Ontology Modeling and Analysis of COVID-19 Associated Acute Kidney Injury and Its Underlying Molecular Mechanisms

Easheta Shah[1*], Roshan Desai[1*], Suyuan Peng[2,3], Luxia Zhang[2,4,5], and Yongqun He [1]

[1] *University of Michigan, Ann Arbor, MI, USA*
[2] *National Institute of Health Data Science, Peking University, No.38 Xueyuan Rd., Beijing, China*
[3] *School of Public Health, Peking University, No.38 Xueyuan Rd., Beijing, China*
[4] *Department of Medicine, Renal Division, Peking University First Hospital, Peking University Institute of Nephrology, No.8 Xishiku St., Beijing, China*
[5] *Advanced Institute of Information Technology, Peking University, Hangzhou, China*
* *Contributed equally.*

Abstract

Acute kidney injury (AKI) is found to be common among COVID-19 patients. In this study, we performed extensive literature mining and used the BioGRID COVID-19 interaction data to bridge the mechanistic and molecular link between COVID-19 and AKI. DAVID GO enrichment analysis of the BioGRID data allowed for further filtration of COVID-19 related interactors by their relevance to untoward kidney manifestations. Key physiological processes involved in this pathway include Renin-Angiotensin system (RAS) activation, complement activation, and most importantly, systemic inflammation. Discovered interactors like CD147, CD209, CypA, and MASP2 were found to be heavily implicated in the mentioned processes. The Coronavirus Infectious Disease Ontology (CIDO) was used to represent our analyzed results, leading to further understanding of the COVID-19 associated AKI mechanisms.

Keywords

COVID-19, Acute Kidney Injury (AKI), Coronavirus Infectious Disease Ontology (CIDO)

1. Introduction

The SARS-CoV-2 virus has many different phenotypic outcomes associated with the failure of organs and organ systems. This multi-organ effect heavily involves untoward manifestations in the kidney, ranging from mild proteinuria to progressive Acute kidney injury (AKI). It has been reported that over 30 percent of hospitalized patients in New York with COVID-19 developed AKI as defined by KDIGO criteria [1].

Many resources and ongoing research efforts are available to study SARS-CoV-2 pathogenesis, including thousands of peer-reviewed journal articles that have been published with experimentally verified scientific insights. Many databases have also been developed in this context. The BioGRID database has compiled many coronavirus proteins and their interactions with specific host molecules [2]. For SARS-CoV-2, BioGRID provides 32 viral proteins and the different host interactors each protein is involved with. Additionally, the Coronavirus Infectious Disease Ontology (CIDO) has been developed to represent an aggregate of experimentally verified knowledge [3]. CIDO provides important representations of coronavirus diseases with relation to their etiology, transmission, pathogenesis, host-coronavirus interactions, and vaccine treatments [3].

In this study, we hypothesize that AKI is manifested through molecular interactions between SARS-CoV-2 proteins and host proteins. BioGRID/literature mining and ontology modeling were used to explore the possible protein-protein interactions (PPIs) leading to AKI in COVID-19 patients.

International Conference on Biomedical Ontologies 2021, September 16–18, 2021, Bozen-Bolzano, Italy
EMAIL: shaheash@umich.edu (A. 1); roshand@umich.edu (A. 2); yongqunh@med.umich.edu (A. 3); peng.suyuan@bjmu.edu.cn (A. 4); luxia_zhang@163.com (A. 5)
ORCID: 0000-0001-9396-0796 (A. 1); 0000-0001-5153-0712 (A. 2); 0000-0001-9189-9661 (A. 3); 0000-0002-8221-7574 (A. 4); 0000-0003-2349-2936 (A. 5)

2. Methods

2.1. Literature Mining and Annotation

PubMed and PubMed Central were utilized in the search for peer-reviewed results in the context of SARS-CoV-2, AKI and host-coronavirus interactions. The resulting discoveries of kidneyrelated phenotypic outcomes were then collected in a highly specific Excel format in which each interaction cause and outcome were precisely categorized. Specific focuses were on S and N proteins and their interactions with host proteins. Specific pathways such as the Renin-Angiotensin system (RAS) and complement activation were systematically studied.

