

RESEARCH

Open Access



# DNAm landscape up to 4 months post SARS-CoV-2 infection: insights from four population-based cohorts

Pamela R. Matías-García<sup>1,2,3\*</sup>, Liye Lai<sup>1,2,4</sup>, Thomas Delerue<sup>1,2</sup>, Jochen Ohnmacht<sup>5</sup>, Klaus J. Stark<sup>6</sup>, Robert Warmerdam<sup>7,8</sup>, Alexey Kolodkin<sup>5</sup>, Julia Six-Merker<sup>2</sup>, Nina Mangold<sup>2,9</sup>, Kathrin Günther<sup>10</sup>, Marc P. O'Sullivan<sup>5</sup>, Armin Rauschenberger<sup>5,11</sup>, Barbara Bohn<sup>12</sup>, Klaus Berger<sup>13</sup>, Julia Fricke<sup>14,15</sup>, Peter Ahnert<sup>16</sup>, Lude Franke<sup>7,8</sup>, Rejko Krüger<sup>5,11,17</sup>, Olaf Gefeller<sup>18</sup>, Klaus Überla<sup>19</sup>, Iris M. Heid<sup>6</sup>, Ralf Wagner<sup>20,21</sup>, Lifelines Corona Research Initiative, TiKoCo Working Group, CON-VINCE Consortium, NAKO Investigators, on behalf of the ORCHESTRA working group, Melanie Waldenberger<sup>1,2,3\*</sup> and Annette Peters<sup>2,3,4</sup>

## Abstract

**Background** Evidence for persistent epigenetic changes in individuals who had a mild SARS-CoV-2 infection is limited, as most DNA methylation (DNAm) studies to date have focused on either the acute phase of infection or on the months following infection in severe cases requiring hospitalization.

**Methods and results** Using the Infinium Human MethylationEPIC BeadChip, we investigated blood DNA methylation (DNAm) up to four months after SARS-CoV-2 infection in cases and controls from four population-based cohorts (NAKO, Lifelines, CON-VINCE, and TiKoCo;  $n = 675$ ) within the framework of the ORCHESTRA Consortium. We observed DNAm changes at 16 differentially methylated positions (DMPs) and 21 differentially methylated regions (DMRs), with 89% of these DMPs/DMRs hypomethylated in cases compared to age- and sex-matched controls. Genes mapped to these CpGs were annotated with Gene Ontology terms and pathways related to immune responses to viral infection. eQTM analyses in whole blood from an independent cohort (KORA FF4 study) produced 49 significant CpG–transcript pairs, including *IFI44L* and *GNA12*. Despite inter-individual variability and cohort heterogeneity, our findings regarding four DMPs (*IFI44L*, *MX1*, *DDX60*, and *RABGAP1L*) and two DMRs (*PARP9* and *GNA12*) replicate changes described both in the acute phase of infection and at long-term follow-up. Differential methylation at other novel loci may reflect the systemic nature of post-infection epigenetic changes.

**Conclusion** Our findings suggest moderate but persistent epigenetic changes up to four months after SARS-CoV-2 infection in mild cases from population-based cohorts. These changes partially mirror those reported during the acute phase of both mild and severe COVID-19 and overlap with pathways dysregulated in autoimmune, metabolic

\*Correspondence:

Pamela R. Matías-García  
pamela.matiasgarcia@helmholtz-munich.de  
Melanie Waldenberger  
melanie.waldenberger@helmholtz-munich.de

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



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

and neurological disease. Future research should examine epigenetic changes associated with persisting symptoms in long COVID, investigate downstream effects of DNAm changes on other -omics, and consider longer follow-up periods to further elucidate the molecular mechanisms underlying SARS-CoV-2 induced epigenetic changes.

**Keywords** COVID-19, SARS-CoV-2, EWAS, DNA methylation, ORCHESTRA cohort, Epigenetic changes, Population-based

## Introduction

Changes in DNA methylation (DNAm) levels have been widely investigated in epigenome-wide association studies (EWAS) during the acute phase of SARS-CoV-2 infection, particularly in critically ill patients requiring hospitalization. Studies have described DNAm signatures correlating with disease progression or distinguishing between cases and controls [1–7]. These studies have demonstrated significant shifts in white blood cell composition and associated epigenetic and transcriptomic changes that drive COVID-19 disease progression [1, 2, 8]. Some epigenome-wide studies have focused on identifying DNAm changes predictive of COVID-19 outcomes. Castro de Moura et al. identified a 44-CpG epigenetic signature predictive of disease severity in mild and severe COVID-19 patients without comorbidities [3]. Balnis et al. identified a COVID-19-specific signature with 77 differentially methylated regions (DMR) predictive of disease severity in hospitalized patients [4]. Konigsberg and colleagues reported 13,033 CpGs associated with case status and DNAm signatures linked to disease progression and hospitalization [5]. Similarly, Calzari et al. identified DNAm markers predictive of severe infection in comparison to mild disease in individuals with comorbidities [7]. Additionally, several studies have sought to distinguish COVID-19 specific DNAm signatures by comparing them with DNAm profiles from individuals with other respiratory and infectious diseases [4, 6]. Despite differences in study designs – such as sample sizes, patient populations, and analytical methods –, common findings across these studies indicate that widespread hypomethylation in severe cases [1, 2], as well as changes in genes and pathways related to interferon response to viral infection [1–3, 5, 6].

In contrast, considerably fewer studies have investigated epigenetic changes associated with milder forms of COVID-19 [8, 9] and/or with longer follow-up periods post-infection [9–13]. A longitudinal study examining asymptomatic and mildly symptomatic infection in 133 young adults found that differentially methylated sites persisted for several months [8]. The study with the longest follow-up assessed epigenetic changes in 15 patients one year after hospital discharge identifying 71 persistently dysregulated differentially methylated regions (DMRs) out of 1,505 DMRs associated with acute illness at baseline [4, 10]. Two additional EWAS have identified changes three to six months post-infection [11, 13]:

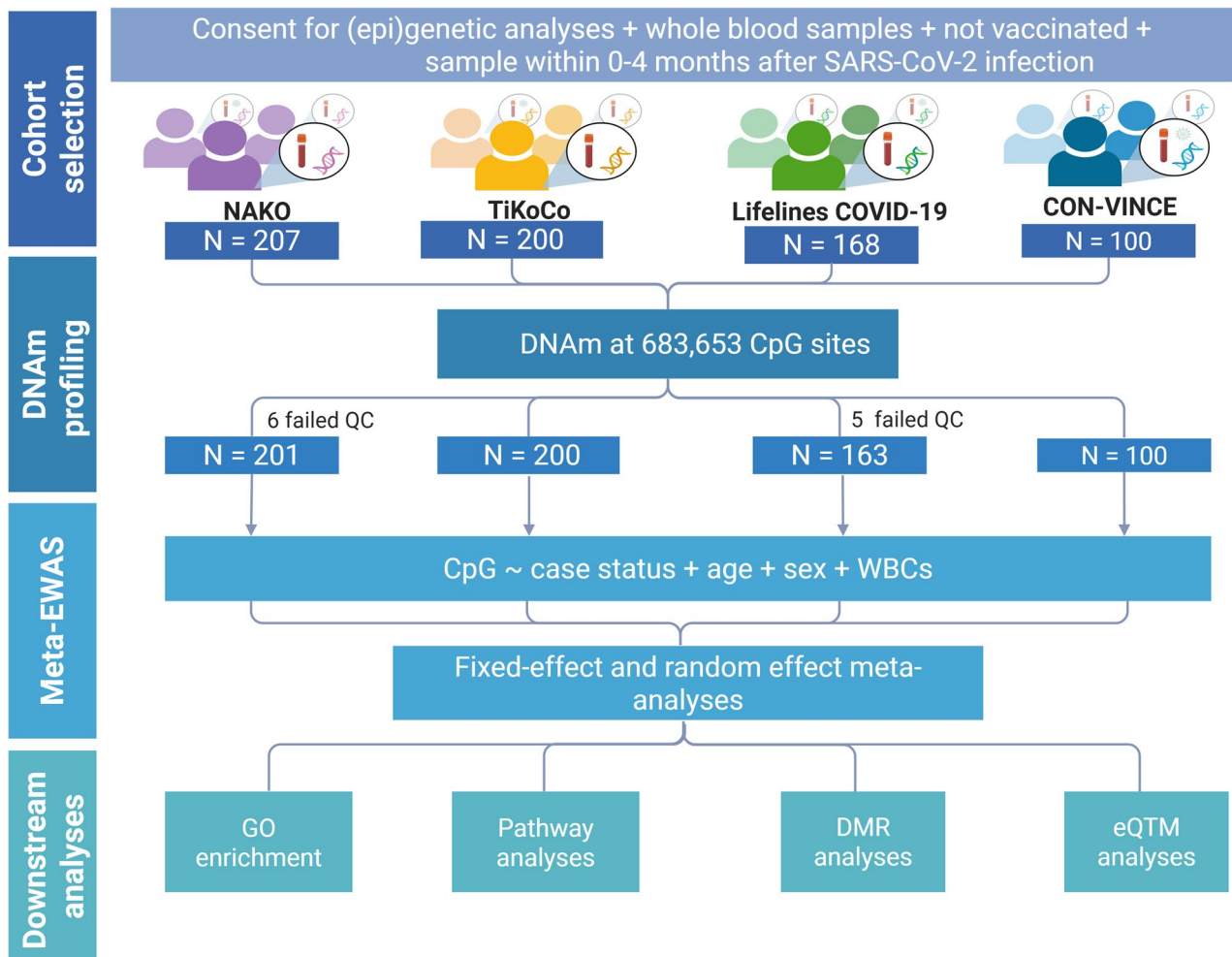
one study, analysing blood samples collected 8–12 weeks after SARS-CoV-2 infection in 109 patients and 73 controls, identified three hypomethylated CpGs in cases [11]. Another, examining 96 post-COVID individuals six months after infection alongside 191 pre-pandemic controls, found 42 differentially methylated CpGs [13].

