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

Jun Ding¹, David Earl Hostallero², Mohamed Reda El Khili², Gregory Fonseca³, Simon Millette⁴, Nuzha Noorah³, Myriam Guay-Belzile³, Jonathan Spicer⁵, Noriko Daneshtalab⁶, Martin Sirois⁷, Karine Tremblay⁸, Amin Emad^{2,*} and Simon Rousseau^{3,*}

¹Computational Biology Department, Carnegie Mellon University, Pittsburgh, PA, USA, 15204

²Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada.

³The Meakins-Christie Laboratories at the Research Institute of the McGill University Health Centre Research Institute, 1001 Boul. Décarie, Montréal, H4A 3J1, Canada.

⁴Goodman Cancer Research Centre, McGill University

⁵Division of Thoracic and Upper Gastrointestinal Surgery, McGill University Health Centre Research Institute, 1001 Boul. Décarie, Montréal, H4A 3J1, Canada.

⁶School of Pharmacy, Memorial University of Newfoundland, 300 Prince Philip Drive, Health Sciences Center, St. John's, Newfoundland, Canada, A1B 3V6

⁷Montreal Heart Institute and Department of pharmacology and physiology, Faculty of medicine, Université de Montréal.

⁸Pharmacology-physiology Department, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Centre intégré universitaire de santé et de services sociaux du Saguenay–Lac-Saint-Jean (Chicoutimi University Hospital) Research Center, Saguenay, QC, Canada.

*Address correspondence to:

Simon Rousseau, The Meakins-Christie Laboratories at the RI-MUHC, 1001 Boul. Décarie, Montréal, H4A 3J1, Canada, Phone: +1-514-934-1934 (76394), Email: simon.rousseau@mcgill.ca

or

Amin Emad, 755 McConnell Engineering Building, 3480 University Street, Montreal, Quebec, H3A 0E9, Canada, Phone: +1-514-398-1847, Email: amin.emad@mcgill.ca

Abstract

Abnormal coagulation and an increased risk of thrombosis are features of severe COVID-19, with parallels proposed with hemophagocytic lymphohistiocytosis (HLH), a life-threatening condition associated with hyperinflammation. The presence of HLH was described in severely ill patients during the H1N1 influenza epidemic, presenting with pulmonary vascular thrombosis. We tested the hypothesis that genes causing primary HLH regulate pathways linking pulmonary thromboembolism to the presence of SARS-CoV-2 using novel network-informed computational algorithms. This approach led to the identification of Neutrophils Extracellular Traps (NETs) as plausible mediators of vascular thrombosis in severe COVID-19 in children and adults. Taken together, the network-informed analysis led us to propose the following model: the release of NETs in response to inflammatory signals acting in concert with SARS-CoV-2 damage the endothelium and direct platelet-activation promoting abnormal coagulation leading to serious complications of COVID-19. The underlying hypothesis is that genetic and/or environmental conditions that favor the release of NETs may predispose individuals to thrombotic complications of COVID-19 due to an increase risk of abnormal coagulation. This would be a common pathogenic mechanism in conditions including autoimmune/infectious diseases, hematologic and metabolic disorders.

Introduction

Early reports from China^{1,2} and France^{3,4}, have highlighted abnormal coagulation and a high risk of thrombosis in severe COVID-19. Evidence of both micro-thrombosis⁵ and macro-thrombosis⁶ have emerged. Understanding the mechanism by which SARS-CoV-2 infections lead to coagulopathy is an important goal of the current research effort to properly identify and treat individuals at risk of severe complications of COVID-19.

A strong pro-inflammatory but ineffective anti-viral response contributes to COVID-19⁷. This can lead to a hyperinflammatory state described in a subset of COVID-19 adult patients with severe and often deadly complications⁸. This hyperinflammation in severe COVID-19 patients includes elevated levels of C-reactive protein (CRP), ferritin, fibrinogen and D-dimers⁹⁻¹¹. Parallels have been drawn between hematologic cytokine storms and severe COVID-19 pointing to hemophagocytic lymphohistiocytosis (HLH) as a putative link¹². HLH is a life-threatening condition associated with hyperinflammation resulting in varying clinical manifestations such as fever, organomegaly, coagulopathy, cytopenia, neurologic dysfunction, and elevations of acute phase reactants. The presence of HLH was also described in severely ill patients during the novel H1N1 influenza epidemic of 2009, presenting with peripheral pulmonary vascular thrombosis¹³⁻¹⁶. Interestingly, the post-mortem lungs from patients with COVID-19 showed severe endothelial injury associated with the presence of intracellular virus and alveolar capillary microthrombi. This was 9 times more prevalent in COVID-19 than H1N1 lungs analysed following autopsy¹⁷. Accordingly, in one study, COVID-19 Acute Respiratory Distress Syndrome (ARDS) developed significantly more thrombotic complications than non-COVID-19 ARDS⁴. HLH has also been observed in children suffering from Kawasaki Disease (KD), a form of systemic vasculitis in children¹⁸. A 30-fold increase of incidence of KD-like illness has been observed in Bergamo, Italy following the SARS-CoV-2 outbreak¹⁹, as well in Paris, France, preprint²⁰ and the UK²¹. In the KD-like illness described in the Italian study, half of the patients were diagnosed with Macrophage-Activation Syndrome (MAS), a condition closely related to HLH. Interestingly, hyperinflammation, rather than the hemophagocytosis appears to be the driving cause of pathology²². Therefore, vascular injury resulting from hyperinflammation in the alveoli in response to SARS-CoV-2 may be an important pathophysiological mechanism of severe illness in children and adults.

Based on the above reports, we hypothesized that genes causing the genetic forms of HLH and associated syndromes regulate pathways linking the risk of pulmonary microvascular thromboembolism to the presence of SARS-CoV-2. To test this hypothesis quickly in view of the world-wide pandemic, we used network-informed approaches, including a novel computational tool to explore the link between HLH genes and a network of human proteins identified to interact with SARS-CoV-2²³. This approach led to the identification of Neutrophils Extracellular Traps (NETs) as plausible mediators of vascular thrombosis in severe COVID-19.

Results

HLH genes are significantly enriched within the SARS-CoV-2 host protein interactome

In the case of the SARS-Cov-2 pandemic, with widespread impact across the world, there is an urgency that requires the adaptation of different strategies to understand COVID-19. In this paper, we exploited the knowledge existing within protein interaction networks to identify the molecular pathways underpinning thrombotic complications of COVID-19 using advanced computational algorithms. As described in the introduction, a subset of patients suffering from severe complications of COVID-19 present clinically with symptoms similar to HLH. Therefore, we have assembled a list of candidate genes responsible for primary HLH and associated syndromes to explore their relationships with COVID-19^{24,25} (**Supplementary Table S1**).

The first question asked was whether these HLH genes had potential interactions with SARS-CoV-2. We assembled a protein interaction network between the SARS-CoV-2 host interaction protein network recently published²³ and the HLH genes using an algorithm that we created for this purpose, GeneList2COVID19. The algorithm establishes the shortest path between the candidate genes and the known host interacting proteins with SARS-CoV-2 and calculates an overall connectivity score for the network (a smaller value represents a greater connectivity) (**Fig 1 and Supplementary Table S1**). We computationally validated the predictions of the GeneList2COVID19 to identify significant interactions. To demonstrate that the method can assign significant connectivity scores to genes associated with COVID-19, we obtained a list of 10 confirmed COVID-19 related genes²⁶, which are differentially expressed in severe COVID19 patients (**Supplementary Table S1**). We then calculated the “COVID-19” connectivity score for those 10 genes (SA) as well as all the genes (SB) using GeneList2COVID19. We found that SA is significantly ($p\text{-value}=0.017$) smaller than SB, which indicates that those 10 COVID-19 related genes are indeed “significantly connected” to SARS-CoV-2 proteins (**Fig. 2**). To show the specificity of the method, we also calculated the “COVID19” connectivity score for 100 randomly selected genes (SC) and compared it to the connectivity score of all genes (SB). We found that SC is NOT significantly smaller than B (the background) ($p\text{-value}=0.106$) (**Fig. 2**). In other words, those 100 random genes are not “significantly connected” to SARS-CoV2 proteins, which reflects the fact that those genes were randomly picked. As an additional control, we repeated the analysis using genes linked to male infertility²⁷, a condition that has not been associated with COVID-19 (**Supplementary Table S1**). The connectivity score was not significantly different from all other genes ($p\text{-value}=0.872$), further demonstrating the specificity of the GeneList2COVID19, which is not restricted to random genes but can also discriminate gene lists associated with other conditions (**Fig. 2**). After the method was validated, we compared the “connectivity” score for HLH genes listed above with all genes that connect to SARS-CoV-2 proteins through our assembled protein-protein interaction network (**Fig. 2**). We found that the score for the HLH marker genes is significantly smaller compared to all other genes ($p\text{-value}=0.0082$, one-sided rank sum test) (**Fig. 2**). As an additional control, we compared the HLH genes to a list of vascular angiogenesis genes linked to both H1N1 and SARS-CoV-2 pulmonary infections¹⁷ (**Supplementary Table S1 & Fig. 2**). The HLH genes’ connectivity score was smaller, which means that those genes had closer theoretical interactions to SARS-CoV-2. This suggests that HLH genes and their associated pathways are of high interest in the study of SARS-CoV-2 infections.