2.2. BioGRID Data Analysis

As of Jun 5, 2021, BioGRID collected 643 proteins interacting with SARS-CoV-2 S protein, and 345 proteins interacting with SARS-CoV-2 N protein. These interactors were downloaded and exposed for DAVID Gene Ontology (GO) enrichment analysis. The bioinformatics resource DAVID is used to create charts and lists of interactors and interactor functions for further functional analysis [4]. DAVID allows for additional filtration by relevance to the kidney. By inputting BioGRID's 20 viral interactors into DAVID, we were able to view specific interactors and determine which ones are related to the kidneys and AKI by searching for kidney-related terms like "renal failure" and "nephropathy".

2.3. Ontology Modeling

Each interaction pair was tagged by its reference, and all together, this data modeled the ontological structure required for infectious disease bioinformatics. These interactions were compared to rapidly developing and existing COVID-19 ontologies like BioGRID and CIDO. Protege-OWL editor [5] was used for CIDO modeling. The results are available at the CIDO GitHub website: https://github.com/CIDO-ontology/cido.

3. Results

3.1. BioGRID Data Analysis and Further Literature Analysis

BioGRID provides a list of known viral-host interactions for the SARS-CoV-2 virus. Each interactor is listed in accordance with its relation to the S or N viral protein and each description is supported by two sources regarding its specific coronavirus interaction and kidney manifestation. The provided table is a compilation of these findings from BioGRID, DAVID, and further literature analysis.

Table 1
Selected AKI-related host proteins that also interact with viral S/N proteins

Protein Name	Viral Protein	Description	PMID
ACE2	S	Primary S protein receptor at RBD site, involved in RAS activation, present on kidney cells	32264791 19020433
CD209 (DCSIGN)	S	Mediates ACE2-dependent viral infection, limited expression, initiation of inflammation	34015061 19273246
CD207	S	C-type lectin receptor binding to NTD site of S protein, associated with ESRD	15860669 27234728
CD147	S	AKI biomarker, kidney/distant organ crosstalk, associated with Lupus Nephritis	33277466 25248362
ASGR1	S	Binds to RBD and NTD sites of S protein, may mediate cell entry with KREMEN1	33883951 12119473

		and ACE2, human renal tubular epithelial cells	
KREMEN1	S	Involved in cell entry mediation with ASGR1 and ACE2	Ref. [6]
NRP1	S	SARS-CoV-2 entry mediator, expressed in olfactory epithelium and lung tissue of COVID-19 patients, involved in diabetic nephropathy	33082294 18974107
MASP2	N	Key enzyme involved in lectin pathway of complement activation when activated, heavily involved in renal ischemia	33671334 24868011
GSK3B	N	Binding activation provokes systemic inflammation, CRAD marker	32948440 18786229
PPIA (CyPa)	N	Renal inflammation associated with Lupus Nephritis	34109257 25580061

Abbreviations: CRAD: Chronic renal allograft dysfunction; End-Stage Renal Disease (ESRD); NTD: N terminal domain; RBD: Receptor binding domain.

3.2. Integrative Analysis of AKI Generation with Literature Report

Our literature search found two key phenotypic processes in circulation that are associated with COVID-19 related AKI: RAS activation and complement activation. Both processes are catalyzed by specific viral-host interactions, currently identified as follows: ACE2 and RAS activation and MASP2 and complement activation. These two processes can both induce inflammation.

The renin-angiotensin system (RAS) is one of the major control systems for blood pressure and fluid balance. It plays an important role in the physiological regulation of the kidneys, heart, and blood vessels. The activation of this system is a central part of many common pathological conditions, including hypertension, heart failure, and kidney disease. One of the enzymes in RAS, ACE2, has been identified as a cellular receptor for SARS-CoV-2. This pathway is activated by the binding of viral S protein and the ACE2 human cell receptor, the latter of which is present in various organs, including the kidney [7]. Viral entry may result in kidney involvement via systemic RAS activation and organ crosstalk. These untoward effects result in kidney injury when left untreated.

