# Multi-layered transcriptomic analyses reveal an immunological overlap between COVID-19 and hemophagocytic lymphohistiocytosis associated with disease severity

Lena F. Schimke[a,5], Alexandre H.C. Marques[a], Gabriela Crispim Baiocchi[a], Caroline Aliane de Souza Prado[b], Dennyson Leandro M. Fonseca[b], Paula Paccielli Freire[a], Desirée Rodrigues Plaça[b], Igor Salerno Filgueiras[a], Ranieri Coelho Salgado[a], Gabriel Jansen-Marques[c], Antonio Edson Rocha Oliveira[b], Jean Pierre Schatzmann Peron[a], José Alexandre Marzagão Barbuto[a,d], Niels Olsen Saraiva Camara[a], Vera Lúcia Garcia Calich[a], Hans D. Ochs[e], Antonio Condino-Neto[a], Katherine A. Overmyer[f,g], Joshua J. Coon[h,i], Joseph Balnis[j,k], Ariel Jaitovich[j,k], Jonas Schulte-Schrepping[l], Thomas Ulas[m], Joachim L. Schultze[l,m], Helder I. Nakaya[b], Igor Jurisica[n,o,p], Otavio Cabral-Marques[a,b,q]

[a]Departamento de Imunologia, Instituto de Ciencias Biomedicas, Universidade de São Paulo, São Paulo, SP, Brazil. ID Ringgold USP + ICB: 54544, ISNI: 0000000406355304
[b]Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, SP, Brazil.
[c]Information Systems, School of Arts, Sciences and Humanities, University of Sao Paulo, São Paulo, SP, Brazil.
[d]Laboratory of Medical Investigation in Pathogenesis and targeted therapy in Onco-Immuno-Hematology (LIM-31), Department of Hematology, Hospital das Clínicas HCFMUSP, Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, Brazil.
[e]Department of Pediatrics, University of Washington School of Medicine, and Seattle Children's Research Institute, Seattle, WA, USA.
[f]National Center for Quantitative Biology of Complex Systems, Madison, WI 53562, USA
[g]Morgridge Institute for Research, Madison, WI 53562, USA
[h]Department of Biomolecular Chemistry, University of Wisconsin, Madison, WI 53506, USA Department of [i]Chemistry, University of Wisconsin, Madison, WI 53506, USA.
[j]Division of Pulmonary and Critical Care Medicine, Albany Medical Center, Albany, NY 12208, USA
[k]Department of Molecular and Cellular Physiology, Albany Medical College, Albany, NY 12208, USA
[l]Life and Medical Sciences (LIMES) Institute, University of Bonn, Germany.
[m]German Center for Neurodegenerative Diseases (DZNE), PRECISE Platform for Genomics and Epigenomics at DZNE, and University of Bonn, Bonn, Germany.
[n]Osteoarthritis Research Program, Division of Orthopedic Surgery, Schroeder Arthritis Institute and Data Science Discovery Centre for Chronic Diseases, Krembil Research Institute, University Health Network, Toronto, Canada.
[o]Departments of Medical Biophysics and Computer Science, University of Toronto, Toronto, Canada.
[p]Institute of Neuroimmunology, Slovak Academy of Sciences, Bratislava, Slovakia.
[q]Network of Immunity in Infection, Malignancy, and Autoimmunity (NIIMA), Universal Scientific Education and Research Network (USERN), Sao Paulo, Brazil.
[5]Lead contact

**Corresponding authors:**

Lena F. Schimke, MD
Department of Immunology
Institute of Biomedical Sciences - University of São Paulo
Lineu Prestes Avenue, 1730,
São Paulo, SP, CEP 05508-900, Brazil
e-mail: lenaschimke@hotmail.com
phone.: +55 (11) 3091-7387

Otavio Cabral-Marques, MSc, PhD
Department of Immunology
Institute of Biomedical Sciences - University of São Paulo
Lineu Prestes Avenue, 1730,
São Paulo, SP, CEP 05508-900, Brazil
e-mail: otavio.cmarques@usp.br
phone: +55 11 974642022

**ABSTRACT**

Clinical and hyperinflammatory overlap between COVID-19 and hemophagocytic lymphohistiocytosis (HLH) has been reported. However, the underlying mechanisms are unclear. Here we show that COVID-19 and HLH have an overlap of signaling pathways and gene signatures commonly dysregulated, which were defined by investigating the transcriptomes of 1253 subjects (controls, COVID-19, and HLH patients) using microarray, bulk RNA-sequencing (RNAseq), and single-cell RNAseq (scRNAseq). COVID-19 and HLH share pathways involved in cytokine and chemokine signaling as well as neutrophil-mediated immune responses that associate with COVID-19 severity. These genes are dysregulated at protein level across several COVID-19 studies and form an interconnected network with differentially expressed plasma proteins which converge to neutrophil hyperactivation in COVID-19 patients admitted to the intensive care unit. scRNAseq analysis indicated that these genes are specifically upregulated across different leukocyte populations, including lymphocyte subsets and immature neutrophils. Artificial intelligence modeling confirmed the strong association of these genes with COVID-19 severity. Thus, our work indicates putative therapeutic pathways for intervention.

**Keywords:**

COVID-19, Hemophagocytic lymphohistiocytosis, common signaling pathways, gene signatures

## INTRODUCTION

More than one year of Coronavirus disease 2019 (COVID-19) pandemic caused by the severe acute respiratory syndrome Coronavirus (SARS-CoV)-2, more than 197 million cases and 4,2 million deaths have been reported worldwide (July 30th 2021, WHO COVID-19 Dashboard). The clinical presentation ranges from asymptomatic to severe disease manifesting as pneumonia, acute respiratory distress syndrome (ARDS), and a life-threatening hyperinflammatory syndrome associated with excessive cytokine release (hypercytokinaemia)[1–3]. Risk factors for severe manifestation and higher mortality include old age as well as hypertension, obesity, and diabetes[4]. Currently, COVID-19 continues to spread, new variants of SARS-CoV-2 have been reported and the number of infections resulting in death of young individuals with no comorbidities has increased the mortality rates among the young population[5,6]. In addition, some novel SARS-CoV-2 variants of concern appear to escape neutralization by vaccine-induced humoral immunity[7]. Thus, the need for a better understanding of the immunopathologic mechanisms associated with severe SARS-CoV-2 infection.

Patients with severe COVID-19 have systemically dysregulated innate and adaptive immune responses, which are reflected in elevated plasma levels of numerous cytokines and chemokines including granulocyte colony-stimulating factor (GM-CSF), tumor necrosis factor (TNF), interleukin (IL)-6, IL-6R, IL18, CC chemokine ligand 2 (CCL2) and CXC chemokine ligand 10 (CXCL10)[8–10], and hyperactivation of lymphoid and myeloid cells[11]. Notably, the hyperinflammation in COVID-19 shares similarities with cytokine storm syndromes such as those triggered by sepsis, autoinflammatory disorders, metabolic conditions and malignancies[12–14], often resembling a hematopathologic condition called hemophagocytic lymphohistiocytosis (HLH)[15]. HLH is a life-threatening progressive systemic hyperinflammatory disorder characterized by multi-organ involvement, fever flares, hepatosplenomegaly, and cytopenia due to hemophagocytic activity in the bone marrow[15–17] or within peripheral lymphoid organs such as pulmonary lymph nodes and spleen. HLH is marked by aberrant activation of B and T lymphocytes and monocytes/macrophages, coagulopathy, hypotension, and ARDS. Recently, neutrophil hyperactivation has been shown to also play a critical role in HLH development[18,19]. This is in agreement with the observation that the HLH-like phenotype observed in severe COVID-19 patients is due to an innate neutrophilic hyperinflammatory response associated with

virus-induced hypercytokinaemia which is dominant in patients with an unfavorable clinical course[17]. Thus, HLH has been proposed as an underlying etiologic factor of severe COVID-19[1,3,20]. HLH usually develops during the acute phase of COVID-19[1,20–27]. However, a case of HLH that occurred two weeks after recovery from COVID-19 has recently been reported as the cause of death during post-acute COVID-19 syndrome[28].

The familial form of HLH (fHLH) is caused by inborn errors of immunity (IEI) in different genes encoding proteins involved in granule-dependent cytotoxic activity of leukocytes such as AP3B1, LYST, PRF1, RAB27A, STXBP2, STX11, UNC13D[29–31]. In contrast, the secondary form (sHLH) usually manifests in adults following a viral infection (e.g., adenovirus, EBV, enterovirus, hepatitis viruses, parvovirus B19, and HIV)[32,33], or in association with autoimmune/rheumatologic, malignant, or metabolic conditions that lead to defects in T/NK cell functions and excessive inflammation[16,31]. fHLH and sHLH affect both children and adults, however, the clinical and genetic distinction of HLH forms is not clear since immunocompetent children can develop sHLH [34,35], while adult patients with sHLH may also have germline mutations in HLH genes[36]. Of note, germline variants in *UNC13D* and *AP3B1* have also been identified in some COVID-19 patients with HLH phenotype[37], thus, indicating that both HLH forms may be associated with COVID-19.

Here, we characterized the signaling pathways and gene signatures commonly dysregulated in both COVID-19 and HLH patients by investigating the transcriptomes of 1253 subjects (controls, COVID-19, and HLH patients) assessed by microarray, bulk RNA-sequencing (RNAseq), and single-cell RNAseq (scRNAseq) (**Table 1**). We found shared gene signatures and cellular signaling pathways involved in cytokine and chemokine signaling as well as neutrophil-mediated immune responses that associate with COVID-19 severity.

## RESULTS

### The transcriptional overlap between COVID-19 and HLH

Considering the fact that COVID-19 and HLH[14,21] share clinical and biological features (**Figure 1A**) we suspected that these diseases have a transcriptional overlap underlying their phenotypic convergence. First, we obtained differentially expressed genes (DEGs) for each dataset from different high-throughput transcriptome technologies present in peripheral blood lymphocytes (PBLs), peripheral blood mononuclear cells (PBMCs), and nasopharyngeal swabs from COVID-19 patients, HLH patients and controls (**Figure 1B** and **Supp. Table S1**). Then, we defined the transcriptome overlap between DEGs from COVID-19 cohorts and HLH by cross-technology comparisons, performed enrichment analysis and association studies between specific DEGs and severity status of disease (**Figure 1C**). To identify the common DEGs we divided the datasets into three subgroups based on type of samples and RNAseq platforms: Overlap 1 (HLH and COVID-19 blood transcriptomes), Overlap 2 (HLH and COVID-19 nasopharyngeal swab transcriptomes), and Overlap 3 (HLH and COVID-19 scRNAseq transcriptomes) (**Supp. Figure 1A** and **1B,** and **Supp. Table S2** and **S3).** We found a total of 239 unique common DEGs between HLH and all COVID-19 datasets, most of them (237 DEGs) up-regulated (**Figure 1D)**. Hereafter, we focused on the implications of the up-regulated genes, since the 2 common down-regulated genes (granulysin or *GNLY;* myomesin 2 *or MYOM2*) alone did not enrich any significant pathway. However, this might also indicate a defect in cytotoxicity activity, typical of HLH[31], that will require future investigation. The 237 common up-regulated DEGs encode proteins mainly involved in immune system, metabolic and signaling processes, forming a highly connected biological network based on physical protein-protein interactions (PPI, **Figure 1E**). Among them are important genes encoding molecules involved in activation of inflammatory immune responses (e.g., PGLYRP1, OLR1, FFAR2), cytokine and chemokine mediated immune pathways (e.g., IL1R2, CXCR2, CXCR8, CCL4, CCL2), and neutrophil activation (e.g., CD177, MPO, ELANE). Of note, the transcriptional overlap between HLH and COVID-19 contains several molecules interacting with 7 genes causing fHLH due to IEI which itself were not among our DEGs (**Figure 1E**).

**Cytokine/chemotaxis and neutrophil signatures predominate in COVID-19 and HLH**

We next dissected the biological functions enriched by the 237 common up-regulated DEGs between COVID-19 and HLH patients by performing enrichment analysis of biological processes (BPs) and cellular components (CCs) by these 237 DEGs. The top 20 most enriched BPs are demonstrated in **Figure 2A,** which encompass cytokine/chemotaxis and neutrophil-mediated innate immune responses, ranging from response to IL-1 to neutrophil activation, degranulation, and migration (for all BPs see **Supp. Table S4**). The CCs enriched (**Figure 2B**) include several compartments such as secretory granule lumen and membrane, azurophil tertiary and specific granules, as well as collagen-containing extracellular matrix, phagocytic vesicle, and primary lysosome (**Supp. Table S5**).

Cytokine/chemotaxis and neutrophil signatures predominate in the COVID-19 and HLH multi-layered transcriptional overlap. A total of 25, 34, and 58 DEGs are assigned to cytokine, chemotaxis, and neutrophil signatures, respectively (**Figure 2C-E:** complete categorization can be seen in **Supp. Table S6 and S7**). Several genes play pleiotropic roles in these gene ontology (GO) categories such as *CEACAM8, IL-1β*, *IL-6*, *EDN1*, *NFKB1* and *PDE4B* (**Supp. Table S8**). For clarity in data visualization, we assigned these genes to a unique category (based on their predominant immunological function according to literature and GeneCards[38]; the human gene database). Among these genes, there are chemokines and chemokine receptors that attract both lymphocytes and myelocytes to inflammation sites (CCL20, CCL2, CXCR1, CXCR3, CXCL8)[39,40], pro-inflammatory cytokines (IL-1B, IL-1R1, NFKB1, IFNG, IL-6, TNF)[41,42] that promote the activation of immune cells, and several proteins/granules with antimicrobial activity (MPO[43], AZU1[44], ELANE[45,46], DEFA4[47]). Moreover, there are metalloproteinases (MMP8 and MMP9) involved in degradation of extracellular matrix (ECM) for neutrophil migration[8,49] into the airways and in the regulation of cytokine activity. Of note, hierarchical clustering analysis of these genes indicated a cross-study grouping of closer functional-related molecules. For instance, *IFNG*, *IL6*, and *TNF*; *IL1A* and *IL1B*; as well as signaling molecules involved in the nuclear factor-κB (NF-κB) such as *NFKB1*, *SPHK1*, and *RIPK2* clustered together in the cytokine group. Likewise, *GYPA* and *GYPB*; *CEACAM6* and *CEACAM8*; as well as *CCL* and *CXCL* chemokines

in the chemotaxis group and antimicrobial-related peptides such as *AZU1*, *MPO*, *CAMP*, *DEFA4*, *LCN2*, *ELANE*, *OLFM4*, and *CD177* in neutrophil-mediated immunity. However, we cannot exclude that this clustering pattern just represent an aleatory tendency due to upstream GO categorization.

