Reviewer 1

The authors discuss the nature of virus protein-protein interaction information that is currently available in various databases, and how these large-scale data might be mined. I have several issues with this manuscript.

Firstly, according to the request to review that I received, it is my understanding that this paper is an invited review article, yet it appears to contain primary research material – namely the analyses presented in figure 1,2 and 3. The structure of the paper is consistent with a review article, but without a fully described method section, I feel it would be inappropriate to publish these analyses in the current format.

The writing is often less formal than is found in many scientific manuscripts. In most cases, I enjoy this because it improves the readability. However, there are some phrases, I would consider revising: ‘It turns out…’ and ‘Virus proteins attack…’ are two I find too colloquial. I also found that acronyms are frequently not defined, and are sometimes defined on the second use. In particular, PPI in the abstract is not defined anywhere. As many viruses form protein-protein interactions with peptidyl-prolyl isomerases, PPI as an acronym is confusing if not defined. I would suggest, as it is not used consistently throughout the manuscript, that PPI be removed. I also find some of the comments in the introduction and conclusions to be generic and unhelpful (if not inaccurate). The assertion in the introduction that viruses are more difficult to control than bacteria due to the use of antibiotics is a very debatable statement, and given that the paper is not about bacteria, this statement is unnecessary. In the conclusion: ‘Viruses evolve much quicker than their hosts, especially RNA viruses, hence viruses can also adapt their virus-host interactions faster than a host can react by mutating its target proteins, including its immune system.’ This statement is at odds with the Red Queen hypothesis – that virus and host co-evolve to maintain mutual survival. Indeed the adaptive arm of the host immune system can, in many cases, actively evolve to eliminate infection faster than the virus can respond. Again, I’m not sure of the value of this statement in the context of the conclusions.

I am also concerned about the oversimplification of some of the biology in this paper. In Table 1, viruses are classified as either DNA or RNA (or ‘anim’ or ‘plant’ is some cases – why?). The Baltimore system of classification of viruses defines seven types of genome: dsDNA, ssDNA, dsRNA, ssRNA(+ve), ssRNA(-ve), ssRNA-RT, and dsDNA-RT. These different groups have different replication requirements which impacts on their mutation rate (amongst other things). In table 1, hepdnaviridae are labeled as DNA viruses – which isn’t strictly true as they have an RNA intermediate. Likewise, retroviridae have a DNA intermediate. So again, I am unsure as to the value of the describing the virus families in terms of being RNA or DNA, as is done in Table 1.

In Figure 2 (and the accompanying section ‘The topology of human-virus interactions’) the authors describe critical nodes in interaction networks as statistically more likely to be targets from protein interactions. Intuitively, this seems reasonable, and I like it as a concept. But could it also be that these critical nodes are also overrepresented in false-positives in large-scale interactome analyses?

In Figure 3, I am intrigued to see that retroviridae and orthomyxoviridae have a negative fold change with respect to cell wall proteins, while herpesviridae and papillomarviridae have a positive fold change. None of these families of virus infect an organism with a cell wall. This calls into question whether the authors understand the systems they are studying. Also, there appears to be a phylogenetic analysis on both axes of the heat map, but this is not described or mentioned. I also question the value of putting ‘other’ in the heat map. Given the assertion that different viruses have different interaction preferences, surely grouping everything that is not of the aforementioned families together would result in a largely featureless map, averaged around no change – which is close to what is shown.

On the whole, I find this paper to be strangely constructed. It is a hybrid review/original research piece, with a smattering of ideas and concepts, with limited flow between sections. The conclusions of the paper are particularly unremarkable. The statement ‘…suggesting that different viral families use different strategies to invade a human host cell’ to be one of the most unprofound statements I have seen in a manuscript. As the conclusions read, the ‘take home message’ appears to be that: there are many interactome databases; some viruses are overrepresented and full of false positives; other viruses are underrepresented; and viruses are different from each other. So I would question what the value of such a paper would have to the scientific community.

