nature research

Peer Review File

Manuscript Title: The Pathogenicity of SARS-CoV-2 in hACE2 Transgenic Mice

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee #1 (Remarks to the Author):

In this manuscript, Bao et al present data showing that transgenic expression of hACE2 renders mice susceptible to 2019-nCoV and that mice develop histological changes after infection that are consistent with pneumonia. Viral antigen can be detected in cells expressing hACE2. The infection is very mild so mice do not develop clinical disease except for 5% weight loss. While the experiments demonstrate that 2019-nCoV causes pneumonia on histological examination, clinical disease is not demonstrated so that Koch's postulates are only partly fulfilled. The results are preliminary in other aspects as well.

Specific comments.
1. The experiments are performed with a minimum number of mice and several important controls/pieces of data are missing.
a. Mice should be examined for infectious virus and genomic RNA at 24 hours p.i., to demonstrate that virus levels detected at later times reflect active replication and not residual input inoculum.
b. Virus titers in the lungs at days 3 and 5 should be shown, in addition to the RT-PCR results. It appears that 2 mice were assessed for virus at days 3 and at days 5. Data from both mice should be shown.
c. Figure 1b-The strongest PCR product appears to be in the liver but according to the text, viral RNA and infectious virus were detected only in the lungs (which has a less prominent band). PCR product was identified in many other tissues as well. The text states that only the lung PCR product was virus-specific. This requires clarification. Since viral RNA was detected in many tissues, the data in Figure 1b should be quantified and data from all (day 3 and day 5) mice that were studied should be shown.
2. According to the literature, hACE2 transgenic mice infected with SARS-CoV developed a lethal encephalitis. It is surprising that Infection with 2019-nCoV does not cause a lethal encephalitis. This lack of encephalitis may occur because the virus does not bind as well to the ACE2 receptor. This should be discussed.
3. Line 46 and methods-The transgenic hACE2 mice should be described in more detail. In what tissue was hACE2 expressed and what are the relative levels of expression?
4. Why were 6-11 month old mice used? Did this reflect availability?
5. Figure 3-Histological changes a day 3 are shown in the figure. Were similar findings observed at day 5? Were similar histological changes found in all infected mice?

Referee #2 (Remarks to the Author):

This manuscript by Bao and colleagues describes the first data from an animal model for 2019-nCoV. Animal models are an essential part of outbreak response as they are crucial for the understanding of the disease and the testing of the efficacy of medical countermeasures. However, even during a public health emergency, we have to ensure the quality of data that are published. This is especially true of animal studies. Despite the importance of animal model development for 2019-nCoV, I find the data in this manuscript lacking sufficient quality to warrant publication in Nature. I have specified some of the issues below.

It is not clear which mice were used in this study, nor how they were generated. The description in lines 113-115 is not clear. Were two different genetic backgrounds of mice used in these experiments? If so, how were they divided over the experiments? It is also not clear what genetic background was used for the WT mice. A different genetic background could explain why the WT-HB-01 mice are gaining weight so much faster than the ACE2-PBS mice.

The animal number is too low. Two animals per timepoint makes it impossible to perform statistical analysis.

The authors state that they determine 'viral load'. They did not. They only confirmed the presence of viral RNA in tissues. The authors do not address why there is a much bigger band in the liver than in the lungs. Was there any histological evidence of virus replication outside the respiratory tract?

It would have been interesting if swabs had been collected from these animals to look at virus shedding from the respiratory and intestinal tract.

How did the authors select mice for euthanasia on 3 and 5 dpi? Based on the bodyweight data it appears as though the animals with weight loss were specifically selected.
It would be better to display weights of individual animals. The error bars are quite large, suggesting that some of the infected ACE2 mice may not have lost weight, whereas others may have lost more than 5%.

The gross lung lesion pointed out in Figure 2 is not very convincing. How many animals showed gross lesions on each time point? Is this the only 'lesion' that was observed?

