# Peer Review File

**Manuscript Title:** A SARS-CoV-2 Protein Interaction Map Reveals Targets for Drug-Repurposing

## Reviewer Comments & Author Rebuttals

**Reviewer Reports on the Initial Version:**

Referee #1 (Remarks to the Author):

Review:
A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug-Repurposing

A. Summary of Key Results:

David E. Gordon and colleagues have cloned and transiently expressed 26 tagged SARS-CoV-2 proteins in human kidney cells. Using affinity-purification mass spectrometry, the authors generated a network of 22153 SARS-CoV-2-human protein-protein interactions (PPIs), from which they identify 332 high confidence PPIs using the SAINTexpress and MiST scoring algorithms. A series of bioinformatic analyses were then employed to provide insight into the biological processes and pathways enriched in their interaction network. A combination of chemoinformatic drug search and target- and pathway-specific literature search identified candidate therapeutics for treating COVID-19 based on existing data reflecting their modulation of 332 host factors interacting with SARS-CoV-2 proteins.

B. Originality and significance:

Previous attempts to characterise a SARS-CoV-2-human protein-protein interaction network were performed based solely on modelling of SARS-CoV-2 protein–protein interaction complexes (Srinivasan et al, 2020; PMID: 32218151) or pooling of protein-protein interaction data for viral proteins from other coronaviruses (Zhou et al, 2020; PMID: 32194980).

Thus, the data presented in this manuscript is highly original and constitutes the first SARS-CoV-2-host interactome derived from experimental evidence obtained using SARS-CoV-2 proteins as baits for affinity-purification mass spectrometry. This dataset and in particular the identification of candidate drugs that target host factors interacting with SARS-CoV-2 proteins may be important for the development of therapies for COVID-19. Given the evolving COVID pandemic this data is thus an important development in the field.

C. Data and methodology:

The data analysis and methodology used in this paper is an accepted standard. AP-MS has successfully identified host-pathogen interactions and this analysis is appropriate for identifying SARS-CoV-2 interactions with host proteins. One weakness to the approach is that all proteins of SARS-CoV-2 have been over-expressed independently to identify interaction partners in the absence of viral infection. The opportunity to study whole complexes involving multiple viral proteins is therefore limited. Additionally, there is no overlap between prey of different viral baits. This could suggest that 1.) other, or additional interactions might be observed if all viral proteins were expressed in addition to the individual bait. For example, the authors state that Nsp7 and Nsp8 are recognized to interact as part of the primase complex. It might therefore have been expected to observe some of the same prey between each of these baits. 2.) virus protein overexpression may result in host-protein interactions that may not be relevant. The authors do hint to the latter point when they discuss their observation that Orf3b is not expressed in their system and could indeed require other viral proteins (lines 148-150). These points should be discussed.

D. Appropriate use of statistics:
The authors use statistics appropriately. p-values are reported where appropriate, and links to R scripts for data analysis are indicated in methods. We could not find the Nature Research Reporting summary (detailing elements of experimental and analytical design) for this manuscript.

E. Conclusions: robustness, validity, reliability
The conclusions made here are valid based on the AP-MS and chemoinformatic analysis performed in this study. Many of the conclusions would be made much more robust if follow-up experiments were performed to verify some of the host-pathogen interactions, and also the utility of at least one of the drugs identified in modifying viral replication. Depending on the availability of live viral replication assays, the latter point may of course be difficult in a short time-frame. In general, wwhat is the consequence of the interaction of a drug with a host protein?
Overall, this study is clearly highly timely given the evolving pandemic and as such, early publication of the data would be appropriate.

F. Suggested improvements; experiments, data for possible revision

Major:

- In addition to the comments in the text above, the authors should seek to validate a small number of interactions. It would be particularly relevant to validate interactions that are further discussed in the paper such as the interaction between SARS-CoV-2 E protein and BRD2/BRD4.

- Several compounds have been identified in this study that are capable of targeting host factors interacting with SARS-CoV-2 proteins. The authors mention that they are currently testing these compounds for antiviral activity. The inclusion of the validation of the effect of at least one drug in a revised version of this manuscript would substantially increase the impact. For example, the authors could show whether specific BRD inhibitors have a functional effect.

- The authors should comment on the effect of over-expression of viral proteins and whether this system might also be facilitating interactions that wouldn't normally take place during infection. (See also section C)

- A comment should be made on the reasons why NSP5 and its mutant do not have the same high-confidence interactors given that the difference lies in a single-point mutation?

Minor:

- Line 137: The codon optimisation of SARS-CoV-2 proteins is not described in the methods (sequences and method needed).

- The reason for the absence of Nsp3 and Nsp16 from this dataset should be outlined in the main text.

- Despite the exclusion of the data for Orf7b due to high background, this data could still be included in Supplementary Table 1.

- References 21 and 33 are the same and no authors are mentioned. Please remove the duplication and add the missing information to the reference.

- On line 154 and 190 the instances of ‘332’ should be accompanied by ‘high confidence interactions’.

- Line 190 – it is not clear how the control sample of genes was defined for this analysis. Could the authors please clarify this in the main text or methods?

- On line 224, please expand the acronym TLE.

