

ARTICLE

Open Access

# Looking for pathways related to COVID-19: confirmation of pathogenic mechanisms by SARS-CoV-2–host interactome

Francesco Messina<sup>1</sup>, Emanuela Giombini<sup>1</sup>, Chiara Montaldo<sup>1</sup>, Ashish Arunkumar Sharma<sup>2</sup>, Antonio Zoccoli<sup>3</sup>, Rafick-Pierre Sekaly<sup>2</sup>, Franco Locatelli<sup>4</sup>, Alimuddin Zumla<sup>5</sup>, Markus Maeurer<sup>6,7</sup>, Maria R. Capobianchi<sup>1</sup>, Francesco Nicola Lauria<sup>1</sup> and Giuseppe Ippolito<sup>1</sup>

## Abstract

In the last months, many studies have clearly described several mechanisms of SARS-CoV-2 infection at cell and tissue level, but the mechanisms of interaction between host and SARS-CoV-2, determining the grade of COVID-19 severity, are still unknown. We provide a network analysis on protein–protein interactions (PPI) between viral and host proteins to better identify host biological responses, induced by both whole proteome of SARS-CoV-2 and specific viral proteins. A host-virus interactome was inferred, applying an explorative algorithm (Random Walk with Restart, RWR) triggered by 28 proteins of SARS-CoV-2. The analysis of PPI allowed to estimate the distribution of SARS-CoV-2 proteins in the host cell. Interactome built around one single viral protein allowed to define a different response, underlining as ORF8 and ORF3a modulated cardiovascular diseases and pro-inflammatory pathways, respectively. Finally, the network-based approach highlighted a possible direct action of ORF3a and NS7b to enhancing Bradykinin Storm. This network-based representation of SARS-CoV-2 infection could be a framework for pathogenic evaluation of specific clinical outcomes. We identified possible host responses induced by specific proteins of SARS-CoV-2, underlining the important role of specific viral accessory proteins in pathogenic phenotypes of severe COVID-19 patients.

## Introduction

Whilst COVID-19 predominantly affects the respiratory system, it is a multisystem disease, with a wide spectrum of clinical presentations from asymptomatic, mild and moderate, to severe, fulminant disease<sup>1</sup>. Host conditions and comorbidities (age, obesity, diabetes, hypertension, organ damages, inflammation and coagulation dysfunctionality), represented risk factors for severe and fatal disease courses<sup>2</sup>, but the mechanisms of host-SARS-CoV-2 interaction,

activating pathological pathways and influencing severity, are still unknown.

Recently, several studies described many mechanisms of SARS-CoV-2 infection at cell and tissue level. It was observed that the replication of SARS-CoV-2, as well as all +RNA viruses, occurs in the cytoplasm of the host cell, inducing a membrane rearrangement of rough endoplasmic reticulum (ER) membranes into double-membrane vesicles<sup>3,4</sup>. NSP8, NSP7 and NSP12, yield the RNA polymerase activity of NSP8, and are assembled into the replicase–transcriptase complex, generating anti-sense (–) RNAs, templates for positive-sense genome (+) and mRNA transcripts<sup>5,6</sup>. During virus entry, a well-known process is the cleavage of S protein by FURIN on the cell membrane, which lead to the split S protein into two subunits, S1 and S2, which the last can interact with ACE2 (refs. <sup>7,8</sup>). However, the SARS-CoV-2–host

Correspondence: Giuseppe Ippolito ([giuseppe.ippolito@inmi.it](mailto:giuseppe.ippolito@inmi.it))

<sup>1</sup>National Institute for Infectious Diseases “Lazzaro Spallanzani” IRCCS, Rome, Italy

<sup>2</sup>Department of Pathology, Emory University School of Medicine, Atlanta, GA, USA

Full list of author information is available at the end of the article

These authors contributed equally: Francesco Nicola Lauria, Giuseppe Ippolito.

Edited by M. Piacentini

© The Author(s) 2021



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

interaction is not restricted to local infection, but it triggers a systemic reaction, including the activation of the Bradykinin Storm, as described in many severe COVID-19 patients<sup>9</sup>. Indeed, SARS-CoV-2 infection causes from one side a decrease of ACE level in the lung cells and, on the other side, an increase of ACE2 level, leading to increase Bradykinin (BK) level<sup>9</sup>. BK is produced from an inactive pre-protein Kininogen-1 (KNG1) through the activation by the serine protease kallikrein<sup>10</sup>. Furthermore, the excess of BK can lead to vasodilatation, hypotension and hypokalaemia<sup>9,11</sup>, which is associated with arrhythmia<sup>12,13</sup>. All these clinical conditions have been widely reported in COVID-19 patients<sup>14–16</sup>. Furthermore, microvascular injuries, due to systemic inflammatory response and endothelial dysfunction, were frequently found in severe COVID-19 patients<sup>17</sup>. Moreover, myocardial injuries were found to be linked to the risk of fatal outcome in COVID-19 patients<sup>15,18,19</sup>. 43% of the patients with severe COVID-19 (21% of all COVID-19 patients) presented myocardial injury, showing an increased risk of myocardial injury of 4.74-fold in severe compared with non-severe patients<sup>20</sup>. Although microvascular injuries and microthrombi formation frequently occurred in severe COVID-19 patients, the role of SARS-CoV-2 in this phenotype is not completely explained yet.

Although many aspects of COVID-19 pathogenesis and mechanisms of SARS-CoV-2 have been investigated, only few papers describe the interactions among SARS-CoV-2 proteins by wet experiments. The influence of SARS-CoV-2 on transcriptome, proteome, ubiquitinome and phosphoproteome of a lung-derived human cell line, was described through a multi-omics approach.

Virus–host interactome by the computational approach has been applied to COVID-19 for drug repurposing<sup>21</sup>, allowing the identification of new drug targets<sup>22</sup> and contributing to explain clinical manifestations<sup>23,24</sup>. The structural information on SARS-CoV-2 proteins and their interactions with human proteins and other viral proteins, allowed to better understand the mechanisms of SARS-CoV-2 infection, also comparing it with SARS-CoV<sup>25</sup>.

One of the most interesting models of SARS-CoV-2–host interactions was carried out on Spike-receptor interactions in other Human Coronaviruses (H-CoV). This network-based analysis of H-CoV–host interaction has provided a theoretic model for H-CoV infections applicable to SARS-CoV-2 pathogenesis, revealing biologically and clinically relevant molecular targets of three H-CoV infections<sup>26</sup>.

The interactome based on PPI and gene expression data, have been applied also to uncover the molecular origins of phenotypes of other complex diseases<sup>27,28</sup>, in order to better define the mechanisms of COVID-19 pathogenesis.

The understanding of all mechanisms of SARS-CoV-2 infection also passes by overall visualization of biological reactions and pathways involved in COVID-19 and H-CoV infections<sup>29–31</sup>.