These studies provide insights into the long-term epigenetic changes associated with severe COVID-19 in comparison to mild disease or pre-pandemic controls. However, most lack the statistical power to detect smaller effects linked to mild disease. Consequently, evidence on the epigenetic signature of mild COVID-19 remains scarce, particularly when compared to exposed controls. Furthermore, it remains unclear how these changes correlate with those observed during the acute phase of infection and whether they persist for months post-infection. This study aims to thus identify differentially methylated CpGs in individuals from population-based cohorts in the largest EWAS to date in order to characterize the epigenetic landscape up to four months following mild SARS-CoV-2 infection.

## Methods

### Cohort descriptions and participant selection

Population-based cohorts with available blood and/or genomic DNA samples for DNAm profiling were identified within the framework of the European Orchestra consortium [14]: German National Cohort (NAKO) [15], COVID-19 National survey for assessing Viral spread by Non-affected Carriers (CON-VINCE) [16, 17], The Lifelines Corona Research Initiative [18, 19], and Tirschenreuth Kohorte COVID-19 (TiKoCo) [20, 21] (Fig. 1). Participants from these cohorts were selected if they met the following inclusion criteria: (i) attended an examination where demographic/clinical data was collected and whole blood sampling was done, (ii) provided informed consent for (epi)genetic analyses, (iii) had (self-reported) data on COVID-19 testing within the last 4 months prior to the examination/blood draw, (iv) had not received a COVID-19 vaccine, and (v) were 18 years or older. Cases were identified as individuals with positive serology results and/or reporting positive COVID-19 testing (PCR, rapid test, self-test). Controls were selected from the pool of individuals with negative self-reported COVID-19 status or negative serology results; controls were matched to cases based on sex and age (see Supplemental Note 1 for further details).



**Fig. 1** Overview of study design. Participants from four population-based cohorts were selected according to predefined criteria (light blue box). DNA was extracted from blood samples, and DNA methylation (DNAm) at ~850,000 CpG sites was measured using the Illumina EPIC array. DNAm data underwent preprocessing, quality control and harmonization across cohorts; 683,653 CpG sites were available for analysis after these steps. Epigenome-wide analyses were performed with DNAm as the outcome and case status as the exposure, adjusting for age, sex, and blood cell proportions. Study-level results were combined in fixed- and random-effect meta-analyses. Significant CpGs were carried forward to downstream analyses, including GO/pathway enrichment, differentially methylated region (DMR) analyses and expression quantitative trait methylation (eQTM) analyses. Created in BioRender. Waldenberger, M. (2026) <https://BioRender.com/lzuyisr>

### DNA methylation profiling and data processing

Whole blood or DNA samples from the selected cohorts were shipped to Helmholtz Munich for processing and DNA-methylation (DNAm) profiling. Samples meeting basic quality measures were used for DNAm analyses. For each sample from all included cohorts, 750 ng of genomic DNA was used for bisulfite conversion with the EZ-96 DNA Methylation Kit (Zymo Research, Orange, CA, USA). Methylation profiling was done following using the iScan platform (San Diego, CA, USA) and the Infinium MethylationEPIC BeadChip following standard protocols defined by Illumina. Initial quality control of the assay and generation of export data was done in GenomeStudio (v2011.1, RRID: SCR\_010973) with Methylation Module version 1.9.0.

Quality control, preprocessing of the data and statistical analyses were performed in R v4.1.3 (R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>).

Data was processed using the package minfi v1.40.0 [22] following the CPACOR pipeline [23], as described in [24]. In brief, raw intensities were read into R and background corrected. The cut-off for the detection rate in each study was determined individually, taking into account results from two quality control steps: (i) comparison of predicted sex, based on DNA methylation patterns, with reported sex to identify potential sample mislabelling, and (ii) assessment of median intensity values. The goal is to optimize the sample size while maintaining data quality, so that the cohorts' range for this threshold was

80%–95% (see Suppl. Note 2 for additional details). Prior to normalization, samples failing sex prediction or falling below median intensity were removed. Likewise, probes with known issues were removed (cross-reactive, with minor allele frequency minor allele > 5% at the CG or the single base extension positions, and > 5% missing values); additionally, probes from EPICv1 not represented in the EPICv2 and probes in sex chromosomes were removed. Quantile normalization (QN, R package limma v3.50.3, RRID: SCR\_010943 [25]) was performed as described in [24]; briefly, this was done separately on the signal intensities divided into the 6 probe types [23]. QN was performed for all samples together for the autosomes, and the transformed intensities were then used to generate methylation beta values: a measure from 0 to 1 indicating the percentage of cells methylated at a given locus. A total of 683,653 CpGs passed quality control measures across all profiled cohorts.

#### Epigenome-wide association analysis (EWAS)

The association between DNA-methylation (DNAm) at 683,653 CpGs and SARS-CoV-2 infection (case/control) was examined in an epigenome-wide association study (EWAS) of differentially methylated positions (DMPs) using linear regression models in each cohort separately. As methylation levels in blood can be strongly influenced by leukocyte composition, we performed white blood cell deconvolution from bulk blood DNAm samples, using the R package FlowSorted.BloodExtended.EPIC and the function estimateCellCounts2() on the raw intensities and default parameters. This produced estimates for 12 cell types: neutrophils (Neu), eosinophils (Eos), basophils (Bas), monocytes (Mono), B naïve cells (Bnv), B memory cells (Bmem), T-helper CD4+ naïve cells (CD4nv), T-helper CD4+ memory cells (CD4mem), T regulatory cells (Treg), T-cytotoxic CD8+ naïve cells (CD8nv), T-cytotoxic memory CD8+ cells (CD8mem), and natural killer cells (NK) [26]. DNAm levels were defined as the dependent variable and case status as the predictor variable, where one model was run per CpG site. Variables included in the main model to address potential confounding were age, sex and 11 variables from the deconvoluted white blood cell proportions; analyses in NAKO additionally included a technical covariate (array version). A second model additionally adjusted for smoking (current smoker, former smoker, never smoker) and prevalent chronic disease (cardiovascular disease, lung disease, and metabolic disease; see Suppl. Note 1 for details on cohort-specific definitions of this variable). Analyses were done using the function lm() from the package stats v4.2.2.

Results were quality controlled using the R package QCEWAS v1.2–3; during this QC, results files are checked for data integrity and validity (e.g. no negative

standard errors or p-values), cohort-specific outlier detection and removal of sex chromosomes [27]. Quantile–Quantile (Q–Q) plots and additional diagnostic plots (histogram for effect sizes and standard errors, volcano plot with distribution of effect sizes in relation to p-values and Manhattan plots) were generated to assess over/under-significance of the results and comparability of effect sizes. Conventional genomic inflation factors ( $\lambda$ ) were calculated to evaluate inflation in each EWAS.

#### Meta-analyses

Cohort-level results were meta-analyzed for 667,707 CpGs with available EWAS results across all cohorts using both fixed-effects (FE) and random-effects (RE) models in METAL (RRID: SCR\_002013) [28]. The FE model combined study-level estimates under the assumption of a common underlying effect, weighting by the inverse of their variance. The RE model incorporated both within- and between-study variance ( $\tau^2$ ), with the latter estimated using the DerSimonian–Laird method. For each CpG, measures of heterogeneity included the chi-squared statistic for heterogeneity (HetChiSq), its p-value (HetPVal), the  $I^2$  statistic (0–100%), and  $\tau^2$ , a measure of between-study variance.

Associations with consistent direction of effect across cohorts and meta-analysis p-values below the Bonferroni-corrected threshold ( $\alpha = 0.05$ ;  $p < 7.48 \times 10^{-8}$ ) were considered statistically significant, while associations with  $p < 1 \times 10^{-5}$  were considered suggestive. For CpG–case status associations showing substantial heterogeneity ( $I^2 > 70\%$ ), leave-one-out meta-analyses were conducted as sensitivity analyses to identify influential studies using the meta and dmetar R packages [29–31]. An additional sensitivity analysis was performed in a subset of participants with serology-confirmed case status (TiKoCo and CONVINCE;  $N = 292$ ).

#### Differentially methylated region (DMR) analyses

Differentially methylated regions (DMRs) represent genomic regions with consistently altered DNA methylation across multiple adjacent CpG sites. To complement the single-site differentially methylated position (DMP) analysis, DMR analyses were done with the meta-analyses results and at the cohort-level using the comb-p function (Enmix v1.38.01; dist.cutoff = 1000, seed = 1e-2 and bin.size = 310) [32]. Sensitivity analyses varying one parameter at a time using more stringent or lenient values were conducted (SA1 with seed = 1-e03, SA2 with d = 500, SA3 with b = 500, SA4 b = 500 and s = 1e-03). Coordinates from DMRs were converted to GRanges objects and annotated using the R packages GenomicRanges (v1.54.01) [33] and the EPIC annotation dataset from chAMP (2.32.0) [34]. Meta-DMRs and cohort-level

DMRs were compared by coordinate overlap and shared CpG probes to assess replication.

### Enrichment analyses and annotation

Gene ontology (GO) and KEGG pathway enrichment analyses (RRID: SCR\_012773) were performed using the goMeth function from the missMethyl R package (v.1.32.1) [35]. Additional potentially relevant pathways were queried in WikiPathways (RRID: SCR\_002134) [36]. To explore whether the identified genes showed cell type-specific expression, single-cell RNA-seq data from 18 sorted immune cell types was queried in the Human Protein Atlas [37, 38]. Annotation to regulatory regions and chromatin states was done using the PBMC Roadmap reference epigenome (E062) [39]. Association with other phenotypes was queried in the EWAS Catalog [40]. An overrepresentation analysis (ORA) was done using WebGestalt (RRID: SCR\_006786) to explore enrichment of genes in Pathway Figure OCR (PFOCR) [41, 42]. The influence of genetic variants in the identified CpGs, namely the overlap with methylation quantitative trait loci (meQTL), was queried in the BIOS QTL database and GoDMC [43, 44].

### Expression quantitative trait methylation (eQTM) analyses

To investigate the functional relevance of these methylation changes, expression quantitative trait methylation (eQTM) analyses were done to identify associations between the complete set of identified CpG sites (DMP and DMR analyses) and gene expression using DNAm and RNA sequencing (RNA-seq) data from the KORA FF4 cohort (see Suppl. Note 3). After quality control, the data were available for 1,543 individuals. Gene expression probes within a 500 kb window surrounding significant CpGs were analyzed using the MatrixEQTL package (version 2.3) [45]. Linear models were adjusted for age, sex, measured white blood cell proportions (neutrophils, monocytes, basophils, and eosinophils), and technical variation, with multiple testing controlled by Bonferroni correction.