Reviewer 2

In this manuscript, Goodacre and colleagues broadly examine protein-protein interactions (PPIs) in the context of human virus infections and discuss the physiological relevance of these interactions. While a large number of PPIs have been identified, the authors raise the question of how many interactions are physiological and required for viral infection. Interestingly, there is considerable overlap in the human proteins targeted by different virus families. The authors describe the databases available for the study of human-virus PPIs and provide a useful summary of the strengths and weaknesses of the databases and their potential uses. Overall, this work is likely to be of interest to a broad readership in virology and addresses some critical questions as the use of systems biology approaches becomes more prevalent.

Specific comments:

1. When questioning the reliability of published PPIs, the authors describe cases where the predicted interactions were validated by co-immunoprecipitation and GST pull-down experiments. However, another critical question is whether these interactions been shown to have a functional effect on viral infection (i.e. by examining viral replication when the interaction is blocked by knockdown, small molecule inhibitors, or genetics approaches). Extending this section to describe and provide examples of functional effects would strengthen the manuscript.

2. Could the authors comment on potential effects of different cell lines in terms of PPIs? One problem is that the cell lines used to study viral infection are not always physiological (e.g., the use of cancer cell lines, where signaling pathways and protein expression levels may be disregulated). One can imagine that PPIs identified in such cell lines may be even less likely to accurately reflect the physiological situation. Some discussion of this, or providing some examples, would be appreciated.

3. Many viruses infect different types of hosts (e.g., Dengue is spread by mosquitoes, Influenza infects birds and pigs). Have any studies compared PPIs in different hosts for viruses such as these? While the reviewer appreciates this manuscript is focused on human disease, it may be interesting to include a brief section about PPIs in other hosts for this group of viruses, many of which are medically important human pathogens.

4. More discussion of the outlook and future directions would be useful for researchers in the field. How can we now exploit this collection of PPIs to target viral infections or diseases associated with the viral interactome? How could databases be improved? How can we better ensure that identified virus-human PPIs are biologically relevant?

Minor comments:

1. Abstract, line 10: insert comma “…relevant, for many…”

2. Page 2, Introduction, paragraph 1, line 4: insert comma “…and structure, viruses…”

3. Page 3, Virus-host interactions in humans, line 9: change to “…more than 100 interactions…” instead of “more than a 100”

4. Page 6, line 2: “Betweenes centrality” should be “Betweenness centrality”?

5. Page 7, Virus interactions with the host metabolome, line 4: change to “…with the cancer protein…” instead of “th cancer”

Reviewer 3

This manuscript provides a global survey of virus-host PPIs in literature and databases. Recent advances and limitations of these resources are discussed. This is an update of the current situation, but some paragraphs have significant weaknesses. Several points should be addressed.

Point 1) In page 2 (second paragraph), some sentences are misleading. Indeed, it could be wrongly understood that « medical interest » is somehow related to higher « infectivity », which could be correlated with higher « pathogenicity » (“In fact, …”). This is confusing and should be clarified.

Point 2) Table 1. Which numbers were used to calculate the U/C ratio? For example, 1008 total genomes are available for Paramyxoviridae and clustered into 327 groups. However, the U/C ratio is 10.06.

Point 3) Table 1: Arenaviridae are responsible for major human diseases such as Lassa fever, and should not be restricted to “animals”.

Point 4) In the legend of Table 1, it is said “Clustered sequenced were clustered at ≥98% sequence identity”. This is different from the main text (Page 2; sequences are clustered and those genomes combined which have less than 5% sequence…”).

Point 5) Page 2: “We have thus compiled data on the diversity of human viruses by looking both at sequence diversity (Table 1)…”. Table 1 compiled data on sequence diversity inside viral families, not “viruses” or “human viruses”. This should be corrected.

Point 6) Table 2: The viral family should be indicated for each virus.

