonstrated increased *ELF5* expression within respiratory epithelial cells as a risk factor for severe COVID-19 [58]. However, they also noted substantially diminished *ELF5* expression in the injured olfactory mucosa in individuals who experienced rapid death secondary to severe COVID-19 as compared to healthy controls. This finding of decreased *ELF5* expression in the olfactory mucosa is in keeping with our noted hypermethylation of *ELF5* targets in these nasal mucosal swabs of individuals with severe COVID-19.

Within our dataset, relatively hypermethylated targets of *ELF5* in individuals with severe as compared to mild COVID-19 included *C1orf116* and *PLAC8* (mediators of viral entry), and *IFRD1* (an interferon-stimulated gene). Interestingly and in contrast to our data, COVID-19 studies using bronchoalveolar lavage samples, cell cultures, and in silico models, *C1orf116*, *PLAC8*, and *IFRD1* have been overexpressed in severe cases [59–61]. It is possible that these differences arise due to differential cell populations of study (i.e., nasal mucosa as compared to the more distal respiratory tract). The nasopharynx represents the initial interface between host and the SARS-CoV-2 virus, and as such, the nasal mucosa plays a critical role in inducing early innate and acquired immune responses [85]. It is possible that hypermethylation of these *ELF5* targets in the nasal mucosa contribute to impairments in early encounters between host and virus. This could result in delayed local and systemic responses, thereby providing the virus opportunity to infect distal airways before meeting host defenses. Similarly, it is possible that the hypermethylation of these *ELF5* targets could arise as the result of severe disease could leave the host

vulnerable to additional respiratory insults, thereby worsening their disease courses.

Our study does have some limitations. We have a relatively small sample size, particularly with regard to our severe COVID-19 cohort. In this way, we may have been limited in our ability to identify significant associations of methylation and disease severity, as these may have been masked by interindividual variability within the severely affected group. Similarly, our small sample size has the potential of being underpowered, however, we were unable to perform a power analysis as this requires knowledge of effect size (i.e., the expected magnitude of association between an epigenetic variant specific to the nasal mucosa and COVID-19 severity) which is currently ill-defined. We overcame this limitation through use of a conservative definition of DMR, average methylation difference >10% and q -value < 0.01. However, as we only included the top 10,000 most significant DMRs in our analysis as determined by q -value, our threshold for inclusion was substantially more stringent with evaluated DMRs having a q -value < 1.20×10^{-15} . We further strengthened the confidence in our results through the corroboration of our findings with multiple publicly available genomic and transcriptomic datasets. As our samples were collected at the time of diagnosis, it is difficult to discern whether the methylation differences we appreciated were the cause of severe versus mild disease outcomes, or if these differences were the result of having severe as compared to mild disease. Finally, though we can extrapolate patterns of gene expression based on our knowledge of methylation and based on the transcription analyses of others, we do not have direct measures of gene expression for the individuals within our dataset.

Conclusions

This whole genome interrogation of the nasal methylome suggests that methylation is linked to the host immune response to SARS-CoV-2 infection. It is difficult to discern whether these methylation differences between severe and mild disease are contributory to the severity of COVID-19, or if these epigenetic changes occur in response to the severity of illness. Differences in the nasal methylome between individuals with severe as compared to mild COVID-19 appear to modulate innate immunity through disruptions in neutrophil adhesion, localization, and degranulation. In the adaptive immune response, differential methylation between individuals with severe and mild disease may lead to alterations in T cell populations. In part these differences in immune response and differential regulation of inflammatory pathways (e.g., PI3K/Akt pathway) could be associated with immune tolerance. Further, impairments in the early immune defenses of the nasal mucosa may be related to COVID-19 severity.

These findings highlight the continued need for exploration into potential causative pathways as we seek to gain understanding of the SARS-CoV-2 viral pathophysiology and gives evidence supporting investigation of these paths as putative therapeutic targets. Further, this study emphasizes the need to expand studies more broadly to enhance statistical power, and to perform longitudinal studies that include individuals prior to first SARS-CoV-2 infection so as to better elucidate whether these identified mechanisms are the cause of severe disease or if they reflect a response to disease severity.

Methods

WGBS sample characteristics

Salvage nasal mucosa derived from patients presenting to the emergency department at University Health Truman Medical Center were accessed and collected from mid-turbinate nasal flocked swabs as part of routine testing for SARS-CoV-2 infection. Individuals were defined as being positive for COVID-19 if routine clinical PCR-based testing for SARS-CoV-2 yielded a positive result; individuals were defined as COVID-19 negative if clinical PCR-based testing for SARS-CoV-2 yielded a negative result. Positive subjects were defined as having severe disease if hospital admission was required, and were defined as having mild disease if they did not require hospitalization. Samples were obtained from 4 individuals with severe COVID-19, 57 with mild disease, and two pools of COVID-19 negative individuals ($n=8$ and $n=7$, respectively). Samples were stored in 3 mL of Universal Transport Medium where 200 μ L of each specimen was tested for SARS-CoV-2 and remaining aliquot was saved in -80°C freezer.

