Computational modeling of human-nCoV protein-protein interaction network

Sovan Saha [a,1], Anup Kumar Halder [b,1], Soumyendu Sekhar Bandyopadhyay [c,e],
Piyali Chatterjee [d], Mita Nasipuri [e], Subhadip Basu [e,*]

[a] Department of Computer Science & Engineering, Institute of Engineering & Management, Salt Lake Electronics Complex, Kolkata 700091, West Bengal, India
[b] Department of Computer Science & Engineering, University of Engineering & Management, Kolkata 700156, West Bengal, India
[c] Department of Computer Science & Engineering, School of Engineering and Technology, Adamas University, Kolkata 700126, West Bengal, India
[d] Department of Computer Science & Engineering, Netaji Subhash Engineering College, Garia, Kolkata, West Bengal 700152, India
[e] Department of Computer Science & Engineering, Jadavpur University, Jadavpur, Kolkata, West Bengal 700032, India

ARTICLE INFO

ABSTRACT

Novel coronavirus (SARS-CoV2) replicates the host cell's genome by interacting with the host proteins. Due to this fact, the identification of virus and host protein–protein interactions could be beneficial in understanding the disease transmission behavior of the virus as well as in potential COVID-19 drug identification. International Committee on Taxonomy of Viruses (ICTV) has declared that nCoV is highly genetically similar to the SARS-CoV epidemic in 2003 (~89% similarity). With this hypothesis, the present work focuses on developing a computational model for the nCoV-Human protein interaction network, using the experimentally validated SARS-CoV-Human protein interactions. Initially, level-1 and level-2 human spreader proteins are identified in the SARS-CoV-Human interaction network, using *Susceptible-Infected-Susceptible* (SIS) model. These proteins are considered potential human targets for nCoV bait proteins. A gene-ontology-based *fuzzy* affinity function has been used to construct the nCoV-Human protein interaction network at a ~99.98% specificity threshold. This also identifies 37 level-1 human spreaders for COVID-19 in the human protein-interaction network. 2474 level-2 human spreaders are subsequently identified using the SIS model. The derived host-pathogen interaction network is finally validated using six potential FDA-listed drugs for COVID-19 with significant overlap between the known drug target proteins and the identified spreader proteins.

1. Introduction

COVID-19 evolved in the Chinese city of Wuhan (Hubei province) [1]. The first case of human species affected by nCoV was observed on 31st December 2019 [2]. Soon it expands its adverse effect on almost all nations within a brief period [3]. World Health Organization (WHO) observes that the massive disastrous outbreak of nCoV is mainly due to mass community spreading and declares a global health emergency on 30th January 2020 [4]. After proper assessment, WHO presumes its fatality rate to be 4% [5] which urges the global researchers to work together to discover an appropriate treatment for this pandemic [6,7].

Coronaviridae is the family to which a coronavirus belongs. It also infects birds and mammals besides affecting human beings. Though the common symptoms of the coronavirus are common cold, cough, etc., it is accompanied by severe acute, chronic respiratory disease and multiple organ failure leading to human death. Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) were the two major outbreaks in 2003 and 2012 before SARS-CoV2. The source of origin of SARS was located in Southern China. Its fatality rate was within 14%–15% [8], due to which 774 people lost their lives among 8804 affected cases. Saudi Arabia was marked as the base for the commencement of MERS. 858 persons among 2494 infected cases were defeated in their battle against the MERS virus. Thus it generated a much higher fatality rate of 34.4% [9] when compared to that of SARS.

All three epidemic creators SARS, MERS, and SARS-CoV2, are biologically included in the genus Beta coronavirus under the Coronaviridae. Both structural and non-structural proteins are involved in the formation of SARS-CoV2. Out of the two, structural proteins like the envelope (E) protein, membrane (M) protein, nucleocapsid (N) protein and the spike (S) protein play a significant role in transmitting the disease by binding with the receptors after entering into the human body [10]. So, there is an urgent need to understand and analyze the mechanism of disease transmission of this new virus.

In this research work, Protein-Protein Interaction Networks (PPIN) are the most significant attribute to study the disease propagation mechanism from SARS-CoV2 to humans. It plays a crucial role in identifying essential proteins [11–15] responsible for various diseases. They are also significant in identifying protein functions [16–19]. According to Lotem *et al.* [20], though human PPIN is constantly expanding, very little information is available about the human PPIN, which gets generated in disease conditions. With the enhancement in the availability of the human PPIN data, the primary focus of research has been shifted from the basic understanding of the PPIN to the study of the PPIN underlying various kind of human diseases [21]. According to the work of Ideker *et al.* [22], PPIN study mainly falls under four categories: 1) Identification of human disease genes based on network analysis, 2) Implication of additional genes involved in the disease by using protein networks, 3) Identification of protein subnetworks involved in diseases and 4) Classification of case-control studies based on protein PPIN.

Host-pathogen PPINs are significant for understanding the mechanism of transmission of infection, which is essential for developing new and more effective therapeutics, including rational drug design. Progression of infection and disease results due to the interaction of proteins in between pathogen and host. Pathogen plays an active role in spreading infection. Pathogen and host PPIN permit pathogenic microorganisms to utilize host capabilities by manipulating the host mechanisms to abscond from the host's immune responses [23–25]. Detection of target proteins, through the analysis of pathogen and host, PPIN is the central point of research [13,26,27]. Topologically significant proteins having a higher degree of interactions are generally found to be an important drug targets. However, proteins with fewer interactions or topologically not substantial may be involved in the mechanism of infection because of some biological pathway relevance.

However, clinically validated Human-nCoV protein interaction data is limited in the current literature. This has motivated us to develop a new computational model for the nCoV-Human PPI network. We have subsequently validated the proteins involved in the host-pathogen interactions with respect to potential Food and Drug Administration (FDA) drugs for COVID-19 treatment. Key aspects of this research work are highlighted below: It has been reported that SARS-CoV has ~89% [28,29] genetic similarities with nCoV. SARS-CoV-Human protein–-protein interaction network has also been studied widely and available in the literature [30–32]. Recently, we developed a computational model to identify potential spreader proteins in a HumanSARS CoV interaction network using the SIS model [14]. In addition, sequence information of 29 nCoV proteins has been released [33]. Gene ontological (GO) information (Biological Process (BP), Molecular Function (MF), Cellular Component (CC)) of 14 of the nCoV proteins are available [33,34]. We have recently developed a method to predict interaction affinity between proteins from the available GO graph [35]. Assessment of interaction affinity between nCoV proteins with potential Human target/bait proteins, which are susceptible to SARS-CoV infection, has been done. *Fuzzy* affinity thresholding is done to detect High-Quality nCoV-Human PPIN. The selected human proteins are considered as level-1 human spreader nodes of nCoV. Level 2 spreader nodes in nCoV-Human PPIN are detected using the spreadability index and validated by SIS [14,36] model. Our developed model is validated for the target proteins of the potential FDA drugs for COVID-19 treatment [37]. All the related terminologies referred in this work like Ego Network [38] (please see Fig S1 in the supplementary document), spreadability index

[14], Node weight [39], Edge ratio [38], Neighborhood density [38], Betweenness Centrality (BC) [40], Closeness centrality (CC) [41], Degree centrality (DC) [42] and Local average centrality (LAC) [43] are discussed in the supplementary document.

2. Methodology

Our developed computational model for nCoV-Human PPIN consists of two crucial methodologies 1) identification of spreader nodes by spreadability index along with the validation of SIS model and 2) *Fuzzy* PPI model. First, the SIS model identifies spreader nodes of SARS-CoV proteins (candidate set of nCoV interactors). Then, the *Fuzzy* PPI model is applied to extract the nCoV-Human interactions, and finally, nCoV spreaders are identified using the SIS model.

2.1. Identification of spreader nodes by spreadability index along with the validation by SIS model

In nCoV-Human PPIN, the former acts as a pathogen/bait while the host, the human, acts as 'Prey'. The transmission of infection starts when a pathogen enters a host body and infects its protein, affecting its directly or indirectly connected neighbor proteins. Considering this method of transmission, PPIN of humans and SARS-CoV are considered to detect spreader nodes. Spreader nodes are those nodes/proteins which transmit the disease fast among their neighbors. But not all the nodes in a PPIN are spreaders. So, proper detection of spreader nodes is crucial. Thus, spreader nodes are identified by the spreadability index, which measures the transmission capability of a node/protein. Furthermore, the compactness of PPIN and its transferal capability are evaluated using centrality analysis. Nodes with high centrality values are usually considered spreader nodes or the most critical node in a network.

The spreadability index [14] is one of the centrality-based measures that combines three major topological neighborhood-based features of a network. They are 1) Node weight [39] 2) Edge ratio [38] and 3) Neighborhood density [38]. Nodes having a high spreadability index are considered spreader nodes. The spreader nodes thus identified are also validated by the SIS model [36]. The SIS model is implemented to design the SARS-CoV and SARS-CoV2 outbreak into a disease model consisting of proteins based on their present infection status. A protein can be in either of the three states: 1) S: Susceptible, which means that every protein is initially susceptible though not yet infected but at risk of getting infected by the disease. 2) I: Infected, which means that the disease already infects the protein and 3) S: Susceptible, which means proteins again become susceptible after getting recovered from the infected state. This model is implemented to generate the overall infection capability of a node after a certain range of iterations. Thus the sum of the infection capability of the top selected spreader nodes is computed by this model, which is compared against the sum obtained for the selected top critical nodes by other existing centrality measures like Betweenness Centrality (BC) [40], Closeness centrality (CC) [41], Degree centrality (DC) [42] and Local average centrality (LAC) [43] (Please see the supplementary document for more details).

