



## Article

# Methylation of Host Genes Associated with Coronavirus Infection from Birth to 26 Years

Rutu Rathod <sup>1,†</sup>, Aniruddha Rathod <sup>1,†</sup>, Parnian Kheirkhah Rahimabad <sup>1,†</sup>, Jiasong Duan <sup>1</sup>, Hongmei Zhang <sup>1</sup> , S. Hasan Arshad <sup>2,3,4</sup> and Wilfried Karmaus <sup>1,\*</sup> 

<sup>1</sup> Division of Epidemiology, Biostatistics, and Environmental Health Sciences, School of Public Health, University of Memphis, Memphis, TN 38111, USA; rpatel15@memphis.edu (R.R.); abrathod@memphis.edu (A.R.); Parnian.K@memphis.edu (P.K.R.); jduan@memphis.edu (J.D.); hzhang6@memphis.edu (H.Z.)

<sup>2</sup> Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton SO17 1BJ, UK; S.H.Arshad@soton.ac.uk

<sup>3</sup> David Hide Asthma and Allergy Research Centre, Isle of Wight, Newport PO30 5TG, UK

<sup>4</sup> NIHR Southampton Biomedical Research Centre, University Hospital Southampton, Southampton SO17 1BJ, UK

\* Correspondence: karmaus1@memphis.edu; Tel.: +1-901-678-2491

† Equal contribution.

**Abstract:** DNA methylation (DNAm) patterns over time at 1146 CpGs on coronavirus-related genes were assessed to understand whether the varying differences in susceptibility, symptoms, and the outcomes of the SARS-CoV-2 infection in children and young adults could be explained through epigenetic alterations in a host cell's transcriptional apparatus to coronaviruses. DNAm data from the Isle of Wight birth cohort (IOWBC) at birth, 10, 18, and 26 years of age were included. Linear mixed models with repeated measurements stratified by sex were used to examine temporal patterns, and cluster analysis was performed to identify CpGs following similar patterns. CpGs on autosomes and sex chromosomes were analyzed separately. The association of identified CpGs and expression of their genes were evaluated. Pathway enrichment analyses of the genes was conducted at FDR = 0.05. DNAm at 635 of the 1146 CpGs on autosomes showed statistically significant time effects (FDR = 0.05). The 635 CpGs were classified into five clusters with each representing a unique temporal pattern of DNAm. Of the 29 CpGs on sex chromosomes, DNAm at seven CpGs in males and eight CpGs in females showed time effects (FDR = 0.05). Sex-specific and non-specific associations of DNAm with gene expression were found at 24 and 93 CpGs, respectively. Genes which mapped the 643 CpGs represent 460 biological processes. We suggest that the observed variability in DNAm with advancing age may partially explain differing susceptibility, disease severity, and mortality of coronavirus infections among different age groups.

**Keywords:** coronavirus; DNA methylation; SARS-CoV2; COVID-19; epigenetics; IoW cohort



**Citation:** Rathod, R.; Rathod, A.; Rahimabad, P.K.; Duan, J.; Zhang, H.; Arshad, S.H.; Karmaus, W. Methylation of Host Genes Associated with Coronavirus Infection from Birth to 26 Years. *Genes* **2021**, *12*, 1198. <https://doi.org/10.3390/genes12081198>

Academic Editor: Taichiro Goto

Received: 8 May 2021

Accepted: 29 July 2021

Published: 31 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Coronaviruses are a group of zoonotic pathogens that were previously considered to cause relatively benign infections in humans [1]. However, the emergence of severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002 in China, and the Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012 in Middle Eastern countries, proved coronaviruses' ability to cause severe respiratory disorders [1]. Recently, for the third time, a novel coronavirus called SARS-CoV-2 crossed the interspecies barrier and caused a global health crisis [2].

Epidemiological studies have shown a significant difference between adult and pediatric populations in terms of incidence and symptomatology of SARS-CoV-2 infection. Children seem to be less susceptible to develop the disease and are more likely to present with milder symptoms if infected, while adult subjects are prone to develop severe forms

with higher mortality rates [3]. Data from around the world suggests that COVID-19 susceptibility [4], positivity [4–8], hospitalization [6], and mortality rates [9] increases with age, i.e., individuals at younger age are at lower risk compared to those at older age. Studies on SARS and MERS infections also show a similar trend of milder symptoms and lower mortality rates in children compared to adults [10,11]. Recent studies have identified the role of host factors at a molecular level to explain the gap. Children and adults have shown different distribution and functioning of angiotensin-converting enzyme-2 (ACE2), the receptor coded on the X-chromosome and used by the SARS virus, SARS-CoV-2, and human coronavirus-NL63 (HCoV-NL63) [12]. In addition, cellular and molecular components, of both the innate and adaptive immune systems, and their ability to coordinate an effective immune response deteriorates drastically with age, explaining poorer outcomes in the elderly [13].

Although an individual's genetic information is stable, its epigenetics can change significantly over time. Growing evidence revealed epigenetics, particularly DNA methylation (DNAm) at cytosine-phosphate-guanine (CpG) sites, to be one of the crucial mechanisms underlying the aging process [14]. The importance of epigenetics and DNAm have been lately emphasized in the pathogenesis of several viral infections, including coronaviruses [15–18]. DNAm is an important regulator that alters host expression patterns. These changes have important implications for the activity of the virus itself since it relies on the host cell to replicate its genetic material and continue to proliferate [18–20]. Recently, Corley et al. reported differential DNAm at CpGs associated with ACE2 in distinct age groups, possibly explaining the gap of SARS-CoV2 infection risk in children and adults [21]. Given the dynamic nature of DNAm during aging and its potential role in coronavirus infection, we suggest that changes in DNAm over time could explain the differences in susceptibility, symptoms, and the outcomes of the SARS-CoV-2 infection in children and young adults. In this study, we assessed DNAm levels of CpGs on candidate genes that were assigned to be associated with SARS CoV-2 infections from birth to 26 years using the Isle of Wight (IoW) birth cohort. To differentiate between systematic age-related and random variation of DNAm, we compared the candidate gene DNAm over the years with those in housekeeping and other immune-related genes. We also assessed the association of identified dynamic DNAm with expression of their mapped genes at 26 years. The results of our study may contribute to a better understanding of the underlying mechanisms of differences in hosts' susceptibility and responses to SARS-CoV-2 over time. In addition, epigenetic targets may be identified for the future preventive and therapeutic measures.

## 2. Materials and Methods

### 2.1. Study Population

The Isle of Wight birth cohort (IoWBC) was established to study natural history of allergic disorders in a semi-rural island near the UK mainland [22]. Children born from 1 January 1989 to 28 February 1990 were included in the cohort. From the 1536 pregnancies in this period, 1456 parents with live births consented for recruitment and follow-up at 1, 2, 4, 10, 18 and 26 years. Demographic information was obtained using hospital records at birth and detailed questionnaires provided at each follow-up session. Blood samples were collected for DNAm measurement at birth (Guthrie cards) and at 10, 18, and 26 years (peripheral blood).

### 2.2. Coronavirus-Related Genes and CpGs

Genes potentially related to SARS-CoV-2 pathogenesis were identified from GeneCards ([www.genecards.org/](http://www.genecards.org/) accessed on 28 March 2020) [23] using the keywords 'Coronavirus' and 'Coronavirus silent sweep', i.e., genes associated with silent infection of coronavirus leading to improved immunity resulting in a subclinical disease. The genes were selected for the study based on the score for each gene obtained from GeneCards, i.e., the number of times that a gene was shown to be associated with coronavirus in the literature. In particular, a scree plot of the scores in descending order was implemented and genes with

scores showing large decreases before flattening out were selected. CpGs on these genes were identified using the Illumina manifestation file and were included in the analysis of this study. Additionally, information on the chromosomes that these CpGs are on, their locations relative to CpG islands, and locations on genes were extracted from the manifestation file.

