



# Role of lymphoid lineage cells aberrantly expressing alarmins S100A8/A9 in determining the severity of COVID-19

Joongho Lee<sup>1</sup> · Hanbyeol Kim<sup>1</sup> · Minsoo Kim<sup>1</sup> · Seokhyun Yoon<sup>1,2</sup> · Sanghun Lee<sup>3</sup>

Received: 27 April 2022 / Accepted: 8 July 2022 / Published online: 15 September 2022  
© The Author(s) under exclusive licence to The Genetics Society of Korea 2022

## Abstract

**Background** Alarmins S100A8 and S100A9 are recognized as hallmarks of severe COVID-19 and are primarily produced in myeloid cells, such as monocytes and neutrophils. As single-cell RNA-sequencing (scRNA-seq) data from patients with COVID-19 revealed the expression of *S100A8/A9* in lymphoid cells in patients with severe COVID-19.

**Objective** We investigated the characteristics of lymphoid cells expressing S100A8/A9 in COVID-19 patients.

**Methods** Publicly available scRNA-seq data from patients with mild (N = 12) or severe (N = 7) COVID-19 were reanalyzed. The data were further divided into the following two groups based on the time of sample collection (from infection-onset): within 6 days (early phase) and after 6 days (late phase). Differential expression and gene set enrichment analyses were performed between S100A8/A9<sup>High</sup> and S100A8/A9<sup>Low</sup> lymphoid cells. Finally, cell-cell interaction analysis was performed to investigate the role of lymphoid cells expressing high levels of *S100A8/A9* in COVID-19.

**Results** *S100A8/A9* overexpression was observed in lymphoid cells, including B cells, T cells, and NK cells, in patients with severe COVID-19 (compared to patients with mild COVID-19). Cells exhibiting strong interferon/cytokine responses were found to be associated with the severity of COVID-19. Furthermore, differences in S100A8/A9-TLR4/RAGE interactions were confirmed between patients with severe and mild disease.

**Conclusions** Lymphoid cells overexpressing *S100A8/A9* contribute to the dysregulation of the innate immune response in patients with severe COVID-19, specifically during the early phase of infection. This study fosters a better understanding of the hyper-induction of pro-inflammatory cytokine expression and the generation of a cytokine storm in response to COVID-19 infection.

**Keywords** COVID-19 · Single-cell RNA-sequencing · S100A8/A9 · Lymphoid cells

## Introduction

The coronavirus disease 2019 (COVID-19) pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an ongoing global threat (Zhao

et al. 2020). The clinical spectrum of SARS-CoV-2 infection ranges from asymptomatic to mild to a severe disease requiring intensive care, which may include life-threatening pneumonia and acute respiratory distress syndrome (ARDS) (Nishiura et al. 2020; Yang et al. 2020). The dysregulated and excessive secretion of pro-inflammatory cytokines, i.e., cytokine storm, due to over-activation of immune cells has been recognized as a hallmark of severe COVID-19. Many cytokines, including IL-6, IL-8, and TNF-alpha, have previously been described as severity indicators for COVID-19. Several single-cell RNA sequencing (scRNA-seq) analyses conducted on COVID-19 patients have revealed profound alterations in myeloid cells, such as the disappearance of non-classical CD14<sup>Low</sup>CD16<sup>High</sup> monocytes, accumulation of HLA-DR<sup>Low</sup> classical monocytes, and accumulation of immature neutrophils with calprotectin plasma levels, all serving as candidate biomarkers of severe COVID-19 (Ren

✉ Seokhyun Yoon  
syoona@dankook.ac.kr

✉ Sanghun Lee  
shlee92@dankook.ac.kr

<sup>1</sup> Department of Computer Science and Engineering, Graduate School, Dankook University, Yongin-si, Republic of Korea

<sup>2</sup> Department of Electronics and Electrical Engineering, College of Engineering, Dankook University, Yongin-si, Republic of Korea

<sup>3</sup> Department of Bioconvergence Engineering, Graduate School, Dankook University, Yongin-si, Republic of Korea

**Table 1** Demographic characteristics and clinical features of study subjects

Group	Severity (count)	Gender (count)	Age (mean)	Days after symptom onset (mean)
Early phases after COVID-19 infection	Mild (3)	Male (1)/female (2)	39.67	5.33
	Severe (2)	Male (2)	55.5	2
Late phases after COVID-19 infection	Mild (9)	Male (5)/female (4)	35.44	10.56
	Severe (5)	Male (2)/female (3)	42.4	10.8

et al. 2021). Growing evidence suggests that the systemic upregulation of alarmins S100A8/A9, accompanied by changes in the proportion of monocyte and neutrophil subsets in the peripheral blood, is a key factor in the strong immune response to SARS-CoV-2 (Guo et al. 2021; Ren et al. 2021).

Alarmins S100A8/A9 (also known as myeloid-related proteins 8 and 14, respectively) are immune-activating proteins mainly derived from myeloid cells under infection-induced inflammatory conditions and exist as heterodimeric complexes called calprotectin (Foell et al. 2007; Wang et al. 2018). Much attention has been given to the role of S100A8/A9 as an endogenous ligand of Toll-like receptor 4 (TLR4) and receptor for advanced glycation end product (RAGE) in cells of the myeloid lineage, including neutrophils, macrophages, and monocytes (Ometto et al. 2017). However, little is known about the role of S100A8/A9 in lymphoid lineage cells following COVID-19 infection. This study aimed to investigate how the proportion of lymphoid cells, including B cells, T cells, and NK cells, differs based on COVID-19 severity using previously published scRNA-seq data. We also aimed to study the characteristics of aberrant S100A8/A9<sup>High</sup> lymphoid cells in COVID-19 infection, considering their potential role in the pathogenesis of COVID-19.