Differential expression of HLH genes in health conditions related to COVID-19.

We next investigated whether the expression of HLH genes in lung tissue were highly regulated (up or down) in conditions associated with COVID-19 (Sex, Smoking, Lung Cancer, Chronic

Obstructive Pulmonary Disease (COPD), Diabetes, Hypertension and Age) using transcriptomic data sets publicly available^{10,28,29}. All the genes available in each dataset were first sorted under each condition, based on the log fold change (Condition vs. Control) with the top 25% genes classified as High (H), having the largest fold change (higher in Condition), and the bottom 25% genes classified as Low (L), having the smallest fold change (much higher in Control). Then, the 11 HLH genes were assessed to determine whether they are among the H or L genes in each condition.

We hypothesized that HLH genes, that may play an active role in thrombotic complications of COVID-19, are more likely to be regulated in co-morbid conditions. *RAB27A* expression was found altered in all studied conditions, while *AP3B1* expression was also found altered in all conditions except one, lung cancer (**Fig 3**). The protein encoded by the *RAB27A* gene has emerged as a central regulator of the neutrophil response through its regulation of vesicular trafficking and degranulation including azurophilic granule exocytosis^{30,31}. *RAB27A* expression was low in females, aged individuals and hypertension whereas it was highly expressed in smoking, cancer and diabetes. The protein encoded by the *AP3B1*, part of the AP-3 complex that is ubiquitously expressed, mediates proteins sorting from the endosomal and trans-Golgi network to lysosomes and endosome-lysosome-related organelles³². *AP3B1* expression was low in females, COPD and diabetes but high in smoking, hypertension and aged individuals. It is always hazardous to infer changes in activity based on changes in gene expression, but the pattern of results shows that genes causing primary HLH have their expression regulated in many of the co-morbid conditions of COVID-19. Moreover, by looking at the clustering pattern emerging from the HLH genes, the following conditions were clustered: smoking and lung cancer, COPD and diabetes, hypertension and aging. Overall, this defines potentially shared risk for disease severity within these pairs in regard to their pathophysiology. Treatment approaches aimed at decreasing the risk of thrombosis may differ between these groups.

HLH genes are functionally associated with neutrophils degranulation

The next crucial question was to determine what function regulated by HLH genes may be linked to the risk of thrombosis in COVID-19. Gene ontology (GO) enrichment analysis was done to identify plausible biological functions regulated by HLH genes using the KnowEnG analytical platform (<https://knoweng.org/analyze>)³³. KnowEnG is a computational system for analysis of 'omic' datasets in light of prior knowledge in the form of various biological networks. After correcting for multiple hypothesis (Benjamini-Hochberg method), the standard mode (no network used) did not find any significant GO terms (False discovery rate (FDR)>0.05). However, the knowledge-guided mode that used three different networks (protein-protein interaction (PPI), co-expression network from the STRING database³⁴ and HumanNet Integrated network³⁵ identified several GO terms (**Table 1 and Supplementary Table S2**). Neutrophil degranulation was the top GO term identified with all three networks above.

While HLH is traditionally associated with hyper activation of T-cells and macrophages, an investigation published in 2019 showed that in addition to T-cells, targeting neutrophils provided benefits in a murine model, indicating an underappreciated role of neutrophils in this syndrome³⁶. Other GO terms point to the importance of vesicular trafficking/exocytosis (intracellular protein transport, exocytic vesicle, positive regulation of exocytosis, positive regulation of regulated secretory pathway, regulation of mast cell degranulation, melanosome organization, SNARE complex, lysosome and Weibel-Palade body) and host defense response (cellular defense response and defense response to virus) as would be expected for genes associated

with diseases that presents with coagulation and inflammation related processes. A main finding of this study is a proposed role for neutrophil degranulation in HLH and by extension in COVID-19 coagulopathies. Amongst the important functions associated with neutrophil degranulation is the release of NETs³⁷.

NET-release, a clue to the risk of SARS-CoV-2 infection complications in rheumatologic pediatric diseases?

Emerging evidence has suggested an association between KD-like illness (KD) and severe complications of COVID-19 in children^{21,38}. Of note, Neutrophil Extracellular Traps (NETs) formation is increased in KD-derived neutrophils³⁹. NETs are protein-coated DNA extruded from neutrophils that directly induce dose-dependent platelet aggregation and dense granule secretion⁴⁰, playing an important role in coagulopathies⁴¹. NET formation can occur via two mechanisms, either a cell death-mediated process (NETosis) occurring over a period of hours or a non-lytic release mechanism via degranulation and expulsion of chromatin that is initiated within minutes of activation³⁷. HLH has been linked to NET induction and production of cytokines such as IL-1 β , IL-6, IL-8 and IL-17⁴², elevated in both KD and COVID-19⁴³. Reactive HLH presents clinically as primary HLH (of genetic origin) and has been described in other pediatric rheumatologic diseases where NETs could also play key pathogenic roles such as in systemic juvenile idiopathic arthritis (sJIA) and lupus^{44,45}. NETs have been proposed to be part of the pathogenesis of sJIA or MAS secondary to sJIA⁴⁶. We did not have access to transcriptomic data from KD-derived neutrophils. However, we were able to access such a dataset from sJIA⁴⁷, to investigate the expression of the HLH-SARS-CoV-2 interactome (**Supplementary Table S1**). The majority of the genes were expressed in neutrophils and 7 of them (UNC13D, LYST, AP3B1, MAGT1, RAB8A, GOLGA2 and G3BP1) were significantly elevated in either inactive or active sJIA compared to control neutrophils (**Fig. 4**). It is worth noting that the most connected gene to SARS-CoV-2, AP3B1 that can directly interact with SARS-CoV-2 E protein, is in the list of up-regulated genes in neutrophils derived from sJIA. An important message stemming from our discovery that NETs may be drivers of coagulopathies in reactive HLH is the potential susceptibility of a subset of the pediatric population, identifying them as at risk of severe complication of COVID-19. This led us to the next step, predicting potential vulnerable populations to thrombotic complications of COVID-19 based on their susceptibility to release NETs.

Identifying potentially vulnerable populations to COVID-19 based on NETs release.

The World Health Organization has established that identifying vulnerable populations is an urgent public health priority in the context of the COVID-19 pandemic⁴⁸. Based on the analyses above, NETs may play an important role in promoting thrombosis in COVID-19. The role of neutrophils in coagulopathies is becoming increasingly recognised and particularly that of NETs⁴¹. Therefore, we hypothesized that health conditions associated with increased release of NETs could be a predictive factor for thrombotic complications of COVID-19. Based on this hypothesis and in order to identify vulnerable populations, we developed a method called foRWaRD (informative Random Walk for Ranking Diseases) to rank different diseases associated with NETs based on their relevance to a gene set of interest (here genes in the HLH-SARS-CoV-2 network) (see Methods for details).

