

1 **Epigenome-wide DNA methylation profiling of healthy COVID-19 recoverees reveals a**
2 **unique signature in circulating immune cells**

3 Johanna Huoman, PhD^{1*}, Shumaila Sayyab, PhD^{1*}, Eirini Apostolou, PhD², Lovisa Karlsson,
4 MSc¹, Lucas Porcile, BSc¹, Muhammad Rizwan, MSc², Sumit Sharma, PhD³, Jyotirmoy Das,
5 PhD¹, Anders Rosén², PhD and Maria Lerm, PhD^{1**}

6 *Shared first authorship

7 **Corresponding author

8

9 ¹Division of Inflammation and Infection, Department of Biomedical and Clinical Sciences,
10 Linköping University, Linköping, Sweden

11 ²Division of Cell Biology, Department of Biomedical and Clinical Sciences, Linköping
12 University, Linköping, Sweden

13 ³Division of Molecular Medicine and Virology, Department of Biomedical and Clinical
14 Sciences, Linköping University, Linköping, Sweden

15

16 Corresponding author:

17 Maria Lerm, PhD, Professor
18 Department of Biomedical and Clinical Sciences
19 Division of Inflammation and Infection, Lab1, floor 12
20 Faculty of Medicine and Health Sciences
21 SE-581 85 Linköping, Sweden
22 Cell phone: +46-73-2707786
23 Visit us at: <https://liu.se/en/employee/marle69>

24

25 **Word count:** 6070 (2404 without M&M and figure legends)

26 **Figures and tables:** 4 Figures, 12 Supplementary Figures, 8 Supplementary Tables

27 **ORCID-ID:s/Twitter accounts:**

28 Johanna Huoman 0000-0003-2509-2418 @HuomanJohanna

29 Shumaila Sayyab 0000-0002-6048-775X @shum_sayyab

30 Lovisa Karlsson 0000-0003-2704-1788 @LovisaElisabet

31 Lucas Porcile 0000-0002-9164-4113

32 Muhammad Rizwan 0000-0002-1725-8337

33 Sumit Sharma 0000-0003-0299-1285

34 Jyotirmoy Das 0000-0002-5649-4658 @Jyotirmoy21

35 Anders Rosén 0000-0001-5082-6423

36 Maria Lerm 0000-0002-5092-9892 @LermLab

37

38

39

40

41

42

43

44

45 **ABSTRACT**

46 **Background:** Epigenetic alterations upon microbial challenge have been described as both a
47 defence strategy and a result of pathogenic manipulation. While most COVID-19 studies
48 focus on inflammatory and immune-mediated responses, little is known about epigenetic
49 modifications in response to SARS-CoV-2 infection.

50 **Methods:** Epigenome-wide DNA methylation patterns from COVID-19 convalescents were
51 compared to uninfected controls from before and after the pandemic. Peripheral blood
52 mononuclear cell (PBMC) DNA was extracted from uninfected controls, COVID-19
53 convalescents and symptom-free individuals with SARS-CoV-2-specific T cell-responses, as
54 well as from PBMCs stimulated *in vitro* with SARS-CoV-2. Subsequently, the Illumina
55 MethylationEPIC 850K array was performed, and statistical/bioinformatic analyses comprised
56 differential DNA methylation, pathway over-representation and module identification
57 analyses.

58 **Results:** Differential DNA methylation patterns distinguished COVID-19 convalescents from
59 uninfected controls, with similar results in an experimental SARS-CoV-2 infection model. A
60 SARS-CoV-2-induced module was identified *in vivo*, comprising 66 genes of which six
61 (*TP53*, *INS*, *HSPA4*, *SPI1*, *ESR1* and *FAS*) were present in corresponding *in vitro* analyses.
62 Over-representation analyses revealed involvement in Wnt, muscarinic acetylcholine receptor
63 signalling and gonadotropin-releasing hormone receptor pathways. Furthermore, numerous
64 differentially methylated and network genes from both settings interacted with the SARS-
65 CoV-2 interactome.

66 **Conclusions:** Altered DNA methylation patterns of COVID-19 convalescents suggest
67 recovery from mild-to-moderate SARS-CoV-2 infection leaves longstanding epigenetic
68 traces. As *in vitro* SARS-CoV-2 infection corroborated *in vivo* exposure results, this indicates

69 DNA methylation is involved in immune cell responses to challenge with this virus. Future
70 studies should determine whether this reflects host-induced protective antiviral defence or
71 targeted viral hijacking to evade host defence.

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88 **Keywords:** COVID-19, DNA methylation, interactome, *in vitro* stimulation, mild-to-
89 moderate, module identification, network analysis, PBMC, SARS-CoV-2

90 INTRODUCTION

91 Since the end of 2019, the Corona virus disease 19 (COVID-19) pandemic has claimed lives
92 of millions world-wide, highlighting the global challenges in detecting, monitoring, and
93 treating novel viral infections. While efficacious vaccines are available at present, still a lot
94 remains to be uncovered regarding underlying mechanisms of the interaction between the
95 severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus and its host.

96 SARS-CoV-2 is a single stranded enveloped RNA virus belonging to the *Coronaviridae*
97 family¹, which similarly to other viruses hijacks host functions for its own advantage.²⁻⁵

98 Evidence suggest that epigenetic mechanisms, *i.e.* processes regulating transcriptional
99 accessibility of genomes without altering the nucleic acid sequence, are involved in the
100 hijacking process.^{6, 7} DNA methylation (DNAm) of CpG sites is considered to be the most
101 stable epigenetic modification, as it ensures heritability throughout cell division, although it is
102 at the same time highly dynamic in response to environmental stimuli.⁸ The malleability and
103 flexibility of the DNA methylome decreases with increasing age⁹, and environmental factors
104 such as smoking and nutrition may alter DNAm patterning in various cell types, including
105 different immune cells.^{10, 11} This could have important implications in the course of COVID-
106 19, as *e.g.* smokers, overweight and elderly patients are more susceptible to become severely
107 ill if contracting SARS-CoV-2.^{12, 13} Furthermore, DNAm patterns may also become altered
108 upon microbial¹⁴ or viral infection.¹⁵⁻¹⁷ In line with this, we have previously observed that
109 immune cells of asymptomatic, tuberculosis-exposed individuals carry a lasting DNAm
110 signature that is linked to protection against mycobacterial infection.¹⁸⁻²⁰

111 A majority (40-80%) of individuals infected with SARS-CoV-2 show no or mild symptoms of
112 COVID-19 and proceed into convalescence thereafter, while a smaller, but non-negligible,
113 proportion of individuals show severe or life-threatening manifestations.^{21, 22} Tolerant
114 immune responses have been observed in transcriptomic and immune profiling comparisons
115 of asymptomatic and symptomatic COVID-19 patients²³, and as epigenetic mechanisms
116 regulate differentiation of *e.g.* T cells⁸, it is conceivable that epigenetic mechanisms may be
117 implicated in combating SARS-CoV-2 infection. However, few studies have thus far
118 addressed whether and how the epigenome is altered in subjects with a recent mild-to-
119 moderate SARS-CoV-2 infection. In this study, we set out to examine epigenome-wide
120 DNAm patterns in convalescent COVID-19 (CC19) subjects, after a mild-to-moderate disease
121 course. Understanding how convalescent individuals have mounted an epigenetic response
122 against new viruses such as SARS-CoV-2, for which no pre-existent immunity was present,
123 may reveal how a functional defence strategy towards the virus is prepared, and guide
124 development of novel diagnostic and preventive measures. Indeed, we found that differential
125 DNAm patterns separated those who have not been infected with SARS-CoV-2 from those
126 who have recovered from mild COVID-19, suggesting that epigenetic mechanisms are at play
127 during SARS-CoV-2 infection. The observations could be replicated in *in vitro* experiments,
128 further underpinning our findings.

129

130

131

132

133

134

135

136

137

138 **MATERIALS AND METHODS**

139 **Study population**

140 In this study, participants were enrolled between May 29th and July 10th 2020 during the first
141 wave of the SARS-CoV-2 pandemic in Linköping, Sweden. Individuals who had recovered
142 from and individuals who had not experienced COVID-19 were recruited after
143 announcements with leaflets. Exclusion criteria were the existence of current active SARS-
144 CoV-2 infection and/or other infectious disease symptoms, as well as being younger than 18
145 years. The study participants voluntarily entered the study in a consecutive manner. The study
146 was conducted on blood and saliva samples from in total 38 individuals from three different
147 groups (Figure 1); non-infected controls (Con, n=18), COVID-19 convalescents (CC19,
148 n=14) and symptom-free individuals with SARS-CoV-2-specific T cell responses (SFT, n=6).
149 Additionally, blood samples from anonymous healthy blood donors from the blood bank at
150 Linköping University Hospital before 2020 were included as a separate group in the analyses
151 (pre20, n=5), collected between 2014-2019 prior to the outbreak of the pandemic. CC19
152 participants presented with either mild or asymptomatic initial infection, and none was
153 admitted to hospital. Cons were defined as neither having any positive circulating IgG-
154 antibody or T cell responses to SARS-CoV-2, while CC19s were defined by the presence of
155 SARS-CoV-2-specific IgG antibodies in plasma using suspension multiplex immunoassay
156 (SMIA), some of which were positive for IgG in saliva, rapid test and in T cell responses as
157 well. From the included individuals, the following information was retrieved using health
158 questionnaires: self-reported COVID-19 symptoms (if applicable, one or several of the