## Materials and methods

### Data collection and data preprocessing

We obtained raw scRNA-seq data of patients with COVID-19 from the Gene Expression Omnibus (GEO) under accession numbers GSE158055 (Ren et al. 2021), which contains GSE145926 (Liao et al. 2020). The dataset included 95 patients with convalescence, 76 patients with disease progression, and 25 healthy controls. For sample types, 76 samples with disease progression consist of peripheral blood mononuclear cell (PBMC) samples (N = 47), bronchoalveolar lavage fluid (BALF) samples (N = 12), Sputum samples (N = 11) and sorted B cell samples (N = 6). We selected 47 PBMC samples from patients with disease progression. For samples from patients with progression data, the sampling

time (days post-symptom onset) varied from 1 day to 68 days in the 76 patients; therefore, we limited our analyses to samples obtained within 15-day post-symptom onset. Patients older than 61 years were excluded to reduce age-related effects in the analysis. We also discarded samples with extremely few lymphoid cells (T cells, B cells, and NK cells). We retained data from 12 mild and 7 severe COVID-19-infected patients. The severe condition was determined with clinical signs of pneumonia (fever, cough, dyspnea) plus one of the following: respiratory rate > 30 breaths/min; severe respiratory distress; or SpO<sub>2</sub> < 90% on room air. In order to analyze gene expression patterns throughout the progression of COVID-19 infection, this dataset was divided into two groups based on sampling time: within 6 days [early phase of COVID-19 infection; PBMC(E)] and after 6 days [late phase of COVID-19 infection; PBMC(L)]. We chose 6 days as the point of division by considering similarities in the patients' gene expression and the cell type profiles. This ultimately yielded 5 samples (3 mild cases and 2 severe cases) in the early phase and 14 samples (9 mild cases and 5 severe cases) in the late phase for use in this study (Table 1 and Supplementary Table 1).

We used Scanpy (Python package) (Wolf et al. 2018) for data preprocessing; cells were selected based on three conditions: (1) total number of genes expressed is more than 400 and less than 4000; (2) total number of unique molecular identifiers (UMI) is less than 15,000; and (3) mitochondrial gene percentage is less than 20%. Subsequently, counts per million (CPM) normalization and natural logarithm (using a pseudo-count of 1) were applied before the analysis of differentially expressed genes (DEG).

### Cell-type identification procedure

We combined previously described cell-type identification methods (Aran et al. 2019; de Kanter et al. 2019; Guo and Li 2021; Kim et al. 2021; Kiselev et al. 2018; Pliner et al. 2019), in which known marker genes or reference data were used to determine cell type. The cell-type identification procedure used in this study is illustrated in Supplementary Fig. 1. First, we predicted the cell type using cell type identifiers and then performed majority voting using the prediction results

from the cell type identifiers. We set the majority voting count threshold to 3, and cells that did not meet the threshold were excluded from the analysis. To identify the specific T cell type, we used known marker genes from a previous study for marker-based identifiers (Zhang et al. 2019) and annotated scRNA-seq data for reference-based identifiers (Puram et al. 2017; Stoeckius et al. 2017).

### Differentially expressed genes analysis and gene set enrichment analysis

We used Scanpy to identify differentially expressed genes (DEG). After performing the Wilcoxon rank-sum test, we selected genes with adjusted *p*-values (Benjamini–Hochberg correction) of less than 0.05. We used the web-based Metascape (Zhou et al. 2019) to assess the functional enrichment of the genes (up to 120 genes) differentially expressed between the S100A8/A9<sup>High</sup> and S100A8/A9<sup>Low</sup> lymphoid cell subsets in patients with severe COVID-19.

### Pro-inflammatory cytokine and chemokine expression in myeloid cells

The expression of various cytokines and chemokines, including *IL-1A*, *IL-1B*, *IL-6*, *IL-7*, *TNF-alpha*, *interferon type I* and *II*, and the chemokines *CCL2*, *CCL3*, and *CXCL10*, in myeloid cells such as monocytes and dendritic cells were analyzed (Arunachalam et al. 2020; Hadjadj et al. 2020; Lucas et al. 2020). Violin plots were constructed to visualize the expression of the cytokines (*TNF-alpha*, *IL1B*) and chemokines (*CXCL8*, *CCL5*, *CCL3*, *CCL3L1*, *CCL4*, *CCL4L2*). Uniform manifold approximation and projection (UMAP) were also used to visualize the expression of several genes of interest.

### Cell-to-cell interaction analysis

CellPhoneDB (Efremova et al. 2020) was used for cell-to-cell interaction analysis of cells expressing high levels of S100A8/A9. We added ligand–receptor interactions between S100A8/A9 and TLR4/RAGE to CellPhoneDB, as the original CellPhoneDB lacked these interactions. We then extracted ligand–receptor interactions with *p* < 0.05. Cell–cell interaction analysis was performed for each sample separately. Interactions commonly found in a sample group were identified, and the mean and *p*-value of these interactions were obtained by taking the mean and max of a sample group (i.e., mild and severe patient groups).