DNA Isolation

Nasal specimens were stored at -80°C and were brought to room temperature. DNA was isolated with a DNeasy Blood and Tissue Kit (Qiagen, Cat No. 69504) with the following modifications to kit protocol: 8 μ L of RNase A was used instead of 4 μ L during the optional RNase A step and the lysis incubation time at 56°C was increased to at least 3 h to ensure complete lysis of the specimens. After isolation, the DNA concentration of each sample was determined using a Qubit dsDNA HS Assay Kit (Fisher, Cat No. Q32851).

WGBS library preparation and sequencing

A minimum of 100 ng of DNA was aliquoted from each sample. Unmethylated λ DNA was added to each sample at 0.5% w/v and the samples were sheared mechanically using a Covaris LE220-plus system to a length of 350 bp, using the settings recommended by the manufacturer. The sizing was

determined by a High Sensitivity D1000 ScreenTape and Reagents (Agilent, Cat. No. 5067–5584 and 5067–5585) on the TapeStation platform. Once the input DNA was at the proper fragment size, the samples were concentrated with a SpeedVac to a volume of 20 μ L. The samples then underwent bisulfite conversion with an EZ DNA Methylation- Gold kit (Zymo, Cat. No. D5006). The samples were eluted off the spin columns with 15 μ l of low EDTA TE buffer (Swift, Cat. No. 30024) before library preparation.

The low-input libraries were prepared using an ACCEL-NGS Methyl-Seq Library kit (Swift, Cat. No. 30024) with a Methyl-Seq Set A Indexing Kit (Swift, Cat. No. 36024), following the protocol associated with the library kit. During the protocol, bead cleanup steps were performed with SPRIselect beads (Beckman Coulter, Cat. No. B23318). Following the recommendation of the kit, 6 PCR cycles were performed to amplify the samples. The final libraries were quantified with a Qubit dsDNA HS Assay Kit and the size was determined by using a BioAnalyzer High Sensitivity DNA Kit (Agilent, Cat. No. 5067–4626). The libraries were then sequenced on the Illumina NovaSeq6000 System using 150 bp paired-end sequencing.

WGBS data processing

WGBS data was processed using the Epigenome Pipeline available from the DRAGEN Bio-IT platform (Edico Genomics/Illumina). Sequence reads were demultiplexed into FASTQ files using Illumina's bcl2Fastq2-2.19.1 software and trimmed for quality (phred33 \geq 20) and Illumina adapters using trimgalore v.0.4.2 (<https://github.com/FelixKrueger/TrimGalore>). Reads were then aligned to the bisulfite-converted GRCh38 reference genome using DRAGEN EP v2.6.3 in paired-end mode using the directional/Lister methylation protocol presets. Alignments were calculated for both Watson and Crick strands and the highest quality unique alignment was retained. Duplicated reads were removed using picard v 2.17.8 [86]. A genome-wide cytosine methylation report was generated by DRAGEN to record counts of methylated and unmethylated cytosines at each cytosine position in the genome. Methylation counts were provided for the CpG, CHG and CHH cytosine contexts but only CpG was considered in the study. To avoid potential biases in downstream analyses, CpGs were further filtered by removing CpGs: covered by five or less reads, and located within genomic regions that are known to have anomalous, unstructured, high signal/read counts as reported in DAC blacklisted regions (DBRs) or Duke excluded regions (DERs) generated by the ENCODE project [87].

Differential methylation analysis

Filtered methylation data from all nasal samples were merged according to disease severity. Only CpGs covered by at least 10 reads and present in at least 2 samples per

group (50% of the severe sample size) were kept. DMRs of destrand autosomes were evaluated in an overlapping tiling window analysis with window size 500 bp and step size 250 bp through a logistic regression analysis with age, gender, and race included as covariates using the R-package, methylKit [88]. *P*-values were adjusted to *q*-values using SLIM method.

Comparative analysis of differential proportions of hypo- versus hypermethylated regions was carried out using Chi-square test of independence.

Gene annotation

Following differential methylation analysis, gene annotation was limited to those bins with a *q*-value $<$ 0.01 and an absolute average methylation difference of $>$ 10% between comparison groups. In the case of evaluation of DMRs between hospitalized versus non-hospitalized subjects, gene annotation was limited to those 10,000 most statistically significant differences by *q*-value. Initial gene annotation was performed using Genomic Regions Enrichment of Annotations Tool (GREAT) algorithm [33, 34] (Association rule: Single nearest gene: 5000 bp max extension, curated regulatory domains included). Subsequent gene ontology was further explored using Coronascape with hypo- and hypermethylated DMRs evaluated separately. Additional pathway analysis was performed using QIAGEN Ingenuity Pathway Analysis (IPA) version 01–21–03 (QIAGEN, Venlo, Netherlands) examining all gene-associated loci with a *q*-value of $<$ 0.05.