We obtained two NET gene signatures from previous studies^{49,50} and combined them to obtain a list of NET-associated genes (**Supplementary Table S1**). Then, we obtained the list of diseases associated with these genes from the DisGeNET database⁵¹. Entries with gene-disease

association (GDA) <0.4 , were filtered out to arrive at a set of 99 diseases. We used the set of 24 genes in the HLH-SARS-CoV-2 interaction network (**Fig. 1**) as the query set and used the HumanNet Integrated network³⁵ as the gene interaction network in foRWARD. The full list of 99 diseases, ranked using foRWARD, is provided in **Supplementary Table S3**. **Table 2** illustrates the diseases with a Normalized Disease Score (NDS) greater than 0.5, meaning that these diseases are enriched above the background probabilities. Most of the 10 top-ranked diseases associated with genes linked to NETs can be sub-grouped in 4 major categories 1- immune/infectious (Alzheimer's disease, Immunodeficiency 8); 2- cardiovascular (Myocardial reperfusion injury, Hemolytic anemia due to G6PI, Bleeding disorder type; 15); 3-metabolic (Diabetes) and 4) Cancer (Liver carcinoma). It is important to note that some of these diseases would be putatively associated with NET-deficiency such as Immunodeficiency 8, more often presenting clinically with bleeding. Whether patients suffering from these diseases are protected from thrombotic-complications of COVID-19 remains to be determined. However, diseases associated with increased NET release are expected to yield greater risk of thrombosis and may identify vulnerable populations to severe thrombotic complications of COVID-19.

Discussion

Based on recent literature, we hypothesized that severe pulmonary thrombotic complications of COVID-19 are associated with a hematologic cytokine storm that could be, in part, defined using genes causing HLH. The network-informed analysis presented in this paper, revealed that 1) the top GO biological function associated with HLH genes is neutrophil degranulation, consistent with a recent report highlighting the undervalued role of neutrophils in HLH³⁶; 2) HLH genes are significantly enriched with the SARS-CoV-2 human interactome; 3) the top-ranked HLH gene, *AP3BI*, has roles in cargo loading of type II pneumocytes, where it may interact with SARS-CoV-2 to disturb surfactant physiological functions to promote inflammation/pro-coagulation activities; 4) diseases/syndromes-associated with increased release of Neutrophil Extracellular Traps (NETs) may predict vulnerable populations, including those affecting children.

Taken together, the network-informed analysis led us to propose the following model: the release of NETs in response to inflammatory signals acting in concert with SARS-CoV-2 damage the endothelium and direct platelet-activation promoting abnormal coagulation leading to serious complications of COVID-19 in susceptible individuals (**Fig. 5**). The underlying hypothesis is that genetic and/or environmental conditions that favor the release of NETs may predispose individuals to thrombotic complications of COVID-19 due to an increase risk of abnormal coagulation. This would be a common pathogenic mechanism amongst numerous conditions including autoimmune/infectious diseases, hematologic and metabolic disorders.

The role of neutrophils in coagulopathies is becoming increasingly recognised and particularly that of NETs⁴¹. Interestingly, elevated Neutrophils count is the best single leukocyte predictor of cardiovascular risk⁵², bettered only by the combination of high neutrophils to low lymphocytes ratio⁵², a clinical feature of COVID-19¹⁰. NET release can be triggered by various inflammatory mediators found elevated in severe COVID-19, including CRP, IL-1 β , IL-6 and IL-8⁴³. There is also a positive correlation between circulating serum of IL-6, IL-8, CRP and NET levels⁵³. NETs are found in a variety of conditions such as infection, malignancy, atherosclerosis, and autoimmune diseases with reports now emerging that describe their presence in COVID-19^{9,54-56}. Amongst the known diseases associated with NETs, several are related to children including

Cystic Fibrosis⁵⁷, Meningococcal Sepsis⁵⁸, Lyme Neurobiellosis⁵⁹; Juvenile Dermatomyositis⁶⁰ and pediatric inflammatory bowel diseases⁶¹. In pediatric sepsis, NETs levels were elevated and correlated with disease severity, mirroring results in mice where higher NETs levels in response to lipopolysaccharides are found in infant mice compared to adults⁶².

One of the quickest ways to decrease the burden of COVID-19 on the health care systems throughout the world is to identify at-risk populations to emphasize the importance of infection prevention measures for those individuals. Since these measures incur high personal, social and economic costs, a precise knowledge is essential. We presented a novel computational algorithm that enabled us to identify potential diseases linked with NETs (**Table 2**). Interestingly, amongst the identified diseases, Diabetes, a well-established comorbidity of COVID-19⁶³, is ranked 4th and 7th. Our study provides additional insight into the potential mechanisms involved, with increase NETs formation resulting from the underlying chronic inflammation as a key factor promoting coagulopathies in diabetics suffering from COVID-19. As for the top ranked disease, Alzheimer's Disease, whether NETs in the brain can lead to an increased risk of systemic thrombosis looks less likely than the reverse, that SARS-CoV-2 infection may increase NETs release in the brain that could exacerbate Alzheimer's disease-driven pathology including a greater risk of stroke. This may be an important question for future studies due to the susceptibility and severity of the elderly to COVID-19 and notably the extreme mortality seen in long-term care home arounds the world where cognitive impairment is highly prevalent^{64,65}.

It has been suggested that COVID-19 should be added to this list of hyperferritinemic syndromes, which includes adult-onset Still's disease, septic shock, catastrophic anti-phospholipid syndrome, and MAS (reactive HLH)⁶⁶. Collectively, these diseases may share similar underlying factors of complications, including an underappreciated role of NETs leading to coagulopathies. It is possible that individuals can unfortunately contract SARS-CoV-2 infection in addition to other factors that underlie any of these conditions (other viruses for example), which may lead to further amplification loop. PCR-negative SARS-CoV-2 patients presenting with clinical symptoms of hyperferritinemic syndrome should be considered highly vulnerable and appropriate infection control measures should be put in place.

Disorders associated with bleeding should decrease the risk of thrombotic complications of COVID-19 (however they may still lead to severe COVID-19 via other mechanism). Nevertheless, they can be informative on pathophysiology. The strongest connectivity to SARS-CoV-2 E protein was AP3B1 (**Fig. 1**). Loss of function of *AP3B1* leads to Hermansky-Pudlak syndrome type 2 that is associated with bleeding and coagulation defects^{67,68}. The SARS-CoV (2002-2003 strain) E (envelop) protein is mainly localised in cells at the ER-Golgi compartment of cells⁶⁹, where it participates in assembly, budding, and intracellular trafficking of viruses⁷⁰. Therefore, both proteins have a coherent subcellular localization supporting their potential interaction. Moreover, in post-mortem immunohistochemical analysis of lung tissue, the SARS-CoV-2 S (spike) and E proteins were found to localize with the respiratory epithelia, the interalveolar, and the septal capillaries⁵. In addition, septal and intra-alveolar neutrophilia was observed⁵, colocalizing some of the key players of a neutrophil-driven SARS-CoV-2 enhanced coagulation cascade in COVID-19 (**Fig. 5**). Whether SARS-CoV-2 E protein can directly or indirectly penetrate neutrophils and/or platelets remains unknown, as these cells are not reported to highly express ACE2+/TMPRSS2+, the two key host proteins for viral entry. However, both of these proteins are highly expressed on type II pneumocytes⁷¹, where AP3B1 is important for cargo loading of lamellar bodies⁷². A postmortem examination in a COVID-19 patient who succumbed

to a sudden cardiovascular accident revealed SARS-CoV-2-viruses present in pneumocytes despite PCR-negative nasal swabs⁷³, indicating a prolonged risk in the lower airways for complications.

Immunodeficiency 8, resulting from the loss of function of coronin-1, also leads to bleeding. Coronin-1 plays key functions in PMN trafficking⁷⁴ in part via its interaction with the integrin $\beta 2$. $\beta 2$ integrin-mediated systemic NET release is a viral mechanism of immunopathology in hantavirus-associated disease such as kidney and lung damage⁷⁵, similar to the immunopathology in severe COVID-19. Overall, diseases associated with a putative loss-of-function of NETs suggest mechanistic roles for AP3B1, coronin-1 and integrin $\beta 2$ in regulating NET-mediated coagulopathies in the lung alveolar and peri-alveolar areas (Fig. 5).