159 following: fever, headache, shortness of breath, loss of smell/taste, cough, fatigue, muscle
160 pain, nausea, sinusitis/congestion), date of self-reported symptoms, weeks between symptoms
161 and sampling, age, sex, smoking, weight, height, comorbidities as well as medications. The
162 blood and saliva from the study participants was processed in a Biosafety level-2 facility. The
163 present study is an exploratory pilot study on the effects of mild-to-moderate SARS-CoV-2
164 infection on DNA methylation patterns in PBMCs. The sample size could not be determined
165 beforehand, as the SARS-CoV-2 infection rate in society was not known at the time of sample
166 collection. Hence, all individuals fulfilling inclusion criteria, consenting to participation and
167 providing both blood and saliva samples were included in the study. However, the sample size
168 is similar to previous studies on the effect of BCG vaccination^{18, 24}, where meaningful
169 differences in DNA methylation upon tuberculosis infection were shown. Likewise, the
170 belonging to the different groups described above was determined after performance of the
171 DNA methylation analyses, and hence handling, extraction and experimental procedures
172 performed on the samples were performed in a blinded fashion. For samples from the natural
173 exposure cohort, all participants provided written informed consent, and the present study was
174 approved by the Regional Ethics Committee for Human Research in Linköping (Dnr. 2019-
175 0618). Regarding the anonymous blood samples used for *in vitro* experiments, informed
176 consent was given by the healthy donors at the time of blood donation and the use of the
177 donated blood for research purposes was guaranteed as per the guidelines of Regional Ethics
178 Committee for Human Research in Linköping and the Helsinki Declaration.

179 **PBMC and plasma isolation from whole blood**

180 Peripheral blood was collected in three 10 ml EDTA tubes (BD Vacutainer, 10331254, Fisher
181 Scientific, Sweden). Up to 20 ml of whole blood was used for PBMC isolation after Ficoll-
182 Paque Plus gradient centrifugation (GE17-1440-03, GE Healthcare Life Sciences, Sigma-
183 Aldrich, Sweden) with SepMateTM tubes (85450, StemCell technologies, France) according

184 to the manufacturer's protocol. Cells were frozen in 10% DMSO (10103483, Fischer
185 Scientific, Sweden) in fetal bovine serum (FBS) (10270106, Gibco, Fischer Scientific,
186 Sweden) and kept at -150°C until analysis. After thawing, the cells were washed twice in cell
187 culture medium (RPMI medium 1640, 31870-025, 10% fetal bovine serum, 1%
188 penicillin/streptomycin, 15140, 1% L-glutamine, 25030081, all from Gibco, Fischer
189 Scientific, Sweden) further on termed as complete culture medium, prior to further
190 processing. Up to 10 ml of whole blood was used for plasma separation by centrifugation
191 (2000g for 15min, 4°C) and aliquots were stored at -80°C till further analysis.

192 **Measurements of SARS-CoV-2-specific T cell responses using ELISpot**

193 Peptides for the spike (S) protein of SARS-CoV-2 were obtained from Mabtech (3629-1,
194 Sweden) and were reconstituted with di-methyl-sulphoxide (DMSO) at a concentration of 200
195 µg/ml according to the manufacturer's instructions. The SARS-CoV-2 S1 scanning pool
196 contains 166 peptides consisting of 15-mers, overlapping with 11 amino acids, covering the
197 S1 domain of the spike S1 protein (amino acid 13-685). The peptides were combined into one
198 pool. IFN-γ ELISpot Plus kit was purchased from Mabtech (3420-4HST-10, Sweden).
199 Briefly, the pre-coated wells were plated with unfractionated PBMCs at counts of 300 000
200 cells/well, and the cells were cultured with peptides for the S protein of SARS-CoV-2 at a
201 final concentration of 2 µg/ml (diluted in complete culture medium) for 20 to 22 hrs in a
202 37°C, 5% CO₂ incubator. Cells cultured with medium alone were used as negative controls.
203 Stimulation with anti-CD3 antibody at a concentration of 1 µg/ml was used as a positive
204 control for each subject. Anti-CD28 antibody (3608-1-50, Mabtech, Sweden) was included at
205 a final concentration of 0.1 µg/ml as a co-stimulator. All experiments were conducted in
206 duplicates and results represent the mean of the duplicates. The plates were then processed
207 according to the manufacturer's protocol. Estimation of specific T cell numbers was
208 expressed as spot-forming cells per 1x10⁶ PBMCs (SFC). SFC were counted using an

209 automated reading system (BioSys Bioreader 5000 Pro-F beta, Bio-sys GmbH, Germany) and
210 assessed with the Bioreader 5000 analyser. A stimulation index was calculated by dividing the
211 SFC elicited by a SARS-CoV-2 stimulus by the SFC present in the negative control wells. An
212 increment value was calculated by subtracting the SFC from the negative control wells from
213 the SFC of the stimulated wells. A stimulus was considered to be positive when the
214 stimulation index was >2 , and the increment value was >10 .

215 **Saliva samples**

216 Prior to saliva collection, participants were required to rinse their mouth with water and
217 confirmed they did not show documented oral disease or injury, that they had fasted, refrained
218 from smoking, chewing a gum, taking oral medication, tooth brushing for a minimum of 1
219 hour before sampling and that no dental work had been performed within 24 hours prior to
220 sample collection. Donors were asked to provide a 5 ml sample of saliva in a 50 ml sterile
221 conical tube by passive drooling.

222 All saliva samples were stored/transported on ice upon receipt of the laboratory for processing
223 to preserve sample integrity. Samples were centrifuged (2500g for 20 minutes at 4°C) to
224 pellet cells and insoluble matter. The supernatant was collected and samples were
225 complemented with cOmplete™ protease (#11836170001, Sigma) and Pierce™ phosphatase
226 inhibitor cocktails (#88667, Thermo Scientific), aliquoted and frozen/stored at -80°C on the
227 same day. On the day of the assay, samples were thawed and micro-centrifuged (2500g for 10
228 minutes at 4°C) prior to analysis.

229 **Antibody responses in plasma and saliva using Suspension Multiplex Immunoassay** 230 **(SMIA)**

231 MagPlex-C microspheres (Luminex Corp., Austin, TX, USA) were used for the coupling of
232 antigens according to the manufacturer's protocol as previously described.²⁵ Briefly, 200 µl of

233 the stock microsphere solution (1.25×10^7 beads/ml) were coupled by adding 10 μ g of
234 recombinant SARS-CoV-2 Spike protein RBD His-Tag (#40592-V08B, SinoBiological Inc.,
235 USA). After the coupling, beads were incubated in phosphate buffered saline (PBS: 0.15 M
236 sodium chloride, 10 mM sodium phosphate, pH 7.4) containing 0.05% (v/v) Tween 20 (PBS-
237 T) for 15 min on a rocking shaker at RT. The beads were then washed with 0.5 ml
238 StabilGuard solution (SurModics, Eden Prairie, MN, USA, #SG01-1000) using a magnetic
239 separator (Milliplex® MAG handheld magnetic separation block for 96-well plates, Millipore
240 Corp. Missouri, USA. Cat. #40-285) and resuspended in 400 μ l of StabilGuard solution. The
241 coupled beads were stored at 4°C in the dark until further use.

242 For plasma samples, 50 μ l of plasma diluted 1:1000, and for saliva samples 50 μ l of sample
243 diluted 1:2 in PBS-T containing and 1% (v/v) BSA (Sigma-Aldrich Sweden AB, Stockholm,
244 Sweden, #Sigma-Aldrich-SRE0036) (PBS-T + 1% BSA) was added per well of a flat bottom,
245 96-well μ Clear non-binding microtiter plate (Greiner Bio-One GmbH, Frickenhausen,
246 Germany, #Greiner-655906). Fifty microliters of a vortexed and sonicated antigen-coupled
247 bead mixture suspended in PBS-T + 1% BSA (~50 beads/ μ l) was then added to each well.
248 The plate was incubated in the dark at 600 rpm for 1h at RT. The wells were then washed
249 twice with 100 μ l of PBS using a magnetic plate separator. The beads were resuspended in
250 100 μ l of 1 μ g/ml goat anti-human IgG-PE labelled antibody (Southern BioTech,
251 Birmingham, AL, USA. Cat. #2040-09) in PBS-T + 1% BSA and incubated for 30 min at RT
252 in the dark with rotation at 600 rpm. The beads were subsequently washed twice with PBS,
253 resuspended in 100 μ l of PBS and analysed in a FlexMap 3D® instrument (Luminex
254 Corporation, Austin, TX, USA) according to the manufacturer's instructions. A minimum of
255 100 events for each bead number was set to read and the median value was obtained for the
256 analysis of the data. All sample analyses were repeated three times. A naked, non-antigen-
257 coupled bead was included as a blank along with PBS-T + 1% BSA as a negative control.

258 ***In vitro* stimulation with SARS-CoV-2**

259 PBMC samples from four healthy blood donors, frozen in 2019 in -150°C in foetal bovine
260 serum (FBS) with 10% DMSO, were thawed and added to 10 ml of Gibco Dulbecco's
261 Modified Eagle Medium (DMEM) (Thermo Fisher Scientific, Waltham, US) containing 1%
262 L-glutamine (Cat no: 25030-024, Gibco, Waltham, Massachusetts, USA), 1% penicillin-
263 streptomycin (Cat no: 15140148 Gibco) and 10% normal human serum (NHS) (pooled from 5
264 donors) filtered through a 40 µm strainer and pre-heated to 37°C. The cells were washed two
265 times by centrifugation at 330g for 10 min. The pellet was resuspended in 1.5 ml medium and
266 2 million per donor were seeded in six-well plates and incubated for 16-24 h. The cell culture
267 media were collected, and centrifugated at 330g for 5 min to pellet the non-adherent cells.