When evaluating the supportive transcriptomic datasets of Gómez-Carballea et al. (GSE 183071, accessible from: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE183071>) [62] and Rombauts et al. (GSE212865, accessible from: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE212865>) [63], DEGs were evaluated using GEO2R software. GEO2R is an interactive web tool that allows comparison of two or more groups of samples in a GEO series to identify differentially expressed genes across experimental conditions. In the case of RNA-seq data, GEO2R uses the R package, *DESeq2* [89], to perform differential expression analysis using NCBI-computed raw count matrices as input. *DESeq2* uses negative binomial generalized linear models and has features that offer consistent performance over a large range of data types. In the case of the dataset of Rombauts et al., differentially expressed genetic loci rather than differentially expressed genes were provided. As such, the GREAT algorithm was applied to significantly differentially expressed loci using the same association rule as described above.

Methylation segmentation

UMRs and LMRs for each samples set were called based on a pooled aggregate methylation profile across the

samples using MethylSeekR package (v 1.38) [90] from Bioconductor (v 3.16) [91, 92].

MethylSeekR: Burger L, Gaidatzis D, Schubeler D, Stadler MB (2013). "Identification of active regulatory regions from DNA methylation data." *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkt599>, <http://nar.oxfordjournals.org/content/early/2013/07/04/nar.gkt599.long>.

Bioconductor: <http://www.nature.com/nmeth/journal/v12/n2/abs/nmeth.3252.html>, <https://genomebiology.biomedcentral.com/articles/https://doi.org/10.1186/gb-2004-5-10-r80>.

Annotation of regulatory elements

Genomic regions were further annotated for overlaps with the entire DNase I Hypersensitive Site (DHS) vocabulary using the *intersect* function in the Bedtools suite (v 2.30.0) [93] with minimum overlap of 1 nucleotide. The DHS coordinates were accessed from https://zenodo.org/record/3838751/files/DHS_Index_and_Vocabulary_hg38_WM20190703.txt.gz using 16 different vocabulary representatives as outlined in Meuleman et al., 2020 [32].

BedTools: Quinlan AR and Hall IM, 2010. *BEDTools: a flexible suite of utilities for comparing genomic features*. *Bioinformatics*. 26, 6, pp. 841–842.

Transcription factor binding analysis

Transcription factor binding site (TFBS) motif analysis was performed using the Homer software (HOMER find-MotifsGenome.pl v4.11.1) [90] using the central 200 bp of regions. Motif analysis was performed using HOMER software examining hypo- and hypermethylated regions separately, however, DMRs were expanded to include all regions with an absolute methylation difference of > 10% between groups and a q -value 1×10^{-5} .

Abbreviations

SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
COVID-19	Coronavirus disease 2019
ACE 2	Angiotensin 2
GWAS	Genome-wide association studies
SNV	Single nucleotide variants
IFN	Interferon
DMR	Differentially methylated region
EWAS	Epigenome wide association studies
WGBS	Whole-genome bisulfite sequencing
PI3K/Akt	Phosphoinositide 3-kinase/serine-threonine kinase
NF- κ B	Nuclear factor kappa B
UMR	Unmethylated regions
LMR	Low methylated regions
DHS	DNase I hypersensitive site
GREAT	Genomic Regions Enrichment of Annotations Tool
DMG	Differentially methylated gene
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
scRNA-seq	Single cell RNA sequencing
IFN-I	Type I interferon
IFN-III	Type III interferon
ISG15	Interferon-stimulated gene 15
MERS-CoV	Middle East respiratory syndrome coronavirus
HCoVs	Human coronaviruses

RSV	Respiratory syncytial virus
DBR	DAC blacklisted region
DER	Duke excluded region
IPA	Ingenuity Pathway Analysis
DEG	Differentially expressed gene
TFBS	Transcription factor binding site
IRB	Institutional Review Board

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12920-025-02125-4>.

Additional file 1.

Additional file 2.

Acknowledgements

The authors would like to thank Daniel Louiselle and Rebecca Biswell for their work in sample processing. EG holds the Roberta D. Harding & William F. Bradley, Jr. Endowed Chair in Genomic Research. BLS was supported in part by The Sam and Helen Kaplan Research Fund in Pediatric Nephrology, and The McLaughlin Family Endowed Chair in Nephrology.

Authors' contributions

The project was conceptualized by EG and RS. Clinical coordination was performed by DB and RS. Data and bioinformatics analyses were done by BS and BK. BS and RM prepared all figures. BS wrote the original manuscript draft with contributions from all authors. All authors read and approved the final manuscript. EG supervised the project and was responsible for funding acquisition.