Of note, the classical literature has emphasized that the 7 genes (*AP31B, LYST, PRF1, RAB27A, STX11, STXBP2, UNC13D*) known to cause fHLH (classically defined as familial HLH syndromes and hypopigmentation syndromes)[14,29] suggest a dominant role for T and NK cells in the development of HLH[50–52]. However, these genes also enrich several CCs (secretory vesicles, azurophilic granules or specific granules) and BPs (neutrophil degranulation) involved in the neutrophil immune responses (**Supp. Figure 2).** This result is in agreement with the role of these genes in a variety of neutrophil functions such as degranulation and formation of neutrophil extracellular traps (NET)[53–57]. Altogether, these data indicate gene signatures and signaling pathways commonly dysregulated in COVID-19 and HLH patients that involve a network of regulatory (production of chemokines/cytokines by T cells and macrophages) and effector (neutrophil hyperactivation, NK and T cell cytotoxicity) immune functions.

**The relationship between cytokine/chemotaxis and neutrophil-mediated immunity gene signatures**

We next analyzed the relationship pattern and degree between the transcriptional signatures related to cytokine signaling and chemotaxis with those involved in neutrophil-mediated immune responses. We chose the COVID-19_PBL dataset from Overmyer et al.[58], which contains transcripts from 100 individuals with COVID-19 and 26 individuals with respiratory symptoms but negative for COVID-19 serving as control group (further explored in the next session). We performed canonical-correlation analysis (CCA), which is a multivariate statistical method to determine the linear relationship between two groups of variables[59]. In accordance with the cross-study hierarchical clustering, CCA revealed a strong association between several cytokine/chemotaxis related genes (e.g., *CXCL8, CEACAMs [1/6/8], IL1RAP, IL1R1, IL1B, NFKB1*) with those involved in neutrophil-mediated immune responses (e.g., *CTSG, ELANE, MMP8, TCN1*) in both patients with COVID-19 and controls (**Figure 3A** and **3B)**. Bivariate correlation

analysis showed a similar phenomenon (**Supp. Figure 3)**. However, these correlation patterns partially changed when comparing COVID-19 with the control group. For instance, while reducing the correlation between molecules including *IL-10, CXCL8, NFKB1, ARG1*, and *SOD2*, new strong associations appeared between *ELANE, DEFA4, AZU1, CTSG*, and *LCN2*, with an overall tendency to higher relationships amid neutrophil-mediated immunity related genes in COVID-19 patients. **Figure 3C** illustrates this observation by scatter plots for some of these variables.

**Transcripts stratifying severe COVID-19 from other respiratory diseases are also highly dysregulated in HLH**

Next, we sought to determine which genes of chemotaxis/cytokine signaling and neutrophil-mediated immune responses discriminate COVID-19 patients according to disease severity. We further investigated the COVID-19_PBL dataset (GSE157103)[58] comparing COVID-19 patients admitted to the intensive care unit (COVID-19_ICU) with those admitted to non-ICU units (COVID-19_nonICU). Severity of critical illness at ICU admission was defined based on APACHE II and SOFA scores[60] according to Overmyer et al.[58] (**Figure 4A**). Among all genes, 25 (15 up-regulated and 10 down-regulated genes) were differentially expressed between COVID-19_ICU and COVID19_nonICU patients (**Supp. Table S9**). These 25 genes were also present and up-regulated in peripheral blood leukocytes of COVID-19 patients from two recently published works, which were unpublished when we started our investigation (GSE163151[61] and GSE152641[62], **Supp Table S10**). Of note, most of these 25 genes have also been identified at protein level as dysregulated in COVID-19 patients across different studies (published during the development of our study; **Supp. Table S11**). In addition, these 25 genes might belong to a systemic immune network of molecules induced by SARS-CoV-2 since they are also highly interconnected with 158 proteins (**Supp. Table 12**) significantly dysregulated in the plasma of COVID-19_ICU when compared to COVID-19_nonICU patients. Thus, they show several interactions and functional overlap (**Figure 4B)** with plasma proteins involved in neutrophil degranulation and neutrophil-mediated immunity (**Supp. Figure 4)**.

To investigate the stratification power of these 25 DEGs, we performed principal component analysis (PCA) using a spectral decomposition approach[63,64], which examines the covariances/correlations between variables. This approach revealed that these DEGs clearly divide COVID-19_ICU, COVID-19_nonICU, Control_ICU and Control_nonICU (due to other respiratory illness but negative for SARS-CoV-2) groups (**Figure 4C** and **Supp. Figure 5A and 5C)**. Likewise, these 25 genes stratified HLH patients from healthy controls (**Figure 4D** and **Supp. Figure 5B and 5D)**. As seen in these PCA graphics, 8 (*AZU1, CEACAM8, CTSG, DEFA4, ELANE, LCN2, OLFM4, and MMP8*) of these 25 DEGs strongly associate with COVID-19_ICU. Bivariate correlation analysis based on these 25 genes showed that while controls and COVID-19_nonICU patients have a similar general cluster distribution, COVID-19_ICU patients tend to differ of them, maintaining only the 8 genes with high positive correlations (**Figure 4E** and **Figure 4F)**. Combined, these results suggest that these genes are associated with COVID-19 severity and highlight the importance of neutrophil-mediated immunity related molecules in severe COVID-19 on both transcriptomic and proteomic level.

**Multi-layered transcriptomic analysis associates COVID-19 and HLH common genes with disease severity**

Since scRNAseq allows comparison of the transcriptomes of individual cells, we next sought to investigate the distribution patterns of these 25 genes associated with COVID19 severity. We analyzed the scRNAseq dataset (EGAS00001004571) reported by Schulte-Schrepping et al.[65] (schematic overview of study group **Figure 5A**) and found that 21 of the 25 genes associated with COVID-19 severity and HLH development are DEGs among the top 2,000 variable genes in the COVID-19 cohort compared to controls (**Figure 5B** and **Supp. Fig 6A and B**). These 21 genes exhibited cell-type-specific expression patterns. For instance, CCL4 (a chemoattractant and stimulator of T-cell immune responses[66,67]) was mainly produced by CD8+ T and NK cells, CD83 (B, T and dendritic cell activation marker[68,69]) by B cells and monocytes. CXCL8 was mostly present in monocytes and low-density neutrophils/granulocytes (LDGs; also frequently reported as immature neutrophils[70–72]), which are neutrophils remaining in the PBMC fraction after density gradient separation. Among these 21 genes, 11 genes (among them also the 8 genes

described above) were differentially expressed when comparing patients with mild and severe COVID-19 (**Figure 5C**). Thus, indicating a network of cell-type-specific expression patterns that may contribute to the clinical similarities between COVID-19 and HLH.

Of note, these 11 genes encode proteins that are crucial for several pathways involved in neutrophil-mediated immunity, and are associated with diseases that increase the risk of severe COVID-19[73,74] such as chronic obstructive pulmonary disease (COPD)[75,76] and ulcerative colitis[77,78] (**Supp. Figure 6E** and **Supp. Table S14**). Similarly, these 11 genes are also significantly different between COVID-19_ICU and COVID-19_nonICU (**Figure 5E**) in the bulk RNAseq dataset (GSE157103, Overmyer et al. 2020[58]). Indicating that these genes are consistently associated with COVID-19 severity across different patient cohorts. Moreover, these 11 genes were differentially expressed in HLH patients compared to healthy controls (**Figure 5D**), reinforcing the transcriptomic overlap of severe COVID-19 and HLH.

We used random forest method[79] to rank the importance of these 11 genes based on their ability to discriminate between COVID-19_ICU and COVID-19_nonICU, hence evaluating the association of these genes with the COVID-19 severity. This approach showed error rate (out of bag or OOB) of 27,03% and an area under the ROC curve of 82,4% for both groups (**Figure 6A** and **6B**). Follow-up analysis indicated that ARG1 was the most significant predictor for ICU admission followed by CD177, MCEMP1, LCN2, AZU1, OLFM4, MMP8, ELANE, CTSG, DEFA4, CEACAM8 based on the number of the nodes, gini-decrease, and average depth criteria for measuring gene importance (**Figure 6C** and **6D**). ARG1 exhibited the most relevant interactions with the other genes according to the mean minimal depth criterion, mostly interacting with CD177, AZU1, MCEMP1, and LCN2 (**Figure 6D)**. Altogether, these multi-layered transcriptomic results associate COVID-19 and HLH common genes with disease severity.

## DISCUSSION

The results of our study provide a COVID-19/HLH immune landscape using a multi-layered transcriptomic approach of microarray, bulk and scRNAseq. Our meta-analysis integrates and unravels the consistency of several important individual studies and datasets that also validated the transcriptome data at the protein level in COVID-19 patients[58,65,80]. To the best of our

knowledge, it represents the first attempt to systems characterize the common signaling pathways and molecular networks shared by COVID-19 and HLH. In agreement with the recent observation that neutrophil hyperactivation plays a key role in the severity of COVID-19[81–84] and HLH[18,19], our approach indicates that COVID-19 and HLH have a common transcriptional profile formed predominantly by a group of regulatory molecules related to cytokine/chemotaxis and by a group of effector molecules that are linked to neutrophil hyperactivation and disease severity. These data highlight the dual role of neutrophils in providing essential antimicrobial functions, but also initiating tissue injury caused by immune dysregulation[85,86]. The genes associated with COVID-19 severity are up-regulated across different leukocyte subpopulations such as lymphoid (NK, T and B cells) and myeloid (monocytes, dendritic cells and LDGs) cells. They form a systemic and interconnected network of cell-type-specific expression pattern and signaling networks that may contribute to the clinical similarities between COVID-19 and HLH. Thus, our analysis identified new candidate biomarkers and novel putative molecular pathways for therapeutic intervention for COVID-19 and HLH.

Our work expands the efforts of others[87–91] and our group[92] to identify networks and pathways involved in the pathogenesis of severe COVID-19. Our multi-layered transcriptomics approach is in agreement with the computational model developed by Ding et al.[18], which is based on a network-informed analysis of the interaction of SARS-CoV-2 and HLH related genes. Their model postulates that neutrophil degranulation/neutrophil extracellular traps (NETs) cause endothelial damage, and consequently, thrombotic complications of COVID-19. Ding's and our interpretation is supported by several experimental evidence[81–83,93] for neutrophil hyperactivation and its association with severity of COVID-19, as recently reviewed by Ackermann et al.[84]. As we were able to demonstrate by the multi-omics association between leukocyte and plasma molecules, flow cytometry and proteomic data indicate a systemic and integrated network of molecules associated with neutrophil growth, activation, and mobilization leading to neutrophil dysregulation in severe COVID-19[82,83]. This supports the concept that the pathophysiology of HLH does not only involve T cell, NK cell and macrophage dysregulation but also the hyperactivation of neutrophils. In accordance with our findings, it has recently been demonstrated that neutrophils accumulate in inflamed tissues of HLH and COVID-19 as a consequence of T-cell driven pro-inflammatory cytokine and chemokine release, which

does not return to a homeostatic level due to an ineffective T cytotoxic response[18,19]. Thus, maintaining the immunogenic stimulus for more cytokine/chemokine secretion, which promotes the sustained neutrophil recruitment and consequent tissue damage.

The use of high-throughput techniques to identify biomarkers, molecular pathways and pathophysiological information derived from genetic and transcriptomic data has contributed to the understanding of the immunopathology of diseases[94]. In this regard, patients with severe COVID-19 have been associated with mutations in genes involved in the regulation of type I and III IFN immunity pointing to the role of structural genomics in determining the course of COVID-19[95]. Our functional genomics approach characterized by a multi-layered transcriptomic analysis started with dozens of thousand genes expressed in nasopharyngeal swabs and peripheral blood obtained from 1253 subjects. We identified 25481 DEGs, of which 237 were commonly up-regulated in COVID-19 and HLH. Among them, 25 genes were differentially expressed in COVID-19_ICU in comparison to COVID-19_nonICU, demonstrating the disruption in gene network signatures and reinforcing the concept that alterations in molecule relationships are immunopathological mechanisms also involved in the breakdown of body homeostasis[96]. In this context, 11 genes (*ARG1, AZU1, CD177, CEACAM8, CTSG, DEFA4, ELANE, LCN2, MCEMP1, MMP8*, and *OLFM4*) commonly dysregulated in COVID-19 and HLH specifically stratified COVID-19_ICU from COVID-19_nonICU patients. They encode proteins involved in neutrophil degranulation and contribute to the development of comorbidities that increase the risk of progressing to severe COVID-19[73,74]. Random forest model ranking indicated that these genes accurately distinguish COVID-19_nonICU from COVID-19_ICU patients. For instance, this machine learning approach ranked ARG1 and its interaction with other molecules (CD177, AZU1, MCEMP1 and LCN2) as an important predictor for ICU admission, supporting the role of these molecules as biomarkers for severe COVID-19[80,97].