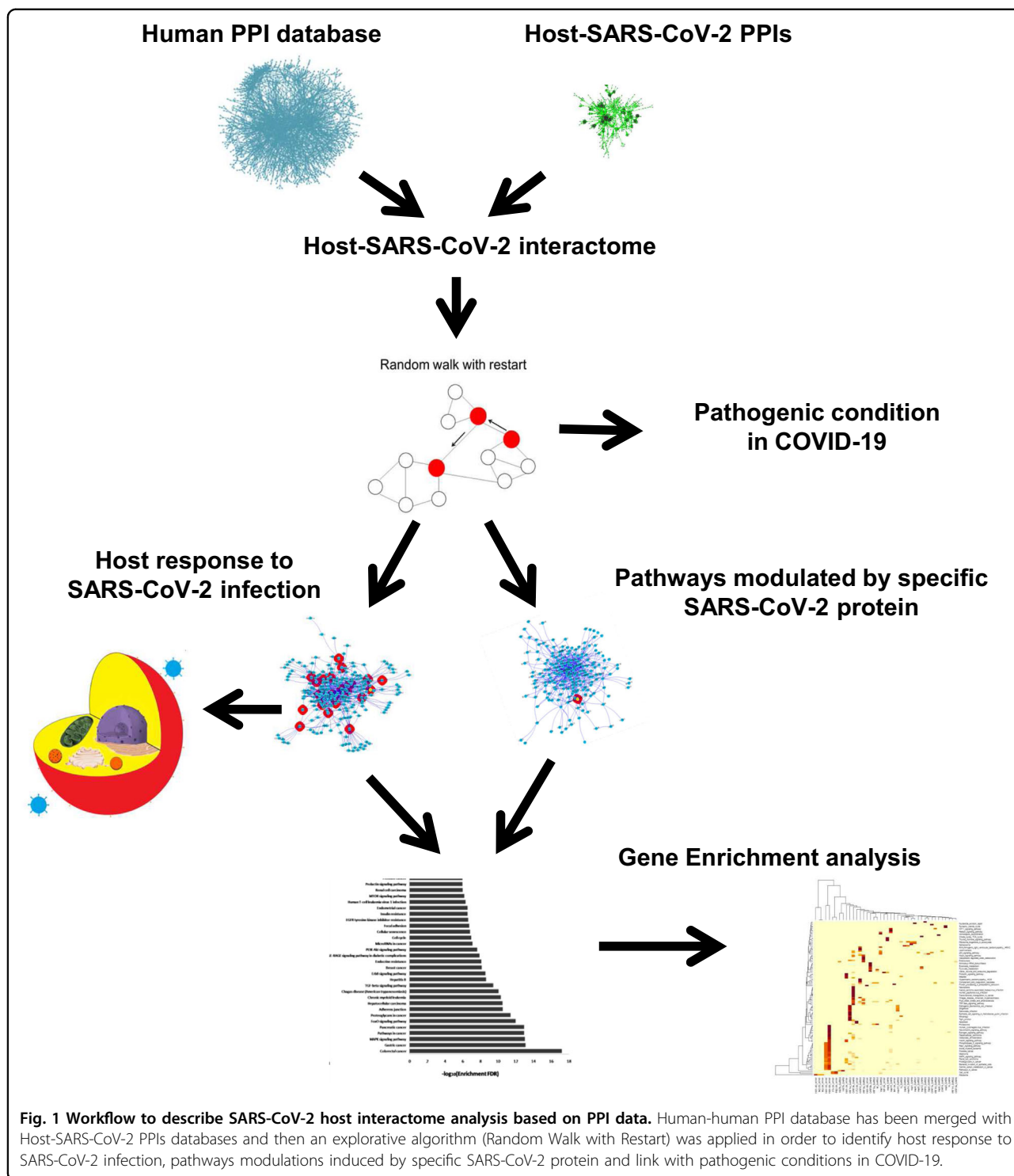
In this context, defining the host response induced by specific viral proteins would be of great importance and can guide the identification of functional viral targets, helping to better define the pathologic phenotypes of the infection.

Here, we carry out a network analysis on PPI to better identify host biological response induced by SARS-CoV-2. Furthermore, the interactome analysis was applied to design network of SARS-CoV-2–host proteins that could lead to a Bradykinin Storm. We used three different applications to identify interacting proteins: 200 proteins that interact closely with 28 SARS-CoV-2 proteins, 50 protein interactomes around each SARS-CoV-2 protein and 200 protein associated with KNG1, the pre-cursor of BK. The proteins found in all three approaches were analysed by gene enrichment and pathways associated with proteins were identified analysis (Fig. 1).

## Results

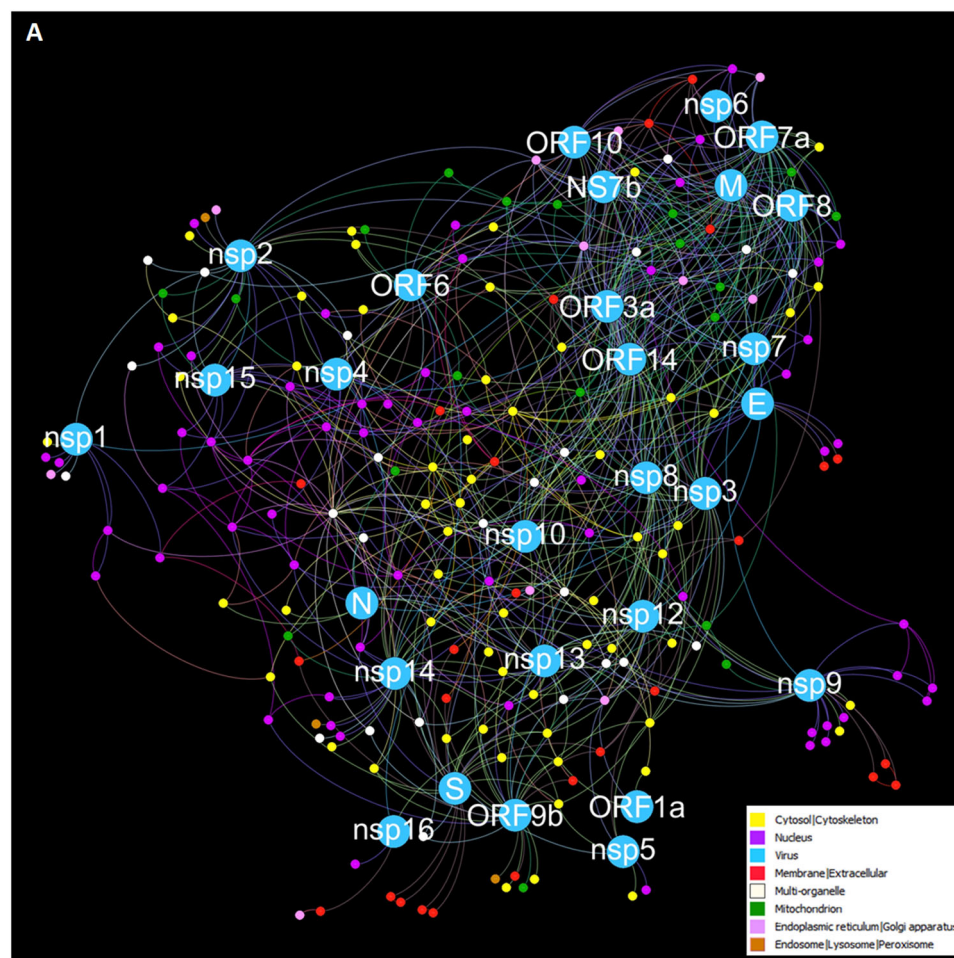
### Interactions with the whole proteome of SARS-CoV-2