268 For *in vitro* infection experiments, SARS-CoV-2 virus previously isolated in a Biosafety level
269 3 lab according to local safety regulations from the nasopharyngeal aspirate of a COVID-19
270 patient (early April 2020) was used.²⁶ The isolated virus was passaged five times in Vero E6
271 cells and for cell infection experiments, freeze-thawed medium supernatants of 4-5 days
272 infected cells or mock supernatants were used. Virus titers were determined using
273 immunoperoxidase assay. In brief, two-day old confluent cells (in a 96-well plate) were first
274 washed with Dulbecco's Modified Eagle's Medium (DMEM) (Gibco, Code: 13345364)
275 containing 100 µg/ml gentamicin, and 100 µl of 10-fold serially diluted SARS-CoV-2 virus
276 lysate was added in quadruplicate. SARS-CoV-2 or mock Vero cell supernatant was added to
277 the PBMC cultures corresponding to a multiplicity of infection of 0.01. 2 hours post infection
278 the cells were washed twice with DMEM and 100 µl of fresh DMEM (containing 2% FBS
279 and 100 µg/ml gentamicin) was added, and the plate was incubated for 8 hours at 37°C in
280 presence of 5% CO₂. After incubation, the supernatant was discarded, and the cells were fixed
281 for 2 hours with 4% formaldehyde. Next, Triton-X (1:500 in phosphate buffered saline, PBS)
282 was added for 15 min, washed once with PBS and incubated for 2 hours at 37°C with PBS

283 containing 3% BSA. Next, the cells were incubated with mouse-anti-dsRNA antibody
284 (Scions, Code: J2 at 1:100 dilution) for 1.5 h followed by detection using horseradish
285 peroxidase–conjugated goat anti-mouse IgG (heavy plus light chain) (Catalog: 1706516, Bio-
286 Rad Laboratories, Hercules, CA, USA) (1:1000) for 1 h. The plates were washed five times
287 with PBS between every incubation, all incubations were done at room temperature and the
288 antibody dilutions were made in PBS containing 1% BSA. Finally, the SARS CoV-2 infected
289 Vero E6 cells were identified using 3-aminoethylcarbazole (AEC) substrate. The spots
290 representing virus-infected cells were counted under the light microscope and the virus lysate
291 was titrated to be 5×10^6 per ml.

292 Cells were monitored in the IncuCyte S3 live cell analysis system (Sartorius, Göttingen,
293 Germany) to allow quantification of cell death in SARS-CoV-2 infected wells versus controls.
294 After 48h incubation the cell culture media was collected from each well and centrifugated at
295 330g for 5 min to collect the non-adherent cells. Lysis buffer (RLT from the AllPrep®
296 DNA/RNA Mini Kit, Qiagen, Hilden, Germany) was added to the wells to lyse adherent cells
297 and the mixture was then added to the pelleted non-adherent cells in order to collect DNA
298 (according to the manufacturer's instructions) from the entire PBMC fraction.

299 **Epigenome-wide DNA methylation analyses**

300 **DNA extraction and quantification**

301 For the performance of epigenome-wide DNA methylation analyses, DNA was extracted
302 from the above isolated PBMCs (approximately 2×10^6 cells) using the AllPrep® DNA/RNA
303 Mini Kit (Cat no: 80204, Qiagen, Hilden, Germany) according to the manufacturer's
304 instructions. Concentrations of extracted DNA were measured using the Qubit® 4.0
305 Fluorometer (Thermo Fisher Scientific, Waltham, Massachusetts, U.S), using dsDNA High

306 Sensitivity (HS) Assay Kit and RNA HS Assay Kit. The measurement was performed
307 according to the manufacturer's instructions.

308 **Illumina MethylationEPIC 850K array**

309 DNA samples were sent to the Bioinformatics and Expression analysis Core facility,
310 Karolinska Institutet, Stockholm, Sweden, where the samples first went through bisulphite
311 conversion on site, followed by the performance of the Illumina Infinium MethylationEPIC
312 850K array. 200 ng of DNA from each sample was analysed.

313 **Statistics**

314 **Descriptive analyses on demographic variables**

315 Initial descriptive analyses of demographic variables were performed on the available
316 information about age, gender, smoking and BMI (kg/m^2). Continuous variables were
317 compared using an unpaired two-tailed t-test and categorical variables were examined using
318 the Pearson χ^2 test or Fisher's exact test (if the number of observations was smaller than five),
319 see Table S1.

320 **DNA methylation analyses**

321 The resulting raw IDAT-files from the MethylationEPIC array analyses were processed in R
322 programming environment (version 4.0.2). The analyses described below were identically
323 performed for the clinical *in vivo* cohort and the *in vitro* experiment, unless stated otherwise.

324 **Pre-processing and quality control**

325 The resulting raw IDAT-files containing the raw DNA methylation profiles for each cell type
326 were analysed in R (version 4.0.2) using the minfi package²⁷ (version 1.36.0) and the data
327 were pre-processed in several steps. The following filters were applied: i) removal of probes

328 with detection p-values above 0.01, ii) removal of non-CpG probes, iii) removal of multi-hit
329 probes, iv) removal of all probes in X and Y chromosomes.

330 *In vivo*

331 We removed the sex chromosomes from our data set, as female X-inactivation skews the
332 distribution of beta values (Figure S1). Of the initial 865 918 probes, 841 524 probes
333 remained upon filtering. After filtering, quality control was performed, and normalisation of
334 the data was done with subset-quantile within array (SWAN) normalisation method.²⁸ The β -
335 values and M-values of the samples were calculated against each probe per sample. The
336 quality of the data was assessed before and after the normalisation (Figure S2). Thereafter, we
337 performed singular value decomposition (SVD) analyses using the ChAMP package²⁹
338 (version 2.19.3) to identify underlying components of variation within the filtered and
339 normalised data set (Figure S3). Significant components consisted of slide, batch and sample
340 groups that contributed to variation within the data set. Corrections were performed for the
341 identified components using ComBat from the SVA package³⁰ (version 3.38.0). As PBMCs
342 consist of multiple nucleated cell types in peripheral blood, we utilised the Houseman method
343 to infer cell type proportions within the samples.³¹ No differences could be determined in cell
344 type proportions between any of the individuals or between sample groups (Table S2),
345 motivating our choice of not correcting for these cell type proportions.

346 *In vitro*

347 In this dataset, we did not have any information on demographic variables, as the samples
348 derived from anonymous donors. However, we still removed the sex chromosomes from our
349 data set, as female X-inactivation skews the distribution of beta values. Of the initial 861 728
350 probes, 837 694 probes remained upon filtering. After filtering, quality control was
351 performed, and normalisation of the data was done with SWAN normalisation method.²⁸ The

352 Houseman method was utilised to infer cell type proportions within the samples³¹, yet again
353 revealing no differences could be determined in cell type proportions between any of the
354 individuals (Table S2), motivating our choice of not correcting for these cell type proportions.
355 The β -values and M-values of the samples were calculated against each probe per sample. The
356 quality of the data was assessed before and after the normalisation (Figure S4). SVA package
357 (version 3.40) was applied to correct the batch effect. Cell deconvolution was performed
358 using FlowSorted.Blood.EPIC package (version 1.11).

359 **Differential DNA methylation analysis**

360 *In vivo*

361 As we were interested in studying CpGs that were differentially methylated between CC19s
362 and non-infected controls from both before and after the start of the COVID-19 pandemic, we
363 performed differential DNA methylation analyses, using the limma package (version 3.46.0).
364 A linear model was fitted to the filtered, normalised and SVD-corrected DNA methylation
365 data. Identified sources of variation that were still present upon SVD correction provided the
366 basis for the inclusion of these variables as co-variates in the models, in this case gender and
367 BMI (Figure S3). For each investigated probe, moderated t-statistics, log₂ Fold Change
368 (logFC) and p-values were computed. The logFC values represent the average beta
369 methylation difference (from hereon referred to as mean methylation difference, MMD)
370 between the CC19s vs. non-infected controls (Cons + Pre20). Differentially methylated CpGs
371 (DMCs) were defined as CpG sites having a nominal p-value of less than 0.01 along with an
372 MMD of > 0.2. As a means to ascertain the quality of the identified DMCs, genomic inflation
373 and pertaining bias were estimated using the BACON package³² (version 1.18.0). As the
374 estimated genomic inflation for the comparison was close to 1 (genomic inflation: 1.20, bias:
375 0.01, Figure S5), this suggested that no major genomic inflation was present in the
376 comparisons, and no correction for this was deemed necessary. The distribution of the DMCs

377 among all investigated DNA methylation sites were illustrated by creating volcano plots
378 (EnhancedVolcano, version 1.8.0). Thereafter, the DMCs were mapped to their corresponding
379 DMGs. DMGs contained at least one DMC, and were considered hyper- or hypomethylated if
380 all DMCs within the gene were hyper- or hypomethylated, respectively. If both hyper- and
381 hypomethylated genes were present in the same gene, the gene was considered having a
382 mixed methylation pattern.