Our proposed method for selecting spreader nodes in SARS-CoV PPIN [14] has performed better than the other existing *state-of-the-art* like Betweenness Centrality (BC) [40], Closeness Centrality (CC) [41], Degree centrality (DC) [42], and Local average centrality (LAC) [43]. The comparison and results are included in our recently published work [14] (Please see Fig. S2 and Tables S1–S5 in the supplementary document for more details). The complete source code is available online here. A synthetic PPIN is considered in Fig. 1 to demonstrate the entire methodology of the spreadability index (see Table 1). In addition, computational analysis of the spreadability index of our proposed model with one of the other methodology LAC (computed by using CytoNCA [44]) has been highlighted in Table 2. $E_{out}^{S_i}$ is the total number of edges that are
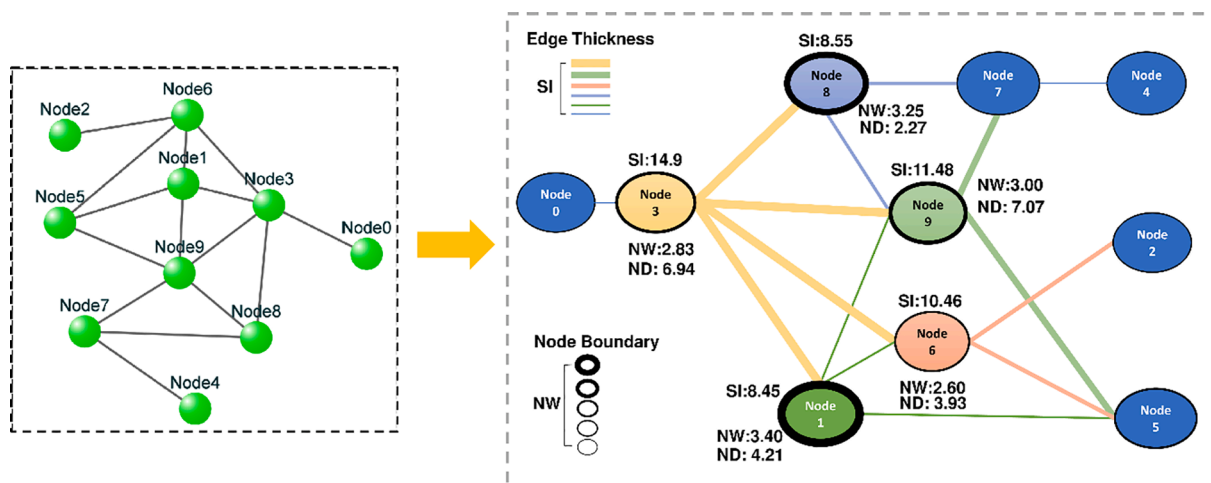
Fig. 1. Synthetic protein–protein interaction network. The network consists of 10 nodes and 25 edges. The neighborhood density (ND), Node Weight (NW), and Spreadability Index (SI) of the top 5 nodes have been highlighted. While the thickness of the edges highlights the rank according to SI, the thickness of the boundary of the nodes highlights the rank according to NW.

Table 1
Computation of spreadability index of Fig. 1 and validation of selected top 5 spreader nodes by the SIS model.

Rank	Proteins	$E_{out}^{S_i}$	$E_{in}^{S_i}$	Edge Ratio	Neighborhood Density	Node Weight	Spreadability Index	Sum of SIS infection rate of top 5 nodes
1	Node 3	6	3	1.75	6.94	2.83	14.99	**1.19**
2	Node 9	5	4	1.20	7.07	3.00	11.48	
3	Node 6	5	2	2.00	3.93	2.60	10.46	
4	Node 8	6	2	2.33	2.27	3.25	8.55	
5	Node 1	5	4	1.20	4.21	3.40	8.45	

Table 2
Computation of the LAC of the synthetic network (Fig. 1) and validation of selected top 5 spreader nodes by the SIS model.

Rank	Proteins	LAC	Sum of SIS infection rate of top 5 nodes
1	Node 1	2.00	**0.86**
2	Node 9	1.60	
3	Node 5	1.33	
4	Node 8	1.33	
5	Node 3	1.20	

outgoing from the ego network S_i [14] (for details please see the supplementary document) whereas $E_{in}^{S_i}$ is denoted as the total number of interconnections in the neighbor of node i [38]. $E_{out}^{S_3}$ of node 3 is 6 while $E_{in}^{S_3}$ of node 3 is 3, which highlights that node 3 has the highest transmission ability from its ego network to outside when compared to other nodes. Node 3 also has the highest spreadability index. But LAC failed to rank node 3 in the first position. The same scenario can be observed for some other nodes in the synthetic network too. Besides SIS validation result shows that the selected top-ranked spreader nodes in this proposed model have the highest infection capability compared to the other ranked nodes.

2.2. Fuzzy PPI model for potential SARS-CoV2-human interaction identification

The binding affinity between any two interacting proteins can be estimated by combining the semantic similarity scores of the GO terms associated with the proteins [26,35,45–47]. A greater number of semantically similar GO annotations between any protein pair indicate higher interaction affinity. The *fuzzy* PPI model is a hybrid approach [35] that utilizes both the topological [48] features of the GO graph and information contents [47,49,50] of the GO terms.

GO is organized in three independent directed acyclic graphs (DAGs): molecular function (MF), biological process (BP), and cellular component (CC) [34]. The nodes in each GO graph represent GO terms, and the edges represent different hierarchical relationships. In this work, the two most essential relations, *'is_a'* and *'part_of,'* have been used for GO relations [51].

The semantic similarity between any two proteins is estimated by considering the similarities between their all

pairs of annotating gene ontology (GO) terms belonging to a particular ontological graph. The similarity of a GO term pair is determined by considering specific topological properties (shortest path length) of the GO graph and the average information content (IC) [52] of the disjunctive common ancestors (DjCAs) [45,46] of the GO terms as proposed in [35]. *Fuzzy* PPI first relies on a *fuzzy* clustering of the GO graph where the selection of GO terms as cluster center is based on the level of association of that GO term in the GO graph. Then, the cluster centers are selected based on the proportion measure of GO terms. The proportion measure for any GO term t is computed as

$$PrM(t) = \frac{|An(t)| + |Dn(t)|}{|N_O|} \qquad (1)$$

where (t), $Dn(t)$ represents the ascendant and descendant of term t, and N_O is the total number of GO terms in ontology O. A higher value of the proportion measure (*PrM*) signifies higher coverage of ascendants and descendants associated with the specific node. Finally, the GO terms for which this proportion measure is above a predefined threshold are selected as cluster centers. In this work, the cluster centers are chosen based on the threshold values as suggested in [26,35].

After selecting the cluster centers, the degree of membership of a GO term to each of the selected cluster centers

is calculated using its respective shortest path lengths to the corresponding cluster centers. The membership of the GO term to a cluster decreases with an increase in its shortest path length to the cluster

center. The membership function is defined as

$$MmFc(t) = e^{\frac{-(x-c_i)^2}{2k^2}} \qquad (2)$$

where c_i is $i-th$ center and k is the width of the membership function, and x is the shortest path length from t to c_i. The difference $D(t_i, t_j)$ in membership values between the GO pair t_i and t_j for each cluster center, is computed to find the weight parameter. The weight parameter is defined as

$$Wt(t_i, t_j) = 1 - \max D(t_i, t_j) \qquad (3)$$

This weight value determines how different two GO terms can be with respect to the cluster centers. Next, the shared information content (SIC) is computed using the average information content (IC) [52] of the $DjCAs$ of the GO term pair (t_i, t_j) for all three GO graphs. The SIC is defined as

$$SIC(t_i, t_j) = \frac{\sum_{a \in DjCA(t_i, t_j)} IC(a)}{|DjCA(t_i, t_j)|} \qquad (4)$$

where $DjCA(t_i, t_j)$ represents the disjunctive common ancestors of GO-term t_i and t_j. The semantic similarity (SS) between the GO term pair t_i and t_j defined as

$$SS(t_i, t_j) = Wt(t_i, t_j) SIC(t_i, t_j) \qquad (5)$$

The semantic similarity of protein pair (P_i, P_j) for each GO-type (CC, MF, and BP), is estimated by utilizing the maximum similarity of all possible GO pairs from the annotations of proteins P_i and P_j for each type of GO. The interaction affinity of protein pair (P_i, P_j) is defined as the average of CC, MF, and BP-based semantic similarity.

This work uses the available ontological information to calculate the fuzzy interaction affinity score between the protein pairs of SARS-CoV2 and spreader human proteins (please see Fig. 3). Here, the SARS-COV's level-1 and level-2 spreader proteins are employed as the primary target for the proposed fuzzy PPI model for interaction affinity computation. A bipartite relation of GO pairs is primarily generated from each pair of proteins for each type of GO annotations (CC, MF, and BP) independently (Fig. 3A). To reduce the computational overhead and time, semantic similarity scores are previously computed between all GO pairs belonging to a particular GO type using equation (5) [35]. The semantic similarity is computed by exploring the topological properties of the GO subgraphs. For each type of GO subgraphs, a different set of cluster center nodes (GO terms) are identified based on proportion measure (equation (1)) that rely on the annotation score and GO relationship graph hierarchy. The GO semantic similarity is estimated with a distance-based measure between the target GO pair by exploring the membership score (equation (2), 3) and values (equation (4)) compared to respective cluster canters of each GO subgraphs (Fig. 3B). For each GO type, the max of all possible scores of the bipartite links in a particular GO subgraph is considered the final semantic score of that type of GO.