## Results

### Cohort description

Table 1 shows population characteristics for the 675 participants from 4 cohorts included in this study. The cohorts largely consisted of similar age distributions and sex ratios. The average age ranged from 41 to 57 years, with a balanced representation of males and females across cases and controls. Smoking prevalence between cases and controls was similar. No significant differences were found in BMI, smoking groups or prevalence of chronic disease (cardiovascular, metabolic and lung disease) between both groups across all cohorts. Twelve blood cell types were deconvoluted from DNAm data and

estimated for all participants from the studied cohorts; the only statistically significant difference observed was higher levels of Basophils in cases compared to controls in NAKO ( $p = 0.009$ ).

### Epigenome-wide association analysis (EWAS)

Epigenome-wide association analyses were conducted in all studies individually (Supplemental Tables 1–4, Supplemental Figs. 1–4). EWAS results from all population-based cohorts showed comparable effect sizes and reasonable precision-sample size ratios (Suppl. Figure 5); genomic inflation values ranged from  $\lambda = 0.931$  to  $\lambda = 1.007$ , so no genomic control correction was done. No deviations from expected precision in the obtained estimates was observed (Suppl. Figures 1–5). Two CpG sites, cg13452062 and cg03607951, both annotated to *IFI44L*, reached Bonferroni-corrected significance in the Lifelines study. No significant signals were observed in the CONVINCE, NAKO, or TiKoCo studies.

### Meta-EWAS results

Differences in DNAm levels between individuals who had a SARS-CoV-2 infection in the prior 4 months and controls were identified in 16 CpGs annotated to 12 different genes (10 protein-coding and 2 encoding for long non-coding RNA transcripts) at suggestive threshold ( $p < 1 \times 10^{-5}$ ); 3 of these CpGs were significant after Bonferroni correction: cg03607951, cg24678928 and cg19397320 (Fig. 2). Notably, 75% of these identified CpGs were hypomethylated in individuals who had a SARS-CoV-2 infection (Table 2), and cohort-specific estimates were consistent in direction and comparable in magnitude (Fig. 3). Most of these CpGs are located in promoter regions (7 CpGs [43.7%] in TSS1500, 1 [6%] in 1st exon, and 2 [12.5%] in 5' UTR), followed by 6 in non-promoter regions (5 CpGs [31.2%] in IGR) and 1 in the gene body (1 CpG [6.2%]) (Table 2). EWAS summary statistics are provided in Suppl. Table 5. DNAm at cg03607951 (annotated to *IFI44L*) shows the strongest association with case status (effect =  $-0.024$  [ $-0.032$ ,  $-0.016$ ],  $p = 1.50E-09$ ); the direction of this effect was consistent across studies (Fig. 3). Although heterogeneity was high given the variability in the effect sizes observed across cohorts ( $I^2 = 78.9\%$ ), low variability between studies was observed (Suppl. Table 6). Similar results were observed for cg24678928 (*DDX60*; effect =  $-0.017$  [ $-0.023$ ,  $-0.011$ ],  $p = 6.37E-09$ ,  $I^2 = 76.3\%$ ). In the case of cg03607951, lowest heterogeneity and smallest pooled effect sizes were observed when Lifelines was omitted from analysis ( $I^2 = 2\%$  and effect =  $-0.02$  [ $-0.03$ ,  $-0.01$ ]; Suppl. Figure 6). The omission of NAKO had a similar effect in reducing heterogeneity in the association for cg24678928 ( $I^2 = 47\%$ ; Suppl. Figure 7). In the case of cg19397320 (*BTBD3*; effect =  $-0.005$  [ $-0.006$ ,  $-0.003$ ],



**Table 1** (continued)

Cohort	NAKO			TiKoCo			Lifelines COVID-19			CONVINCE		
	Case	Control	p	Case	Control	p	Case	Control	p	Case	Control	p
Positive	NA	NA		0 (0.0)	100 (100.0)		NA	NA		40 (80.0)	0 (0.0)	
Borderline	NA	NA		0 (0.0)	0 (0.0)	<0.001	NA	NA		2 (4.0)	0 (0.0)	
IgA_status (%)												
Negative	NA	NA		NA	NA		NA	NA		18 (36.0)	47 (94.0)	<0.001
Positive	NA	NA		NA	NA		NA	NA		29 (58.0)	3 (6.0)	
Borderline	NA	NA		NA	NA		NA	NA		3 (6.0)	0 (0.0)	

Summary of demographics, blood cell proportions, and serology-based status of participants from the four cohorts. Continuous variables are presented as mean (SD) with p-values from t-tests; categorical variables are shown as counts (%) with p-values from chi-squared tests

$p = 6.21E-08$ ), no heterogeneity was observed and thus estimates from the random-effects model were the same. All the identified associations were slightly attenuated by adjusting for prevalent chronic disease and smoking, but the estimated effect sizes in both models were highly correlated (Spearman's rank  $\text{corr} = 0.99$ ,  $p = 1.037E-05$ ) and remained significant at  $p < 1e-03$  (Suppl. Table 6). Effect sizes estimated at the epigenome-wide level in both models also showed a high correlation (Pearson's  $\text{corr} = 0.948$ ,  $p < 2.2E-16$ ; full summary statistics from model further adjusting for smoking and prevalent chronic disease in Suppl. Table 7). Similarly, the effects estimated in the sensitivity analysis including only cohorts with a serology-based case definition (TiKoCo and CONVINCE,  $N = 292$ ) robustly replicated those of the main analyses with lower heterogeneity (Suppl. Table 8).

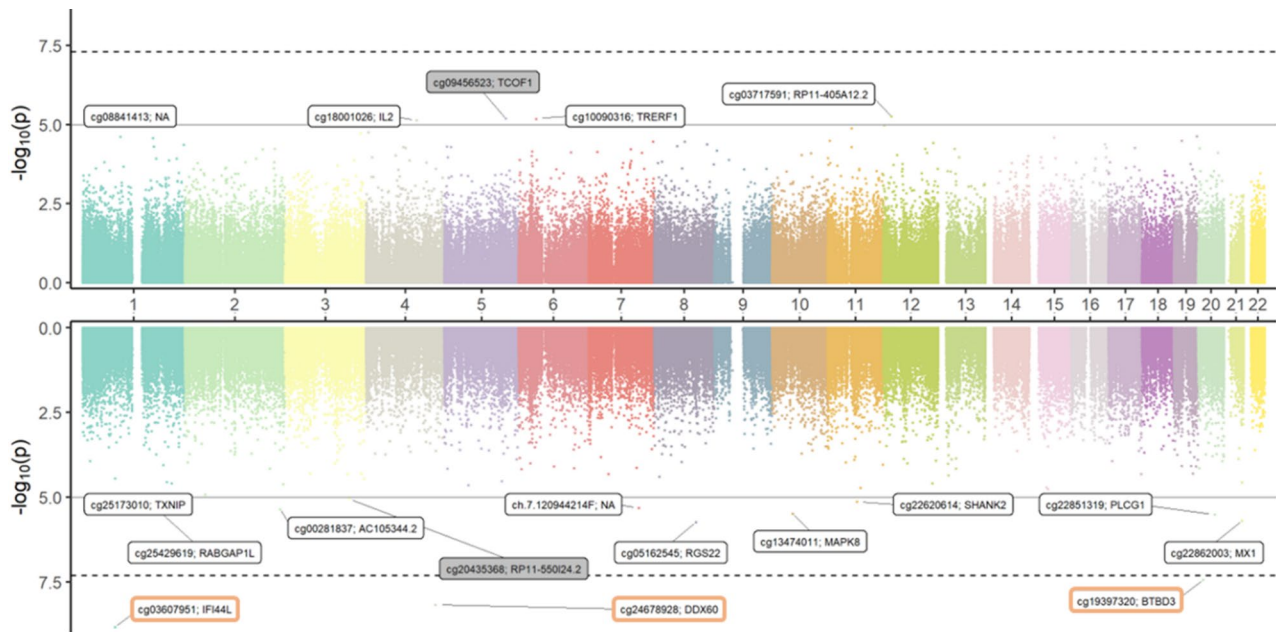
**DMR results**

Table 3 shows 21 DMRs from the meta-EWAS (Sidak  $p < 0.05$ ), composed by 101 CpG sites and overlapping 10 genes and 2 pseudogenes; eight are in proximity to promoter regions, followed by nine in non-promoter regions. All DMRs were hypomethylated, in full concordance with the DMP effects observed across cohorts (Suppl. Table 9). The top two CpGs in the DMP analyses were also identified in the DMR analyses but lacked enough adjacent CpGs to meet DMR criteria. Sensitivity analyses using different parameters to call DMRs produced 8 to 30 DMRs (Suppl. Table 10); eight DMRs were consistently identified across all analyses (Suppl. Table 11). Of note, two of the sensitivity analyses allowing for larger DMRs identified an additional *IFI44L*-annotated CpG (Suppl. Table 12), thus producing a 3-CpG-DMR overlapping DMP results.

In the DMR analyses conducted across four cohorts, a total of 51 distinct DMRs were identified (20 in NAKO, 13 in TiKoCO and Lifelines each, 5 in CONVINCE; Suppl. Table 13). Five of the meta-DMRs were replicated in one cohort each: DMR3 and DMR21 in TiKoCo, DMR5 and DMR18-*THRB* in NAKO, and DMR10-*PARP9* in Lifelines (Suppl. Table 14). There was no overlap of Sidak significant DMRs across the four cohorts (Suppl. Table 15).