Funding

This work was supported by a CTSA grant from NCATS awarded to the University of Kansas for Frontiers: University of Kansas Clinical and Translational Science Institute (# UL1TR002366) The contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH or NCATS.

Data availability

The datasets generated and/or analysed during the current study are available in the BioProject and Gene Expression Omnibus (GEO) repositories. WGBS data from SARS-CoV-2 positive individuals are available under BioProject ID PRJNA1162448 (<http://www.ncbi.nlm.nih.gov/bioproject/1162448>). WGBS data from SARS-CoV-2 negative individuals are available GEO accession number: GSE168254 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE168254>).

Declarations

Ethics approval and consent to participate

The Office of Research Integrity at the Children's Mercy Research Institute has determined that the proposed activity under this study (STUDY00001258) does not involve human subjects as defined by DHHS regulations (45 CFR 46.102(e)). Therefore, the Office of Research Integrity has determined that this project does not meet the definition of research involving human subjects under 45 CFR 46.102(e) and IRB review and approval by this organization are not indicated.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Pediatrics, University of Wisconsin School of Medicine and Public Health, 600 Highland Ave, Madison, WI 53792, USA. ²Department of Pediatrics, Genomic Medicine Center, Children's Mercy Kansas City, 2401 Gillham Rd, Kansas City, MO 64108, USA. ³Department of Pathology

and Laboratory Medicine, Children'S Mercy Kansas City, 2401 Gillham Rd, Kansas City, MO 64108, USA. ⁴Department of Pathology, University Health, University of Missouri- Kansas City School of Medicine, 2411 Holmes St, Kansas City, MO 64108, USA.

