Conventional dendritic cell 2 links the genetic causal association from allergic asthma to COVID-19: a Mendelian randomization and transcriptomic study

Hua Liu[5], Siting Huang[2], Liting Yang[1,2,6], Hongshu Zhou[1,2,6], Bo Chen[1,3*] , Lisha Wu[4*] and Liyang Zhang[1,2,6*]

*Correspondence:
bochen1997@connect.
hku.hk; lisaent@csu.edu.cn;
zhangliyang@csu.edu.cn

[1] Department of Neurosurgery,
Xiangya Hospital, Central South
University, Jiangxi Hospital.,
Nanchang, No.266, Fenghe North
Road, 330038, Jiangxi, China
[2] Hypothalamic Pituitary
Research Center, Xiangya
Hospital, Central South
University, 87 Xiangya Road,
Changsha 410008, Hunan, China
[3] Division of Neurosurgery,
Department of Surgery, LKS
Faculty of Medicine, The
University of Hong Kong, Queen
Mary Hospital, 102 Pokfulam
Road, Hong Kong, People's
Republic of China
[4] Department
of Otorhinolaryngology Head
and Neck Surgery, Xiangya
Hospital, Central South
University, 87 Xiangya Road,
Changsha 410008, Hunan, China
[5] Department of Nuclear
Medicine, Xiangya Hospital,
Central South University, 87
Xiangya Road, Changsha 410008,
Hunan, China
[6] National Clinical Research
Center for Geriatric Disorders,
Xiangya Hospital, Central South
University, Changsha, Hunan,
People's Republic of China

Abstract

Recent evidence suggests that allergic asthma (AA) decreases the risk of Coronavirus Disease 2019 (COVID-19). However, the reasons remain unclear. Here, we systematically explored data from GWAS (18 cohorts with 11,071,744 samples), bulk transcriptomes (3 cohorts with 601 samples), and single-cell transcriptomes (2 cohorts with 29 samples) to reveal the immune mechanisms that connect AA and COVID-19. Two-sample Mendelian randomization (MR) analysis identified a negative causal correlation from AA to COVID-19 hospitalization (OR = 0.968, 95% CI 0.940–0.997, P = 0.031). This correlation was bridged through white cell count. Furthermore, machine learning identified dendritic cells (DCs) as the most discriminative immunocytes in AA and COVID-19. Among five DC subtypes, only conventional dendritic cell 2 (cDC2) exhibited differential expression between AA/COVID-19 and controls (P < 0.05). Subsequently, energy metabolism, intercellular communication, cellular stemness and differentiation, and molecular docking analyses were performed. cDC2s exhibited more differentiation, increased numbers, and enhanced activation in AA exacerbation, while they showed less differentiation, reduced number, and enhanced activation in severe COVID-19. The capacity of cDC2 for differentiation and SARS-CoV-2 antigen presentation may be enhanced through ZBTB46, EXOC4, TLR1, and TNFSF4 gene mutations in AA. Taken together, cDC2 links the genetic causality from AA to COVID-19. Future strategies for COVID-19 prevention, intervention, and treatment could be stratified according to AA and guided with DC-based therapies.

Highlights of study

1. Investigating the correlation between allergic asthma (AA) and Coronavirus Disease 2019 (COVID-19) through Mendelian randomization (MR), bulk and single-cell transcriptome, and molecular docking.
2. AA exhibited a negative causal correlation to COVID-19 hospitalization, which is bridged by cDC2.

Liu *et al. Journal of Big Data*   (2024) 11:26

Page 2 of 18

3. cDC2s exhibited better differentiation, increased numbers, and enhanced activation in AA exacerbation, while COVID-19 decreased the differentiation and number of cDC2s.

4. The capacity of cDC2s for differentiation and SARS-CoV-2 antigen presentation may be enhanced through ZBTB46, EXOC4, TLR1, and TNFSF4 gene mutations in AA.

Keywords: Allergic asthma (AA), Coronavirus disease 2019 (COVID-19), Causal effect, Conventional dendritic cell 2 (cDC2), Differentiation, Antigen presentation, Mendelian randomization (MR), Transcriptomic analyses

Graphical Abstract
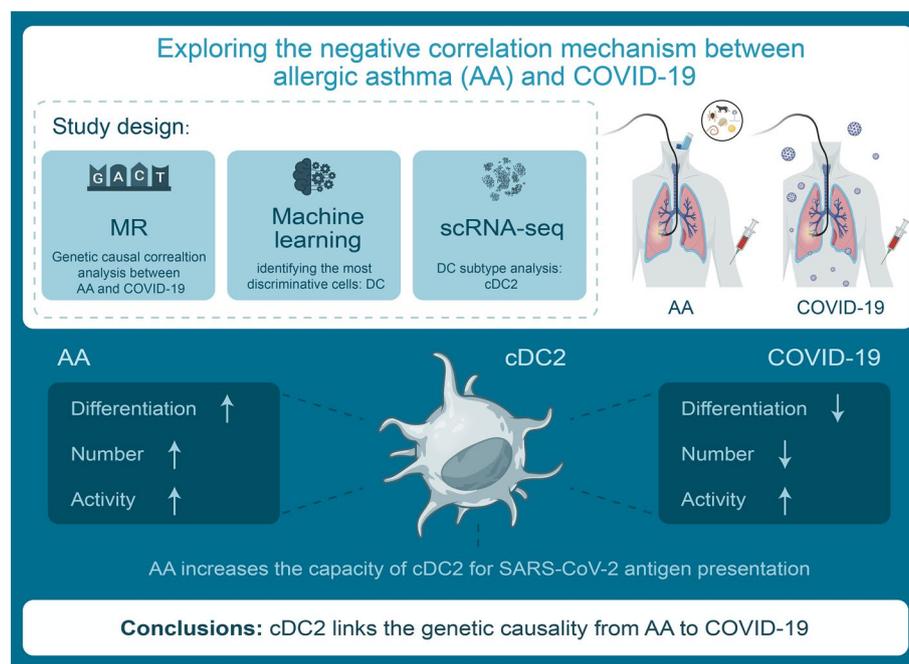
Introduction

The Coronavirus Disease 2019 (COVID-19) pandemic, caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), has imposed a huge burden on global health and the economy [1–4]. According to the World Health Organization, there have been 767,518,723 cumulative COVID-19 cases and 6,947,192 cumulative COVID-19 deaths worldwide as of July 3rd, 2023 (https://covid19.who.int/). A collection of risk and protective factors have been identified which correlate to susceptibility, hospitalization, and severity of COVID-19 [5–9]. Among these, the correlations between COVID-19 and allergic diseases have been repeatedly reported [10–13]. Allergic asthma (AA) is commonly defined as asthma associated with sensitization to aeroallergens [14]. Approximately 11% of the general population suffer from AA [15]. Previous studies reveal that AA decreases the risk of hospitalization and death in COVID-19 patients compared with non-AA [16, 17]. Patients with AA have a lower

probability of SARS-CoV-2 test positivity than those with non-AA [16]. However, the associated mechanism between these two diseases remains largely unknown.

Immune cells play crucial roles in regulating AA attacks and defending against COVID-19 infections [18, 19]. As the most efficient antigen-presenting cells, dendritic cells (DCs) link innate and adaptive immunity [20]. There are five main subtypes of DCs in human lung tissue: conventional DC1 (cDC1), conventional DC2 (cDC2), plasmacytoid DC (pDC), monocyte-derived DC (Mo-DC), and Pre-DC [21]. Among these, cDC2 is derived from Pre-DCs, accounting for the most prominent subset of DCs [22]. The molecular markers of cDC2 include CD1C, ITGAM, CLEC10A, and others [21, 22]. In AA lungs, the number of cDC2s is significantly up-regulated [23]. cDC2s can promote T helper (Th) cell differentiation and regulate humoral immunity in asthma [19, 24]. After SARS-CoV-2 infection, cDC2s are activated and accumulate in patients lungs [25]. The activated cDC2s can induce a positive immune response in COVID-19 [20]. Furthermore, in some severe infections, the number and maturation of DCs are reduced, indicating DC dysfunction or failure [20]. Although some studies explored the immune landscape in AA and COVID-19, to our knowledge, the shared immune genetic architectures have not been investigated.

Herein, we used 18 GWAS datasets to examine the causal links between allergic diseases and COVID-19 by employing bi-directional two-sample Mendelian randomization (MR) analysis and considering blood parameters as risk factors. Among these, AA and white cell count stood out. We further performed machine learning using 3 bulk transcriptome datasets to screen the shared discriminative immunocytes in AA and COVID-19, and identified cDC2s. Moreover, the quantity, activity, intercellular communication, and differentiation of cDC2s were investigated in 2 single-cell transcriptome datasets, and the antigen presentation process of cDC2s was exhibited using molecular docking analysis. In summary, this study proposes for the first time, that cDC2s link the genetic causal association between AA and COVID-19.