- Lines 275-288 are recapitulated almost exactly in lines 327-339. Please remove the duplication.

- On line 477 an NSP3 mutant is mentioned but not used or discussed anywhere else in the manuscript. Can the authors make a comment to its removal from the dataset or remove it completely from the manuscript?

- On lines 525 – 527, please clarify which instrument was used as the second mass spectrometer.

- On line 540 please indicate the version of Cytoscape used in this manuscript.

- On line 542, orf9c should be Orf9c.

- Line 543: "an" should be "a"

- On lines 542-546, the authors state that a larger in-house database of 87 baits was used to perform MiST scoring for the interactors of Orf8, Nsp8, Nsp13 and Orf9c. The raw files for this database of 87 baits does not appear to be included in the PRIDE submission for this manuscript. To ensure reproducibility, it would be pertinent to add these files to the same PRIDE entry. It is not otherwise possible for others to reproduce their conclusions from these data.

- On page 18 of the supplementary discussion, line 8 should be formatted as a title.

- Can the authors comment on the additional band that is observed for the NSP5 mutant compared to NSP5 on Figure 1C?

- On the legend for Figure 2C, the source for the expression data should be stated.

- In the legend for Figure 3, the source for the human protein interactions should be indicated.

- On Figure 4, the drugs depicted should be mentioned in the legend of 4-a-II and 4-b-I.

- In Extended figure 1 the degree of variability between biological replicates varies depending on samples, and while the majority do seem to cluster as expected there's a high degree of variability for the triplicates of Orf6. The authors should include a comment about this variability in the manuscript.

- In the Extended data Figure 4 legend: "that" should be "than"

- The graphs in Extended Figure 6 do not clearly depict a significant difference between SARS-CoV-2 interacting proteins and 'All RefSeq Genes', particularly for missense mutations. We suggest that the authors display these results differently to allow a clear visualisation of this result and explain the conclusions.

- In Supplementary Tables 1 and 2, please explain some of the column headings. In particular, explain 'Spec' and 'CtrlCounts' – are these spectral counts? These would be particularly useful to allow readers to produce similar plots as the ones made in Figure 3.

- Besides the worksheet name on the Excel file, no apparent difference is observed between the Supplementary Tables 5 and 6. Can the authors confirm the correct files have been uploaded and highlight the differences between the two files?

- Figure 4-B-III: The amino acid labels in the overlaid structure could be more clearly labeled to indicate which labels belong to which protein.

- Figure 4-C-II: The predicted alpha helix for Orf10 seems a bit out of place. Are the authors trying to suggest that Orf10 has structure? This is unclear.

- Figure 4-C-III: This model is very speculative. The schematic may be just as useful as the homology model considering the authors have not identified where Orf10 binds to ZYG11B.

G. References: appropriate credit to previous work
The authors do an exceptional job of reviewing the literature, especially in the extended discussion. This is a comprehensive and useful resource for interpreting the data of the main paper and for using as a vantage point for follow-up work on SARS-CoV-2.

H. Clarity and context: lucidity of abstract/ summary, appropriateness of abstract, introduction and conclusions
The abstract is clear and appropriate. The introduction is well-written and the conclusions based on the data presented are acceptable if somewhat speculative.

Luis Nobre, Cassie Zerbe and Mike Weekes

Referee #2 (Remarks to the Author):

Gordon and colleagues investigated cellular interaction partners of SARS-CoV-2 proteins and determined in silico whether known drugs could block these interactions. For this, they cloned the viral proteins, expressed them in 293T cells and identified interaction partners via affinity purification-mass spectrometry. The binding partners were enriched in certain protein domains and for several binding partners increased expression was found in the lung. Moreover, the results allowed to generate hypotheses regarding SARS-CoV-2 host cell interactions critical for viral spread and pathogenesis as well as modulation of the innate immune system. Finally, further in silico studies identified approved drugs that may inhibit protein-protein interactions potentially important for viral spread and pathogenesis. The manuscript is well written and provides the field with drug candidates and hypotheses regarding SARS-CoV-2 spread and pathogenesis that should be tested.

Major

A considerable weakness of this manuscript is the lack of proof of concept – in the sense that none of the compounds or interactions identified in silico/in vitro were analyzed for relevance for SARS-CoV-2 infection. This reviewer realizes that work with SARS-CoV-2 is limited to BSL3. However, the virus is available in many labs and, for instance, the 5-10 most promising drugs (identified in silico) can be tested within a week for antiviral activity. This should be done and the results provided.

Minor

"The identification of host dependency factors….". The screen might also detect restriction factors.

"SARS-CoV-2 appears to spread more efficiently, making it difficult to contain and increasing its pandemic potential." It should also be stated that SARS-CoV-2 infection is associated with a reduced case-fatality rate as compared to SARS-CoV and MERS-CoV infection.

"nucleotide analog RNA-dependent RNA Polymerase (RdRP) inhibitor remdesivir". "Nucleotide" should be exchanged by "nucleoside".