## Results

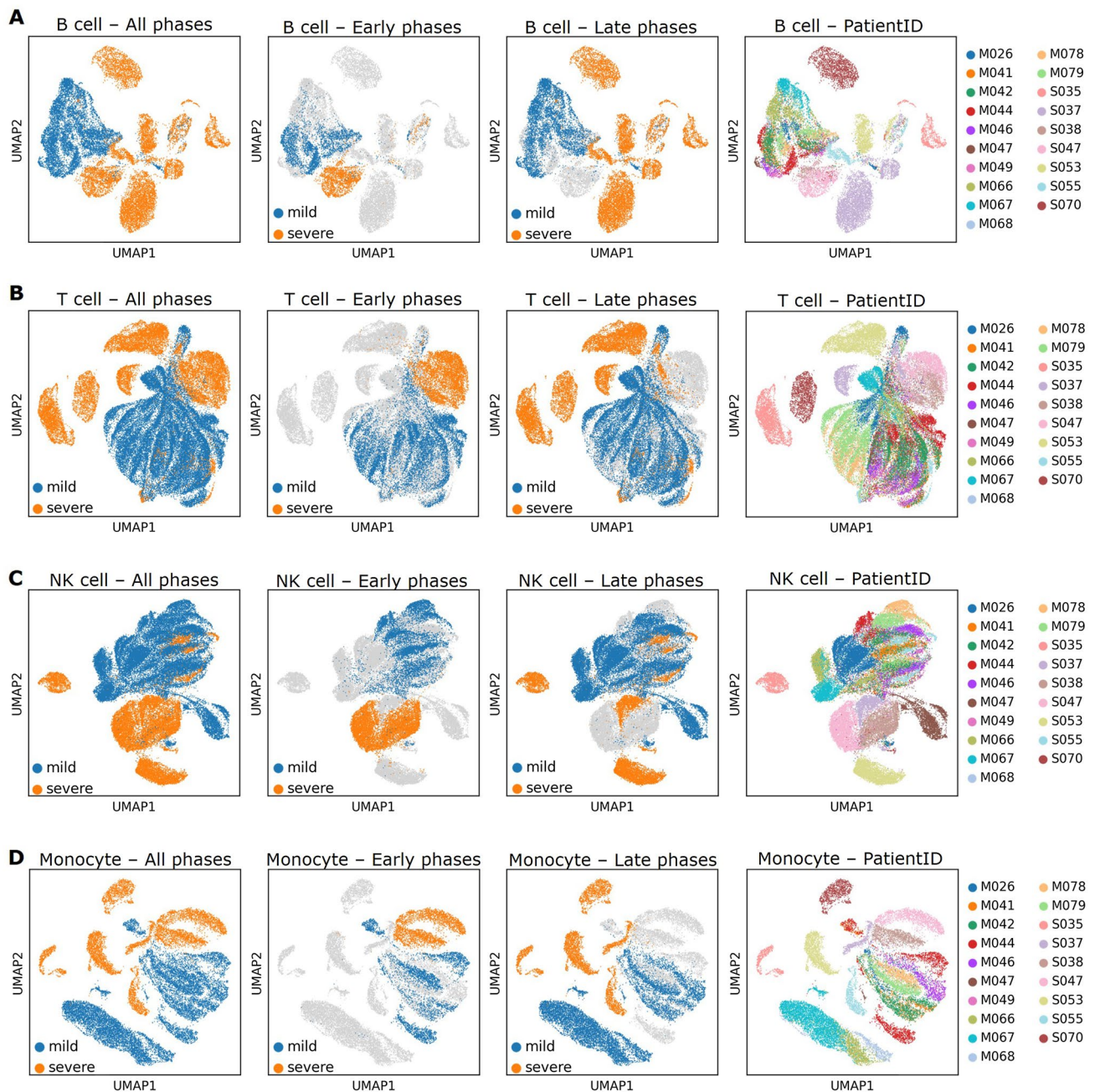
### Cell annotation according to cell-type specific marker genes

We used a combination of previously described cell-type annotation methods to annotate the dataset and verified our annotation using known marker genes. The annotated B cells expressed *CD79A*, of which plasma cells expressing the *MS4A1* gene were well-clustered (Supplementary Fig. 2A). The *CD3D* gene was expressed in almost all T cells. *CD4*, *CD8A*, and *CD8B* genes, important marker genes for classifying T cell subsets, were expressed in these T cell clusters (Supplementary Fig. 2B). Similarly, annotation of monocytes was verified based on the expression of *CD14* and *FCGR1A* and that of NK cells based on the expression of *KLRB1*, *NCRI*, *CD244*, and *ITGAM* (Supplementary Fig. 2C, 2D). We annotated 11 cell types and subsequently divided each cell type into two groups according to the expression of S100A8/A9. Neutrophils were excluded from the analysis because the data used in the analysis included many frozen PBMC samples; neutrophils are not commonly found in PBMC (He et al. 2019) and do not tolerate the cell freezing process (Wilk et al. 2020).

### Differences in B cells, T cells, and NK cells according to COVID-19 severity

We found that lymphoid cells (including B cells, T cells, and NK cells) in PBMC samples from patients with mild disease generally had similar gene expression patterns, while those from patients with severe disease displayed heterogeneous expression (Fig. 1a–c). The typical features of the lymphoid cells in patients with the severe disease tend to be definite in the late phase of COVID-19 infection. The gene expression patterns of monocytes in patients with mild disease did not demonstrate similarity to that of the lymphoid cells in these patients (Fig. 1d). To characterize the immune properties according to lymphoid cells in COVID-19 infection, we divided the PBMC samples into two groups based on sampling time: samples obtained at the early phase of infection (PBMC(E)) and samples obtained at the late phase of infection (PBMC(L)) (Fig. 2a). We compared the cell type proportions in patients with severe and mild disease. The cell-type proportions in the PBMC(E) samples were similar within each patient group. Patients with mild disease had more CD8<sup>+</sup> and CD4<sup>+</sup> T cells than patients with severe disease. In contrast, the natural killer (NK) cell subset was over-represented in patients with severe disease (Fig. 2b). The