### 2.3. Housekeeping and Immune Genes and Their CpGs

Five housekeeping (HK) genes were randomly identified that are not known to be related to immunity, coronavirus, or silent sweep infection. Additionally, genes potentially related to immunity were identified from GeneCards using the keyword 'immune', and five immune genes were randomly selected that are not known to be related to coronavirus or silent sweep infection. These HK and immune-related genes were selected to compare their DNAm patterns with patterns seen in coronavirus-related genes.

### 2.4. DNA Methylation (DNAm)

DNA was extracted from Guthrie cards (blood collected within 5 days of birth) using a procedure previously described by Beyan et al. [24], and from peripheral blood samples at 10, 18 and 26 years using a standard salting out procedure [25]. DNA concentration was determined by PicoGreen dsDNA quantitation (Molecular Probes, INC. Eugene, OR, USA) or Qubit (ThermoFisher, Waltham, MA, USA). For each sample, about 1 µg DNA was treated with bisulfite to convert cytosine to thymine using the EZ 96-DNA Kit (Zymo Research, Irvine, CA, USA). DNAm levels were assessed using the Infinium HumanMethylation450 BeadChips and MethylationEPIC BeadChips (Illumina, Inc., San Diego, CA, USA) using a standard protocol [26] arrays that were processed with multiple identical control samples allocated to each bisulfite-converted batch to determine assay variability. A BeadStation was used to scan the beadchips. DNAm level, i.e., the  $\beta$  value for each queried CpG locus, was assessed using BeadStudio software (Methylation module).

CPACOR pipeline was used for quality control of the DNAm data [27]. Specifically, the DNA methylation intensity data were quantile-normalized using the R package, *minfi* [28], and ComBat was applied to remove the batch effects [29]. DNAm levels are presented in  $\beta$  values calculated using proportions of intensity of methylated ( $M$ ) over the sum of methylated and unmethylated ( $U$ ) sites ( $\beta = M/[c + M + U]$ , where  $c$  is a constant to prevent division by zero if  $M + U$  is too small). Since  $\beta$  values ranged between 0 and 1, a base-2 logit transformation was applied to  $\beta$  values (denoted as  $M$  values) to avoid severe heteroscedasticity [30]. CpGs with probe-SNPs within ten base pairs and with minor allele frequency (MAF) greater than 0.007 were excluded as they may influence DNA methylation measurements. CpGs on the sex chromosomes were analyzed separately from CpGs on autosomes and were stratified by sex.

Since blood is composed of functionally and developmentally different cell populations [31], we adjusted the cell type proportions to remove the potential confounding effect of cell heterogeneity in DNAm measured from blood samples [32]. Cell type proportions were estimated using the Bioconductor *minfi* package [22], a method proposed by Jaffe and Irizarry [27] and adapted from Houseman et al. [28]. We included the estimated cell type proportions as adjusting factors in the analyses.

### 2.5. Genome-Wide RNA-Seq Gene Expression Data Generation

Peripheral blood samples obtained at 26 years were used for the assessment of gene expression levels using paired-end ( $2 \times 75$  bp) RNA sequencing with the Illumina TruSeq Stranded mRNA Library Preparation Kit with IDT for Illumina Unique Dual Index (UDI) barcode primers according to the manufacturer's instructions. We sequenced all samples twice using the same protocol and combined the output from both runs. The quality of the FASTQ files (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> accessed on 8 May 2021) were assessed by running FASTQC. We mapped the reads against human genome (GRCh37 version 75) using HISAT2 (v2.1.0) aligner [33]. The sequence

alignment map (SAM) format produced the alignment files which were subsequently converted into the Binary Alignment Map (BAM) format using SAMtools (v1.3.1) [34]. The reads mapped to each gene were counted using HTseq (v0.11.1) in the same reference genome used for alignment [35]. We used the countToFPKM package (<https://github.com/AAlhendi1707/countToFPKM> accessed on 8 May 2021) to calculate the normalized read count fragments per kilobase of transcript per million mapped reads (FPKM) which were further log-transformed and used in the data analysis.

### 2.6. Statistical Analysis

Linear mixed models with repeated measurements were implemented to characterize trends in CpG sites, such that DNAm levels changed over time from birth to adulthood. DNAm of CpG sites expressed as M-values at birth, 10, 18, and 26 years were included in the model as the dependent variable and time (with birth as the reference group) and gender (with males as the reference group) as predictors. The estimated cell-type proportions of CD4+ T cells, natural killer cells, neutrophils, B cells, monocytes, and eosinophils were included in the analyses as adjusting factors. CpGs with statistically significant time effects were considered as being dynamic CpGs. Multiple testing was corrected on the overall F tests on time effects by controlling for FDR at 0.05 level. An identical analysis was performed for CpGs of HK and immune-related genes.

Cluster analysis was performed on the regression coefficients at 10, 18, and 26 years to identify CpGs showing similar patterns (trajectories) in DNAm over time using proc fastclus (SAS). Three to ten clusters were assessed, and the highest pseudo-F statistic and cubic clustering criterion were selected to decide on the number of clusters. The profile of each cluster was visualized using cluster medians of regression coefficients at birth, 10, 18, and 26 years to display the group patterns. CpGs on sex chromosome were analyzed separately using the same statistical analysis, stratified by sex. The analysis was performed in SAS version 9.4.

The biologic significance of the identified dynamic CpGs were assessed by evaluating the associations between gene expression of their mapped genes and DNAm at 26 years. Linear regressions were used where gene expression ( $n = 139$ ) was the outcome, and DNAm (in M-values) and sex were the exposure variables. An interaction term of DNAm  $\times$  sex was included in the model as we previously have found the association between gene expression and DNAm to be different in both males and females [36]. Statistical significance of the interaction effects was set at  $p$ -value  $< 0.05$ . If the interaction was not statistically significant, the main effects of DNAm were evaluated after adjusting for sex. For CpGs on sex chromosomes, whose values may vary by the number of X chromosomes, the analysis was stratified by sex.

### 2.7. Pathway Enrichment Analyses

Genes annotated to the dynamic CpGs were identified based on the Illumina manifest file. Pathway enrichment analysis of the genes was conducted using Toppfun (<https://toppgene.cchmc.org/enrichment.jsp> accessed on 24 July 2021). Additionally, we performed two more pathway analysis using candidate genes and non-dynamic candidate genes (i.e., genes with CpGs not showing dynamic patterns) to compare biological processes with dynamic genes with whole genome as a background. Multiple testing was adjusted by controlling Bonferroni  $p$ -value of 0.05.