Referee #3 (Remarks to the Author):

Gordon et al., A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug-Repurposing

This is an interesting experience, and I suspect something of a harbinger of a new mode of publication for the life sciences, to be reviewing a paper that I already read in detail at the preprint stage, and one that was publicized internationally, featured in The New York Times, and that has already been used on an extraordinarily short time scale for helping guide the choice of medicines for COVID-19. I have to congratulate the authors on putting together a remarkable paper on such a short timeline with such a pressing international need. My assessment of the paper is that it is extremely strong, not just given the circumstances, but moving forward I suspect it will hold up quite well.

The approach itself is straightforward, using previously established affinity purification-mass spectrometry protocols to isolate human proteins that interact with SARS-nCoV-2 proteins that had been transiently expressed in HEK293T cells. As is typical with many AP/MS survey experiments, the interaction assays are performed in cells that are easy to grow and handle, rather than the physiologically relevant cells for the infection. Nonetheless, it's not unreasonable to expect that most of the interactions will be among core proteins shared across cell types. The authors use well-established and robust statistics for identifying interacting proteins and calculating significant interactors. The authors then leverage these interaction data to produce testable hypotheses prioritizing pharmaceutically available compounds and drugs for their potential to modulate COVID-19. The work is remarkable for taking a tried-and-true pipeline, applying it on an extraordinarily short time scale, and producing testable hypotheses about medications to address a pressing need.

Specific comments:

The authors analyze the human interacting partners for evolutionary conservation and conclude that the human proteins have depleted missense and premature stop mutations. The distributions plotted in Extended Data Fig 6 do not demonstrate this finding. Specifically, the missense distribution of human interacting proteins is indistinguishable from background even though the t-test performed shows significance. I therefore question the biological significance of this result. I would suggest the authors just leave that out. It is not relevant to the drug discovery story.

Some very exciting predictions that come from this work are the targets of Nps5 protease. I feel however a clearer description of the Nps5 cleavage sequence specificity is required. This will guide the reader to better interpret the predicted targets of Nps5. Specifically, it is unclear what the Nps5 cleavage specificity is and how well the HDAC2 and TRMT1 sequences match. A sequence logo of the sequence specificity would suffice.

Minor edits/typos:

"Chemoinformatics searches of the literature yielded 15 approved drugs, four investigational new drugs (clinical), and 18 pre-clinical candidates (Table 1a), while specialist knowledge revealed 12 approved drugs, 10 investigational new drugs (clinical), and 10 preclinical candidates (Table 1b)." It is unclear from this wording whether the searches were done on the complete set of human interactors.

"Similarly, the Sigma2 receptor interacted with the vial protein orf9." Should be "viral protein".

The Extended Data Figure legend mentions Coronanet 1.0 but there is no citation/link.

The authors used a standard mass spectrometry experiment protocol for their protein identification. However it's unclear in their methods when the authors state, "...on two different mass spectrometers" they refer to two Q-Exactive Plus mass spectrometers that were run using the same methods or if different instruments and methods were used. The authors might better clarify the machines used in their experiments.

Many times it is unclear when a gene or the protein is being referenced. For example there seems to be mixed usage of nsp5 and Nsp5. It would be nice if this were kept consistent in the text.

Paragraph duplication in section, "Identification of existing drugs targeting SARS-CoV-2 human host factors". Specifically the paragraph starting with, "To identify small molecules targeting…"

**Author Rebuttals to Initial Comments:**

**Referee #1 (Remarks to the Author):**

**Review:**

**A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug-Repurposing**

**A. Summary of Key Results:**

**David E. Gordon and colleagues have cloned and transiently expressed 26 tagged SARS-CoV-2 proteins in human kidney cells. Using affinity-purification mass spectrometry, the authors generated a network of 22153 SARS-CoV-2-human protein-protein interactions (PPIs), from which they identify 332 high confidence PPIs using the SAINTexpress and MiST scoring algorithms. A series of bioinformatic analyses were then employed to provide insight into the biological processes and pathways enriched in their interaction network. A combination of chemoinformatic drug search and target- and pathway-specific literature search identified candidate therapeutics for treating COVID-19 based on existing data reflecting their modulation of 332 host factors interacting with SARS-CoV-2 proteins.**

**B. Originality and significance:**

**Previous attempts to characterise a SARS-CoV-2-human protein-protein interaction network were performed based solely on modelling of SARS-CoV-2 protein–protein interaction complexes (Srinivasan et al, 2020; PMID: 32218151) or pooling of protein-protein interaction data for viral proteins from other coronaviruses (Zhou et al, 2020; PMID: 32194980).**

**Thus, the data presented in this manuscript is highly original and constitutes the first SARS-CoV- 2-host interactome derived from experimental evidence obtained using SARS-CoV-2 proteins as baits for affinity-purification mass spectrometry. This dataset and in particular the identification of candidate drugs that target host factors interacting with SARS-CoV-2 proteins may be important for the development of therapies for COVID-19. Given the evolving COVID pandemic this data is thus an important development in the field.**

**We thank the reviewer for their kind words of support.**

**C. Data and methodology:**

***The data analysis and methodology used in this paper is an accepted standard. AP-MS has successfully identified host-pathogen interactions and this analysis is appropriate for identifying SARS-CoV-2 interactions with host proteins. One weakness to the approach is that all proteins of SARS-CoV-2 have been over-expressed independently to identify interaction partners in the absence of viral infection. The opportunity to study whole complexes involving multiple viral proteins is therefore limited. Additionally, there is no overlap between prey of different viral baits. This could***