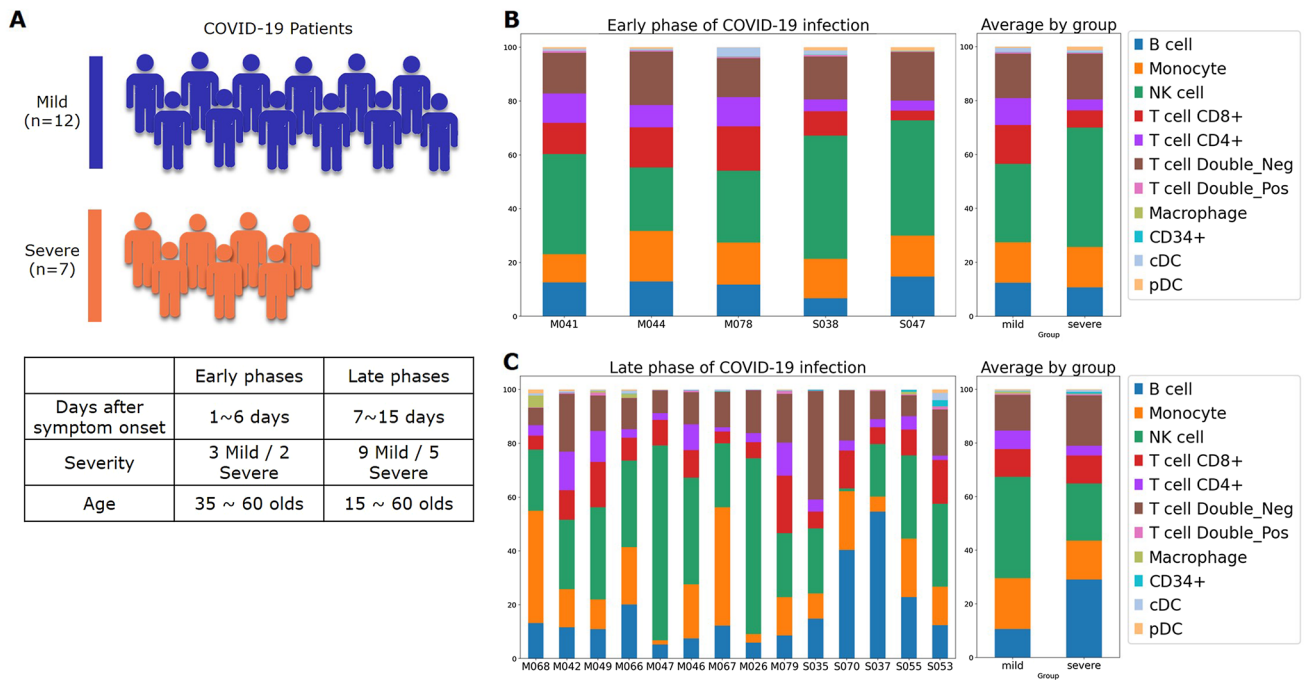
**Fig. 1** Lymphoid cell subset profiling from COVID-19 patients. **a** Uniform manifold approximation and projection (UMAP) of B cells in peripheral blood mononuclear cells (PBMCs), colored according

to disease severity, sampling time and patient. **b** UMAP of T cells in PBMCs, similar to **a**. **c** UMAP of NK cells in PBMCs, similar to **a**. **d** UMAP of monocytes in PBMCs, similar to **a**

cell-type proportions in PBMC(L) samples demonstrated heterogeneity even within the same patient group. The number of CD8<sup>+</sup> and CD4<sup>+</sup> T cells was higher in patients with mild disease than that in patients with severe disease, while the number of B cells was higher in patients with severe disease than that in patients with mild disease (Fig. 2c).

### Comparison of *S100A8/A9*<sup>High</sup> immune cells between mild and severe patients

The gene expression of *S100A8/A9* in lymphoid cells was significantly different between patients with severe and mild disease, whereas the expression of *S100A8/A9* in myeloid cells was not significantly different between the two patient



**Fig. 2** Cell-type proportions in COVID-19 patients. **a** Brief clinical information of COVID-19 patients. **b** Proportions of each cell type by patient in the early phase of COVID-19 infection. **c** Proportions of each cell type by patient in the late phase of COVID-19 infection

groups (Fig. 3a, b). Monocytes are divided into three subsets (classical, non-classical, and intermediate) according to the expression levels of *CD14* and *CD16* (*FCGR3A*) (Bassler et al. 2019; Kapellos et al. 2019). Of these, the non-classical  $CD14^{Low}CD16^{High}$  monocytes were more prevalent in patients with mild disease than those in patients with severe disease, similar to the findings of a previous study (Silvin et al. 2020). However, this finding was not consistent across all patients with mild disease. Further, non-classical monocytes were also present in some patients with severe disease. Therefore, the presence of non-classical monocytes alone was insufficient to distinguish patients with mild disease from those with severe disease. Although the classical  $CD14^{High}CD16^{Low}$  monocytes were the most common cell type among the monocyte subset, the difference in their numbers between the two patient groups was not significant.