383 ***In vitro***

384 To evaluate the difference between the mock and infected samples, the fold change was
385 calculated using the cut-off obtained from the density plot (M-value $>|2|$; Figure S6) for each
386 CpG site. Only those CpGs with higher values than the cut-off, were selected for further
387 analysis. Venn analysis was performed among the samples using the ggVennDiagram
388 (version 1.1) package in R (version 4.0.3) and bioconductor (version 3.12).

389 **Pathway over-representation analyses**

390 To make biological sense of the putatively SARS-CoV-2-induced DNA methylation
391 differences, we performed PANTHER pathway over-representation test analyses using the
392 PANTHER database (version 16.0). The Fisher's exact test was used for generation of
393 nominal p-values (significance level set to p-value of < 0.05), in case false discovery rate
394 correction was too stringent. The significantly enriched pathways were displayed in dot plots
395 generated in R using ggplot2 package (version 3.3.3).

396 **Network analyses**

397 Network analyses were conducted to generate further and wider biological insight about the
398 DMGs generated in the *in vivo*. An input object was constructed using the pre-2020 (Pre20,
399 n=5) and post-2020 (Con, n=18) non-exposed controls and COVID-19 convalescents (CC19,
400 n=14), as a two-column data frame containing gene annotation and P-value of the significant

401 DMGs (n=54). The graph clustering algorithm MCODE³³ was used to identify molecular
402 complexes and create a large disease module, which was then fitted to a protein-protein
403 interaction network, and both were analysed and rendered in Cytoscape (version 3.8.0). High
404 confidence interactions with a STRINGdb confidence value >0.7 were displayed in the
405 network. Centrality measurements of degree, betweenness and closeness were used to expose
406 the most central nodes in the network. Finally, a functional enrichment of the genes present
407 within the module was carried out using StringDB.³⁴ In addition, the inference of modules
408 was performed with two other methods from the MODifieR package (DIAMOnD and
409 WGCNA)³⁵ to study whether it was possible to condense the module genes to fewer genes of
410 particular interest within the network, for both the *in vivo* and the *in vitro* setting.

411 **Overlap to SARS-CoV-2 interactome**

412 A publicly available protein-protein interaction (PPI) network of SARS-COV-2 and human
413 genes curated by BioGRID (version 4.4.197) was downloaded from the Network Data
414 Exchange in Cytoscape (version 3.8.0). The DMGs from the *in vivo* and *in vitro* setting
415 alongside the gene list from the module generated by MCODE were overlapped onto the PPI
416 network to visualise their respective distributions.

417

418

419

420

421

422

423

424

425

426

427

428

429 **RESULTS**

430 **COVID-19 convalescents display altered DNAm patterns compared to non-infected** 431 **controls**

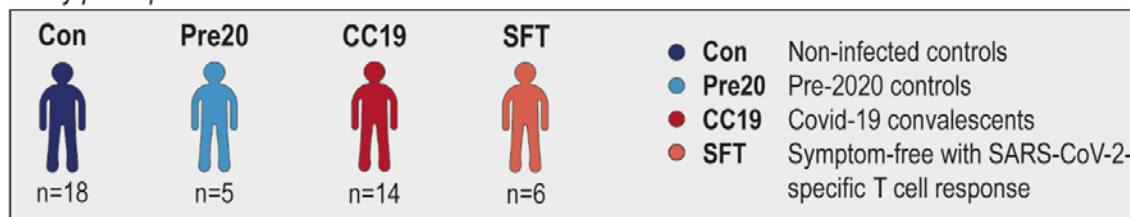
432 We compared epigenome-wide DNAm patterning in peripheral blood mononuclear cells
433 (PBMC) from non-infected controls (Con, n=19), COVID-19 convalescents who had
434 recovered from mild or moderate symptoms (CC19, n=14), donor blood collected before the
435 pandemic (Pre20, n=5) and from asymptomatic individuals presenting with SARS-CoV-2-
436 specific T cell responses (SFT, n=6, Figure 1). Comparisons of demographic variables
437 revealed no significant differences between any of the groups (Table S1). To examine any
438 inherent differences in the DNA methylome between the different sample groups, principal
439 component analyses (PCA) were performed. Three principal components (PC) were identified
440 as both contributing to the variation within the DNAm data and correlating with the sample
441 groups (Figure S7). A three-dimensional illustration of these three most contributing
442 components revealed that the CC19 subjects are distinct from the Con, Pre20 and SFT
443 subjects, whose centroids clustered more closely together (Figure 2a, Figure S8). The
444 observed methylome-wide differences prompted us to identify differentially methylated CpGs
445 (DMCs), which we defined as CpG sites with a nominal p-value of <0.01 along with a mean
446 methylation difference (MMD) of >0.2 . We found 87 DMCs, when comparing the DNA

447 methylomes of CC19s to the merged groups of Cons and Pre20s (Figure 2b, Table S3a). This
 448 identified DMC signature could furthermore distinguish the CC19s from Cons, Pre20s and
 449 SFTs (Figure 2c), suggesting that a past SARS-CoV-2 infection may have resulted in
 450 modulation of the epigenome that persists at least a couple of months after the virus is
 451 eliminated from the body.

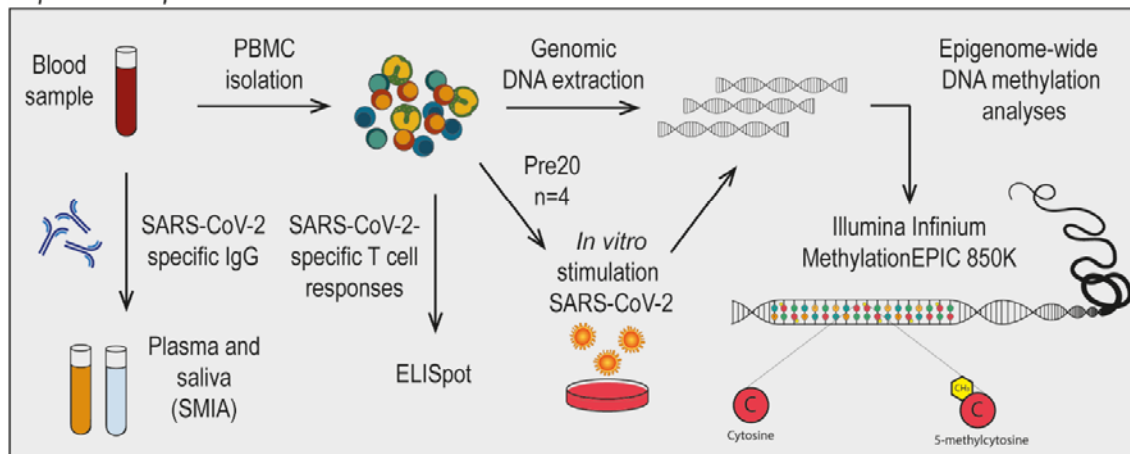
452

453 **Figure 1. Outline of included participants, experimental procedures as well as statistical**
 454 **and bioinformatic approaches utilised in the present study.**

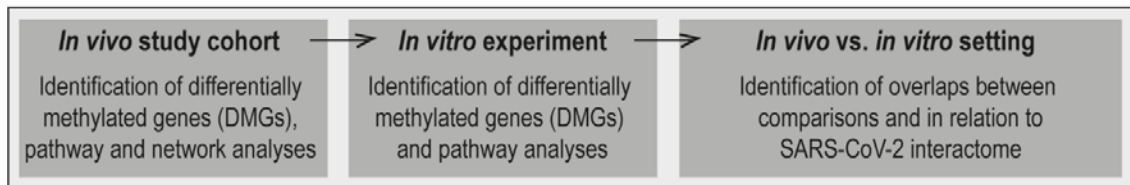
Study participants



Experimental procedure



Statistics and bioinformatics



455

456 **Figure 1.** A brief overview of the present study. CC19 – convalescent COVID-19, Con – non-infected
 457 control, DMG – differentially methylated gene, Pre20 – Pre-2020 non-infected control, SFT –

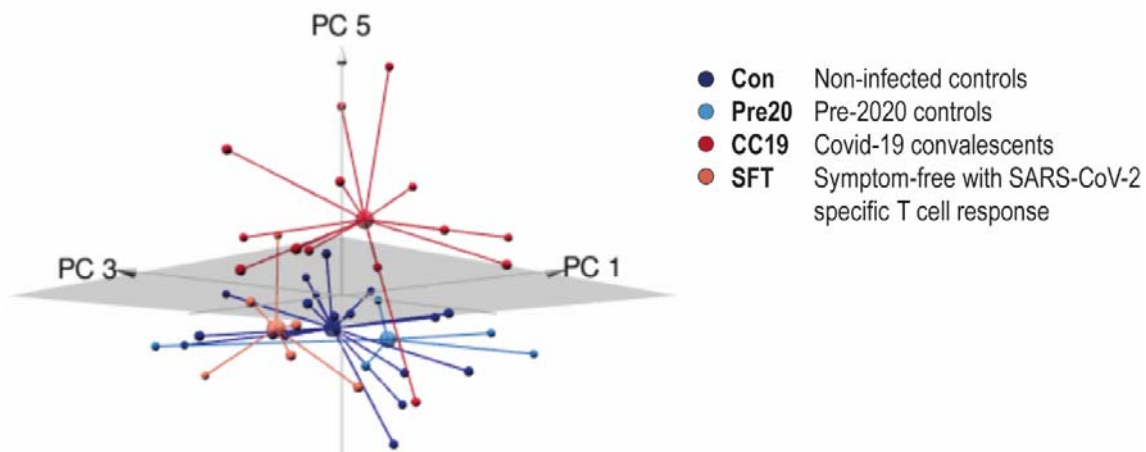
458 symptom-free individuals with SARS-CoV-2-specific T cell response, SMIA – suspension multiplex
459 immunoassay.