To define the massive effect of SARS-CoV-2 infection on the host cell, an interactome between the entire set of SARS-CoV-2 proteins and host was carried out. Observing the colours' distribution, it is possible to distinguish three different areas, corresponding to as many sub-cellular districts, where SARS-CoV-2 proteins seem to be distributed (Fig. 2). In fact, N and some non-structural proteins (nsp1, nsp4 and nsp15) are posed around nuclear proteins, while S, along with nsp5, nsp10, nsp12, nsp13, nsp14, nsp16, ORF1a and ORF9b, were among cytosolic and membrane proteins, at the border of the interactome. Protein M and many accessory and not-structural proteins (NS7b, nsp6, nsp7, ORF3a, ORF7a, ORF8, ORF10, and ORF14) are around mitochondrial and endoplasmic proteins. At first, Gene Enrichment Analysis on Gene Ontology databases highlighted the processes DNA replication, synthesis of RNA primer (GO:0006269) and peptide antigen transport (GO:0046968) were the most enriched (66.33 fold enrichment and False Discovery Rate, FDR < 5%). To further dissect the interactions with the entire proteome of SARS-CoV-2 and to better understand which pathways could be involved, enrichment analysis was carried out on proteins reported in the interactome, using WikiPathways and Kyoto Encyclopaedia of Genes and Genomes (KEGG) databases. WikiPathways gene enrichment analysis revealed biological pathways of DNA replication, ubiquitination, and proteasome, with high significance (FDR < 0.01%). KEGG pathway enrichment analysis revealed DNA and RNA replication



pathways as the most significant pathways (FDR < 0.01%), as well as signalling pathways and viral infection pathways (Supplementary Fig. 1). We described the high values of betweenness centrality and degree for nine host proteins: two proteins multi-organelle amyloid-beta precursor protein (APP) and dual-specificity protein

phosphatase (PTEN); the membrane protein Sodium/potassium-transporting ATPase subunit alpha-1 (ATP1A1); two cytoplasmic proteins, 14-3-3 protein theta (YWHAQ) and Ubiquitin (UBC); two proteins of ER and Golgi apparatus, Sarcoplasmic/ER calcium ATPase 2 (ATP2A2) and Unconventional myosin-VI



**Fig. 2** PPI interactome, based on human PPI and SARS-CoV-2-host interactions, with top 200 closest proteins identified by RWR, using together 28 proteins of SARS-CoV-2 as seeds for only one RWR run. Different colours of node and edges represent different locations in the cell, which are reported in the legend and in Supplementary Table 1. This network was represented by Force Atlas algorithm.

(MYO6); the mitochondrial ATP synthase subunit alpha (ATP5F1A); one nuclear receptor for export of RNAs, Exportin-1 (XPO1), suggesting their main role in this infection (Supplementary Fig. 2, Supplementary Table 1).

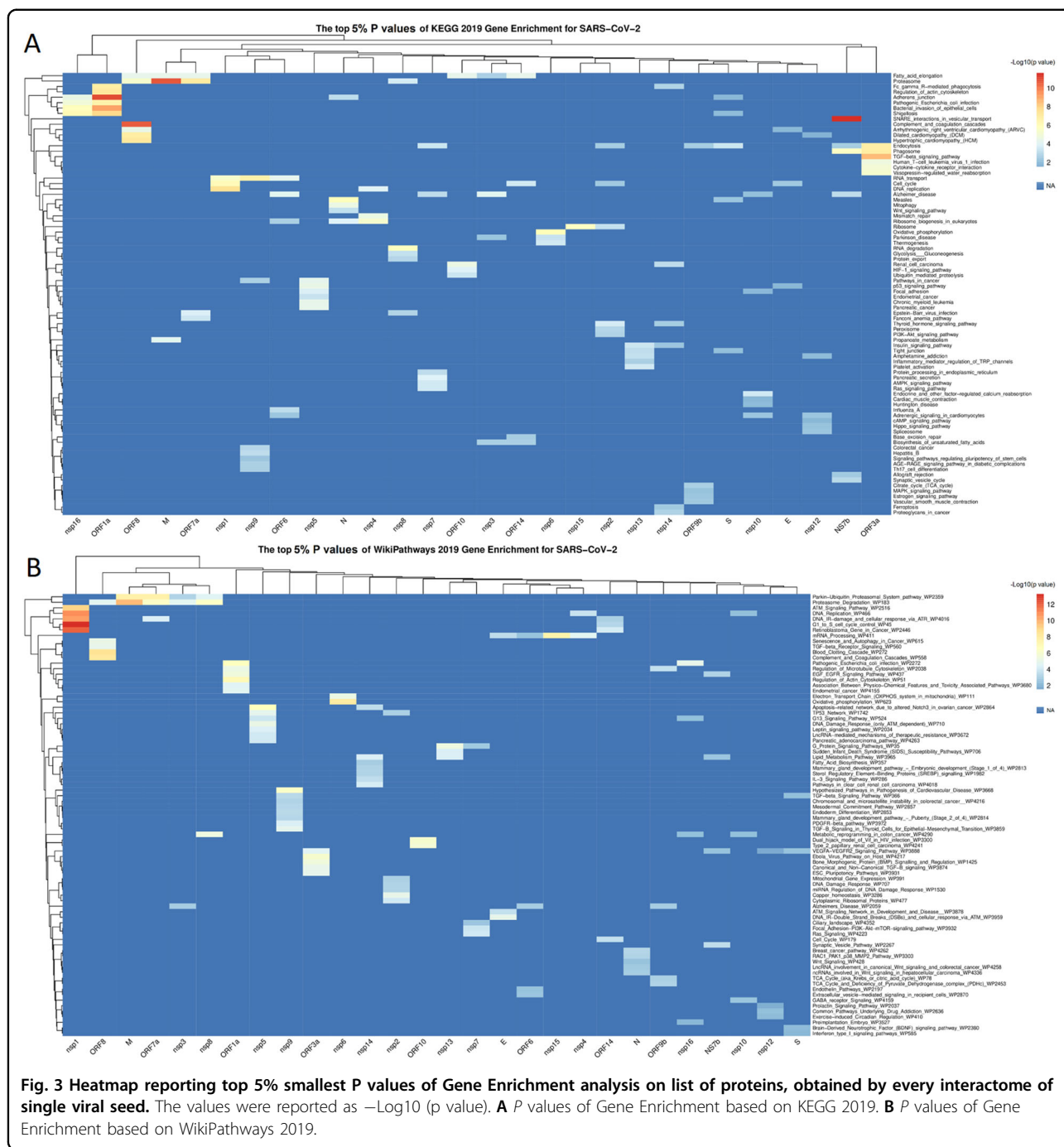
#### Host response induced by the specific protein of SARS-CoV-2