*suggest that 1.) other, or additional interactions might be observed if all viral proteins were expressed in addition to the individual bait. For example, the authors state that Nsp7 and Nsp8 are recognized to interact as part of the primase complex. It might therefore have been expected to observe some of the same prey between each of these baits. 2.) virus protein overexpression may result in host-protein interactions that may not be relevant. The authors do*

***hint to the latter point when they discuss their observation that Orf3b is not expressed in their system and could indeed require other viral proteins (lines 148-150). These points should be discussed.***

**We thank the reviewer for raising a number of important points.**

**The reviewer rightly points out that there is no overlap in preys detected between any two viral baits. This may partly be a consequence of our protein scoring methodology, a component of which emphasizes the uniqueness of interactions to minimize promiscuous background and focus on high-confidence interactions. In addition, while it might seem logical that individual components of the same viral complex would pull down similar host proteins, we question whether all subunits of a viral complex would bind the same host protein interactor with high affinity -- we don't believe this is a foregone conclusion, due to the likely heterogeneous contributions of protein interaction surfaces among the quaternary structure of the viral protein complexes in question. In addition, viral proteins often have multiple functions, and these may extend their interaction landscape outside their known interactions with other viral proteins.**

**Nevertheless, the reviewers rightly point out that assembly of complexes between SARS- CoV-2 proteins will likely augment their interactions with host molecular machinery. The following text that has been added to the supplemental discussion to address these points:**

**"A number of SARS-CoV-2 proteins are known to assemble and function in multi-subunit complexes (e.g. Nsp7, 8 and 12) (Kirchdoerfer and Ward 2019), and it is likely that co- expression of specific proteins will highlight other important biochemical interactions with the host. In order to address this, we are generating a library of untagged transient expression constructs for co-transfection with the tagged library used in this study. In parallel we will also integrate endogenous affinity tags into SARS-CoV-2 replicons and full- length recombinant SARS-CoV-2 genomes to study interactions in the context of infection. Performing affinity purifications in the context of infection will also address concerns regarding protein overexpression, which may allow the identification of weak interactions, but may also lead to false positives. Replication of our AP-MS overexpression experiments is ongoing in other cell types, including cultured and primary lung epithelial cells, and this data will also help clarify the most high-confidence host interactions."**

**D. Appropriate use of statistics:**

**The authors use statistics appropriately. p-values are reported where appropriate, and links to R scripts for data analysis are indicated in methods. We could not find the Nature Research Reporting summary (detailing elements of experimental and analytical design) for this manuscript.**

**Thank you for your patience, the Nature Research Reporting summary is attached to this resubmission.**

**E. Conclusions: robustness, validity, reliability**

**The conclusions made here are valid based on the AP-MS and chemoinformatic analysis performed in this study. Many of the conclusions would be made much more robust if follow-up experiments were performed to verify some of the host-pathogen interactions, and also the utility of at least one of the drugs identified in modifying viral replication. Depending on the availability**

**of live viral replication assays, the latter point may of course be difficult in a short time-frame. In general, what is the consequence of the interaction of a drug with a host protein?**

**Overall, this study is clearly highly timely given the evolving pandemic and as such, early publication of the data would be appropriate.**

**To address the reviewer's request for functional validation of our drug predictions, we now include data describing the impact of small molecules on SARS-CoV-2 infection in cultured cells in the revised manuscript.**

**F. Suggested improvements; experiments, data for possible revision**

**Major:**

***In addition to the comments in the text above, the authors should seek to validate a small number of interactions. It would be particularly relevant to validate interactions that are further discussed in the paper such as the interaction between SARS-CoV-2 E protein and BRD2/BRD4.***

***Several compounds have been identified in this study that are capable of targeting host factors interacting with SARS-CoV-2 proteins. The authors mention that they are currently testing these compounds for antiviral activity. The inclusion of the validation of the effect of at least one drug in a revised version of this manuscript would substantially increase the impact. For example, the authors could show whether specific BRD inhibitors have a functional effect.***

**For verification purposes, we conducted antiviral assays independently by our colleagues at Mount Sinai (37 drugs tested) and at the Pasteur Institute (44 drugs tested), who were among the first in the world to be able to run SARS-CoV-2 infectivity assays. We are excited to report that two classes of drugs stood out as effective at reducing viral infectivity: the protein translation inhibitors Zotatavin, Ternatin 4, and PS3061, and multiple ligands of the Sigma1 and Sigma2 receptors, including the approved drugs clemastine, cloperastine, and haloperidol, the pre-clinical molecules PB28 and PD-144418, and hydroxychloroquine, which has recently been linked to COVID-19. We provide additional *in vitro* evidence to suggest that several of these ligands show higher affinity for the Sigma1/2 receptors over the hERG receptor (linked to cardiotoxicity) when compared to hydroxychloroquine, suggesting therapeutic strategies that are less toxic.**

***The authors should comment on the effect of over-expression of viral proteins and whether this system might also be facilitating interactions that wouldn't normally take place during infection. (See also section C)***