Point 7) Page 3: “Coronaviruses, are highly pathogenic with a lot of victims, but have received rather scant attention”. This is contradictory with data presented in Table 2, since coronaviruses were responsible for a very small number of victims (<10.000 for SARS-CoV and MERS-CoV over several years). In fact, dengue virus, hRSV or rabies virus that are not cited in Table 2 have a much higher impact on human populations.

Point 8) Table 2: “Unless otherwise indicated, figures are yearly”. Data presented for SARS-CoV and MERS-CoV are not yearly. This should be corrected.

Point 9) Page 3. It is never specified if human-virus PPIs from databases correspond to direct or indirect physical interactions. Authors need to discriminate these two situations in their review. Indeed, a vast majority of PPIs identified by co-affinity purification followed by MS identification correspond to large protein complexes and indirect interactions. Furthermore, this is directly related to a key question raised by the authors on page 3: “…we need to ask if it is biologically meaningful or even possible that a protein has >100 interactions. ». The answer is yes if interactions were identified by co-AP.

Point 10) Table 3. DenHunt (http://proline.biochem.iisc.ernet.in/DenHunt/) or the most recent DenvInt (https://denvint.000webhostapp.com) are dedicated databases for dengue virus and should be included in Table 3.

Point 11) Page 3. “however small” should be deleted.

Point 12) Page 3. Authors cite their work on KSHV-human interactions as an example of “validation experiment”. This is irrelevant because interactions were predicted by computational analysis, and then tested by co-IP. This is not an example of PPIs identified by high-throughput screening, and validated with a secondary assay. Other works should be cited where large datasets of virus-host PPIs identified in a primary screen were validated using either biochemical (PMID:22898364; PMID:21994455 ; PMID:18985028 ; PMID:23816991) or functional (PMID:25464832) secondary assays. In addition, the overall quality of PPIs obtained by different methods or available in literature and databases has been evaluated (PMID:19060904; PMID:19116613). This should be discussed as the overall quality of virus-host PPIs is probably equivalent to host-host PPIs available in databases or literature.

Point 13) “How many interactions does a virus require?”. It is surprising that authors used bacteriophages as a model for this estimation, as the review is essentially focusing on human viruses. There is no attempt to use the “large number” of host-interactions available for human viruses, as stated by the authors, to propose some estimation. Furthermore, the statement that “they are likely to contain a large number of false positives” is gratuitous and not supported. Finally, it should be mentioned that viral proteins are highly enriched for disordered regions that are known to favor interactions with many partners (PMID:19062293).

Point 14) Page 5. « Of the over 17,000 HIV-1 – human PPIs reported in HIV-1db as of August, 2017, fewer than 7,000 are direct physical interactions. Many interactions are indirect, e.g. genetic interactions or inferred from mutagenesis experiments.». This is confusing because physical interactions can be direct or indirect (see Point 9). This needs to be rephrased.

Point 15) Page 5. “Notably, we collected 5,495 interactions with human host proteins involving proteins of orthomyxoviridae.” Again, are these interactions essentially direct or indirect?

Point 16) Since Fig. 2A, 2C and 2D correspond to enrichment values, it is expected that values >1 correspond to enrichment, whereas values <1 correspond to an under representation. This is apparently not the case, and should be better explained in the figure legend. Did the authors used log-values in these three figures?

Point 17) Page 7. “The virus interactome-diseasome connection”. Authors should discuss in this chapter the paper from Rozenblatt-Rosen O. & al. (PMID:22810586) where this interactome-diseasome connection is directly addressed.

Point 18) Page 7. “Virus interactions with the host metabolome”. NS5A from HCV was shown to bind and to activate the hexokinase 2, thus increasing glucose uptake and glycolysis. This is an example of a direct interaction between a viral protein and a metabolic enzyme (PMID:24390321).