**Enrichment analyses and annotation**

The set of 117 CpG sites identified in either DMP- or DMR-analyses were annotated to 22 protein-coding genes (Suppl. Table 13). Even though no statistically significant enrichment of gene-sets in the GO or KEGG pathway databases were found, these analyses are helpful in contextualizing our findings in relation to biological functions. Top GO terms are related to response to viral infections (Suppl. Figure 8 and Suppl Table 16). Forty-seven KEGG pathways were represented in this list of



**Fig. 2** Associations between genome-wide CpG methylation levels and case status from meta-EWAS Manhattan plot of meta-EWAS results for CpG methylation and case status. The X-axis shows genomic position by chromosome; the Y-axis shows  $-\log_{10}$  p-values from the IWW FE meta-analysis. The upper half of the plot represents CpG sites with positive effect estimates, while the lower half shows CpG sites with negative effect estimates. Each point represents a CpG, colored by chromosome. 16 probes significant at the suggestive threshold ( $p < 1 \times 10^{-5}$ ) are labeled, with Bonferroni-significant CpGs highlighted in orange. Grey-labeled CpGs (cg09456523, cg20435368) showed inconsistent effect directions across cohorts and were not considered robust

CpGs (Suppl. Table 17), which included pathways related to immune and inflammatory diseases. Likewise, *IFI44L*, *MX1* and *MAPK8* were assigned to pathways involved in downstream molecular events triggered by SARS-CoV-2 infection (WikiPathways WP5115). Five genes showed immune-cell specific expression (Suppl. Table 18); for example, *THRB* had higher expression in basophils and naïve B-cells, *SLC35F3* expression was specific to dendritic cells and *PF4* specific to neutrophils. Additional analyses using WebGestalt showed a high overlap between the gene set identified in these analyses and those identified in an epigenetic signature of Systemic Lupus Erythematosus and HIV immune control (Suppl. Figure 9; Suppl. Table 19). meQTL were found in GoDMC for a few of the identified CpGs (e.g. cg25173010 in *TXNIP* and cg22862003 in *MX1*). The set of 16 identified CpGs were annotated to six chromatin states (Suppl. Table 20), of which TssBiv was suggestive of enrichment (OR = 7.76,  $p = 0.04$ ; Suppl. Figure 10).

#### Association between DNA methylation and gene expression

Several of the CpGs hypomethylated in cases were associated with expression of nearby genes. We identified 49 CpG–transcript pairs in whole blood in an eQTM analysis using RNA-seq and methylation data from KORA FF4 (Suppl. Table 21). Most CpG–gene transcript pairs had a negative association (69.39%, average effect size

of  $-1.92$ ), where hypomethylation was associated with increased gene expression; the remaining pairs showed increasing methylation levels with higher gene expression (average effect size of 0.47). Cg03607951 (*IFI44L*), the top CpG from the DMP analysis, showed strong negative associations with its gene transcript (Suppl. Figure 11 A). Notably, all seven CpGs in DMR4 (*GNAI2*) were negatively associated with its gene (Suppl. Figure 11B); likewise, all 3 CpGs in DMR10 (*PARP9*) were associated with multiple transcripts (*PARP9*, *PARP14* and *DXTL30*) and positively associated with *KPNA1* (Suppl. Table 21; Suppl. Figure 11 C).

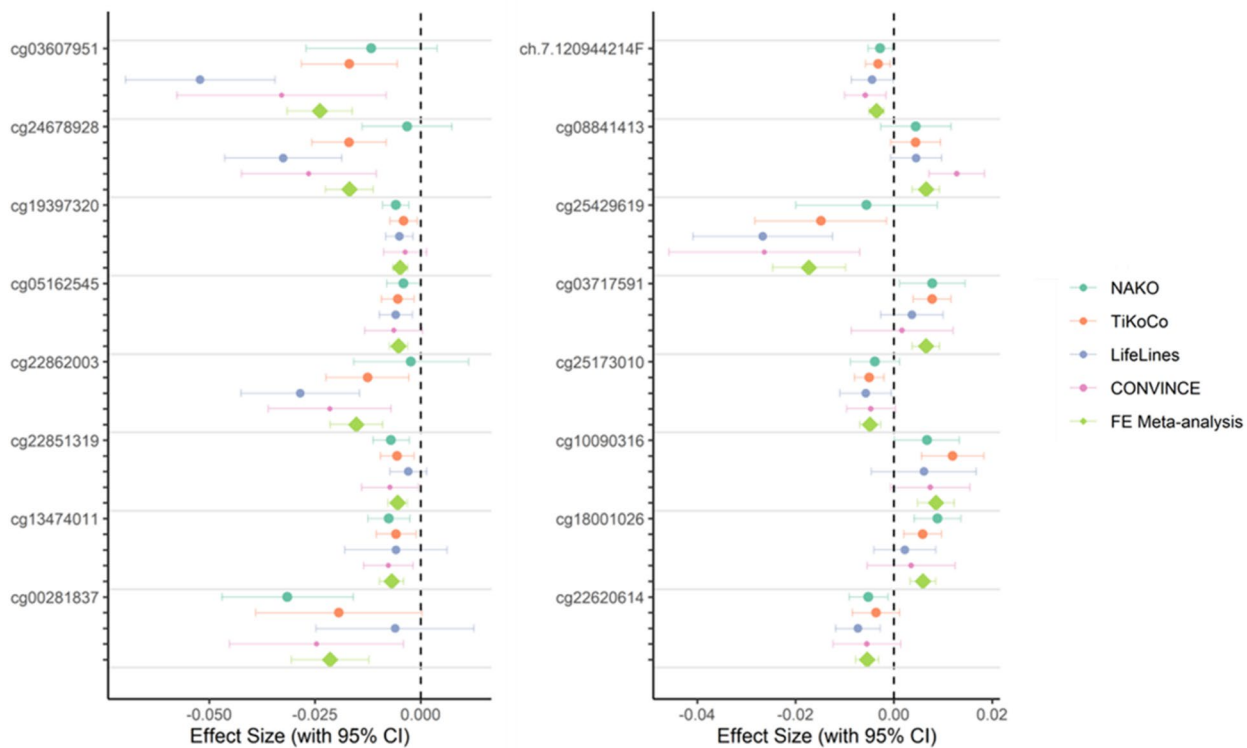
#### Discussion

We conducted the largest EWAS to date assessing DNAm changes up to four months after SARS-CoV-2 infection in 675 individuals from four population-based European cohorts within the ORCHESTRA Consortium [14]. DNAm changes were identified in 16 DMP and 21 DMRs, most of which were located in promoter regions or body of protein-coding genes; most DMPs and all DMRs were hypomethylated in cases vs. controls. Three DMPs reached genome-wide significance after Bonferroni-correction ( $p < 7.48E-08$ ): the strongest association was observed for cg03607951 (*IFI44L*), followed by cg24678928 (*DDX60*) and cg19397320 (*BTBD3*). The top two DMPs were also identified in DMR sensitivity analyses, supporting the robustness and complementarity

**Table 2** Differentially methylated positions (DMPs) associated with case status in the meta-EWAS

CpG	Genomic location	Gene	Chr	Dir	FE Meta-analysis			RE Meta-analysis				
					Effect	SE	pval	I <sup>2</sup>	HetPVal	Effect ARE	SE ARE	Pvalue ARE
cg03607951 <sup>#</sup>	TSS1500	<i>IFI44L</i>	1	---	-0.024	0.004	1.50E-09	78.9	0.003	-0.027	0.009	0.0027
cg24678928 <sup>#</sup>	TSS1500	<i>DDX60</i>	4	---	-0.017	0.003	6.73E-09	76.3	0.005	-0.019	0.006	0.0024
cg19397320 <sup>*</sup>	1stExon	<i>BTBD3</i>	20	---	-0.005	0.001	3.71E-08	0	0.83	-0.005	0.001	3.71E-08
cg05162545 <sup>*</sup>	TSS1500	<i>RG522</i>	8	---	-0.005	0.001	1.84E-06	0	0.907	-0.005	0.001	1.84E-06
cg22862003 <sup>#</sup>	TSS1500	<i>MX1</i>	21	---	-0.015	0.003	2.00E-06	62.3	0.047	-0.016	0.005	0.003
cg22851319 <sup>*</sup>	TSS1500	<i>PLCG1</i>	20	---	-0.005	0.001	3.12E-06	0	0.557	-0.005	0.001	3.12E-06
cg13474011 <sup>*</sup>	IGR	<i>MAPK8</i>	10	---	-0.007	0.001	3.24E-06	0	0.951	-0.007	0.001	3.24E-06
cg00281837	IGR	<i>AC105344.2</i>	2	---	-0.021	0.005	4.44E-06	31.1	0.225	-0.021	0.006	0.0002
ch.7.120944214 F <sup>*</sup>	IGR		7	---	-0.004	0.001	4.81E-06	0	0.635	-0.004	0.001	4.81E-06
cg08841413	IGR		1	++++	0.007	0.001	5.12E-06	52	0.1	0.007	0.002	0.0017
cg25429619 <sup>#</sup>	5'UTR	<i>RABGAP1L</i>	1	---	-0.017	0.004	5.32E-06	42.2	0.158	-0.018	0.005	0.0005
cg03717591 <sup>*</sup>	IGR	<i>RP11-405A12.2</i>	12	++++	0.007	0.001	5.52E-06	0	0.529	0.007	0.001	5.52E-06
cg25173010 <sup>*</sup>	TSS1500	<i>TXNIP</i>	1	---	-0.005	0.001	5.91E-06	0	0.964	-0.005	0.001	5.91E-06
cg10090316 <sup>*</sup>	5'UTR	<i>TREFF1</i>	6	++++	0.009	0.002	6.57E-06	0	0.646	0.009	0.002	6.57E-06
cg18001026	TSS1500	<i>IL2</i>	4	++++	0.006	0.001	7.45E-06	2.5	0.38	0.006	0.001	1.14E-05
cg22620614 <sup>*</sup>	Body	<i>SHANK2</i>	11	---	-0.005	0.001	7.52E-06	0	0.752	-0.005	0.001	7.52E-06

Sixteen CpG probes associated with prior SARS-CoV-2 infection ( $p < 1 \times 10^{-5}$ ) in a meta-analysis of 675 samples from four cohorts. Bonferroni-significant CpGs ( $p < 7.48 \times 10^{-8}$ , FE) are in bold; those also significant in RE are marked with \*. CpG: probe; Genomic location: Genomic annotation categories in relation to regulatory features (TSS200: CpG located within 200 bp upstream of transcription start site (TSS), TSS1500: CpG 200–1500 bp upstream of TSS, 5'UTR: located in 5' untranslated region, 1st exon: CpG located in first coding exon, Body: within gene body, IGR: intergenic region); I<sup>2</sup>: statistic measuring heterogeneity on scale of 0–100%; HetPVal: P-value for heterogeneity statistic. Previously reported CpG sites associated with COVID-19 were labeled as “#”.