Another pathway to kidney injury is complement activation which is activated by binding of viral N protein and the MASP-2 cell receptor, a mannan-binding lectin protease. Complement activation is a signaling cascade event that can be activated by lectin binding, and it results in lung injury, which will indirectly affect the kidney, and endothelial injury and thrombotic microangiopathy which will directly affect the kidney [8, 9].

By incorporating literature review results with our BioGRID and DAVID filtration analysis, we were able to generate an integrative model of the AKI formation (Fig. 1). This model integrates systemic inflammation through the processes of complement activation and RAS activation. Out of the ten proteins in Table 1, three notable interactors were found to have strong relevance to the previously mentioned systemic processes and SARS-CoV-2 infection: CD147 (BSG) and CD209 (DC-SIGN), S protein interactors, and PPIA (CyPA), an N protein interactor (Table 1). Of note is the binding relationship between CD147 and CyPa, which manifests into Acute Kidney Injury independent of viral entry. This interaction manifestation may also be implicated with the SARS-CoV-2 virus, but this part of the pathway requires further investigation. While explicit mechanisms leading to AKI remain unclear, a large portion of the story has been uncovered by performing further literature studies on the articles regarding these three interactors and combining the information to formulate a cohesive conclusion. For DC-SIGN and BSG, in their respective sources, there was a strong correlation between the molecular presence and the likelihood of AKI [10, 11]. Furthermore, all three interactors noted are heavily involved in stimulating an inflammatory response, leading to the understanding that inflammation is a common connection leading from Viral Infection to AKI.