Reviewer 4

While the purpose of this article is to provide a thorough overview about virus–host protein-protein interactions (PPIs) and on relevant sources of biological data for studying virus–host PPIs, I find the description of the various sections incomplete, general, confusing and sometime irrelevant, without providing the current state of knowledge. Overall I find that the article unstructured, inaccurate with description of studies that seem to be selected with a weak rational for expressing an opinion. Few section are correctly emphasizing conclusion of previous review studies with limited added value. The section on bacteriophages as model is adding to the confusion of better defining the number of physiological interactions per proteins from data of large scale studies. Indeed, it may not be a good model for human virus as highlighted by the high false negative rate due to protein post-translational modifications and high protein maturation required in bacteriophages. The section is closing with a statement that no comprehensive analysis or even a review of published interactions among human and virus proteins has attempted to evaluate these interactions for their plausibility or physiological relevance. Finally the conclusion and outlook are hardly based on the description of the article.

More specifically,

 The abstract as it stands is incomprehensible with many sentences that are meaningless.

 The Introduction section is quite irrelevant without providing any current state of knowledge for the different topics of the review and updated references. While the introduction about virus diversity makes the point that for some viruses like Zika there are no extensive data, Table 1 highlights Zika in the 20 best-studied viruses of the Flaviviridae (and how about HCV?).

The following sections:

Diversity of human viruses and Virus-host interactions in humans

- Unstructured and no rational for the description of the text based mainly on ref. 1

- Does not take into consideration blood borne viral infections

- Doubt on the accuracy of Tables 1, 2 and 4 (why is the number of interactions per individual viruses is often higher when added than the number in the column HPIs?)

- The numbers of PPIs are not up-to-date (for Flaviviridae it is much higher than 575 HPI)

- There is no description of the various orthogonal approaches or limitations of proteomics studies

How reliable are published virus-host interactions?

Authors discuss the value of homologous interactions to validate 68% PPIs and describe one PPI study but claimed a biased validation of Y2H data set. It is not clear what the authors want to achieve in this section and their conclusion. HTS screens generate unlikely data of identified PPIs

**Decision Letter that was sent:**

Ref: YSCDB\_2017\_345  
Title: Virus-host protein-protein interactions and human disease  
Journal: Seminars in Cell and Developmental Biology

Dear Dr. Uetz,

Thank you for submitting your manuscript to Seminars in Cell and Developmental Biology. I apologize for the delay in getting back to you. I have now received comments from 4 different reviewers on your manuscript. They raised a number of major issues with your manuscript and some recommended rejection. However, I will consider a revised manuscript  in light of the appended reviewer comments, that addresses the major points.

Please submit your revised manuscript by the 1st of March. When resubmitting your manuscript, please carefully consider all issues mentioned in the reviewers' comments, outline every change made point by point, and provide suitable rebuttals for any comments not addressed. I would also understand if you are unwilling to make such substantial changes in this time period.

**What happens next?**

* Log into EVISE® at: <http://www.evise.com/evise/faces/pages/navigation/NavController.jspx?JRNL_ACR=YSCDB>
* Locate your manuscript under the header 'My Submissions that need Revisions' on your 'My Author Tasks' view
* Confirm whether you want to 'Agree to Revise' or 'Decline to Revise'.
* If you choose to 'Agree to Revise', a pop up will appear. Please click on the confirm button on this pop up.
* After you 'Agree to Revise' you can either start submitting your revision or you can logout and submit your revision at any time within your due date.

While revising your manuscript please consider to incorporate the necessary data and files that are supported by the [Content Innovations](https://www.elsevier.com/books-and-%20journals/content-innovation) available for Seminars in Cell and Developmental Biology.