It is very hard to draw any conclusions from the histology pictures. The lungs were not insufflated before fixation resulting in all the lungs being collapsed making it very hard to analyze these panels. The tissues were fixed inappropriately which appears to have resulted in some autolysis. This makes it hard to distinguish cell necrosis due to autolysis (resulting in many visible nuclei) from the influx of inflammatory cells (usually observed as many purple nuclei). Normally, one would also observe perivascular and peribronchiolar edema in case of bronchointerstitial pneumonia, but this cannot be observed in the slides from nCoV-infected ACE2 mice. The authors have included staining specific for 2019-nCoV, but when defining a new model it is essential to show that the expression of viral antigen overlaps with specific lesions. So nonfluorescent antibody stain in a counterstained slide needs to be included instead of the histology panels.

The language needs attention. E.g. the term 'infection' is used inappropriately in several instances.

Referee #3 (Remarks to the Author):

Key results: In this manuscript, the authors experimentally inoculate genetically modified mice expressing the receptor for SARS-CoV (human ACE2) on their cell surfaces, in order to determine the pathogenicity of SARS-CoV-2 (the formal name of the new coronavirus) and to fulfill Koch's postulates for COVID-19 (the formal name of the disease caused by infection with SARS-CoV-2). The most outstanding features of this work are the demonstration of body weight loss and pneumonia associated with SARS-CoV-2 infection.

Validity: The manuscript has important flaws.
1) The authors do not reach their aims, which are to clarify the pathogenicity and to determine whether Koch's postulates are fulfilled. They make no statement about the pathogenicity (that is, the virulence) of SARS-CoV-2. They conclude that the results of their study fulfill Koch's

postulates, but they do neither specify what these postulates are nor indicate how the results of their study fulfill them. To elaborate on these Koch's postulates, they were originally used for bacteria, not viruses. Thomas Rivers adapted them for viruses in a 1937 paper in J. Bact. Although not very clearly indicated in the paper, the adapted postulates could be summarized as:
1. isolation of virus from diseased hosts
2. cultivation in host cells
3. proof of filterability
4. production of comparable disease in the original host species or a related one
5. re-isolation of the virus
6. detection of a specific immune response to the virus
Although the house mouse is not very closely related to the human being (postulate 4), it is a mammal, and the authors do show a comparable disease (pneumonia) as seen in affected people (ground-glass opacities in the lung on CT scan, consistent with interstitial pneumonia). The authors also re-isolate the virus from lung tissue of the inoculated mice. Therefore, if the mouse is considered sufficiently related to the human being for the purposes of these postulates, they have fulfilled postulates 4 and 5 of above list. Postulates 1 to 3 have been reached by others prior to this study, and postulate 6 is not yet fulfilled.

2) The results of pathological analysis, which are the main results of this manuscript, are inadequately performed and described. The use of frozen sections, embedding in Cryo-Gel, and sectioning at 10 micrometers results in sections that are too thick to see enough histological details. A better method would be formalin fixation, paraffin embedding, and sectioning at 3 micrometers. In addition, panels d to i of Figure 2 indicate that the haematoxylin and eosin staining is not correct. Specifically, the eosin staining is too weak. Correct eosin staining would show the erythrocytes, which must be present in these sections, as bright red circles, but they are not visible at all in these photographs. Because of the thick sections and the poor staining, it is not possible to describe the details of the pathological changes at sufficient detail. For example, it is not possible to determine whether there is oedema fluid in the alveolar lumina (which would be expected in such a lesion), nor the identity of the inflammatory cells that are presumably infiltrating the affected tissue.

3) There is inadequate support from pathological analyses for SARS-CoV-2 as the cause of the observed pneumonia. The key evidence for this is colocalisation of virus antigen expression and histopathological changes. The authors do not show this. All they show, in Figure 3q is colocalisation of virus antigen expression and hACE2 expression.