The analysis in this study is based on a new algorithm that we develop (freely available at <https://github.com/phoenixding/genelist2covid19>). GeneList2COVID19 can systematically evaluate the connection of any given gene list to SARS-CoV-2 proteins both within-host proteins and between host-viral proteins. Therefore, it can be used to study a wide variety of biological problems associated with COVID-19, especially in circumstances where experimental data on COVID-19 (e.g. transcriptomics or genomics) is not yet available for the problems of interest. The algorithm was found effective, on positive (proven to be associated with COVID-19) and negative (irrelevant to COVID-19) gene lists. In terms of limitation, GeneList2COVID19 is dependent on the prior knowledge of the protein interactome within the host, and between the host and virus. Currently, we have a well-established protein-protein interactome for the human species. However, the interactome between the SARS-CoV-2 proteins and human proteins is relatively limited²³, since such an interactome is far from complete. For example, there are no reported interactions for ACE2 and TMPRSS2, which are critical to SARS-CoV-2 infection. We provided an option (-v) in GeneList2COVID19 to utilize any new host-viral protein interactome data when it becomes available. While, GeneList2COVID19 is good at telling whether an input gene list is associated with COVID-19, it cannot test the mechanistic hypothesis generated. It can provide the network that connects the genes in the list of the SARS-COV-2 proteins, but it cannot determine which nodes/edges in the network is more critical (and when they are activated). At this stage, the most useful information is derived from considering the entire network. As the availability of COVID-19 related “-omics” data increases, we will extend the method into a joint-model that integrates all those omics data for a more comprehensive, high-definition network model that can provide additional and more precise insights for the role of genes in COVID-19. The second computational algorithm provided, forWaRD (<https://github.com/ddhostallero/forWaRD>), is also limited by the requirement of known gene-disease associations. As a result, this method does not work with syndromes/conditions that do not have a well-established cause (and associated genes) or are from infectious origins, an important trigger of NETs formation.

Conclusion: NETs as a determinant of severe thrombotic complications of COVID-19

The results presented herein points to neutrophils and the release of NETs as important mediators of coagulopathies in COVID-19. In a murine model of the 2002-2003 SARS-CoV infection, respiratory failure was associated with lung neutrophilia⁷⁶. Accordingly, in a study investigating autopsy-derived lung tissue of COVID-19, septal and intra-alveolar neutrophilia was observed⁵, which could contribute to pulmonary embolism atypical from other ARDS. In addition, the levels of NETs correlate with poor prognosis in severe influenza A viral infections⁷⁷. Taken together, this supports a central role of NETs in the risk of thrombotic complications of COVID-19, that may be particularly relevant for children suffering from underlying rheumatologic or infectious conditions. Further studies in well-defined cohorts of COVID-19 patients are mandatory to confirm the relevance of the observations highlighted in the present study. Such knowledge may

be of importance in novel COVID-19 severity biomarkers identification that will be needed in the management of individuals at risk of complications.

Methods

Datasets

For this study we used the following datasets: The interactions between the SARS-CoV-2 proteins and host proteins²³, reporting 332 interactions that involve 26 SARS-CoV-2 proteins and 332 human host proteins. Each interaction in the map was assigned an interaction score that represents the strength of the interaction (a score between 0 and 1). We also collected the protein-protein interactions (with interaction scores) between all human host proteins from the HIPPIE database⁷⁸. We obtained a list of Highly/Lowly (H/L) expressed genes under different health conditions that potentially associated with COVID19^{10,28,29}. To identify the vulnerable populations using foRWARD, the full list of genes associated with these diseases were downloaded from the DisGeNET database (www.disgenet.org)⁵¹. We downloaded the HumanNet Integrated network³⁵ from KnowEnG's Knowledge Network version 17.06 (https://github.com/KnowEnG/KN_Fetcher/blob/master/Contents.md) and used as input to foRWARD.

Network-guided gene set enrichment analysis using KnowEnG

We used the gene set characterization pipeline of KnowEnG³³ (www.knowng.org/analyze) in the standard mode and in the knowledge-guided mode to identify GO terms associated with the HLH genes. The standard mode performs Fisher's exact test, while its knowledge-guided mode is an implementation of DRaWR⁷⁹, a method that utilizes RWRs in order to incorporate gene-level biological networks in the enrichment analysis to improve identification of important pathways and GO terms. For the knowledge-guided mode, we used three biological networks to augment the GO enrichment analysis: the experimentally verified protein-protein interaction (PPI) and the co-expression networks from the STRING database³⁴ as well as the HumanNet Integrated network³⁵. For the analysis, we did not use the bootstrapping option, selected homo sapiens as 'species', and used default values for all other parameters. We obtained GO terms with a "difference score" above 0.5. This score represents the normalized difference between the query probabilities and the baseline probabilities in the RWR algorithm, with the best score observed as 1 (Table 1).

Building the HLH-SARS-CoV-2 interaction network

We first built a network that connects all the SARS-CoV-2 proteins and the human host proteins based on the collected protein interaction data. The edges connecting different proteins are weighted based on the interaction scores obtained from the original datasets above. Next, we inferred the signaling paths from SARS-CoV-2 proteins down to a list of proteins (genes) of interest.

A few key assumptions must be made before we can make such inference. First, since collected protein interactions within the host (and between the host and the SARS-CoV-2 virus) do not have directions, the reconstructed network graph is undirected. Here, we assumed that the information (i.e., infection) flows from the SARS-CoV-2 proteins to the proteins that directly interact with SARS-CoV-2 proteins, next to other intermediate signaling proteins, and finally to the target genes (proteins) of interest. There might be multiple intermediate proteins residing

between the direct SARS-CoV-2 interacting proteins and the target proteins of interest. Second, we did not allow loops in our path from SARS-CoV-2 proteins to the target proteins to reduce the computation complexity. Although feedback loops have been reported in previous studies⁸⁰, they are still relatively rarely observed in the human protein-protein network⁸¹. Last, we assumed that the interaction score between two proteins is proportional to the strength (or the likelihood) of their interaction. A larger interaction score represents either a stronger or more likely interaction, which results in a “stronger” connection edge in both cases.

The objective of the analysis was to find the strongest (or most likely) “connecting” path from SARS-CoV-2 proteins to the target proteins (genes) of interest in the constructed network, where the connection strength was quantified by a “connectivity score”. We formulated the above problem as:

$$MP(s, t) = \underset{P \in Q}{\operatorname{argmax}} \prod_{e \in P} \Pr(e) \quad (1)$$

where $MP(s, t)$ is the most likely path from source node s to the target node t , Q is the union set of all possible paths P from s to t , and e is the edge on the path P . $\Pr(e)$ is the probability of the edge e , which is the interaction score (range $[0, 1]$) between two proteins (ends) of the edge. Since the protein-protein interaction scores were all in the range of $[0, 1]$, we directly used them to represent the interacting probabilities. The above equation can be re-written as the following maximization problem:

$$MP(s, t) = \underset{P \in Q}{\operatorname{argmax}} \prod_{e \in P} \Pr(e)$$

$$= \underset{P \in Q}{\operatorname{argmin}} \prod_{e \in P} \frac{1}{\Pr(e)}$$

$$= \underset{P \in Q}{\operatorname{argmin}} \sum_{e \in P} \log_2 \frac{1}{\Pr(e)}$$

$$Connectivity(s, t) = \min_{P \in Q} \sum_{e \in P} \log_2 \frac{1}{\Pr(e)} \quad (2)$$

Note that $\Pr(e) > 0$, as all the reported interaction scores are strictly positive. This minimization problem denotes the “shortest path problem”, which we solved using Dijkstra’s algorithm (with a quadratic time complexity in the number of vertices).