Similarly, all three different scores are evaluated and averaged to find the interaction affinity for the annotated protein pair. Then, the fuzzy score of interaction affinity is computed by normalizing the interaction affinity using max–min normalization. Finally, with high specificity threshold (please see Fig. 6), high-quality interactions (78 interactions involving 37 human level-1 spreaders) are extracted for human-SARSCoV2.

3. Dataset description

SARS-CoV-Human PPIN serves as a baseline for our model. The potential level-1 and level-2 human spreaders of SARS-CoV become the possible candidate set for selecting level-1 human spreaders of SARS-CoV2. Various datasets have been curated for this purpose which has been outlined below:

3.1. Human PPIN

The dataset [53,54] consists of all possible interactions between human proteins experimentally documented in humans. Human
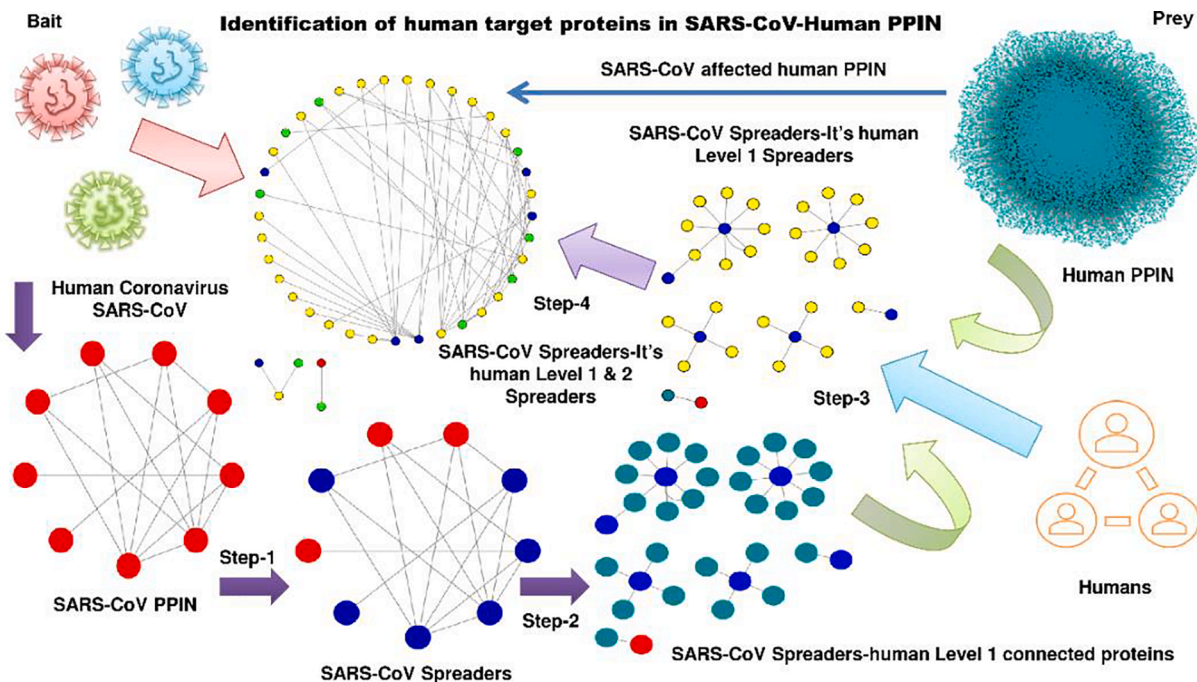

Fig. 2. A computational model for the selection of spreader nodes in Human-SARS-CoV PPIN by spreadability index. Red-coloured nodes represent SARS-CoV proteins, while blue-colored nodes are the selected spreader nodes in it. Deep green colored nodes represent level-1 human connected proteins with SARS-CoV proteins, while yellow-coloured nodes represent the selected human spreaders. Light green colored nodes represent level-2 human spreaders of SARS-CoV. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
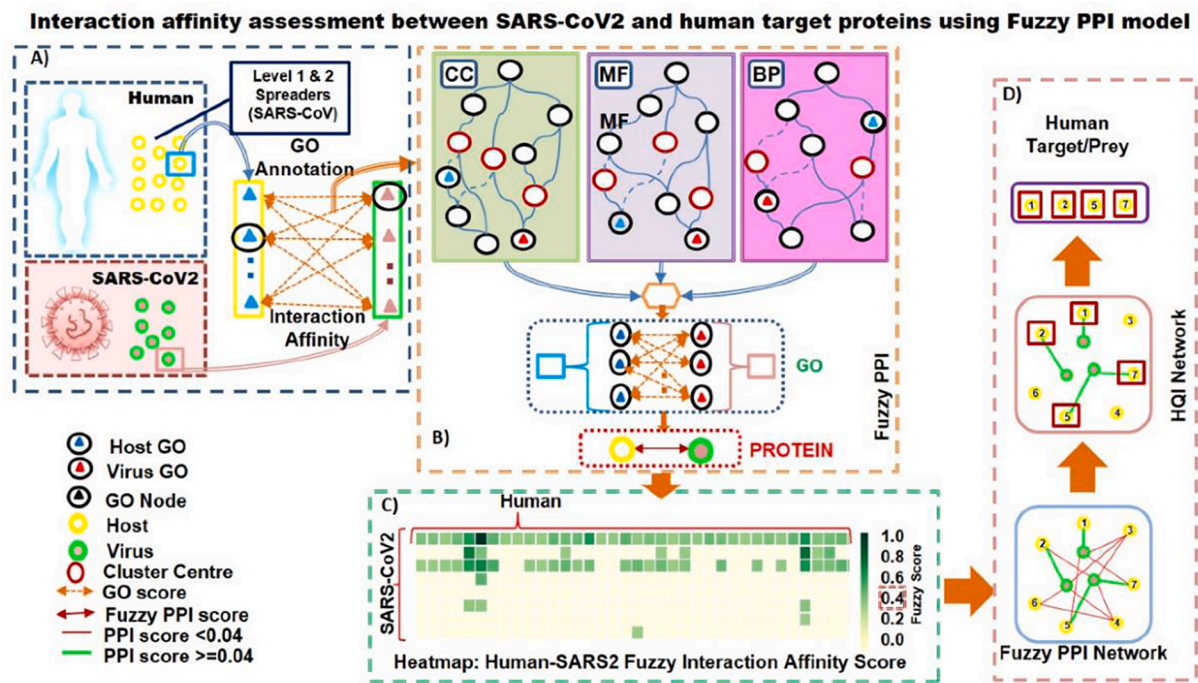
Fig. 3. Schematic diagram of Fuzzy PPI model. A) The fuzzy PPI model finds the interaction affinity between the SARS-CoV2 and Human proteins (L1 and L2 spreader of SARS-CoV) using ontological gene information. B) All GO pair-wise interaction affinities are assessed from three independent GO-relationship graphs CC, MF, and BP. The fuzzy interaction affinity of a protein pair is computed from all three pair-wise scores of all GO-pair affinities. C) Heatmap representation of Fuzzy PPI score. D) Network representation of Human and SARS-CoV2 proteins with 0.2 onward thresholds of Fuzzy PPI score at high specificity. Finally, high-quality interactions are extracted to retrieve the potential human prey for SARS-CoV2 at the 0.4 threshold.

proteins are represented as nodes, while edges represent the physical interactions between proteins. It is a collection of 21,557 nodes and includes 342,353 edges/interactions.

3.2. SARS-CoV PPIN

The dataset [30] consists of interactions between SARS-CoV proteins. It contains 7 unique proteins and the involvement of 17 interacting edges. Only the densely connected proteins are considered rather than the isolated ones since the former play a more active role in the transmission of infection than the latter.

3.3. SARS-CoV-human PPIN

The dataset [30] comprises 118 interactions between SARS-CoV and humans. It is used to fetch the level-1 human interactions of SARS-CoV.

3.4. SARS-CoV2 proteins

This data is collected from the pre-released dataset of available SARS-CoV2 protein from UniProtKB [33 55], which includes 14 reviewed SARS-CoV2 proteins.

3.5. GO graph and protein-GO annotations

GO graph types (CC, MF, and BP) are collected from GO Consortium [34,51]. In addition, the protein to GO-annotation map is retrieved from the UniProtKB database.

3.6. Potential COVID-19 FDA drugs

Six potential FDA drugs: Lopinavir [56], Ritonavir [57], Azithromycin [58], Remdesivir [59–61], Favipiravir [62,63], and Darunavir [64] have been identified from the DrugBank [65] published

white paper [37] which have been used for validation in our proposed model.

4. Results and discussion

Our developed computational model of nCoV-Human PPIN contains high-quality interactions (HQI) and proteins identified by *Fuzzy* affinity thresholding and spreadability index validated by the SIS model. The sources of input and the generated results always play a crucial role in any computational model, which is also true for our proposed model.