**Fig. 3** Forest plots of CpG~case status associations. Circles (cohorts) show effect sizes with 95% CI, scaled by sample size; diamonds show pooled FE meta-analysis estimates. Probes are ordered by significance

of the DMR analyses. We also demonstrate that SARS-CoV-2-associated DNAm alterations at both the DMP and DMR levels are linked to transcriptional differences, with downstream analyses indicating these regions are involved in immune response pathways and pro-inflammatory processes.

Cell deconvolution methods were applied to estimate WBC proportions across cohorts; no significant differences in their distribution were observed between cases and controls, except for basophil levels in NAKO. Two prior studies on DNAm changes 3–6 months post-infection also found no evidence of altered immune system composition in individuals with mild infection [11, 13], in contrast to significant changes correlated with COVID-19 severity observed during the acute phase of infection [1, 2]. In line with this and our WBC-adjusted analyses, most genes mapped to the top DMP/DMR associations did not show cell-type-specific expression in the HPA datasets. Adjusting for smoking and prevalent chronic disease attenuated the estimated effects, but the associations remained robust. This was expected, as both smoking and chronic disease are known risk factors for severe COVID-19 [46]. Furthermore, recent evidence suggests that smoking has short- and long-term effects mediated by DNAm on immune regulation that persist after smoking cessation [47]. Despite the high heterogeneity

observed in two of the three strongest CpG-case status associations, the direction of effect was consistent across studies, and the associations remained robust in the leave-one-out and serology-based sensitivity analyses.

Our DMP and DMR results are in line with epigenetic changes seen in severe COVID-19 during the acute phase of infection [1–7]. The majority of probes associated with case status in this study exhibited lower DNA methylation levels, consistent with hypomethylation associated with severe COVID-19 during early disease [1, 4, 5] and three to twelve months post-infection [9, 11, 48]. Several interferon-stimulated genes (ISG) identified in the DMP analyses –namely *IFI44L*, *MX1* and *DDX60* – have been consistently identified in severe [1–3, 5] and early mild COVID-19 [6]. *IFI44L* is involved in response to type I interferon (IFN), a key pathway in COVID-19 disease development and severity [49–51]. IFN are signalling cytokines involved in the regulation of the immune response to both viral and bacterial infections [52] and autoimmune disease [53]. *MX1* (MX Dynamin Like GTPase 1) is an interferon-induced gene (ISG) encoding for a GTPase that inhibits viral replication; genetic variation in *MX1* is associated with differential risk to developing severe COVID-19 [54] and its differential methylation is linked to disease progression [55]. Our DMP and eQTM results are in line with previous

**Table 3** Differentially methylated regions (DMRs) associated with case status in the meta-EWAS

DMR ID	DMR description	N probes	Sidak pval	Overlapping CpGs	Genomic annotation
1*	chr13:24519920–24,520,508	5	3.05E-10	cg02002217;cg08405447;cg11450541;cg16014862;cg18976044	[IGR]; <i>AN-KRD20A19P</i> [body]
2*	chr17:79495006–79,495,468	7	1.69E-08	cg09084279;cg13468858;cg13700073;cg15774065;cg16507479;cg18919720;cg20932150	<i>FSCN2</i> [TSS200, TSS1500, 5'UTR]
3*	chr13:23309689–23,310,226	7	2.35E-08	cg01863042;cg03042692;cg05215994;cg08083251;cg15973954;cg20395040;cg26361286	[IGR]
4*	chr7:2802554–2,803,066	8	7.66E-07	cg00947599;cg03459839;cg05793240;cg12102607;cg12444411;cg20884522;cg23285459;cg23333125	<i>GNA12</i> [TSS200, TSS1500, body]
5*	chr3:127634188–127,634,460	6	6.32E-06	cg05768427;cg08488841;cg10961186;cg12311882;cg13859958;cg14701491	[IGR]
6	chr14:70690287–70,690,705	6	4.94E-05	cg04204356;cg08322580;cg21484985;cg23442650;cg23618713;cg26146732	[IGR]
7*	chr1:234367443–234,367,587	3	8.10E-05	cg09119043;cg23445321;cg24723561	<i>SLC35F3</i> [body]
8*	chr19:45885800–45,886,079	4	0.000153258	cg04683509;cg16655626;cg21219373;cg23397216	<i>PPP1R13L</i> [body]
9	chr8:102236522–102,236,832	5	0.000166719	cg02168442;cg03250224;cg13211302;cg16058797;cg23042086	[IGR]
10*	chr3:122281939–122,282,158	3	0.000198913	cg00959259;cg07815522;cg08122652	<i>PARP9</i> [5' UTR]
11	chr6:2876762–2,877,000	4	0.000279189	cg03500459;cg07704716;cg11678426;cg20958467	<i>SERPINB9P1</i> [TSS200, TSS1500]
12	chr2:121338498–121,338,608	3	0.000524689	cg03792788;cg07218647;cg09408902	<i>STX2</i> [body]
13	chr12:131303016–131,303,248	4	0.000653375	cg05407909;cg10832093;cg11011512;cg20050828	[IGR]
14	chr6:134436460–134,436,741	3	0.001424293	cg07710266;cg12383699;cg14997321	<i>PF4</i> [TSS200, TSS1500, 1st exon]
15	chr4:74847710–74,848,017	7	0.001831057	cg02530824;cg05509609;cg06834998;cg13126871;cg15398841;cg16072462;cg21043213	<i>UCN3</i> [TSS200, 1st exon]
16	chr10:5406890–5,407,120	7	0.003325681	cg01134643;cg11218842;cg12428514;cg15671083;cg17158564;cg26644395;cg27349333	<i>UCN3</i> [TSS200, 1st exon]
17	chr7:156735383–156,735,657	4	0.004911592	cg03930209;cg17604655;cg22871949;cg27539527	<i>THRB</i> [TSS1500]
18	chr3:24536765–24,536,890	4	0.005301857	cg04928005;cg25829666;cg26139133;cg26146027	[IGR]
19	chr11:69286231–69,286,353	5	0.009814467	cg00695387;cg13293618;cg15442037;cg18498241;cg21408915	[IGR]
20	chr1:95008041–95,008,118	3	0.013555894	cg03460527;cg09549015;cg25893275	<i>F3</i> [TSS1500]
21	chr12:7781093–7,781,289	3	0.020072455	cg10578777;cg14906510;cg25828445	[IGR]

DMR ID: Identifier assigned to each differentially methylated region detected in the meta-EWAS; DMR description — Genomic coordinates of the region (chromosome, start, end) summarizing the span of the DMR; N probes — Number of Illumina EPIC CpG probes constituting the region; Sidak pval: Sidak-adjusted region-level p-value; Overlapping CpGs: individual CpG probe IDs located within the DMR boundaries; genomic annotation: annotation of the DMR based on Illumina EPIC definitions, including annotated gene and gene context (TSS200, TSS1500, 5'UTR, 1st Exon, Body, 3'UTR, IGR). DMRs replicated across sensitivity analyses are marked with a star (\*) next to their ID

reports of hypomethylation in *IFI44L* both during acute infection and persisting for months thereafter [3, 8, 11]. Hypomethylation of *MXI* was identified across COVID-19 severity groups [2, 6], was specific to COVID-19 patients in comparison to patients with other respiratory diseases [4, 5] and was still present at 12 months after infection [8, 10]. In addition to cg03607951-*IFI44L* and cg22862003-*MXI*, cg24678928 (*DDX60*) and cg25429619 (*RABGAP1L*) were hypomethylated in COVID-19 patients during the acute-phase of disease [5]; likewise, altered DNAm in cg24678928-*DDX60* and cg22862003-*MXI* was observed up to 3 months post-infection [11]. *DDX60* (DEXD/H-Box Helicase 60), encodes for an RNA

helicase acting as an antiviral factor involved in the interferon inducible response to viral infection, and it was included in DMP signatures in COVID-19 cases [5, 56]. Interestingly, hypomethylation at cg24678928-*DDX60* was correlated with worse quality of life in convalescent COVID-19 patients with Post-Acute COVID-19 Syndrome (PASC) [57].

These DMPs partially overlap DNAm changes in autoimmune and inflammatory disease [8, 11] – e.g. systemic lupus erythematosus [58], primary Sjögren's Syndrome [59] and rheumatoid arthritis [60] – but also feature in interferon response to COVID-19 [61]. Importantly, individuals with autoimmune disease were excluded from

analyses in cohorts where this information was available (see Supplemental Note 1). This points to shared dysregulation of the innate immune system [5, 6], a cross-talk between acute COVID-19 pathogenesis and autoimmune disease [62, 63], and further supports DNAm of ISGs to be involved in long-term dysregulation [8].

DMR analyses identified ten additional protein-coding genes associated with case status. Of these, DMRs in *PARP9* and *GNA12* show the strongest and most robust association with case status, as seen across sensitivity analyses, replication in at least one cohort, and supportive eQTM data. Although expressed in immune cells, their expression is not cell-type-specific. The location of DMR4-*GNA12* (in proximity to TSS) and DMR10-*PARP9* (5' UTR) suggests these changes might regulate gene expression. In line with this, our eQTM results show hypomethylation across CpGs within DMR10 correlates with higher expression of several genes (*PARP9*, *PARP14*, *DTX3L*), and hypomethylation at DMR4 with expression of *GNA12*. *PARP9-DT3XL* is an interferon-induced module crucial to regulating *PARP14* activity in a cascade of antiviral signaling [64], and *GNA12* encodes a G-protein subunit in signaling hub involved in cancer progression, vascular dysfunction and inflammation [65]. Our *PARP9* DMR and eQTM results replicate observations in COVID patients versus non-COVID respiratory controls [4]; likewise, DMR4 is in proximity to a DMP identified across COVID severity groups [6] and to two DMRs persistently hypomethylated one year post-infection [48]. Interestingly, this DMR almost perfectly replicates the COVID-19-specific DMR reported in [56]. Hypomethylation of *GNA12* and its corresponding increased gene expression have been reported one year post-infection in hospitalized patients [48]. DMR18 (*THRB*) was also replicated in at least one cohort and across all sensitivity analyses but SA1 (most strict seed value). *THRB* encodes for a receptor involved in thyroid-hormone signaling in metabolism and endocrine homeostasis, while its expression in immune cells points to endocrine-immune cross-talk [66]. Additional DMRs mapped to genes involved in endocrine-metabolic dysregulation and stress (*UCN3*, *SLC35F3*), as well as in inflammatory signaling (*FSCN12*, *GNA12*, *PPP1R13L*) and platelet activation and coagulation (*PF4*, *F3*, *STX2*) (Suppl. Table 22). Similar pathways were identified in a study on differential DNAm in mucosal nasal tissue of COVID-19 patients during the acute infection phase [67]. This study is the first to report persistent differential DNAm in these genes months after a SARS-CoV-2 infection.