The above optimization strategy relies heavily on the strong edges (interactions with high scores). The preference of high score edges may lead to over-sized paths, composed of only high score edges. To avoid oversized paths, we penalized/constrained the length of the path (# of edges in the path) while minimizing the connectivity score (a smaller connectivity score represents stronger connectivity). Here, we revised the aforementioned optimization problem into a 2-pass strategy. In the first pass, we find all the shortest paths $X(s, t)$ (with the same path length) that connect SARS-CoV-2 proteins to the target proteins of interest, without considering the edge

weights (interaction scores) in the graph. In the second pass, we find the path $x(s,t)$ in $X(s,t)$ that produces the minimal connectivity score $\text{Connectivity}_{\text{constrained}}(s,t)$ by taking the weight scores into considerations for only the selected candidate paths from the first run.

$$\begin{aligned} X(s,t) &= \operatorname{argmin}_{P \in Q} \sum_{e \in P} 1 \\ \text{Connectivity}_{\text{constrained}}(s,t) &= \min_{P \in X(s,t)} \sum_{e \in P} \log_2 \frac{1}{\Pr(e)} \\ x(s,t) &= \operatorname{argmin}_{P \in X(s,t)} \sum_{e \in P} \log_2 \frac{1}{\Pr(e)} \quad (3) \end{aligned}$$

We have packaged all the code into a tool named GeneList2COVID19, which is freely available for academic uses.

Ranking diseases using foRWaRD

We developed foRWaRD to rank a set of diseases (with known associated genes) based on their relevance to a set of genes (here HLH genes). This method works on the principles of Random Walk with Restarts (RWRs) for ranking genes and gene sets on heterogeneous networks^{33,79,82}, and enables integration of gene-level interactions to rank a set of diseases, with known associated genes, based on their relevance.

foRWaRD, requires three types of inputs (**Sup. Fig. 1**): 1) a set of diseases along with genes associated with each disease and the score of gene-disease associations (optional), 2) a gene interaction network (e.g. co-expression, protein-protein interaction, etc.), and 3) a set of query genes. Using these inputs, foRWaRD first generates a heterogeneous network comprising of gene-gene edges and disease-gene edges, with normalized edge weights representing the strength of the gene-gene interaction and the strength of evidence for gene-disease interaction (e.g. from the DisGeNET database). Then, the query set is superimposed on this network and is used as the restart set in an RWR algorithm. Using RWR in this algorithm allows us to capture topological information within the network both locally (the neighborhood surrounding the query set) and globally. After the convergence of the RWR, the steady-state probabilities of the disease nodes represent their relevance to the query set. In order to correct for the network bias (i.e. to avoid diseases with a large number of associated genes be ranked highly independent of their relevance to the query set), we run the RWR one more time with all the genes in the network as the restart set, providing a background steady-state probability for each disease node. The difference between the steady state probabilities of these two RWRs are then normalized between 0 and 1. More specifically, let d_i represents the difference between the steady state probabilities of the two RWRs for disease i , where $i = 1, 2, \dots, n$ and n is the total number of diseases to be ranked (note that $-1 \leq d_i \leq 1$). Also, let $d_{\max} = \max_i (|d_i|)$. The normalized disease score (NDS) for the i -th disease is:

$$D_i = \frac{d_i}{2d_{\max}} + 0.5.$$

It is important to note that NDS above 0.5 reflect diseases whose similarity score with respect to the query set is larger than their similarity score with respect to all genes (i.e. background).

The RWR (which we used in foRWaRD) is an algorithm for scoring the similarity between any given node of a weighted network and a query set of nodes. Starting from some initial node,

at each step the random walker moves to an adjacent node with a probability proportional to the edge weight connecting the two nodes and with some probability (known as probability of restart) it jumps to one of the nodes in the query set (also known as the restart set). The restart probability p controls the influence of the local topology of the network (surrounding the query set) and its global topology. We used $p = 0.5$ to balance the influence of these two factors.

Software Availability

The software GeneList2COVID19 is written in Python, available as an open source tool at GitHub (<https://github.com/phoenixding/genelist2covid19>). An implementation of the software foRWaRD is available in Python and is freely available on GitHub (<https://github.com/ddhostallero/foRWaRD>). These GitHub repositories include the source code as well as detailed instructions on how to install and use the methods.

Acknowledgements

This work was partly supported by the NSERC Discovery grant RGPIN-2019-04460 (AE), and a grant from McGill Initiative in Computational Medicine (MiCM) and McGill Interdisciplinary Initiative in Infection and Immunity (MI4) (AE and SR). SR is supported by a Chercheur-Boursier Senior salary award from the Fonds de Recherche du Québec – Santé. KT is supported by a Chercheur-Junior 1 salary award from the Fonds de Recherche du Québec – Santé.

Authors' contributions

JD, GF, AE and SR contributed to the conception or design of the work; JD, DEH, MREK, GF, NN, MGB, AE and SR contributed to the acquisition, analysis, or interpretation of data; JD, DEH, MREK, and AE contributed to the creation of new software used in the work; JD, GF, SM, JS, ND, MS, KT, AE and SR have drafted the work or substantively revised it. All the authors have approved the submitted version have agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Tables

Table 1. GO terms associated with the HLH gene set, identified by KnowEnG's network-guided gene set characterization pipeline using HumanNet Integrated network

GO ID ^a	GO Term Description ^a	Difference Score ^b
GO:0043312	Neutrophil degranulation	1
GO:0006886	Intracellular protein transport	0.689
GO:0070382	Exocytic vesicle	0.675
GO:0045921	Positive regulation of exocytosis	0.672
GO:1903307	Positive regulation of regulated secretory pathway	0.660
GO:0043304	Regulation of mast cell degranulation	0.645
GO:0032438	Melanosome organization	0.634
GO:0006968	Cellular defense response	0.625
GO:0031201	SNARE complex	0.616
GO:0005764	Lysosome	0.613
GO:0051607	Defense response to virus	0.611
GO:0042127	Regulation of cell proliferation	0.610
GO:0033093	Weibel-Palade body	0.578

^a From KnowEnG³³ (www.knowng.org/analyze)

^b See methods section.

Table 2. Diseases associated with NETs and HLH genes, identified using forWaRD

Rank	Disease	Disease Score	NET ^a	Remarks
1	Alzheimer's Disease	1	+	Increased NET in the proximity of Amyloid beta deposit ⁸³ . Increase risk of thrombosis ⁸⁴ and stroke ⁸⁵ .
2	Liver carcinoma	0.599	+	NETs have been described in multiple malignancies ^{83,55} , including liver carcinoma ⁸⁶ , which is also associated with an increased risk of VTE ⁸⁷
3	Myocardial Reperfusion Injury	0.516	+	Neutrophils play a major role in reperfusion injury that result in microvascular damage ⁸⁸ . Reperfusion injury leads to the release of NETs in animal models ⁸⁹ .
4	Diabetes Mellitus	0.508	+	In COVID-19, well-controlled blood glucose was associated with lower mortality compared to individuals with poorly controlled blood glucose ⁶³ . Accordingly, proteomic analysis of peripheral blood polymorphonuclear cells (PBMCS) reveals alteration of NET components in uncontrolled diabetes ⁹⁰ . Interestingly, high glucose cooperates with IL-6 and homocysteine to form NETs in the context of Type 2 diabetes (T2D) ^{91,92} . T2D is a major COVID-19 co-morbidity ⁶³ .
5	Congenital Fiber Type Disproportion	0.506	(-)	None
6	Immunodeficiency 8	0.502	-	Results from the loss of function of coronin-1, leads to bleeding. Coronin-1 plays key functions in PMN trafficking in innate immunity ⁷⁴ .
7	Diabetes Mellitus, Insulin-Dependent	0.501	+	See entry above
8	Cataract 30	0.501	(-)	None
9	Hemolytic anemia, nonspherocytic, due	0.500	(-)	None

	to glucose phosphate isomerase deficiency			
10	Bleeding disorder, platelet-type, 15	0.500	(-)	None

Abbreviations used: NETs = Neutrophil Extracellular Traps.

^a A “+” sign indicates demonstrated NET release in the disease, a “-” sign a deficiency in NETs. Brackets “()” around the “+” or “-” signs indicate prediction without published data.