It is important to mention that while our work has the strengths of a robust cross-study and cross-tissue network analysis, it has limited value to mechanistically explore the role of specific molecules in the outcome of COVID-19. However, several of these dysregulated molecules shared by COVID-19 and HLH have been successfully investigated for the treatment of SARS-CoV-2 infection, supporting our findings. For instance, inhibition of the CCR5-CCL4 axis by Leronlima (anti-CCR5 monoclonal antibody)[98], or blockade of cytokine signaling by

Tocilizumab (anti-IL-6R)[99], Adalimumab (anti-TNF)[100], or Anakinra (anti-IL1R)[101] have been shown to ameliorate, in some cases, severe COVID-19. Furthermore, Ruxolitinib, a JAK1/JAK2 inhibitor acting downstream of JAK-dependent chemokines/cytokines such as IFN-γ, IL-1β, IL-6, TNF, G-CSF, CXCL9, and CXCL10 [19,102] has shown promising results in treating COVID-19[103]. Of note, several approaches targeting neutrophils to treat SARS-CoV-2 complications have entered clinical trials, including the disruption of signaling via CXCR2, IL-8, IL-17A, or the use of phosphodiesterase (PDE) inhibitors[104]. Moreover, in agreement with our data, inhibition of neutrophil-derived anti-microbial proteins are being actively investigated in clinical trials by exploring the mechanistic and clinical effects of alvelestat, an oral neutrophil elastase inhibitor (COVID-19 Study of Safety and Tolerability of Alvelestat, ClinicalTrials.gov). Meanwhile other proteases (AZU1 and CTSG) and the inhibition of NETs have been suggested to alleviate SARS-CoV-2 symptoms[80,84,105].

In conclusion, our comprehensive multi-layered transcriptomic and cross-tissue analysis of nasopharyngeal swabs and peripheral blood leukocytes indicates systemic communalities among severe COVID-19 and in HLH. Our functional genomic work shows that the overlapping molecular pathways between COVID-19 and HLH are characterized by an interconnected cytokine/chemokine profile that hyper stimulates and systemically attracts adaptive and innate immune cells, culminating in the hyperactivation of neutrophils and leading to life-threatening clinical complications. Thus, our work suggests novel putative molecular pathways that can be exploited for therapeutic intervention for both COVID-19 and HLH.

**ACKNOWLEDGMENTS**

We acknowledge the Latin American Society of Immunodeficiencies (LASID) for providing the research funding of LFS (LASID Fellowship award 2020), and the São Paulo Research Foundation (FAPESP grants 2018/18886-9, 2020/01688-0, and 2020/07069-0 to OCM) for financial support. Computational analysis was supported by FAPESP and partially by the grants from Ontario Research Fund (#34876), Natural Sciences Research Council (NSERC #203475), Canada Foundation for Innovation (CFI #29272, #225404, #33536), and IBM granted to IJ, as well as the National Institutes of Health (NHLBI) through award HL130704 granted to AJ. This study was financed in part by the coordination for the improvement of higher education personnel – Brazil (CAPES) – finance code 001. We thank Prof. Luis Carlos de Souza Ferreira for discussion and suggestions to develop this project.

**AUTHOR CONTRIBUTIONS**

LFS, and OCM wrote the manuscript, conceived the study design, bioinformatics analyses and performed scientific insights; LFS, AHCM, GCB, CASP, DLMF, PPF, DRP, ISF, RCS, GJM, AERO, and IJ, co-wrote the manuscript and performed bioinformatics analyses; KAO, JJC, JB, AJ, JSS, TU and JLS generated transcriptome and proteome data; JPSP, JAMB, NOSC, HDO, VLGC, ACN, KAO, JJC, JB, AJ, JSS, TU, JLS, HIN, IJ and OCM revised and edited the final manuscript and provided scientific input; OCM supervised the project.

**COMPETING INTEREST STATEMENT**

The authors declare no competing financial and/or non-financial interests concerning the work described.

**FIGURE LEGENDS:**

**Graphic Abstract:** Schematic view of datasets and results obtained indicating new candidate biomarkers and therapeutic targets for COVID-19 and HLH (Figure created using BioRender.com).

**Figure 1: Transcriptional overlap between COVID-19 and HLH. (A)** Phenotypic convergence between severe COVID-19 and HLH. **(B)** Number of differentially expressed genes (DEGs, up- and down-regulated) by dataset. **(C)** Study design for multi-transcriptomics analysis. **(D)** Circos plot showing 237 common up-regulated DEGs (or transcripts) between HLH and the different COVID-19 datasets (red lines), divided in 3 overlapping subgroups (Detailed in **Supp. Table 2** and **3**). The thickness of each line represents the number of genes shared between the different datasets. **(E)** Protein-protein interaction network among the 237 transcripts and the 7 genes causing HLH due to inborn errors of immunity (IEI). Node colours denote Gene Ontology biological processes. The label (gene name) colours represent transcripts from *Overlap 1* (green), *Overlap 2* (red), and *Overlap 3* (blue). Center circle and side circles represent common molecules across all 3 or 2 overlapping datasets, respectively. The upper left subnetwork represents the interactions between the 7 genes associated with fHLH and those from overlaps are bold. The circle on upper left (gene names not shown) contains 1329 proteins connected by 217 interactions with the 7 HLH/IEI-associated genes. The full network comprises 1538 proteins and 2522 direct physical interactions obtained from IID database ver. 2021-05[106].

**Figure 2: Cytokine/chemotaxis and neutrophil-associated transcriptional signatures predominate in the COVID-19 and HLH overlap. (A)** Dot plot showing the most significant biological processes enriched by the 237 common up-regulated transcripts of COVID-19 and HLH datasets. The dot size is proportional to the number of genes enriching the ontology term and color proportional to adjusted p value (green > significant than blue). **(B)** Network highlighting genes and cellular component associations. Only enriched terms with adjusted *p* value <0.05 are shown. The degree of associations is displayed by edge color and thickness (e.g., lighter color and thinner edges signify fewer connections). Node color represents different GO CCs. Both enriched CCs and BPs were analyzed using ClusterProfiler with R programming. **(C-E)** Bubble heatmaps showing the hierarchical clustering based on Euclidian distance of expression patterns of genes associated to **(C)** cytokine signaling, **(D)** chemotaxis**,** and **(E)** neutrophil-mediated immunity in COVID-19 and HLH datasets. The color of circles corresponds to log2 fold change (log2FC). Pleiotropic genes belonging to more than one category are bold (**Supp. Table S8**).

**Figure 3: Disease state impacts the correlation between cytokine/chemotaxis and neutrophil-mediated immunity genes.** (**A and B**) Estimated correlations of cytokine signaling/chemotaxis and neutrophil-mediated immunity molecules versus their corresponding first 2 canonical variates (x-CV1 and x-CV2, for cytokine/chemotaxis related genes; y-CV1 and y-CV2 for neutrophil-mediated immunity genes in (**A**) controls and (**B**) COVID-19 patients. Cytokine/chemotaxis and neutrophil-mediated immunity genes with a correlation of ≥ 0.7 or ≤ -0.7 are colored in green and blue, respectively, while those with a correlation of < 0.7 or > -0.7 are gray in both groups. **(C)** Scatter plots with marginal boxplots display the relationship

between variables (genes). Correlation coefficient (ρ) and significance level (p-value) for each correlation are shown within each graph.

**Figure 4: Transcripts stratifying severe COVID-19 from other respiratory diseases are also highly dysregulated in HLH. (A)** Schematic overview of study design and patient classification of dataset GSE157103 reported by Overmyer et al.[58]. **(B)** Protein-protein interaction (PPI) network among 158 proteins and the 25 genes significant for severe COVID-19_ICU. Node colour denotes Gene Ontology biological process terms. Genes (lower right side half circle) and proteins (left circle and upper right half circle) are denoted by different symbols as explained in the figure legend. Blue circled symbols represent molecules involved in neutrophil-mediated immunity. **(C)** Principal Component Analysis (PCA) with spectral decomposition shows the stratification of COVID-19_ICU from COVID-19_nonICU and other respiratory diseases (Control_nonICU and Control_ICU). Variables with positive correlation are pointing to the same side of the plot, contrasting with negative correlated variables, which point to opposite sides. Confidence ellipses are shown for each group/category. **(D)** PCA displaying the stratification of HLH and healthy controls based on the same 25 DEGs as in (B). **(E)** Correlation matrices of the 25 DEGs (Controls, left matrix; COVID-19_nonICU, middle matrix; and COVID-19_ICU, right matrix). The color scale bar represents the Pearson's correlation coefficient, containing negative and positive correlations from -1 to 1, respectively. **(F)** Scatter plots with marginal boxplots display the relationship between the eight genes stratifying severe COVID-19. Correlation coefficient (ρ) and significance level (p-value) for each correlation are shown within each graph.

**Figure 5: Multi-layered transcriptomic analysis associates COVID-19 and HLH common genes with disease severity.** (**A**) Schematic overview of sample cohort and classification of scRNAseq dataset obtained by Schulte-Schrepping et al.[65] and used for the following analysis. **(B)** Heatmap showing scRNAseq expression of differentially expressed genes (DEGs) associated with disease severity. Cells and cohorts (controls, mild and severe COVID-19) are indicated by different colors in the legends. **(C)** Box plots of scRNAseq expression demonstrating that 11 from the 21 genes identified in **B** are up-regulated when comparing severe and mild COVID-19 patients. **(D)** Box plots of the 11 transcripts stratifying COVID-19_ICU patients from COVID-19_nonICU patients obtained from the bulk RNAseq dataset from Overmyer et al.[58]. Significant differences between groups are indicated by asterisks (* $p \leq 0.05$, ** $p \leq 0.01$ and *** $p \leq 0.001$). **(E)** Box plots of microarray data illustrating that the disease severity association of COVID-19 detected by scRNAseq corresponds to the expression differences between HLH patients and controls. Significant differences between groups are indicated by asterisks (* $p \leq 0.05$, ** $p \leq 0.01$ and *** $p \leq 0.001$).

**Figure 6: Random Forest prediction analysis suggests potential biomarker for severe COVID-19. (A)** Receiver operating characteristics (ROC) curve of 11 genes from COVID-19_ICU compared to COVID-19_nonICU patients with an area under the curve (AUC) of 82,4% for both groups. 1=COVID-19_nonICU; 2=COVID-19_ICU. **(B)** Stable Curve showing number of trees and error rate (out of bag or OOB) with medium of 27,03%. 1=COVID-19_nonICU; 2=COVID-19_ICU.

**(C)** Variable importance scores plot based on gini_decrease and number(no)_of_nodes for each variable showing which variables are more likely to be essential in the random forest's prediction. **(D)** Ranking of the top 10 variables according to mean minimal depth (vertical bar with the mean value in it) calculated using trees. The blue color gradient reveals the min and max minimal depth for each variable. The range of the x-axis is from zero to the maximum number of trees for the feature. **(E)** Mean minimal depth variable interaction plot showing most frequent occurring interactions between the variables on the left side with light blue color, and least frequent occurring interactions on the right side of the graph with dark blue color. The red horizontal line indicates the smallest mean minimum depth and the black lollipop represent the unconditional mean minimal depth of a variable.

**TABLES:**

**Table 1: Dataset Information and sample size used for transcriptome analysis**

| Database | Dataset ID | Seq. Method | Sample type | Number of patients | Number of Controls | Original study |
|---|---|---|---|---|---|---|
| GEO | GSE26050 | microarray | PBMCs | 11 | 33 | Sumegi et al. 2011[107] |
| GEO | GSE152418 | bulk-RNA seq | PBMCs | 17 | 17 | Arunachalam et al. 2020[108] |
| GEO | GSE152075 | bulk-RNA seq | nph swab | 430 | 54 | Liebermann et al. 2020[109] |
| GEO | GSE156063 | bulk-RNA seq | nph swab | 93 | 141[a] | Mick et al. 2020[110] |
| GEO | GSE163151 | bulk-RNA seq | nph swab<br>PBL | 138<br>7 | 11<br>20 | Ng et al. 2021[61] |
| GEO | GSE152641 | bulk-RNA seq | PBL | 62 | 24 | Thair et al. 2021[62] |
| GEO | GSE157103 | bulk-RNA seq | PBL | 100 | 26[b] | Overmyer et al. 2020[58] |
| EGA | EGAS00001004571 | scRNA seq | PBL/PBMCs | Cohort1 18<br>Cohort2 17 | 21<br>13 | Schulte-Schrepping et al. 2020[65] |

[a]individuals with non-infectious respiratory diseases (NIRD; n=100) and with other respiratory infectious diseases (OIRD; n=41)

[b]individuals negative for SARS-CoV-2 but admitted to ICU (n=16) and nonICU (n=10) units due to respiratory symptoms

nph=nasopharyngeal

## METHODS

### RESOURCE AVAILABILITY

*Data and code availability*

This paper analyzes existing, publicly available data. The accession numbers for the datasets are listed in the key resources table.

All original codes used for data analysis have been deposited at github (https://github.com/lschimke/COVID19-and-HLH-paper) and are publicly available as of the date of publication. R packages are listed in the key resources table.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

### METHOD DETAILS

#### Data Curation

We searched in public functional genomics data repositories (Gene Expression Omnibus[111] and Array Express[112] for human transcriptome data from patients with HLH and COVID-19 published until February 2021. Inclusion criteria for datasets were transcriptome data from minimum n=10 patients without any treatment, availability of gene counts, blood or nasopharyngeal swab samples, and availability of control groups for comparison. This resulted in 7 studies, 6 from COVID-19 patients and one from HLH patients (containing 4 patients with germline mutations and 7 without identified mutation) with transcriptome data generated from different platforms (**Table 1**). Furthermore, we included the scRNAseq study obtained from the European Genome-phenome Archive[113] (EGA; EGAS00001004571) of blood leukocytes from two independent cohorts of patients with COVID-19 and healthy controls, respectively[65]. Combined, we investigated a total of 1253 transcriptome samples.

**Differential expression analysis, meta-analysis and visualization of multiple gene expression data sets from microarray and bulk RNAseq**

Read counts were transformed (log2 count per million or CPM) and differentially expressed genes (DEGs) between groups were identified through the webtool NetworkAnalyst 3.0[114] using limma-voom pipeline[115]. Statistical cut-offs to define DEGs are described below in the section **statistical analysis.** Shared DEGs among all datasets were displayed using Venn diagram[116] and Circos Plot[117] online tools.

**Single cell RNAseq analysis**

The Seurat Object containing the scRNAseq published by Schulte-Schrepping et al.[65] and deposited at the EGA (EGAS00001004571) were used for single cell analysis. We followed the Seurat pipeline[118] as previously described by Stuart et al.[119] to perform differential expression analysis and data visualization, i.e., UMAP, dotplot, and heatmap construction. Regression for the number of UMIs and scaling were performed as previously described[65].