Materials and methods

Ethics approval

An overview of the study design is presented in Fig. 1. As this study utilized pre-existing publications and public databases, no further ethical approval or consent was necessary.

Data source

GWAS data for COVID-19, including 3 phenotypes (susceptibility, severity, and hospitalization), were extracted from the COVID-19 Host Genetics (hg) GWAS meta-analyses (Round 7, April 2022) (https://www.covid19hg.org/results/r7/). GWAS data for allergy, consisting of 8 phenotypes, and for blood, comprising 7 parameters, were obtained from the FinnGen database (https://www.finngen.fi/en/access_results) and GWAS Catalog/ IEU Open GWAS Project (https://www.ebi.ac.uk/gwas, http://gwas.mrcieu.ac.uk), separately. All GWAS was restricted to individuals of European ancestry. Moreover, transcriptome data of blood and airway mucosa in COVID-19 and AA cohorts were accessed from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/). A total of 630 samples from 5 datasets (including 3 bulk and 2 single-cell datasets) were collected. Bulk-tissue datasets included blood samples from AA/
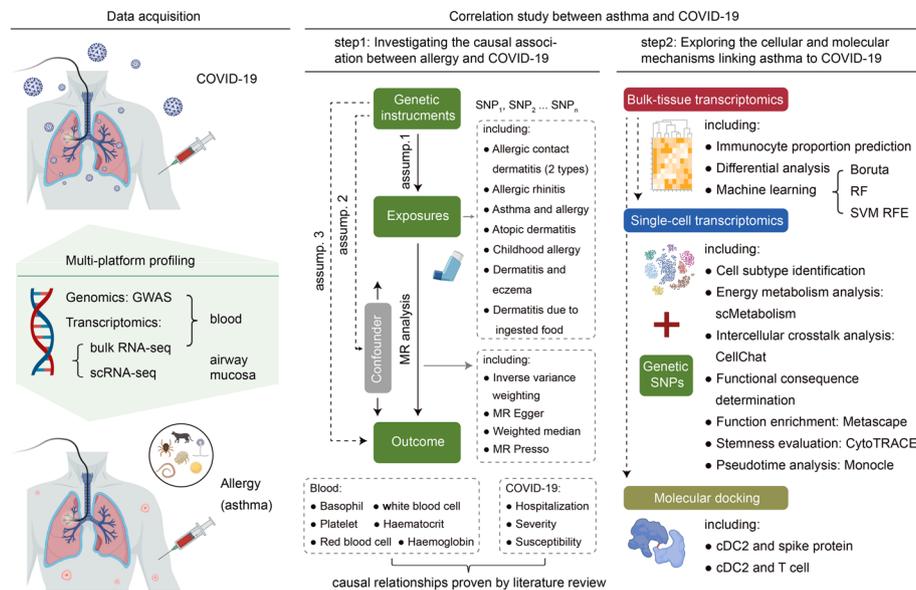
Fig. 1 Study design overview. Genomic and transcriptomic data of COVID-19 and allergic diseases were collected from public platforms. The causal correlations between allergy and COVID-19 were detected using the Mendelian randomization (MR) analysis, and blood parameters were used as mediating variables. Based on MR results, the study further applied bulk-tissue transcriptome, single-cell transcriptome, and molecular docking analyses to explore the potential cellular and molecular mechanism that bridges the allergic asthma (AA) and COVID-19

COVID-19 patients and healthy controls (HCs). Single-cell datasets incorporated airway mucosa samples from AA patients and AA controls (AC) with or without allergen (Ag) challenge, as well as airway mucosa samples from severe (S) COVID-19 patients and HCs. For detailed information see Additional file 2:: Table S1. Additionally, 3 protein structures were downloaded from the Protein Data Bank (PDB, https://www1.rcsb.org/), including the COVID-19 spike glycolprotein (6VXX), TLR1 (6NIH), and OX40L-OX40 (TNFSF4-TNFRSF4) complex (2HEV).

MR analysis

The bi-directional two-sample MR study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement [26]. The study design fulfilled 3 crucial assumptions [26]: (1) instrumental variables should be strongly correlated to exposures; (2) instrumental variables should be independent of underlying confounders; (3) instrumental variables influence the outcome only through the risk factors (Fig. 1). For forward analysis, allergic diseases were designated as exposure factors, while COVID-19 was set as the outcome factor. Reverse analysis was conducted for the causal relationships identified by the forward analysis, where COVID-19 (hospitalization) was designated as the exposure factor and allergic disease (AA) was set as the outcome factor. Next, instrumental variable SNPs were selected based on the following criteria: (1) high exposure correlation: $P < 5 \times 10^{-6}$; (2) high independence: linkage disequilibrium (LD) $r^2 < 0.001$; (3) high statistical strength: $F > 10$; (4) no direct outcome correlation: $P > 5 \times 10^{-6}$; (5) no palindromic and incompatible structure: harmonization analysis [27]. The harmonization analysis between exposure and outcome data was conducted

Liu *et al. Journal of Big Data*   (2024) 11:26

Page 5 of 18

using the function 'harmonise_data' of the R package 'TwoSampleMR'. MR-Steiger filtering [28] was applied to exclude SNPs which implied reverse causal direction. Moreover, since some phenotypes had a limited number of SNPs to meet the criteria, we relaxed their correlation threshold to $P < 5 \times 10^{-5}$. [29] Furthermore, 4 MR analytical algorithms were applied to evaluate the causal effects between exposures and outcomes. Among these, inverse-variance weighted (IVW) functioned as the primary method, while the other three, namely MR Egger, Weighted Median (WM), and MR PRESSO, served as supplementary tools [30]. Statistical significance was determined, only when the P-value of IVW was < 0.05 and the OR directions of the other three algorithms were consistent with that of IVW.

Sensitivity analyses were performed, which included tests for pleiotropy and heterogeneity and leave-one-out analysis. Directional pleiotropy was measured by the Egger-intercept test [31] (no pleiotropy: $P > 0.05$). Heterogeneity was assessed through the Cochran Q test (no heterogeneity: $P > 0.05$) and funnel pot. Leave-one-out analysis was conducted to determine whether the causal correlation was driven by any single SNP (no driven: results are all on one side). Furthermore, to explore the mechanisms genetically linking allergy and COVID-19, we evaluated some underlying mediators, namely seven blood parameters that have been previously reported to exhibit causal correlations to COVID-19 [32–35]. Allergy phenotypes were treated as exposures, and blood parameters were designated as outcomes. Their MR analyses were conducted based on the methods mentioned above.

Bulk transcriptome analysis

The immunocyte proportion of blood in AA and COVID-19 cohorts was estimated using 4 independent algorithms: CIBERSORT [36], MCPCounter [37], ssGSEA [38], and xCell [39]. The discriminating abilities of immunocytes between AA/COVID-19 patients and HCs were evaluated and ranked through three machine learning algorithms: Boruta, Random Forest, and Support Vector Machine Recursive Feature Elimination (SVM RFE) [40, 41]. Next, we intersected the immunocytes that exhibited high discriminating abilities across three algorithms. The overlapping cells between the AA and COVID-19 were identified as target cells.

Single-cell transcriptome analysis

The data were initially normalized using the functions 'NormalizeData' and 'ScaleData', and its dimensionality was reduced by the functions 'FindVariableFeatures' and 'Run-PCA' (R package 'Seurat') [42]. Batch effects were corrected with the function 'Run-Harmony' (R package 'harmony') [43]. Cells were clustered and visualized through the functions 'FindClusters' and 'UMAP', separately (R package 'Seurat') [42]. Cell lineage identity was annotated according to the previous studies [23, 44]. DCs were further divided into different subtypes (pDC, cDC1, cDC2, Mo DC, and Pre DC) using sets of marker genes [21]. Furthermore, the degree of cell activation was assessed by quantifying energy metabolism using the R package 'scMetabolism' [45]. Intercellular cross-talk was calculated by the R package 'CellChat' [46]. Cell differentiation was evaluated through pseudotime analysis with the R package 'monocle3´ [47] and stemness analysis with the R package 'cytoTRACE' [48]. Additionally, the corresponding genes and

functional consequences of instrumental variable SNPs for allergy were obtained from the dbSNP website (https://www.ncbi.nlm.nih.gov/snp/). For these genes, functional enrichment was conducted using the Metascape website (https://metascape.org/), and their role in cell differentiation was assessed.