## 3. Results

From GeneCards, genes with a score of  $>2.8$  (the top 66 genes) were selected based on a scree plot for coronavirus and 14 genes for coronavirus silent sweep infection. In total, 1555 CpGs located on these 78 genes were extracted from the Illumina manifestation file. Of these, four CpGs with 'ch' prefix were excluded, and DNAm of the remaining 1551 CpGs at different time points (i.e., birth, 10, 18, 26 years) were analyzed. Of the 1551 CpGs, 29 CpGs

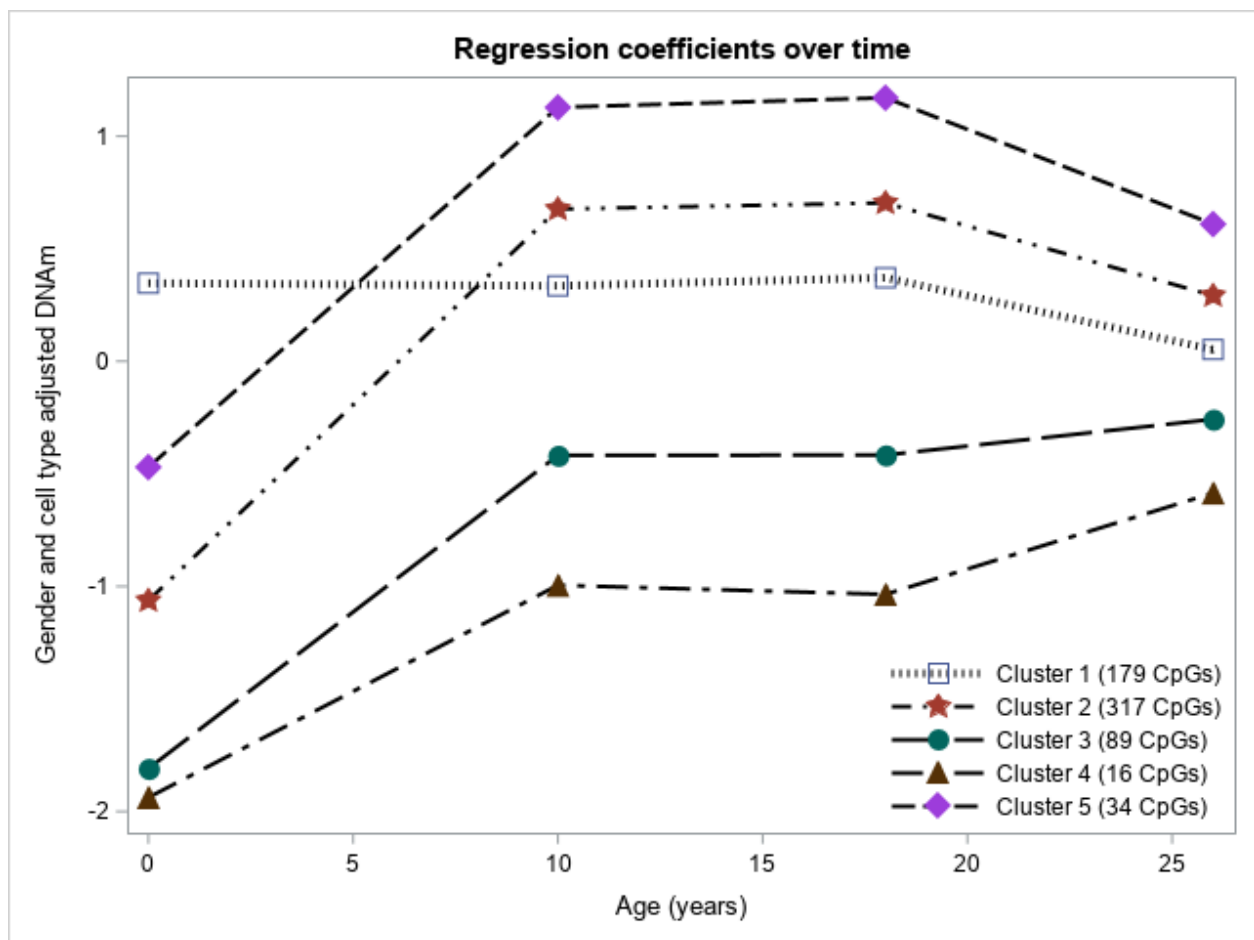
were found to be located on sex chromosome on gene *ACE2* and *CD40LG*, and the analysis was stratified by gender for those CpGs.

Of 1522 CpGs on autosomes, DNAm data was available for 988 CpGs at birth ( $n = 796$ ), 688 CpGs at 10 years ( $n = 330$ ), 688 CpGs at 18 years ( $n = 476$ ), and 677 CpGs at 26 years ( $n = 242$ ), in the IoW cohort (in total, 1146 CpGs). Identified through use of linear mixed models, DNAm at 635 of the 1146 CpGs (55.4%) showed statistically significant time effects at the FDR = 0.05 level (Table 1 and Supplement Table S1a). We observed both linear and non-linear time effects on DNAm. The pattern of these CpGs/genes was called dynamic. For 100 identified CpGs on five randomly selected HK genes, DNAm data was available for 47 CpGs at birth, 40 CpGs at 10 years, 40 CpGs at 18 years, and 33 CpGs at 26 years (58 CpGs in total). DNAm at only one of the 58 CpGs showed significant time effects at the FDR = 0.05 level (1.72%, Supplement Table S1b). For 125 identified CpGs on five randomly selected immune genes, DNAm data were available for 69 CpGs at birth, 41 CpGs at 10 years, 41 CpGs at 18 years, and 39 CpGs at 26 years (76 CpGs in total). DNAm at 11 of the 76 CpGs showed significant time effects at the FDR = 0.05 level (14.47%, Supplement Table S1c, Supplement Figure S1).

**Table 1.** DNAm for CpGs on autosomes showing significant time effects (dynamic) along with the information on locations of the identified CpG sites. Only results for top 10 CpGs showing the most statistically significant time effects at FDR = 0.05 level was shown.

CpG Site	F Value	FDR $p$ -Value	Gene	Chromosome Number	CpG Islands	Gene Location
cg04791421	1218.77	$4.15 \times 10^{-308}$	<i>CCL5</i>	17		TSS1500
cg21278129	1176.06	$9.26 \times 10^{-303}$	<i>CLEC4G</i>	19	S_Shore	TSS1500
cg11694510	1127.39	$1.85 \times 10^{-296}$	<i>IFITM1</i>	11	S_Shore	TSS1500
cg03589230	1031.87	$2.2 \times 10^{-283}$	<i>MYOM2</i>	8		TSS1500
cg21686213	1000.14	$7.48 \times 10^{-279}$	<i>IFITM1</i>	11	N_Shore	3'UTR
cg20545776	972.67	$7.67 \times 10^{-275}$	<i>GPT</i>	8	N_Shore	TSS1500
cg12876900	957.15	$1.51 \times 10^{-272}$	<i>IFITM3</i>	11	S_Shore	1st Exon
cg16628205	922.08	$3.25 \times 10^{-267}$	<i>TFR2</i>	7		TSS1500
cg14231966	914.28	$4.97 \times 10^{-266}$	<i>FURIN</i>	15	Island	Body
cg03038262	907.51	$5.37 \times 10^{-265}$	<i>IFITM1</i>	11	N_Shore	3'UTR

For each CpG site, results from linear mixed models enabled us to estimate DNAm at each time point with gender and cell type heterogeneity adjusted. These adjusted DNAm were then used to cluster the CpG sites to reveal different temporal DNAm patterns. Based on the pseudo-F statistic and cubic clustering criterion, we grouped the 635 CpG sites into five clusters (Figure 1). In clusters 2 (317 CpGs) and 5 (34 CpGs), DNAm on average showed similar patterns over time (increase from birth to 10 years, minimal changes from 10 to 18 years, and slight decrease from 18 to 26 years), but DNAm at CpGs in cluster 5 tends to be higher than that in cluster 2. Average DNAm patterns in clusters 3 (89 CpGs) and 4 (16 CpGs) also showed similar patterns over time (increase from birth to 10 years, minimal changes from 10 to 18 years, and an increase from 18 to 26 years) with cluster 3 having a higher DNAm on average. The temporal pattern of average DNAm in the CpG cluster 1 is unique in that it is stable overall with a slight decrease from 18 to 26 years. In terms of average DNAm across all the five clusters, DNAm levels were lower on average for CpGs in clusters 3 and 4 compared to CpGs in clusters 2 and 5, while average DNAm in cluster 1 was in the middle starting from age 10 but the highest at birth (Figure 1).