**Interactome analysis**

For more comprehensive Protein–Protein Interaction (PPI) analyses, we used NAViGaTOR 3.0.14[120] to visualize genes commonly dysregulated in COVID-19 and HLH datasets, highlighting the biological processes enriched by each gene. Prior to visualization, DEGs were used as input into Integrated Interactions Database (IID version 2021-05; http://ophid.utoronto.ca/iid)[106,121] to identify direct physical protein interactions. Then, the resultant network was annotated, analysed, and visualized using NAViGaTOR 3.0.14[120].

**Enrichment analysis and data visualization**

We used ClusterProfiler[122] R package to obtain dot plots of enriched signaling pathways. Elsevier Pathway Collection analysis for selected gene lists (7 genes underlying fHLH/IEI and 11 genes associated with severe COVID-19) was carried out using Enrichr webtool[123–125]. Set of genes

associated with cytokine/chemotaxis and neutrophil-mediated immunity from each dataset were visualized in bubble-based heat maps applying one minus cosine similarity using Morpheus[126]. Circular heatmaps were generated using R version 4.0.5 (The R Project for Statistical Computing. https://www.r-project.org/) and R studio Version 1.4.1106 (RStudio. https://www.rstudio.com/) using the circlize R package. Box plots were generated using the R packages ggpubr, lemon, and ggplot2.

**Correlation Analysis**

Principal Component Analysis (PCA) of genes associated with COVID-19 severity (25 transcripts) was performed using the R functions prcomp and princomp, through factoextra package[127]. Canonical Correlation Analysis (CCA)[128] of genes associated with cytokines/chemokines and neutrophil-mediated immune responses was performed using the packages CCA and whitening[128]. In addition, we used the corrgram, psych, and inlmisc R packages to build correlograms. In addition, multilinear regression analysis for combinations of different variables (genes) was performed using the R package ggpubr, ggplot2 and ggExtra.

**Proteome Data Analysis**

We also evaluated the proteomics data obtained from plasma samples of COVID-19 patients previously reported by Overmyer et al.[58]. Briefly, raw LFQ abundance values were quantified, normalized and log2 transformed as previously described[58]. Differences in protein expression between COVID-19_ICU and COVID-19_nonICU were calculated as described below in the section **statistical analysis**. Enrichment of proteins significant for COVID-19 ICU was performed using Enrichr webtool[123–125] and most significant enriched pathways were displayed by dot plot created with R using tidyverse, viridis and ggplot2 packages while Circos Plot of gene-pathway association was built using Circos online tool[117].

### Decision-tree classification and machine learning model predictors

We employed random forest model to construct a classifier able to discriminate between COVID-19_nonICU and COVID-19_ICU highlighting the most significant predictors for ICU admission. We trained a Random Forest model using the functionalities of the R package *randomForest* (version 4.6.14)[79]. Five thousand trees were used, and the number of variables resampled were equal to three. Follow-up analysis used the Gini decrease, number of nodes, and mean minimum depth as criteria to determine variable importance. Interaction between pair of variables was assessed by minimum depth as criterion. The adequacy of the Random Forest model as a classifier was assessed through out of bags error rate and ROC curve. For cross-validation, we split the dataset in training and testing samples, using 75% of the observations for training and 25% for testing

## STATISTICAL ANALYSIS

### Differentially expressed genes and proteins

To determine DEGs of each dataset we applied the statistical cut-offs of log2 fold-change > 1 (up-regulated), log2 fold-change < -1 (down-regulated), and adjusted p-value < 0.05. We used the Fisher´s method to combine p values from multiple studies for information integration[114]. Differences in protein expression between COVID-19_ICU and COVID-19_nonICU was calculated using the nonparametric MANOVA (multivariate analysis of variance) test[129] followed by analysis of nonparametric Inference for Multivariate Data[130] using the R packages npmv, nparcomp, and ggplot2.

**SUPPLEMENTAL INFORMATION**

All supplemental figures, titles, and legends are provided in separate document file.

Supplemental Tables S1-S14 are provided in a separate document file:

Table S1: Up- and Down-regulated DEGs of all datasets

Table S2: List of common DEGs

Table S3: Genelist Overlaps

Table S4: Biological Process (BP) enriched terms

Table S5: Cellular components (CC) enriched terms

Table S6: Cytokine and Chemotaxis genes

Table S7: Neutrophil-mediated immunity genes

Table S8: Pleiotropic genes

Table S9: DEGs between COVID-19_ICU and COVID-19_nonICU

Table S10: DEGs in additional datasets

Table S11: 25 Genes in proteomic/protein expression datasets

Table S12: Significant plasma proteins comparing COVID-19_ICU versus COVID-19_nonICU patients

Table S13: Enriched BP GO terms of 158 significant plasma proteins

Table S14: Enriched Elsevier Pathway collections terms of 11 genes significant for severe COVID-19

**REFERENCES**


1. Mehta, P. *et al.* COVID-19: consider cytokine storm syndromes and immunosuppression. *The Lancet* vol. 395 1033–1034 (2020).
2. Chen, L. Y. C. & Quach, T. T. T. COVID-19 cytokine storm syndrome: a threshold concept. (2021) doi:10.1016/j.blre.2020.100707.
3. Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506 (2020).
4. England, J. T. *et al.* Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID- 19 . The COVID-19 resource centre is hosted on Elsevier Connect , the company ' s public news and information . (2020).
5. Ørskov, S., Frost Nielsen, B., Føns, S., Sneppen, K. & Simonsen, L. The COVID-19 pandemic: Key considerations for the epidemic and its control. *APMIS* (2021) doi:10.1111/apm.13141.
6. Pandit, B., Bhattacharjee, S. & Bhattacharjee, B. Association of clade-G SARS-CoV-2 viruses and age with increased mortality rates across 57 countries and India. *Infect. Genet. Evol.* **90**, (2021).
7. Garcia-Beltran, W. F. *et al.* Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. *Cell* **184**, 2372-2383.e9 (2021).
8. Mathew, D. *et al.* Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science (80-. ).* **369**, (2020).
9. Laing, A. G. *et al.* A dynamic COVID-19 immune signature includes associations with poor prognosis. *Nat. Med.* **26**, 1623–1635 (2020).
10. Azkur, A. K. *et al.* Immune response to SARS-CoV-2 and mechanisms of immunopathological changes in COVID-19. *Allergy: European Journal of Allergy and Clinical Immunology* vol. 75 1564–1581 (2020).
11. Koutsakos, M. *et al.* Integrated immune dynamics define correlates of COVID-19 severity and antibody responses. *Cell Reports Med.* **2**, (2021).
12. Webb, B. J. *et al.* Clinical criteria for COVID-19-associated hyperinflammatory syndrome: a cohort study. *Lancet Rheumatol.* **2**, e754–e763 (2020).
13. Fajgenbaum, D. C. & June, C. H. Cytokine Storm. *N. Engl. J. Med.* **383**, 2255–2273 (2020).
14. Yongzhi, X. COVID-19-associated cytokine storm syndrome and diagnostic principles: an old and new Issue. *Emerging Microbes and Infections* vol. 10 266–276 (2021).
15. George, M. Hemophagocytic lymphohistiocytosis: review of etiologies and management. *J. Blood Med.* 69 (2014) doi:10.2147/jbm.s46255.
16. Soy, M., Atagündüz, P., Atagündüz, I. & Sucak, G. T. Hemophagocytic lymphohistiocytosis: a review inspired by the COVID-19 pandemic. *Rheumatol. Int.* **41**, 7–18 (2021).

17. Lorenz, G. *et al.* Title: Cytokine release syndrome is not usually caused by secondary hemophagocytic lymphohistiocytosis in a cohort of 19 critically ill COVID-19 patients. *Sci. Rep.* **10**, 18277 (2020).

18. Ding, J. *et al.* A network-informed analysis of SARS-CoV-2 and hemophagocytic lymphohistiocytosis genes' interactions points to Neutrophil extracellular traps as mediators of thrombosis in COVID-19. *PLoS Comput. Biol.* **17**, e1008810 (2021).

19. Albeituni, S. *et al.* Mechanisms of action of ruxolitinib in murine models of hemophagocytic lymphohistiocytosis. *Blood* **134**, 147–159 (2019).

20. Prilutskiy, A. *et al.* SARS-CoV-2 infection–Associated hemophagocytic lymphohistiocytosis an autopsy series with clinical and laboratory correlation. *Am. J. Clin. Pathol.* **154**, 466–474 (2020).

21. Opoka-Winiarska, V., Grywalska, E. & Roliński, J. Could hemophagocytic lymphohistiocytosis be the core issue of severe COVID-19 cases? *BMC Med.* **18**, 1–11 (2020).

22. Harris, C. K., Hung, Y. P., Nielsen, G. P., Stone, J. R. & Ferry, J. A. Bone Marrow and Peripheral Blood Findings in Patients Infected by SARS-CoV-2. *Am. J. Clin. Pathol.* **155**, 627–637 (2021).

23. Retamozo, S. *et al.* Haemophagocytic syndrome and COVID-19. *Clinical Rheumatology* vol. 40 1233–1244 (2021).

24. von der Thüsen, J. H. *et al.* Case report: a fatal combination of hemophagocytic lymphohistiocytosis with extensive pulmonary microvascular damage in COVID-19 pneumonia. *J. Hematop.* **14**, 79–83 (2021).

25. Lima, R. *et al.* Hemophagocytic syndrome and COVID-19. *Respir. Med. Case Reports* **31**, (2020).

26. Tholin, B., Hauge, M. T., Aukrust, P., Fehrle, L. & Tvedt, T. H. Hemophagocytic lymphohistiocytosis in a patient with COVID-19 treated with tocilizumab: a case report. *J. Med. Case Rep.* **14**, (2020).

27. Schnaubelt, S. *et al.* Hemophagocytic lymphohistiocytosis in COVID-19: Case reports of a stepwise approach. *Medicine (Baltimore).* **100**, e25170 (2021).

28. Naous, E., Nassani, B. M., Yaghi, C., Nasr, F. & Medlej, R. Hemophagocytic lymphohistiocytosis, a new cause of death during 'post-acute COVID-19 syndrome?' A case report. *J. Hematop.* **1**, 3 (2021).

29. Cabral-Marques, O. *et al.* Flow Cytometry Contributions for the Diagnosis and Immunopathological Characterization of Primary Immunodeficiency Diseases With Immune Dysregulation. *Frontiers in Immunology* vol. 10 2742 (2019).

30. Janka, G. E. Familial and acquired hemophagocytic lymphohistiocytosis. *Annual Review of Medicine* vol. 63 233–246 (2012).

31. Usmani, G. N., Woda, B. A. & Newburger, P. E. Advances in understanding the pathogenesis of HLH. *British Journal of Haematology* vol. 161 609–622 (2013).

32. Tiab, M., Mechinaud, F. & Harousseau, J. L. Haemophagocytic syndrome associated with infections. *Bailliere's Best Pract. Res. Clin. Haematol.* **13**, 163–178 (2000).

33. Ramos-Casals, M., Brito-Zerón, P., López-Guillermo, A., Khamashta, M. A. & Bosch, X. Adult haemophagocytic syndrome. in *The Lancet* vol. 383 1503–1516 (Elsevier B.V., 2014).

34. Simon, A. C., Delhi Kumar, C. G., Basu, D. & Ramesh Kumar, R. Hemophagocytic Lymphohistiocytosis in Children: Clinical Profile and Outcome. *J. Pediatr. Hematol. Oncol.* **42**, e281–e285 (2020).

35. Veerakul, G., Tanphaichitr, V. S., Jirarattanasopa, N., Sanpakit, K. & Mahasandana, C. Secondary hemophagocytic lymphohistiocytosis in children: An analysis of etiology and outcome. *J. Med. Assoc. Thail.* **85**, S530-41 (2002).

36. Miao, Y. *et al.* Pathogenic gene mutations or variants identified by targeted gene sequencing in adults with hemophagocytic lymphohistiocytosis. *Front. Immunol.* **10**, 395 (2019).

37. Luo, H. *et al.* Germline variants in UNC13D and AP3B1 are enriched in COVID-19 patients experiencing severe cytokine storms. *Eur. J. Hum. Genet.* (2021) doi:10.1038/s41431-021-00886-x.

38. Stelzer, G. *et al.* The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinforma.* **54**, 1.30.1-1.30.33 (2016).

39. Lebre, M. C. *et al.* Differential expression of inflammatory chemokines by Th1- and Th2-cell promoting dendritic cells: A role for different mature dendritic cell populations in attracting appropriate effector cells to peripheral sites of inflammation. *Immunol. Cell Biol.* **83**, 525–535 (2005).

40. Charo, I. F. & Ransohoff, R. M. The Many Roles of Chemokines and Chemokine Receptors in Inflammation. *N. Engl. J. Med.* **354**, 610–621 (2006).

41. Blandino-Rosano, M., Perez-Arana, G., Mellado-Gil, J. M., Segundo, C. & Aguilar-Diosdado, M. Anti-proliferative effect of pro-inflammatory cytokines in cultured β cells is associated with extracellular signal-regulated kinase 1/2 pathway inhibition: Protective role of glucagon-like peptide -1. *J. Mol. Endocrinol.* **41**, 35–44 (2008).

42. Zhang, J. M. & An, J. Cytokines, inflammation, and pain. *International Anesthesiology Clinics* vol. 45 27–37 (2007).

43. Aratani, Y. Myeloperoxidase: Its role for host defense, inflammation, and neutrophil function. *Archives of Biochemistry and Biophysics* vol. 640 47–52 (2018).

44. Pohl, J., Pereira, H. A., Martin, N. M. & Spitznagel, J. K. Amino acid sequence of CAP37, a human neutrophil granule-derived antibacterial and monocyte-specific chemotactic glycoprotein structurally similar to neutrophil elastase. *FEBS Lett.* **272**, 200–204 (1990).

45. Korkmaz, B., Horwitz, M. S., Jenne, D. E. & Gauthier, F. Neutrophil elastase, proteinase 3, and cathepsin G as therapeutic targets in human diseases. *Pharmacological Reviews* vol. 62 726–759 (2010).

46. Tralau, T., Meyer-Hoffert, U., Schröder, J. M. & Wiedow, O. Human leukocyte elastase and cathepsin G are specific inhibitors of C5a-dependent neutrophil enzyme release and chemotaxis. *Exp. Dermatol.* **13**, 316–325 (2004).