4) There is inadequate support from the photos of the gross lesions (panels a to c in Figure 2) for the presence of pneumonia. The arrow in panel c is pointing at a supposed area of pneumonia, but this cannot be identified as such. It appears that the lungs in panels a to c are collapsed, a normal phenomenon that occurs after the lungs are taken out of the thoracic cavity (which is under vacuum during life) at autopsy. This makes identification of subtle gross lesions difficult. A possible solution for this, in future autopsies, is to close the trachea by use of a small haemostat clamp before opening the thoracic cavity and removing the lungs. In this way, the lungs remain filled with air and it is much easier to distinguish normal from abnormal lung tissue. The description of the gross lesions is inadequate, both qualitatively and quantitatively. Qualitatively, description of the gross lesion should include colour and whether lesions were raised, level, or depressed compared to surrounding tissue. Quantitatively, the affected area of lung and the weight of the lungs (compared to the negative controls) would be suitable measures to estimate the proportion of affected lung tissue.

Originality and significance: The results are original. SARS-CoV-2 is a recently identified virus, and to my knowledge no studies on experimentally infected laboratory animals have been published. The results of such studies, if performed adequately, would be of interest to people from several disciplines.

Data & methodology: There is inadequate support from the results that "viral RNA was positive by RT-PCR (Figure 1b) and identified by sequencing only in the lung of infected-hACE2 mice at 3 dpi and 5 dpi". The cited panel (b of Figure 1) shows bands at the same level as the positive control not only for the lung sample, but also for the spleen and liver samples. Whether there are bands at the same level as the positive control for intestine, kidney, and brain samples is not clear because the resolution of the photo is too low. The authors state "identified by sequencing", which presumably indicates that they sequenced the PCR products in the bands cut out of the gel, but they do not state the sequencing method in the methods section. Because these mice presumably express hACE2 on cells throughout the body, it is possible that SARS-CoV-2 replicates in other tissues than the lungs, and therefore it would not be unexpected to detect viral RNA in other tissues. In this regard, it is important to also specifically state what the results of histopathology and immunofluorescence in all the organs tested by RT-PCR. The authors state that "the major organs of mice were examined by histopathology and immunofluorescence", but only show results for the lung.

Appropriate use of statistics and treatment of uncertainties: P-values are indicated clearly in Figure 1A. Nature requires requires detailed statements on statistical tests, some of which have not been fulfilled. It has not been stated whether the t-test used is one- or two-sided. It also has not been stated that it is assumed that both populations have normal distributions, which is a requirement for Welch's t-test. Also, the test statistic (t) with degrees of freedom and P value (exact value) are not noted.

References:
Since this paper refers to Koch's postulates—and more precisely, River's modifications for viral diseases—a reference to these postulates should be included: Rivers, T. M. J. Bacteriol. 33, 1–12 (1937).

Minor comments:
Line 50: It is not clear to which group of mice the statement "n = 7" refers.
Line 50: Please state the inoculum volume, whether and how mice were anaesthetised for inoculation, and how they were held.
Line 52: Please clarify what is meant by "for 10 dpi".
Line 53: Please clarify whether you used a score card to record clinical signs, and which signs were specifically checked.
Line 54: Please list which major organs were assessed for distribution of SARS-CoV-2.
Line 56: Please clarify what is meant by "by sequencing", and provide methodological details.
Line 57: Please state from which tissue virus was isolated. Please indicate whether it was confirmed, e.g. by RT-PCR, that the cytopathic effect and the observed coronavirus particles corresponded to SARS-CoV-2.
Line 67: Please clarify what is meant by "interstitial hyperplasia". Specifically, which cell types were involved in this hyperplasia?
Line 67: Please specify which cell types comprised these inflammatory cells.
Line 69: Were no pathological changes observed in alveolar epithelial cells?
Line 70: Please specify which cell types comprised these leukocytes.
Line 72: Was is possible to distinguish between type 1 and type 2 alveolar epithelial cells?
Line 161: Please indicate whether negative controls, positive controls, and omission controls were used, and provide details about these controls.
Supplementary figure 1: An explanation for the asterisk is missing.