To define the effect of specific viral proteins on host response, interactomes were built, choosing one specific viral protein as seed, imposing to find the 50 closest proteins. These analyses allowed to produce 28 restricted interactomes, which defined the strictest biological interactions associated to seed proteins both among human proteins and other viral proteins. In Supplementary Fig. 3 the interactomes for structural proteins were reported, while in Supplementary Figs. 4 and 5 the interactomes for accessory and non-structural proteins were plotted.

Most of the interactomes have few viral proteins into the reconstruct PPI network (i.e., M, nsp1, and ORF3a), while into other interactomes the viral seed protein shared own human target proteins with other viral proteins or interacts directly with other viral proteins, such as nsp3, nsp10, ORF6 and ORF10.

The lists of proteins for each reduced interactome were submitted to gene enrichment analysis on WikiPathways and KEGG pathway databases (Supplementary Tables 2 and 3, respectively). For each single viral protein interactome, the top 5% with the smallest p values of gene enrichment analysis was selected, reporting for KEGG and WikiPathways in Fig. 3A and B, respectively.

For KEGG database the gene enrichment analysis on interactomes of NS7b, ORF1a, ORF3a, and ORF8 showed pathway clusters highly significant and consistent with possible pathogenic mechanisms, such as the activation of the complement and of the coagulative cascade<sup>32</sup>, and the



**Fig. 3** Heatmap reporting top 5% smallest P values of Gene Enrichment analysis on list of proteins, obtained by every interactome of single viral seed. The values were reported as  $-\log_{10}(p \text{ value})$ . **A** P values of Gene Enrichment based on KEGG 2019. **B** P values of Gene Enrichment based on WikiPathways 2019.

TGF- $\beta$ -dominated immune response<sup>33</sup>. In particular, the NS7b interactome revealed SNARE interactions in the vesicular transport pathway, describing the mechanisms of intracellular vesicle trafficking and secretion, as the most significant pathway among all interactomes (FDR < 0.00001%). The ORF1a interactome revealed a cluster of pathways, composed by Adherent junctions, Bacteria host infection mechanisms, Regulation of actin cytoskeleton and phagocytosis, all showing the direct involvement of

these accessory proteins in a viral entry in host cells (FDR < 0.001%). The ORF8 interactome showed a pathway cluster involving complement and coagulation cascades and cardiovascular pathology (FDR < 0.3%). The ORF3a interactome revealed the TGF- $\beta$  and Cytokine signalling, endocytosis, and Vasopressin-regulation pathways (FDR < 0.01%). From pathway point of view, Proteasome pathway was involved in the interactome of M protein as well as for ORF7a, ORF8 (FDR < 0.05%) (Fig. 3A).

For WikiPathways database the gene enrichment analysis showed results consistent with KEGG gene enrichment for ORF1a, ORF3a and ORF8, and added new pathway cluster for nsp1 (FDR < 0.00001%) with DNA replication and modification pathways (Fig. 3B).

### Computational investigation on Bradykinin Storm in COVID-19

As the virus-host interactome has been efficient to describe the response to specific viral proteins, this method was applied to reconstruct the possible involvement of SARS-CoV-2 proteins in triggering the Bradykinin Storm during the infection. In this case, KNG1, precursor of bradykinin made by proteolysis cleavage, was considered as seed for RWR, imposing 200 closest proteins as limit to stop the algorithm. The resulting interactome showed NS7b, ORF3a, ORF8, and S as proximal to the seed protein, suggesting their role in the modulation of biological processes around KNG1. Indeed, it would provide a direct influence of NS7b and ORF3a proteins to activation of BK. In fact, both NS7b and ORF3a interact with the same target, endothelin converter enzyme (ECE1), and cell surface endopeptidase that converts big endothelin-1 to pressure peptide endothelin-1 (ET-1) and inactivates BK (Fig. 4A). The enrichment analysis on WikiPathways and KEGG databases revealed biological pathways of Complement and coagulation cascades (FDR < 0.000001%). Finally, to quantify the direct effect of S, ORF8, ORF3a, NS7b on Complement and coagulation cascades pathways, all proteins that belonged to the most significantly enriched pathways, were highlighted on the interactome: Human Complement System WP2806 for WikiPathways (22 of 97 genes; adj *P* value 2.9E−21) and Complement and coagulation cascades for KEGG (24 of 79 genes; adj *P* value 5.96E−27). ORF8 showed the direct interactions with Fibrinogen alpha chain (FGA), described in both pathways, and urokinase-type plasminogen activator (PLAU), tissue-type plasminogen activator (PLAT) and Plasminogen activator inhibitor 1 (SERPINE1).

### Discussion

In this study, we built a functional interactome between SARS-CoV-2 and human proteome through RWR algorithm, to identify biological mechanisms and cell responses during SARS-CoV-2 infection and to propose a model of infection contributing to a better understanding of COVID-19 pathogenesis.

SARS-CoV-2–host interactome allowed to mirror likely mechanisms of infections in different subcellular districts.

The high likelihood of our model to *in vivo* real mechanisms strengthens the representative power of the interactome and the explorative algorithm to simulate biological processes.

Moreover, the proximity among specific viral proteins into the network, such as nsp7, nsp8, nsp9, nsp10, nsp12, nsp13, nsp14 involved in the viral RNA replication, can be a good way to investigate their possible function in viral infection. The closeness among S, nsp5, nsp16, ORF9b in the network is difficult to interpret, but S, nsp5 and ORF9b, along with N, had significant positive responses of IgG antibody in sera of COVID-19 patients<sup>34</sup>. In SARS-CoV infected cell culture, the location of nsp2 in the cytoplasm and to some extent in the nucleus, as well as ORF3a, ORF7b, ORF6 and M in ER, seems to be consistent with their location among nuclear and cytoplasmic, and ER proteins respectively<sup>35</sup>.