Molecular docking

Protein docking was conducted using the ZDOCK module of Discovery Studio software (version 2019, available at https://discover.3ds.com/discovery-studio-visualizer-download). The docking structures with the top ZDOCK score were selected. Their 3D interaction diagram was visualized using the PyMOL software (version 1.8, available at http://www.pymol.org/pymol), and the 2D interaction figure was generated using the LigPlot+software (version 2.2.8, available at https://www.ebi.ac.uk/thornton-srv/software/LIGPLOT/).

Statistical analysis

Statistical analysis was conducted using the R software (version 4.1.2, available at https://www.r-project.org/). The Shapiro-Wilks test was applied for the normality test. The two-tailed t-test (normal distribution) and Mann–Whitney test (non-normal distribution) were utilized to compare the continuous variables between the two groups, separately. Additionally, MR analysis was conducted using the R package 'TwoSampleMR'. Forest plots were produced through the R package 'forestplot'. Bubble, bar, and violin plots were generated by the R package 'ggplot2'. Venn diagrams were drawn through the R package 'VennDiagram'. NS indicates no statistical significance, while '*' indicates statistical significance, namely $P < 0.05$.

Results

AA exhibits a negative causal correlation to COVID-19

Instrumental variable SNPs were screened for different allergy phenotypes and COVID-19 hospitalization, which has been detailed in Additional file 2: Table S2. The number of SNPs varied from 5 to 76, and their summary data are shown in Additional file 2: Tables S3-S11. Among the allergy phenotypes, IVW analysis showed that AA decreased the risk of COVID-19 hospitalization (OR=0.968, 95% CI 0.940—0.997, P=0.031). The result directions of the other 3 MR methods were consistent (Fig. 2A, B, and Additional file 2: Table S12). Additionally, no reverse causal relationship was detected from COVID-19 hospitalization to AA (Additional file 2: Table S14). We also failed to observe causal associations between the other allergy phenotypes and COVID-19 risk (Fig. 2A and Additional file 2: Table S12). To evaluate the robustness and reliability of the above results, we conducted multiple types of sensitivity analyses. MR-Egger intercept tests indicated that no horizontal pleiotropy existed (all P > 0.05) (Additional file 2: Tables S13 and S14). However, heterogeneity was observed in Cochran's Q test among some allergy and COVID-19 phenotypes, like the association from AA to COVID-19 (Q=52.293, P=0.030) (Fig. 2C and Additional file 2: Tables S13 and S14). Despite the presence of heterogeneity, the MR estimates in the current study were not invalidated, as the random-effect IVW method was employed [49, 50]. Moreover, the existence of

Liu *et al. Journal of Big Data*   (2024) 11:26

Page 7 of 18
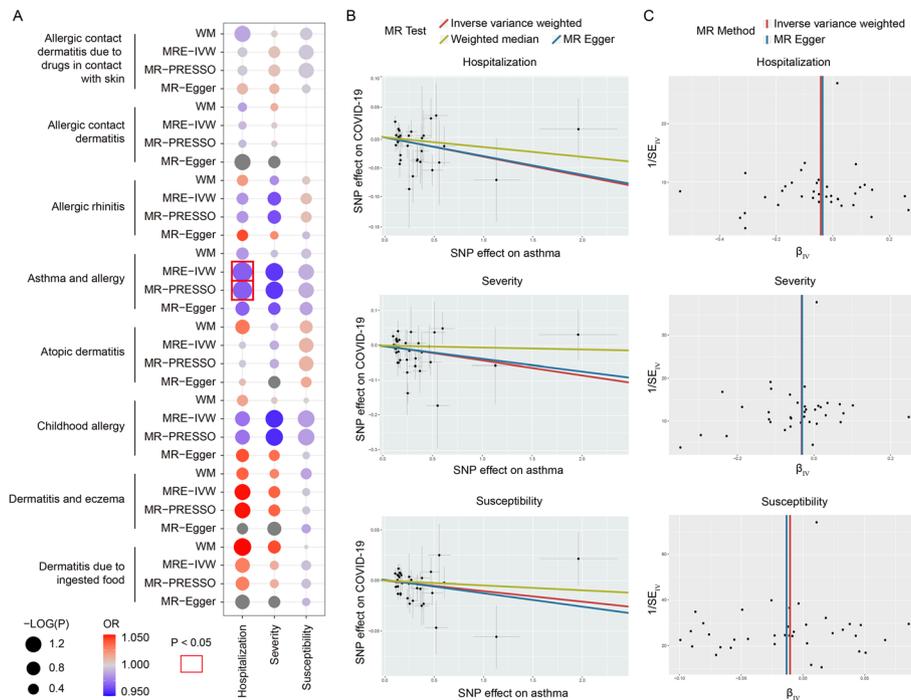
Fig. 2 AA exhibits a negative causal correlation to COVID-19 hospitalization. **A** Bubble plot showing the causal correlation between allergic diseases and COVID-19 phenotypes. **B** Scatter plots displaying the causal correlation between AA and COVID-19 phenotypes. **C** Funnel plots assessing the heterogeneity of the correlation between AA and COVID-19 phenotypes
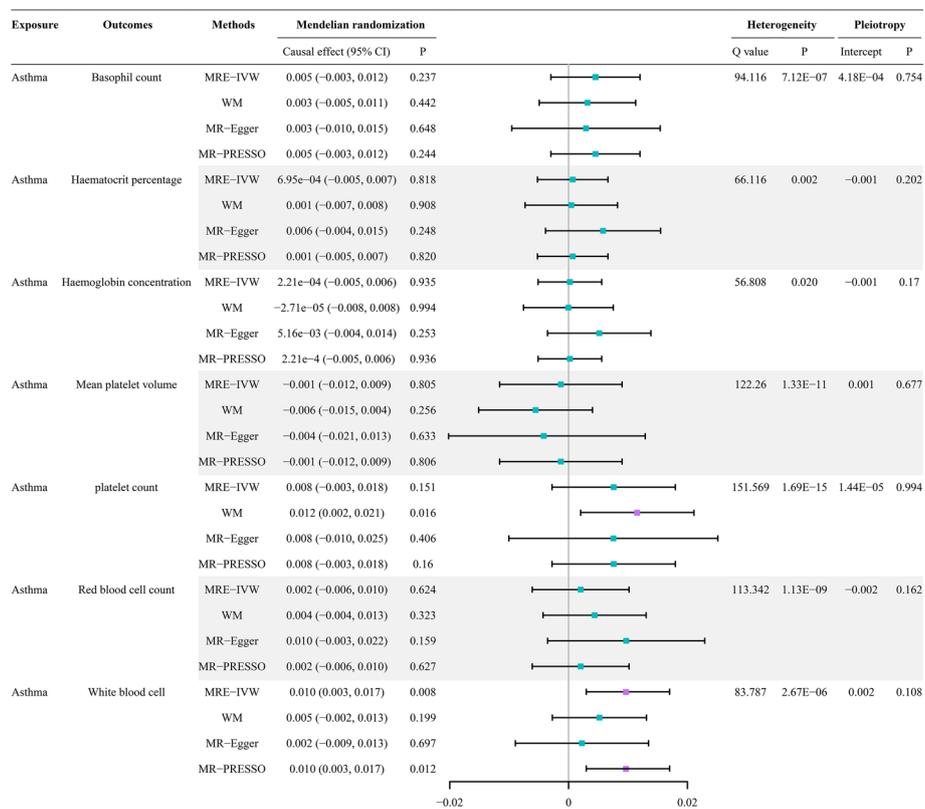
heterogeneity did not introduce any pleiotropic bias to MR estimates. Furthermore, the leave-one-out analysis demonstrated that no single SNP drove the results, indicating that none of the estimates were violated (Additional file 1: Fig. S1A).

White cell counts bridge the correlation between AA and COVID-19

Based on logical reasoning, the correlation between AA and COVID-19 hospitalization may not be direct. To explore the underlying mediators linking AA and COVID-19 hospitalization, we evaluated the effect of AA on COVID-19-related blood variables through MR analysis. AA exhibited a positive correlation to white cell count (IVW, $\beta = 0.010$, 95% CI 0.003—0.017, $P = 0.008$), which was consistent with the direction of results obtained from the other three MR methods (Fig. 3). However, no significant association was observed between AA and other blood variables (Fig. 3). Furthermore, the sensitivity analysis did not reveal any evidence of pleiotropy but reflected the presence of heterogeneity, which was accepted under the random-effect IVW method (Fig. 3). The leave-one-out analysis further excluded significant impact from a single SNP on the MR estimates (Additional file 1: Fig. S1B). These data suggest that white cell count bridged AA and COVID-19.