**We have inserted the following text into the supplemental discussion:**

**"A number of SARS-CoV-2 proteins are known to assemble and function in multi-subunit complexes (e.g. Nsp7, 8 and 12) (Kirchdoerfer and Ward 2019), and it is likely that co- expression of specific proteins will highlight other important biochemical interactions with the host. In order to address this, we are generating a library of untagged transient expression constructs for co-transfection with the tagged library used in this study. In parallel we will also integrate endogenous affinity tags into SARS-CoV-2 replicons and full- length recombinant SARS-CoV-2 genomes to study interactions in the context of infection. Performing affinity purifications in the context of infection will also address concerns**

**regarding protein overexpression, which may allow the identification of weak interactions, but may also lead to false positives. Replication of our AP-MS overexpression experiments is ongoing in other cell types, including cultured and primary lung epithelial cells, and this data will also help clarify the most high-confidence host interactions."**

**To the reviewer's second point, it is correct that viral protein overexpression may result in the immunoprecipitation of promiscuous or otherwise irrelevant host proteins. We address this by imposing a stringent scoring pipeline which weighs the abundance, specificity and reproducibility of each host interacting protein, with the goal of removing spurious interactions from our dataset. To be identified as high confidence, an interaction must pass three distinct scoring thresholds, one based on MIST (>= 0.7), one based on SAINT (BFDR <= 0.05), and one based on protein abundance in the immunoprecipitate (average spectral count >= 2). This requires high confidence interactions to have high specificity across baits as well as high abundance compared to negative controls (empty vector and GFP).**

***A comment should be made on the reasons why NSP5 and its mutant do not have the same high- confidence interactors given that the difference lies in a single-point mutation?***

**We reason that the C145A catalytically dead mutation may interfere with the HDAC2 interaction with wild type Nsp5. We have added the following text to the supplemental discussion:**

**"We typically utilize catalytically dead mutants to stabilize protease interactions with substrates, thus it's no surprise that we identify more interactors for the C145A Nsp5 mutant. The fact that HDAC2 specifically interacts with wild-type Nsp5 may suggest that this interaction is unusually stable, and perhaps that the C145A mutation may interfere with this interaction."**

**Minor:**

***Line 137: The codon optimisation of SARS-CoV-2 proteins is not described in the methods (sequences and method needed).***

**We have added our annotated genome as well as genbank files of all SARS-CoV-2 mammalian expression plasmids generated for this study.**

**Regarding codon optimization, we have added a section on cloning to our Materials and Methods section, including the following text:**

**"Open reading frames and proteolytically mature Nsps annotated in the SARS-CoV-2 genome were human codon optimized using the IDT codon optimization tool (https://www.idtdna.com/codonopt) and internal EcoRI and BamHI sites eliminated. Start and stop codons were added as necessary to Nsps 1-16, a Kozak sequence was added before each start codon, and a 2x-Strep tag with linker was added to either the N- or C- terminus."**

***The reason for the absence of Nsp3 and Nsp16 from this dataset should be outlined in the main text.***

**In the new cloning section under Materials and Methods we have added the following text:**

**“Nsp16 displayed multiple mutations which could not be repaired prior to the time- sensitive preparation of this manuscript, and Nsp3 was too large to be synthesized in time to be included in this study. Strep-tagged constructs encoding Nsp3, Nsp3 C857A (catalytic mutant), and Nsp16 will be used in future AP-MS experiments. “**

***Despite the exclusion of the data for Orf7b due to high background, this data could still be included in Supplementary Table 1.***

**Our belief is that the high signal observed for Orf7b was the result of a technical issue. We determined that the higher than average number of peptides/proteins detected from these samples classify it as an outlier, which did not satisfy the assumptions for MiST scoring (namely a similar background proteome detection across samples). Therefore, we prefer to exclude interactions identified for orf7b because we are not confident in technical aspects of the orf7b results. Moving forward, we are optimizing this protein for inclusion in future studies and datasets.**

***References 21 and 33 are the same and no authors are mentioned. Please remove the duplication and add the missing information to the reference.***

**Thank you for pointing this out. The duplicated reference has been removed and the reference information has been corrected.**

***On line 154 and 190 the instances of ‘332’ should be accompanied by ‘high confidence interactions’.***

**This text referring to the 332 high confidence interactions has been modified accordingly.**

***Line 190 – it is not clear how the control sample of genes was defined for this analysis. Could the authors please clarify this in the main text or methods?***

**Per suggestion of Reviewer 3, Extended Data Figure 6 and the corresponding paragraph in the main text have been removed from the manuscript.**

***On line 224, please expand the acronym TLE.***

**We now define the TLE acronym: “transducin-like enhancer of split (TLE) family…” in the supplemental discussion.**

***Lines 275-288 are recapitulated almost exactly in lines 327-339. Please remove the duplication.***

**We thank the reviewer for pointing this out; the duplicated paragraph has now been removed.**

***On line 477 an NSP3 mutant is mentioned but not used or discussed anywhere else in the manuscript. Can the authors make a comment to its removal from the dataset or remove it completely from the manuscript?***