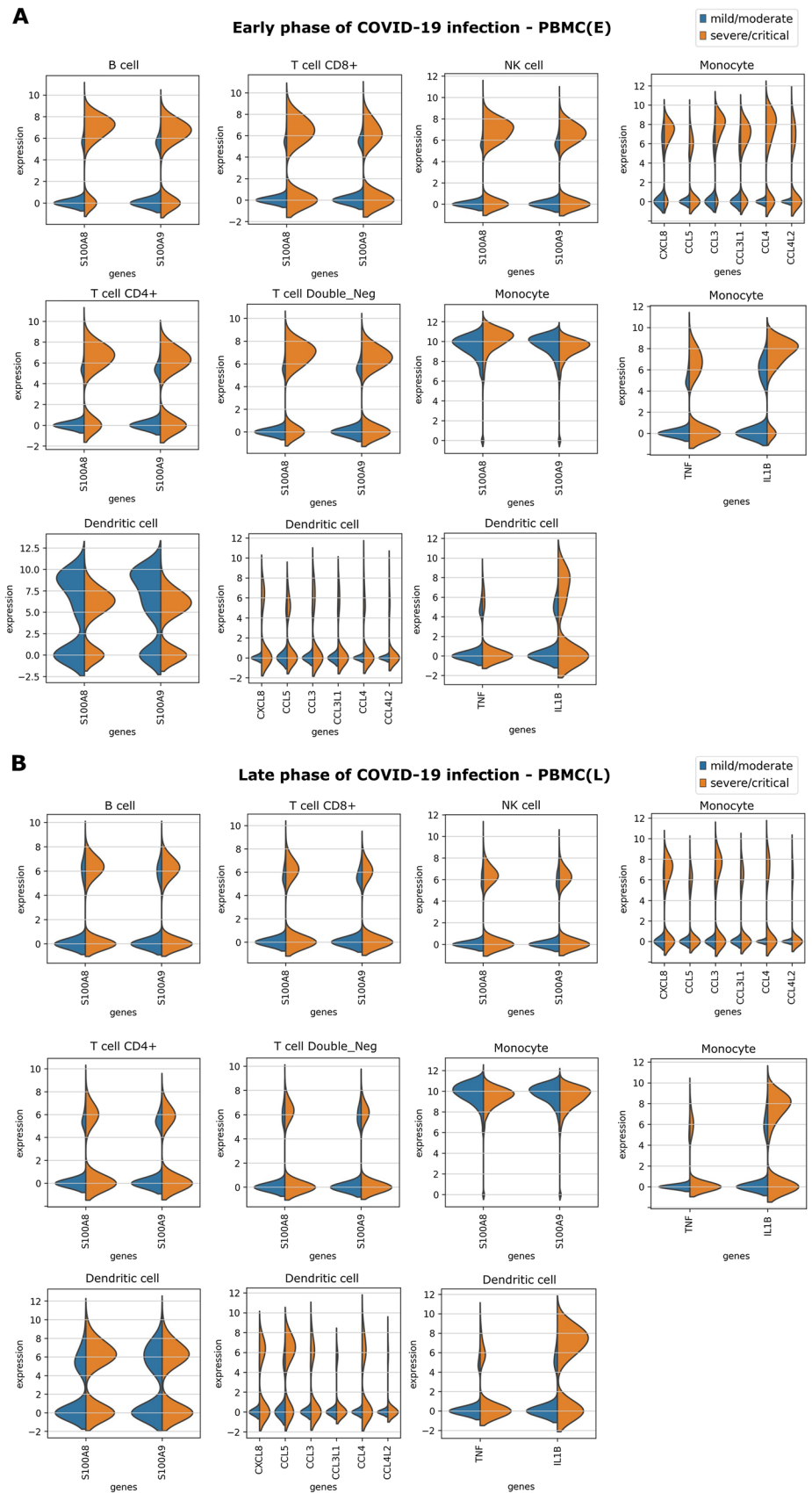
We observed that the expression level of *S100A8/A9* within monocytes was high in all COVID-19 patients, regardless of the monocyte subset or disease severity (Supplementary Fig. 3). Differentially expressed genes of monocytes between severe and mild patients included chemokines (*CXCL8*, *CCL5*, *CCL3*, *CCL3L1*, *CCL4*, and *CCL4L2*), which stimulate innate immune cells such as neutrophils. In addition, the expression of pro-inflammatory cytokines was higher in patients with severe disease than in those with mild disease (monocytes in Fig. 3a, b). These chemokines and cytokines were enriched in the

PBMC samples from all patients with severe disease. The *CXCL8*, *CCL5*, and *IL1B* expression in dendritic cells were higher in patients with severe disease than that in patients with mild disease. However, unlike monocytes, we found that *S100A8/A9* expression in dendritic cells differed based on disease severity (dendritic cells in Fig. 3a, b). Lymphoid cells and dendritic cells exhibited two *S100A8/A9* expression states:  $S100A8/A9^{High}$  and  $S100A8/A9^{Low}$ .

### Activation of cytokine/interferon signaling pathways in $S100A8/A9^{High}$ lymphoid cells

To gather further insights on the role of  $S100A8/A9^{High}$  lymphoid cells, we performed DEG analysis and gene set enrichment analysis (GSEA) comparing  $S100A8/A9^{High}$  versus  $S100A8/A9^{Low}$  T cells, NK cells, and B cells from patients with severe disease. GSEA demonstrated that cytokine signaling pathways were involved in  $S100A8/A9$ -positive lymphoid cells, regardless of the phase. Additionally, the interferon signaling pathway was highly activated, especially in the early phase of severe disease (PBMC(E) in Table 2). Therefore, our results support the hypothesis that dysregulated activation of  $S100A8/A9$ -positive lymphoid cells enhances aberrant immune activity during the early phase of COVID-19, leading to a cytokine storm.

**Fig. 3** Distribution of gene expression levels in each cell type by disease severity. **a** Violin plots of gene expression in in PBMC(E) samples from patients with severe and mild disease. **b** Violin plots of gene expression in in PBMC(L) samples from patients with severe and mild disease. *PBMC* peripheral blood mononuclear cells, *E* sampling at early phase of infection, *L* sampling at late phase of infection



**Table 2** Gene set enrichment analysis' result for lymphoid cells in severe patients

	Term	Description	Log-P val
PBMC(E) B cell	R-HSA-913531	Interferon signaling	– 11.2454
	R-HSA-1280215	Cytokine signaling in immune system	– 9.9253
	GO:0034340	Response to type I interferon	– 9.46425
	R-HSA-909733	Interferon alpha/beta signaling	– 9.28055
PBMC(E) T cell	R-HSA-168928	DDX58/IFIH1-mediated induction of interferon-alpha/beta	– 8.85912
	R-HSA-1280215	Cytokine signaling in immune system	– 20.7867
	R-HSA-913531	Interferon signaling	– 15.8979
	GO:0034097	Response to cytokine	– 14.005
PBMC(E) NK cell	R-HSA-909733	Interferon alpha/beta signaling	– 13.5656
	R-HSA-168928	DDX58/IFIH1-mediated induction of interferon-alpha/beta	– 11.1816
	GO:0045087	Innate immune response	– 13.3462
	R-HSA-1280215	Cytokine signaling in immune system	– 12.847
PBMC(L) B cell	GO:1903900	Regulation of viral life cycle	– 10.5171
	R-HSA-913531	Interferon signaling	– 10.2826
	R-HSA-909733	Interferon alpha/beta signaling	– 10.1349
	R-HSA-1280215	Cytokine signaling in Immune system	– 13.2888
PBMC(L) T cell	R-HSA-449147	Signaling by interleukins	– 11.5219
	GO:0071345	Cellular response to cytokine stimulus	– 10.5205
	GO:0034097	Response to cytokine	– 10.4709
	GO:0045087	Innate immune response	– 9.90851
PBMC(L) NK cell	R-HSA-1280215	Cytokine signaling in immune system	– 11.4421
	GO:0034097	Response to cytokine	– 10.5424
	R-HSA-449147	Signaling by interleukins	– 9.81936
	GO:0071345	Cellular response to cytokine stimulus	– 9.76216
PBMC(L) NK cell	WP4754	IL-18 signaling pathway	– 9.61301
	GO:0045087	Innate immune response	– 15.6762
	R-HSA-1280215	Cytokine signaling in immune system	– 14.329
	GO:0034097	Response to cytokine	– 13.2776
	GO:0071345	Cellular response to cytokine stimulus	– 12.5271
	R-HSA-449147	Signaling by interleukins	– 10.8223