4.1. Spreader nodes selection in Human-SARS CoV interaction network using spreadability index

SARS-CoV-Human PPIN (up to level-2) is formed by the combination of SARS-CoV-Human and Human-Human PPIN datasets. SARS-CoV-Human dataset generates the direct level-1 human interactions of SARS-CoV, while the human–human PPIN dataset is used to fetch the corresponding level-2 human interactions. Potential spreader nodes are identified using the spreadability index validated by the SIS model [14]. The entire process of the detection of spreader nodes in SARS-CoV-Human PPIN is depicted in four steps in Fig. 2 (used only for description): 1) Spreader nodes (6 spreaders) in SARS-CoV PPIN are detected by spreadability index. 2) Corresponding level-1 human proteins of the spreader nodes in SARS-CoV PPIN are identified. 3) Spreader nodes (24 spreaders) in level-1 human proteins of the spreader nodes in SARS-CoV PPIN are detected. 4) The same process is repeated, and spreader nodes (9 spreaders) in level-2 human proteins of the spreader nodes in SARS-CoV PPIN are identified. The selected spreader nodes in SARS-COV-Human PPIN are highlighted in additional Table A1, Table A2, and Table A3. The network view of SARS-CoV-Human PPIN at each level and various selected thresholds of spreadability index are also available online (SARS-CoV human spreaders link: L-1, human spreaders at the high threshold of spreadability index link: L1 & L2:high, and human

spreaders at the low threshold link: L1 & L2:low).

4.2. Identification of the nCoV-Human proteins interactions using fuzzy PPI model

The GO information can be helpful to infer the binding affinity of any pair of interacting proteins using three different types of GO hierarchical relationship graphs (CC, MF, and BP) [34]. The *fuzzy* PPI model has been applied to find the interaction affinity between the SARS-CoV2 and Human proteins using GO-based information (please see Fig. 3 and section 2.2 for details). To identify the interactors of SARS-CoV2 on humans using the *Fuzzy* PPI model, a set of candidate proteins are selected, which are identified as the L1 and L2 spreader nodes of SARS-CoV using the SIS model (as depicted in Fig. 2). The *fuzzy* PPI model is constructed from the ontological relationship graphs by evaluating the affinity between all possible GO pairs annotated from any target protein pair. Finally, the *fuzzy* score of interaction affinity of protein pair is computed from these GO pair-wise interaction affinity into a range of [0,1].

We have used experimentally validated human protein interactions (physical only) from publicly available interaction databases, such as HIPPIE [66], STRING [67], BioGRID [68], DIP [69], HuRI [70] for positive data and Negatome 2.0 [71], Trabuco *et al.* [72] for negative data. The positive interactions are also filtered by removing the edges that are common in both positive and negative interaction sets. In each database, Gold standard data is curated by using the scoring scheme provided by the respective databases. The selection criteria are described in Table S6 in the supplementary document.

With this benchmarking data set, the FuzzyPPI Model has been assessed with different fuzzy scoring cut-off values. The performance of this assessment is reported in Table S7 in the supplementary document. In any classification task, specificity signifies the ability to identify a positive sample correctly. In order to identify high-quality positive interactions, we used the specificity metric. With the increasing value of specificity, the number of false-positive (FP) interactions has shown a sharp fall as depicted in the following table. At threshold ≥0.2 and ≥0.4, the FP is 0.0048% and 0.0001% of total negative interactions respectively. Thus, the Specificity threshold is set at ≥0.4. The heatmap representation of fuzzy interaction affinities (with a score ≥of 0.2 for very high specificity ~99%) is shown in additional Fig. A1 and Table A4. The high-quality interaction (HQI) is retrieved at threshold 0.4 (almost ~99:98% Specificity), which results in a total of 78 interactions between SARS-CoV2 and humans (37 human level-1 spreaders). The interaction networks predicted from the *Fuzzy*-PPI model are shown in Fig. 4.

4.3. Identification of human spreader proteins for nCoV

Human proteins present in the high-quality interactions of nCoV-Human PPIN fetched by applying *fuzzy* affinity threshold are considered level-1 spreaders. From these 37 level-1 spreaders, corresponding level-2 human interactions are obtained using the human–human PPIN dataset. Spreadability index is thus computed for these level-2 human proteins for the identification of level-2 human spreader nodes. The SIS model also verifies the selection. The selected spreader nodes in SARS-COV2-Human PPIN (2474 level-2 human spreaders under high threshold) are highlighted in additional Table A4, Table A5, and Table A6. In addition, the computational model of nCoV-Human PPIN under a high threshold has been highlighted online https://kzeumvafuq8hob5bzpsphq-on.drv.tw/www.40highthreshold_sarscov2.com/sarcov2_graph_40_percent_high_threshold.html. It highlights the human level-1 (marked in yellow) and level-2 spreader nodes (marked in green). The network view of SARS-CoV2-Human PPIN at each level and various selected thresholds are also available online (SARS-CoV2 Level-1 human spreaders, Level-1 & Level-2:high spreaders at the high threshold of spreadability index and Level-1 & Level-2:low human spreaders at a low threshold of spreadability index).
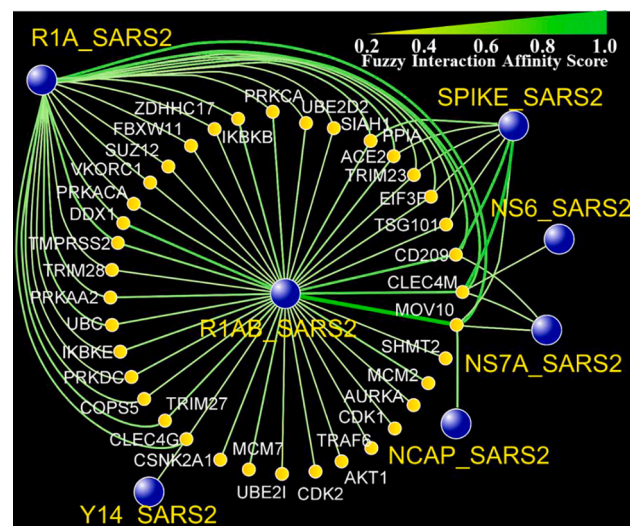

Fig. 4. Network representation of HQIs (score ≥ 0.4) between SARS-CoV2 and human proteins. Blue and yellow spherical nodes represent the SARS-CoV2 and human proteins, respectively. The edge width reflects the fuzzy score of binding affinity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.4. Validation using potential FDA drugs for COVID-19

After proper assessment of all potential drugs as mentioned in the DrugBank [65] white paper [37], six drugs: Lopinavir [56], Ritonavir [57], Azithromycin [58], Remdesivir [59–61], Favipiravir [62,63] and Darunavir [64] are identified which are showing expected results to some extent in the clinical trials done for SARS-CoV2 vaccine. All approved human protein targets for each of the five approved drugs are fetched from the advanced search section [73] of the drug bank [65,74]. When searched in our proposed model of nCoV-Human PPIN, these targets are found to play an active role in spreader nodes. This reveals that the selected spreader nodes are of biological importance in transmitting infection in a network that makes them the protein drug targets of the potential FDA drugs for COVID-19. The target protein hits in our nCoV-Human PPIN for each of the 7 potential FDA drugs are highlighted in Fig. 5. It can be observed that 3 target proteins for Ritonavir, 2 target protein hits for each of Lopinavir, Darunavir, and Azithromycin, and 1 target protein hit for Remdesivir and Favipiravir. Out of these protein targets, ACE2 is the most important one since it is considered one of the crucial receptors of humans for nCoV to transmit infection deep inside the human cell [75–77]. Based on this validation, further research is conducted along with drug repurposing study, docking study, and COVID-19 symptoms-based analysis in our next research work [78] which helps us to identify a possible potential drug for COVID-19 named Fostamatinib [79–81]. Clinical studies involving Fostamatinib are also in progress [82,83]. Though the research is at the initial level, yet it somehow supports our research findings to some extent.

5. Conclusion

In any host-pathogen interaction network, the identification of spreader nodes is crucial for disease prognosis. However, not every protein in an interaction network has an intense disease-spreading capability. In this work, we have used the SARS-CoV-Human PPIN network and the spreader nodes at both level-1 and level-2 using the SIS model. These spreader nodes are considered for computing the protein interaction affinity score to unmask the level-1 human spreaders of nCoV. In addition, GO annotations have also been considered along with PPIN properties to make this model more effective and significant. With the gradual progress of the work, it has been observed that the selected
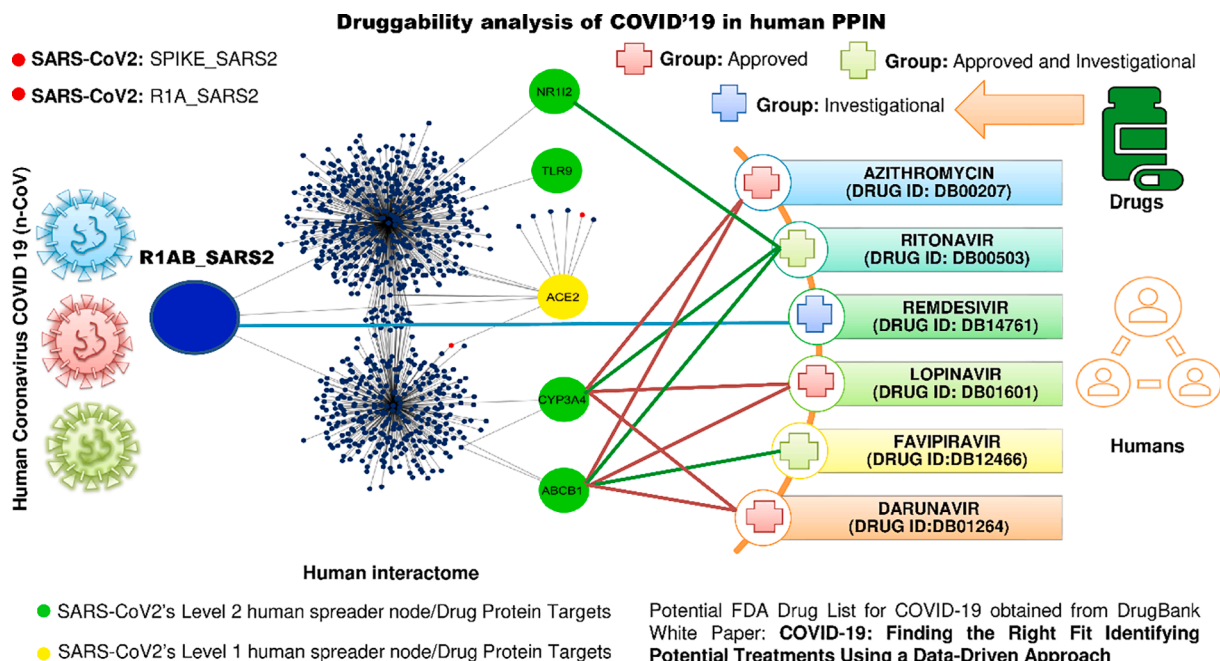
Fig. 5. Validation of our developed computational model with respect to the target proteins of the FDA accepted drugs for COVID-19 treatment. Yellow- and green-colored nodes denote level-1 and level-2 human spreaders of nCoV, which acts as the drug-protein targets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