**Enrich your article to present your research with maximum impact.** This journal supports the following [Content Innovations](http://www.elsevier.com/about/content-innovation):

* [Antibody Data](http://www.elsevier.com/books-and-journals/content-innovation/antibody-data/): Reference application linking to information about the antibodies mentioned in the article – based on the [NIF Antibody Registry](http://antibodyregistry.org/). Please include relevant Antibody Registry accession numbers in your article following [this syntax](http://www.elsevier.com/databaselinking) and other [relevant identifiers](http://www.elsevier.com/books-and-journals/content-innovation/minimal-data-standards/) if appropriate.
* Explain your research in your own words and attract interest in your work using [AudioSlides](http://www.elsevier.com/audioslides" \t "_blank) : 5-minute webcast-style presentations that are displayed next to your published article and can be posted on other websites. You will receive an invitation email to create an AudioSlides presentation within three weeks after your paper has been accepted.
* [Interactive Plots](http://www.elsevier.com/books-and-journals/content-innovation/iplots/): Interactive plot viewer providing easy access to the data behind plots. Please prepare a [.CSV](http://en.wikipedia.org/wiki/Comma-separated_values) file with your plot data and test it online [here](http://authortools.elsevier.com/interactiveplots/verification) before submitting as supplementary material.

I look forward to receiving your revised manuscript as soon as possible.

Kind regards,

Dr Sierecki  
Guest Editor   
Seminars in Cell and Developmental Biology

**Comments from the editors and reviewers:**

-**Reviewer 1**

The authors discuss the nature of virus protein-protein interaction information that is currently available in various databases, and how these large-scale data might be mined.  I have several issues with this manuscript.

Firstly, according to the request to review that I received, it is my understanding that this paper is an invited review article, yet it appears to contain primary research material – namely the analyses presented in figure 1,2 and 3.  The structure of the paper is consistent with a review article, but without a fully described method section, I feel it would be inappropriate to publish these analyses in the current format.

The writing is often less formal than is found in many scientific manuscripts.  In most cases, I enjoy this because it improves the readability.  However, there are some phrases, I would consider revising: ‘It turns out…’ and ‘Virus proteins attack…’ are two I find too colloquial.  I also found that acronyms are frequently not defined, and are sometimes defined on the second use.  In particular, PPI in the abstract is not defined anywhere.  As many viruses form protein-protein interactions with peptidyl-prolyl isomerases, PPI as an acronym is confusing if not defined.  I would suggest, as it is not used consistently throughout the manuscript, that PPI be removed.  I also find some of the comments in the introduction and conclusions to be generic and unhelpful (if not inaccurate).  The assertion in the introduction that viruses are more difficult to control than bacteria due to the use of antibiotics is a very debatable statement, and given that the paper is not about bacteria, this statement is unnecessary.  In the conclusion: ‘Viruses evolve much quicker than their hosts, especially RNA viruses, hence viruses can also adapt their virus-host interactions faster than a host can react by mutating its target proteins, including its immune system.’  This statement is at odds with the Red Queen hypothesis – that virus and host co-evolve to maintain mutual survival.  Indeed the adaptive arm of the host immune system can, in many cases, actively evolve to eliminate infection faster than the virus can respond.  Again, I’m not sure of the value of this statement in the context of the conclusions.

I am also concerned about the oversimplification of some of the biology in this paper.  In Table 1, viruses are classified as either DNA or RNA (or ‘anim’ or ‘plant’ is some cases – why?).  The Baltimore system of classification of viruses defines seven types of genome: dsDNA, ssDNA, dsRNA, ssRNA(+ve), ssRNA(-ve), ssRNA-RT, and dsDNA-RT.  These different groups have different replication requirements which impacts on their mutation rate (amongst other things).  In table 1, hepdnaviridae are labeled as DNA viruses – which isn’t strictly true as they have an RNA intermediate.  Likewise, retroviridae have a DNA intermediate.   So again, I am unsure as to the value of the describing the virus families in terms of being RNA or DNA, as is done in Table 1.

In Figure 2 (and the accompanying section ‘The topology of human-virus interactions’) the authors describe critical nodes in interaction networks as statistically more likely to be targets from protein interactions.  Intuitively, this seems reasonable, and I like it as a concept.  But could it also be that these critical nodes are also overrepresented in false-positives in large-scale interactome analyses?