Author Rebuttals to Initial Comments:

Referee #1 (Remarks to the Author):

In this manuscript, Bao et al present data showing that transgenic expression of hACE2 renders mice susceptible to 2019-nCoV and that mice develop histological changes after infection that are consistent with pneumonia. Viral antigen can be detected in cells expressing hACE2. The infection is very mild so mice do not develop clinical disease except for 5% weight loss. While the experiments demonstrate that 2019-nCoV causes pneumonia on histological examination, clinical disease is not demonstrated so that Koch's postulates are only partly fulfilled. The results are preliminary in other aspects as well.

Author response: Thank you for your insightful comment. We have fully considered your advice and answered your questions point by point.

Specific comments.

1. The experiments are performed with a minimum number of mice and several important controls/pieces of data are missing.

a. Mice should be examined for infectious virus and genomic RNA at 24 hours p.i., to demonstrate that virus levels detected at later times reflect active replication and not residual input inoculum.

Author response: Thank you for your valuable comment. We have supplemented the viral loads systematically including seven primary tissues (Figure 1b). We particularly compared viral titers in lungs from WT-HB-01, ACE2-MOCK and ACE2-HB-01 mice along the timeline (Figure 1c). From the time-dependent changes, it was clearly shown that the virus actively replicates in lung, but not residual input inoculum. In addition, although the viral loads were detected in the intestine at 1 dpi, no virus in intestine was isolated, which was speculated as the residue of the virus input inoculum from the inoculation site of the nasal mucosa to intestines by swallowing.

b. Virus titers in the lungs at days 3 and 5 should be shown, in addition to the RT-PCR results. It appears that 2 mice were assessed for virus at days 3 and at days 5. Data from both mice should be shown.

Author response: Thank you for your valuable comment. During the stage of under-review, we have replenished the relative work and highly improved the abundance of the experiments. We have replenished the experimental animals up to 49. We observed the clinical manifestation

(Figure 1a) and detected the specific antibody (Figure 1e) in 13 mice, including ACE2-HB-01 mice (n=7), ACE2-mock mice (n=3) and WT-HB-01 mice (n=3). We euthanized 12 mice per group at 1 dpi (n=3/group), 3 dpi (n=3/group), 5 dpi (n=3/group) or 7 dpi (n=3/group) to obtain time-dependent viral replication and pathological changes undergone virus infection. We have supplemented the data of viral titers at 1 dpi, 3 dpi, 5 dpi and 7 dpi in the revised manuscript (Figure 1c). We demonstrated individual's data per timepoint to provide as abundant information as we have.

c. Figure 1b-The strongest PCR product appears to be in the liver but according to the text, viral RNA and infectious virus were detected only in the lungs (which has a less prominent band). PCR product was identified in many other tissues as well. The text states that only the lung PCR product was virus-specific. This requires clarification. Since viral RNA was detected in many tissues, the data in Figure 1b should be quantified and data from all (day 3 and day 5) mice that were studied should be shown.

Author response: Thank you for your valuable comment. We have sequenced all PCR products and confirmed that the viral PCR could only been found in tissue from lungs. Remained PCR products were believed as non-specific bands during experiments as it could not validated by sequencing. Moreover, we have presented the time-dependent viral loads (Figure 1b) in all primary tissues along detecting timeline (1 dpi, 3 dpi, 5 dpi and 7 dpi) according to your suggestions.

2. According to the literature, hACE2 transgenic mice infected with SARS-CoV developed a lethal encephalitis. It is surprising that Infection with 2019-nCoV does not cause a lethal encephalitis. This lack of encephalitis may occur because the virus does not bind as well to the ACE2 receptor. This should be discussed.

Author response: Thank you for your insightful advice. We have discussed this issue in the revised manuscript (Line 146-152).

3. Line 46 and methods-The transgenic hACE2 mice should be described in more detail. In what tissue was hACE2 expressed and what are the relative levels of expression?

Author response: Thank you for your valuable comment. The transgenic hACE2 mice have been previously generated and stored in our lab. The details on transgenic mice modeling and validating of hACE2 expression could be found in publication (*Yang et al 2007*). We have cited this publication. Moreover, we have replenished the key details on model construction in Methods section in the revised manuscript (Line 180-186).