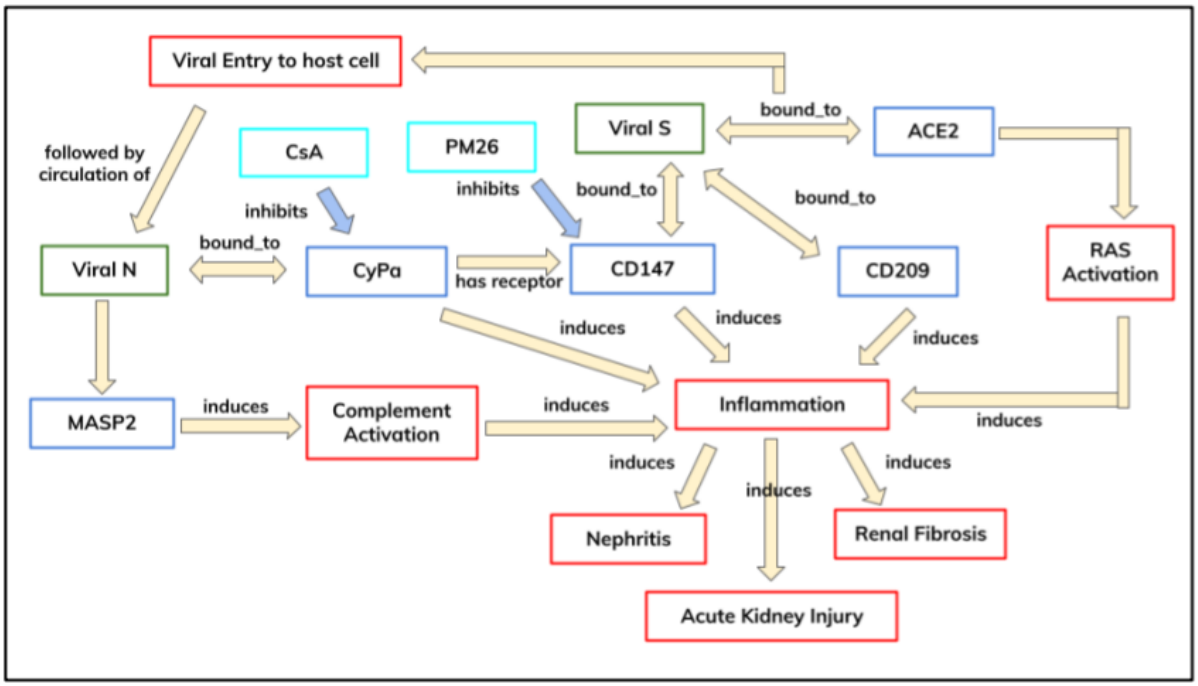

Figure 1: Summarized model of SARS-CoV-2 viral infection leading to AKI. Visualization of findings from literature review and BioGRID analysis with concepts in boxes (red for physiological processes, blue for host proteins, and green for viral proteins) and connectors like "bound_to" and "induces." Cyclosporine A (CsA) and Polyman26 (i.e., PM26) are two drugs targeting CyPa and CD147, respectively.

3.3. CIDO Ontology Modeling of COVID-19 Associated AKI

The terms and relations identified above (Table 1 and Fig. 1) were also represented in the CIDO. After importing all the proteins defined in Table 1 from the Protein Ontology (PRO) [12], we generated new axioms to define the relations between host and viral proteins as demonstrated below:

- ***'S protein of SARS-CoV-2'***: *'capable of binding to' some 'basigin (human)'*
- ***'nucleoprotein (SARS-CoV-2)'***: *'capable of binding to' some 'peptidyl-prolyl cis-trans isomerase A (human)'*
- ***'peptidyl-prolyl cis-trans isomerase A (human)'***: *'has receptor' some 'basigin (human)'*
- ***'cyclosporin A'***: *'chemical has protein target as inhibitor' some 'peptidyl-prolyl cis-trans isomerase A (human)'*
- ***Polyman26***: *'chemical has protein target as antagonist' some 'CD209 antigen (human)'*

Here we needed to define the meanings of the two relations 'capable of binding to' and 'has receptor,' which are not available in Relation Ontology (RO). While the binding of a certain viral protein to some host interactor can induce viral infection, this process will not always occur under varying conditions. Thus, we must use a term like 'capable of binding to' rather than a relation term like bound_to (http://purl.obolibrary.org/obo/OBI_1110119), the latter of which asserts that such binding is always there. Similarly, the binding between CD147 and CyPA is a process that may occur under specific conditions independent of SARS-CoV-2 viral infection. Therefore, we use the related term *'has receptor.'*

4. Discussion

The contributions of this study are multifold. Firstly, by BioGRID data analysis and literature annotation, we identified many human proteins such as CD147, CD209, and CyPA that are capable of

interacting with viral S or N protein and likely contribute to COVID-19-associated AKI induction. Secondly, through literature mining, we identified several important processes, including systemic inflammation, RAS activation, and complement activation, which likely contribute to AKI formation in COVID-19 patients. Thirdly, we were able to generate a new integrative model that explains the COVID-19 associated AKI generation by linking our newly identified proteins and biological processes. Lastly, we modeled and represented the knowledge identified in this study in the CIDO ontology.

While the BioGRID database provides a large number of host-coronavirus interactions, the knowledge does not directly translate to our understanding of COVID-19 associated AKI generation. This is due to the high throughput screening of the present BioGRID data. Our approach strategically deciphers through new knowledge from BioGRID data to support developing hypotheses about this pathogenesis. The hypotheses generated in this study are worth experimental verification.

Ontologies have played significant roles in our study. We used the Gene Ontology (GO) enrichment analysis through the DAVID tool. Later, we used CIDO for our ontology modeling of the host-coronavirus interactions to better understand how AKI occurs in COVID-19 patients. The logical axioms added to CIDO in this project begin to represent the knowledge of COVID-19 related AKI manifestation, but further axioms can be developed to link more downstream processes (e.g., systemic inflammation) of this pathogenesis.