**We now address the omission of Nsp3 wild type and C857A mutant in the cloning section of the Materials and Methods:**

**"...Nsp3 was too large to be synthesized in time to be included in this study. Strep-tagged constructs encoding Nsp3, Nsp3 C857A (catalytic mutant), and Nsp16 will be used in future AP-MS experiments."**

***On lines 525 – 527, please clarify which instrument was used as the second mass spectrometer.***

**We corrected this and now state that all samples were analyzed on one instrument, the Q- Exactive Plus mass spectrometer (Thermo Fisher Scientific)**

***On line 540 please indicate the version of Cytoscape used in this manuscript.***

**The Cytoscape version used in this manuscript is 3.7.1. We have added a note indicating this in the Materials and Methods section.**

***On line 542, orf9c should be Orf9c.***

**We have corrected this in the revised manuscript.**

***Line 543: "an" should be "a"***

**We have corrected this in the revised manuscript.**

***On lines 542-546, the authors state that a larger in-house database of 87 baits was used to perform MiST scoring for the interactors of Orf8, Nsp8, Nsp13 and Orf9c. The raw files for this database of 87 baits does not appear to be included in the PRIDE submission for this manuscript. To ensure reproducibility, it would be pertinent to add these files to the same PRIDE entry. It is not otherwise possible for others to reproduce their conclusions from these data.***

**The in-house dataset of 87 baits used to perform MIST scoring for Orf8, Nsp8, Nsp13, and Orf9c were extracted from a large, multi-year initiative in our laboratory that has yet to be published. The purpose of including the additional**

**data was to provide a more comprehensive background dataset for comparison, introducing increased diversity to more adequately control for Orf8, Nsp8, Nsp13, and Orf9c possessing a large number of interactions which we deemed non-specific. If any other large number of baits are used as background controls, similar results would be obtained.**

***On page 18 of the supplementary discussion, line 8 should be formatted as a title.***

**We have corrected this in the revised manuscript.**

***Can the authors comment on the additional band that is observed for the NSP5 mutant compared to NSP5 on Figure 1C?***

**We thank the reviewer for this intriguing question. We hypothesize that the mutation of the catalytic site may impact the stability or localization of Nsp5. The smaller band may be a degradation product.**

*On the legend for Figure 2C, the source for the expression data should be stated.*

**We have updated the legend of Figure 2C to specify that protein expression data from the Human Protein Atlas was utilized for this analysis. The reference for this data has been added to the manuscript and is listed below:**

**Wang, D. et al. A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol. Syst. Biol. 15, e8503 (2019).**

*In the legend for Figure 3, the source for the human protein interactions should be indicated.*

**Human-human protein-protein interactions were curated from CORUM, IntAct, and Reactome. We have added this sentence to the legend of Figure 3, as suggested by the reviewer.**

*On Figure 4, the drugs depicted should be mentioned in the legend of 4-a-II and 4-b-I.*

**We have added the drugs to the legend of Figure 4.**

*In Extended figure 1 the degree of variability between biological replicates varies depending on samples, and while the majority do seem to cluster as expected there's a high degree of variability for the triplicates of Orf6. The authors should include a comment about this variability in the manuscript.*

**As demonstrated in Extended Data Figure 1, there is a high degree of correlation between replicates for individual baits, which we now state in the main text. Quantification of pairwise correlation between each of the triplicate mass spec runs for each SARS-CoV-2 bait results in a range of values from 0.46 - 0.72. The pairwise correlations of Orf6 replicates are 0.6, 0.57, and 0.49, which are consistent with the rest of the dataset. In most cases the experiments corresponding to each bait cluster together, with the exception of a few baits such as Orf6 with lower numbers of specific host interactions, but the correlations between replicates among all baits including Orf6 appear consistent.**

**To the legend of Extended Data Figure 2 we added: "Correlation between replicates for individual baits ranges from 0.46 - 0.72, and in most cases the experiments corresponding to each bait cluster together, with the exception of a couple of baits with lower numbers of specific host interactions (e.g. E, Nsp2, Orf6, Orf3a, and Orf3b)."**

*In the Extended data Figure 4 legend: "that" should be "than"*

**We have corrected this in the revised manuscript.**

***The graphs in Extended Figure 6 do not clearly depict a significant difference between SARS- CoV-2 interacting proteins and 'All RefSeq Genes', particularly for missense mutations. We suggest that the authors display these results differently to allow a clear visualisation of this result and explain the conclusions.***

**This Extended Data figure has been deleted in response to reviewer 3; please see comment below.**

*In Supplementary Tables 1 and 2, please explain some of the column headings. In particular, explain 'Spec' and 'CtrlCounts' – are these spectral counts? These would be particularly useful to allow readers to produce similar plots as the ones made in Figure 3.*

**We have now added a description to the table legend for clarification:**

**"Supplementary table 1: Scoring results for all baits and all proteins, showing spectral counts of experimental samples (columns AvgSpec and Spec) and empty controls (column Ctrl Counts)", and**

**"Supplementary table 2: SARS-CoV 2 high confidence interactors, showing spectral counts of experimental samples (columns AvgSpec and Spec) and empty controls (column Ctrl Counts)"**