DC is the most discriminative white cell in AA and COVID-19

Because the range of white cells remains too large for the mediator variables between AA and COVID-19, we tried to narrow the scope. We used 4 independent algorithms to

Liu *et al. Journal of Big Data*   (2024) 11:26

Page 8 of 18

Exposure	Outcomes	Methods	Mendelian randomization			Heterogeneity		Pleiotropy	
			Causal effect (95% CI)	P		Q value	P	Intercept	P
Asthma	Basophil count	MRE−IVW	0.005 (−0.003, 0.012)	0.237		94.116	7.12E−07	4.18E−04	0.754
		WM	0.003 (−0.005, 0.011)	0.442					
		MR−Egger	0.003 (−0.010, 0.015)	0.648					
		MR−PRESSO	0.005 (−0.003, 0.012)	0.244					
Asthma	Haematocrit percentage	MRE−IVW	6.95e−04 (−0.005, 0.007)	0.818		66.116	0.002	−0.001	0.202
		WM	0.001 (−0.007, 0.008)	0.908					
		MR−Egger	0.006 (−0.004, 0.015)	0.248					
		MR−PRESSO	0.001 (−0.005, 0.007)	0.820					
Asthma	Haemoglobin concentration	MRE−IVW	2.21e−04 (−0.005, 0.006)	0.935		56.808	0.020	−0.001	0.17
		WM	−2.71e−05 (−0.008, 0.008)	0.994					
		MR−Egger	5.16e−03 (−0.004, 0.014)	0.253					
		MR−PRESSO	2.21e−4 (−0.005, 0.006)	0.936					
Asthma	Mean platelet volume	MRE−IVW	−0.001 (−0.012, 0.009)	0.805		122.26	1.33E−11	0.001	0.677
		WM	−0.006 (−0.015, 0.004)	0.256					
		MR−Egger	−0.004 (−0.021, 0.013)	0.633					
		MR−PRESSO	−0.001 (−0.012, 0.009)	0.806					
Asthma	platelet count	MRE−IVW	0.008 (−0.003, 0.018)	0.151		151.569	1.69E−15	1.44E−05	0.994
		WM	0.012 (0.002, 0.021)	0.016					
		MR−Egger	0.008 (−0.010, 0.025)	0.406					
		MR−PRESSO	0.008 (−0.003, 0.018)	0.16					
Asthma	Red blood cell count	MRE−IVW	0.002 (−0.006, 0.010)	0.624		113.342	1.13E−09	−0.002	0.162
		WM	0.004 (−0.004, 0.013)	0.323					
		MR−Egger	0.010 (−0.003, 0.022)	0.159					
		MR−PRESSO	0.002 (−0.006, 0.010)	0.627					
Asthma	White blood cell	MRE−IVW	0.010 (0.003, 0.017)	0.008		83.787	2.67E−06	0.002	0.108
		WM	0.005 (−0.002, 0.013)	0.199					
		MR−Egger	0.002 (−0.009, 0.013)	0.697					
		MR−PRESSO	0.010 (0.003, 0.017)	0.012					

Fig. 3 White blood cell counts bridge the correlation between AA and COVID-19. Blood parameters were used as mediating variables between AA and COVID-19. Forest plots show the causal correlation between AA and blood parameters. The causal correlation between these blood parameters and COVID-19 has been identified in previous studies

identify the particular type of white blood cell that bridged AA and COVID-19. Overall, compared to healthy subjects, AA/COVID-19 patients had significantly higher levels of granulocytes and monocytes/macrophages, as well as lower levels of B cells and T cells in the blood (Fig. 4A and Additional file 2: Table S15). The expression of DCs was up-regulated in the blood of AA patients and down-regulated in the blood of COVID-19 patients (Fig. 4A and Additional file 2: Table S15). Subsequently, three machine learning algorithms (Boruta, Random Forest, and SVM RFE) were separately applied to identify the sixteen most discriminative immunocytes between AA/COVID-19 patients and healthy subjects (Fig. 4B and Additional file 2: Table S16–S18). Through the intersection, we identified DCs as the target cells (Fig. 4C).
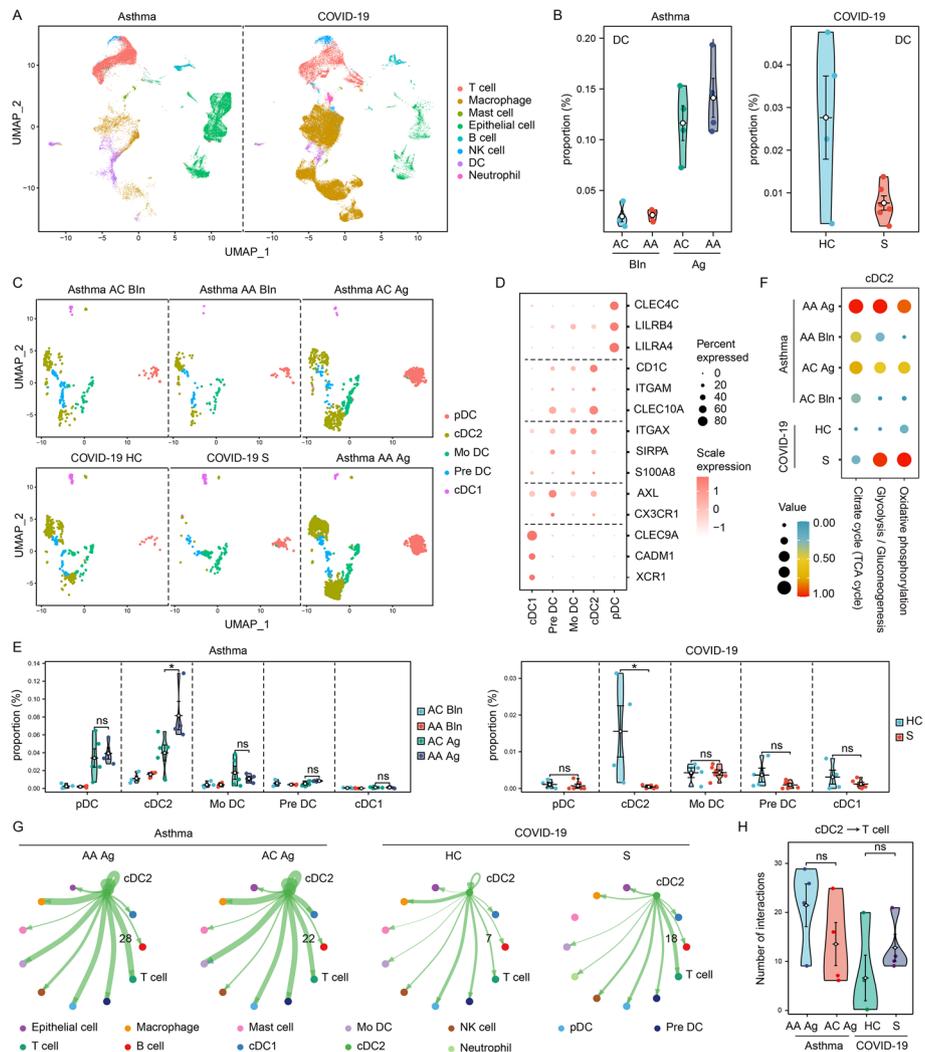
cDC2 is upregulated/activated in AA exacerbation and downregulated/activated in severe COVID-19

To further uncover the correlation between the two diseases at the cellular subtype level, we performed the single-cell RNA sequencing (scRNA-seq) analysis on the airway mucosa of AA and COVID-19 cohorts. First, AA and COVID-19 datasets were integrated and corrected for batch effects (Additional file 1: Fig. S2A and B). Eight types of cells were annotated, including T cells, macrophages, mast cells, epithelial cells, B

Fig. 4 Dendritic cells (DCs) are the most discriminative immune cell in AA and COVID-19. **A** Heatmap showing the expression of immune cells in AA and COVID-19. **B** Three machine learning algorithms separately identify the top 16 discriminative immune cells in AA and COVID-19. **C** Venn plots showing the intersected immunocytes identified by three machine learning algorithms for distinguishing AA/COVID-19 and controls

cells, NK cells, DCs, and neutrophils (Fig. 5A). Their cell markers were confirmed (Additional file 1: Fig. S2C). For AA, AAs with Ag challenge had the highest proportion of DCs, followed by ACs with Ag challenge and AAs/ACs without Ag (samples at baseline). For COVID-19, patients with severe infection exhibited a lower proportion of DCs compared to HCs (Fig. 5B). These two results validated the findings regarding DCs in the bulk sequencing analysis. Subsequently, we further divided the DCs into 5 cellular subtypes, including pDCs, cDC2s, Mo DCs, Pre DCs, and cDC1s (Fig. 5C). Their cell markers were confirmed (Fig. 5D). When comparing the expression of DC subtypes in different groups of AA/COVID-19, only cDC2s exhibited statistically significant disparities (P < 0.05). In terms of AA, AAs with Ag challenge exhibited the highest proportion of cDC2s, followed by ACs with Ag challenge and AAs/ACs without Ag. Regarding COVID-19, patients with severe infection demonstrated a significantly lower proportion

Fig. 5 Conventional dendritic 2 cells (cDC2s) are increased/activated in AA exacerbation and downregulated/activated in severe COVID-19. **A** UMAP plot of total cells from AA and COVID-19, with each cell color coded for cell type. **B** The proportion of DC in AA/COVID-19 and controls. **C** UMAP plot of DC populations from AA and COVID-19, with each cell color coded for cell subtype. **D** Bubble plot showing marker genes for cell subtype annotation. **E** The proportion of DC subtypes in AA/COVID-19 and controls. **F** The level of energy metabolism of cDC2 in AA/COVID-19 and controls. **G** Intercellular communication from cDC2 to other cells. **H** The number of interactions from cDC2 to T cells in AA/COVID-19 and controls. *Bln* samples at baseline, *Ag* samples receiving antigen, *AC* allergic non-asthmatic controls, *HC* healthy controls, *S* patients with severe COVID-19 infection

of cDC2s compared to HCs (Fig. 5E). These results suggest that the level of cDC2s is up-regulated in AA exacerbation and down-regulated in the severe COVID-19.