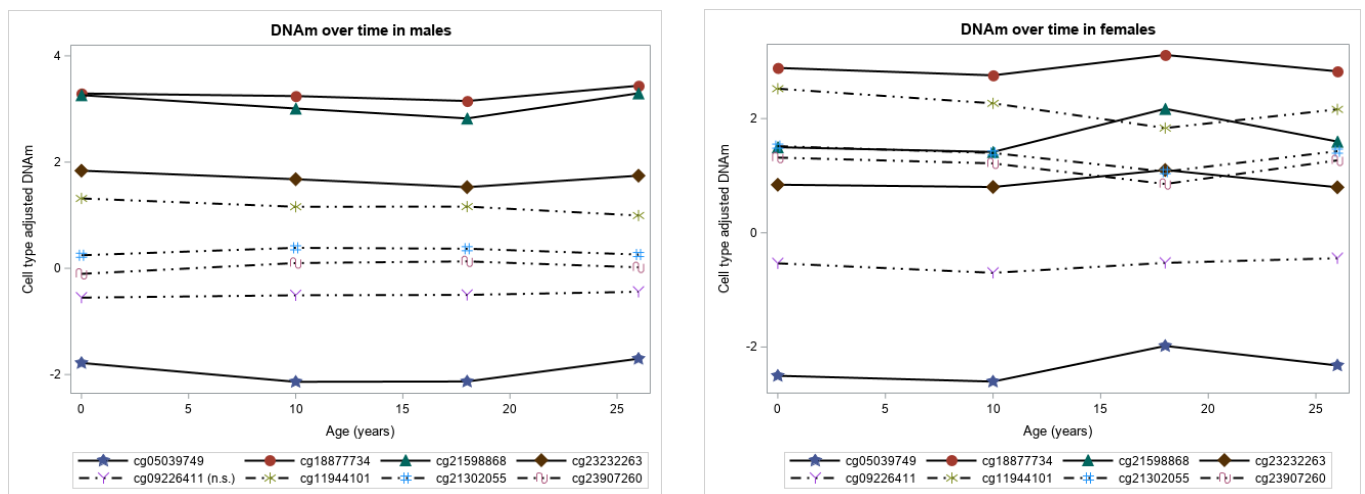
**Figure 1.** DNA methylation profiles of the five CpG clusters represented by the median of gender and cell type heterogeneity adjusted DNA methylation at each age.

Of the 29 CpGs on sex chromosome, DNAm data were available for eight CpGs at birth, 10, 18, and 26 years ( $n = 506$  males, 506 females). In linear mixed models, statistically significant time effects were observed at seven CpGs in males and eight CpGs in females after adjusting for multiple testing at the FDR = 0.05 level (Table 2, Figure 2).

The association between the identified dynamic CpGs and expression of their mapped genes were assessed. Significant effects for the interaction of DNAm and sex were observed at 24 CpGs with 15 genes on autosomes (all  $p$ -values  $< 4.48 \times 10^{-2}$ , Table 3 and Supplement Table S2a). Of the 24 CpGs, the estimates for both sexes were in opposite direction at 22 CpGs (i.e., opposite signs of regression coefficients after combining the main and interaction effects), thereby suggesting a potential gender reversal at these specific CpG sites. For example, with one unit increase in DNAm levels of cg21657705, the expression of *ACE* gene is downregulated by 0.64 units in males while it upregulates *ACE* expression by 0.45 units ( $=1.09-0.64$ ) in females (Table 3 and Supplement Table S2a). At 16 of the 22 CpGs, an increase in DNAm was associated with decreased expression in females while there were increased gene expression levels in males. Whereas, at 6 of the 22 CpGs, an increase in DNAm was associated with increased expression in females but decreased gene expression in males. For CpGs without interaction effects, main effects were assessed and significant association of 93 CpGs with 31 genes were found after adjusting for sex (all  $p$ -values  $< 5 \times 10^{-2}$ , Table 4 and Supplement Table S2b). The models evaluating the association of DNAm and gene expression on sex chromosomes were stratified by sex. An increase in DNAm at CpG site cg23907260 was associated with increased gene expression levels of *CD40LG* in males (estimate: 0.55,  $p$ -value:  $8.40 \times 10^{-3}$ , Supplement Table S2c).

**Table 2.** DNAm for CpGs on sex chromosomes showing significant time effects (dynamic) along with the information on locations of the identified CpG sites. The analysis was stratified by sex.

CpG Site	F Value	FDR <i>p</i> -Value	Gene	Gene Location	Sex
cg21598868	63.77	$4.64 \times 10^{-35}$	<i>ACE2</i>	TSS1500	Females
cg05039749	44.64	$1.31 \times 10^{-25}$	<i>ACE2</i>	Body	Females
cg23232263	38.77	$1.33 \times 10^{-22}$	<i>ACE2</i>	3'UTR	Females
cg18877734	22.13	$1.78 \times 10^{-13}$	<i>ACE2</i>	TSS1500	Females
cg11944101	97.5	$1.22 \times 10^{-49}$	<i>CD40LG</i>		Females
cg23907260	91.36	$2.56 \times 10^{-47}$	<i>CD40LG</i>	TSS200	Females
cg21302055	77.44	$2.53 \times 10^{-41}$	<i>CD40LG</i>	TSS200	Females
cg09226411	18.16	$3.01 \times 10^{-11}$	<i>CD40LG</i>	TSS200	Females
cg21598868	22.59	$3.78 \times 10^{-13}$	<i>ACE2</i>	TSS1500	Males
cg05039749	26.73	$3.58 \times 10^{-15}$	<i>ACE2</i>	Body	Males
cg23232263	13.81	$2.26 \times 10^{-8}$	<i>ACE2</i>	3'UTR	Males
cg18877734	7.99	$3.72 \times 10^{-5}$	<i>ACE2</i>	TSS1500	Males
cg11944101	19.18	$2.23 \times 10^{-11}$	<i>CD40LG</i>		Males
cg23907260	11.63	$3.53 \times 10^{-7}$	<i>CD40LG</i>	TSS200	Males
cg21302055	8.14	$3.52 \times 10^{-5}$	<i>CD40LG</i>	TSS200	Males
cg09226411	1.9	0.13	<i>CD40LG</i>	TSS200	Males

**Figure 2.** DNA methylation patterns of the available CpGs on sex chromosomes in males and females represented by cell type adjusted DNA methylation over time. The solid lines represent temporal patterns of CpGs on gene *ACE2* and dash-dotted lines are for CpGs on gene *CD40LG*.

The 643 identified dynamic CpGs (showing statistically significant time effects) were mapped to 60 genes (referred to as dynamic genes). To better understand the biological function of these 643 CpGs, pathway enrichment analyses was conducted for the 60 dynamic genes, 78 candidate genes, and 18 (=78–60) non-dynamic candidate genes. Using these genes in *ToppFun*, with the whole genome as the background, we identified 460, 524, and 38 biological processes for dynamic, candidate, and non-dynamic candidate genes, respectively, which were enriched after multiple testing adjusted by controlling for Bonferroni *p*-value of 0.05 (Table 5 and Supplement Table S3).

**Table 3.** Sex-specific associations of DNAm with expression of their mapped genes. Five CpGs with the most statistically significant interaction effects of DNAm  $\times$  sex on gene expression were shown. Males are in the reference group. The  $p$ -values are for interaction effects. The full list of CpGs showing significant interaction effects (24 CpGs mapped to 15 genes) are in Supplement Table S2a.