Figure Legends

Figure 1. Reconstructed network paths from SARS-CoV-2 proteins to HLH genes. This network shows all the paths connecting the SARS-CoV-2 proteins to the HLH proteins (genes). The red nodes represent the SARS-CoV-2 proteins, the yellow nodes are the human host proteins that directly interact with SARS-CoV-2 proteins, the green nodes are the intermediate interacting host proteins, and the blue nodes denote the target HLH proteins (genes). The edge weights in the network represent the interaction strength (or probability).

Figure 2. HLH genes are significantly enriched within the SARS-CoV-2 host protein interactome. A connectivity score was calculated for each of the genes of interest (e.g. HLH genes in this work). We further analyzed the network connectivity of all genes to the SARS-CoV2 proteins (or randomly picked genes). With these two analyses, we ended up with two lists of connectivity scores: *SA* (for HLH genes) and *SB* (for all background genes). Then, we calculated the statistical significance (p-value) using a one-sided Mann-Whitney rank test to determine whether *SA* is significantly smaller than *SB* (stronger connectivity). *SA* significant p-value implies that the list of proteins (genes) of interest is “significantly connected” to the SARS-CoV2 proteins. **A)** A list of 10 known COVID19 related genes (differential genes in severe COVID19 patients) have statistically stronger connections to the SARS-CoV-2 proteins compared with all background genes (p-value=0.017). **B)** A list of 100 random genes does not “significantly” connect to the SARS-CoV-2 proteins. **C)** The 23 Male infertility genes do not “significantly” connect to the SARS-CoV-2 proteins. **D)** The 11 HLH genes have statistically (p-value=0.00821) stronger connections to the SARS-CoV-2 proteins (compared with all background genes). **E)** A list of 45 vascular angiogenesis genes linked to both H1N1 and SARS-CoV-2 pulmonary infections significantly (p-value=4.89e-5) connect to the SARS-CoV-2 proteins. The 11 HLH genes have the smallest mean/median connectivity score compared to all the gene lists analyzed. Please note that the p-values here only indicate whether the input gene lists have significantly smaller connectivity scores than all the background genes, and they could be affected by the size of the gene list. To compare the strength of the “connectivity” of input gene lists to SARS-CoV-2 proteins, we should also look at the mean (represented by a green triangle) and the median (represented by a vertical line) connectivity scores.

Figure 3. Differential expression of HLH genes in COVID19 associated health conditions. Gene expression of lung cells under different health conditions that have been associated with COVID19^{10,28,29} was analyzed. The health conditions include Sex (Male Vs. Female healthy individuals), Smoking status (current smokers vs never smokers), Cancer (Lung cancer vs. Normal), COPD (COPD vs. Normal), Diabetes (Type 1 Diabetes vs. Normal), Hypertension (Severe PH: mPAP>40mmHg vs. without PH: mPAP<20mmHg), and Aged (>=60 years vs. <60 years healthy individuals). All the genes in these datasets were sorted under each condition, based on the log fold change (condition vs control). Genes whose fold change was among the top 25% were classified as high (H) and those whose fold change was among the bottom 25% were classified as low (L). Finally, the 11 HLH genes were assessed to determine whether they are among the H or L genes in each condition. The red blocks represent HLH genes that were highly expressed (H, top 25%) in condition (vs. control) while the blue blocks represent HLH genes that were lowly expressed (L, bottom 25%) in condition (vs. control). We compared the HLH genes

and all background genes in terms of H/L expression under various conditions. We first counted the number of H/L (differentially expressed between the condition and control) for each of the HLH genes, and then for each of the background genes. Next, we used a one-sided Mann-Whitney rank test to determine whether the HLH genes have larger absolute fold changes, (*i.e.*, are differentially expressed), in COVID19 associated conditions compared to all the background genes significantly ($p\text{-value} < 0.05$). The average number of H/L conditions for HLH genes (red or blue blocks) is 3.82, which is significantly larger (one-sided rank-sum test $p\text{-value} = 1.01 \times 10^{-4}$) than the average number of H/L conditions for all genes (1.70).

Figure 4. Expression of HLH genes in Control, inactive sJIA and active sJIA Neutrophils.

Gene expression of HLH-SARS-CoV-2 and positive COVID-19 genes (**Supplementary Table 1**) in sJIA was calculated from GEO series GSE122552⁴⁷. Data was mapped to the hg38 genome and normalized by reads per kilobase per million (RPKM). Values for HLH genes were displayed for control and sJIA patients that were either in remission (inactive sJIA) or had active symptoms (active sJIA).

Figure 5. Model of NET-mediated endothelial damage contributing to pulmonary vascular thrombosis in severe COVID-19.

Infection by SARS-CoV-2 in vulnerable population will lead to hyperinflammation either from underlying genetic mutations, specific epigenetic landscapes or external factors, that will result in the increase circulation of acute phase reactants such as CRP and pro-inflammatory cytokines associated with neutrophilia like IL-6, IL-17A/F and CXCL8 (IL-8). IL-17A activates the endothelium to induce neutrophil adhesion⁹³, where the increase in CRP can trigger the release of NETs, resulting in damage to the endothelium as well as aggregation and activation of platelets. Additionally, the presence of SARS-CoV-2 E protein in type II pneumocytes could disturb the surfactant cargo via its interaction with AP3B1, leading to impaired secretion of SP-D and greater NET formation by septal and intra-alveolar neutrophils increasing the risk of thrombosis in the pulmonary microvasculature. In some predisposed patients the combinations of these mechanisms will lead to severe COVID-19 complications. The identification of mediators of this pro-coagulation cascade is essential in achieving the two-fold task of identifying vulnerable populations and developing a personalized medicine approach.

References

1. Tang, N., Li, D., Wang, X. & Sun, Z. Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia. *J. Thromb. Haemost.* **18**, 844–847 (2020).
2. Zhou, F. *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet* **395**, 1054–1062 (2020).
3. Llitjos, J.-F. *et al.* High incidence of venous thromboembolic events in anticoagulated severe COVID-19 patients. *J. Thromb. Haemost.* **n/a**,.
4. Helms, J. *et al.* High risk of thrombosis in patients with severe SARS-CoV-2 infection: a multicenter prospective cohort study. *Intensive Care Med.* 1–10 (2020) doi:10.1007/s00134-020-06062-x.
5. Magro, C. *et al.* Complement associated microvascular injury and thrombosis in the pathogenesis of severe COVID-19 infection: a report of five cases. *Transl. Res.* S1931524420300700 (2020) doi:10.1016/j.trsl.2020.04.007.
6. Oxley, T. J. *et al.* Large-Vessel Stroke as a Presenting Feature of Covid-19 in the Young. *N. Engl. J. Med.* **382**, e60 (2020).
7. Blanco-Melo, D. *et al.* Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* **181**, 1036-1045.e9 (2020).
8. Mehta, P. *et al.* COVID-19: consider cytokine storm syndromes and immunosuppression. *The Lancet* **395**, 1033–1034 (2020).
9. Zuo, Y. *et al.* Neutrophil extracellular traps in COVID-19. *JCI Insight* **5**, (2020).
10. Zhou, F. *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet* **395**, 1054–1062 (2020).