Next, we investigated whether there were differences in the activity of cDC2s among different groups of AA/COVID-19. We found that the levels of energy metabolism were increased in AAs in response to Ag and were higher in AAs in comparison with ACs (Fig. 5F). Additionally, severe COVID-19 exhibited higher levels of energy metabolism than HCs (Fig. 5F). cDC2s have been found to possess the ability to present antigens to T cells [22]. Thus, we examined the cell communication between cDC2s and other cells. In the context of antigen exposure, cDC2s from AAs sent more signals to T cells compared

to that from ACs. Moreover, cDC2s from COVID-19 patients trended towards increased signal transmission to T cells in comparison to HCs, despite lacking statistical significance due to limited sample size (Fig. 5G, H). Collectively, these data suggest the presence of more activated cDC2s in exacerbation status of AA and COVID-19.
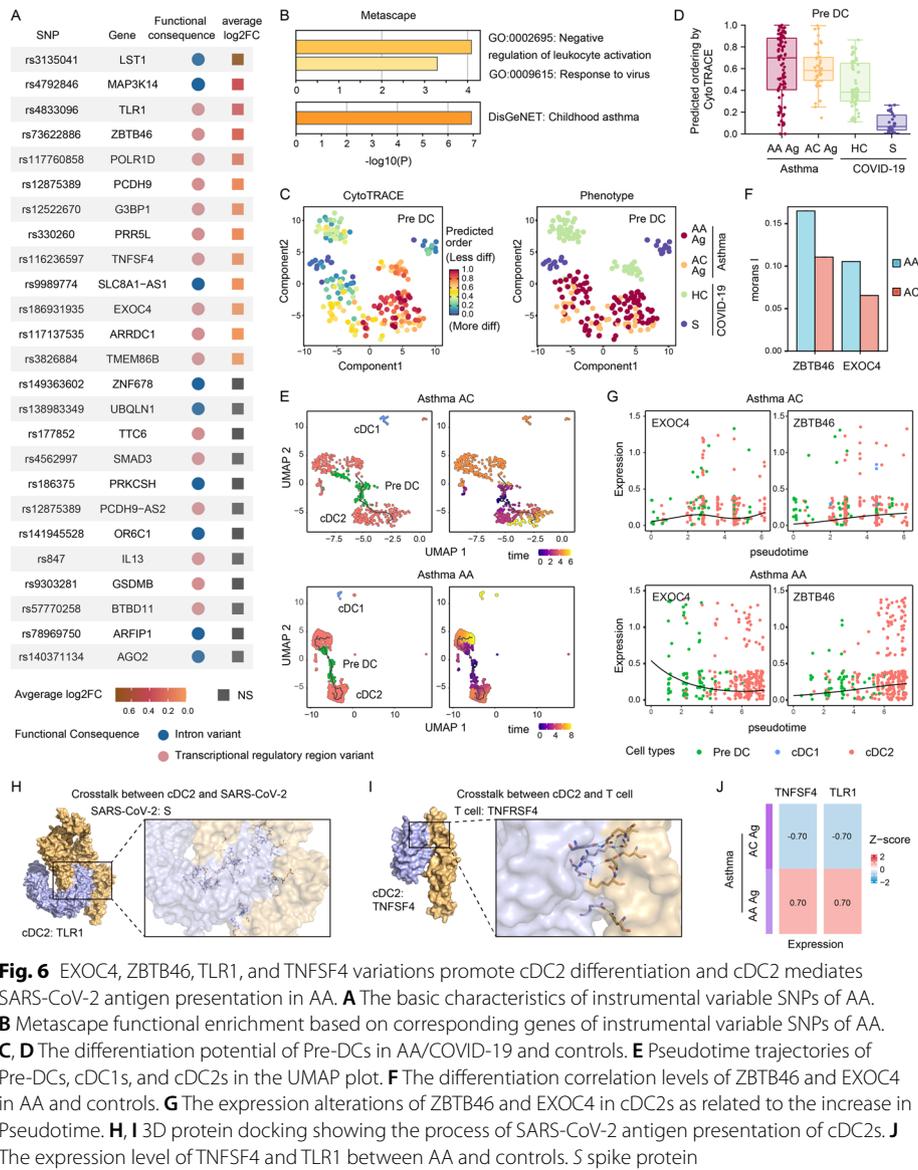
EXOC4, ZBTB46, TLR1 and TNFSF4 variations promote the cDC2 differentiation and cDC2 mediates SARS-CoV-2 antigen presentation in AA

Because the causal correlation between AA and COVID-19 had been identified in the aforementioned results, we further explored the shared genetic mechanism. Within 36 instrumental variable SNPs of AA, a total of 25 had corresponding genes, including TLR1, ZBTB46, EXOC4, and TNFSF4. Among these, 16 SNPs exhibited the functional consequences of transcriptional regulatory region variants, and 13 SNP corresponding genes were highly expressed in DCs (min.pct > 0.01, log2FC > 0.01, and P < 0.05) (Fig. 6A and Additional file 2: Tables S19, S20). Metascape enrichment analysis showed that these 13 genes focused on AA (childhood asthma), virus (response to virus), and immune cells (negative regulation of leukocyte activation) (Fig. 6B).

Next, we investigated the differentiation and antigen presentation of DCs. CytoTRACE analysis predicted a higher developmental potential for Pre-DCs from AA Ag than those from AC Ag (Fig. 6C, D). Pre-DCs from COVID-19 patients exhibited decreased differentiation capacity compared to those from HCs (Fig. 6C, D). Monocle analysis drew the differentiation trajectory from Pre-DC to cDC1/cDC2 in AAs and ACs (Fig. 6E). Differentiation correlations of those genes were presented along the pseudotime trajectories (Additional file 2: Table S20). Among these, both ZBTB46 and EXOC4 exhibited dramatically higher correlations with differentiation in AAs than ACs (Fig. 6F and Additional file 1: Table S20). As pseudotime advanced, ZBTB46 was up-regulated in both AAs and ACs, while EXOC4 displayed down-regulation in AAs but up-regulation in ACs (Fig. 6G). Moreover, protein docking analysis presented the process of antigen presentation mediated by cDC2s, including the identification of SARS-CoV-2 (protein docking between TLR1 of cDC2 and spike protein on SARS-CoV-2) and presentation to T cells (protein docking between TNFSF4 of cDC1 and TNFRSF4 on T cells). Interactive forces and distances were displayed (Fig. 6H, I, and Additional file 1: Fig. S3). When antigen was exposed, both TLR1 and TNFSF4 exhibited higher expression in AAs than ACs, indicating a more pronounced activation of cDC2-mediated SARS-CoV-2 antigen presentation in AAs (Fig. 6J). Collectively, the above findings uncovered that EXOC4, ZBTB46, TLR1, and TNFSF4 variations promote cDC2 differentiation and cDC2 mediated SARS-CoV-2 antigen presentation in AA.