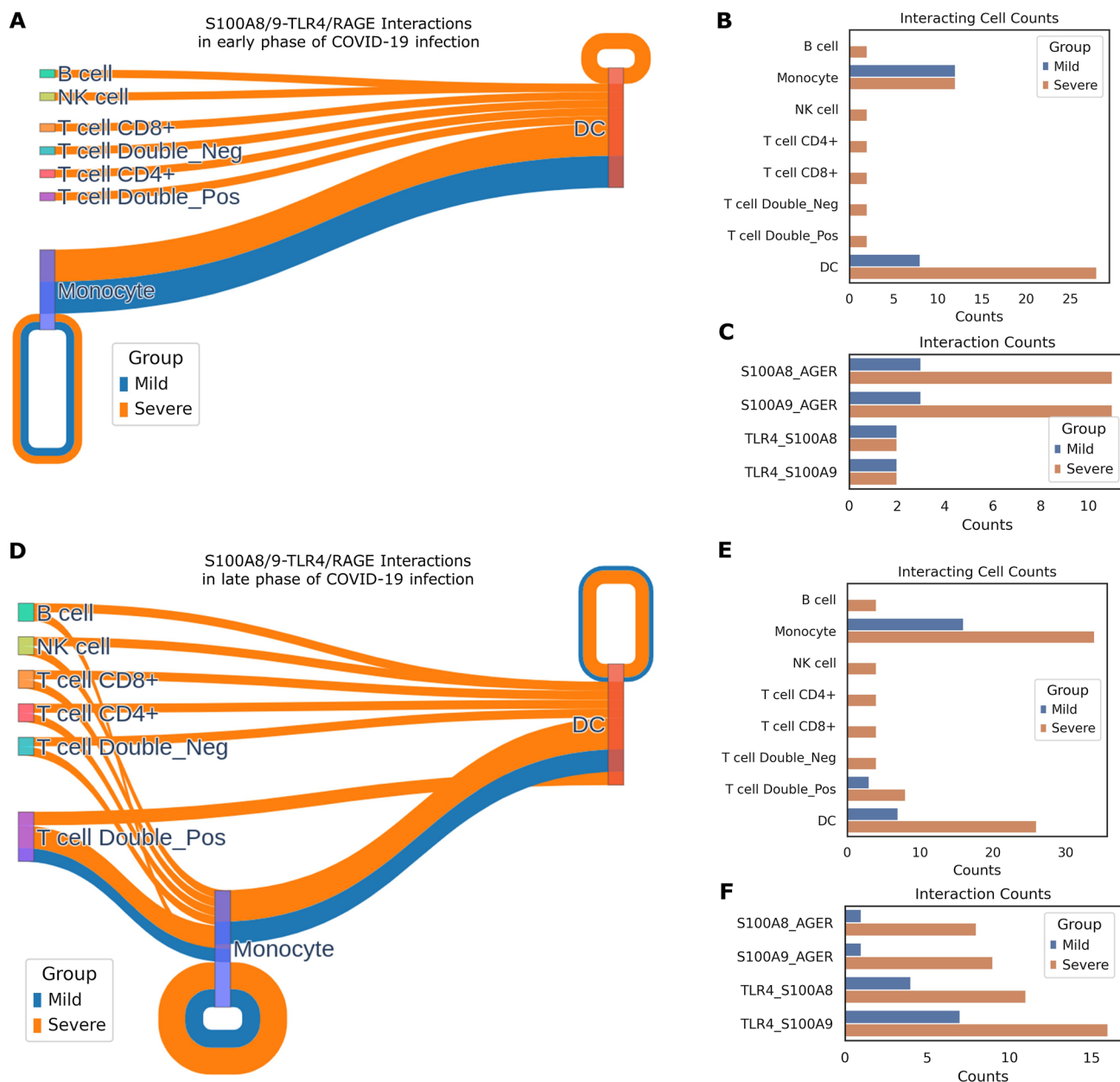
### S100A8/A9-TLR4/RAGE interactions in COVID-19 patients

We extracted ligand–receptor interactions to investigate the characteristics of S100A8/A9-TLR4/RAGE interactions in CellPhoneDB. Monocytes and dendritic cells from patients with severe and mild disease showed several common interactions (Fig. 4a, b, d, e). Despite the small number of patients with severe disease, more interactions were detectable in their samples. Lymphoid cell-associated outgoing signaling was observed only in the PBMC(E) samples from patients with severe disease, and this was linked to the incoming signaling from dendritic cells through RAGE (Fig. 4a, Supplementary Table 2). In addition, RAGE-associated incoming signaling was more dominant than TLR4 receptor-associated signaling in patients with severe disease (Fig. 4c). On the other hand, the outgoing signaling in patients with mild disease originated only from monocytes, despite interactions between monocytes

and dendritic cells being similar in patients with severe and mild disease (Fig. 4a, Supplementary Table 2).