460

461 Interestingly, a majority of CC19s showed positive SARS-CoV-2-specific IgG responses both
462 in the circulation and in saliva (Figure 2c). Individuals who showed T cell responses towards
463 SARS-CoV-2 or presented with SARS-CoV-2-specific antibodies in saliva while being
464 negative for antibodies in plasma, aligned with the uninfected controls in the PCA and
465 unsupervised clustering analyses (Figure 2b-c).

466 **Figure 2. Principal component and differential DNAm analysis of PBMC DNA**
467 **methylomes in COVID-19 convalescents and uninfected controls.**

468 **A**

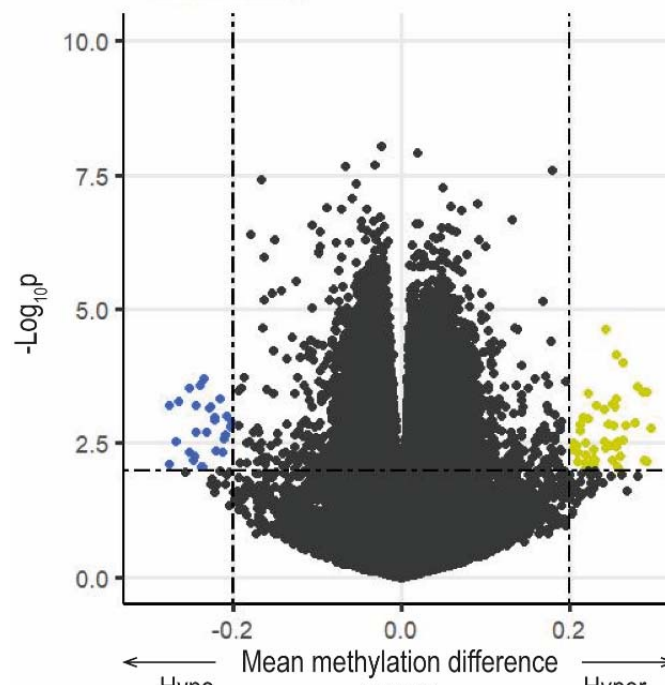


469 **B**

470

471

472

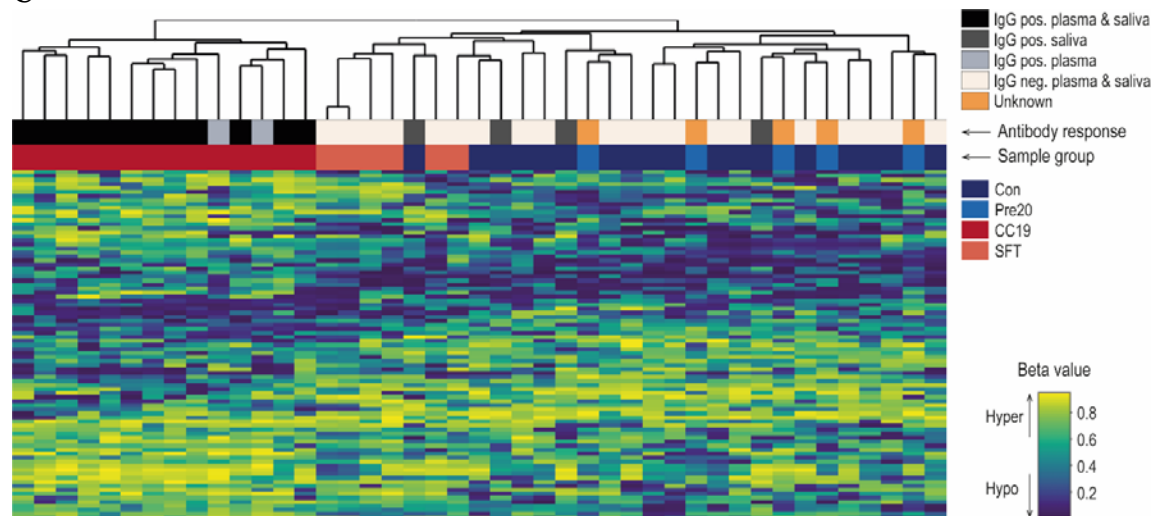


473

474

475

476 C



477

478 **Figure 2.** Upon filtering and normalisation, the DNAm data were subjected to PCA. Panel A shows a
479 3D-PCA plot of principal components (PC)1, PC3 and PC5, where the group means are illustrated as
480 centroids. DMCs were identified comparing CC19s to Cons and Pre20s by computing a linear model
481 on the DNAm data. Panel B illustrates a volcano plot of the CC19 vs. Con + Pre20 DNAm data. The
482 dash-dotted horizontal line represents a nominal p-value cut-off of 0.01, and the vertical lines
483 represent a cut-off in mean methylation difference (MMD) in CC19 vs. Con + Pre20 of $> \pm 0.2$. Panel
484 C shows a heatmap representing an unsupervised hierarchical clustering analysis of individual β
485 values of the 87 identified DMCs in B. The individuals' antibody status is indicated as a grey-scale
486 (unknown = anonymous Pre20 blood donors, orange).

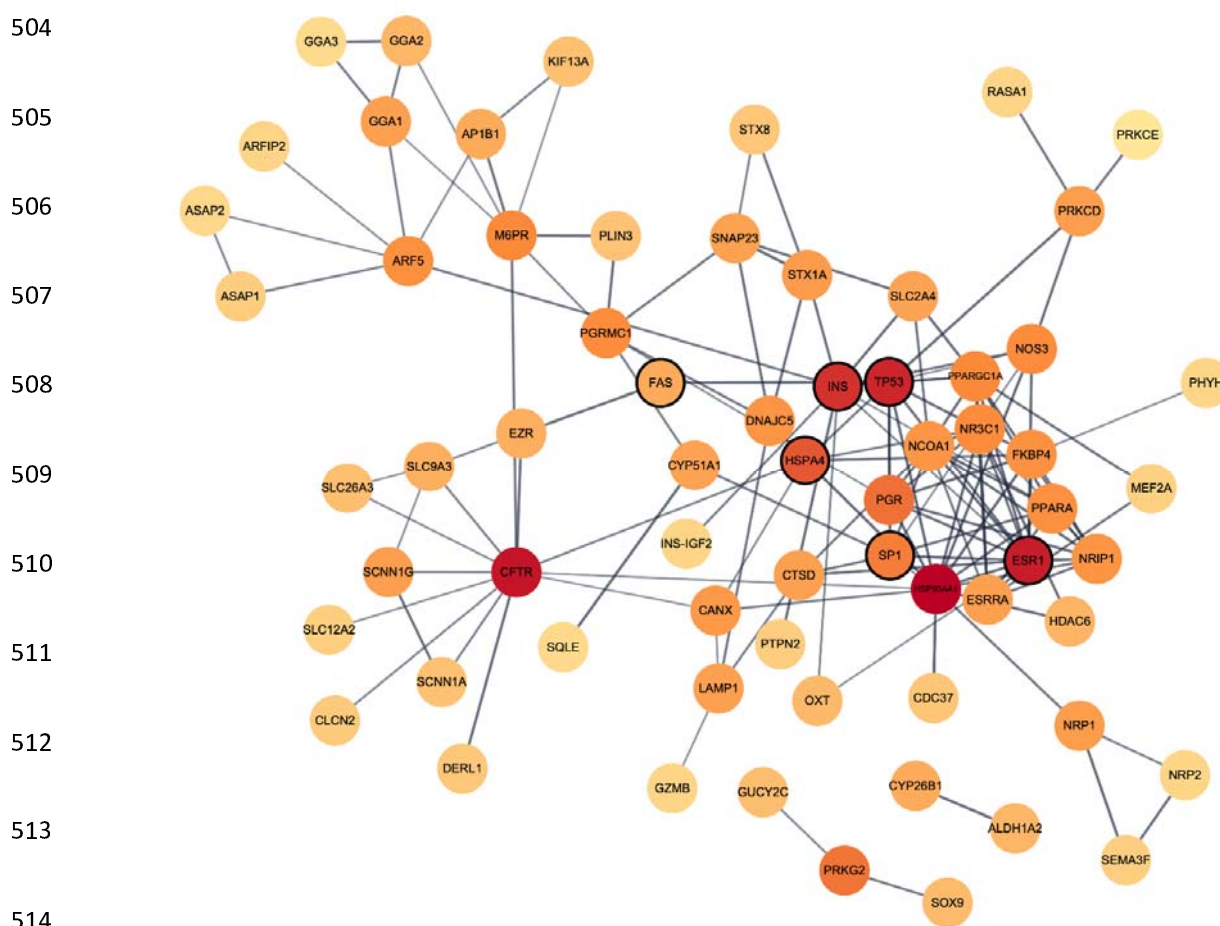
487

488 **Differential methylation of COVID-19 convalescents identifies a putatively SARS-CoV- 489 2-induced module**

490 To further explore the biological impact of SARS-CoV-2 exposure in the CC19 subjects, the
491 identified DMCs were annotated to their respective differentially methylated genes (DMG),
492 resulting in 54 unique genes (Table S3b). Subsequent pathway over-representation analyses
493 revealed involvement in two significantly over-represented pathways (Wnt and integrin
494 signalling pathways, Table S4).