Received: 3 September 2024 Accepted: 12 March 2025

Published online: 01 April 2025

References

- WHO Coronavirus (COVID-19) Dashboard. Available from <https://covid19.who.int/>. Cited 2024 Mar 01
- Xu J, Murphy SL, Kochanek KD, Arias E. Mortality in the United States. *NCHS Data Brief*. 2021;2022:1–8.
- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan. *China Lancet*. 2020;395:497–506.
- Elezkurtaj S, Greuel S, Ihlow J, Michaelis EG, Bischoff P, Kunze CA, Sinn BV, Gerhold M, Hauptmann K, Ingold-Heppner B, et al. Causes of death and comorbidities in hospitalized patients with COVID-19. *Sci Rep*. 2021;11:4263.
- Kermali M, Khalsa RK, Pillai K, Ismail Z, Harky A. The role of biomarkers in diagnosis of COVID-19 - A systematic review. *Life Sci*. 2020;254: 117788.
- Liao M, Liu Y, Yuan J, Wen Y, Xu G, Zhao J, Cheng L, Li J, Wang X, Wang F, et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat Med*. 2020;26:842–4.
- You M, Chen L, Zhang D, Zhao P, Chen Z, Qin EQ, Gao Y, Davis MM, Yang P. Single-cell epigenomic landscape of peripheral immune cells reveals establishment of trained immunity in individuals convalescing from COVID-19. *Nat Cell Biol*. 2021;23:620–30.
- Jiang M, Guo Y, Luo Q, Huang Z, Zhao R, Liu S, Le A, Li J, Wan L. T-Cell Subset Counts in Peripheral Blood Can Be Used as Discriminatory Biomarkers for Diagnosis and Severity Prediction of Coronavirus Disease 2019. *J Infect Dis*. 2020;222:198–202.
- Wen W, Su W, Tang H, Le W, Zhang X, Zheng Y, Liu X, Xie L, Li J, Ye J, et al. Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing. *Cell Discov*. 2020;6:31.
- Wilk AJ, Rustagi A, Zhao NQ, Roque J, Martínez-Colón GJ, McKechnie JL, Ivson GT, Ranganath T, Vergara R, Hollis T, et al. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat Med*. 2020;26:1070–6.
- Ali A, Vijayan R. Dynamics of the ACE2-SARS-CoV-2/SARS-CoV spike protein interface reveal unique mechanisms. *Sci Rep*. 2020;10:14214.
- Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*. 2020;367:1444–8.
- Yang J, Petitjean SJL, Koehler M, Zhang Q, Dumitru AC, Chen W, Derclaye S, Vincent SP, Soumillon P, Alsteens D. Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. *Nat Commun*. 2020;11:4541.
- Matuozzo D, Talouarn E, Marchal A, Zhang P, Manry J, Seeleuthner Y, Zhang Y, Bolze A, Chaldebass M, Milisavljevic B, et al. Rare predicted loss-of-function variants of type I IFN immunity genes are associated with life-threatening COVID-19. *Genome Medicine*. 2023;15:22.
- Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, Ogiishi M, Sabli IKD, Hodeib S, Korol C, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science*. 2020;370(6515):eabd4570.
- Zhang Q, Cobat A, Bastard P, Notarangelo LD, Su HC, Abel L, Casanova JL. Association of rare predicted loss-of-function variants of influenza-related type I IFN genes with critical COVID-19 pneumonia. *J Clin Invest*. 2021;131(15):e152474.
- Li S, Ma F, Yokota T, Garcia G, Jr., Palermo A, Wang Y, Farrell C, Wang YC, Wu R, Zhou Z, et al. Metabolic reprogramming and epigenetic changes of vital organs in SARS-CoV-2-induced systemic toxicity. *JCI Insight*. 2021;6(2):e145027.
- Li Y, Duche A, Sayer MR, Roosan D, Khalafalla FG, Ostrom RS, Totonchy J, Roosan MR. SARS-CoV-2 early infection signature identified potential key infection mechanisms and drug targets. *BMC Genomics*. 2021;22:125.
- Mahmoodpoor A, Sanaie S, Roudbari F, Sabzevari T, Sohrabifar N, Kazeminasab S. Understanding the role of telomere attrition and epigenetic signatures in COVID-19 severity. *Gene*. 2022;811: 146069.
- Pruimboom L. Methylation Pathways and SARS-CoV-2 Lung Infiltration and Cell Membrane-Virus Fusion Are Both Subject to Epigenetics. *Front Cell Infect Microbiol*. 2020;10:290.
- Sen R, Garbati M, Bryant K, Lu Y. Epigenetic mechanisms influencing COVID-19. *Genome*. 2021;64:372–85.
- Shirvallioo M. Epigenomics in COVID-19; the link between DNA methylation, histone modifications and SARS-CoV-2 infection. *Epigenomics*. 2021;13:745–50.
- Spector BL, Harrell L, Sante D, Wyckoff GJ, Willig L. The methylome and cell-free DNA: current applications in medicine and pediatric disease. *Pediatr Res*. 2023;94:89–95.
- Laurent L, Wong E, Li G, Huynh T, Tsirigos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J, Wei CL. Dynamic changes in the human methylome during differentiation. *Genome Res*. 2010;20:320–31.
- Busche S, Shao X, Caron M, Kwan T, Allum F, Cheung WA, Ge B, Westfall S, Simon MM, Barrett A, et al. Population whole-genome bisulfite sequencing across two tissues highlights the environment as the principal source of human methylome variation. *Genome Biol*. 2015;16:290.
- Tsaprouni LG, Yang TP, Bell J, Dick KJ, Kanoni S, Nisbet J, Viñuela A, Grundberg E, Nelson CP, Meduri E, et al. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics*. 2014;9:1382–96.
- Gruenbaum Y, Naveh-Many T, Cedar H, Razin A. Sequence specificity of methylation in higher plant DNA. *Nature*. 1981;292:860–2.
- Straussman R, Nejman D, Roberts D, Steinfeld I, Blum B, Benvenisty N, Simon I, Yakhini Z, Cedar H. Developmental programming of CpG island methylation profiles in the human genome. *Nat Struct Mol Biol*. 2009;16:564–71.
- Balnis J, Madrid A, Hogan KJ, Drake LA, Adhikari A, Vancavage R, Singer HA, Alisch RS, Jaitovich A. Whole-Genome Methylation Sequencing Reveals that COVID-19-induced Epigenetic Dysregulation Remains 1 Year after Hospital Discharge. *Am J Respir Cell Mol Biol*. 2023;68:594–7.
- Corley MJ, Pang APS, Dody K, Mudd PA, Patterson BK, Seethamraju H, Bram Y, Peluso MJ, Torres L, Iyer NS, et al. Genome-wide DNA methylation profiling of peripheral blood reveals an epigenetic signature associated with severe COVID-19. *J Leukoc Biol*. 2021;110:21–6.
- Castro de Moura M, Davalos V, Planas-Serra L, Alvarez-Erriro D, Arribas C, Ruiz M, Aguilera-Albessa S, Troya J, Valencia-Ramos J, Vélez-Santamaría V, et al. Epigenome-wide association study of COVID-19 severity with respiratory failure. *EBioMedicine*. 2021, 66:103339.
- Meuleman W, Muratov A, Rynes E, Halow J, Lee K, Bates D, Diegel M, Dunn D, Neri F, Teodosiadis A, et al. Index and biological spectrum of human DNase I hypersensitive sites. *Nature*. 2020;584:244–51.
- McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*. 2010;28:495–501.
- Tanigawa Y, Dyer ES, Bejerano G. WhichTF is functionally important in your open chromatin data? *PLoS Comput Biol*. 2022;18: e1010378.
- Neidhart M. Chapter 6 - DNA Methylation and Viral Infections. In: Neidhart M, editor. *DNA Methylation and Complex Human Disease*. Oxford: Academic Press; 2016. p. 81–102.
- Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10:1523.
- Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, Ebert D, Feuermann M, Gaudet P, Harris NL, Hill DP, et al. The Gene Ontology knowledgebase in 2023. *Genetics*. 2023;224(1):iyad031.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium Nat Genet*. 2000;25:25–9.
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27–30.
- Milacic M, Beavers D, Conley P, Gong C, Gillespie M, Griss J, Haw R, Jassal B, Matthews L, May B, et al. The Reactome Pathway Knowledgebase 2024. *Nucleic Acids Res*. 2024;52:D672–8.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene

- set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci*. 2005;102:15545–50.
42. Ruepp A, Brauner B, Dunger-Kaltenbach I, Frishman G, Montrone C, Stransky M, Waegle B, Schmidt T, Doudieu ON, Stümpflen V, Mewes HW. CORUM: the comprehensive resource of mammalian protein complexes. *Nucleic Acids Res*. 2008;36:D646–650.
 43. Aschenbrenner AC, Mouktaroudi M, Krämer B, Oestreich M, Antonakos N, Nuesch-Germano M, Gkizeli K, Bonaguro L, Reusch N, Baßler K, et al. Disease severity-specific neutrophil signatures in blood transcriptomes stratify COVID-19 patients. *Genome Med*. 2021;13:7.
 44. Chua RL, Lukassen S, Trump S, Hennig BP, Wendisch D, Pott F, Debnath O, Thürmann L, Kurth F, Völker MT, et al. COVID-19 severity correlates with airway epithelium-immune cell interactions identified by single-cell analysis. *Nat Biotechnol*. 2020;38:970–9.
 45. Schulte-Schrepping J, Reusch N, Paclik D, Baßler K, Schlickeiser S, Zhang B, Krämer B, Krammer T, Brumhard S, Bonaguro L, et al. Severe COVID-19 Is Marked by a Dysregulated Myeloid Cell Compartment. *Cell*. 2020;182:1419–1440.e1423.
 46. Diao B, Wang C, Tan Y, Chen X, Liu Y, Ning L, Chen L, Li M, Liu Y, Wang G, et al. Reduction and Functional Exhaustion of T Cells in Patients With Coronavirus Disease 2019 (COVID-19). *Front Immunol*. 2020;11:827.
 47. Karawajczyk M, Douhan Håkansson L, Lipcsey M, Hultström M, Pauksens K, Frithiof R, Larsson A. High expression of neutrophil and monocyte CD64 with simultaneous lack of upregulation of adhesion receptors CD11b, CD162, CD15, CD65 on neutrophils in severe COVID-19. *Ther Adv Infect Dis*. 2021;8:20499361211034064.
 48. Laing AG, Lorenc A, Barrio DMD, I, Das A, Fish M, Monin L, Muñoz-Ruiz M, McKenzie DR, Hayday TS, Francos-Quijorna I, et al. A dynamic COVID-19 immune signature includes associations with poor prognosis. *Nat Med*. 2020;26:1623–35.
 49. Mukund K, Nayak P, Ashokkumar C, Rao S, Almeda J, Betancourt-García MM, Sindhri R, Subramaniam S. Immune Response in Severe and Non-Severe Coronavirus Disease 2019 (COVID-19) Infection: A Mechanistic Landscape. *Front Immunol*. 2021;12: 738073.
 50. Rydzynski Moderbacher C, Ramirez SI, Dan JM, Grifoni A, Hastie KM, Weiskopf D, Belanger S, Abbott RK, Kim C, Choi J, et al. Antigen-Specific Adaptive Immunity to SARS-CoV-2 in Acute COVID-19 and Associations with Age and Disease Severity. *Cell*. 2020;183:996–1012.e1019.
 51. Vitte J, Diallo AB, Boumazza A, Lopez A, Michel M, Allardet-Servent J, Mezouar S, Sereme Y, Busnel JM, Miloud T, et al. A Granulocytic Signature Identifies COVID-19 and Its Severity. *J Infect Dis*. 2020;222:1985–96.
 52. Wargodsky R, Dela Cruz P, LaFleur J, Yamane D, Kim JS, Benjenk I, Heinz E, Ironi OO, Farrar K, Toma I, et al. RNA Sequencing in COVID-19 patients identifies neutrophil activation biomarkers as a promising diagnostic platform for infections. *PLoS ONE*. 2022;17: e0261679.
 53. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res*. 2002;12:996–1006.
 54. Monaco G, Lee B, Xu W, Mustafah S, Hwang YY, Carré C, Burdini N, Visan L, Ceccarelli M, Poidinger M, et al. RNA-Seq Signatures Normalized by mRNA Abundance Allow Absolute Deconvolution of Human Immune Cell Types. *Cell Rep*. 2019;26:1627–1640.e1627.
 55. Uhlen M, Karlsson MJ, Zhong W, Tebani A, Pou C, Mikes J, Lakshminanth T, Forsström B, Edfors F, Odeberg J, et al. A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. *Science*. 2019;366(6472):eaax9198.
 56. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38:576–89.
 57. Nain Z, Rana HK, Liò P, Islam SMS, Summers MA, Moni MA. Pathogenic profiling of COVID-19 and SARS-like viruses. *Brief Bioinform*. 2021;22:1175–96.
 58. Pietzner M, Chua RL, Wheeler E, Jechow K, Willett JDS, Radbruch H, Trump S, Heidecker B, Zeberg H, Heppner FL, et al. ELF5 is a potential respiratory epithelial cell-specific risk gene for severe COVID-19. *Nat Commun*. 2022;13:4484.
 59. Haslbauer JD, Savic Prince S, Stalder AK, Matter MS, Zinner CP, Jahn K, Obermann E, Hanke J, Leuzinger K, Hirsch HH, Tzankov A. Differential Gene Expression of SARS-CoV-2 positive Bronchoalveolar Lavages: A Case Series. *Pathobiology*. 2024;91(2):158–68.