4. Why were 6-11 month old mice used? Did this reflect availability?

Author response: Thank you for your valuable comment. We have performed the experiments using 4-5 months old mice after the manuscript was submitted. We have observed the similar clinical manifestation and pathogenic changes. We compared the preliminary data between 6-11-month mice and 4-5-month mice and found no significant disparity. Herein, 6-11-month mice were used in our experiment.

5. Figure 3-Histological changes a day 3 are shown in the figure. Were similar findings observed at day 5? Were similar histological changes found in all infected mice?

Author response: Thank you for your valuable comment. We have replenished the histopathological changes at 5 dpi according to your suggestion. All the ACE2-HB-01 mice at 3 dpi and 5 dpi displayed comparable gross lesions and histopathological changes, respectively (Figure 2a)(in the revised manuscript). The lesions were progressed from 3 dpi to 5 dpi revealing different stages of the pathogenesis after SARS-CoV-2 infection. Specifically, consistent with the gross lesions, lung tissues from ACE2-HB-01 mice at 3 dpi had moderate, multifocal, interstitial pneumonia, and the lung progressed into coalescing interstitial pneumonia with diffuse lesions at 5 dpi.

Referee #2 (Remarks to the Author):

This manuscript by Bao and colleagues describes the first data from an animal model for 2019-nCoV. Animal models are an essential part of outbreak response as they are crucial for the understanding of the disease and the testing of the efficacy of medical countermeasures. However, even during a public health emergency, we have to ensure the quality of data that are published. This is especially true of animal studies. Despite the importance of animal model development for 2019-nCoV, I find the data in this manuscript lacking sufficient quality to warrant publication in Nature. I have specified some of the issues below.

Author response: Thank you for your critical comment. We apologize that we have submitted preliminary data with insufficient abundance as the we attempt to provide valuable information to an emergent epidemic as soon as we could. During the stage of under-review, we have replenished the relative work and highly improved the abundance of the experiments. We highly appreciated the insightful comments you have addressed and we have revised the manuscript completed. We are very appreciated if you could consider the updated edition for another time. Thank you for your patience.

1) It is not clear which mice were used in this study, nor how they were generated. The description in lines 113-115 is not clear. Were two different genetic backgrounds of mice used in these experiments? If so, how were they divided over the experiments? It is also not clear what genetic background was used for the WT mice. A different genetic background could explain why the WT-HB-01 mice are gaining weight so much faster than the ACE2-PBS mice.

Author response: Thank you for your valuable comment. The transgenic hACE2 mice have been previously generated and stored in our lab. The details on transgenic mice modeling and validating of hACE2 expression could be found in publication (*Yang et al 2007*). We have cited this publication. We have replenished the key details on model construction in Methods section in the revised manuscript (Line 180-186). We also confirmed the control group (WT mice) harbored the same genetic background (ICR mice) with the hACE2 transgenic mice to enhance the validation of our findings. Due to the increase in the number of mice and the longer observation time, there are no significant difference between the WT-HB-01 mice and ACE2-mock in the revised manuscript (Figure 1a).

2) The animal number is too low. Two animals per timepoint makes it impossible to perform statistical analysis.

Author response: Thank you for your critical comment. During the stage of under-review, we have replenished the relative work and highly improved the abundance of the experiments. We have replenished the experimental animals up to 49. We observed the clinical manifestation and detected the specific antibody in 13 mice, including ACE2-HB-01 mice (n=7), ACE2-mock mice (n=3) and WT-HB-01 mice (n=3). We euthanized 12 mice per group at 1 dpi (n=3/group), 3 dpi (n=3/group), 5 dpi (n=3/group) or 7 dpi (n=3/group) to obtain time-dependent viral replication and pathological changes undergone virus infection. After we have added the experimental animals, we can detect the viral loads and viral titers in updated 3 animals per group to perform statistical analysis. Also, we addressed the individual-dependent data and time-dependent data along the timeline of observation and measurements to improve the sufficiency of the finding.