CpG	Gene	DNAm Effect	Sex $\times$ DNAm Interaction Effect	$p$ -Value (Sex $\times$ DNAm)	Gene Location
cg12455187	<i>CCL5</i>	1.25	−1.38	$7.00 \times 10^{-4}$	TSS1500
cg20559158	<i>MYOM2</i>	1.81	−2.94	$8.42 \times 10^{-3}$	Body
cg00162643	<i>DDX58</i>	−0.68	0.83	$7.07 \times 10^{-3}$	TSS200
cg15096505	<i>IL10</i>	−3.17	2.30	$3.11 \times 10^{-3}$	Body
cg21657705	<i>ACE</i>	−0.64	1.09	$2.82 \times 10^{-3}$	Body

**Table 4.** Sex-unspecific associations of DNAm with expression of their mapped genes. Five CpGs with the most statistically significant main effects (sex unspecific) of DNAm on gene expression are shown. Males are in the reference group. The full list of CpGs showing significant interaction effects (93 CpGs mapped to 31 genes) are in Supplement Table S2b.

CpG	Gene	DNAm Effect	$p$ -Value	Gene Location
cg10315334	<i>CCL5</i>	−1.79	$2.98 \times 10^{-17}$	1st Exon, 5'UTR
cg02867514	<i>CCL5</i>	−1.62	$6.88 \times 10^{-15}$	1st Exon, 5'UTR
cg08656816	<i>CCL5</i>	−1.72	$6.03 \times 10^{-15}$	TSS200
cg15055101	<i>SH2D3A</i>	−1.31	$3.69 \times 10^{-14}$	5'UTR
cg15353603	<i>DPP4</i>	−0.90	$5.01 \times 10^{-10}$	Body

**Table 5.** Top 10 most significant biological processes from pathway enrichment analysis of dynamic and non-dynamic candidate genes.

Dynamic Genes		Non-Dynamic Candidate Genes	
Biological Processes	Bonferroni $p$ -Value	Biological Processes	Bonferroni $p$ -Value
<b>immune effector process</b>	$1.17 \times 10^{-24}$	response to cytokine	$1.85 \times 10^{-5}$
defense response	$1.34 \times 10^{-24}$	defense response	$5.74 \times 10^{-5}$
response to biotic stimulus	$9.40 \times 10^{-24}$	cell activation	$1.66 \times 10^{-4}$
response to cytokine	$1.45 \times 10^{-20}$	cellular response to cytokine stimulus	$1.81 \times 10^{-4}$
response to other organism	$2.41 \times 10^{-20}$	entry into host	$2.51 \times 10^{-4}$
response to external biotic stimulus	$2.51 \times 10^{-20}$	response to other organism	$2.57 \times 10^{-4}$
cellular response to cytokine stimulus	$4.59 \times 10^{-19}$	response to external biotic stimulus	$2.60 \times 10^{-4}$
cytokine-mediated signaling pathway	$1.63 \times 10^{-18}$	response to biotic stimulus	$3.22 \times 10^{-4}$
<b>defense response to other organism</b>	$3.82 \times 10^{-18}$	movement in host environment	$4.83 \times 10^{-4}$
viral process	$9.63 \times 10^{-17}$	viral life cycle	$4.85 \times 10^{-4}$

Note: Two biological processes in bold font are unique to dynamic genes, i.e., were not obtained in pathways of non-dynamic candidate genes.



#### 4. Discussion

We examined the development of methylation levels of CpGs potentially associated with coronavirus infection using four measurements spanning from birth to age 26 years. DNAm levels at 635 CpGs on autosomes showed significant time effects. For most of the CpGs on autosomes, increasing age was associated with a rise in DNAm levels from birth to pre-adolescence period, no change in DNAm levels from pre- to post-adolescence, and a decrease in DNAm levels from post-adolescence to adulthood. About 55.4% (=635/1146) of CpGs on coronavirus-related genes were identified as dynamic compared to 1.72% (=1/58) in random samples of HK genes and 14.47% (=11/76) of immune-related genes. In addition, the temporal patterns in DNAm were consistent across all the identified dynamic CpGs on immune-related genes, which were different from the patterns revealed by the dynamic CpGs on coronavirus-related genes. These findings suggest that the observed variability in DNAm levels with advancing age may, in part, explain differing susceptibility, disease severity, and mortality of coronavirus infections among distinct age groups. More specifically, lower DNAm levels at most of CpGs from birth to pre-adolescence compared to other ages may provide protection against SARS-CoV-2 infection. An increase in DNAm levels at pre-adolescence and no change from pre- to post-adolescence might explain the increase in the number of coronavirus cases from 10 to 18 years compared to earlier ages. From post-adolescence to adulthood, DNAm levels decrease at some CpGs and increase at some CpGs, which may explain high susceptibility, morbidity, and mortality among adults compared to children. The sex chromosome, DNAm, at seven CpGs (four on *ACE2* and three on *CD40LG*) in males and eight CpGs (four on *ACE2* and four on *CD40LG*) in females showed significant time effects.

In this study, we observed DNAm levels of CpGs on the *ACE2* gene to be higher in males compared to females, while on *CD40LG* gene, DNAm levels were higher in females compared to males from birth to adulthood. It has been shown previously that *CD40LG* (CD40 ligand) on the X chromosome is involved in response to infections and escapes X inactivation in some cells, contributing to gender differences in immune responses [37]. In addition, the identified CpGs on autosomes and sex chromosomes suggest that there is a possibility of epigenetic regulation on gene activities, and that the observed sex-specific associations of gene expression and DNAm at some CpGs may be linked to the observed gender gap in incidence and mortality of coronaviruses, although in-depth assessment is certainly needed.

The *CCL5* gene on autosome drew most of our attention. One of the CpG site on this gene showed the most statistically significant change over time (with  $p$ -value at the level of  $10^{-308}$ ). In addition, DNAm at multiple CpGs on this gene was shown to be associated with expression of *CCL5*. In particular, three CpGs (cg10315334, cg02867514, cg08656816) have the strongest biological relevance among all the CpGs identified, reflected by their largest effects on gene expression as well as the smallest  $p$ -values, regardless of sex. The *CCL5* gene has been previously linked to the susceptibility and pathogenesis of SARS-CoV [38,39] and SARS-CoV-2 [40,41].

Recently, Corley et al. showed DNAm levels of CpGs associated with *ACE2* in airway epithelial cells to be age-dependent, potentially explaining the host differences in children and adults [21]. Our findings indicated that age-dependency of DNAm at CpGs on this gene was also present in blood. However, since *ACE2* is not expressed in blood [42–45], we were not able to assess the biological relevance of the identified CpGs.

The pathways identified using dynamic candidate and non-dynamic candidate genes in the pathway analysis yielded a different number of biological processes. Of the 460 biological processes identified based on dynamic genes, 429 were unique compared to the processes identified based on non-dynamic genes, indicating different underlying biological functionalities between these two sets of genes. Pathway enrichment analysis revealed biological processes involved in host immune function (e.g., immune effector process, cellular response to cytokine stimulus, cytokine-mediated signaling pathway, response to biotic stimulus, cytokine, other organism, external biotic stimulus,) and the response to

viruses (e.g., viral process, defense response to virus and other organism). These biological processes have been suggested to have critical roles in the pathogenesis of coronavirus infections [46–48]. Lymph nodes maintain and coordinate new immune responses to control the viruses, although age-related lymph nodes changes reduce the ability of B and T cells to proliferate and differentiate in lymph nodes. Because of this reason, new immune responses are blunted, with significantly less effector cells that are less well prepared by antimicrobial molecules. It is assumed that this mechanism in older adults renders them less effective in defending against SARS-CoV-2 infection [13]. Diao. et al. demonstrated a reduction and functional exhaustion of surviving T cells in COVID-19 patients [49]. Lymphocytes are significantly reduced in the SARS-CoV-2 infected subjects which is directly affected by the viral load [50], possibly due to SARS-CoV-2-induced activation of apoptosis [51]. It was shown that there is an association of COVID-19 pathogenesis and excessive cytokine release [51]. Issa et al. identified six functional domains (I to VI) in the SARS-CoV-2 3a protein that were linked to virulence, infectivity, ion channel formation, and virus release in SARS-CoV-2 [52]. In addition, immune response differences in children compared to adults includes lower production of proinflammatory cytokines, higher production of immunomodulatory cytokines, decreased infiltration of neutrophils, and a predominance of CD4+ T cells [53]. These pathophysiological differences in children and adults are believed to underlie lower susceptibility of children to coronavirus infections and their diminished immune mediated lung injury [53].