In Figure 3, I am intrigued to see that retroviridae and orthomyxoviridae have a negative fold change with respect to cell wall proteins, while herpesviridae and papillomarviridae have a positive fold change.  None of these families of virus infect an organism with a cell wall.  This calls into question whether the authors understand the systems they are studying.  Also, there appears to be a phylogenetic analysis on both axes of the heat map, but this is not described or mentioned.  I also question the value of putting ‘other’ in the heat map.  Given the assertion that different viruses have different interaction preferences, surely grouping everything that is not of the aforementioned families together would result in a largely featureless map, averaged around no change – which is close to what is shown.

On the whole, I find this paper to be strangely constructed.  It is a hybrid review/original research piece, with a smattering of ideas and concepts, with limited flow between sections.  The conclusions of the paper are particularly unremarkable.  The statement ‘…suggesting that different viral families use different strategies to invade a human host cell’ to be one of the most unprofound statements I have seen in a manuscript.  As the conclusions read, the ‘take home message’ appears to be that: there are many interactome databases; some viruses are overrepresented and full of false positives; other viruses are underrepresented; and viruses are different from each other.  So I would question what the value of such a paper would have to the scientific community.

-**Reviewer 2**

In this manuscript, Goodacre and colleagues broadly examine protein-protein interactions (PPIs) in the context of human virus infections and discuss the physiological relevance of these interactions. While a large number of PPIs have been identified, the authors raise the question of how many interactions are physiological and required for viral infection. Interestingly, there is considerable overlap in the human proteins targeted by different virus families. The authors describe the databases available for the study of human-virus PPIs and provide a useful summary of the strengths and weaknesses of the databases and their potential uses. Overall, this work is likely to be of interest to a broad readership in virology and addresses some critical questions as the use of systems biology approaches becomes more prevalent.

Specific comments:

1.   When questioning the reliability of published PPIs, the authors describe cases where the predicted interactions were validated by co-immunoprecipitation and GST pull-down experiments. However, another critical question is whether these interactions been shown to have a functional effect on viral infection (i.e. by examining viral replication when the interaction is blocked by knockdown, small molecule inhibitors, or genetics approaches). Extending this section to describe and provide examples of functional effects would strengthen the manuscript.

2.     Could the authors comment on potential effects of different cell lines in terms of PPIs? One problem is that the cell lines used to study viral infection are not always physiological (e.g., the use of cancer cell lines, where signaling pathways and protein expression levels may be disregulated). One can imagine that PPIs identified in such cell lines may be even less likely to accurately reflect the physiological situation. Some discussion of this, or providing some examples, would be appreciated.

3.     Many viruses infect different types of hosts (e.g., Dengue is spread by mosquitoes, Influenza infects birds and pigs). Have any studies compared PPIs in different hosts for viruses such as these? While the reviewer appreciates this manuscript is focused on human disease, it may be interesting to include a brief section about PPIs in other hosts for this group of viruses, many of which are medically important human pathogens.

4.     More discussion of the outlook and future directions would be useful for researchers in the field. How can we now exploit this collection of PPIs to target viral infections or diseases associated with the viral interactome? How could databases be improved? How can we better ensure that identified virus-human PPIs are biologically relevant?

Minor comments:

1.     Abstract, line 10: insert comma “…relevant, for many…”

2.     Page 2, Introduction, paragraph 1, line 4: insert comma “…and structure, viruses…”

3.      Page 3, Virus-host interactions in humans, line 9: change to “…more than 100 interactions…” instead of “more than a 100”

4.     Page 6, line 2: “Betweenes centrality” should be “Betweenness centrality”?

5.     Page 7, Virus interactions with the host metabolome, line 4: change to “…with the cancer protein…” instead of “th cancer”

-**Reviewer 3**

This manuscript provides a global survey of virus-host PPIs in literature and databases. Recent advances and limitations of these resources are discussed. This is an update of the current situation, but some paragraphs have significant weaknesses. Several points should be addressed.