47. Lehrer, R. I. & Lu, W. α-Defensins in human innate immunity. *Immunological Reviews* vol. 245 84–112 (2012).

48. Lin, W. C. & Fessler, M. B. Regulatory mechanisms of neutrophil migration from the circulation to the airspace. *Cellular and Molecular Life Sciences* vol. 1 3 (2021).

49. Bradley, L. M., Douglass, M. F., Chatterjee, D., Akira, S. & Baaten, B. J. G. Matrix metalloprotease 9 mediates neutrophil migration into the airways in response to influenza virus-induced toll-like receptor signaling. *PLoS Pathog.* **8**, e1002641 (2012).

50. Marsh, R. A. *et al.* STX11 mutations and clinical phenotypes of familial hemophagocytic lymphohistiocytosis in North America. *Pediatr. Blood Cancer* **55**, 134–140 (2010).

51. Wada, T. *et al.* Down-regulation of CD5 expression on activated CD8+ T cells in familial hemophagocytic lymphohistiocytosis with perforin gene mutations. *Hum. Immunol.* **74**, 1579–1585 (2013).

52. Janka, G. E. Familial and acquired hemophagocytic lymphohistiocytosis. *European Journal of Pediatrics* vol. 166 95–109 (2007).

53. Dell'Acqua, F. *et al.* Hermansky-Pudlak syndrome type II and lethal hemophagocytic lymphohistiocytosis: Case description and review of the literature. *J. Allergy Clin. Immunol. Pract.* **7**, 2476-2478.e5 (2019).

54. Zhao, X. W. *et al.* Defects in neutrophil granule mobilization and bactericidal activity in familial hemophagocytic lymphohistiocytosis type 5 (FHL-5) syndrome caused by STXBP2/Munc18-2 mutations. *Blood* **122**, 109–111 (2013).

55. D'Orlando, O. *et al.* Syntaxin 11 is required for NK and CD8+ T-cell cytotoxicity and neutrophil degranulation. *Eur. J. Immunol.* **43**, 194–208 (2013).

56. Zhang, Y. *et al.* A Network of interactions enables CCM3 and STK24 to coordinate UNC13D-driven vesicle exocytosis in neutrophils. *Dev. Cell* **27**, 215–226 (2013).

57. Catz, S. D. The role of Rab27a in the regulation of neutrophil function. *Cell. Microbiol.* **16**, 1301–1310 (2014).

58. Overmyer, K. A. *et al.* Large-Scale Multi-omic Analysis of COVID-19 Severity. *Cell Syst.* (2020) doi:10.1016/j.cels.2020.10.003.

59. Rickman, J. M., Wang, Y., Rollett, A. D., Harmer, M. P. & Compson, C. Data analytics using canonical correlation analysis and Monte Carlo simulation. *npj Comput. Mater.* **3**, 26 (2017).

60. Ferreira, F., Bota, D., Bross, A., Mélot, C. & Vincent, J. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA* **286**, 1754–1758 (2001).

61. Ng, D. *et al.* A diagnostic host response biosignature for COVID-19 from RNA profiling of nasal swabs and blood. *Sci. Adv.* **7**, (2021).

62. Thair, S. A. *et al.* Transcriptomic similarities and differences in host response between SARS-CoV-2 and other viral infections. *iScience* **24**, (2021).

63. Lever, J., Krzywinski, M. & Altman, N. Points of Significance: Principal component analysis. *Nature Methods* vol. 14 641–642 (2017).

64. Ringnér, M. What is principal component analysis? *Nature Biotechnology* vol. 26 303–304 (2008).

65. Schulte-Schrepping, J. *et al.* Severe COVID-19 Is Marked by a Dysregulated Myeloid Cell Compartment. *Cell* (2020) doi:10.1016/j.cell.2020.08.001.

66. Moser, B. T-cell memory: the importance of chemokine-mediated cell attraction. *Current biology : CB* vol. 16 (2006).

67. Castellino, F. *et al.* Chemokines enhance immunity by guiding naive CD8+ T cells to sites of CD4+ T cell-dendritic cell interaction. *Nature* **440**, 890–895 (2006).

68. Pinho, M. P., Migliori, I. K., Flatow, E. A. & Barbuto, J. A. M. Dendritic cell membrane CD83 enhances immune responses by boosting intracellular calcium release in T lymphocytes. *J. Leukoc. Biol.* **95**, 755–762 (2014).

69. Krzyzak, L. *et al.* CD83 Modulates B Cell Activation and Germinal Center Responses. *J. Immunol.* **196**, 3581–3594 (2016).

70. Silvestre-Roig, C., Fridlender, Z. G., Glogauer, M. & Scapini, P. Neutrophil Diversity in Health and Disease. *Trends in Immunology* vol. 40 565–583 (2019).

71. Ostendorf, L. *et al.* Low-Density Granulocytes Are a Novel Immunopathological Feature in Both Multiple Sclerosis and Neuromyelitis Optica Spectrum Disorder. *Front. Immunol.* **10**, 2725 (2019).

72. Carmona-Rivera, C. & Kaplan, M. J. Low-density granulocytes: A distinct class of neutrophils in systemic autoimmunity. *Seminars in Immunopathology* vol. 35 455–463 (2013).

73. Singh, A. K., Jena, A., Kumar-M, P., Sharma, V. & Sebastian, S. Risk and outcomes of coronavirus disease in patients with inflammatory bowel disease: A systematic review and meta-analysis. *United Eur. Gastroenterol. J.* **9**, 159–176 (2021).

74. Gerayeli, F. V. *et al.* COPD and the risk of poor outcomes in COVID-19: A systematic review and meta-analysis. *EClinicalMedicine* **33**, (2021).

75. Almansa, R. *et al.* Critical COPD respiratory illness is linked to increased transcriptomic activity of neutrophil proteases genes. *BMC Res. Notes* **5**, 1–8 (2012).

76. Sng, J. H. J., Prazakova, S., Thomas, P. S. & Herbert, C. MMP-8, MMP-9 and Neutrophil Elastase in Peripheral Blood and Exhaled Breath Condensate in COPD. *COPD J. Chronic Obstr. Pulm. Dis.* **14**, 238–244 (2017).

77. Gao, S., Zhu, H., Zuo, X. & Luo, H. Cathepsin g and its role in inflammation and autoimmune diseases. *Archives of Rheumatology* vol. 33 748–749 (2018).

78. Asad, S. *et al.* Proteomics-Informed Identification of Luminal Targets For In Situ Diagnosis

of Inflammatory Bowel Disease. *J. Pharm. Sci.* **110**, 239–250 (2021).

79. Liaw, A. & Wiener, M. Classification and Regression by randomForest. **2**, (2002).

80. Akgun, E. *et al.* Proteins associated with neutrophil degranulation are upregulated in nasopharyngeal swabs from SARS-CoV-2 patients. *PLoS One* **15**, e0240012 (2020).

81. Reusch, N. *et al.* Neutrophils in COVID-19. *Frontiers in Immunology* vol. 12 952 (2021).

82. Metzemaekers, M. *et al.* Kinetics of peripheral blood neutrophils in severe coronavirus disease 2019. *Clin. Transl. Immunol.* **10**, e1271 (2021).

83. Meizlish, M. L. *et al.* A neutrophil activation signature predicts critical illness and mortality in COVID-19. *Blood Adv.* **5**, 1164–1177 (2021).

84. Ackermann, M. *et al.* Patients with COVID-19: in the dark-NETs of neutrophils. *Cell Death Differ.* **11**, 12 (2021).

85. Margraf, A., Ley, K. & Zarbock, A. Neutrophil Recruitment: From Model Systems to Tissue-Specific Patterns. *Trends in Immunology* vol. 40 613–634 (2019).

86. Mortaz, E., Alipoor, S. D., Adcock, I. M., Mumby, S. & Koenderman, L. Update on neutrophil function in severe inflammation. *Frontiers in Immunology* vol. 9 2171 (2018).

87. Delorey, T. M. *et al.* COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature* (2021) doi:10.1038/s41586-021-03570-8.

88. Andonegui-Elguera, S. *et al.* Molecular Alterations Prompted by SARS-CoV-2 Infection: Induction of Hyaluronan, Glycosaminoglycan and Mucopolysaccharide Metabolism. *Arch. Med. Res.* **51**, 645–653 (2020).

89. Li, Y. *et al.* SARS-CoV-2 early infection signature identified potential key infection mechanisms and drug targets. *BMC Genomics* **22**, (2021).

90. Gardinassi, L. G., Souza, C. O. S., Sales-Campos, H. & Fonseca, S. G. Immune and Metabolic Signatures of COVID-19 Revealed by Transcriptomics Data Reuse. *Front. Immunol.* **11**, (2020).

91. Lucas, C. *et al.* Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **584**, 463–469 (2020).

92. Freire, P. P. *et al.* The relationship between cytokine and neutrophil gene network distinguishes SARS-CoV-2–infected patients by sex and age. *JCI Insight* **6**, (2021).

93. Protasio Veras, F. *et al.* SARS-CoV-2-triggered neutrophil extracellular traps mediate COVID-19 pathology. *J. Exp. Med.* **217**, (2020).

94. Su, C., Tong, J. & Wang, F. Mining genetic and transcriptomic data using machine learning approaches in Parkinson's disease. *npj Parkinson's Disease* vol. 6 1–10 (2020).

95. Zhang, Q. *et al.* Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science (80-. ).* **370**, (2020).

96. Cabral-Marques, O. *et al.* GPCR-specific autoantibody signatures are associated with physiological and pathological immune homeostasis. *Nat. Commun.* **9**, 1–14 (2018).

97. Derakhshani, A. *et al.* Arginase 1 (Arg1) as an Up-Regulated Gene in COVID-19 Patients: A Promising Marker in COVID-19 Immunopathy. *J. Clin. Med.* **10**, 1051 (2021).

98. Agresti, N. *et al.* Disruption of CCR5 signaling to treat COVID-19-associated cytokine storm: Case series of four critically ill patients treated with leronlimab. *Journal of Translational Autoimmunity* vol. 4 100083 (2021).

99. Guaraldi, G. *et al.* Tocilizumab in patients with severe COVID-19: a retrospective cohort study. *Lancet Rheumatol.* **2**, e474–e484 (2020).

100. Robinson, P. C. *et al.* The Potential for Repurposing Anti-TNF as a Therapy for the Treatment of COVID-19. *Med* **1**, 90–102 (2020).

101. Huet, T. *et al.* Anakinra for severe forms of COVID-19: a cohort study. *Lancet Rheumatol.* **2**, e393–e400 (2020).

102. Gozzetti, A., Capochiani, E. & Bocchia, M. The Janus kinase 1/2 inhibitor ruxolitinib in COVID-19. *Leukemia* vol. 34 2815–2816 (2020).

103. Rizk, J. G. *et al.* Pharmaco-Immunomodulatory Therapy in COVID-19. *Drugs* **80**, 1267–1292 (2020).

104. Chiang, C. C., Korinek, M., Cheng, W. J. & Hwang, T. L. Targeting Neutrophils to Treat Acute Respiratory Distress Syndrome in Coronavirus Disease. *Frontiers in Pharmacology* vol. 11 1576 (2020).

105. Korkmaz, B., Lesner, A., Marchand-Adam, S., Moss, C. & Jenne, D. E. Lung Protection by Cathepsin C Inhibition: A New Hope for COVID-19 and ARDS? *Journal of Medicinal Chemistry* vol. 63 13258–13265 (2020).

106. Kotlyar, M., Pastrello, C., Malik, Z. & Jurisica, I. IID 2018 update: Context-specific physical protein-protein interactions in human, model organisms and domesticated species. *Nucleic Acids Res.* **47**, D581–D589 (2019).

107. Sumegi, J. *et al.* Gene expression profiling of peripheral blood mononuclear cells from children with active hemophagocytic lymphohistiocytosis. *Blood* **117**, (2011).

108. Arunachalam, P. S. *et al.* Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science (80-. ).* **369**, 1210–1220 (2020).

109. Lieberman, N. A. P. *et al.* In vivo antiviral host transcriptional response to SARS-CoV-2 by viral load, sex, and age. *PLoS Biol.* **18**, (2020).

110. Mick, E. *et al.* Upper airway gene expression differentiates COVID-19 from other acute respiratory illnesses and reveals suppression of innate immune responses by SARS-CoV-2. *medRxiv Prepr. Serv. Heal. Sci.* (2020) doi:10.1101/2020.05.18.20105171.

111. Clough, E. & Barrett, T. The Gene Expression Omnibus database. in *Methods in Molecular Biology* vol. 1418 93–110 (Humana Press Inc., 2016).

112. Athar, A. *et al.* ArrayExpress update - From bulk to single-cell expression data. *Nucleic Acids Res.* **47**, D711–D715 (2019).

113. Lappalainen, I. *et al.* The European Genome-phenome Archive of human data consented

for biomedical research. *Nature Genetics* vol. 47 692–695 (2015).

114. Zhou, G. *et al.* NetworkAnalyst 3.0: A visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* **47**, W234–W241 (2019).

115. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).

116. Bardou, P., Mariette, J., Escudié, F., Djemiel, C. & Klopp, C. Jvenn: An interactive Venn diagram viewer. *BMC Bioinformatics* **15**, (2014).

117. Krzywinski, M. *et al.* Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).

118. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *bioRxiv* 2020.10.12.335331 (2020) doi:10.1101/2020.10.12.335331.

119. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).

120. Brown, K. R. *et al.* NAViGaTOR: Network Analysis, Visualization and Graphing Toronto. *Bioinformatics* **25**, 3327–3329 (2019).

121. Pastrello, C., Kotlyar, M. & Jurisica, I. Informed use of protein–protein interaction data: A focus on the integrated interactions database (IID). in *Methods in Molecular Biology* vol. 2074 125–134 (Humana Press Inc., 2020).

122. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* **16**, 284–287 (2012).

123. Chen, E. Y. *et al.* Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, (2013).

124. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).

125. Xie, Z. *et al.* Gene Set Knowledge Discovery with Enrichr. *Curr. Protoc.* **1**, e90 (2021).

126. Starruß, J., de Back, W., Brusch, L. & Deutsch, A. Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics* **30**, 1331–2 (2014).