The strength of this study is the availability of DNAm from birth to 26 years, enabling us to examine changes from birth to adulthood for CpGs on autosomes and sex chromosome separately. To our knowledge, this is the first study to examine the epigenetics of genes potentially associated with the coronavirus to explain differences in susceptibility, morbidity, and mortality among children and adults. Our effort was to try to explain the susceptibility, morbidity, and mortality through DNAm levels in children and adults. However, we were unable to directly assess changes in DNAm among subjects infected with coronavirus due to the lack of COVID-19 data. It would also be preferable to look at the individual DNAm patterns (trajectories), including older adults; however, to our knowledge, no cohort currently has DNAm data covering birth, childhood, and early and late adulthood. In our study, DNA was extracted from peripheral blood cells. Coronaviruses affect many different cell types, primarily cells of the respiratory tract [54]. In our analysis, we extracted candidate genes from the literature, potentially related to SARS-CoV-2 infection, and did not have information on tissues or cells which these genes were identified in. It has been shown that DNAm of blood cells has concordance with that of the respiratory system cells [55], although some differences exist between the two. Almost all coronavirus genes included in the study are immune-related genes; however, not all immune-related genes are coronavirus-related genes. Future studies linking gene with alterations of methylation are needed to directly assess association of DNAm with coronavirus infections in older adults and elderly populations. Large cohorts with DNAm data before and after the recent coronavirus pandemic can be used to (a) estimate risks related to differences in DNAm of the exposed subjects who did and did not develop the infection, (b) estimate differences in DNAm of the patients who presented with mild, moderate, or severe symptoms, and (c) demonstrate the effects of the infection on the epigenome of the host. Our findings showed that changes in DNAm levels from birth to adulthood in some genes might explain COVID-19 susceptibility and severity differences in children and young adults. Thus, our results are informative for the COVID-19 disease map [56], as it provides details on possible involvement in changing susceptibility to the coronavirus based on DNAm of various genes.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/genes12081198/s1>, Table S1a: DNAm for CpGs on autosomes of coronavirus-related genes showing significant time effects along with the information on locations of the identified CpG sites, Table S1b: DNAm for CpGs on housekeeping genes showing significant time effects along with the information on locations of the identified CpG sites, Table S1c: DNAm for CpGs on immune-related

genes showing significant time effects along with the information on locations of the identified CpG sites, Table S2a: Association of DNAm at 24 CpGs with their mapped genes' expression levels on autosomes that are sex-specific. Only results on CpGs showing statistically significant interaction effects of DNAm  $\times$  sex on gene expression were shown. Males are in the reference group. The *p*-values are for interaction effects, Table S2b: Association of DNAm at 93 CpGs with their mapped genes' expression levels on autosomes that are sex-nonspecific. Only results on CpGs showing statistically significant DNAm effects on gene expression were shown. Males are in the reference group, Table S2c: Association of DNAm with their mapped genes' expression levels on sex chromosomes, Table S3: Significant GO terms and its biological processes from pathway enrichment analysis of genes annotated to the dynamic CpGs, Figure S1: DNA methylation (DNAm) levels of the 11 CpGs on immune-related genes at each age demonstrates consistent temporal patterns at all the 11 CpGs (i.e., parallel to each other) except for cg17498272 which remains roughly the same over time.

**Author Contributions:** Conceptualization, W.K.; Data curation, S.H.A.; Formal analysis, R.R., A.R. and J.D.; Funding acquisition, H.Z., S.H.A. and W.K.; Methodology, H.Z.; Project administration, H.Z. and W.K.; Software, R.R., A.R. and J.D.; Supervision, H.Z., S.H.A. and W.K.; Visualization, R.R.; Writing—original draft, R.R., A.R. and P.K.R.; Writing—review & editing, H.Z., S.H.A. and W.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Institutes of Health (NIH) research fund, grant number R01AI121226, R01HL132321 and R01AI091905. "The Isle of Wight Birth Cohort assessments was funded by NIH research fund, grant number R01HL082925; Asthma UK, grant number 364; and the David Hide Asthma and Allergy Research Trust".

**Institutional Review Board Statement:** The IoW birth cohort study was conducted according to the guidelines of the Declaration of Helsinki, and approved by Isle of Wight, Portsmouth and Hampshire Local Research Ethics Committee (now known as the National Research Ethics Service, NRES Committee South Central—Hampshire A) (06/Q1701/34) (09/H0504/129 and IRB ID 32824) and the Institutional Review Board of the University of Memphis (FWA00006815 and IRB ID 2423 and 3917).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study. Written informed consent was obtained from parents during in-person visits. For participants assessed by phone interview, consent was documented on the consent form with the name of the person giving consent, and the name and signature of the person taking the form were recorded.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available because the cohort was established with specific focus on asthma and allergy related studies. The data access is restricted to ensure that the use of data is in compliance with participants' consent and ethics approval.

**Acknowledgments:** The authors are thankful to the nurses and staff at the David Hide Asthma & Allergy Research Centre, Isle of Wight, UK, for their help in recruitment and sample collections, and are thankful to all the cohort participants.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Habibzadeh, P.; Stoneman, E.K. The Novel Coronavirus: A Bird's Eye View. *Int. J. Occup. Environ. Med.* **2020**, *11*, 65–71. [[CrossRef](#)] [[PubMed](#)]
2. Perlman, S. Another Decade, Another Coronavirus. *N. Engl. J. Med.* **2020**, *382*, 760–762. [[CrossRef](#)] [[PubMed](#)]
3. Yuki, K.; Fujiogi, M.; Koutsogiannaki, S. COVID-19 pathophysiology: A review. *Clin. Immunol.* **2020**, *215*, 108427. [[CrossRef](#)]
4. Davies, N.G.; Klepac, P.; Liu, Y.; Prem, K.; Jit, M.; Eggo, R.M. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* **2020**, *26*, 1205–1211. [[CrossRef](#)]
5. Coronavirus (COVID-19) Latest Insights: Overview; Office for National Statistics. Available online: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19/latestinsights> (accessed on 29 July 2021).
6. Bialek, S.; Gierke, R.; Hughes, M.; McNamara, L.A. Coronavirus Disease 2019 in Children—United States, February 12–April 2, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **2020**, *69*, 422–426. [[CrossRef](#)]