More work is currently underway. Some COVID-19 patients experience early AKI symptoms (AKI onset before multiorgan dysfunction) while other patients do not. Previously, we found that this AKI generation might be due to direct viral infection or indirect influences from the viral infection of other organs [13]. It is hypothesized that early AKI is a result of direct viral infection to the kidney. We are currently extending our current methods to better address the hypothesis.

5. Acknowledgements

This study was supported by a grant to LZ and YH from Michigan Medicine–Peking University Health Sciences Center Joint Institute for Clinical and Translational Research (71017Y2027) and PKUBaidu Fund (2020BD032) to SP.

6. References

[1] J. S. Hirsch, J. H. Ng, D. W. Ross, P. Sharma, H. H. Shah, R. L. Barnett, et al., Acute kidney injury in patients hospitalized with COVID-19, Kidney Int, vol. 98, pp. 209-218, Jul 2020.

[2] R. Oughtred, C. Stark, B. J. Breitkreutz, J. Rust, L. Boucher, C. Chang, et al., The BioGRID interaction database: 2019 update, Nucleic Acids Res, vol. 47, pp. D529-D541, Jan 8 2019.

[3] Y. He, H. Yu, E. Ong, Y. Wang, Y. Liu, A. Huffman, et al., CIDO, a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis, Sci Data, vol. 7, p. 181, Jun 12 2020.

[4] W. Huang da, B. T. Sherman, and R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources, Nat Protoc, vol. 4, pp. 44-57, 2009.

[5] D. L. Rubin, N. F. Noy, and M. A. Musen, Protege: a tool for managing and using terminology in radiology applications, J Digit Imaging, vol. 20 Suppl 1, pp. 34-46, Nov 2007.

[6] Y. Gu, J. Cao, X. Zhang, H. Gao, Y. Wang, J. Wang, et al., Interaction network of SARSCoV-2 with host receptome through spike protein, bioRxiv, 2020.

[7] I. Hamming, W. Timens, M. L. Bulthuis, A. T. Lely, G. Navis, and H. van Goor, Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis, J Pathol, vol. 203, pp. 631-7, Jun 2004.

[8] F. Pfister, E. Vonbrunn, T. Ries, H. M. Jack, K. Uberla, G. Lochnit, et al., Complement Activation in Kidneys of Patients With COVID-19, Front Immunol, vol. 11, p. 594849, 2020.

[9] E. Asgari, C. A. Farrar, N. Lynch, Y. M. Ali, S. Roscher, C. Stover, et al., Mannan-binding lectin-associated serine protease 2 is critical for the development of renal ischemia reperfusion injury and mediates tissue injury in the absence of complement C4, FASEB J, vol. 28, pp. 3996-4003, Sep 2014.

[10] T. Kosugi, K. Maeda, W. Sato, S. Maruyama, and K. Kadomatsu, CD147 (EMMPRIN/Basigin) in kidney diseases: from an inflammation and immune system viewpoint, Nephrol Dial Transplant, vol. 30, pp. 1097-103, Jul 2015.
[11] T. Zhou, X. Li, J. Zou, M. Cai, G. Sun, Y. Zhang, et al., Effects of DC-SIGN expression on renal tubulointerstitial fibrosis in nephritis, Front Biosci (Landmark Ed), vol. 14, pp. 2935-43, Jan 1 2009.
[12] D. A. Natale, C. N. Arighi, J. A. Blake, C. J. Bult, K. R. Christie, J. Cowart, et al., Protein Ontology: a controlled structured network of protein entities, Nucleic Acids Res, vol. 42, pp. D415-21, Jan 2014.
[13] S. Peng, H. Y. Wang, X. Sun, P. Li, Z. Ye, Q. Li, et al., Early versus late acute kidney injury among patients with COVID-19-a multicenter study from Wuhan, China, Nephrol Dial Transplant, vol. 35, pp. 2095-2102, Dec 4 2020.