Discussion

AA has been suggested to be a protective factor for COVID-19, with the potential mechanism being largely unknown [5, 16, 17]. In this study, we found that AA exhibits negative genetic correlations to COVID-19 hospitalization, which is mediated by white cell count. Among them, cDC2s are the most discriminative immune cells between AA/COVID-19 and controls. When diseases are exacerbated, cDC2s exhibit upregulation, activation, and more differentiation in AA, while they show downregulation, activation, and poorer differentiation in COVID-19. ZBTB46, EXOC4, TLR1,

Fig. 6 EXOC4, ZBTB46, TLR1, and TNFSF4 variations promote cDC2 differentiation and cDC2 mediates SARS-CoV-2 antigen presentation in AA. **A** The basic characteristics of instrumental variable SNPs of AA. **B** Metascape functional enrichment based on corresponding genes of instrumental variable SNPs of AA. **C**, **D** The differentiation potential of Pre-DCs in AA/COVID-19 and controls. **E** Pseudotime trajectories of Pre-DCs, cDC1s, and cDC2s in the UMAP plot. **F** The differentiation correlation levels of ZBTB46 and EXOC4 in AA and controls. **G** The expression alterations of ZBTB46 and EXOC4 in cDC2s as related to the increase in Pseudotime. **H**, **I** 3D protein docking showing the process of SARS-CoV-2 antigen presentation of cDC2s. **J** The expression level of TNFSF4 and TLR1 between AA and controls. *S* spike protein

and TNFSF4 gene mutations promoted cDC2 differentiation and SARS-CoV-2 antigen presentation in AA.

The causal correlation between asthma and COVID-19 has been previously investigated. Wang et al. reported that asthma was negatively associated with susceptibility to and severe respiratory symptoms of COVID-19 [13]. Baranova et al. indicated that asthma decreased the risk both for COVID-19 hospitalization and infection [8]. Similarly, our study revealed a negative causal correlation from AA to COVID-19 hospitalization. However, the correlation between AA and COVID-19 susceptibility/severity was not observed. Such disparities might be due to the use of different datasets, which may generate varying instrumental variables and yield different MR

results. Moreover, hospitalization represents a particularly severe condition. Therefore, in subsequent transcriptome analyses, we focused on the exacerbation of AA and COVID-19.

Cellular immunity plays many crucial roles in AA and COVID-19 diseases [18, 19]. In AA, epithelial cell-derived cytokines drive the type 2 immune response, which promotes the infiltration of lung tissue with eosinophils, Type 2 T helper (Th2) cells, inflammatory DCs, and others. The accumulation of these cells leads to the acute response and chronic features of AA [19]. In COVID-19, the number of immunocytes, including DC, NK, B, CD4[+] T, CD8[+] T, and other cells are dramatically altered. Defects in cytotoxicity of NK and T cells lead to cytokine storm, causing tissue damage and organ dysfunction [51]. In our study, AA exhibited a positive causal correlation to white cell count, which has been previously reported to possess causal protective effects in COVID-19 [35]. Accordingly, we inferred that the white cell count bridged the correlation between AA and COVID-19. Through machine learning, we further identified DCs, as the most discriminative immunocytes between AA/COVID-19 and controls. As the most efficient antigen-presenting cells, DCs connect innate immunity and adaptive immunity. After penetrating the submucosa, Ag or SARS-CoV-2 are captured by DCs, undergo processing, and are subsequently presented to T cells [19, 20]. Previous studies have reported that AA increased the number of DCs in the peripheral blood [52]. However, the quantity and maturation of DCs decreases in severe COVID-19 [20]. Similarly, we found that the expression of DCs was up-regulated in AA blood and airway mucosa and down-regulated in those of COVID-19.

DCs have five subtypes, mainly including cDC1s, cDC2s, pDCs, Mo-DCs, and Pre-DCs [21]. Among these, cDC2s constitute a significant proportion of the DCs [22]. Our single-cell analysis found that only cDC2s were differentially expressed in the airway mucosa between AA/COVID-19 exacerbation and controls. The number of cDC2s was upregulated in AA, whereas it was downregulated in COVID-19. Consistent with our findings, previous studies have revealed that AA patients had a higher percentage of cDC2s in bronchial tissues, while hospitalized COVID-19 patients had a lower proportion of cCD2s in blood [25, 53]. cDC2s are derived from the differentiation of Pre-DCs [21]. Accordingly, we conducted stemness analysis, and found that the Pre-DCs of AA exacerbation possessed higher differentiation potential, whereas the Pre-DCs of COVID-19 exacerbation had lower differentiation potential. The stemness disparity of pre-DCs may, at least partially, account for the number difference of cDC2s in these two diseases. In addition, pseudotime analysis revealed that the expression of AA mutation genes ZBTB46 and EXOC4 were highly correlated to cDC2 differentiation. Among them, the role of ZBTB46 in DC differentiation has been previously reported. ZBTB46, a zinc finger transcription factor, is a marker for conventional DCs and committed DC precursors [54, 55]. Satpathy et al. found that ZBTB46 overexpression could promote cDC development [56]. Wang et al. revealed that ZBTB46 characterizations were potentially involved in DC maturation and activation [57]. In addition, EXOC4 encodes the protein belonging to the exocyst complex component family. Yi eti al. reported that EXOC2 knockdown inhibits COVID-19 infection. However, the study of EXOC4 in DC differentiation is lacking and deserves further investigation.

Liu *et al. Journal of Big Data*   (2024) 11:26

Page 14 of 18

cDC2s respond to multiple danger signals ranging from nucleotides to polysaccharides, and can promote a variety of immune responses, especially the CD4+T cell response and Th cell response [20, 22]. In our study, more activated cDC2s were found in both exacerbated AA and severe COVID-19. These cells sent more signals to T cells, promoting the production of adaptive immune responses. However, it is worth noting that the roles of cDC2s in AA and COVID-19 were mainly opposite. In AA, cDC2s responded to Ag, which contributed to disease occurrence, while in COVID-19, cDC2s resisted the virus invasion, which inhibited the disease progression. Furthermore, we investigated the mutation genes of AA, and found that the TLR1 protein of cDC2s (encoded by AA mutation gene TLR1) could strongly bind with the spike protein of SARS-CoV-2, and the TNFSF4 protein of cDC2s (encoded by AA mutation gene TNFSF4) could bind with its receptor TNFRSF4 on T cells. Both TLR1 and TNFSF4 were more highly expressed in cDC2s of AA compared to those of controls, suggesting an increased capacity of cDC2s for antigen presentation in AA. Among these, TLR1, belonging to the Toll-like receptor (TLR) family, plays a crucial role in pathogen recognition and activation of innate immunity [58]. Previous studies have confirmed the TLR1 gene mutation is associated with asthma [59]. The TLR family can mediate antiviral DC responses to resist SARS-CoV-2 infection [60]. Additionally, TNFSF4, belonging to the tumor necrosis factor (TNF) ligand family, functions in the interaction between antigen-presenting cells and T cells [61]. Previous studies have confirmed that SNP polymorphisms in TNFSF4 are correlated to asthma [62]. TNFSF4 is primarily expressed on DCs, and its receptor TNFRSF4 is preferentially expressed by T cells [63]. TNFSF4/TNFRSF4 signaling plays a crucial role in resisting the infection of bacteria and viruses [64].

According to the study results, we proposed a potential theory explaining why patients with AA are less susceptible to COVID-19 hospitalization. For COVID-19, SARS-CoV-2 may inhibit the differentiation of Pre-DCs, reduce the number of cDC2s, attenuate the connection between the innate immune and adaptive immune system, and as a result, increase its ability to infect the body. AA might enhance the capacity of Pre-DC differentiation through EXOC4 and ZBTB46 gene mutation, and therefore increase the number of cDC2s, especially the number of active cDC2s. These cells highly express TLR1 and TNFSF4, recognize more SARS-CoV-2, and present more signals to T cells, which induces a stronger adaptive immune response to resist SARS-CoV-2 infection.

Our study has some limitations. First, the MR analysis was performed among European ancestry participants, while bioinformatic analysis was conducted among American, Asian, and European ancestry participants. Race differentiation belongs to confounding factors, which may affect the accuracy of results. Second, apart from blood parameters, there was a lack of analysis for other risk factors that bridged AA and COVID-19. Third, this study is lacking the experimental validation.

Conclusions

In conclusion, this study suggests that cDC2s link the genetic causality from AA to COVID-19. Future strategies of COVID-19 prevention, intervention, and treatment for the population may be stratified according to AA and guided with dendritic cell (DC)-based therapies.

Liu *et al. Journal of Big Data*   (2024) 11:26

Page 15 of 18

Abbreviations

AA	Allergic asthma
AC	Allergic asthma control
Ag	Allergen
cDC1 (DC1)	Conventional dendritic cell 1
cDC2 (DC2)	Conventional dendritic cell 2
COVID-19	Coronavirus disease 2019
DC	Dendritic cell
GEO	Gene expression omnibus
HC	Healthy controls
Hg	Host genetics
IVW	Inverse-variance weighted
LD	Linkage disequilibrium
Mo-DC	Monocyte-derived dendritic cell
MR	Mendelian randomization
pDC	Plasmacytoid dendritic cell
STROBE	Strengthening the reporting of observational studies in epidemiology
SVM RFE	Support vector machine recursive feature elimination
Th	T helper
Th2	Type 2T helper
WM	Weighted median

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40537-024-00881-1.