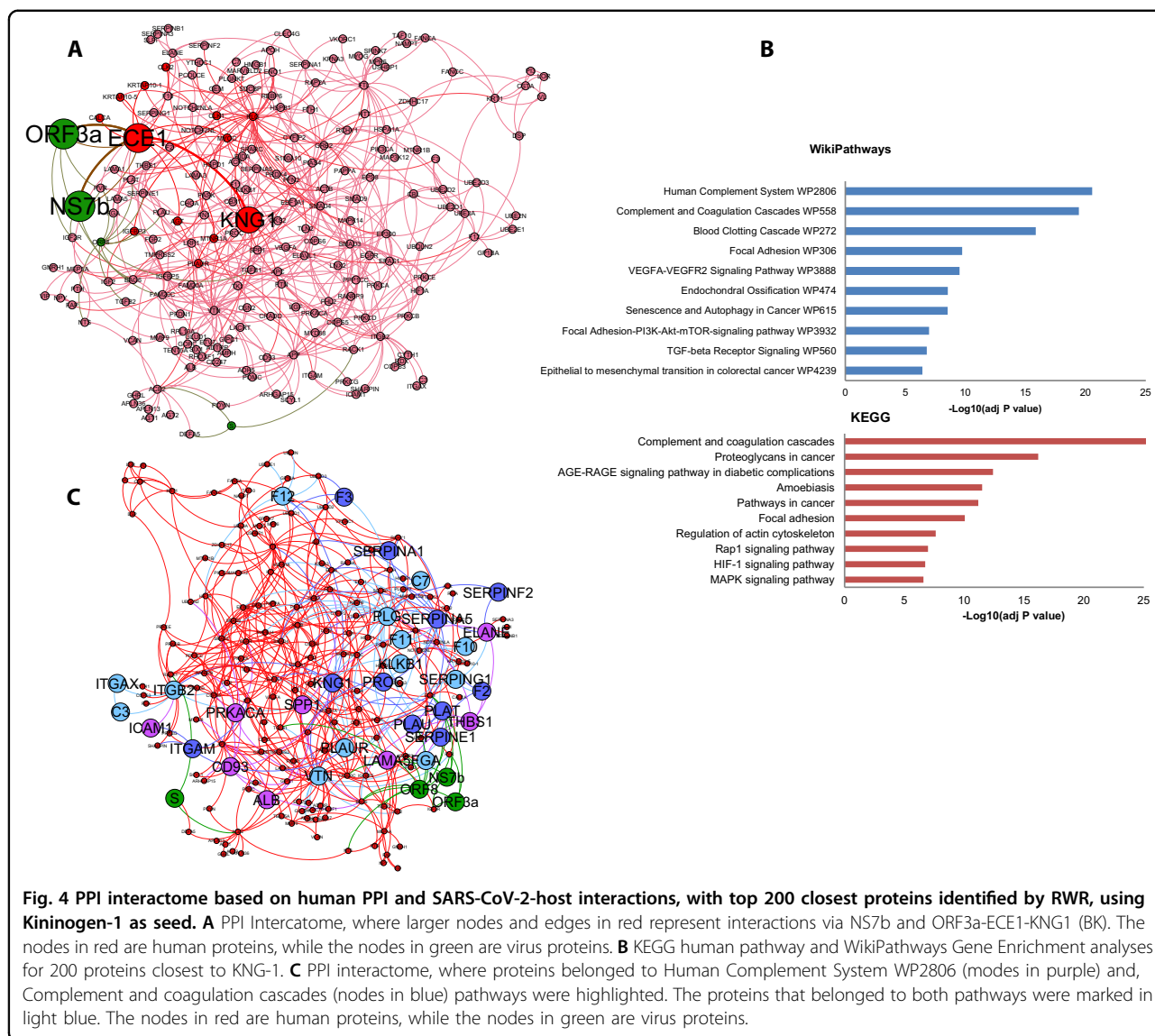
The involvement of proteasome and ubiquitination pathways, along with RNA replication, represent the principal pathways activated for assembly and replication of SARS-CoV-2 (ref. <sup>36</sup>). In fact, strong increases in RNA-modifying proteins were revealed in cell culture after infection with SARS-CoV-2 (ref. <sup>37</sup>). The ubiquitin-proteasome system deletes viral proteins to control the infection, but the virus can use them for its propagation<sup>38</sup>.

The interactomes, built around a single protein of SARS-CoV-2, allowed to draw effects on cell, involving specific pathways. Vesicular transport mechanism by SNARE interactions, identified for NS7b, is used by pathogens to penetrate host cells through their membranes and in particular in SARS-CoV<sup>39,40</sup>.

The pathways showed in ORF3a and ORF8 interactomes suggested a possible modulation of the coagulation cascade and cardiovascular pathology in COVID-19 and the involvement in Cytokine storm and water homeostasis in endothelial tissue, respectively.

Binding of ORF3a and ORF8 to TGF-β-associated factors (TGFB1, TGFB2, LTBP1, TGFBR2, FURIN, BAMBI) supports a strong involvement to begging pro-inflammation state, while the direct interactions between ORF8 and components of the Coagulation pathway, such as FGB, FGA, C5, PLAU, PROS1, SERPINE1, PLAT and CLU<sup>21,36</sup> suggest a modulation of coagulation cascade by SARS-CoV-2. Moreover, a multi-omics analysis highlighted the role of ORF3a and ORF8 during SARS-CoV-2 infection: autophagy mechanisms regulation by ORF3a and nsp6; the modulation of innate immunity by M, ORF3a and NS7b; the Integrin-TGF-β-EGFR-RTK signalling perturbation by ORF8 (ref. <sup>36</sup>).

Although the biological function of the ORF8 protein of SARS-CoV-2 remains unclear, the role of ORF8 in severe COVID-19 outcome might be supported by SARS-CoV-2 variant with a 382-nucleotide deletion (Δ382) found in Singapore in January–February 2020 linked to mild forms of COVID-19. This deletion truncates ORF 7b and locks ORF8 transcription and would be associated to less cytokine releasing during the acute phase of infection<sup>41,42</sup>.



**Fig. 4** PPI interactome based on human PPI and SARS-CoV-2-host interactions, with top 200 closest proteins identified by RWR, using Kininogen-1 as seed. **A** PPI Interactome, where larger nodes and edges in red represent interactions via NS7b and ORF3a-ECE1-KNG1 (BK). The nodes in red are human proteins, while the nodes in green are virus proteins. **B** KEGG human pathway and WikiPathways Gene Enrichment analyses for 200 proteins closest to KNG-1. **C** PPI interactome, where proteins belonged to Human Complement System WP2806 (nodes in purple) and, Complement and coagulation cascades (nodes in blue) pathways were highlighted. The proteins that belonged to both pathways were marked in light blue. The nodes in red are human proteins, while the nodes in green are virus proteins.

The systemic inflammatory response and the related endothelial dysfunction might be responsible for the microvascular injuries resulting in the formation of microthrombi, as well as a featuring microvascular thrombosis and haemorrhage pathology, viewed in the lung of COVID-19 patients<sup>43</sup>. The main pathways described in ORF3a and ORF8 interactomes, are consistent with the microvascular injuries and the development of cardiovascular complications, i.e., heart failure, myocarditis, pericarditis, vasculitis and cardiac arrhythmias observed in COVID-19 patients<sup>44</sup>. Such featuring symptoms, suggest specific mechanisms of host-virus interactions, and reinforce the role of ORF3a and ORF8 in the COVID-19 pathogenesis.