495 As a means to elaborate on the wider interaction context in which the DMGs act with other
496 proteins, the DMGs (n=54) were used as seed genes in the identification of SARS-CoV-2-
497 induced modules in network analyses. The resulting module consisted of 66 genes from the
498 protein-protein interaction network, with 139 interactions, which is significantly more
499 interactions than expected (34 interactions) for a network of that size (Figure 3, Table S5). Six
500 of these genes were present in at least two module identification methods (*INS*, *HSPA4*, *SP1*,
501 *ESR1*, *TP53* and *FAS*), and they were all located in the centre of the module. The four genes

502 **Figure 3. Network illustration of SARS-CoV-2-induced module genes from the *in vivo***
503 **comparison.**



515 **Figure 3.** A network module constructed by means of the graph clustering algorithm MCODE with the 54
516 DMGs from the *in vivo* setting as input. Nodes (n=66) represent genes and connecting lines represent high-
517 confidence protein-protein interactions within the network (STRING combined score > 0.7). Combined ranked

518 scores of centrality quantification of degree, betweenness and closeness is visualised as a colour (light orange to
519 dark red) continuum, with dark red nodes constituting the most central parts of the network. Nodes that were also
520 found both when utilising two other module identifying methods (DIAMOND and WGCNA) and when
521 performing the same analyses on the *in vitro* data set using MCODE are enclosed with a black line.

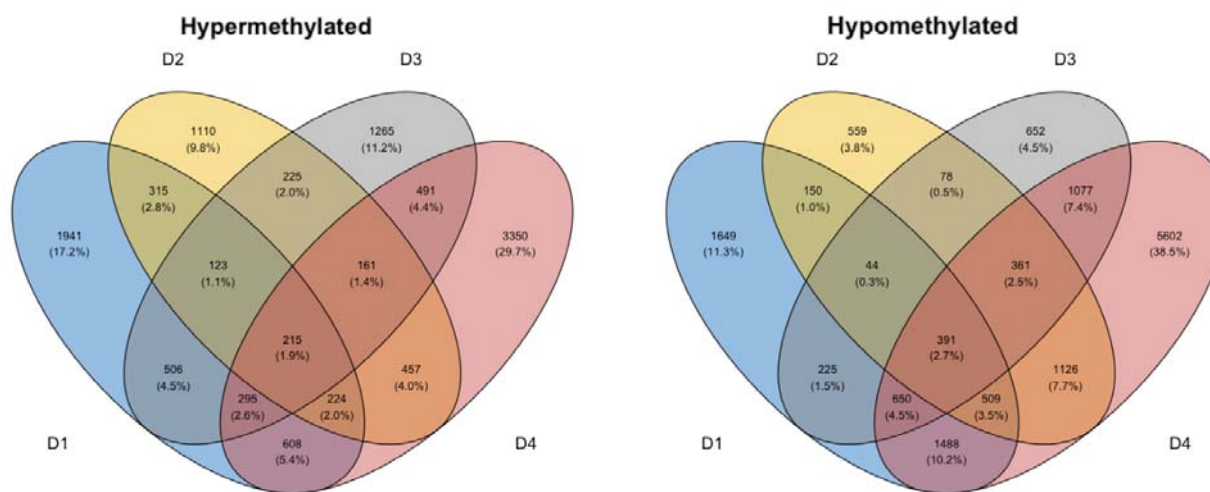
522 with the highest combined centrality scores were *HSP90AA1*, *TP53*, *INS* and *CFTR*. Pathway
523 over-representation analyses of the 66 module genes revealed involvement in pathways such
524 as apoptosis signalling, muscarinic acetylcholine receptor 1 and 3 signalling and
525 gonadotropin-releasing hormone receptor pathway (Figure S9).

526 SARS-CoV-2 stimulated PBMCs *in vitro* reveal overlaps with *in vivo* differential 527 methylation, network analyses and SARS-CoV-2 interactome

528 In the present study, we only had access to self-reported time-after-onset of COVID-19
529 symptoms (Table S6), thus making the immediate effects of SARS-CoV-2 exposure on the
530 epigenome impossible to analyse. Moreover, as the virus-induced DNAm patterns in the
531 CC19's may fade over time, we set out to examine SARS-CoV-2-induced DNAm patterns in
532 an *in vitro* setting. To this end, we exposed PBMCs collected from blood donors in 2019 to
533 SARS-CoV-2 at a low multiplicity of infection for 48h to mimic immediate *in vivo* exposure

534 Figure 4. Differential DNAm analyses of PBMCs stimulated *in vitro* with SARS-CoV-2.

535



536 **Figure 4.** Venn diagrams depicting the overlap of DMCs from the SARS-CoV-2 *in vitro* stimulated PBMCs.
537 Intra-individual comparisons of differential DNAm were performed in treated vs. untreated PBMCs from four
538 different blood donors (D1-D4) collected before the start of the COVID-19 pandemic (2014-2019). DMCs were

539 defined as a fold change in M-value $>|2|$. These DMCs were further mapped to their corresponding annotated
540 genes (DMGs, n=542).

541 to the virus (Figure S10). Exploring the intra-individual DNAm differences between
542 stimulated and unstimulated cells, a set of DMCs (n=3693, Figure 4) were identified to be
543 shared between all four individuals (Table S7a-b). These DMCs mapped to in total 606
544 DMGs (542 unique genes, Table S7c), which were significantly over-represented in a number
545 of pathways including several glutamate receptor pathways, muscarinic acetylcholine receptor
546 1 and 3 signalling pathway, as well as the Wnt and cadherin signalling pathways (Figure S11).

547 As similar pathways were revealed in the findings from the *in vivo* study and the SARS-CoV-
548 2 stimulations, we wanted to explore further similarities in DNAm between the *in vivo* and *in*
549 *vitro* settings. Analyses of the overlap of shared DMGs identified in the two comparisons
550 revealed eight overlapping DMGs (*OR12D3*, *PCSK6*, *INPP5A*, *RAD51B*, *CDH4*, *PHACTR3*,
551 *CDH13*, *SFTA2*). Additionally, to understand the biological context of the genes identified in
552 the *in vitro* comparison, we performed network analyses in the same manner as for the *in vivo*
553 comparison. These analyses found a module consisting of six genes (*TP53*, *INS*, *HSPA4*, *SPI1*,
554 *ESR1* and *FAS*), which were among the previously identified module genes from the *in vivo*
555 setting and were also identical to those that had been identified by more than two module
556 identification methods (Figure 3). Furthermore, explorations of the overlap between identified
557 genes in the differential DNAm analyses and network module analyses to the genes from a
558 publicly available SARS-CoV-2 interactome identified numerous interactions in the *in vivo*
559 (n=11/54), *in vitro* (n=100/542) and network module setting (n=33/66) (Figure S12, Table
560 S8).

561

562

563

564

565

566 **DISCUSSION**

567 The epigenetic events triggered during a mild COVID-19 disease course are largely
568 unexplored, despite the fact that these individuals make up a majority of all SARS-CoV-2-
569 infected individuals. The main finding of our study was an observed DNAm signature that
570 was evident several months after recovery in CC19s compared to non-infected individuals.
571 Although this has not, to our knowledge, previously been described, further investigations
572 need to prove whether this particular signature is a remnant from the time of active infection.
573 Studies of DNA methylomes in circulating cells of COVID-19 patients have so far focused on
574 the hospitalisation phase for moderate-to-severe disease or at discharge³⁶⁻⁴⁰, and none of these
575 studies report comparisons upon convalescence from a mild-to-moderate disease course. Not
576 surprisingly, most of these studies mainly identify engagement of several antiviral immune-
577 related pathways as well as inflammatory responses in severely ill COVID-19 patients
578 compared to controls.^{39, 40} In contrast, our pathway over-representation analyses revealed the
579 involvement of distinct, previously unappreciated pathways such as the Wnt signalling and
580 the muscarinic acetylcholine receptor 1 and 3 signalling pathways. The Wnt signalling
581 pathway has been implicated in several aspects of COVID-19, including development of
582 inflammation, cytokine storms, as well as pulmonary fibrosis.⁴¹ Furthermore, potential viral
583 hijacking of host Wnt targets has been suggested upon SARS-CoV-2 infection in multi-omics
584 studies.⁴² The muscarinic acetylcholine receptor 1 and 3 signalling pathway was present in the
585 module identification analyses from both the natural *in vivo* exposure and the *in vitro*
586 stimulations. Interestingly, in post-viral fatigue patients, including post-SARS-CoV and
587 myalgic encephalomyelitis/chronic fatigue syndrome patients, this signalling pathway is
588 dysfunctional, which has been tentatively attributed to the development of anti-muscarinic

589 receptor autoantibodies.^{43, 44} Although this was not investigated in our study, this could
590 suggest that these pathways found may be implicated in the development of post-acute
591 COVID-19 syndrome, as the effects we observe have persisted for months after the initial
592 exposure to the virus.