In PBMC(L) samples, TLR4-associated incoming signaling was more dominant than the signaling through RAGE (Fig. 4f). Notably, the outgoing signaling from lymphoid cells was linked to both monocytes and dendritic cells (Fig. 4d, Supplementary Table 3). The total number of S100A8/A9-related interactions increased from 26 in the early phase to 43 in the late phase of the disease. These cell-to-cell interactions were more pronounced in patients with severe disease than those in patients with mild disease, indicating that S100A8/A9-related interactions increase over time post-infection in patients with severe disease.





**Fig. 4** Characterization of S100A8/A9-TLR4/RAGE interactions. **a** S100A8/A9-TLR4/RAGE interactions in PBMC(E) samples. Edge thickness indicates the sum of signals between populations [from outgoing (left) to incoming (right)], colored according to disease severity. **b** Cell counts for the entire S100A8/A9-TLR4/RAGE interaction in PBMC(E) samples. **c** Ligand–receptor counts for the entire S100A8/A9-TLR4/RAGE interaction in PBMC(E) samples

**d** S100A8/A9-TLR4/RAGE interactions in PBMC(L) samples, similar to **a**. **e** Cell counts for the entire S100A8/A9-TLR4/RAGE interaction in PBMC(L) samples. **f** Ligand–receptor counts for the entire S100A8/A9-TLR4/RAGE interaction in PBMC(L) samples. *PBMC* peripheral blood mononuclear cells, *E* sampling at early phase of infection, *L* sampling at late phase of infection, *TLR4* Toll-like receptor 4, *RAGE* receptor for advanced glycation end product

## Discussion

S100A8/A9, a  $\text{Ca}^{2+}$ -binding protein from the S100 family, has been identified as a possible biomarker for diagnosis and follow-up and a predictor of therapeutic responses in inflammation-related disorders. These proteins have recently received much attention as key alarmins involved

in the inflammatory response after COVID-19 infection. By engaging with TLR4 or RAGE, S100A8/A9-mediated NF- $\kappa$ B/MAPK activation generates intracellular signaling cascades, activating pro-inflammatory cytokines/chemokines for the recruitment and activation of inflammatory cells (Leclerc et al. 2009; Ma et al. 2017; Murray 2014; Nishikawa et al. 2017). The TLR4 pathway



was found to be more essential for S100A8/A9-mediated signaling than the RAGE pathway in mouse models of COVID-19 (Guo et al. 2021). Here, we identified key characteristics of lymphoid cells expressing high levels of S100A8/A9 based on the severity of COVID-19.

Compared to lymphoid cells that did not express S100A8/A9, S100A8/A9<sup>High</sup> lymphoid cells had a higher IFN or cytokine response. We found that S100A8/A9-positive lymphoid cells communicate with myeloid cells (monocytes and dendritic cells) via S100A8/A9-TLR4/RAGE interactions. These interactions were dominant at later stages of infection in patients with severe disease than in patients with mild disease. Generally, interferon signaling pathways are important for defending against invading viruses and clearing the infection; however, several studies have already shown that delayed or destructive IFN-I responses enhance viral invasion in the early phases of the COVID-19 (Cao 2020; Wen et al. 2020). Our study showed that S100A8/A9-positive lymphoid cells are implicated in the dysregulated interferon response during the early phases of COVID-19 infection, especially in severe disease.

In our analysis, pro-inflammatory and chemotaxis-related gene sets were upregulated in monocytes and dendritic cells from patients with severe disease. CellPhoneDB analysis suggests that S100A8/A9 secreted from lymphoid cells plays a role in activating monocytes and dendritic cells via the TLR4/RAGE pathway in severe disease even though the upregulation of these gene sets cannot be attributed solely to S100A8/A9-TLR4/RAGE interactions.

Our study has some limitations; it was difficult to acquire and compare immune cells common to all patients due to the variable proportions of annotated cell types. This can be attributed to a number of factors, including genetic predisposition, age, sex, effects from treatments in the course of COVID-19, and different sampling times. Therefore, larger sample size is required. Additionally, analyzing the same patient at multiple time points post-infection will enable a stronger comparison of the early and late stages of COVID-19 infection. A meta-analysis integrating multiple RNA-seq datasets from different studies would also provide reliable and reproducible results.

In conclusion, we discovered that lymphoid cells expressing high levels of S100A8/A9 play a significant role in the dysregulation of innate immune responses in patients with severe COVID-19. We hope that this study contributes to a better understanding of the COVID-19 infection-induced elevation of pro-inflammatory cytokines, known as the cytokine storm.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s13258-022-01285-2>.

**Acknowledgements** This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) and funded by the Ministry of Education, Science and Technology (NRF-2020R1F1A1066320 and NRF-2021R1A4A5032463).

**Author contributions** JL, SY, and SL conceptualized and designed the study; JL performed data analysis, including cell type annotation and statistical analyses, with which HK and MK assisted. SL curated the results and provided feedback. JL, SL and SY drafted the manuscript. All the authors wrote, read, and approved the manuscript.