Point 1) In page 2 (second paragraph), some sentences are misleading. Indeed, it could be wrongly understood that « medical interest » is somehow related to higher « infectivity », which could be correlated with higher « pathogenicity » (“In fact, …”). This is confusing and should be clarified.

Point 2) Table 1. Which numbers were used to calculate the U/C ratio? For example, 1008 total genomes are available for Paramyxoviridae and clustered into 327 groups. However, the U/C ratio is 10.06.

Point 3) Table 1: Arenaviridae are responsible for major human diseases such as Lassa fever, and should not be restricted to “animals”.

Point 4) In the legend of Table 1, it is said “Clustered sequenced were clustered at ≥98% sequence identity”. This is different from the main text (Page 2; sequences are clustered and those genomes combined which have less than 5% sequence…”).

Point 5) Page 2: “We have thus compiled data on the diversity of human viruses by looking both at sequence diversity (Table 1)…”. Table 1 compiled data on sequence diversity inside viral families, not “viruses” or “human viruses”. This should be corrected.

Point 6) Table 2: The viral family should be indicated for each virus.

Point 7) Page 3: “Coronaviruses, are highly pathogenic with a lot of victims, but have received rather scant attention”. This is contradictory with data presented in Table 2, since coronaviruses were responsible for a very small number of victims (<10.000 for SARS-CoV and MERS-CoV over several years). In fact, dengue virus, hRSV or rabies virus that are not cited in Table 2 have a much higher impact on human populations.

Point 8) Table 2: “Unless otherwise indicated, figures are yearly”. Data presented for SARS-CoV and MERS-CoV are not yearly. This should be corrected.

 Point 9) Page 3. It is never specified if human-virus PPIs from databases correspond to direct or indirect physical interactions. Authors need to discriminate these two situations in their review. Indeed, a vast majority of PPIs identified by co-affinity purification followed by MS identification correspond to large protein complexes and indirect interactions. Furthermore, this is directly related to a key question raised by the authors on page 3: “…we need to ask if it is biologically meaningful or even possible that a protein has >100 interactions. ». The answer is yes if interactions were identified by co-AP.

Point 10) Table 3. DenHunt (<http://proline.biochem.iisc.ernet.in/DenHunt/>) or the most recent DenvInt (*https://denvint.000webhostapp.com*) are dedicated databases for dengue virus and should be included in Table 3.

Point 11) Page 3. “however small” should be deleted.

Point 12) Page 3. Authors cite their work on KSHV-human interactions as an example of “validation experiment”. This is irrelevant because interactions were predicted by computational analysis, and then tested by co-IP. This is not an example of PPIs identified by high-throughput screening, and validated with a secondary assay. Other works should be cited where large datasets of virus-host PPIs identified in a primary screen were validated using either biochemical (PMID:22898364; PMID:21994455 ; PMID:18985028 ; PMID:23816991) or functional (PMID:25464832) secondary assays. In addition, the overall quality of PPIs obtained by different methods or available in literature and databases has been evaluated (PMID:19060904; PMID:19116613). This should be discussed as the overall quality of virus-host PPIs is probably equivalent to host-host PPIs available in databases or literature.

Point 13) “How many interactions does a virus require?”. It is surprising that authors used bacteriophages as a model for this estimation, as the review is essentially focusing on human viruses. There is no attempt to use the “large number” of host-interactions available for human viruses, as stated by the authors, to propose some estimation. Furthermore, the statement that “they are likely to contain a large number of false positives” is gratuitous and not supported. Finally, it should be mentioned that viral proteins are highly enriched for disordered regions that are known to favor interactions with many partners (PMID:19062293).