593 In the present study, DNA methylome analysis of PBMCs identified a number of genes that
594 were shared between the natural *in vivo* infection and following *in vitro* stimulation, which
595 were further confirmed by several module identification methods. One of these genes was
596 tumour protein 53 (TP53), an evolutionarily conserved protein that is one of the most well-
597 studied hub genes in cell signalling due to its central role in cancer.⁴⁵ TP53 has in several
598 other studies previously been identified as a hub gene, in whole blood from
599 COVID-19 patients⁴⁶, and been shown to interact with ACE2 in SARS-CoV-2-infected
600 human induced pluripotent stem cell-derived cardiomyocytes.⁴⁷ Moreover, transcriptomic
601 analyses of PBMCs from a small group of patients infected with SARS-CoV-2 revealed
602 involvement of apoptosis and p53 signalling pathways⁴⁸, a finding that was further supported
603 by studies of the SARS-CoV-2 interactome, where TP53 was identified as a central player in
604 apoptosis-mediated pathways.⁴⁹ Two additional genes, both members of the heat shock
605 protein family, *HSP90AA1* and *HSPA4* stand out in the network derived from our *in vivo* and
606 *in vitro* data. Interestingly, reports on differentially expressed genes overlapping between
607 acute respiratory distress syndrome and venous thromboembolism datasets identified *TP53*
608 and *HSP90AA1* as central genes, among the top ranked hub genes in their networks.⁵⁰
609 *HSP90AA1* was previously shown to be upregulated in bronchial cells of patients with mild
610 COVID-19 disease, as compared to those with a severe disease course⁵¹, suggesting that this
611 gene may be of particular importance in the mounting of a protective antiviral response.
612 Although our study does not provide any evidence for a protective role of the observed
613 epigenetic alterations, HSP70 family members have been discussed as both anti-viral defence

614 components^{52, 53}, as well as anti-viral drug targets, against SARS-CoV-2.⁵⁴ Altogether, our
615 findings on the network centrality of the hub genes that we derived from the *in vivo* and *in*
616 *vitro* data suggests that they may be of particular importance in the interaction with
617 epigenetically modulated genes upon SARS-CoV-2 infection. Nevertheless, further studies
618 are needed to elucidate the mechanistic role of these genes during infection and recovery from
619 COVID-19.

620 A limitation of the study is the lack of validation of the DNAm findings on a transcriptional
621 level. Hence, whether the observed DNAm patterns are indeed associated or causally linked to
622 host protective or host detrimental immune responses still needs to be addressed in future
623 studies, with more well-designed, larger cohorts and consecutive sample materials from the
624 onset of SARS-CoV-2 infection. The investigation of epigenetic modifications in mild-to-
625 moderately ill COVID-19 patients enabled us to discern DNAm differences that otherwise
626 would have been masked by overriding inflammatory responses. Though these subtle changes
627 may not primarily be relevant to immune response severity towards SARS-CoV-2, they may
628 be insightful for the identification of both effective host protective mechanisms at play, or
629 ensuing deliberating conditions such as long-COVID. The presentation of longstanding
630 symptoms in long-COVID could be attributed to detrimental alterations in DNAm patterns,
631 though originally triggered as a short-term anti-viral response.

632 In conclusion, we found epigenome-wide differences in DNAm patterns of individuals that
633 had recovered from a mild-to-moderate disease course of COVID-19 compared to non-
634 infected controls. This study suggests that DNAm is one of several epigenetic mechanisms
635 that is altered upon SARS-CoV-2 infection, and further investigations should elaborate on
636 whether they are induced by protective host responses or constitute virally-induced hi-jacking
637 processes. Pinpointing these matters will aid the development of efficacious diagnostic tools
638 and treatments of COVID-19 in the future.

639

640

641 **ACKNOWLEDGMENTS**

642 We would like to thank the Bioinformatics and Expression analysis Core facility at
643 Karolinska Institutet for their fruitful collaboration and the performance of the Illumina
644 Infinium MethylationEPIC 850K arrays described in this paper. Additionally, we
645 acknowledge the Swedish National Infrastructure for Computing (SNIC) at National
646 Supercomputing Centre (NSC), Linköping University for the computing systems
647 enabling the data handling, partially funded by the Swedish Research Council (Grant
648 No. 2018-05973).

649 **DECLARATIONS OF INTEREST STATEMENT**

650 M.L., S.Sa. and J.D. have prepared and filed a patent based on the findings from the present
651 study. None of the remaining authors declare any competing interests.

652 **FUNDING**

653 This work was supported by the Swedish Heart and Lung Foundation under grants 20200319,
654 20200067 and 20210067 (M.L.); the Swedish Research Council under grant Covid-19/biobank
655 210202#1 (A.R.) and the Open Medicine foundation under grant OMF190626 (A.R.).

656 **AVAILABILITY OF DATA AND MATERIALS**

657 The datasets used and/or analysed in the presented work will be available upon publication on
658 GeneExpression Omnibus, due to a pending patent application. The datasets comprise filtered
659 and preprocessed DNA methylation data from deidentified individual samples in the study.
660 The dataset will until publication be available using a secure token, provided by the authors
661 upon request. Please, refer to GEO-ID: GSE178962 for further information on the data set.

662 Utilised scripts for performing the described statistical analyses within the paper, as well as
663 for creating graphs, will be available on the following GitHub account upon publication
664 (<https://github.com/Lerm-Lab/Covid19>).

665 **AUTHOR CONTRIBUTIONS**

666 *Conceptualisation:* M.L. – equal, A.R. – equal, *Data curation:* S.Sa. – lead, J.D. – support,
667 J.H. – support, *Formal analysis:* S.Sa. – lead, J.D. – support, L.K. – support, L.P. – support,
668 J.H. – support, *Funding acquisition:* M.L. – equal, A.R. – equal, *Investigation:* E.A. – equal,
669 M.R. – equal (ELISpot/SMIA), M.L. – equal, L.K. – equal, S.Sh. – support (*in vitro* SARS-
670 CoV-2 stimulation of PBMCs), L.K. – lead (DNA extraction PBMC samples), *Methodology:*
671 M.L. – equal, L.K. – equal, A.R. – equal, E.A. – equal, M.R. – equal, *Project administration:*
672 M.L. – equal, A.R. – equal, J.H. – equal, S. Sa. – equal, *Resources:* A.R. – equal, M.L. – equal,
673 *Software:* S.Sa. – lead, J.D. – support, L.K. – support, L.P. – support, *Supervision:* M.L. –
674 equal, A.R. – equal, J.H. – equal, S. Sa. – equal, *Validation:* Not applicable, *Visualisation:*
675 S.Sa. – equal, J.H. – equal, L.K. – support, J.D. – support, L.P. – support, *Writing – original*
676 *draft:* J.H. – lead, all other authors – support, *Writing – review and editing:* J.H. – lead, all
677 other authors – support.

678

679

680

681

682

683

684

685

686

687 REFERENCES

- 688 1. Atlante S, Mongelli A, Barbi V, Martelli F, Farsetti A, Gaetano C. The epigenetic implication in
689 coronavirus infection and therapy. *Clinical Epigenetics*. 2020;12(1):156.
- 690 2. O'Donoghue SII, Schafferhans A, Sikta N, *et al.* SARS-CoV-2 structural coverage map reveals
691 viral protein assembly, mimicry, and hijacking mechanisms. *Molecular systems biology*.
692 2021;17(9).
- 693 3. Zhang Y, Guo R, Kim SH, *et al.* SARS-CoV-2 hijacks folate and one-carbon metabolism for
694 viral replication. *Nature communications*. 2021;12(1):1676.
- 695 4. Singh KK, Chaubey G, Chen JY, Suravajhala P. Decoding SARS-CoV-2 hijacking of host
696 mitochondria in COVID-19 pathogenesis. *American Journal of Physiology-Cell Physiology*.
697 2020;319(2).
- 698 5. Menachery VD, Schäfer A, Burnum-Johnson KE, *et al.* MERS-CoV and H5N1 influenza virus
699 antagonize antigen presentation by altering the epigenetic landscape. *Proceedings of the*
700 *National Academy of Sciences*. 2018;115(5).
- 701 6. Salgado-Albarrán M, Navarro-Delgado EI, Del Moral-Morales A, *et al.* Comparative
702 transcriptome analysis reveals key epigenetic targets in SARS-CoV-2 infection. *NPJ systems*
703 *biology and applications*. 2021;7(1):21.
- 704 7. Oriol-Tordera B, Berdasco M, Llano A, *et al.* Methylation regulation of Antiviral host factors,
705 Interferon Stimulated Genes (ISGs) and T-cell responses associated with natural HIV control.
706 *PLoS pathogens*. 2020;16(8).
- 707 8. Schmidl C, Delacher M, Huehn J, Feuerer M. Epigenetic mechanisms regulating T-cell
708 responses. *The Journal of allergy and clinical immunology*. 2018;142(3):728-43.
- 709 9. Mueller AL, McNamara MS, Sinclair DA. Why does COVID-19 disproportionately affect older
710 people? *Aging*. 2020;12(10):9959-81.
- 711 10. Martin EM, Fry RC. Environmental Influences on the Epigenome: Exposure-Associated DNA
712 Methylation in Human Populations. *Annual Review of Public Health*. 2018;39(1):1-25.
- 713 11. Öst A, Lempradl A, Casas E, *et al.* Paternal diet defines offspring chromatin state and
714 intergenerational obesity. *Cell*. 2014;159(6):1352-64.
- 715 12. Booth A, Reed AB, Ponzio S, *et al.* Population risk factors for severe disease and mortality in
716 COVID-19: A global systematic review and meta-analysis. *PloS one*. 2021;16(3).
- 717 13. Taylor EH, Marson EJ, Elhadi M, *et al.* Factors associated with mortality in patients with
718 COVID-19 admitted to intensive care: a systematic review and meta-analysis. *Anaesthesia*.
719 2021;76(9):1224-32.
- 720 14. Qin W, Scicluna BP, van der Poll T. The Role of Host Cell DNA Methylation in the Immune
721 Response to Bacterial Infection. *Frontiers in immunology*. 2021;12:696280.
- 722 15. Schäfer A, Baric R. Epigenetic Landscape during Coronavirus Infection. *Pathogens*.
723 2017;6(1):8.
- 724 16. Sen R, Garbati M, Bryant K, Lu Y. Epigenetic mechanisms influencing COVID-19. *Genome*.
725 2021;99(999):1-14.
- 726 17. Chlamydas S, Papavassiliou AG, Piperi C. Epigenetic mechanisms regulating COVID-19
727 infection. *Epigenetics*. 2020:1-8.
- 728 18. Das J, Idh N, Pehrson I, Paues J, Lerm M. A DNA methylome biosignature in alveolar
729 macrophages from TB-exposed individuals predicts exposure to mycobacteria. *MedRxiv : the*
730 *preprint server for health sciences*. 2021.
- 731 19. Karlsson L, Das J, Nilsson M, *et al.* A Differential DNA Methylome Signature of Pulmonary
732 Immune Cells from Individuals Converting to Latent Tuberculosis Infection. *MedRxiv : the*
733 *preprint server for health sciences*. 2021.
- 734 20. Pehrson I, Das J, Idh N, *et al.* DNA methylomes derived from alveolar macrophages display
735 distinct patterns in latent tuberculosis - implication for interferon gamma release assay status
736 determination. 2021.
- 737 21. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease
738 2019 (COVID-19) Outbreak in China. *JAMA*. 2020;323(13):1239-42.