human spreader nodes, identified by our proposed model, emerge as the potential protein targets of the FDA-approved drugs for COVID-19.

The primary hypotheses of the work may be listed as follows: 1) There is a genetic overlap of ~89% [84] between SARS-CoV and SARS-CoV2, which also lead to a significant overlap in spreader proteins between human-SARS-COV and human-SARSCOV2 protein-interaction networks. 2) *Fuzzy* PPI approach can assess protein interaction affinities at very high specificity with respect to benchmark datasets, as shown in Fig. 6. High specificity signifies a meager false-positive rate at a given threshold. Thus, at a 0.4 threshold (~99:9% specificity), the proposed model evaluates high-quality positive interactions in Human-nCoV

PPIN.

Finally, we propose that the developed computational model effectively identifies Human-nCoV PPIs with high specificity. The nCoV-Human interactions are inferred from another pandemic initiator SARS-CoV, which is highly genetically similar to nCoV. We also recognize the spreadability index of the human spreader proteins (up to level-2), validated through the SIS model. Due to high network density in human interaction networks, the number of proteins increases with the transition from one level to another. So, our proposed model can also identify human spreader proteins in level-2 by using the spreadability index validated by the SIS model.
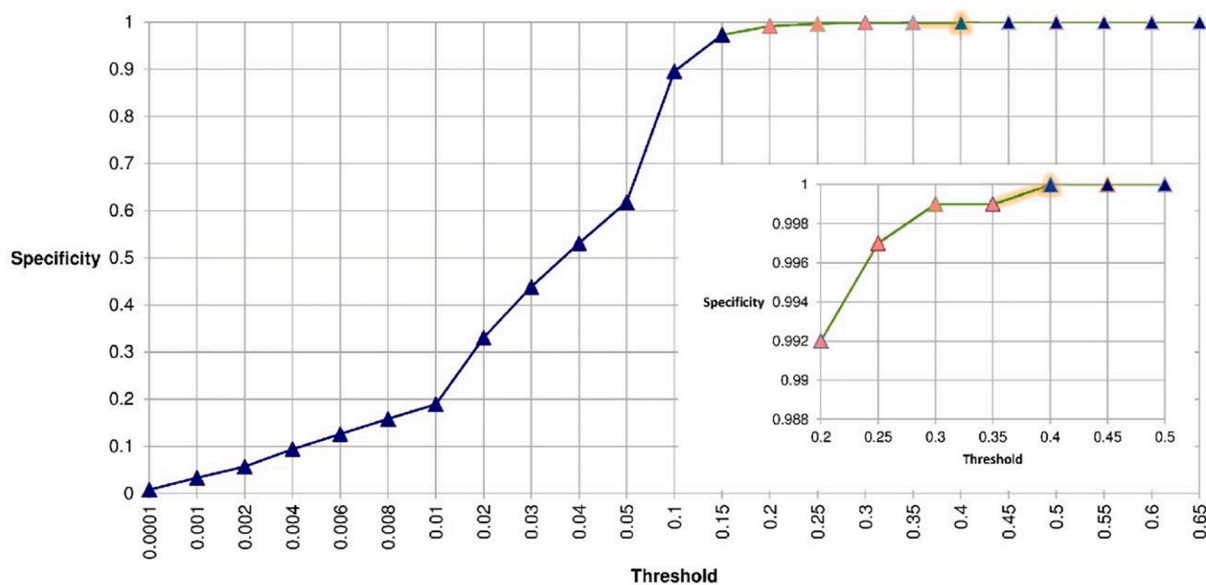

Fig. 6. Specificity at different threshold (x-axis) of binding affinity obtained from Fuzzy PPI model for complete human proteome interaction network. At 0.2 onward threshold, it produces high specificity with respect to benchmark positive and negative interaction data. High-quality interactions are extracted at a 0.4 threshold with ~99:9% specificity.

Our proposed method has identified the ACE2 and TMPRSS2 as an interactor of SARS-CoV2 proteins, which is essential for entry into the human host. SARS-CoV2 interacts with the SARS-CoV entry receptor ACE2 as SARS-CoV2 preserves those amino acid residues of SARS-CoV that are essential for ACE2 binding [85]. However, the binding strength of SARS-CoV2 with ACE2 is 10 to 20 times more than the SARS-CoV2-ACE2 attachment [86]. This is because several changes occur in the receptor-binding domains (RBDs) of SARS-CoV2 spike protein [87]. In addition, the cellular serine protease TMPRSS2 primes SARS-CoV2 for host entry, and a Serine protease inhibitor blocks SARS-CoV2 infection of lung cells [85,88]. Thus, TMPRSS2 activity is essential for viral spread and pathogenesis in the infected host [85,89].

In a recent study [90], Gordon *et al.* have identified 332 high-confidence SARS-CoV2-human protein–protein interactions where they have worked on the sequence analysis of SARS-CoV2 isolates. They cloned, tagged, and expressed 26 of the 29 SARS-CoV2 proteins in human cells and identified the human proteins that were physically associated with each using affinity-purification mass spectrometry (AP-MS). However, while comparing their seminal work with ours, we found that the SARS-CoV2 protein sequences used by Gordon *et al.* do not map directly with the available UniProt accession ids. In our case, we have worked only on the UniProt listed SARS-CoV2 proteins and applied a mathematical model of binding affinity assessment on a subset of UniProt listed reviewed Human proteins. Therefore, direct comparison and validation could not be possible with respect to Gordon *et al.*, primarily because of the unavailability of direct mapping of SARS-CoV2 proteins into corresponding UniProt accession ids. However, an attempt has been made to map UniProt ids of SARS-CoV2 proteins of Gordon *et al.*, from COVID-19 UniProtKB reference data [55] (please see Table S8 in the supplementary document). It is clear from Table S8 in the supplementary document that though UniProt ids are available for some of them but GO annotations for most of them are missing. Another interesting observation is that the entries marked in green have been also taken into consideration in this research work as well. It should be noted here that the current work depends heavily on the underlying GO Network of the host-pathogen PPIN. As evident from Table S8, GO annotations are often missing in the new protein list. Therefore we are working on a new strategy for the computational prediction of GO annotations for the set of proteins [16–19] in the Gordon's list as well new mutant variants.

One of the key highlights of our study may be underlined by the fact that the target proteins of the potential FDA drugs for COVID-19 overlap with the spreader nodes of the proposed nCoV-Human protein interaction network. Target proteins of six potential FDA drugs: Lopinavir [56], Ritonavir [57], Azithromycin [58], Remdesivir [59–61], Favipiravir [62,63], and Darunavir [64] for COVID-19 as mentioned in the Drug-Bank white paper [37] overlap with the spreader nodes of the proposed in silico nCoV-Human protein interaction model (see Fig. 5). Though clinical trials for the COVID-19 vaccine are on their way to date, three out of the six repurposed drugs, *i.e.,* Remdesivir [91] and Favipiravir [92] are found to be the most promising as well as effective ones. Our proposed model successfully identified their protein targets R1AB SARS2, TLR9, ACE2, CYP3A4, and ABCB1 as spreader nodes. This assessment reveals the fact that these spreader nodes indeed have biological relevance relative to disease propagation. It also motivates us to further do a drug repurposing study on the generated SARS-CoV2-human PPIN in our subsequent research work [78], which highlights that the drug Fostamatinib/R406 might be one of the potential drugs to be used for SARS-CoV2.

CRediT authorship contribution statement

Sovan Saha: Conceptualization, Data curation, Methodology, Writing – original draft, Software. **Anup Kumar Halder:** Conceptualization, Data curation, Methodology, Writing – original draft, Software. **Soumyendu Sekhar Bandyopadhyay:** Data curation, Writing – original draft, Visualization. **Piyali Chatterjee:** Supervision, Investigation, Formal analysis, Writing – review & editing. **Mita Nasipuri:** Supervision, Project administration, Investigation, Formal analysis, Writing – review & editing. **Subhadip Basu:** Supervision, Project administration, Investigation, Formal analysis, Data curation, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ymeth.2021.12.003.