To study the pathogenesis of COVID-19 in a systemic context, Bradykinin Storm was investigated by the

interactome approach. We showed NS7b and ORF3a interacting with ECE1, which can inactivate BK<sup>45</sup>. Such observation suggests specific mechanisms of host-virus interactions, occurring in severe COVID-19 patients and reinforce the role of ORF3a to enhance the bradykinin dysregulation.

These host-virus interactions could enhance a better viral fitness during the infection, as suggested by the interaction already observed with SARS-CoV between NS7b in ECE1 (ref. <sup>36</sup>).

This interaction system might suggest a hypothesis that can explain the possible enhancing effect on Bradykinin Storm in COVID-19: the locking of ECE1 activity due to NS7b and especially ORF3a interaction, could reduce the activation of the vasoconstrictor endothelin-1, and amplify the vasodilatation effect of BK.

The ECE1 gene is much more expressed in all the tissues than ACE2 and ACE, especially in the lungs, where ECE1 is 128 fold and 3.8 fold expressed compared to ACE2 and ACE, respectively<sup>46</sup>. Significant down-expression of ECE1 protein was reported in endothelial cells infected by SARS-CoV-2 in vitro, resulting clearly consistent with effect of SARS-CoV-2 infection on BK modulation<sup>47</sup>.

BK activation is associated with coagulation disorders: F12 activates plasma kallikrein (PK), which releases bradykinin by cleavage<sup>48</sup>.

The results of our network-based model for the pathogenesis of SARS-CoV-2 mirror likely clinical manifestations in severe COVID-19 patients, noting that analysis of published data about COVID-19 allows to synthesise and describe complex aspect of pathogenesis, as reported for inflammatory cytokines storm in COVID-19-induced respiratory failure<sup>49</sup>.

This theoretical approach, combined with proteomic analysis in COVID-19 patients, might suggests new targets for specific treatments, as described for interaction between nsp10 and NKRF to mediate IL-8 expression, proposing this molecular mechanism as possible therapeutic target for anti-IL-8 and JUK inhibitor monoclonal antibodies<sup>50</sup>.

In conclusion, we developed a network-based model for SARS-CoV-2 infection, which could be a framework for pathogenic evaluation of specific clinical outcomes.

Here, the PPI interactomes were used to identify mechanistic processes of the viral molecular machinery during SARS-CoV-2 infection, for viral survival and replication within the host. With this knowledge, proteins interactions crucial for pathogenesis could be discerned. We identified different host response induced by specific proteins of SARS-CoV-2, underlining the important role of ORF3a and ORF8 in phenotypes of severe COVID-19 patients. The interactome approach applied to the identification of biological reactions around KNG1, allowed to view NS7b and ORF3a interactions with ECE1, which might have a role to enhance the Bradykinin storm. This network-based model of SARS-CoV-2–host interaction could guide to develop novel treatments against specific viral proteins.

## Materials and methods

The virus–host interactome was made by merging SARS-CoV-2–host interaction data (PPI) from Intact<sup>51</sup>, with data from human PPI databases, such as BioGrid, InnateDB-All, IMEx, IntAct, MatrixDB, MBIInfo, MINT, Reactome, Reactome-FIs, UniProt, VirHostNet, BioData, obtained by R packages PSICQUIC and biomaRt<sup>52,53</sup>. For SARS-CoV-2–host interaction experimentally obtained, 1407 protein interactors and 2305 interactions were downloads from IntAct at 12 September 2020. 28

SARS-CoV-2 proteins are used for all the analyses: E, M, N, S, nsp1, nsp2, nsp3, nsp4, nsp5, nsp6, nsp7, nsp8, nsp9, nsp10, nsp12, nsp13, nsp14, nsp15, nsp16, ORF1a, ORF3a, ORF6, ORF7a, NS7b, ORF8, ORF9b, ORF10, ORF14.

To obtain functional information about PPI in vitro, to simulate biological reactions in infected cells and to explore cell response against viral infection, we employed a RWR algorithm, a state-of-the-art guilt-by-association approach<sup>54</sup>. This algorithm allows to establish a proximity network from a given protein (seed), to study its functions, based on the premise that nodes related to similar functions tend to lie close to each other in the network. For this study, two types of interactome were performed: the whole SARS-CoV-2–host interactome and an interactome per single SARS-CoV-2 protein.

In the first imputation, every protein of SARS-CoV-2 proteome was used as seeds, with the limit of 200 closest host's proteins to every SARS-CoV-2 protein, every protein of SARS-CoV-2 proteome was chosen as seeds for RWR algorithm, imposing the limit of 200 closest host proteins to every SARS-CoV-2 protein. This analysis allowed to design SARS-CoV-2–host interactome, where the different colours correspond to the localization of the proteins within the cell.

For the second kind of interactome, one SARS-CoV-2 per time was used as seed, lowering to the closest 50 proteins to define the induced biological response as better as possible.

For each node, a score was computed as a measure of proximity to the seed protein<sup>26</sup>. In total, a large PPI interaction database was assembled, including 13334 nodes and 73584 interactions. Graphical representations of networks were performed by GEPHI 0.9.2 (ref. <sup>55</sup>). To identify hub protein in the SARS-CoV-2–host interactome, the values of betweenness centrality and degree were plotted. Betweenness centrality score measure how a specific node is in-between other nodes and then can be considered a hub, while the degree of node corresponds to number of connections.

Pathways of proteins involved in host response were tested by gene enrichment analysis on KEGG human pathways and WikiPathways databases<sup>56</sup>, the biological processes were identified by PANTHER GO<sup>57</sup>. To allow gathering of results for every running, the R package enrichR was used, an R interface to web-based tool 'Enrichr' for analysing gene sets<sup>58</sup>. The Enrichr analysis was performed using these statistical parameters: *p*-value (Fisher exact test), *q*-value (adjusted *p*-value for False Discovery Rate, FDR). Results for KEGG and WikiPathways were considered significant with a revised *p*-value < 0.05. To infer pathways involved in a single viral interactome, gene enrichment analyses for each viral interactome were collected along with *p*-values, as reported in enrichR package output. *P*-values of every single



enrichment analyses were transformed by the function  $x = -\log_{10}(p\text{-value})$  and the 5% of  $x$  values were plotted on heatmap by R package pheatmap<sup>59</sup>.

### Limitations

There are many experimental platforms for deriving such physical interactions, such as affinity purification mass-spectrometry (AP-MS) and yeast-two-hybrid (Y2H), which enable the accurate identification of interactions with a relatively long time.