 60. Hossain MS, Tonmoy MIQ, Fariha A, Islam MS, Roy AS, Islam MN, Kar K, Alam MR, Rahaman MM. Prediction of the Effects of Variants and Differential Expression of Key Host Genes ACE2, TMPRSS2, and FURIN in SARS-CoV-2 Pathogenesis: An In Silico Approach. *Bioinform Biol Insights*. 2021;15:11779322211054684.
 61. Ugalde AP, Bretones G, Rodríguez D, Quesada V, Llorente F, Fernández-Delgado R, Jiménez-Clavero M, Vázquez J, Calvo E, Tamargo-Gómez I, et al. Autophagy-linked plasma and lysosomal membrane protein PLAC8 is a key host factor for SARS-CoV-2 entry into human cells. *Embo j*. 2022;41: e110727.
 62. Gómez-Carballa A, Rivero-Calle I, Pardo-Seco J, Gómez-Rial J, Rivero-Velasco C, Rodríguez-Núñez N, Barbeito-Castiñeiras G, Pérez-Freixo H, Cebery-López M, Barral-Arca R, et al. A multi-tissue study of immune gene expression profiling highlights the key role of the nasal epithelium in COVID-19 severity. *Environ Res*. 2022;210: 112890.
 63. Rombauts A, Bódalo Torruella M, Abelda-Alonso G, Perera-Bel J, Ferrer-Salvador A, Acedo-Terrades A, Gabarrós-Subirà M, Oriol I, Gudiol C, Nonell L, Carratalà J. Dynamics of gene expression profiling and identification of high-risk patients for severe COVID-19. *Biomedicines*. 2023;11(5):1348.
 64. A first update on mapping the human genetic architecture of COVID-19. *Nature*. 2022, 608:E1–e10.
 65. Speir ML, Bhaduri A, Markov NS, Moreno P, Nowakowski TJ, Papatheodorou I, Pollen AA, Raney BJ, Seninge L, Kent WJ, Haussler M. UCSC Cell Browser: visualize your single-cell data. *Bioinformatics*. 2021;37:4578–80.
 66. Ganji A, Farahani I, Khansarinejad B, Ghazavi A, Mosayebi G. Increased expression of CD8 marker on T-cells in COVID-19 patients. *Blood Cells Mol Dis*. 2020;83: 102437.
 67. Rha MS, Shin EC. Activation or exhaustion of CD8(+) T cells in patients with COVID-19. *Cell Mol Immunol*. 2021;18:2325–33.
 68. Wyler E, Mösbauer K, Franke V, Diag A, Gottula LT, Arsiè R, Klironomos F, Kopstein D, Hönzke K, Ayoub S, et al. Transcriptomic profiling of SARS-CoV-2 infected human cell lines identifies HSP90 as target for COVID-19 therapy. *iScience*. 2021;24(3):102151.
 69. Blanco-Melo D, Nilsson-Payant BE, Liu WC, Uhl S, Hoagland D, Møller R, Jordan TX, Oishi K, Panis M, Sachs D, et al. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell*. 2020;181:1036–1045.e1039.
 70. Galani IE, Rovina N, Lampropoulou V, Triantafyllia V, Manioudaki M, Pavlos E, Koukaki E, Fragkou PC, Panou V, Rapti V, et al. Untuned antiviral immunity in COVID-19 revealed by temporal type I/III interferon patterns and flu comparison. *Nat Immunol*. 2021;22:32–40.
 71. Hadjadj J, Yatim N, Barnabei L, Corneau A, Boussier J, Smith N, Péré H, Charbit B, Bonnet V, Chenevier-Gobeaux C, et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science*. 2020;369:718–24.
 72. Jiang Y, Zhao T, Zhou X, Xiang Y, Gutierrez-Castrellon P, Ma X. Inflammatory pathways in COVID-19: Mechanism and therapeutic interventions. *MedComm*. 2020;2022(3): e154.
 73. Consortia CZIS-CC-, Ballestar E, Farber DL, Glover S, Horwitz B, Meyer K, Nikolić M, Ordovas-Montanes J, Sims P, Shalek A, et al. Single cell profiling of COVID-19 patients: an international data resource from multiple tissues. *medRxiv* 2020:2020.2011.2020.20227355.
 74. Munnur D, Teo Q, Eggermont D, Lee HHY, Thery F, Ho J, van Leur SW, Ng WWS, Siu LYL, Beling A, et al. Altered ISGylation drives aberrant macrophage-dependent immune responses during SARS-CoV-2 infection. *Nat Immunol*. 2021;22:1416–27.
 75. Szabo A, Rajnavolgyi E. Finding a fairy in the forest: ELF4, a novel and critical element of type I interferon responses. *Cell Mol Immunol*. 2014;11:218–20.
 76. You F, Wang P, Yang L, Yang G, Zhao YO, Qian F, Walker W, Sutton R, Montgomery R, Lin R, et al. ELF4 is critical for induction of type I interferon and the host antiviral response. *Nat Immunol*. 2013;14:1237–46.
 77. Khezri MR, Vazandeh R, Ghasemnejad-Berenji M. The probable role and therapeutic potential of the PI3K/AKT signaling pathway in SARS-CoV-2 induced coagulopathy. *Cell Mol Biol Lett*. 2022;27:6.
 78. Kindrachuk J, Ork B, Hart BJ, Mazur S, Holbrook MR, Frieman MB, Traynor D, Johnson RF, Dyall J, Kuhn JH, et al. Antiviral potential of ERK/MAPK and PI3K/AKT/mTOR signaling modulation for Middle East respiratory syndrome coronavirus infection as identified by temporal kinome analysis. *Antimicrob Agents Chemother*. 2015;59:1088–99.