*Besides the worksheet name on the Excel file, no apparent difference is observed between the Supplementary Tables 5 and 6. Can the authors confirm the correct files have been uploaded and highlight the differences between the two files?*

**We would like to thank the reviewers for spotting this. We have now corrected our tables as follows: Supplementary Tables 5 and 6 show the raw chemical associations of prey proteins per IUPHAR/BPS Guide to Pharmacology (2020-3-12) (Table S5) or ChEMBL25 (Table S6), as stated in the Table legend.**

*Figure 4-B-III: The amino acid labels in the overlaid structure could be more clearly labeled to indicate which labels belong to which protein.*

**We have clarified this in the revised manuscript.**

*Figure 4-C-II: The predicted alpha helix for Orf10 seems a bit out of place. Are the authors trying to suggest that Orf10 has structure? This is unclear.*

**We used JPRED (https://www.compbio.dundee.ac.uk/jpred/index.html) to predict an alpha helical region in Orf10, which suggests that Orf10 is not a random unfolded peptide and likely has a secondary structure. This structure could be adopted in the isolated polypeptide or when it is in complex with host factors. Upon double-checking the position of the predicted alpha helix, we have adjusted the corresponding figure to be more accurate.**

***Figure 4-C-III: This model is very speculative. The schematic may be just as useful as the homology model considering the authors have not identified where Orf10 binds to ZYG11B.***

**We agree with the reviewer and have removed panel 4-C-III.**

***G. References: appropriate credit to previous work***

***The authors do an exceptional job of reviewing the literature, especially in the extended discussion. This is a comprehensive and useful resource for interpreting the data of the main paper and for using as a vantage point for follow-up work on SARS-CoV-2.***

**We thank the reviewer for their kind words of support.**

*H. Clarity and context: lucidity of abstract/ summary, appropriateness of abstract, introduction and conclusions*

***The abstract is clear and appropriate. The introduction is well-written and the conclusions based on the data presented are acceptable if somewhat speculative.***

**To address the reviewer's request for functional validation of our drug predictions, we include data describing the impact of small molecules on SARS-CoV-2 infection in cultured cells in the revised manuscript.**

**Referee #2 (Remarks to the Author):**

**Gordon and colleagues investigated cellular interaction partners of SARS-CoV-2 proteins and determined in silico whether known drugs could block these interactions. For this, they cloned the viral proteins, expressed them in 293T cells and identified interaction partners via affinity purification-mass spectrometry. The binding partners were enriched in certain protein domains and for several binding partners increased expression was found in the lung. Moreover, the results allowed to generate hypotheses regarding SARS-CoV-2 host cell interactions critical for viral spread and pathogenesis as well as modulation of the innate immune system. Finally, further in silico studies identified approved drugs that may inhibit protein-protein interactions potentially important for viral spread and pathogenesis. The manuscript is well written and provides the field with drug candidates and hypotheses regarding SARS-CoV-2 spread and pathogenesis that should be tested.**

**Major**

***A considerable weakness of this manuscript is the lack of proof of concept – in the sense that none of the compounds or interactions identified in silico/in vitro were analyzed for relevance for SARS-CoV-2 infection. This reviewer realizes that work with SARS-CoV-2 is limited to BSL3. However, the virus is available in many labs and, for instance, the 5-10 most promising drugs (identified in silico) can be tested within a week for antiviral activity. This should be done and the results provided.***

**For verification purposes, we conducted antiviral assays independently by our colleagues at Mount Sinai (37 drugs tested) and at the Pasteur Institute (44 drugs tested), who were among the first in the world to be able to run SARS-CoV-2 infectivity assays. We are excited to report that two classes of drugs stood out as effective at reducing viral infectivity: the protein translation inhibitors Zotatavin, Ternatin 4, and PS3061, and multiple ligands of the Sigma1 and Sigma2 receptors, including the approved drugs clemastine, cloperastine, and haloperidol, the pre-clinical molecules PB28 and PD-144418, and hydroxychloroquine, which has recently been linked to COVID-19. We provide additional *in vitro* evidence to suggest that several of these ligands show higher affinity for the Sigma1/2 receptors**

**over the hERG receptor (linked to cardiotoxicity) when compared to hydroxychloroquine, suggesting therapeutic strategies that are less toxic.**

**Minor**

***"The identification of host dependency factors….". The screen might also detect restriction factors.***

**We agree with the reviewer that our PPI map likely includes both host-dependency and host-restriction factors. The abstract now refers to host factors which are the targets of small molecules that inhibit SARS-CoV-2 infection.**

***"SARS-CoV-2 appears to spread more efficiently, making it difficult to contain and increasing its pandemic potential." It should also be stated that SARS-CoV-2 infection is associated with a reduced case-fatality rate as compared to SARS-CoV and MERS-CoV infection.***

**We now address this comment in the corresponding sentence: "Compared to MERS-CoV or SARS-CoV, SARS-CoV-2 has a lower case-fatality rate but appears to spread more efficiently6,7, making it difficult to contain."**

***"nucleotide analog RNA-dependent RNA Polymerase (RdRP) inhibitor remdesivir". "Nucleotide" should be exchanged by "nucleoside".***