References

[1] C. Wang, P.W. Horby, F.G. Hayden, G.F. Gao, A novel coronavirus outbreak of global health concern, Lancet 395 (2020) 470–473.
[2] World-Health-Organization Coronavirus disease (COVID-19) outbreak, 2020. https://www.who.int/emergencies/diseases/novel-coronavirus-2019. (Accessed 28-02-2021 Access 2020).
[3] World Map | CDC, 2020. https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/world-map.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Flocations-confirmed-cases.html. (Accessed 28-02-2020 Access 2020).
[4] Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV), 2020. https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov). (Accessed 28-02-2020 Access 2020).
[5] Statement on the meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus 2019 (n-CoV) on 23 January 2020, 2020. https://www.who.int/news-room/detail/23-01-2020-statement-on-the-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov). (Accessed 28-02-2020 Access 2020).
[6] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, Lancet 395 (10223) (2020) 497–506.
[7] D.L. Heymann, Data sharing and outbreaks: best practice exemplified, The Lancet 395 (2020) 469–470.
[8] S. Ruan, Likelihood of survival of coronavirus disease 2019, Lancet. Infect. Dis 20 (6) (2020) 630–631.
[9] WHO | Middle East respiratory syndrome coronavirus (MERS-CoV), 2020. https://www.who.int/emergencies/mers-cov/en/. (Accessed 28-02-2020 Access 2020).
[10] Y. Chen, Q. Liu, D. Guo, Emerging coronaviruses: genome structure, replication, and pathogenesis, J. Med. Virol. 92 (4) (2020) 418–423.
[11] J. Zhong, C. Tang, W. Peng, M. Xie, Y. Sun, Q. Tang, Q. Xiao, J. Yang, A novel essential protein identification method based on PPI networks and gene expression data, BMC Bioinf. 22 (1) (2021) 248.
[12] X. He, L. Kuang, Z. Chen, Y. Tan, L. Wang, Method for identifying essential proteins by key features of proteins in a novel protein-domain, Network 12 (2021) 1081.

[13] S. Saha, K. Sengupta, P. Chatterjee, S. Basu, M. Nasipuri, Analysis of protein targets in pathogen–host interaction in infectious diseases: a case study on Plasmodium falciparum and Homo sapiens interaction network, Brief. Funct. Genomics 17 (6) (2017) 441–450.

[14] S. Saha, P. Chatterjee, S. Basu, M. Nasipuri, Detection of spreader nodes and ranking of interacting edges in Human-SARS-CoV protein interaction network, PeerJ 9 (2021) 2020.04.12.038216-2020.04.12.038216.

[15] S.N. Basak, A.K. Biswas, S. Saha, P. Chatterjee, S. Basu, M. Nasipuri, Target protein function prediction by identification of essential proteins in protein-protein interaction network, Comput. Intelligence, Commun., Bus. Anal. (2019) 219–231.

[16] S. Saha, P. Chatterjee, S. Basu, M. Nasipuri, D. Plewczynski, Funpred 3.0: Improved protein function prediction using protein interaction network, PeerJ 2019(5) (2019) e6830.

[17] S. Saha, A. Prasad, P. Chatterjee, S. Basu, M. Nasipuri, Protein function prediction from dynamic protein interaction network using gene expression data, J. Bioinf. Comput. Biol. 17 (4) (2019).

[18] S. Saha, A. Prasad, P. Chatterjee, S. Basu, M. Nasipuri, Protein function prediction from protein-protein interaction network using gene ontology based neighborhood analysis and physico-chemical features, 16.

[19] S. Saha, A. Prasad, P. Chatterjee, S. Basu, M. Nasipuri, Modified FPred-Apriori: improving function prediction of target proteins from essential neighbours by finding their association with relevant functional groups using Apriori algorithm, Int. J. Adv. Intelligence Paradigms 19 (1) (2021) 61–83.

[20] E. Yeger-Lotem, R. Sharan, Human protein interaction networks across tissues and diseases, 6(257) (2015).

[21] M.G. Kann, Protein interactions and disease: computational approaches to uncover the etiology of diseases, Briefings Bioinf. 8 (5) (2007) 333–346.

[22] T. Ideker, R. Sharan, Protein networks in disease, Genome Res. 18(1088-9051 (Print)) 644-652.

[23] M.D. Dyer, T.M. Murali, B.W. Sobral, Computational prediction of host-pathogen protein-protein interactions, Bioinformatics (Oxford, England) 23 (13) (2007) i159–i166.

[24] P. Dutta, A.K. Halder, S. Basu, M. Kundu, A survey on Ebola genome and current trends in computational research on the Ebola virus, Brief Funct. Genomics 17 (6) (2018) 374–380.

[25] M.D. Dyer, T.M. Murali, B.W. Sobral, Supervised learning and prediction of physical interactions between human and HIV proteins, Infection, Genetics Evolut. 11 (5) (2011) 917–923.

[26] A.K. Halder, P. Dutta, M. Kundu, S. Basu, M. Nasipuri, Review of computational methods for virus-host protein interaction prediction: a case study on novel Ebola-human interactions, Brief. Funct. Genomics 17 (6) (2018) 381–391.

[27] A.K. Halder, S.S. Bandyopadhyay, P. Chatterjee, M. Nasipuri, D. Plewczynski, S. Basu, JUPPI: A multi-level feature based method for PPI prediction and a refined strategy for performance assessment, IEEE/ACM Trans. Comput. Biol. Bioinf. (2020) 1-1.

[28] China releases genetic data on new coronavirus, now deadly | CIDRAP, 2020. https://www.cidrap.umn.edu/news-perspective/2020/01/china-releases-genetic-data-new-coronavirus-now-deadly. (Accessed 28-02-2020 Access 2020).

[29] J.F.W. Chan, K.H. Kok, Z. Zhu, H. Chu, K.K.W. To, S. Yuan, K.Y. Yuen, Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan, Emerging Microbes Infect. 9 (1) (2020) 221–236.

[30] S. Pfefferle, J. Schöpf, M. Kögl, C.C. Friedel, M.A. Müller, J. Carbajo-Lozoya, T. Stellberger, E. von Dall'Armi, P. Herzog, S. Kallies, D. Niemeyer, V. Ditt, T. Kuri, R. Züst, K. Pumpor, R. Hilgenfeld, F. Schwarz, R. Zimmer, I. Steffen, F. Weber, V. Thiel, G. Herrler, H.J. Thiel, C. Schwegmann-Weßels, S. Pöhlmann, J. Haas, C. Drosten, A. von Brunn, The SARS-Coronavirus-host interactome: identification of cyclophilins as target for pan-Coronavirus inhibitors, PLoS Pathogens 7 (10) (2011).

[31] A. von Brunn, C. Teepe, J.C. Simpson, R. Pepperkok, C.C. Friedel, R. Zimmer, R. Roberts, R. Baric, J. Haas, Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFeome, PLoS One 2 (5) (2007).

[32] A. Bateman, M.J. Martin, C. O'Donovan, M. Magrane, E. Alpi, R. Antunes, B. Bely, M. Bingley, C. Bonilla, R. Britto, B. Bursteinas, H. Bye-Ajee, A. Cowley, A. Da Silva, M. De Giorgi, T. Dogan, F. Fazzini, L.G. Castro, L. Figueira, P. Garmiri, G. Georghiou, D. Gonzalez, E. Hatton-Ellis, W. Li, W. Liu, R. Lopez, J. Luo, Y. Lussi, A. MacDougall, A. Nightingale, B. Palka, K. Pichler, D. Poggioli, S. Pundir, L. Pureza, Q. Qi, S. Rosanoff, R. Saidi, T. Sawford, A. Shypitsyna, E. Speretta, E. Turner, N. Tyagi, V. Volynkin, T. Wardell, K. Warner, X. Watkins, R. Zaru, H. Zellner, I. Xenarios, L. Bougueleret, A. Bridge, S. Poux, N. Redaschi, L. Aimo, G. ArgoudPuy, A. Auchincloss, K. Axelsen, P. Bansal, D. Baratin, M.C. Blatter, B. Boeckmann, J. Bolleman, E. Boutet, L. Breuza, C. Casal-Casas, E. De Castro, E. Coudert, B. Cuche, M. Doche, D. Dornevil, S. Duvaud, A. Estreicher, L. Famiglietti, M. Feuermann, E. Gasteiger, S. Gehant, V. Gerritsen, A. Gos, N. Gruaz-Gumowski, U. Hinz, C. Hulo, F. Jungo, G. Keller, V. Lara, P. Lemercier, D. Lieberherr, T. Lombardot, X. Martin, P. Masson, A. Morgat, T. Neto, N. Nouspikel, S. Paesano, I. Pedruzzi, S. Pilbout, M. Pozzato, M. Pruess, C. Rivoire, B. Roechert, M. Schneider, C. Sigrist, K. Sonesson, S. Staehli, A. Stutz, S. Sundaram, M. Tognolli, L. Verbregue, A.L. Veuthey, C.H. Wu, C.N. Arighi, L. Arminski, C. Chen, Y. Chen, J.S. Garavelli, H. Huang, K. Laiho, P. McGarvey, D.A. Natale, K. Ross, C.R. Vinayaka, Q. Wang, Y. Wang, L.S. Yeh, J. Zhang, UniProt: The universal protein knowledgebase, Nucleic Acids Res. 45 (D1) (2017) D158–D169.