Point 14) Page 5. « Of the over 17,000 HIV-1 – human PPIs reported in HIV-1db as of August, 2017, fewer than 7,000 are direct physical interactions. Many interactions are indirect, e.g. genetic interactions or inferred from mutagenesis experiments.». This is confusing because physical interactions can be direct or indirect (see Point 9). This needs to be rephrased.

Point 15) Page 5. “Notably, we collected 5,495 interactions with human host proteins involving proteins of orthomyxoviridae.” Again, are these interactions essentially direct or indirect?

Point 16) Since Fig. 2A, 2C and 2D correspond to enrichment values, it is expected that values >1 correspond to enrichment, whereas values <1 correspond to an under representation. This is apparently not the case, and should be better explained in the figure legend. Did the authors used log-values in these three figures?

Point 17) Page 7. “The virus interactome-diseasome connection”. Authors should discuss in this chapter the paper from Rozenblatt-Rosen O. & al. (PMID:22810586) where this interactome-diseasome connection is directly addressed.

Point 18) Page 7. “Virus interactions with the host metabolome”. NS5A from HCV was shown to bind and to activate the hexokinase 2, thus increasing glucose uptake and glycolysis. This is an example of a direct interaction between a viral protein and a metabolic enzyme (PMID:24390321).

-**Reviewer 4**  
  
  - While the purpose of this article is to provide a thorough overview about virus–host protein-protein interactions (PPIs) and on relevant sources of biological data for studying virus–host PPIs, I find the description of the various sections incomplete, general, confusing and sometime irrelevant, without providing the current state of knowledge. Overall I find that the article unstructured, inaccurate with description of studies that seem to be selected with a weak rational for expressing an opinion. Few section are correctly emphasizing conclusion of previous review studies with limited added value. The section on bacteriophages as model is adding to the confusion of better defining the number of physiological interactions per proteins from data of large scale studies. Indeed, it may not be a good model for human virus as highlighted by the high false negative rate due to protein post-translational modifications and high protein maturation required in bacteriophages. The section is closing with a statement that no comprehensive analysis or even a review of published interactions among human and virus proteins has attempted to evaluate these interactions for their plausibility or physiological relevance. Finally the conclusion and outlook are hardly based on the description of the article.

More specifically,  
• The abstract as it stands is incomprehensible with many sentences that are meaningless.  
• The Introduction section is quite irrelevant without providing any current state of knowledge for the different topics of the review and updated references. While the introduction about virus diversity makes the point that for some viruses like Zika there are no extensive data, Table 1 highlights Zika in the 20 best-studied viruses of the Flaviviridae (and how about HCV?).  
The following sections:  
Diversity of human viruses and Virus-host interactions in humans  
- Unstructured and no rational for the description of the text based mainly on ref. 1  
- Does not take into consideration blood borne viral infections  
- Doubt on the accuracy of Tables 1, 2 and 4 (why is the number of interactions per individual viruses is often higher when added than the number in the column HPIs?)  
- The numbers of PPIs are not up-to-date (for Flaviviridae it is much higher than 575 HPI)  
- There is no description of the various orthogonal approaches or limitations of proteomics studies  
How reliable are published virus-host interactions?  
Authors discuss the value of homologous interactions to validate 68% PPIs and describe one PPI study but claimed a biased validation of Y2H data set. It is not clear what the authors want to achieve in this section and their conclusion. HTS screens generate unlikely data of identified PPIs

**Have questions or need assistance?**  
For further assistance, please visit our [Customer Support](http://help.elsevier.com/app/answers/list/p/9435/) site. Here you can search for solutions on a range of topics, find answers to frequently asked questions, and learn more about EVISE® via interactive tutorials. You can also talk 24/5 to our customer support team by phone and 24/7 by live chat and email.

-------------------------------------------------------------

Copyright © 2018 Elsevier B.V. | [Privacy Policy](https://www.elsevier.com/legal/privacy-policy)

Elsevier B.V., Radarweg 29, 1043 NX Amsterdam, The Netherlands, Reg. No. 33156677.