11. Ruan, Q., Yang, K., Wang, W., Jiang, L. & Song, J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive Care Med.* **46**, 846–848 (2020).
12. England, J. T. *et al.* Weathering the COVID-19 storm: Lessons from hematologic cytokine syndromes. *Blood Rev.* 100707 (2020) doi:10.1016/j.blre.2020.100707.
13. Zhang, X.-Y. *et al.* Hemophagocytic Lymphohistiocytosis Induced by Severe Pandemic Influenza A (H1N1) 2009 Virus Infection: A Case Report. *Case Rep. Med.* **2011**, 951910 (2011).
14. Yöntem, Y. *et al.* Analysis of fatal cases of pandemic influenza A (H1N1) virus infections in pediatric patients with leukemia. *Pediatr. Hematol. Oncol.* **30**, 437–444 (2013).
15. Ozdemir, H. *et al.* Hemophagocytic lymphohistiocytosis associated with 2009 pandemic influenza A (H1N1) virus infection. *J. Pediatr. Hematol. Oncol.* **33**, 135–137 (2011).
16. Shrestha, B., Omran, A., Rong, P. & Wang, W. Report of a Fatal Pediatric Case of Hemophagocytic Lymphohistiocytosis Associated with Pandemic Influenza A (H1N1) Infection in 2009. *Pediatr. Neonatol.* **56**, 189–192 (2015).
17. Ackermann, M. *et al.* Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. *N. Engl. J. Med.* (2020) doi:10.1056/NEJMoa2015432.
18. Ramos-Casals, M., Brito-Zerón, P., López-Guillermo, A., Khamashta, M. A. & Bosch, X. Adult haemophagocytic syndrome. *The Lancet* **383**, 1503–1516 (2014).
19. Verdoni, L. *et al.* An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. *Lancet Lond. Engl.* (2020) doi:10.1016/S0140-6736(20)31103-X.

20. Toubiana, J. *et al.* Outbreak of Kawasaki disease in children during COVID-19 pandemic: a prospective observational study in Paris, France. *medRxiv* 2020.05.10.20097394 (2020) doi:10.1101/2020.05.10.20097394.
21. Viner, R. M. & Whittaker, E. Kawasaki-like disease: emerging complication during the COVID-19 pandemic. *Lancet Lond. Engl.* (2020) doi:10.1016/S0140-6736(20)31129-6.
22. Weaver, L. K. & Behrens, E. M. Hyperinflammation, rather than hemophagocytosis, is the common link between macrophage activation syndrome and hemophagocytic lymphohistiocytosis. *Curr. Opin. Rheumatol.* **26**, 562–569 (2014).
23. Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* 1–13 (2020) doi:10.1038/s41586-020-2286-9.
24. Al-Samkari, H. & Berliner, N. Hemophagocytic Lymphohistiocytosis. *Annu. Rev. Pathol. Mech. Dis.* **13**, 27–49 (2018).
25. Brisse, E., Wouters, C. H. & Matthys, P. Advances in the pathogenesis of primary and secondary haemophagocytic lymphohistiocytosis: differences and similarities. *Br. J. Haematol.* **174**, 203–217 (2016).
26. Bost, P. *et al.* Host-Viral Infection Maps Reveal Signatures of Severe COVID-19 Patients. *Cell* S0092867420305687 (2020) doi:10.1016/j.cell.2020.05.006.
27. Wacławska, A. & Kurpisz, M. Key functional genes of spermatogenesis identified by microarray analysis. *Syst. Biol. Reprod. Med.* **58**, 229–235 (2012).
28. Zhao, X. *et al.* Incidence, clinical characteristics and prognostic factor of patients with COVID-19: a systematic review and meta-analysis. *medRxiv* 2020.03.17.20037572 (2020) doi:10.1101/2020.03.17.20037572.

29. Ding, J., Lugo-Martinez, J., Yuan, Y., Kotton, D. N. & Bar-Joseph, Z. Reconstructing SARS-CoV-2 response signaling and regulatory networks. *bioRxiv* 2020.06.01.127589 (2020) doi:10.1101/2020.06.01.127589.
30. Catz, S. D. The role of Rab27a in the regulation of neutrophil function. *Cell. Microbiol.* **16**, 1301–1310 (2014).
31. Johnson, J. L. *et al.* Rab27a and Rab27b regulate neutrophil azurophilic granule exocytosis and NADPH oxidase activity by independent mechanisms. *Traffic Cph. Den.* **11**, 533–547 (2010).
32. Peden, A. A. *et al.* Localization of the AP-3 adaptor complex defines a novel endosomal exit site for lysosomal membrane proteins. *J. Cell Biol.* **164**, 1065–1076 (2004).
33. Blatti, C. *et al.* Knowledge-guided analysis of ‘omics’ data using the KnowEnG cloud platform. *PLoS Biol.* **18**, e3000583 (2020).
34. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
35. Hwang, S. *et al.* HumanNet v2: human gene networks for disease research. *Nucleic Acids Res.* **47**, D573–D580 (2019).
36. Albeituni, S. *et al.* Mechanisms of action of ruxolitinib in murine models of hemophagocytic lymphohistiocytosis. *Blood* **134**, 147–159 (2019).
37. Papayannopoulos, V. Neutrophil extracellular traps in immunity and disease. *Nat. Rev. Immunol.* **18**, 134–147 (2018).
38. Jones, V. G. *et al.* COVID-19 and Kawasaki Disease: Novel Virus and Novel Case. *Hosp. Pediatr.* **10**, 537–540 (2020).

39. Yoshida, Y. *et al.* Enhanced formation of neutrophil extracellular traps in Kawasaki disease. *Pediatr. Res.* **87**, 998–1004 (2020).
40. Elaskalani, O., Abdol Razak, N. B. & Metharom, P. Neutrophil extracellular traps induce aggregation of washed human platelets independently of extracellular DNA and histones. *Cell Commun. Signal. CCS* **16**, 24 (2018).
41. Jorch, S. K. & Kubes, P. An emerging role for neutrophil extracellular traps in noninfectious disease. *Nat. Med.* **23**, 279–287 (2017).
42. Krishnamoorthy, N. *et al.* Neutrophil cytoplasts induce TH17 differentiation and skew inflammation toward neutrophilia in severe asthma. *Sci. Immunol.* **3**, (2018).
43. Sproston, N. R. & Ashworth, J. J. Role of C-Reactive Protein at Sites of Inflammation and Infection. *Front. Immunol.* **9**, (2018).
44. Oguz, M. M. *et al.* Secondary hemophagocytic lymphohistiocytosis in pediatric patients: a single center experience and factors that influenced patient prognosis. *Pediatr. Hematol. Oncol.* **36**, 1–16 (2019).
45. Huttenlocher, A. & Smith, J. A. Neutrophils in pediatric autoimmune disease. *Curr. Opin. Rheumatol.* **27**, 500–504 (2015).
46. Hu, X., Xie, Q., Mo, X. & Jin, Y. The role of extracellular histones in systemic-onset juvenile idiopathic arthritis. *Ital. J. Pediatr.* **45**, 14 (2019).
47. Brown, R. A. *et al.* Neutrophils From Children With Systemic Juvenile Idiopathic Arthritis Exhibit Persistent Proinflammatory Activation Despite Long-Standing Clinically Inactive Disease. *Front. Immunol.* **9**, 2995 (2018).
48. R&D Blueprint and COVID-19. <https://www.who.int/teams/blueprint/covid-19>.

49. O'Donoghue, A. J. *et al.* Global substrate profiling of proteases in human neutrophil extracellular traps reveals consensus motif predominantly contributed by elastase. *PloS One* **8**, e75141 (2013).
50. Urban, C. F. *et al.* Neutrophil extracellular traps contain calprotectin, a cytosolic protein complex involved in host defense against *Candida albicans*. *PLoS Pathog.* **5**, e1000639 (2009).
51. Piñero, J. *et al.* The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* **48**, D845–D855 (2020).
52. Horne, B. D. *et al.* Which White Blood Cell Subtypes Predict Increased Cardiovascular Risk? *J. Am. Coll. Cardiol.* **45**, 1638–1643 (2005).
53. Vulesevic, B. *et al.* CRP Induces NETosis in Heart Failure Patients with or without Diabetes. *ImmunoHorizons* **3**, 378–388 (2019).
54. Barnes, B. J. *et al.* Targeting potential drivers of COVID-19: Neutrophil extracellular traps. *J. Exp. Med.* **217**, (2020).
55. Rayes, R. F. *et al.* Primary tumors induce neutrophil extracellular traps with targetable metastasis promoting effects. *JCI Insight* **5**, (2019).
56. Cools-Lartigue, J. *et al.* Neutrophil extracellular traps sequester circulating tumor cells and promote metastasis. *J. Clin. Invest.* (2013) doi:10.1172/JCI67484.
57. Khan, M. A., Ali, Z. S., Swezey, N., Grasemann, H. & Palaniyar, N. Progression of Cystic Fibrosis Lung Disease from Childhood to Adulthood: Neutrophils, Neutrophil Extracellular Trap (NET) Formation, and NET Degradation. *Genes* **10**, (2019).