We identified biologically plausible changes at the DMP and DMR level in whole blood of mild cases up to four months post-infection, with independent eQTM data supporting the potential functional impact in gene regulation of these subtle DNAm changes. Our findings

suggest that subtle but persistent DNAm changes contribute to long-term epigenetic remodeling following a SARS-CoV-2 infection that overlaps pathways involved in metabolic, inflammatory and autoimmune disease. Other changes we identified in intergenic regions and pseudogenes are less directly related to specific molecular pathways, but potentially reflect widespread changes in the epigenome [68]. Despite multiple levels of heterogeneity in studies assessing long-term epigenetic changes (populations assessed, follow-up time and disease severity groups), DNAm changes at the DMP- and DMR-level partially replicate those observed in multiple independent studies. Our findings suggest some epigenetic injuries acquired during infection are persistent months after infection, although the epigenetic signature reported here is associated with mild infection in the general population and is therefore more subtle than the large effects observed during severe infections in hospitalized patients. A recent review on DNAm in persistent COVID-19 suggests infection alters the epigenetic landscape [69], with these changes potentially contributing to the molecular signature of long COVID [70].

This study is the largest multi-cohort cross-sectional EWAS in the months following a SARS-CoV-2 infection in population-based cohorts. Associations were consistent across multiple European cohorts, and the inclusion of non-vaccinated individuals allowed to control for confounding by vaccination-induced changes. Unlike earlier studies detecting large effects in small sample sizes, this study was designed to detect small sized effects (Cohen's  $d=0.2$ ), which enabled the identification of less pronounced DNAm changes months after infection. These findings suggest that subtle but persistent epigenetic changes in blood mirror alterations in well-described immune and inflammatory pathways, help further understand the multi-systemic nature of long-term biological dysregulation following SARS-CoV-2 infection and offer initial insights into potential biomarkers for risk stratification or therapeutic intervention.

This study has several limitations. Time between infection and blood sampling (DNAm) was variable within and across cohorts, which contributed to case heterogeneity and is an important limitation inherent to cohorts' study designs. Case definition in each cohort relied on differing SARS-CoV-2 detection methods (test-based or self-reported), so it may have been affected by recall and misclassification bias; however, sensitivity analyses using a serology-based case definition in a subset of the participants produced same conclusions. NAKO and Lifelines were identified as likely influential studies, potentially due to larger sample sizes and varying case definitions, yet leave-one-out sensitivity analyses produced similar results. Although inflation at the study-level was not significant, residual confounding is inherent

to observational studies like this one. Beyond cohort heterogeneity, underlying inter-individual variation in disease severity/symptomatology and recovery cloud the interpretation of the results in terms of molecular mechanisms and pathophysiology. Additional limitations include limited coverage of the epigenome by using array-based methods, assessing autosomes only and profiling mixed white blood cell populations instead of other potentially more informative tissues [67, 71]. Study designs across cohorts did not allow for longitudinal analyses, assessment of concomitant gene expression or genetic variation, or the examination of epigenetic changes considering long-COVID and related outcomes; these aspects should be considered by future studies. External validation of our findings in independent cohorts will be important to confirm the robustness and generalizability of the DNAm changes identified, as this study only included European cohorts representative from the early pandemic.

## Conclusion

This study represents the first large-scale EWAS examining the long-term epigenetic landscape of mild SARS-CoV-2 infection. Our findings indicate moderate but persistent changes in whole blood up to 4 months post-infection. This study adds to the growing evidence on the long-term epigenetic signatures of COVID-19 infection and its overlap with inflammation, metabolic and immune dysregulation. We confirmed associations previously reported for *IFI44L*, *MX1*, *DDX60*, *PARP9* and *GNAI2*, and identify additional regions warranting further research. This study identifies persistent epigenetic injuries acquired during acute phase of disease and presents initial evidence suggesting a few of these changes to be correlated with quality of life following an infection. Future longitudinal studies having concomitant DNAm and omics data should investigate the downstream effects of epigenetic changes and their change over time; the study setup of future studies should also consider appropriate collection of data on clinical symptoms long after an infection to contribute to molecular understanding of epigenetic changes associated with post-acute sequelae of SARS-CoV-2 infection.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-026-02166-1>.

Supplementary Material 1

## Acknowledgements

We would like to extend our appreciation to the members of the ORCHESTRA Advisory Panel: Prof. Marc Lallemand, Prof. Leonard Leibovici, Prof. Jaques Scheres, Prof. Ranjeeta Thomas, Prof. Maria Vehreschild and Dr. Catherine Cohet (EMA) and Dr. Anastasia Pharris (ECDC) as well as to the Ethics Board:

Prof. Greetleven, Prof. Nikolaus Forgó and Prof. Satinder Aneja for their support and guidance throughout. Methylation analysis and initial quality control was performed at the Core Facility Genotyping (Helmholtz Munich) by Nadine Lindemann under the supervision of Dr. Melanie Waldenberger. Further methylation data quality control and preprocessing was performed by Dr. Thomas Delerue, Research Unit Molecular Epidemiology, Institute of Epidemiology, Helmholtz Munich. Special thanks to Dr. Bärbel Fösel and Dr. Sapna Sharma, from the Research Unit Molecular Epidemiology, Institute of Epidemiology, Helmholtz Munich, for providing mentoring and scripts to conduct quality control of GSA data. We thank all participants who took part in the NAKO study and the staff in this research program. The authors also wish to acknowledge the services of the Lifelines Cohort Study, the contributing research centres delivering data to Lifelines, and all the study participants. We would like to give special thanks to all participants of the CON-VINCE study and to each volunteer that helped in sample processing or sample kit preparation. We acknowledge the joint effort of the CON-VINCE team involved in sample and data collection. We also would like to thank the Ministry of Health of Luxembourg as well as the Directorate of Health of Luxembourg for their support. We would like to thank the Research Luxembourg COVID-19 Task Force (Frank Glod, Paul Wilmes, Lars Geffers, Jasmin Schulz, Henry-Michel Cauchie, Ulf Nehrbass, Rudi Balling) for their overall support of the CON-VINCE study. A special thanks also to the data protection officers Sandrine Munoz and Laurent Prévotat. Furthermore, we would like to acknowledge the whole Communication teams involved within CON-VINCE, especially Sabine Schmitz, Arnaud D'Agostini, Didier Gossens, Hélène Jacuszin for their excellent work and support during the implementation and execution of CON-VINCE. We would like to thank Philippe Lamesch for important and successful fundraising for research on COVID-19 in Luxembourg.

## Author contributions

Authors' list with affiliations and contributions were compiled using the online tool tenzing [72, 73]. Conceptualization: P.R.M.-G., M.W., A.P., I.M.H., J.O., R.K., M.P.O.S., R. W., L.F., J.S.-M., N.M., K.G., B.B., K.B., J.F., and P.A. Data curation: P.R.M.-G., I.M.H., J.O., A.R., M.P.O.S., A.K., R. W., J.S.-M., N.M., K.G., B.B., K.B., J.F., and P.A. Formal analysis: P.R.M.-G., L.L., T.D., and I.M.H. Funding acquisition: A.P., K.Ü., R. W., R.K., and L.F. Investigation: P.R.M.-G., O.G., R. W., L.F., B.B., K.B., J.F., and P.A. Methodology: P.R.M.-G., M.W., L.L., T.D., A.P., I.M.H., K.J.S., O.G., K.Ü., R. W., J.O., L.F., J.S.-M., N.M., K.G., B.B., K.B., J.F., and P.A. Project administration: P.R.M.-G., M.W., A.P., K.J.S., O.G., K.Ü., R. W., J.O., R.K., R. Warmerdam, L.F., J.S.-M., N.M., K.G., B.B., K.B., J.F., and P.A. Resources: M.W., A.P., I.M.H., K.J.S., O.G., K.Ü., R. W., J.O., R.K., L.F., J.S.-M., N.M., and K.G. Supervision: P.R.M.-G., M.W., A.P., I.M.H., K.J.S., O.G., K.Ü., R. W., R.K., and L.F. Validation: P.R.M.-G., L.L., and T.D. Visualization: P.R.M.-G., L.L., and T.D. Writing - original draft: P.R.M.-G., M.W., and A.P. Writing - review & editing: P.R.M.-G., M.W., L.L., T.D., A.P., I.M.H., K.J.S., O.G., K.Ü., R. Wagner, J.O., R.K., A.R., M.P.O.S., A.K., R. W., L.F., J.S.-M., N.M., K.G., B.B., K.B., J.F., and P.A.

## Funding

Open Access funding enabled and organized by Projekt DEAL. The present study was carried out in the context of the ORCHESTRA Project ([www.orchestra-cohort.eu](http://www.orchestra-cohort.eu)). ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101016167. The views expressed in this abstract are the sole responsibility of the authors and the Commission is not responsible for any use that may be made of the information it contains. This project was conducted with data from the German National Cohort (NAKO, [www.nako.de](http://www.nako.de)). The NAKO is funded by the Federal Ministry of Education and Research (BMBF, project funding reference numbers: 01ER1301A/B/C and 01ER1511D), federal states and the Helmholtz Association with additional financial support by the participating universities and the institutes of the Leibniz Association. The TiKoCo study was supported by the Bavarian Ministry of Science and Arts as well as by the National Research Network of the University Medicine. The Lifelines initiative has been made possible by subsidy from the Dutch Ministry of Health, Welfare and Sport, the Dutch Ministry of Economic Affairs, the University Medical Center Groningen (UMCG), Groningen University and the Provinces in the North of the Netherlands (Drenthe, Friesland, Groningen). FES (Fonds Economische Structuurversterking), SNN (Samenwerkingsverband Noord Nederland) and REP (Ruimtelijk Economisch Programma). Furthermore, this project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101057553. The CON-VINCE Study is funded by the National Research Fund Luxembourg (14716281/CONVINCE/Kruger) and the André Losch Foundation (Luxembourg).