127. Kassambara, A. & Mundt, F. *Multivariate Analysis II, Practical Guide to Principal Component Methods in R*. (STHDA).

128. Jendoubi, T. & Strimmer, K. A whitening approach to probabilistic canonical correlation analysis for omics data integration. *BMC Bioinformatics* **20**, 15 (2019).

129. Konietschke, F., Placzek, M., Schaarschmidt, F. & Hothorn, L. A. nparcomp: An R Software Package for Nonparametric Multiple Comparisons and Simultaneous Confidence Intervals. *J. Stat. Softw.* **64**, 1–17 (2015).

130. Burchett, W. W., Ellis, A. R., Harrar, S. W. & Bathke, A. C. Nonparametric Inference for Multivariate Data: The R Package npmv. *J. Stat. Softw.* **76**, 1–18 (2017).

**Graphic abstract**

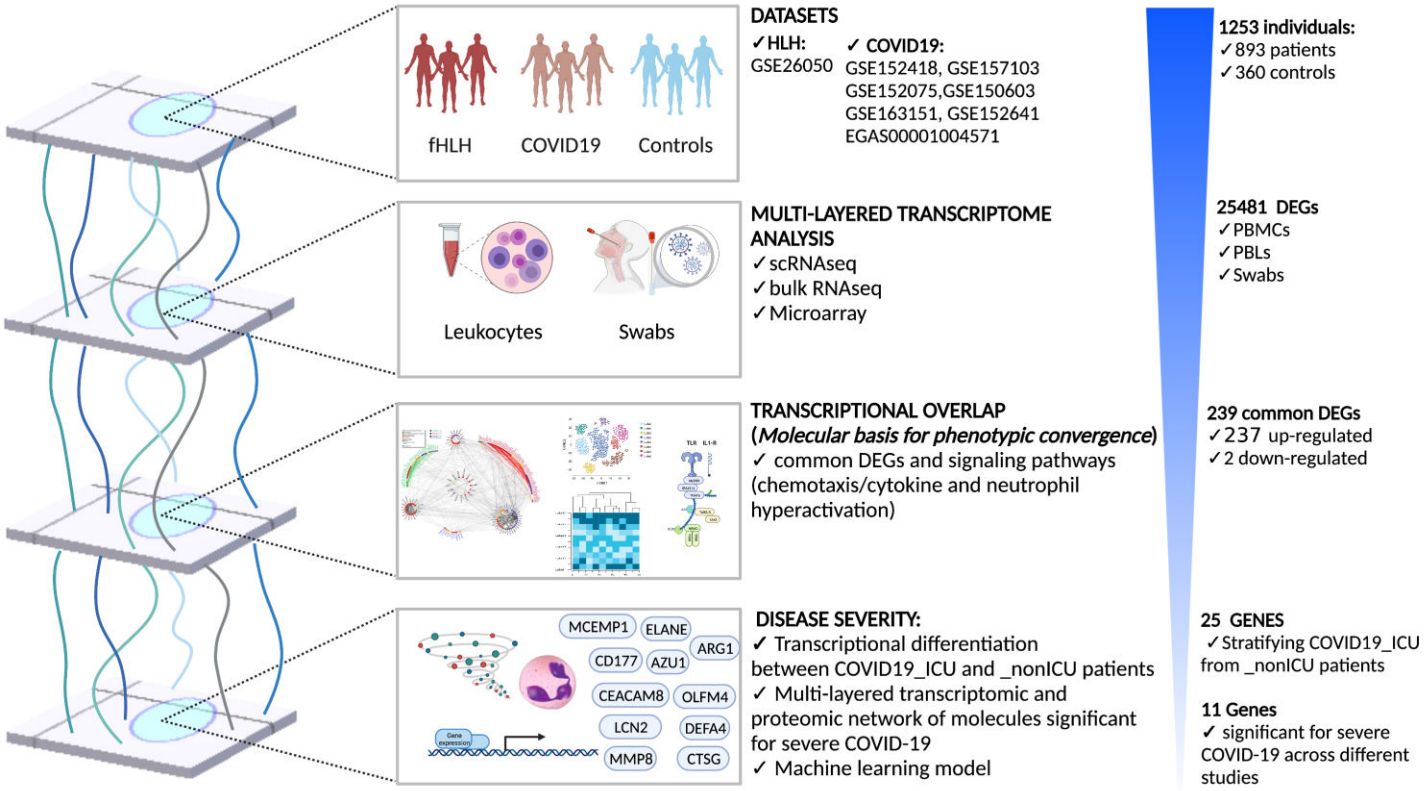
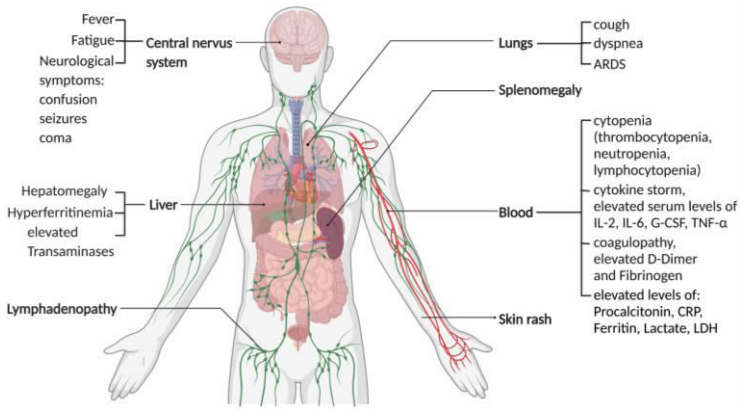


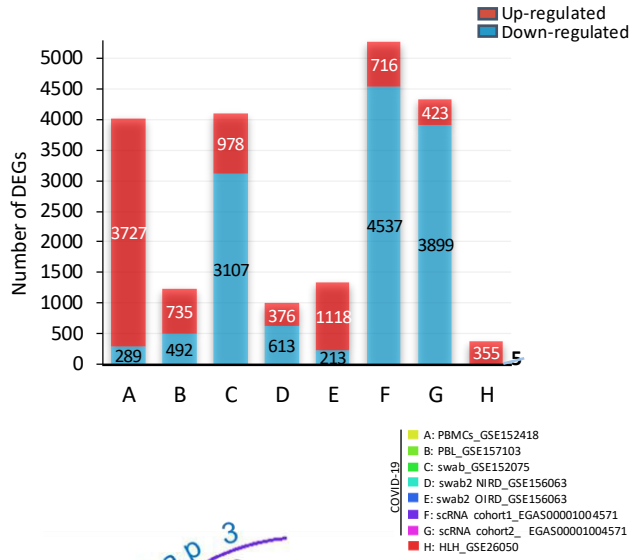
**Fig 1**

**A**

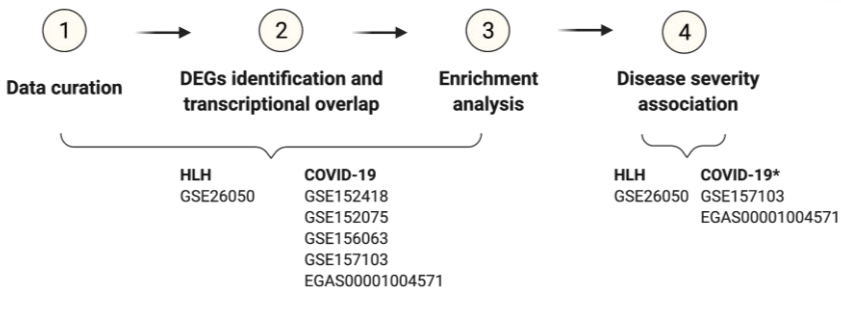
Phenotype convergence of COVID-19 and HLH