**Data availability** The datasets analyzed during the current study are available in the GEO dataset (GSE158055) repository, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE158055>.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR et al (2019) Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* 20:163–172
- Arunachalam PS, Wimmers F, Mok CKP, Perera R, Scott M, Hagan T, Sigal N, Feng Y, Bristow L, Tak-Yin Tsang O et al (2020) Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science* 369:1210–1220
- Bassler K, Schulte-Schrepping J, Warnat-Herresthal S, Aschenbrenner AC, Schultze JL (2019) The myeloid cell compartment-cell by cell. *Annu Rev Immunol* 37:269–293
- Cao X (2020) COVID-19: immunopathology and its implications for therapy. *Nat Rev Immunol* 20:269–270
- de Kanter JK, Lijnzaad P, Candelli T, Margaritis T, Holstege FCP (2019) CHETAH: a selective, hierarchical cell type identification method for single-cell RNA sequencing. *Nucleic Acids Res* 47:e95
- Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R (2020) CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat Protoc* 15:1484–1506
- Foell D, Wittkowski H, Roth J (2007) Mechanisms of disease: a ‘DAMP’ view of inflammatory arthritis. *Nat Clin Pract Rheumatol* 3:382–390
- Guo H, Li J (2021) scSorter: assigning cells to known cell types according to marker genes. *Genome Biol* 22:69
- Guo Q, Zhao Y, Li J, Liu J, Yang X, Guo X, Kuang M, Xia H, Zhang Z, Cao L et al (2021) Induction of alarmin S100A8/A9 mediates activation of aberrant neutrophils in the pathogenesis of COVID-19. *Cell Host Microbe* 29:222–235 e224
- Hadjadj J, Yatim N, Barnabei L, Corneau A, Boussier J, Smith N, Pere H, Charbit B, Bondet V, Chenevier-Gobeaux C et al (2020) Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* 369:718–724
- He D, Yang CX, Sahin B, Singh A, Shannon CP, Oliveria JP, Gauvreau GM, Tebbutt SJ (2019) Whole blood vs PBMC: compartmental differences in gene expression profiling exemplified in asthma. *Allergy Asthma Clin Immunol* 15:67

- Kapellos TS, Bonaguro L, Gemund I, Reusch N, Saglam A, Hinkley ER, Schultze JL (2019) Human monocyte subsets and phenotypes in major chronic inflammatory diseases. *Front Immunol* 10:2035
- Kim H, Lee J, Kang K, Yoon S (2022) MarkerCount: A stable, count-based cell type identifier for single-cell RNA-seq experiments. *Comput Struct Biotechnol J* 20:3120–3132. PMID: 35782735
- Kiselev VY, Yiu A, Hemberg M (2018) scmap: Projection of single-cell RNA-seq data across data sets. *Nat Methods* 15:359–362
- Leclerc E, Fritz G, Vetter SW, Heizmann CW (2009) Binding of S100 proteins to RAGE: an update. *Biochim Biophys Acta* 1793:993–1007
- Liao M, Liu Y, Yuan J, Wen Y, Xu G, Zhao J, Cheng L, Li J, Wang X, Wang F et al (2020) Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat Med* 26:842–844
- Lucas C, Wong P, Klein J, Castro TBR, Silva J, Sundaram M, Ellingson MK, Mao T, Oh JE, Israelow B et al (2020) Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* 584:463–469
- Ma L, Sun P, Zhang JC, Zhang Q, Yao SL (2017) Proinflammatory effects of S100A8/A9 via TLR4 and RAGE signaling pathways in BV-2 microglial cells. *Int J Mol Med* 40:31–38
- Murray PJ (2014) Obesity corrupts myelopoiesis. *Cell Metab* 19:735–736
- Nishikawa Y, Kajiura Y, Lew JH, Kido JI, Nagata T, Naruishi K (2017) Calprotectin induces IL-6 and MCP-1 production via Toll-like receptor 4 signaling in human gingival fibroblasts. *J Cell Physiol* 232:1862–1871
- Nishiura H, Kobayashi T, Miyama T, Suzuki A, Jung SM, Hayashi K, Kinoshita R, Yang Y, Yuan B, Akhmetzhanov AR et al (2020) Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *Int J Infect Dis* 94:154–155
- Ometto F, Friso L, Astorri D, Botsios C, Raffeiner B, Punzi L, Doria A (2017) Calprotectin in rheumatic diseases. *Exp Biol Med (Maywood)* 242:859–873
- Pliner HA, Shendure J, Trapnell C (2019) Supervised classification enables rapid annotation of cell atlases. *Nat Methods* 16:983–986
- Puram SV, Tirosh I, Parkh AS, Patel AP, Yizhak K, Gillespie S, Rodman C, Luo CL, Mroz EA, Emerick KS et al (2017) Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* 171:1611 e1624–1624 e1624
- Ren X, Wen W, Fan X, Hou W, Su B, Cai P, Li J, Liu Y, Tang F, Zhang F et al (2021) COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell* 184:1895 e1819–1913 e1819
- Silvin A, Chapuis N, Dunsmore G, Goubet AG, Dubuisson A, Derosa L, Almire C, Henon C, Kosmider O, Droin N et al (2020) Elevated calprotectin and abnormal myeloid cell subsets discriminate severe from mild COVID-19. *Cell* 182:1401 e1418–1418 e1418
- Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, Satija R, Smibert P (2017) Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* 14:865–868
- Wang S, Song R, Wang Z, Jing Z, Wang S, Ma J (2018) S100A8/A9 in inflammation. *Front Immunol* 9:1298
- Wen W, Su W, Tang H, Le W, Zhang X, Zheng Y, Liu X, Xie L, Li J, Ye J et al (2020) Immune cell profiling of COVID-19 patients in the recovery stage by single-cell sequencing. *Cell Discov* 6:31
- Wilk AJ, Rustagi A, Zhao NQ, Roque J, Martinez-Colon GJ, McKechnie JL, Ivison GT, Ranganath T, Vergara R, Hollis T et al (2020) A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nat Med* 26:1070–1076
- Wolf FA, Angerer P, Theis FJ (2018) SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* 19:15
- Yang X, Yu Y, Xu J, Shu H, Xia J, Liu H, Wu Y, Zhang L, Yu Z, Fang M et al (2020) Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 8:475–481
- Zhang X, Lan Y, Xu J, Quan F, Zhao E, Deng C, Luo T, Xu L, Liao G, Yan M et al (2019) CellMarker: a manually curated resource of cell markers in human and mouse. *Nucleic Acids Res* 47:D721–D728
- Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, Wang X, Yuan J, Li T, Li J et al (2020) Antibody responses to SARS-CoV-2 in patients with novel coronavirus disease 2019. *Clin Infect Dis* 71:2027–2034
- Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 10:1523

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.