[33] G.O. Consortium, others, The Gene Ontology (GO) database and informatics resource, Nucleic Acids Res 32 (suppl 1) (2004) D258–D261.

[34] P. Dutta, S. Basu, M. Kundu, Assessment of semantic similarity between proteins using information content and topological properties of the gene ontology graph, IEEE/ACM Trans. Comput. Biol. Bioinf. 15 (3) (2018) 839–849.

[35] N.T.J. Bailey, The mathematical theory of infectious diseases and its applications, 2nd ed., Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13, 6LE.1975..

[36] L. Chin, J. Cox, S. Esmail, M. Franklin, D. Le, COVID-19 : Finding the Right Fit Identifying Potential Treatments Using a Data-Driven Approach, Drugbank White Paper (2020).

[37] N. Samadi, A. Bouyer, Identifying influential spreaders based on edge ratio and neighborhood diversity measures in complex networks, Computing 101 (8) (2019) 1147–1175.

[38] S. Wang, F. Wu, Detecting overlapping protein complexes in PPI networks based on robustness, Proteome Sci. 11 (Suppl 1) (2013) 1–8.

[39] J.M. Anthonisse, The rush in a directed graph, Stichting Mathematisch Centrum (1971).

[40] G. Sabidussi, The centrality index of a graph, Psychometrika 31 (4) (1966) 581–603.

[41] H. Jeong, S.P. Mason, A.L. Barabási, Z.N. Oltvai, Lethality and centrality in protein networks, Nature 411 (6833) (2001) 41–42.

[42] M. Li, J. Wang, X. Chen, H. Wang, Y. Pan, A local average connectivity-based method for identifying essential proteins from the network level, Comput. Biol. Chem. 35 (3) (2011) 143–150.

[43] Y. Tang, M. Li, J. Wang, Y. Pan, F.-X. Wu, CytoNCA: A cytoscape plugin for centrality analysis and evaluation of protein interaction networks, Biosystems 127 (2015) 67–72.

[44] F.M. Couto, M.J. Silva, P.M. Coutinho, Semantic similarity over the gene ontology: family correlation and selecting disjunctive ancestors, CIKM '05, in: Proceedings of the 14th ACM international conference on Information and knowledge management, 2005, pp. 343–344.

[45] F.M. Couto, M.J. Silva, P.M. Coutinho, Measuring semantic similarity between Gene Ontology terms, Data Knowl. Eng. 61 (1) (2007) 137–152.

[46] P. Resnik, Using information content to evaluate semantic similarity in a taxonomy, Proceedings of the 14th international joint conference on Artificial intelligence – Volume 1 (1995) 448–453.

[47] S. Jain, G.D. Bader, An improved method for scoring protein-protein interactions using semantic similarity within the gene ontology, BMC Bioinf. 11 (1) (2010) 562.

[48] D. Lin, An Information-Theoretic Definition of Similarity, Proceedings of the Fifteenth International Conference on Machine Learning (1998) 296–304.

[49] D. Lin, An Information-Theoretic Definition of Similarity, Proceedings of the Fifteenth International Conference on Machine Learning (1998) 296–304.

[50] J.J. Jiang, D.W. Conrath, Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy, Proceedings of the 10th Research on Computational Linguistics International Conference (1997) 19–33.

[51] M. Ashburner, Gene ontology: tool for the unification of biology, Nat. Genet. 25 (1) (2000) 25–29.

[52] C.E. Shannon, A mathematical theory of communication, ACM SIGMOBILE Mobile Computing and Communications Review 5(1) (2001) 3–55.

[53] M. Agrawal, M. Zitnik, J. Leskovec, Large-Scale Analysis of Disease Pathways in the Human Interactome, bioRxiv (2017) 189787–189787.

[54] BioSNAP: Network datasets: Human protein-protein interaction network, 2021. https://snap.stanford.edu/biodata/datasets/10000/10000-PP-Pathways.html. (Accessed 28-02-2021 Access 2021).

[55] COVID-19 UniProtKB 2021. https://covid-19.uniprot.org/. (Accessed 15-04-2021 2021).

[56] C. Harrison, Coronavirus puts drug repurposing on the fast track, Nat. Biotechnol. (2020).

[57] B. Cao, Y. Wang, D. Wen, W. Liu, J. Wang, G. Fan, L. Ruan, B. Song, Y. Cai, M. Wei, X. Li, J. Xia, N. Chen, J. Xiang, T. Yu, T. Bai, X. Xie, L. Zhang, C. Li, Y. Yuan, H. Chen, H. Li, H. Huang, S. Tu, F. Gong, Y. Liu, Y. Wei, C. Dong, F. Zhou, X. Gu, J. Xu, Z. Liu, Y. Zhang, H. Li, L. Shang, K. Wang, K. Li, X. Zhou, X. Dong, Z. Qu, S. Lu, X. Hu, S. Ruan, S. Luo, J. Wu, L. Peng, F. Cheng, L. Pan, J. Zou, C. Jia, J. Wang, X. Liu, S. Wang, X. Wu, Q. Ge, J. He, H. Zhan, F. Qiu, L. Guo, C. Huang, T. Jaki, F.G. Hayden, P.W. Horby, D. Zhang, C. Wang, A Trial of Lopinavir–Ritonavir in Adults Hospitalized with Severe Covid-19, New England J. Med. (2020).

[58] P. Gautret, J.-C. Lagier, P. Parola, V.T. Hoang, L. Meddeb, M. Mailhe, B. Doudier, J. Courjon, V. Giordanengo, V.E. Vieira, H.T. Dupont, S. Honoré, P. Colson, E. Chabrière, B. La Scola, J.-M. Rolain, P. Brouqui, D. Raoult, Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial, Int. J. Antimicrobial Agents (2020) 105949–105949.

[59] E. de Wit, F. Feldmann, J. Cronin, R. Jordan, A. Okumura, T. Thomas, D. Scott, T. Cihlar, H. Feldmann, Prophylactic and therapeutic remdesivir (GS-5734) treatment in the rhesus macaque model of MERS-CoV infection, PNAS 117 (12) (2020) 6771–6776.

[60] Emergency Access to Remdesivir Outside of Clinical Trials, 2021. https://www.gilead.com/purpose/advancing-global-health/covid-19/emergency-access-to-remdesivir-outside-of-clinical-trials. (Accessed 28-02-2021 Access 2021).

[61] Remdesivir Clinical Trials, 2021. https://www.gilead.com/purpose/advancing-global-health/covid-19/remdesivir-clinical-trials. (Accessed 28-02-2021 Access 2021).

[62] China approves antiviral favilavir to treat coronavirus - UPI.com, 2021. https://www.upi.com/Health_News/2020/02/17/China-approves-antiviral-favilavir-to-treat-coronavirus/5291581953892/. (Accessed 28-02-2021 Access 2021).

[63] Taiwan synthesizes anti-viral drug favilavir for COVID-19 patients - Focus Taiwan, 2021. https://focustaiwan.tw/sci-tech/202003020012. (Accessed 28-02-2021 Access 2021).

[64] Efficacy and Safety of Darunavir and Cobicistat for Treatment of COVID-19 - Full Text View - ClinicalTrials.gov, 2021. https://clinicaltrials.gov/ct2/show/NCT04252274. (Accessed 28-02-2021 Access 2021).

[65] D.S. Wishart, C. Knox, A.C. Guo, D. Cheng, S. Shrivastava, D. Tzur, B. Gautam, M. Hassanali, DrugBank: a knowledgebase for drugs, drug actions and drug targets, Nucleic acids research 36(Database issue) (2008) D901-D906.

[66] M.H. Schaefer, A. Fontaine Jf Fau – Vinayagam, P. Vinayagam A Fau - Porras, E.E. Porras P Fau - Wanker, M.A. Wanker Ee Fau – Andrade-Navarro, M.A. Andrade-Navarro, HIPPIE: Integrating protein interaction networks with experiment based quality scores, (1932-6203 (Electronic)).

[67] D. Szklarczyk, A.L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N.T. Doncheva, J.H. Morris, P. Bork, L.J. Jensen, Christian v. Mering, STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, Nucleic Acids Res. 47 (D1) (2018) D607–D613.

[68] R. Oughtred, C. Stark, B.J. Breitkreutz, J. Rust, L. Boucher, C. Chang, N. Kolas, L. O'Donnell, G. Leung, R. McAdam, F. Zhang, S. Dolma, A. Willems, J. Coulombe-Huntington, A. Chatr-Aryamontri, K. Dolinski, M. Tyers, The BioGRID interaction database: 2019 update, (1362-4962 (Electronic)).

[69] I. Xenarios, L. Rice Dw Fau - Salwinski, M.K. Salwinski L Fau - Baron, E.M. Baron Mk Fau - Marcotte, D. Marcotte Em Fau - Eisenberg, D. Eisenberg, DIP: the database of interacting proteins, (0305-1048 (Print)).