The scenario reported in this study refers to few experimental data available on public databases and could be different respect to real phenotypes of COVID-19 patients.

The pathways' analysis did not consider tissue and cell-type diversity. Finally, the low threshold established for the number of nodes found by RWR (200) limited the reconstruction of the entire pathways. However, this was a software-imposed threshold. Although such a network-based approach showed great potential in identifying mechanisms not yet observed, experimental tests will be necessary to confirm what we have described. This study does not report clinical data or the outcomes of patients, but our results resulted highly consistent with clinical cases reported in literature<sup>43,44,49</sup> and in vitro studies<sup>21,36,37,47</sup>. Moreover, this approach should be used to design adaptive clinical or experimental trials.

### Acknowledgements

We gratefully acknowledge collaborators members of COVID-19 INMI Network Medicine for IDs Study Group. All the members of the COVID-19 INMI Network Medicine for IDs Study Group: Abbate Isabella, Agrati Chiara, Al Moghazi Samir, Ascoli Bartoli Tommaso, Bartolini Barbara, Capobianchi Maria Rosaria, Capone Alessandro, Goletti Delia, Rozera Gabriella, Nisii Carla, Gagliardini Roberta, Ciccossanti Fabiola, Fimia Gian Maria, Nicastrì Emanuele, Giombini Emanuela, Lanini Simone, D'Abramo Alessandra, Rinonapoli Gabriele, Girardi Enrico, Montaldo Chiara, Marconi Raffaella, Addis Antonio, Maron Bradley, Bianconi Ginestra, De Meulder Bertrand, Kennedy Jason, Khader Shabaana Abdul, Luca Francesca, Maeurer Markus, Piacentini Mauro, Merler Stefano, Pantaleo Giuseppe, Rafick-Pierre Sekaly, Sanna Serena, Segata Nicola, Zumla Alimuddin, Messina Francesco, Vairo Francesco, Lauria Francesco Nicola, Ippolito Giuseppe.

### Author details

<sup>1</sup>National Institute for Infectious Diseases "Lazzaro Spallanzani" IRCCS, Rome, Italy. <sup>2</sup>Department of Pathology, Emory University School of Medicine, Atlanta, GA, USA. <sup>3</sup>National Institute of Nuclear Physics (INFN), Bologna, Italy. <sup>4</sup>Department of Pediatric Hematology and Oncology IRCCS Ospedale Pediatrico Bambino Gesù, Rome, Italy. <sup>5</sup>Department of Infection, Division of Infection and Immunity, University College London, and National Institutes of Health and Research Biomedical Research Centre, University College London Hospitals NHS Foundation Trust, London, UK. <sup>6</sup>Champalimaud Centre for the Unknown, Lisbon, Portugal. <sup>7</sup>Medical Clinic, University of Mainz, Mainz, Germany

### Author contributions

F.M. and F.N.L. performed study concept and design; F.M., E.G., M.R.C., F.N.L., A.Z., A.Z., A.A.S., C.M., R.P.S., F.L., M.M., and G.I. performed development of methodology and writing, review, and revision of the paper; F.M., E.G., R.P.S., A.Z., and F.N.L. provided acquisition, analysis and interpretation of data, and statistical analysis; E.G. and A.A.S. provided technical and material support. All authors read and approved the final paper.

### Funding statement

National Institute for Infectious Diseases Lazzaro Spallanzani-IRCCS received financial support funded by Italian Ministry of Health, grants: Ricerca Corrente program 1- Emerging and re-emerging infections; It-IDRIN CCR-2017-23669075; and by projects COVID-2020-12371675 and COVID-2020-12371817. This work was supported also by Findus Italia, part of the Nomad Foods. Giuseppe Ippolito and Alimuddin Zumla are co-Principal Investigators of the Pan-African Network on Emerging and Re-Emerging Infections (PANDORA-ID-NET – <https://www.pandora-id.net/>) funded by the European and Developing Countries Clinical Trials Partnership the EU Horizon 2020 Framework Programme. Sir Zumla is in receipt of a National Institutes of Health Research senior investigator award and is a Mahathir Science Award Laureate.

### Code availability

R scripts and pipelines were stored on Github ([https://github.com/INMIbioinfo/SARS2\\_interactome](https://github.com/INMIbioinfo/SARS2_interactome)).

### Ethics statement

No ethics requirement was requested for the paper.

### Competing interests

All authors have an interest in infectious diseases epidemics and public health. The authors declare no competing interests.

### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41419-021-03881-8>.

Received: 26 March 2021 Revised: 21 May 2021 Accepted: 24 May 2021  
Published online: 12 August 2021

### References

- European Centre for Disease Prevention and Control (2020). Clinical characteristics of COVID-19.
- Wolff D, Nee S, Hickey N. S. & Marscholke M. Risk factors for Covid-19 severity and fatality: a structured literature review. *Infection* **49**, 15–28 (2021).
- Wolff G, Melia C. E., Snijder E. J. & Barcena M. Double-membrane vesicles as platforms for viral replication. *Trends Microbiol.* **28**, 1022–1033 (2020).
- Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
- Hillen, H. S. et al. Structure of replicating SARS-CoV-2 polymerase. *Nature* **584**, 154–156 (2020).
- te Velthuis, A. J., van den Worm, S. H. & Snijder, E. J. The SARS-coronavirus nsp7 +nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res.* **40**, 1737–1747 (2012).
- Hoffmann, M., Kleine-Weber, H. & Pohlmann, S. A Multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells. *Mol. Cell* **78**, 779–784.e775 (2020).
- Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e278 (2020).
- Garvin, M. R. et al. A mechanistic model and therapeutic interventions for COVID-19 involving a RAS-mediated bradykinin storm. *Elife* **9**, e59177 (2020).
- Schmaier, A. H. The contact activation and kallikrein/kinin systems: pathophysiological and physiologic activities. *J. Thromb. Haemost.* **14**, 28–39 (2016).
- Zhang, D. D. et al. Bradykinin stimulates renal Na(+) and K(+) excretion by inhibiting the K(+) channel (Kir4.1) in the distal convoluted tubule. *Hypertension* **72**, 361–369 (2018).
- Kjeldsen, K. Hypokalemia and sudden cardiac death. *Exp. Clin. Cardiol.* **15**, e96–e99 (2010).
- Skogestad, J. & Aronsen, J. M. Hypokalemia-induced arrhythmias and heart failure: new insights and implications for therapy. *Front. Physiol.* **9**, 1500 (2018).
- Huang, C. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506 (2020).