**Nucleotide has been changed to nucleoside.**

**Referee #3 (Remarks to the Author):**

**Gordon et al., A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug-Repurposing**

**This is an interesting experience, and I suspect something of a harbinger of a new mode of publication for the life sciences, to be reviewing a paper that I already read in detail at the preprint stage, and one that was publicized internationally, featured in The New York Times, and that has already been used on an extraordinarily short time scale for helping guide the choice of medicines for COVID-19. I have to congratulate the authors on putting together a remarkable paper on such a short timeline with such a pressing international need. My assessment of the paper is that it is extremely strong, not just given the circumstances, but moving forward I suspect it will hold up quite well.**

**The approach itself is straightforward, using previously established affinity purification-mass spectrometry protocols to isolate human proteins that interact with SARS-nCoV-2 proteins that had been transiently expressed in HEK293T cells. As is typical with many AP/MS survey experiments, the interaction assays are performed in cells that are easy to grow and handle, rather than the physiologically relevant cells for the infection. Nonetheless, it's not unreasonable to expect that most of the interactions will be among core proteins shared across cell types. The authors use well-established and robust statistics for identifying interacting proteins and calculating significant interactors. The authors then leverage these interaction data to produce testable hypotheses prioritizing pharmaceutically available compounds and drugs for their potential to modulate COVID-19. The work is remarkable**

**for taking a tried-and-true pipeline, applying it on an extraordinarily short time scale, and producing testable hypotheses about medications to address a pressing need.**

**We thank the reviewer for their kind words of support.**

**Specific comments:**

*The authors analyze the human interacting partners for evolutionary conservation and conclude that the human proteins have depleted missense and premature stop mutations. The distributions plotted in Extended Data Fig 6 do not demonstrate this finding. Specifically, the missense distribution of human interacting proteins is indistinguishable from background even though the t- test performed shows significance. I therefore question the biological significance of this result. I would suggest the authors just leave that out. It is not relevant to the drug discovery story.*

**We would like to thank the reviewer for their suggestion. In the revised version of the manuscript, we have deleted Extended Data Fig. 6 and the corresponding paragraph in the main text.**

*Some very exciting predictions that come from this work are the targets of Nps5 protease. I feel however a clearer description of the Nps5 cleavage sequence specificity is required. This will guide the reader to better interpret the predicted targets of Nps5. Specifically, it is unclear what the Nps5 cleavage specificity is and how well the HDAC2 and TRMT1 sequences match. A sequence logo of the sequence specificity would suffice.*

**Thank you for your suggestion. Since SARS-CoV-2 nsp5 is 98.7% similar to SARS-CoV nsp5, we have included sequence logos based on the 11 nsp5 cleavage sites in each virus demonstrating the nsp5 cleavage sequence; see Extended Data Fig. 6.**

**Minor edits/typos:**

*"Chemoinformatics searches of the literature yielded 15 approved drugs, four investigational new drugs (clinical), and 18 pre-clinical candidates (Table 1a), while specialist knowledge revealed 12 approved drugs, 10 investigational new drugs (clinical), and 10 preclinical candidates (Table 1b)." It is unclear from this wording whether the searches were done on the complete set of human interactors.*

**We thank the reviewer for the helpful comment. We have clarified that this analysis was done on the complete set of interactors in the text.**

*"Similarly, the Sigma2 receptor interacted with the vial protein orf9." Should be "viral protein".*

**We have corrected this in the revised manuscript.**

*The Extended Data Figure legend mentions Coronanet 1.0 but there is no citation/link.*

**We have corrected Coronanet to NetCorona and added the corresponding reference in the revised Supplementary Methods.**

***The authors used a standard mass spectrometry experiment protocol for their protein identification. However it's unclear in their methods when the authors state, "...on two different mass spectrometers" they refer to two Q-Exactive Plus mass spectrometers that were run using the same methods or if different instruments and methods were used. The authors might better clarify the machines used in their experiments.***

**We have corrected this and now state that samples were analyzed on one instrument, the Q-Exactive Plus mass spectrometer (Thermo Fisher Scientific).**

***Many times it is unclear when a gene or the protein is being referenced. For example there seems to be mixed usage of nsp5 and Nsp5. It would be nice if this were kept consistent in the text.***

**We have revised the manuscript to systematically name SARS-CoV-2 proteins with the first letter uppercase, remaining letters lowercase.**

***Paragraph duplication in section, "Identification of existing drugs targeting SARS-CoV-2 human host factors". Specifically the paragraph starting with, "To identify small molecules targeting…"***

**We thank the reviewer for spotting this, it has been corrected in the revised manuscript.**

**Reviewer Reports on the First Revision:**

Referee #2 (Remarks to the Author):

The authors have adequately addressed the questions raised by this reviewer. They still might want to state that Vero E6 kidney cells were used for testing antiviral activity and that lung cells, which are more relevant, might have yielded different results for some inhibitors. In addition, OC435 should read OC43.

**Author Rebuttals to First Revision:**

In response to Referee #2, we now include the following statement in our methods section: "The Vero E6 cell line used in this study is a kidney cell line; therefore, we cannot exclude that lung cells yield different results for some inhibitors"