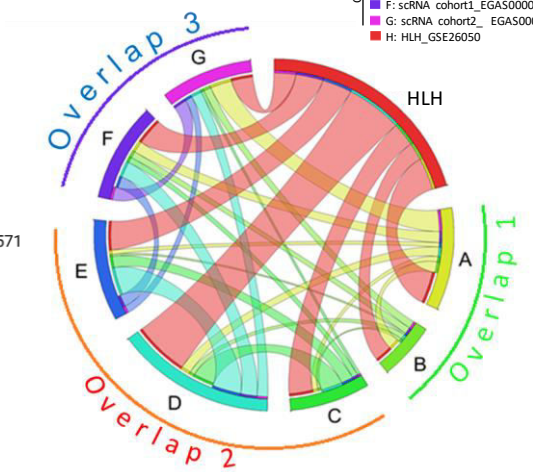

**B**




**C**




**D**




**E**

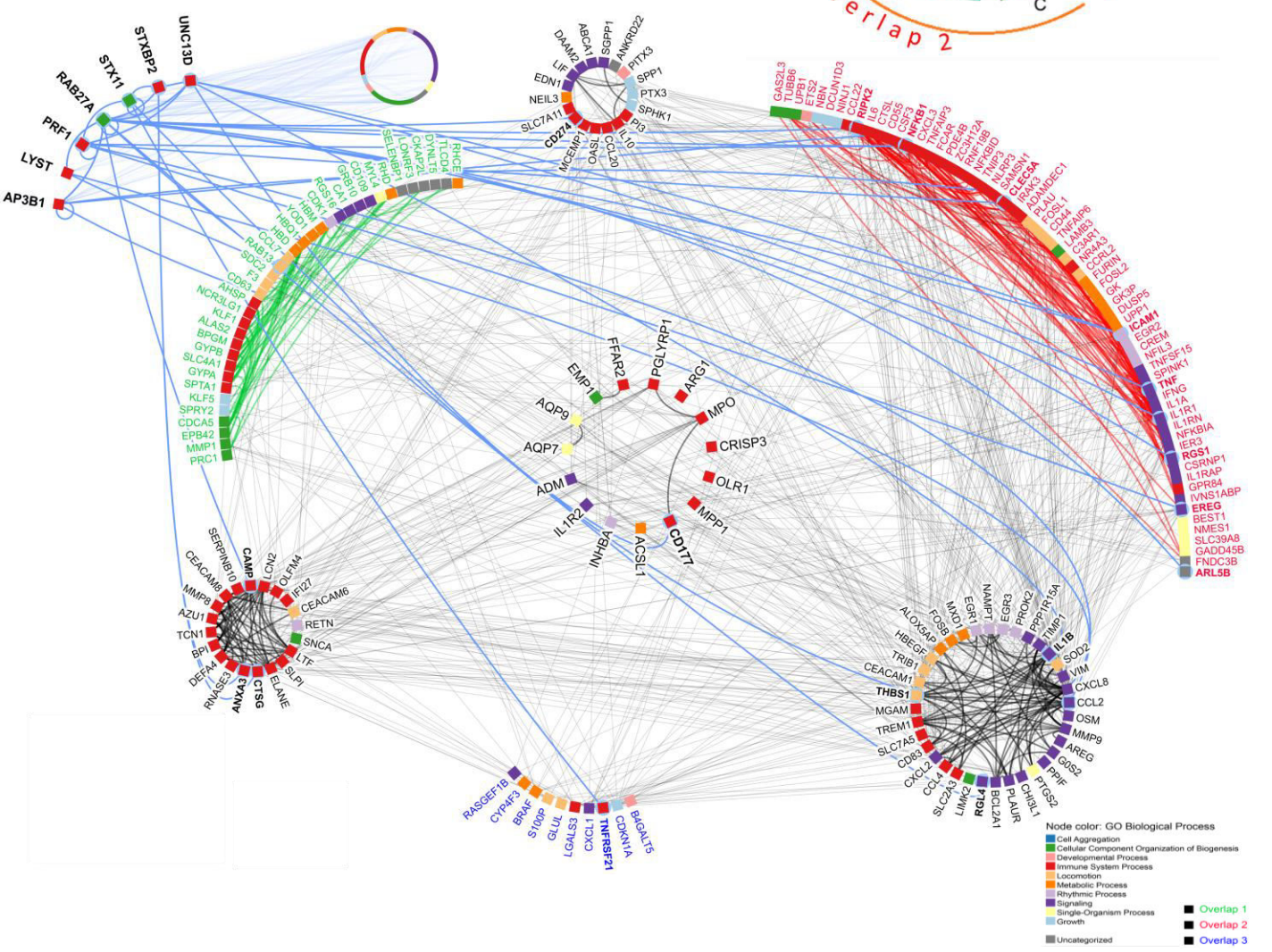


Fig 2

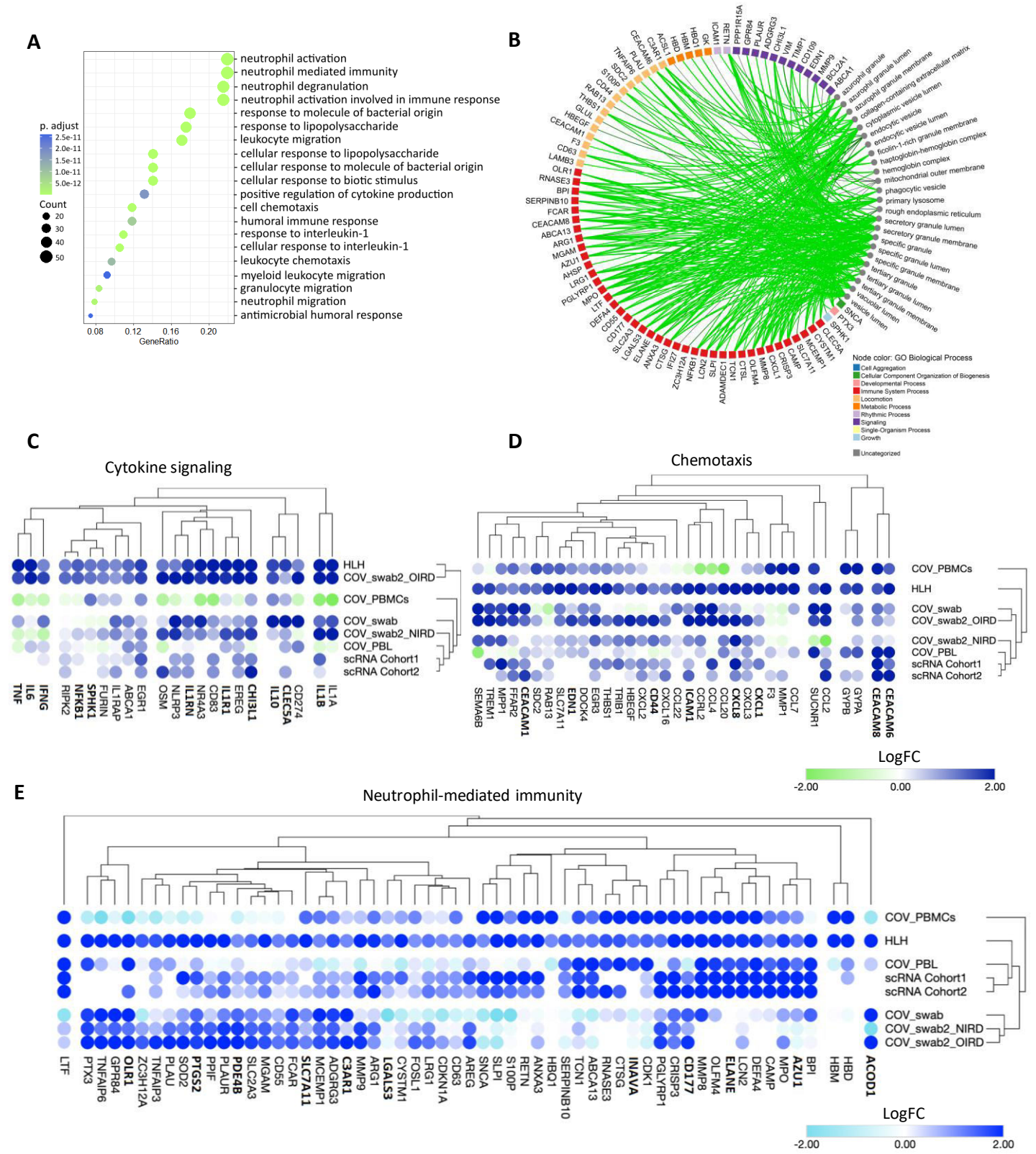

A
p. adjust
2.5e-11
2.0e-11
1.5e-11
1.0e-11
5.0e-12
Count
20
30
40
50
neutrophil activation
neutrophil mediated immunity
neutrophil degranulation
neutrophil activation involved in immune response
response to molecule of bacterial origin
response to lipopolysaccharide
leukocyte migration
cellular response to lipopolysaccharide
cellular response to molecule of bacterial origin
cellular response to biotic stimulus
positive regulation of cytokine production
cell chemotaxis
humoral immune response
response to interleukin-1
cellular response to interleukin-1
leukocyte chemotaxis
myeloid leukocyte migration
granulocyte migration
neutrophil migration
antimicrobial humoral response
0.08
0.12
0.16
0.20
GeneRatio
B
azurophil granule
azurophil granule lumen
azurophil granule membrane
azurophil granule extracellular matrix
collagen-containing extracellular matrix
cytoplasmic vesicle lumen
cytoplasmic vesicle
endocytic vesicle lumen
endocytic vesicle membrane
ficolin-1-rich granule membrane
haptoglobin-hemoglobin complex
hemoglobin complex
mitochondrial outer membrane
phagocytic vesicle
primary lysosome
rough endoplasmic reticulum
secretory granule lumen
secretory granule membrane
specific granule
specific granule lumen
specific granule membrane
tertiary granule
tertiary granule lumen
tertiary granule membrane
vacuolar lumen
vesicle lumen
Node color: GO Biological Process
Cell Aggregation
Cellular Component Organization of Biogenesis
Developmental Process
Immune System Process
Locomotion
Metabolic Process
Rhythmic Process
Signaling
Single-Organism Process
Growth
Uncategorized
C
Cytokine signaling
HLH
COV_swab2_OIRD
COV_PBMCs
COV_swab
COV_swab2_NIRD
COV_PBL
scRNA Cohort1
scRNA Cohort2
D
Chemotaxis
COV_PBMCs
HLH
COV_swab
COV_swab2_OIRD
COV_swab2_NIRD
COV_PBL
scRNA Cohort1
scRNA Cohort2
LogFC
-2.00
0.00
2.00
E
Neutrophil-mediated immunity
COV_PBMCs
HLH
COV_PBL
scRNA Cohort1
scRNA Cohort2
COV_swab
COV_swab2_NIRD
COV_swab2_OIRD
LogFC
-2.00
0.00
2.00

**Fig 3**

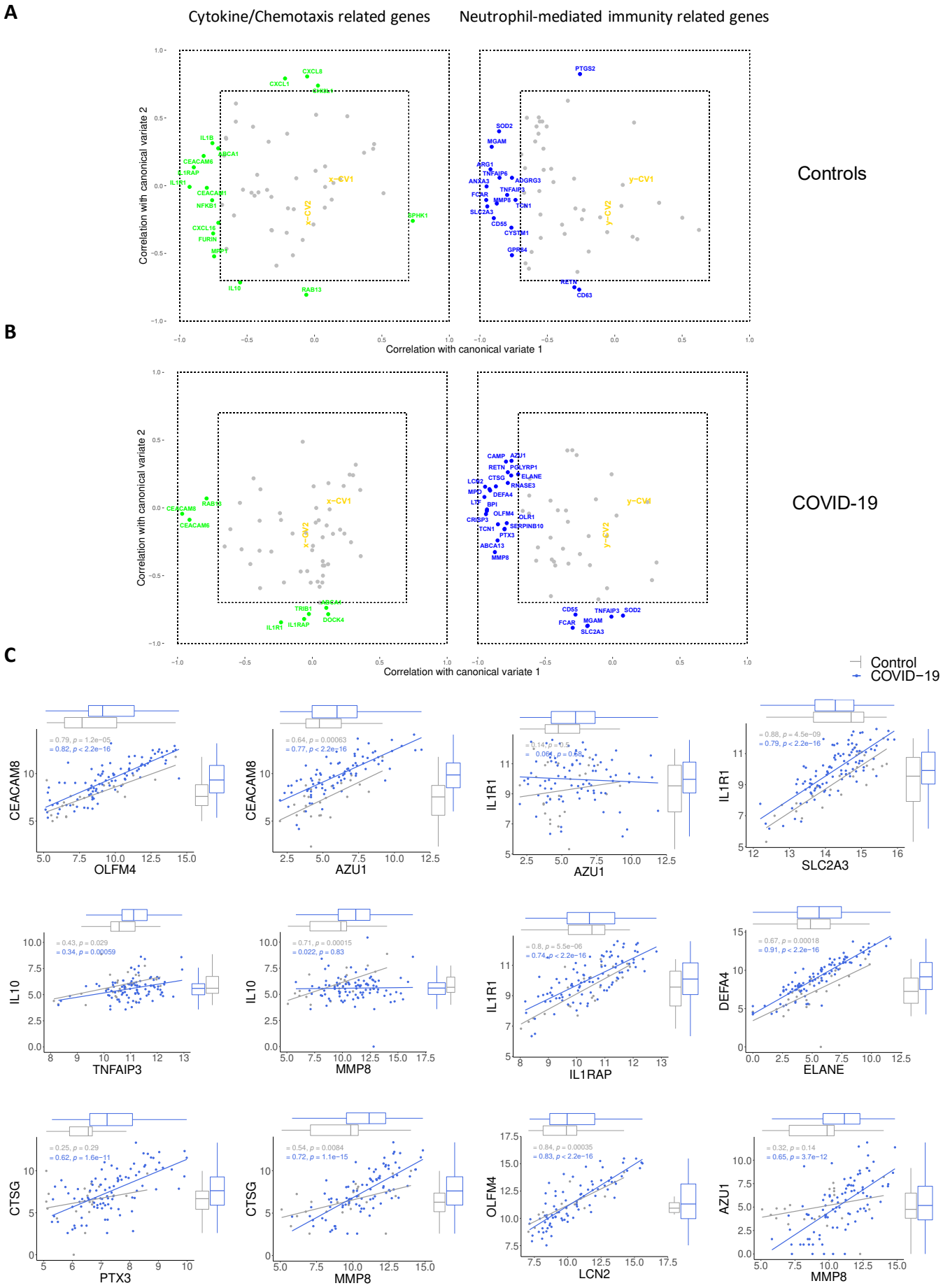

A
Cytokine/Chemotaxis related genes
Neutrophil-mediated immunity related genes
Controls
Correlation with canonical variate 2
Correlation with canonical variate 1
x−CV1
x−CV2
y−CV1
y−CV2
B
COVID-19
C
Control
COVID−19
CEACAM8
OLFM4
AZU1
IL1R1
SLC2A3
IL10
TNFAIP3
MMP8
IL1RAP
DEFA4
ELANE
CTSG
PTX3
LCN2

Fig 4

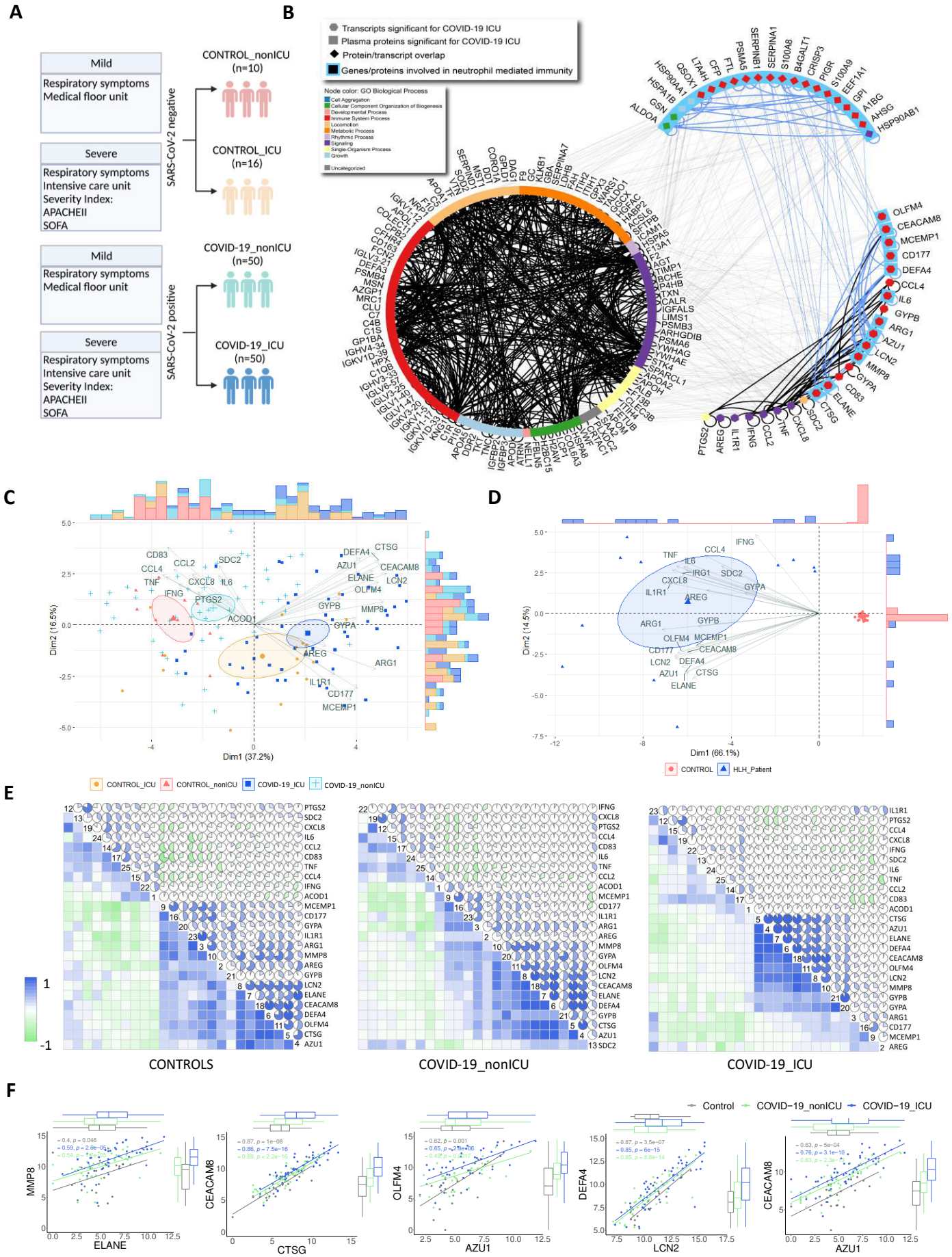

A
Mild
Respiratory symptoms
Medical floor unit
Severe
Respiratory symptoms
Intensive care unit
Severity Index:
APACHEII
SOFA
SARS-CoV-2 negative
CONTROL_nonICU
(n=10)
CONTROL_ICU
(n=16)
Mild
Respiratory symptoms
Medical floor unit
Severe
Respiratory symptoms
Intensive care unit
Severity Index:
APACHEII
SOFA
SARS-CoV-2 positive
COVID-19_nonICU
(n=50)
COVID-19_ICU
(n=50)
B
Transcripts significant for COVID-19 ICU
Plasma proteins significant for COVID-19 ICU
Protein/transcript overlap
Genes/proteins involved in neutrophil mediated immunity
Node color: GO Biological Process
Cell Aggregation
Cellular Component Organization of Biogenesis
Developmental Process
Immune System Process
Locomotion
Metabolic Process
Rhythmic Process
Signaling
Single-Organism Process
Growth
Uncategorized
C
Dim1 (37.2%)
Dim2 (16.5%)
CONTROL_ICU
CONTROL_nonICU
COVID-19_ICU
COVID-19_nonICU
D
Dim1 (66.1%)
Dim2 (14.5%)
CONTROL
HLH_Patient
E
CONTROLS
COVID-19_nonICU
COVID-19_ICU
F
Control
COVID-19_nonICU
COVID-19_ICU
MMP8
ELANE
CEACAM8
CTSG
OLFM4
AZU1
DEFA4
LCN2
CEACAM8
AZU1