### Data availability

NAKO: Data of the NAKO are generally not available to the public due to strict data protection regulations. However, scientists can apply for data use according to the official usage regulation specifications. Please refer to <http://transfer.nako.de> for further information. Lifelines: Data may be obtained from a third party and are not publicly available. Researchers can apply to use the Lifelines data used in this study. More information about how to request Lifelines data and the conditions of use can be found on their website. CON-VINCE: The dataset for this manuscript is not publicly available as it is linked to the CON-VINCE Study and its internal regulations. Any requests for accessing the dataset can be directed at [con-vince@ih.lu](mailto:con-vince@ih.lu). All data of the manuscript will be provided upon reasonable request and approval by the ethics committee. TiKoCo: The dataset for this manuscript is not publicly available as it is linked to the TiKoCo Study and its internal regulations. Any requests for accessing the dataset can be directed to the Study Pls, Prof. Dr. Ralf Wanger ([ralf.wagner@ukr.de](mailto:ralf.wagner@ukr.de)) and Prof. Dr. med. Klaus Überla ([klaus.ueberla@uk-erlangen.de](mailto:klaus.ueberla@uk-erlangen.de)). DNAm datasets: DNAm datasets are stored independently by each cohort and therefore protected by the same data use and access policies mentioned above.

### Declarations

#### Ethics approval and consent to participate

The NAKO obtained ethics approval from all local committees of the study centers (original ethics approval of the leading ethics committee of the Bayerische Landesärztekammer with protocol code 13023). All participants provided written informed consent. The TiKoCo study was approved by the Ethics Committee of the University of Regensburg, Germany (vote 20-1867-101; 22 May 2020) and adopted by the Ethics Committee of the University of Erlangen (vote 248\_20 Bc; 22 May 2020). The study complies with the 1964 Helsinki Declaration and its later amendments. All participants provided written informed consent. The general Lifelines protocol has been approved by the UMG Medical ethical committee under number 2007/152. This study was conducted according to the Ethical Principles for Medical Research Involving Human Subjects, as stated in the 2013 revised version of the 1964 World Medical Association Declaration of Helsinki. The national research ethics committee (Comité National d'Ethique de Recherche, CNER) and the Luxembourgish Ministry of Health (references 202004/01 and 831 x6ce0d, respectively) approved the study. Trial registration number is NCT04379297. All participants completed an electronic Informed Consent Form and had the right to withdraw from the study at any time.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

#### Author details

- <sup>1</sup>Research Unit Molecular Epidemiology, Helmholtz Munich, Neuherberg, Germany
- <sup>2</sup>Institute of Epidemiology, Helmholtz Munich, Neuherberg, Germany
- <sup>3</sup>German Research Center for Cardiovascular Disease (DZHK), partner site Munich Heart Alliance, Munich, Germany
- <sup>4</sup>Institute for Medical Information Processing, Biometry, and Epidemiology (IBE), Faculty of Medicine, Ludwig-Maximilians-Universität München, Munich, Germany
- <sup>5</sup>Luxembourg Institute of Health, Strassen, Luxembourg
- <sup>6</sup>Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany
- <sup>7</sup>Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
- <sup>8</sup>Onco Institute, Utrecht, The Netherlands
- <sup>9</sup>Department for Diagnostic and Interventional Radiology, University Hospital Augsburg, Augsburg, Germany
- <sup>10</sup>Leibniz Institute for Prevention Research and Epidemiology - BIPS, Bremen, Germany
- <sup>11</sup>Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Esch-sur-Alzette, Luxembourg
- <sup>12</sup>NAKO e.V., Heidelberg, Germany

- <sup>13</sup>Institute of Epidemiology and Social Medicine, University of Münster, Münster, Germany
- <sup>14</sup>Institute of Social Medicine, Epidemiology and Health Economics, Charité – Universitätsmedizin Berlin, Berlin, Germany
- <sup>15</sup>Unit for Municipal Health Strategies for the City of Freiburg and the District of Breisgau-Hochschwarzwald, Freiburg, Germany
- <sup>16</sup>Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany
- <sup>17</sup>Centre Hospitalier du Luxembourg, Strassen, Luxembourg
- <sup>18</sup>Department of Medical Informatics, Biometry and Epidemiology, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
- <sup>19</sup>Institute of Clinical and Molecular Virology, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
- <sup>20</sup>Institute of Medical Microbiology and Hygiene, Molecular Microbiology (Virology), University of Regensburg, Regensburg, Germany
- <sup>21</sup>Institute of Clinical Microbiology and Hygiene, University Hospital Regensburg, Regensburg, Germany

Received: 2 April 2025 / Accepted: 12 May 2026

Published online: 13 June 2026

### References

- Bernardes JP, Mishra N, Tran F, Bahmer T, Best L, Blase JI, et al. Longitudinal multi-omics analyses identify responses of megakaryocytes, erythroid cells, and plasmablasts as hallmarks of severe COVID-19. *Immunity*. 2020;53(6):1296–314.
- Corley MJ, Pang APS, Dody K, Mudd PA, Patterson BK, Seethamraju H, et al. Genome-wide DNA methylation profiling of peripheral blood reveals an epigenetic signature associated with severe COVID-19. *J Leukoc Biol*. 2021;110(1):21–6.
- Castro de Moura M, Davalos V, Planas-Serra L, Alvarez-Errico D, Arribas C, Ruiz M, et al. Epigenome-wide association study of COVID-19 severity with respiratory failure. *EBioMedicine*. 2021;66:103339.
- Balnis J, Madrid A, Hogan KJ, Drake LA, Chieng HC, Tiwari A, et al. Blood DNA methylation and COVID-19 outcomes. *Clin Epigenetics*. 2021;13(1):118.
- Konigsberg IR, Barnes B, Campbell M, Davidson E, Zhen Y, Pallisard O, et al. Host methylation predicts SARS-CoV-2 infection and clinical outcome. *Commun Med*. 2021;1(1):42.
- Barturen G, Carnero-Montoro E, Martínez-Bueno M, Rojo-Rello S, Sobrino B, Porras-Perales Ó, et al. Whole blood DNA methylation analysis reveals respiratory environmental traits involved in COVID-19 severity following SARS-CoV-2 infection. *Nat Commun*. 2022;13(1):4597.
- Calzari L, Zanotti L, Inglese E, Scaglione F, Cavagnola R, Ranucci F, et al. Role of epigenetics in the clinical evolution of COVID-19 disease. Epigenome-wide association study identifies markers of severe outcome. *Eur J Med Res*. 2023;28(1):81.
- Mao W, Miller CM, Nair VD, Ge Y, Amper MAS, Cappuccio A, et al. A methylation clock model of mild SARS-CoV-2 infection provides insight into immune dysregulation. *Mol Syst Biol*. 2023;19(5):e11361.
- Hong P, Waldenberger M, Pritsch M, Gilberg L, Brand I, Bruger J, et al. Differential DNA methylation 7 months after SARS-CoV-2 infection. *Clin Epigenetics*. 2025;17(1):60.
- Balnis J, Madrid A, Hogan KJ, Drake LA, Adhikari A, Vancavage R, et al. Persistent blood DNA methylation changes one year after SARS-CoV-2 infection. *Clin Epigenetics*. 2022;14(1):94.
- Lee Y, Riskedal E, Kalleberg KT, Istre M, Lind A, Lund-Johansen F, et al. EWAS of post-COVID-19 patients shows methylation differences in the immune-response associated gene, *IFI44L*, three months after COVID-19 infection. *Sci Rep*. 2022;12(1):11478.
- Nikesjö F, Sayyab S, Karlsson L, Apostolou E, Rosén A, Hedman K, et al. Defining post-acute COVID-19 syndrome (PACS) by an epigenetic biosignature in peripheral blood mononuclear cells. *Clin Epigenetics*. 2022;14(1):172.
- Calzari L, Dragani DF, Zanotti L, Inglese E, Danesi R, Cavagnola R, et al. Epigenetic patterns, accelerated biological aging, and enhanced epigenetic drift detected 6 months following COVID-19 infection: insights from a genome-wide DNA methylation study. *Clin Epigenetics*. 2024;16(1):112.
- ORCHESTRA - EU Horizon 2020 cohort to tackle COVID-19 internationally. 2020–2024 [cited 2025]. <https://orchestra-cohort.eu/>