79. Appelberg S, Gupta S, Svensson Akusjärvi S, Ambikan AT, Mikaeloff F, Saccon E, Végvári Á, Benfeitas R, Sperk M, Ståhlberg M, et al. Dysregulation in Akt/mTOR/HIF-1 signaling identified by proteo-transcriptomics of SARS-CoV-2 infected cells. *Emerg Microbes Infect.* 2020;9:1748–60.
80. Garcia G Jr, Sharma A, Ramaiah A, Sen C, Purkayastha A, Kohn DB, Parcels MS, Beck S, Kim H, Bakowski MA, et al. Antiviral drug screen identifies DNA-damage response inhibitor as potent blocker of SARS-CoV-2 replication. *Cell Rep.* 2021;35: 108940.
81. Divangahi M, Aaby P, Khader SA, Barreiro LB, Bekkering S, Chavakis T, van Crevel R, Curtis N, DiNardo AR, Dominguez-Andres J, et al. Trained immunity, tolerance, priming and differentiation: distinct immunological processes. *Nat Immunol.* 2021;22:2–6.
82. Agrawal B. Heterologous Immunity: Role in Natural and Vaccine-Induced Resistance to Infections. *Front Immunol.* 2019;10:2631.
83. Habibi MS, Thwaites RS, Chang M, Jozwik A, Paras A, Kirsebom F, Varese A, Owen A, Cuthbertson L, James P, et al. Neutrophilic inflammation in the respiratory mucosa predisposes to RSV infection. *Science.* 2020;370(6513):eaba9301.
84. Kuang M, Zhao Y, Yu H, Li S, Liu T, Chen L, Chen J, Luo Y, Guo X, Wei X, et al. XAF1 promotes anti-RNA virus immune responses by regulating chromatin accessibility. *Sci Adv.* 2023;9:eadg5211.
85. Gallo O, Locatello LG, Mazzoni A, Novelli L, Annunziato F. The central role of the nasal microenvironment in the transmission, modulation, and clinical progression of SARS-CoV-2 infection. *Mucosal Immunol.* 2021;14:305–16.
86. Broad Institute: Picard toolkit. Available from <http://broadinstitute.github.io/picard/>. Accessed 2023 Apr 01. 2019.
87. Amemiya HM, Kundaje A, Boyle AP. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci Rep.* 2019;9:9354.
88. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 2012;13:R87.
89. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.
90. Burger L, Gaidatzis D, Schübeler D, Stadler MB. Identification of active regulatory regions from DNA methylation data. *Nucleic Acids Res.* 2013;41: e155.
91. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004;5:R80.
92. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods.* 2015;12:115–21.
93. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26:841–2.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.