> **Additional file 1: Figure S1.** Leave-one-out analysis for AA on COVID-19 hospitalization (A) and AA on white cell count (B). **Figure S2**. (A, B) The integration of AA and COVID-19 datasets. (C) Bubble plot shows marker genes for cell type annotation. **Figure S3**. (A) The 2D protein docking graph shows the interactive forces and distances between TLR1 and spike protein. (B) The 2D protein docking graph shows the interactive forces and distances between TNFSF4 and TNFRSF4.
>
> **Additional file 2: Table S1.** The datasets used in this study. **Table S2.** The number of instrument variables for exposures. **Table S3.** Instrument variables of allergic contact dermatitis due to drugs in contact with skin. **Table S4.** Instrument variables of allergic contact dermatitis. **Table S5.** Instrument variables of allergic rhinitis. **Table S6.** Instrument variables of asthma and allergy. **Table S7.** Instrument variables of atopic dermatitis. **Table S8.** Instrument variables of childhood allergy. **Table S9.** Instrument variables of dermatitis and eczema. **Table S10.** Instrument variables of dermatitis due to ingested food. **Table S11.** Instrument variables of COVID-19 hospitalization. **Table S12.** MR estimates of the causality from the allergy risk to COVID-19 risk. **Table S13.** Sensitivity analysis of the causality from the allergy risk to COVID-19 risk. **Table S14.** MR estimates and sensitivity analysis of the causality from the COVID-19 hospitalization to allergic asthma. **Table S15.** Comparison of immune cell populations in the blood between AA/COVID-19 patients and control subjects. **Table S16.** The rankings of immune cell power to distinguish AA patients and control subjects in the GSE69683 dataset. **Table S17.** The rankings of immune cell power to distinguish COVID-19 patients and control subjects in the GSE171110 dataset. **Table S18.** The rankings of immune cell power to distinguish COVID-19 patients and control subjects in the GSE196822 dataset. **Table S19.** The genes and functional consequences corresponded to instrumental variable SNPs of AA. **Table S20.** The expression of corresponding genes of AA SNPs in DCs and their role in DC differentiation.

Author contributions
HL, BC, LSW, and LYZ designed and drafted the manuscript; HL, BC, LSW, and LYZ organized the figures and edited the legends; HL, BC, LSW, and LYZ revised the article; HL, BC, LSW, LYZ, STH, LTY, and HSZ conducted the data analysis; and all authors have read and approved the final manuscript.

Availability of data and materials
The datasets generated and/or analyzed during the current study are available in the COVID-19 hg (https://www.covid19hg.org/results/r7/), FinnGen (https://www.finngen.fi/en/access_results), GWAS Catalog (https://www.ebi.ac.uk/gwas/), IEU-OpenGWAS project (https://gwas.mrcieu.ac.uk/), and GEO (https://www.ncbi.nlm.nih.gov/geo/).

Liu *et al. Journal of Big Data*   (2024) 11:26

Page 16 of 18

Declarations

Ethics approval and consent to participate
Genome-wide association studies that provided data for our analysis were approved by the relevant review boards. All analyses were based on publicly available non-individual-level data, and no ethical approval was required from the ethics committee.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References

1. Cao G, Guo Z, Liu J, Liu M. Change from low to out-of-season epidemics of influenza in China during the COVID-19 pandemic: a time series study. J Med Virol. 2023;95(6): e28888. https://doi.org/10.1002/jmv.28888.
2. Lee J, Lee KO. Online listing data and their interaction with market dynamics: evidence from Singapore during COVID-19. J Big Data. 2023;10(1):99. https://doi.org/10.1186/s40537-023-00786-5.
3. Domalewska D. An analysis of COVID-19 economic measures and attitudes: evidence from social media mining. J Big Data. 2021;8(1):42. https://doi.org/10.1186/s40537-021-00431-z.
4. Corti L, Zanetti M, Tricella G, Bonati M. Social media analysis of Twitter tweets related to ASD in 2019–2020, with particular attention to COVID-19: topic modelling and sentiment analysis. J Big Data. 2022;9(1):113. https://doi.org/10.1186/s40537-022-00666-4.
5. Baranova A, Cao H, Teng S, Zhang F. A phenome-wide investigation of risk factors for severe COVID-19. J Med Virol. 2023;95(1): e28264. https://doi.org/10.1002/jmv.28264.
6. Cao H, Baranova A, Wei X, Wang C, Zhang F. Bidirectional causal associations between type 2 diabetes and COVID-19. J Med Virol. 2023;95(1): e28100. https://doi.org/10.1002/jmv.28100.
7. Baranova A, Song Y, Cao H, Zhang F. Causal associations between basal metabolic rate and COVID-19. Diabetes. 2023;72(1):149–54. https://doi.org/10.2337/db22-0610.
8. Baranova A, Cao H, Chen J, Zhang F. Causal association and shared genetics between asthma and COVID-19. Front Immunol. 2022;13: 705379. https://doi.org/10.3389/fimmu.2022.705379.
9. Hssayeni MD, Chala A, Dev R, et al. The forecast of COVID-19 spread risk at the county level. J Big Data. 2021;8(1):99. https://doi.org/10.1186/s40537-021-00491-1.
10. Ren J, Pang W, Luo Y, et al. Impact of allergic rhinitis and asthma on COVID-19 infection, hospitalization, and mortality. J Allergy Clin Immunol Pract. 2022;10(1):124–33. https://doi.org/10.1016/j.jaip.2021.10.049.
11. Cianferoni A, Votto M. COVID-19 and allergy: how to take care of allergic patients during a pandemic? Pediatr Allergy Immunol. 2020;31(Suppl 26):96–101. https://doi.org/10.1111/pai.13367.
12. Gao YD, Agache I, Akdis M, et al. The effect of allergy and asthma as a comorbidity on the susceptibility and outcomes of COVID-19. Int Immunol. 2022;34(4):177–88. https://doi.org/10.1093/intimm/dxab107.
13. Wang Y, Gu X, Wang X, Zhu W, Su J. Exploring genetic associations between allergic diseases and indicators of COVID-19 using mendelian randomization. iScience. 2023;26(6): 106936. https://doi.org/10.1016/j.isci.2023.106936.
14. Akar-Ghibril N, Casale T, Custovic A, Phipatanakul W. Allergic endotypes and phenotypes of asthma. J Allergy Clin Immunol Pract. 2020;8(2):429–40. https://doi.org/10.1016/j.jaip.2019.11.008.
15. Pakkasela J, Ilmarinen P, Honkamäki J, et al. Age-specific incidence of allergic and non-allergic asthma. BMC Pulm Med. 2020;20(1):9. https://doi.org/10.1186/s12890-019-1040-2.
16. Yang JM, Koh HY, Moon SY, et al. Allergic disorders and susceptibility to and severity of COVID-19: a nationwide cohort study. J Allergy Clin Immunol. 2020;146(4):790–8. https://doi.org/10.1016/j.jaci.2020.08.008.
17. Murphy TR, Busse W, Holweg CTJ, et al. Patients with allergic asthma have lower risk of severe COVID-19 outcomes than patients with nonallergic asthma. BMC Pulm Med. 2022;22(1):418. https://doi.org/10.1186/s12890-022-02230-5.
18. Primorac D, Vrdoljak K, Brlek P, et al. Adaptive immune responses and immunity to SARS-CoV-2. Front Immunol. 2022;13: 848582. https://doi.org/10.3389/fimmu.2022.848582.
19. Komlósi ZI, van de Veen W, Kovács N, et al. Cellular and molecular mechanisms of allergic asthma. Mol Aspects Med. 2022;85: 100995. https://doi.org/10.1016/j.mam.2021.100995.
20. Wang X, Guan F, Miller H, et al. The role of dendritic cells in COVID-19 infection. Emerg Microbes Infect. 2023;12(1):2195019. https://doi.org/10.1080/22221751.2023.2195019.
21. Collin M, Bigley V. Human dendritic cell subsets: an update. Immunology. 2018;154(1):3–20. https://doi.org/10.1111/imm.12888.
22. Balan S, Saxena M, Bhardwaj N. Dendritic cell subsets and locations. Int Rev Cell Mol Biol. 2019;348:1–68. https://doi.org/10.1016/bs.ircmb.2019.07.004.
23. Alladina J, Smith NP, Kooistra T, et al. A human model of asthma exacerbation reveals transcriptional programs and cell circuits specific to allergic asthma. Sci Immunol. 2023;8(83): eabq6352. https://doi.org/10.1126/sciimmunol.abq6352.
24. Sakurai S, Furuhashi K, Horiguchi R, et al. Conventional type 2 lung dendritic cells are potent inducers of follicular helper T cells in the asthmatic lung. Allergol Int. 2021;70(3):351–9. https://doi.org/10.1016/j.alit.2021.01.008.
25. Winheim E, Rinke L, Lutz K, et al. Impaired function and delayed regeneration of dendritic cells in COVID-19. PLoS Path. 2021;17(10): e1009742. https://doi.org/10.1371/journal.ppat.1009742.

26. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomization: the STROBE-MR statement. JAMA. 2021;326(16):1614–21. https://doi.org/10.1001/jama.2021.18236.

27. Cai J, He L, Wang H, et al. Genetic liability for prescription opioid use and risk of cardiovascular diseases: a multivariable Mendelian randomization study. Addiction. 2022;117(5):1382–91. https://doi.org/10.1111/add.15767.

28. Hemani G, Tilling K, Davey SG. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. PLoS Genet. 2017;13(11): e1007081. https://doi.org/10.1371/journal.pgen.1007081.

29. Li P, Wang H, Guo L, et al. Association between gut microbiota and preeclampsia-eclampsia: a two-sample Mendelian randomization study. BMC Med. 2022;20(1):443. https://doi.org/10.1186/s12916-022-02657-x.

30. Chen X, Hong X, Gao W, et al. Causal relationship between physical activity, leisure sedentary behaviors and COVID-19 risk: a Mendelian randomization study. J Transl Med. 2022;20(1):216. https://doi.org/10.1186/s12967-022-03407-6.

31. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol. 2017;32(5):377–89. https://doi.org/10.1007/s10654-017-0255-x.

32. Fan Z, Ruan Z, Liu Z, et al. Causal association of the brain structure with the susceptibility, hospitalization, and severity of COVID-19: a large-scale genetic correlation study. J Med Virol. 2023;95(3): e28651. https://doi.org/10.1002/jmv.28651.

33. Cheung CL, Ho SC, Krishnamoorthy S, Li GH. COVID-19 and platelet traits: a bidirectional Mendelian randomization study. J Med Virol. 2022;94(10):4735–43. https://doi.org/10.1002/jmv.27920.

34. Yao Y, Song H, Zhang F, et al. Genetic predisposition to blood cell indices in relation to severe COVID-19. J Med Virol. 2023;95(1): e28104. https://doi.org/10.1002/jmv.28104.

35. Sun Y, Zhou J, Ye K. White blood cells and severe COVID-19: a Mendelian randomization study. J Pers Med. 2021. https://doi.org/10.3390/jpm11030195.

36. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453–7. https://doi.org/10.1038/nmeth.3337.

37. Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17(1):218. https://doi.org/10.1186/s13059-016-1070-5.

38. Xia J, Xie Z, Niu G, et al. Single-cell landscape and clinical outcomes of infiltrating B cells in colorectal cancer. Immunology. 2023;168(1):135–51. https://doi.org/10.1111/imm.13568.

39. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017;18(1):220. https://doi.org/10.1186/s13059-017-1349-1.

40. Zhang N, Zhang H, Liu Z, et al. An artificial intelligence network-guided signature for predicting outcome and immunotherapy response in lung adenocarcinoma patients based on 26 machine learning algorithms. Cell Prolif. 2023. https://doi.org/10.1111/cpr.13409.

41. Laatifi M, Douzi S, Bouklouz A, et al. Machine learning approaches in Covid-19 severity risk prediction in Morocco. J Big Data. 2022;9(1):5. https://doi.org/10.1186/s40537-021-00557-0.

42. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. Cell. 2019;177(7):1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

43. Korsunsky I, Millard N, Fan J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019;16(12):1289–96. https://doi.org/10.1038/s41592-019-0619-0.

44. Liao M, Liu Y, Yuan J, et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. Nat Med. 2020;26(6):842–4. https://doi.org/10.1038/s41591-020-0901-9.

45. Wu Y, Yang S, Ma J, et al. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. Cancer Discov. 2022;12(1):134–53. https://doi.org/10.1158/2159-8290.Cd-21-0316.

46. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using Cell Chat. Nat Commun. 2021;12(1):1088. https://doi.org/10.1038/s41467-021-21246-9.

47. Cao J, Spielmann M, Qiu X, et al. The single-cell transcriptional landscape of mammalian organogenesis. Nature. 2019;566(7745):496–502. https://doi.org/10.1038/s41586-019-0969-x.

48. Gulati GS, Sikandar SS, Wesche DJ, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. Science. 2020;367(6476):405–11. https://doi.org/10.1126/science.aax0249.

49. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557–60. https://doi.org/10.1136/bmj.327.7414.557.

50. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. Stat Med. 2017;36(11):1783–802. https://doi.org/10.1002/sim.7221.

51. Jasim SA, Mahdi RS, Bokov DO, et al. The deciphering of the immune cells and marker signature in COVID-19 pathogenesis: an update. J Med Virol. 2022;94(11):5128–48. https://doi.org/10.1002/jmv.28000.

52. Morianos I, Semitekolou M. Dendritic cells: critical regulators of allergic asthma. Int J Mol Sci. 2020. https://doi.org/10.3390/ijms21217930.

53. El-Gammal A, Oliveria JP, Howie K, et al. Allergen-induced changes in bone marrow and airway dendritic cells in subjects with asthma. Am J Respir Crit Care Med. 2016;194(2):169–77. https://doi.org/10.1164/rccm.201508-1623OC.

54. Shao T, Ji JF, Zheng JY, et al. Zbtb46 controls dendritic cell activation by reprogramming epigenetic regulation of cd80/86 and cd40 costimulatory signals in a zebrafish model. J Immunol. 2022;208(12):2686–701. https://doi.org/10.4049/jimmunol.2100952.

55. Satpathy AT, Brown RA, Gomulia E, et al. Expression of the transcription factor ZBTB46 distinguishes human histiocytic disorders of classical dendritic cell origin. Mod Pathol. 2018;31(9):1479–86. https://doi.org/10.1038/s41379-018-0052-4.

56. Satpathy AT, Kc W, Albring JC, et al. Zbtb46 expression distinguishes classical dendritic cells and their committed progenitors from other immune lineages. J Exp Med. 2012;209(6):1135–52. https://doi.org/10.1084/jem.20120030.

57. Wang J, Wang T, Benedicenti O, et al. Characterisation of ZBTB46 and DC-SCRIPT/ZNF366 in rainbow trout, transcription factors potentially involved in dendritic cell maturation and activation in fish. Dev Comp Immunol. 2018;80:2–14. https://doi.org/10.1016/j.dci.2016.11.007.

Liu *et al. Journal of Big Data*   (2024) 11:26

Page 18 of 18

58. Duan T, Du Y, Xing C, Wang HY, Wang RF. Toll-like receptor signaling and its role in cell-mediated immunity. Front Immunol. 2022;13: 812774. https://doi.org/10.3389/fimmu.2022.812774.

59. Koponen P, Vuononvirta J, Nuolivirta K, Helminen M, He Q, Korppi M. The association of genetic variants in toll-like receptor 2 subfamily with allergy and asthma after hospitalization for bronchiolitis in infancy. Pediatr Infect Dis J. 2014;33(5):463–6. https://doi.org/10.1097/inf.0000000000000253.

60. van der Sluis RM, Cham LB, Gris-Oliver A, et al. TLR2 and TLR7 mediate distinct immunopathological and antiviral plasmacytoid dendritic cell responses to SARS-CoV-2 infection. EMBO J. 2022;41(10): e109622. https://doi.org/10.15252/embj.2021109622.

61. Fu N, Xie F, Sun Z, Wang Q. The OX40/OX40L axis regulates T follicular helper cell differentiation: implications for autoimmune diseases. Front Immunol. 2021;12: 670637. https://doi.org/10.3389/fimmu.2021.670637.

62. Liu Y, Ke X, Kang HY, Wang XQ, Shen Y, Hong SL. Genetic risk of TNFSF4 and FAM167A-BLK polymorphisms in children with asthma and allergic rhinitis in a Han Chinese population. J Asthma. 2016;53(6):567–75. https://doi.org/10.3109/02770903.2015.1108437.

63. Kaur D, Brightling C. OX40/OX40 ligand interactions in T-cell regulation and asthma. Chest. 2012;141(2):494–9. https://doi.org/10.1378/chest.11-1730.

64. Ming S, Zhang M, Liang Z, et al. OX40L/OX40 signal promotes IL-9 production by mucosal MAIT cells during *Helicobacter pylori* infection. Front Immunol. 2021;12: 626017. https://doi.org/10.3389/fimmu.2021.626017.

Publisher's Note