769 58. Hoppenbrouwers, T. *et al.* Neutrophil Extracellular Traps in Children With Meningococcal
770 Sepsis. *Pediatr. Crit. Care Med. J. Soc. Crit. Care Med. World Fed. Pediatr. Intensive Crit.*
771 *Care Soc.* **19**, e286–e291 (2018).

772 59. Appelgren, D. *et al.* Neutrophil Extracellular Traps (NETs) in the Cerebrospinal Fluid
773 Samples from Children and Adults with Central Nervous System Infections. *Cells* **9**, (2019).

774 60. Duvvuri, B. *et al.* Neutrophil Extracellular Traps in Tissue and Periphery in Juvenile
775 Dermatomyositis. *Arthritis Rheumatol. Hoboken NJ* **72**, 348–358 (2020).

776 61. Gottlieb, Y. *et al.* Neutrophil extracellular traps in pediatric inflammatory bowel disease.
777 *Pathol. Int.* **68**, 517–523 (2018).

778 62. Colón, D. F. *et al.* Neutrophil extracellular traps (NETs) exacerbate severity of infant sepsis.
779 *Crit. Care Lond. Engl.* **23**, 113 (2019).

780 63. Zhu, L. *et al.* Association of Blood Glucose Control and Outcomes in Patients with COVID-
781 19 and Pre-existing Type 2 Diabetes. *Cell Metab.* **31**, 1068-1077.e3 (2020).

782 64. Wang, H. *et al.* Dementia care during COVID-19. *Lancet Lond. Engl.* **395**, 1190–1191
783 (2020).

784 65. McMichael, T. M. *et al.* Epidemiology of Covid-19 in a Long-Term Care Facility in King
785 County, Washington. *N. Engl. J. Med.* **382**, 2005–2011 (2020).

786 66. Alunno, A., Carubbi, F. & Rodríguez-Carrio, J. Storm, typhoon, cyclone or hurricane in
787 patients with COVID-19? Beware of the same storm that has a different origin. *RMD Open*
788 **6**, (2020).

789 67. Jung, J. *et al.* Identification of a homozygous deletion in the AP3B1 gene causing
790 Hermansky-Pudlak syndrome, type 2. *Blood* **108**, 362–369 (2006).

68. Dell'Angelica, E. C., Shotelersuk, V., Aguilar, R. C., Gahl, W. A. & Bonifacino, J. S. Altered Trafficking of Lysosomal Proteins in Hermansky-Pudlak Syndrome Due to Mutations in the β 3A Subunit of the AP-3 Adaptor. *Mol. Cell* **3**, 11–21 (1999).
69. Nieto-Torres, J. L. *et al.* Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein. *Virology* **415**, 69–82 (2011).
70. Schoeman, D. & Fielding, B. C. Coronavirus envelope protein: current knowledge. *Virol. J.* **16**, 69 (2019).
71. Ziegler, C. *et al.* SARS-CoV-2 Receptor ACE2 is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Enriched in Specific Cell Subsets Across Tissues. <https://papers.ssrn.com/abstract=3555145> (2020) doi:10.2139/ssrn.3555145.
72. Young, L. R. *et al.* The alveolar epithelium determines susceptibility to lung fibrosis in Hermansky-Pudlak syndrome. *Am. J. Respir. Crit. Care Med.* **186**, 1014–1024 (2012).
73. Yao, X.-H. *et al.* Pathological evidence for residual SARS-CoV-2 in pulmonary tissues of a ready-for-discharge patient. *Cell Res.* **30**, 541–543 (2020).
74. Pick, R. *et al.* Coronin 1A, a novel player in integrin biology, controls neutrophil trafficking in innate immunity. *Blood* **130**, 847–858 (2017).
75. Raftery, M. J. *et al.* β 2 integrin mediates hantavirus-induced release of neutrophil extracellular traps. *J. Exp. Med.* **211**, 1485–1497 (2014).
76. Gralinski, L. E. *et al.* Complement Activation Contributes to Severe Acute Respiratory Syndrome Coronavirus Pathogenesis. *mBio* **9**, (2018).
77. Zhu, L. *et al.* High Level of Neutrophil Extracellular Traps Correlates With Poor Prognosis of Severe Influenza A Infection. *J. Infect. Dis.* **217**, 428–437 (2018).

78. Schaefer, M. H. *et al.* HIPPIE: Integrating Protein Interaction Networks with Experiment Based Quality Scores. *PLOS ONE* **7**, e31826 (2012).
79. Blatti, C. & Sinha, S. Characterizing gene sets using discriminative random walks with restart on heterogeneous biological networks. *Bioinforma. Oxf. Engl.* **32**, 2167–2175 (2016).
80. Varusai, T. M., Kolch, W., Kholodenko, B. N. & Nguyen, L. K. Protein-protein interactions generate hidden feedback and feed-forward loops to trigger bistable switches, oscillations and biphasic dose-responses. *Mol. Biosyst.* **11**, 2750–2762 (2015).
81. Gavenonis, J., Sheneman, B. A., Siegert, T. R., Eshelman, M. R. & Kritzer, J. A. Comprehensive Analysis of Loops at Protein-Protein Interfaces for Macrocycle Design. *Nat. Chem. Biol.* **10**, 716–722 (2014).
82. Emad, A., Cairns, J., Kalari, K. R., Wang, L. & Sinha, S. Knowledge-guided gene prioritization reveals new insights into the mechanisms of chemoresistance. *Genome Biol.* **18**, 153 (2017).
83. Zenaro, E. *et al.* Neutrophils promote Alzheimer’s disease-like pathology and cognitive decline via LFA-1 integrin. *Nat. Med.* **21**, 880–886 (2015).
84. Schmaier, A. H. Alzheimer disease is in part a thrombohemorrhagic disorder. *J. Thromb. Haemost.* **14**, 991–994 (2016).
85. Honig, L. S. *et al.* Stroke and the Risk of Alzheimer Disease. *Arch. Neurol.* **60**, 1707–1712 (2003).
86. van der Windt, D. J. *et al.* Neutrophil extracellular traps promote inflammation and development of hepatocellular carcinoma in nonalcoholic steatohepatitis. *Hepatol. Baltim. Md* **68**, 1347–1360 (2018).

87. Zanetto, A. *et al.* Cancer-Associated Thrombosis in Cirrhotic Patients with Hepatocellular Carcinoma. *Cancers* **10**, (2018).
88. Quintana, M., Kahan, T. & Hjemdahl, P. Pharmacological prevention of reperfusion injury in acute myocardial infarction. A potential role for adenosine as a therapeutic agent. *Am. J. Cardiovasc. Drugs Drugs Devices Interv.* **4**, 159–167 (2004).
89. Ge, L. *et al.* Neutrophil extracellular traps in ischemia-reperfusion injury-induced myocardial no-reflow: therapeutic potential of DNase-based reperfusion strategy. *Am. J. Physiol. Heart Circ. Physiol.* **308**, H500-509 (2015).
90. Soongsathitanon, J., Umsa-Ard, W. & Thongboonkerd, V. Proteomic analysis of peripheral blood polymorphonuclear cells (PBMCS) reveals alteration of neutrophil extracellular trap (NET) components in uncontrolled diabetes. *Mol. Cell. Biochem.* **461**, 1–14 (2019).
91. Joshi, M. B. *et al.* Elevated homocysteine levels in type 2 diabetes induce constitutive neutrophil extracellular traps. *Sci. Rep.* **6**, 36362 (2016).
92. Menegazzo, L. *et al.* NETosis is induced by high glucose and associated with type 2 diabetes. *Acta Diabetol.* **52**, 497–503 (2015).
93. Roussel, L. *et al.* IL-17 promotes p38 MAPK-dependent endothelial activation enhancing neutrophil recruitment to sites of inflammation. *J. Immunol. Baltim. Md 1950* **184**, 4531–7 (2010).

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5