Fig 5

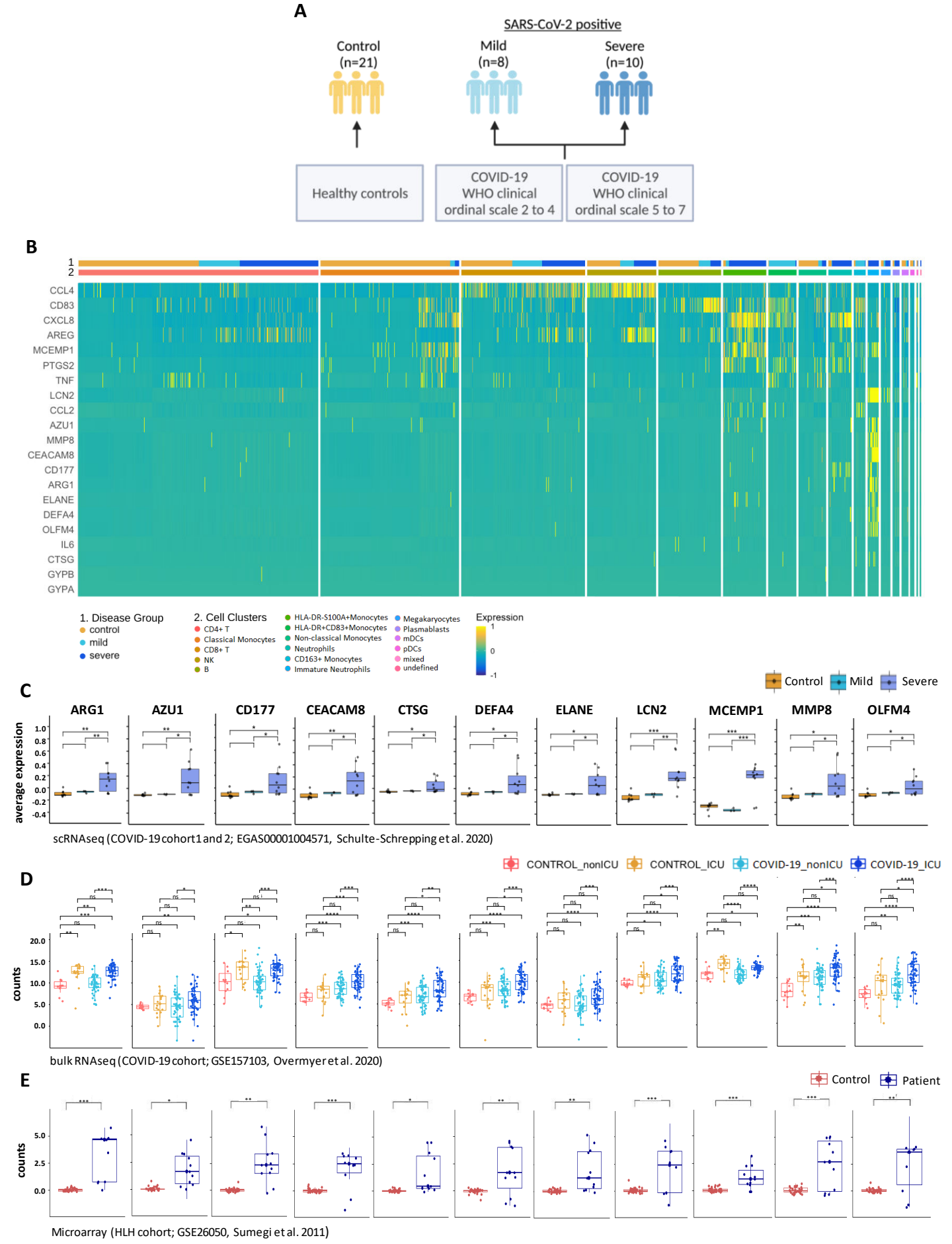
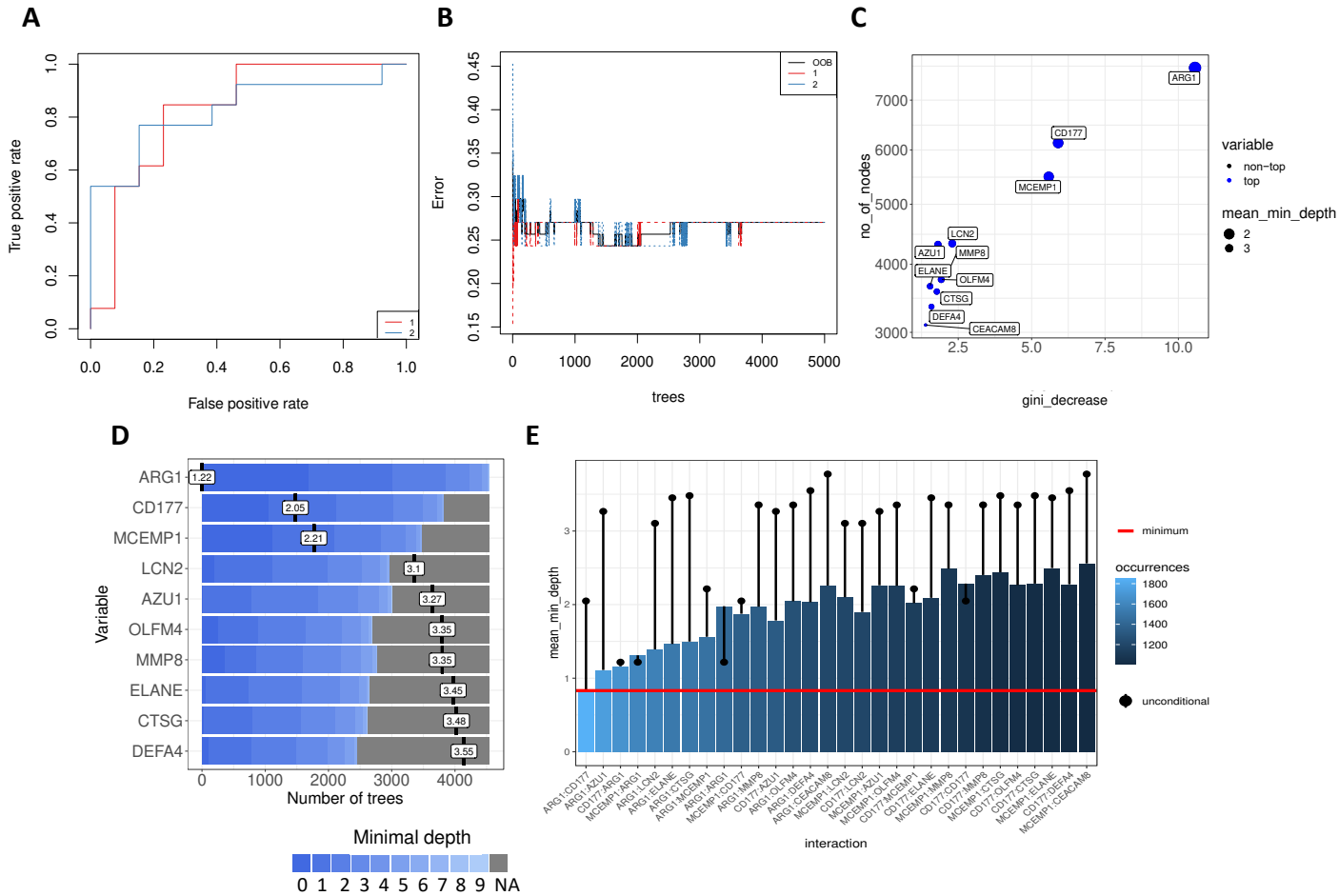

A
SARS-CoV-2 positive
Control (n=21)
Mild (n=8)
Severe (n=10)
Healthy controls
COVID-19 WHO clinical ordinal scale 2 to 4
COVID-19 WHO clinical ordinal scale 5 to 7
B
CCL4
CD83
CXCL8
AREG
MCEMP1
PTGS2
TNF
LCN2
CCL2
AZU1
MMP8
CEACAM8
CD177
ARG1
ELANE
DEFA4
OLFM4
IL6
CTSG
GYPB
GYPA
1. Disease Group
control
mild
severe
2. Cell Clusters
CD4+ T
Classical Monocytes
CD8+ T
NK
B
HLA-DR-S100A+Monocytes
HLA-DR+CD83+Monocytes
Non-classical Monocytes
Neutrophils
CD163+ Monocytes
Immature Neutrophils
Megakaryocytes
Plasmablasts
mDCs
pDCs
mixed
undefined
Expression
C
Control
Mild
Severe
ARG1
AZU1
CD177
CEACAM8
CTSG
DEFA4
ELANE
LCN2
MCEMP1
MMP8
OLFM4
average expression
scRNAseq (COVID-19 cohort1 and 2; EGAS00001004571, Schulte-Schrepping et al. 2020)
D
CONTROL_nonICU
CONTROL_ICU
COVID-19_nonICU
COVID-19_ICU
counts
bulk RNAseq (COVID-19 cohort; GSE157103, Overmyer et al. 2020)
E
Control
Patient
counts
Microarray (HLH cohort; GSE26050, Sumegi et al. 2011)

**Fig 6**


A
True positive rate
False positive rate
1
2
B
Error
trees
OOB
1
2
C
no_of_nodes
gini_decrease
ARG1
CD177
MCEMP1
LCN2
AZU1
MMP8
ELANE
OLFM4
CTSG
DEFA4
CEACAM8
variable
non−top
top
mean_min_depth
2
3
D
Variable
ARG1
CD177
MCEMP1
LCN2
AZU1
OLFM4
MMP8
ELANE
CTSG
DEFA4
1.22
2.05
2.21
3.1
3.27
3.35
3.35
3.45
3.48
3.55
Number of trees
Minimal depth
0 1 2 3 4 5 6 7 8 9 NA
E
mean_min_depth
interaction
minimum
occurrences
1800
1600
1400
1200
unconditional

**KEY RESOURCES TABLE**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Deposited data | | |
| mircoarray data of HLH patients and healthy controls | Sumegi et al. 2011 | GEO: GSE26050 |
| bulk-RNA seq data of COVID-19 patients and controls | Arunachalam et al. 2020 | GEO: GSE152418 |
| bulk-RNA seq data of COVID-19 patients and controls | Liebermann et al. 2020 | GEO: GSE152075 |
| bulk-RNA seq data of COVID-19 patients and patients with other respiratory diseases | Mick et al. 2020 | GEO: GSE156063 |
| bulk-RNA seq data of COVID-19 patients and controls | Ng et al. 2021 | GEO: GSE163151 |
| bulk-RNA seq data of COVID-19 patients and controls | Thair et al. 2021 | GEO: GSE152641 |
| bulk-RNA seq data of COVID-19 patients and patients with other respiratory symptoms | Overmyer et al. 2020 | GEO: GSE157103 |
| scRNA seq data of COVID-19 patients and controls | Schulte-Schrepping et al. 2020 | EGA: EGAS00001004571 |
| Software and algorithms | | |
| NetworkAnalyst 3.0 | Zhou, G. et al. 2019 | https://www.networkanalyst.ca/ |
| Limma voom pipeline | Law, C. W. et al. 2014 | https://doi.org/10.1186/gb-2014-15-2-r29 |
| jvenn | Bardou, P. et al. 2014 | http://jvenn.toulouse.inra.fr/app/example.html |
| Circos online tool | Krzywinski, M. et al. 2009 | http://mkweb.bcgsc.ca/tableviewer/ |
| Seurat V.4 | Hao, Y. et al. 2020 | https://doi.org/10.1016/j.cell.2021.04.048https://satijalab.org/seurat/ |
| NAViGaTOR 3.0.14 | Brown, K. R. et al. 2009 | doi:10.1093/bioinformatics/btp595 |
| Integrated Interactions Database, IID version 2020-05 | Kotlyar M. et al. 2018 | http://ophid.utoronto.ca/iid |
| ClusterProfiler | Yu, G et al. 2012 | doi:10.1089/omi.2011.0118. |
| Enrichr | Chen, E. Y. et al. 2013; Kuleshov, M. V. et al.2016 | https://maayanlab.cloud/Enrichr/ |
| Morpheus | Starruß, J. et al. 2014 | https://software.broadinstitute.org/morpheus/ |
| R version 4.0.5 | R Core Team (2020) | https://www.r-project.org/ |
| R studio Version 1.4.1106 | RStudio Team (2020) | http://www.rstudio.com/. |
| circlize R package | Zuguang Gu et al. 2020 | https://github.com/jokergoo/circlize |
| ggpubr R package | Alboukadel Kassambara, 2020 | https://rpkgs.datanovia.com/ggpubr/ |

| | | |
|---|---|---|
| lemon R package | Stefan McKinnon Edwards et al. 2020 | https://github.com/stefanedwards/lemon |
| ggplot2 R package | Wickham H, 2016 | https://ggplot2.tidyverse.org |
| factoextra R package for PCA analysis | Alboukadel Kassambara and Fabian Mundt, 2017 | https://rpkgs.datanovia.com/factoextra/index.html |
| Canonical Correlation Analysis (CCA) R package | Jendoubi, T. & Strimmer, K., 2019 | https://cran.r-project.org/web/packages/whitening/index.html http://www.strimmerlab.org/software/whitening/ |
| corrgram R package | Kevin Wright, 2021 | https://kwstat.github.io/corrgram/ |
| psych R package | William Revelle, 2021 | https://personality-project.org/r/psych/ |
| inlmisc R package | Jason C Fisher, 2021 | https://github.com/USGS-R/inlmisc |
| tydiverse R package | Hadley Wickham, 2021 | https://github.com/tidyverse/tidyverse/ |
| viridis R package | Simon Garnier et al. 2021 | https://sjmgarnier.github.io/viridis/ |
| randomForest (version 4.6.14) | Andy Liaw and Matthew Wiener 2002 | https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf |
| Fisher´s method | Fisher, R. A., 1932 | Fisher, R. A. Statistical Methods for Research Workers 4th edn. (Oliver and Boyd, 1932). |
| nonparametric MANOVA: nparcomp R package | Konietschke, F. et al., 2015 | doi:10.18637/jss.v064.i09 |
| nonparametric Inference for Multivariate Data: npmv R package | Burchett, W. et al., 2017 | doi:10.18637/jss.v076.i04 |