- 739 22. Li J, Huang DQ, Zou B, *et al.* Epidemiology of COVID-19: A systematic review and
740 meta-analysis of clinical characteristics, risk factors, and outcomes. *Journal of Medical*
741 *Virology*. 2021;93(3):1449-58.
- 742 23. Chan Y-HH, Fong S-WW, Poh C-MM, *et al.* Asymptomatic COVID-19: disease tolerance with
743 efficient anti-viral immunity against SARS-CoV-2. *EMBO molecular medicine*. 2021;13(6).
- 744 24. Verma D, Parasa VR, Raffetseder J, *et al.* Anti-mycobacterial activity correlates with altered
745 DNA methylation pattern in immune cells from BCG-vaccinated subjects. *Scientific reports*.
746 2017;7(1):12305.
- 747 25. Rizwan M, Rönnerberg B, Cistjakovs M, Lundkvist Å, Pipkorn R, Blomberg J. Serology in the
748 Digital Age: Using Long Synthetic Peptides Created from Nucleic Acid Sequences as Antigens
749 in Microarrays. *Microarrays*. 2016;5(3):22.
- 750 26. Nissen K, Hagbom M, Krambrich J, *et al.* Presymptomatic viral shedding and infective ability of
751 SARS-CoV-2; a case report. *Heliyon*. 2021;7(2).
- 752 27. Aryee MJ, Jaffe AE, Corrada-Bravo H, *et al.* Minfi: a flexible and comprehensive Bioconductor
753 package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics (Oxford,*
754 *England)*. 2014;30(10):1363-9.
- 755 28. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for
756 illumina infinium HumanMethylation450 BeadChips. *Genome biology*. 2012;13(6).
- 757 29. Morris TJ, Butcher LM, Feber A, *et al.* ChAMP: 450k Chip Analysis Methylation Pipeline.
758 *Bioinformatics (Oxford, England)*. 2014;30(3):428-30.
- 759 30. Leek JT, Johnson EW, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch
760 effects and other unwanted variation in high-throughput experiments. *Bioinformatics*.
761 2012;28(6):882-3.
- 762 31. Houseman EA, Accomando WP, Koestler DC, *et al.* DNA methylation arrays as surrogate
763 measures of cell mixture distribution. *BMC bioinformatics*. 2012;13:86.
- 764 32. van Iterson M, van Zwet EW, Consortium B, Heijmans BT. Controlling bias and inflation in
765 epigenome- and transcriptome-wide association studies using the empirical null distribution.
766 *Genome Biology*. 2017;18(1):19.
- 767 33. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein
768 interaction networks. *BMC bioinformatics*. 2003;4:2.
- 769 34. Szklarczyk D, Gable AL, Lyon D, *et al.* STRING v11: protein-protein association networks with
770 increased coverage, supporting functional discovery in genome-wide experimental datasets.
771 *Nucleic acids research*. 2019;47(D1).
- 772 35. de Weerd HA, Badam TV, Martínez-Enguita D, *et al.* MODifieR: an ensemble R package for
773 inference of disease modules from transcriptomics networks. *Bioinformatics (Oxford, England)*.
774 2020;36(12):3918-9.
- 775 36. Zhou S, Zhang J, Xu J, *et al.* An epigenome-wide DNA methylation study of patients with
776 COVID-19. *Annals of human genetics*. 2021.
- 777 37. Balnis J, Madrid A, Hogan KJ, *et al.* Blood DNA methylation and COVID-19 outcomes.
778 *Clinical epigenetics*. 2021;13(1):118.
- 779 38. Castro de Moura M, Davalos V, Planas-Serra L, *et al.* Epigenome-wide association study of
780 COVID-19 severity with respiratory failure. *EBioMedicine*. 2021;66:103339.
- 781 39. Corley MJ, Pang A, Dody K, *et al.* Genome-wide DNA methylation profiling of peripheral
782 blood reveals an epigenetic signature associated with severe COVID-19. *Journal of Leukocyte*
783 *Biology*. 2021.
- 784 40. Konigsberg IR, Barnes B, Campbell M, *et al.* Host methylation predicts SARS-CoV-2 infection
785 and clinical outcome. *Communications Medicine*. 2021;1(1).
- 786 41. Vallée A, Lecarpentier Y, Vallée J-NN. Interplay of Opposing Effects of the WNT/ β -Catenin
787 Pathway and PPAR γ and Implications for SARS-CoV2 Treatment. *Frontiers in immunology*.
788 2021;12:666693.
- 789 42. Zheng F, Zhang S, Churas C, Pratt D, Bahar I, Ideker T. HiDeF: identifying persistent structures
790 in multiscale 'omics data. *Genome biology*. 2021;22(1):21.
- 791 43. Blomberg J, Gottfries C-GG, Elfaitouri A, Rizwan M, Rosén A. Infection Elicited
792 Autoimmunity and Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: An Explanatory
793 Model. *Frontiers in immunology*. 2018;9:229.
- 794 44. Skiba MA, Kruse AC. Autoantibodies as Endogenous Modulators of GPCR Signaling. *Trends in*
795 *pharmacological sciences*. 2021;42(3):135-50.
- 796 45. Levine AJ. p53: 800 million years of evolution and 40 years of discovery. *Nature reviews*
797 *Cancer*. 2020;20(8):471-80.

- 798 46. Vastrad B, Vastrad C, Tengli A. Identification of potential mRNA panels for severe acute
799 respiratory syndrome coronavirus 2 (COVID-19) diagnosis and treatment using microarray
800 dataset and bioinformatics methods. *3 Biotech*. 2020;10(10):422.
- 801 47. Wicik Z, Eyileten C, Jakubik D, *et al*. ACE2 Interaction Networks in COVID-19: A
802 Physiological Framework for Prediction of Outcome in Patients with Cardiovascular Risk
803 Factors. *Journal of clinical medicine*. 2020;9(11).
- 804 48. Xiong Y, Liu Y, Cao L, *et al*. Transcriptomic characteristics of bronchoalveolar lavage fluid and
805 peripheral blood mononuclear cells in COVID-19 patients. *Emerging microbes & infections*.
806 2020;9(1):761-70.
- 807 49. Tiwari R, Mishra AR, Gupta A, Nayak D. Structural similarity-based prediction of host factors
808 associated with SARS-CoV-2 infection and pathogenesis. *Journal of biomolecular structure &
809 dynamics*. 2021:1-12.
- 810 50. Mishra A, Chanchal S, Ashraf MZ. Host-Viral Interactions Revealed among Shared
811 Transcriptomics Signatures of ARDS and Thrombosis: A Clue into COVID-19 Pathogenesis.
812 *TH open : companion journal to thrombosis and haemostasis*. 2020;4(4).
- 813 51. Ma D, Liu S, Hu L, *et al*. Single-cell RNA sequencing identify SDCBP in ACE2-positive
814 bronchial epithelial cells negatively correlates with COVID-19 severity. *Journal of cellular and
815 molecular medicine*. 2021.
- 816 52. Kim MY, Shu Y, Carsillo T, *et al*. hsp70 and a novel axis of type I interferon-dependent
817 antiviral immunity in the measles virus-infected brain. *Journal of virology*. 2013;87(2):998-
818 1009.
- 819 53. Kim MY, Ma Y, Zhang Y, Li J, Shu Y, Oglesbee M. hsp70-dependent antiviral immunity
820 against cytopathic neuronal infection by vesicular stomatitis virus. *Journal of virology*.
821 2013;87(19):10668-78.
- 822 54. Tampere M, Pettke A, Salata C, *et al*. Novel Broad-Spectrum Antiviral Inhibitors Targeting
823 Host Factors Essential for Replication of Pathogenic RNA Viruses. *Viruses*. 2020;12(12).
- 824