7. Docherty, A.B.; Harrison, E.M.; Green, C.A.; Hardwick, H.E.; Pius, R.; Norman, L.; Holden, K.A.; Read, J.M.; Dondelinger, F.; Carson, G.; et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Prospective observational cohort study. *BMJ* **2020**, *369*, m1985. [[CrossRef](#)]
8. Wu, Z.; McGoogan, J.M. Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* **2020**, *323*, 1239–1242. [[CrossRef](#)]
9. O'Driscoll, M.; Ribeiro Dos Santos, G.; Wang, L.; Cummings, D.A.T.; Azman, A.S.; Paireau, J.; Fontanet, A.; Cauchemez, S.; Salje, H. Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* **2021**, *590*, 140–145. [[CrossRef](#)] [[PubMed](#)]
10. Leung, T.F.; Wong, G.W.; Hon, K.L.; Fok, T.F. Severe acute respiratory syndrome (SARS) in children: Epidemiology, presentation and management. *Paediatr. Respir. Rev.* **2003**, *4*, 334–339. [[CrossRef](#)]
11. Thabet, F.; Chehab, M.; Bafaqih, H.; Al Mohaimeed, S. Middle East respiratory syndrome coronavirus in children. *Saudi Med. J.* **2015**, *36*, 484–486. [[CrossRef](#)] [[PubMed](#)]
12. Lee, P.I.; Hu, Y.L.; Chen, P.Y.; Huang, Y.C.; Hsueh, P.R. Are children less susceptible to COVID-19? *J. Microbiol. Immunol. Infect.* **2020**, *53*, 371. [[CrossRef](#)] [[PubMed](#)]
13. Nikolich-Zugich, J.; Knox, K.S.; Rios, C.T.; Natt, B.; Bhattacharya, D.; Fain, M.J. SARS-CoV-2 and COVID-19 in older adults: What we may expect regarding pathogenesis, immune responses, and outcomes. *GeroScience* **2020**, *42*, 505–514. [[CrossRef](#)] [[PubMed](#)]
14. Jung, M.; Pfeifer, G.P. Aging and DNA methylation. *BMC Biol.* **2015**, *13*, 7. [[CrossRef](#)] [[PubMed](#)]
15. Pech, M.; Weckmann, M.; König, I.R.; Franke, A.; Heinsen, F.A.; Oliver, B.; Ricklefs, I.; Fuchs, O.; Rabe, K.; Hansen, G.; et al. Rhinovirus infections change DNA methylation and mRNA expression in children with asthma. *PLoS ONE* **2018**, *13*, e0205275. [[CrossRef](#)] [[PubMed](#)]
16. McErlean, P.; Favoreto, S., Jr.; Costa, F.F.; Shen, J.; Quraishi, J.; Biyasheva, A.; Cooper, J.J.; Scholtens, D.M.; Vanin, E.F.; de Bonaldo, M.F.; et al. Human rhinovirus infection causes different DNA methylation changes in nasal epithelial cells from healthy and asthmatic subjects. *BMC Med. Genom.* **2014**, *7*, 37. [[CrossRef](#)] [[PubMed](#)]
17. Bowden, S.J.; Kalliala, I.; Veroniki, A.A.; Arbyn, M.; Mitra, A.; Lathouras, K.; Mirabello, L.; Chadeau-Hyam, M.; Paraskevaidis, E.; Flanagan, J.M.; et al. The use of human papillomavirus DNA methylation in cervical intraepithelial neoplasia: A systematic review and meta-analysis. *EBioMedicine* **2019**, *50*, 246–259. [[CrossRef](#)] [[PubMed](#)]
18. Chai, P.; Yu, J.; Ge, S.; Jia, R.; Fan, X. Genetic alteration, RNA expression, and DNA methylation profiling of coronavirus disease 2019 (COVID-19) receptor ACE2 in malignancies: A pan-cancer analysis. *J. Hematol. Oncol.* **2020**, *13*, 43. [[CrossRef](#)]
19. Schäfer, A.; Baric, R.S. Epigenetic Landscape during Coronavirus Infection. *Pathogens* **2017**, *6*, 8. [[CrossRef](#)]
20. Menachery, V.D.; Schäfer, A.; Burnum-Johnson, K.E.; Mitchell, H.D.; Einfeld, A.J.; Walters, K.B.; Nicora, C.D.; Purvine, S.O.; Casey, C.P.; Monroe, M.E.; et al. MERS-CoV and H5N1 influenza virus antagonize antigen presentation by altering the epigenetic landscape. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E1012–E1021. [[CrossRef](#)]
21. Corley, M.J.; Ndhlovu, L.C. DNA Methylation Analysis of the COVID-19 Host Cell Receptor, Angiotensin I Converting Enzyme 2 Gene (ACE2) in the Respiratory System Reveal Age and Gender Differences. *Preprints* **2020**. [[CrossRef](#)]
22. Arshad, S.H.; Holloway, J.W.; Karmaus, W.; Zhang, H.; Ewart, S.; Mansfield, L.; Matthews, S.; Hodgekiss, C.; Roberts, G.; Kurukulaaratchy, R. Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC). *Int. J. Epidemiol.* **2018**, *47*, 1043i–1044i. [[CrossRef](#)] [[PubMed](#)]
23. Stelzer, G.; Rosen, N.; Plaschkes, I.; Zimmerman, S.; Twik, M.; Fishilevich, S.; Stein, T.I.; Nudel, R.; Lieder, I.; Mazor, Y.; et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinform.* **2016**, *54*, 1–30. [[CrossRef](#)] [[PubMed](#)]
24. Beyan, H.; Down, T.A.; Ramagopalan, S.V.; Uvebrant, K.; Nilsson, A.; Holland, M.L.; Gemma, C.; Giovannoni, G.; Boehm, B.O.; Ebers, G.C.; et al. Guthrie card methylomics identifies temporally stable epialleles that are present at birth in humans. *Genome Res.* **2012**, *22*, 2138–2145. [[CrossRef](#)]
25. Miller, S.A.; Dykes, D.D.; Polesky, H.F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* **1988**, *16*, 1215. [[CrossRef](#)]
26. Bibikova, M.; Fan, J.B. GoldenGate assay for DNA methylation profiling. *Methods Mol. Biol.* **2009**, *507*, 149–163. [[CrossRef](#)] [[PubMed](#)]
27. Lehne, B.; Drong, A.W.; Loh, M.; Zhang, W.; Scott, W.R.; Tan, S.T.; Afzal, U.; Scott, J.; Jarvelin, M.R.; Elliott, P.; 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*, 37. [[CrossRef](#)]
28. Aryee, M.J.; Jaffe, A.E.; Corrada-Bravo, H.; Ladd-Acosta, C.; Feinberg, A.P.; Hansen, K.D.; Irizarry, R.A. Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **2014**, *30*, 1363–1369. [[CrossRef](#)]
29. Johnson, W.E.; Li, C.; Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **2007**, *8*, 118–127. [[CrossRef](#)]
30. Du, P.; Zhang, X.; Huang, C.C.; Jafari, N.; Kibbe, W.A.; Hou, L.; Lin, S.M. Comparison of  $\beta$ -value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinform.* **2010**, *11*, 587. [[CrossRef](#)] [[PubMed](#)]