15. Peters A, Peters A, Greiser KH, Göttlicher S, Ahrens W, Albrecht M, et al. Framework and baseline examination of the German National Cohort (NAKO). *Eur J Epidemiol*. 2022;37(10):1107–24.
16. Snoeck CJ, Vaillant M, Abdelrahman T, Satagopam VP, Turner JD, Beaumont K et al. Prevalence of SARS-CoV-2 infection in the Luxembourgish population – the CON-VINCE study. *medRxiv*. 2020:2020.05.11.20092916.
17. Tsurkalenko O, Bulaev D, O'Sullivan MP, Snoeck C, Ghosh S, Kolodkin A, et al. Creation of a pandemic memory by tracing COVID-19 infections and immunity in Luxembourg (CON-VINCE). *BMC Infect Dis*. 2024;24(1):179.
18. Mc Intyre K, Lanting P, Deelen P, Wiersma HH, Vonk JM, Ori APS, et al. Lifelines COVID-19 cohort: investigating COVID-19 infection and its health and societal impacts in a Dutch population-based cohort. *BMJ Open*. 2021;11(3):e044474.
19. Scholtens S, Smidt H, Swertz MA, Bakker SJ, Dotinga A, Vonk JM, et al. Cohort profile: lifelines, a three-generation cohort study and biobank. *Int J Epidemiol*. 2015;44(4):1172–80.
20. Einhauser S, Peterhoff D, Beileke S, Günther F, Niller HH, Steininger P, et al. Time trend in SARS-CoV-2 seropositivity, surveillance detection- and infection fatality ratio until spring 2021 in the Tirschenreuth County-results from a population-based longitudinal study in Germany. *Viruses*. 2022. <https://doi.org/10.3390/v14061168>.
21. Wagner R, Peterhoff D, Beileke S, Günther F, Berr M, Einhauser S, et al. Estimates and determinants of SARS-Cov-2 seroprevalence and infection fatality ratio using latent class analysis: the population-based Tirschenreuth study in the hardest-hit German county in spring 2020. *Viruses*. 2021. <https://doi.org/10.3390/v13061118>.
22. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363–9.
23. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol*. 2015;16(1):37.
24. Hellbach F, Sinke L, Costeira R, Baumeister SE, Beekman M, Louca P, et al. Pooled analysis of epigenome-wide association studies of food consumption in KORA, TwinsUK and LLS. *Eur J Nutr*. 2023;62(3):1357–75.
25. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
26. Salas LA, Zhang Z, Koestler DC, Butler RA, Hansen HM, Molinaro AM, et al. Enhanced cell deconvolution of peripheral blood using DNA methylation for high-resolution immune profiling. *Nat Commun*. 2022;13(1):761.
27. Van der Most PJ, Küpers LK, Snieder H, Nolte I. QC EWAS: automated quality control of results of epigenome-wide association studies. *Bioinformatics*. 2017;33(8):1243–5.
28. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190–1.
29. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health*. 2019;22(4):153–60.
30. Harrer M, Cuijpers P, Furukawa TA, Ebert DD. Doing meta-analysis with R: a hands-on guide. 1st ed. Boca Raton, FL and London: Chapman & Hall/CRC Press; 2021. 2021.
31. Harrer M, Cuijpers P, Furukawa T, Ebert DD. dmetar: companion R package for the guide 'doing meta-analysis in R 2019'. <http://dmetar.protectlab.org/>
32. Xu Z, Niu L, Taylor JA. The ENmix DNA methylation analysis pipeline for Illumina BeadChip and comparisons with seven other preprocessing pipelines. *Clin Epigenetics*. 2021;13(1):216.
33. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. *PLoS Comput Biol*. 2013;9(8):e1003118.
34. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics*. 2017;33(24):3982–4.
35. Maksimovic J, Oshlack A, Phipson B. Gene set enrichment analysis for genome-wide DNA methylation data. *Genome Biol*. 2021;22(1):173.
36. Agrawal A, Balci H, Hanspers K, Coort SL, Martens M, Slenter DN, et al. WikiPathways 2024: next generation pathway database. *Nucleic Acids Res*. 2023;52(D1):D679–89.
37. Uhlen M, Karlsson MJ, Zhong W, Tebani A, Pou C, Mikes J, et al. A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. *Science*. 2019. <https://doi.org/10.1126/science.aax9198>.
38. Human Protein Atlas. Human Protein Atlas [05.05.2026]. Available from: [www.proteinatlas.org](http://www.proteinatlas.org).
39. Kundaje A, Meuleman W, Ernst J, Bilenyk M, Yen A, Heravi-Moussavi A, et al. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518(7539):317–30.
40. Battram T, Yousefi P, Crawford G, Prince C, Sheikhalil Babaei M, Sharp G, et al. The EWAS Catalog: a database of epigenome-wide association studies. *Wellcome Open Res*. 2022;7:41.
41. Elizarraras JM, Liao Y, Shi Z, Zhu Q, Pico Alexander R, Zhang B. WebGestalt 2024: faster gene set analysis and new support for metabolomics and multi-omics. *Nucleic Acids Res*. 2024;52(W1):W415–21.
42. Shin MG, Pico AR. Using published pathway figures in enrichment analysis and machine learning. *BMC Genomics*. 2023;24(1):713.
43. Bonder MJ, Luijk R, Zernakova DV, Moed M, Deelen P, Vermaat M, et al. Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet*. 2017;49(1):131–8.
44. Min JL, Hemani G, Hannon E, Dekkers KF, Castillo-Fernandez J, Luijk R, et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat Genet*. 2021;53(9):1311–21.
45. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics*. 2012;28(10):1353–8.
46. Booth A, Reed AB, Ponzio S, Yassaee A, Aral M, Plans D, et al. Population risk factors for severe disease and mortality in COVID-19: a global systematic review and meta-analysis. *PLoS ONE*. 2021;16(3):e0247461.
47. Saint-André V, Charbit B, Biton A, Rouilly V, Possémé C, Bertrand A, et al. Smoking changes adaptive immunity with persistent effects. *Nature*. 2024;626(8000):827–35.
48. Balnis J, Madrid A, Hogan KJ, Drake LA, Adhikari A, Vancavage R, et al. 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(5):594–7.
49. Bastard P, Rosen LB, Zhang Q, Michailidis E, Hoffmann H-H, Zhang Y, et al. Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science*. 2020;370(6515):eabd4585.
50. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al. Genetic mechanisms of critical illness in COVID-19. *Nature*. 2021;591(7848):92–8.
51. Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science*. 2020;370(6515):eabd4570.
52. Perry AK, Chen G, Zheng D, Tang H, Cheng G. The host type I interferon response to viral and bacterial infections. *Cell Res*. 2005;15(6):407–22.
53. Barrat FJ, Crow MK, Ivashkiv LB. Interferon target-gene expression and epigenomic signatures in health and disease. *Nat Immunol*. 2019;20(12):1574–83.
54. Martínez-Diz S, Morales-Álvarez CM, García-Iglesias Y, Guerrero-González JM, Romero-Cachinero C, González-Cabezuelo JM, et al. Analyzing the role of ACE2, AR, MX1 and TMPRSS2 genetic markers for COVID-19 severity. *Hum Genomics*. 2023;17(1):50.
55. Zguro K, Fallarini C, Fava F, Furini S, Renieri A. Host genetic basis of COVID-19: from methodologies to genes. *Eur J Hum Genet*. 2022;30(8):899–907.
56. Bińkowski J, Taryma-Leśniak O, Łuczowska K, Niedzwiedz A, Lechowicz K, Strapagiel D, et al. Epigenetic activation of antiviral sensors and effectors of interferon response pathways during SARS-CoV-2 infection. *Biomed Pharmacother*. 2022;153:113396.
57. Balnis J, Madrid A, Drake LA, Vancavage R, Tiwari A, Patel VJ, et al. Blood DNA methylation in post-acute sequelae of COVID-19 (PASC): a prospective cohort study. *EBioMedicine*. 2024. <https://doi.org/10.1016/j.ebiom.2024.105251>.
58. Yeung KS, Chung BH, Choufani S, Mok MY, Wong WL, Mak CC, et al. Genome-wide DNA methylation analysis of Chinese patients with Systemic Lupus Erythematosus identified hypomethylation in genes related to the type I interferon pathway. *PLoS One*. 2017;12(1):e0169553.
59. Imgenberg-Kreuz J, Sandling JK, Almlöf JC, Nordlund J, Signér L, Norheim KB, et al. Genome-wide DNA methylation analysis in multiple tissues in primary Sjögren's syndrome reveals regulatory effects at interferon-induced genes. *Ann Rheum Dis*. 2016;75(11):2029–36.
60. Hillary RF, Ng HK, McCartney DL, Elliott HR, Walker RM, Campbell A, et al. Blood-based epigenome-wide analyses of chronic low-grade inflammation across diverse population cohorts. *Cell Genom*. 2024;4(5):100544.
61. Prasad DTEF, Weitz E, Willighagen E, Coort S, Garde MW, Hasbun JF, Willighagen L, Ono H. Network map of SARS-CoV-2 signaling pathway (WP5115) 2023-06-28. <https://www.wikipathways.org/instance/WP5115>

62. Ehrenfeld M, Tincani A, Andreoli L, Cattalini M, Greenbaum A, Kanduc D, et al. COVID-19 and autoimmunity. *Autoimmun Rev.* 2020;19(8):102597.
63. Sacchi MC, Tamiazzo S, Stobbione P, Agatea L, De Gaspari P, Stecca A, et al. SARS-CoV-2 infection as a trigger of autoimmune response. *Clin Transl Sci.* 2021;14(3):898–907.
64. Ribeiro VC, Russo LC, Hoch NC. PARP14 is regulated by the PARP9/DTX3L complex and promotes interferon  $\gamma$ -induced ADP-ribosylation. *EMBO J.* 2024;43(14):2908–28.
65. Guo P, Tai Y, Wang M, Sun H, Zhang L, Wei W et al. Ga12 and Ga13: versatility in physiology and pathology. *Front Cell Dev Biol.* 2022;10 (2022).
66. Wenzek C, Boelen A, Westendorf AM, Engel DR, Moeller LC, Führer D. The interplay of thyroid hormones and the immune system—where we stand and why we need to know about it. *Eur J Endocrinol.* 2022;186(5):R65–77.
67. Govender M, Das J, Hopkins FR, Svanberg C, Nordgren J, Hagbom M, et al. Altered DNA methylation pattern contributes to differential epigenetic immune signaling in the upper respiratory airway of unvaccinated COVID-19 patients. *Cells.* 2025;14(21):1673.
68. Shekhar Patil M, Richter E, Fanning L, Hendrix J, Wyns A, Barrero Santiago L, et al. Epigenetic changes in patients with post-acute COVID-19 symptoms (PACS) and long-COVID: a systematic review. *Expert Rev Mol Med.* 2024;26:e29.
69. Xiao Y, Vermund SH. DNA methylation in long COVID. *Front Virol.* 2024. <https://doi.org/10.3389/fviro.2024.1371683>.
70. Cao X, Li W, Wang T, Ran D, Davalos V, Planas-Serra L, et al. Accelerated biological aging in COVID-19 patients. *Nat Commun.* 2022;13(1):2135.
71. Li Y-Y, Yuan M-M, Li Y-Y, Li S, Wang J-D, Wang Y-F, et al. Cell-free DNA methylation reveals cell-specific tissue injury and correlates with disease severity and patient outcomes in COVID-19. *Clin Epigenetics.* 2024;16(1):37.
72. Holcombe AO, Kovacs M, Aust F, Aczel B. Documenting contributions to scholarly articles using CRediT and Tenzing. *PLoS ONE.* 2021;15(12):e0244611.
73. Kovacs M, Holcombe A, Aust F, Aczel B, Lawlor B. Tenzing and the importance of tool development for research efficiency. *Inform Serv Use.* 2021;41(1–2):123–30.

### Publisher's note

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