15. Guo, T. et al. Cardiovascular implications of fatal outcomes of patients with coronavirus disease 2019 (COVID-19). *JAMA Cardiol.* **5**, 811–818 (2020).
16. Wang, D. et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* **323**, 1061–1069 (2020).
17. Montone, R. A., Iannaccone, G., Meucci, M. C., Gurgoglione, F. & Niccoli, G. Myocardial and microvascular injury due to coronavirus disease 2019. *Eur. Cardiol.* **15**, e52 (2020).
18. Shi, S. et al. Association of cardiac injury with mortality in hospitalized patients with COVID-19 in Wuhan, China. *JAMA Cardiol.* **5**, 802–810 (2020).
19. Shi, S. et al. Characteristics and clinical significance of myocardial injury in patients with severe coronavirus disease 2019. *Eur. Heart J.* **41**, 2070–2079 (2020).
20. Gu, Z. C. et al. Incidence of myocardial injury in coronavirus disease 2019 (COVID-19): a pooled analysis of 7,679 patients from 53 studies. *Cardiovasc. Diagn. Ther.* **10**, 667–677 (2020).
21. Gordon, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).
22. Hazra, S., Chaudhuri, A. G., Tiwary, B. K. & Chakrabarti, N. Matrix metalloproteinase 9 as a host protein target of chloroquine and melatonin for immunoregulation in COVID-19: a network-based meta-analysis. *Life Sci.* **257**, 118096 (2020).
23. Zhou, Y. et al. A network medicine approach to investigation and population based validation of disease manifestations and drug repurposing for COVID-19. *PLoS Biol.* **18**, e3000970 (2020).
24. Law, J. N. et al. Identifying human interactors of SARS-CoV-2 proteins and drug targets for COVID-19 using network-based label propagation. Preprint at <https://arxiv.org/abs/2006.01968> (2020).
25. Srinivasan, S. et al. Structural genomics of SARS-CoV-2 indicates evolutionary conserved functional regions of viral proteins. *Viruses* **12**, 360 (2020).
26. Messina, F. et al. COVID-19: viral-host interactome analyzed by network based-approach model to study pathogenesis of SARS-CoV-2 infection. *J. Transl. Med.* **18**, 233 (2020).
27. Barabasi, A. L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* **12**, 56–68 (2011).
28. Menche, J. et al. Disease networks. Uncovering disease-disease relationships through the incomplete interactome. *Science* **347**, 1257601 (2015).
29. Ostaszewski, M. et al. COVID-19 Disease Map, building a computational repository of SARS-CoV-2 virus-host interaction mechanisms. *Sci. Data* **7**, 136 (2020).
30. Mazein, A. et al. Systems medicine disease maps: community-driven comprehensive representation of disease mechanisms. *NPJ Syst. Biol. Appl.* **4**, 21 (2018).
31. Ostaszewski, M. et al. COVID-19 Disease Map, a computational knowledge repository of SARS-CoV-2 virus-host interaction mechanisms. Preprint at <https://www.biorxiv.org/content/10.1101/2020.10.26.356014v1> (2020).
32. Pena, A. L. B., Oliveira, R. A., Severo, R. G. & Simoes, E. S. A. C. COVID-19 related coagulopathy: what is known up to now. *Curr. Med. Chem.* **28**, 4207–4225 (2020).
33. Ferreira-Gomes, M. et al. SARS-CoV-2 in severe COVID-19 induces a TGF- $\beta$ -dominated chronic immune response that does not target itself. *Nat. Commun.* **12**, 1961 (2021).
34. Jiang, H. W. et al. SARS-CoV-2 proteome microarray for global profiling of COVID-19 specific IgG and IgM responses. *Nat. Commun.* **11**, 3581 (2020).
35. von Brunn, A. et al. Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFome. *PLoS ONE* **2**, e459 (2007).
36. Stukalov, A. et al. Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. *Nature* **94**, 246–252 (2021).
37. Bojkova, D. et al. Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* **583**, 469–472 (2020).
38. Tang, Q., Wu, P., Chen, H. & Li, G. Pleiotropic roles of the ubiquitin-proteasome system during viral propagation. *Life Sci.* **207**, 350–354 (2018).
39. Wesolowski, J. & Paumet, F. SNARE motif: a common motif used by pathogens to manipulate membrane fusion. *Virulence* **1**, 319–324 (2010).
40. Xu, Y. et al. Crystal structure of severe acute respiratory syndrome coronavirus spike protein fusion core. *J. Biol. Chem.* **279**, 49414–49419 (2004).
41. Su, Y. C. F. et al. Discovery and Genomic Characterization of a 382-Nucleotide Deletion in ORF7b and ORF8 during the Early Evolution of SARS-CoV-2. *mBio* **11**, e01610-20 (2020).
42. Young, B. E. et al. Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. *Lancet* **396**, 603–611 (2020).
43. McGonagle, D., O'Donnell, J. S., Sharif, K., Emery, P. & Bridgewood, C. Immune mechanisms of pulmonary intravascular coagulopathy in COVID-19 pneumonia. *Lancet Rheumatol.* **2**, e437–e445 (2020).
44. 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.* **220**, 1–13 (2020).
45. Padilla, B. E. et al. Endothelin-converting enzyme-1 regulates endosomal sorting of calcitonin receptor-like receptor and beta-arrestins. *J. Cell Biol.* **179**, 981–997 (2007).
46. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
47. Wang, P. et al. A cross-talk between epithelium and endothelium mediates human alveolar-capillary injury during SARS-CoV-2 infection. *Cell Death Dis.* **11**, 1042 (2020).
48. Hofman, Z., de Maat, S., Hack, C. E. & Maas, C. Bradykinin: inflammatory product of the coagulation system. *Clin. Rev. Allergy Immunol.* **51**, 152–161 (2016).
49. Leisman, D. E. et al. Cytokine elevation in severe and critical COVID-19: a rapid systematic review, meta-analysis, and comparison with other inflammatory syndromes. *Lancet Respir. Med.* **8**, 1233–1244 (2020).
50. Li, J. et al. Virus-host interactome and proteomic survey reveal potential virulence factors influencing SARS-CoV-2 pathogenesis. *Medicine* **2**, 99–112 e117 (2021).
51. Hermjakob, H. et al. IntAct: an open source molecular interaction database. *Nucleic Acids Res.* **32**, D452–D455 (2004).
52. Aranda, B. et al. PSICQUIC and PSISCORE: accessing and scoring molecular interactions. *Nat. Methods* **8**, 528–529 (2011).
53. Smedley, D. et al. BioMart—biological queries made easy. *BMC Genomics* **10**, 22 (2009).
54. Valdeolivas, A. et al. Random walk with restart on multiplex and heterogeneous biological networks. *Bioinformatics* **35**, 497–505 (2019).
55. Bastian, M., Heymann, S. & Jacomy, M. Gephi: an open source software for exploring and manipulating networks. In International AAAI Conference on Weblogs and Social Media. (2009).
56. Slenter, D. N. et al. WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res.* **46**, D661–D667 (2018).
57. Mi, H., Muruganujan, A., Ebert, D., Huang, X. & Thomas, P. D. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **47**, D419–D426 (2019).
58. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
59. Kolde R. pheatmap: Pretty heatmaps [Software] (2015).