31. Reinius, L.E.; Acevedo, N.; Joerink, M.; Pershagen, G.; Dahlen, S.E.; Greco, D.; Soderhall, C.; Scheynius, A.; Kere, J. Differential DNA methylation in purified human blood cells: Implications for cell lineage and studies on disease susceptibility. *PLoS ONE* **2012**, *7*, e41361. [[CrossRef](#)] [[PubMed](#)]
32. Koestler, D.C.; Christensen, B.; Karagas, M.R.; Marsit, C.J.; Langevin, S.M.; Kelsey, K.T.; Wiencke, J.K.; Houseman, E.A. Blood-based profiles of DNA methylation predict the underlying distribution of cell types: A validation analysis. *Epigenetics* **2013**, *8*, 816–826. [[CrossRef](#)]
33. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [[CrossRef](#)] [[PubMed](#)]
34. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)] [[PubMed](#)]
35. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*, 166–169. [[CrossRef](#)]
36. Patel, R.; Solatikia, F.; Zhang, H.; Wolde, A.; Kadalayil, L.; Karmaus, W.; Ewart, S.; Arathimos, R.; Relton, C.; Ring, S.; et al. Sex-specific associations of asthma acquisition with changes in DNA methylation during adolescence. *Clin. Exp. Allergy* **2020**, *51*, 318–328. [[CrossRef](#)] [[PubMed](#)]
37. Wang, J.; Syrett, C.M.; Kramer, M.C.; Basu, A.; Atchison, M.L.; Anguera, M.C. Unusual maintenance of X chromosome inactivation predisposes female lymphocytes for increased expression from the inactive X. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E2029–E2038. [[CrossRef](#)] [[PubMed](#)]
38. Ng, M.W.; Zhou, G.; Chong, W.P.; Lee, L.W.; Law, H.K.; Zhang, H.; Wong, W.H.; Fok, S.F.; Zhai, Y.; Yung, R.W.; et al. The association of RANTES polymorphism with severe acute respiratory syndrome in Hong Kong and Beijing Chinese. *BMC Infect. Dis.* **2007**, *7*, 50. [[CrossRef](#)] [[PubMed](#)]
39. Chang, F.Y.; Chen, H.C.; Chen, P.J.; Ho, M.S.; Hsieh, S.L.; Lin, J.C.; Liu, F.T.; Sytwu, H.K. Immunologic aspects of characteristics, diagnosis, and treatment of coronavirus disease 2019 (COVID-19). *J. Biomed. Sci.* **2020**, *27*, 72. [[CrossRef](#)]
40. Patterson, B.K.; Seethamraju, H.; Dhody, K.; Corley, M.J.; Kazempour, K.; Lalezari, J.P.; Pang, A.P.; Sugai, C.; Francisco, E.B.; Pise, A.; et al. Disruption of the CCL5/RANTES-CCR5 Pathway Restores Immune Homeostasis and Reduces Plasma Viral Load in Critical COVID-19. *MedRxiv* **2020**. [[CrossRef](#)]
41. Smatti, M.K.; Al-Sarraj, Y.A.; Albagha, O.; Yassine, H.M. Host Genetic Variants Potentially Associated With SARS-CoV-2: A Multi-Population Analysis. *Front. Genet.* **2020**, *11*, 578523. [[CrossRef](#)]
42. Li, M.Y.; Li, L.; Zhang, Y.; Wang, X.S. Expression of the SARS-CoV-2 cell receptor gene ACE2 in a wide variety of human tissues. *Infect. Dis. Poverty* **2020**, *9*, 45. [[CrossRef](#)]
43. Bao, R.; Hernandez, K.; Huang, L.; Luke, J.J. ACE2 and TMPRSS2 expression by clinical, HLA, immune, and microbial correlates across 34 human cancers and matched normal tissues: Implications for SARS-CoV-2 COVID-19. *J. Immunother. Cancer* **2020**, *8*, e001020. [[CrossRef](#)]
44. Ziegler, C.G.K.; Allon, S.J.; Nyquist, S.K.; Mbanjo, I.M.; Miao, V.N.; Tzouanas, C.N.; Cao, Y.; Yousif, A.S.; Bals, J.; Hauser, B.M.; et al. SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. *Cell* **2020**, *181*, 1016–1035. [[CrossRef](#)] [[PubMed](#)]
45. Hamming, I.; Timens, W.; Bulthuis, M.L.; Lely, A.T.; Navis, G.; van Goor, H. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* **2004**, *203*, 631–637. [[CrossRef](#)] [[PubMed](#)]
46. Blanco-Melo, D.; Nilsson-Payant, B.E.; Liu, W.C.; Uhl, S.; Hoagland, D.; Møller, R.; Jordan, T.X.; Oishi, K.; Panis, M.; Sachs, D.; et al. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* **2020**, *181*, 1036–1045. [[CrossRef](#)]
47. Prasad, K.; Khatoun, F.; Rashid, S.; Ali, N.; AlAsmari, A.F.; Ahmed, M.Z.; Alqahtani, A.S.; Alqahtani, M.S.; Kumar, V. Targeting hub genes and pathways of innate immune response in COVID-19: A network biology perspective. *Int. J. Biol. Macromol.* **2020**, *163*, 1–8. [[CrossRef](#)]
48. Zhang, Y.; Zeng, T.; Chen, L.; Ding, S.; Huang, T.; Cai, Y.D. Identification of COVID-19 Infection-Related Human Genes Based on a Random Walk Model in a Virus-Human Protein Interaction Network. *BioMed Res. Int.* **2020**, *2020*, 4256301. [[CrossRef](#)]
49. Diao, B.; Wang, C.; Tan, Y.; Chen, X.; Liu, Y.; Ning, L.; Chen, L.; Li, M.; Liu, Y.; Wang, G.; et al. Reduction and Functional Exhaustion of T Cells in Patients With Coronavirus Disease 2019 (COVID-19). *Front. Immunol.* **2020**, *11*, 827. [[CrossRef](#)]
50. Allegra, A.; Di Gioacchino, M.; Tonacci, A.; Musolino, C.; Gangemi, S. Immunopathology of SARS-CoV-2 Infection: Immune Cells and Mediators, Prognostic Factors, and Immune-Therapeutic Implications. *Int. J. Mol. Sci.* **2020**, *21*, 4782. [[CrossRef](#)]
51. Xiong, Y.; Liu, Y.; Cao, L.; Wang, D.; Guo, M.; Jiang, A.; Guo, D.; Hu, W.; Yang, J.; Tang, Z.; et al. Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients. *Emerg. Microbes Infect.* **2020**, *9*, 761–770. [[CrossRef](#)] [[PubMed](#)]
52. Issa, E.; Merhi, G.; Panossian, B.; Salloum, T.; Tokajian, S. SARS-CoV-2 and ORF3a: Nonsynonymous Mutations, Functional Domains, and Viral Pathogenesis. *mSystems* **2020**, *5*, e00266–20. [[CrossRef](#)]
53. Lingappan, K.; Karmouty-Quintana, H.; Davies, J.; Akkanti, B.; Harting, M.T. Understanding the age divide in COVID-19: Why are children overwhelmingly spared? *Am. J. Physiol. Lung Cell. Mol. Physiol.* **2020**, *319*, L39–L44. [[CrossRef](#)] [[PubMed](#)]

54. Hui, K.P.Y.; Cheung, M.C.; Perera, R.; Ng, K.C.; Bui, C.H.T.; Ho, J.C.W.; Ng, M.M.T.; Kuok, D.I.T.; Shih, K.C.; Tsao, S.W.; et al. Tropism, replication competence, and innate immune responses of the coronavirus SARS-CoV-2 in human respiratory tract and conjunctiva: An analysis in ex-vivo and in-vitro cultures. *Lancet. Respir. Med.* **2020**, *8*, 687–695. [[CrossRef](#)]
55. Stueve, T.R.; Li, W.Q.; Shi, J.; Marconett, C.N.; Zhang, T.; Yang, C.; Mullen, D.; Yan, C.; Wheeler, W.; Hua, X.; et al. Epigenome-wide analysis of DNA methylation in lung tissue shows concordance with blood studies and identifies tobacco smoke-inducible enhancers. *Hum. Mol. Genet.* **2017**, *26*, 3014–3027. [[CrossRef](#)] [[PubMed](#)]
56. Ostaszewski, M.; Mazein, A.; Gillespie, M.E.; Kuperstein, I.; Niarakis, A.; Hermjakob, H.; Pico, A.R.; Willighagen, E.L.; Evelo, C.T.; Hasenauer, J.; et al. COVID-19 Disease Map, building a computational repository of SARS-CoV-2 virus-host interaction mechanisms. *Sci. Data* **2020**, *7*, 136. [[CrossRef](#)]