[70] K. Luck, D.-K. Kim, L. Lambourne, K. Spirohn, B.E. Begg, W. Bian, R. Brignall, T. Cafarelli, F.J. Campos-Laborie, B. Charloteaux, D. Choi, A.G. Coté, M. Daley, S. Deimling, A. Desbuleux, A. Dricot, M. Gebbia, M.F. Hardy, N. Kishore, J.J. Knapp, I.A. Kovács, I. Lemmens, M.W. Mee, J.C. Mellor, C. Pollis, C. Pons, A.D. Richardson, S. Schlabach, B. Teeking, A. Yadav, M. Babor, D. Balcha, O. Basha, C. Bowman-Colin, S.-F. Chin, S.G. Choi, C. Colabella, G. Coppin, C. D'Amata, D. De Ridder, S. De Rouck, M. Duran-Frigola, H. Ennajdaoui, F. Goebels, L. Goehring, A. Gopal, G. Haddad, E. Hatchi, M. Helmy, Y. Jacob, Y. Kassa, S. Landini, R. Li, N. van Lieshout, A. MacWilliams, D. Markey, J.N. Paulson, S. Rangarajan, J. Rasla, A. Rayhan, T. Rolland, A. San-Miguel, Y. Shen, D. Sheykhkarimli, G.M. Sheynkman, E. Simonovsky, M. Taşan, A. Tejeda, V. Tropepe, J.-C. Twizere, Y. Wang, R.J. Weatheritt, J. Weile, Y. Xia, X. Yang, E. Yeger-Lotem, Q. Zhong, P. Aloy, G.D. Bader, J. De Las Rivas, S. Gaudet, T. Hao, J. Rak, J. Tavernier, D.E. Hill, M. Vidal, F. P. Roth, M.A. Calderwood, A reference map of the human binary protein interactome, Nature 580(7803) (2020) 402–408.

[71] P. Blohm, P. Frishman G Fau - Smialowski, F. Smialowski P Fau - Goebels, B. Goebels F Fau – Wachinger, A. Wachinger B Fau - Ruepp, D. Ruepp A Fau - Frishman, D. Frishman, Negatome 2.0: a database of non-interacting proteins derived by literature mining, manual annotation and protein structure analysis, (1362-4962 (Electronic)).

[72] L.G. Trabuco, R.B. Betts Mj Fau - Russell, R.B. Russell, Negative protein-protein interaction datasets derived from large-scale two-hybrid experiments, (1095-9130 (Electronic)).

[73] Advanced Search - DrugBank, 2021. https://www.drugbank.ca/unearth/advanced/drugs. (Accessed 28-02-2021 Access 2021).

[74] DrugBank, 2021. https://www.drugbank.ca/. (Accessed 28-02-2021 Access 2021).

[75] J.J. Mourad, B.I. Levy, Interaction between RAAS inhibitors and ACE2 in the context of COVID-19, Nature Reviews Cardiology 17 (2020) 313-313.

[76] ACE-2 is shown to be the entry receptor for SARS-CoV-2: R&D Systems, 2021. https://www.rndsystems.com/resources/articles/ace-2-sars-receptor-identified. (Accessed 28-02-2021 Access 2021).

[77] A.B. Patel, A. Verma, COVID-19 and angiotensin-converting enzyme inhibitors and angiotensin receptor blockers: what is the evidence? Am. Med. Associat. (2020).

[78] S. Saha, A.K. Halder, S.S. Bandyopadhyay, P. Chatterjee, M. Nasipuri, D. Bose, S. Basu, Drug repurposing for COVID-19 using computational screening: is fostamatinib/R406 a potential candidate? Methods (San Diego, Calif.) 21 (2021) https://doi.org/10.1016/j.ymeth.2021.08.007 (in this issue).

[79] Drug Approval Package: TAVALISSE (fostamatinib disodium hexahydrate), 2020. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2018/209299Orig1s000TOC.cfm. (Accessed 20-06-2020 Access 2020).

[80] Fostamatinib - DrugBank, 2020. https://www.drugbank.ca/drugs/DB12010. (Accessed 26-08-2020 Access 2020).

[81] FDA approves fostamatinib tablets for ITP | FDA, 2020. https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-fostamatinib-tablets-itp. (Accessed 26-08-2020 Access 2020).

[82] Positive Topline Data Shows Fostamatinib Meets Primary Endpoint of Safety in Phase 2 Clinical Trial in Hospitalized Patients with COVID-19, 2021. https://www.rigel.com/investors/news-events/press-releases/detail/312/positive-topline-data-shows-fostamatinib-meets-primary. (Accessed 21-04-2021 Access 2021).

[83] Double-Blind, Randomized, Placebo-Controlled, Adaptive Design, Multi-Center Phase 3 Study to Evaluate the Efficacy and Safety of Fostamatinib in COVID-19 Subjects, 2020. https://clinicaltrials.gov/ct2/show/NCT04629703. (Accessed 20-12-2020 Access 2020).

[84] A.A.T. Naqvi, K. Fatima, T. Mohammad, U. Fatima, I.K. Singh, A. Singh, S.M. Atif, G. Hariprasad, G.M. Hasan, M.I. Hassan, Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach, Biochim Biophys Acta Mol Basis Dis 1866(10) (2020) 165878–165878.

[85] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N.H. Wu, A. Nitsche, M.A. Müller, C. Drosten, S. Pöhlmann, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor, Cell 181(2) (2020) 271-280.e8.

[86] R. Yadav, J.K. Chaudhary, N. Jain, P.K. Chaudhary, S. Khanra, P. Dhamija, A. Sharma, A. Kumar, S. Handu, Role of structural and non-structural proteins and therapeutic targets of SARS-CoV-2 for COVID-19, Cells 10 (4) (2021).

[87] J. Shang, G. Ye, K. Shi, Y. Wan, C. Luo, H. Aihara, Q. Geng, A. Auerbach, F. Li, Structural basis of receptor recognition by SARS-CoV-2, Nature 581 (7807) (2020) 221–224.

[88] S. Matsuyama, N. Nagata, K. Shirato, M. Kawase, M. Takeda, F. Taguchi, Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2, J. Virol. 84 (24) (2010) 12658–12664.

[89] P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C. L. Huang, H.D. Chen, J. Chen, Y. Luo, H. Guo, R.D. Jiang, M.Q. Liu, Y. Chen, X. R. Shen, X. Wang, X.S. Zheng, K. Zhao, Q.J. Chen, F. Deng, L.L. Liu, B. Yan, F. X. Zhan, Y.Y. Wang, G.F. Xiao, Z.L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (7798) (2020) 270–273.

[90] D.E. Gordon, G.M. Jang, M. Bouhaddou, J. Xu, K. Obernier, K.M. White, M.J. O'Meara, V.V. Rezelj, J.Z. Guo, D.L. Swaney, T.A. Tummino, R. Hüttenhain, R.M. Kaake, A.L. Richards, B. Tutuncuoglu, H. Foussard, J. Batra, K. Haas, M. Modak, M. Kim, P. Haas, B.J. Polacco, H. Braberg, J.M. Fabius, M. Eckhardt, M. Soucheray, M. J. Bennett, M. Cakir, M.J. McGregor, Q. Li, B. Meyer, F. Roesch, T. Vallet, A. Mac Kain, L. Miorin, E. Moreno, Z.Z.C. Naing, Y. Zhou, S. Peng, Y. Shi, Z. Zhang, W. Shen, I.T. Kirby, J.E. Melnyk, J.S. Chorba, K. Lou, S.A. Dai, I. Barrio-Hernandez, D. Memon, C. Hernandez-Armenta, J. Lyu, C.J.P. Mathy, T. Perica, K.B. Pilla, S.J. Ganesan, D.J. Saltzberg, R. Rakesh, X. Liu, S.B. Rosenthal, L. Calviello, S. Venkataramanan, J. Liboy-Lugo, Y. Lin, X.-P. Huang, Y. Liu, S.A. Wankowicz, M. Bohn, M. Safari, F.S. Ugur, C. Koh, N.S. Savar, Q.D. Tran, D. Shengjuler, S.J. Fletcher, M.C. O'Neal, Y. Cai, J.C.J. Chang, D.J. Broadhurst, S. Klippsten, P.P. Sharp, N.A. Wenzell, D. Kuzuoglu-Ozturk, H.-Y. Wang, R. Trenker, J.M. Young, D. A. Cavero, J. Hiatt, T.L. Roth, U. Rathore, A. Subramanian, J. Noack, M. Hubert, R. M. Stroud, A.D. Frankel, O.S. Rosenberg, K.A. Verba, D.A. Agard, M. Ott, M. Emerman, N. Jura, M. von Zastrow, E. Verdin, A. Ashworth, O. Schwartz, C. d'Enfert, S. Mukherjee, M. Jacobson, H.S. Malik, D.G. Fujimori, T. Ideker, C.S. Craik, S.N. Floor, J.S. Fraser, J.D. Gross, A. Sali, B.L. Roth, D. Ruggero, J. Taunton, T. Kortemme, P. Beltrao, M. Vignuzzi, A. García-Sastre, K.M. Shokat, B.K. Shoichet, N.J. Krogan, A SARS-CoV-2 protein interaction map reveals targets for drug repurposing, Nature 583(7816) (2020) 459–468.

[91] Trial shows Covid-19 patients recover with Gilead's remdesivir, 2021. https://www.clinicaltrialsarena.com/news/niaid-trial-remdesivir-covid-19-data/. (Accessed 28-02-2021 Access 2021).

[92] Clinical Trial of Favipiravir Tablets Combine With Chloroquine Phosphate in the Treatment of Novel Coronavirus Pneumonia - Full Text View - ClinicalTrials.gov, 2021. https://clinicaltrials.gov/ct2/show/NCT04319900. (Accessed 28-02-2021 Access 2021).