3) The authors state that they determine 'viral load'. They did not. They only confirmed the presence of viral RNA in tissues. The authors do not address why there is a much bigger band in the liver than in the lungs. Was there any histological evidence of virus replication outside the respiratory tract?

Author response: Thank you for your valuable comment. We have sequenced all PCR products and confirmed that the viral PCR could only been found in tissue from lungs. Remained PCR products were believed as non-specific bands during experiments as it could not validated by sequencing. Moreover, we presented the time-dependent viral loads in all primary tissues along detecting timeline (1 dpi, 3 dpi, 5 dpi and 7 dpi) to confirm the findings (Figure 1b). We also have added the immunohistochemistry stained with anti-Spike antibody in the lung (Figure 2) and the primary tissues outside the respiratory tract (Supp. Fig..4).

4) It would have been interesting if swabs had been collected from these animals to look at virus shedding from the respiratory and intestinal tract.

Author response: Thank you for your valuable comment. We have systematically detected the viral loads and immunohistochemistry in all primary tissues. These findings were highly

consistent. The considerably high viral load could only be shown in lung. A transient low expression in intestine could be supposed as the residue of the virus input inoculum.

5) How did the authors select mice for euthanasia on 3 and 5 dpi? Based on the bodyweight data it appears as though the animals with weight loss were specifically selected.

Author response: Thank you for your valuable comment. We have designed a time-dependent pathological experiments in 12 WT-HB-01 mice, 12 ACE2-mock and 12 ACE2-HB-01 for euthanasia. These 12 mice per group were randomly divided in to the four groups to undergo four time-point measurements (1 dpi, 3 dpi, 5 dpi and 7 dpi). In the revised manuscript, we have replenished these time-dependent data including HE (Figure 2 and Supp. Fig. 3), immunochemistry (Figure 2, Supp. Fig. 2 and 4), viral loads (Figure 1b) and viral titers in tissues (Figure 1c). 13 mice (3 WT-HB-01 mice, 3 ACE2-mock and 7 ACE2-HB-01) were designed for the continuous observation of weight loss (Figure 1a) and the measurements for specific antibody (Figure 1e) in the revised manuscript.

6) It would be better to display weights of individual animals. The error bars are quite large, suggesting that some of the infected ACE2 mice may not have lost weight, whereas others may have lost more than 5%.

Author response: Thank you for your valuable comment. We have addressed the changes of weight loss individually in the revised manuscript (Figure 1a).

7) The gross lung lesion pointed out in Figure 2 is not very convincing. How many animals showed gross lesions on each time point? Is this the only 'lesion' that was observed?

Author response: Thank you for your valuable comment. In order to clarify the histopathological changes, we have improved the abundance of the experiments and added the data, including gross lung lesion of ACE2-HB-01 at 5 dpi (Figure 2a), HE stain (Figure 2 and Supp. Fig. 3), modified Masson's Trichrome stain (Supp. Fig. 1), PAS stain (Supp. Fig. 1) and immunochemistry (Figure 2, Supp. Fig. 2 and 4). All the ACE2-HB-01 mice at 3 dpi and 5 dpi displayed comparable gross lesions and histopathological changes, respectively (Figure 2a) (described in the revised manuscript). The lesions were progressed from 3 dpi to 5 dpi revealing different stages of the pathogenesis after SARS-CoV-2 infection. Specifically, consistent with the gross lesions, lung tissues from ACE2-HB-01 mice at 3 dpi had moderate, multifocal,

interstitial pneumonia, and the lung progressed into coalescing interstitial pneumonia with diffuse lesions at 5 dpi.

8) It is very hard to draw any conclusions from the histology pictures. The lungs were not insufflated before fixation resulting in all the lungs being collapsed making it very hard to analyze these panels. The tissues were fixed inappropriately which appears to have resulted in some autolysis. This makes it hard to distinguish cell necrosis due to autolysis (resulting in many visible nuclei) from the influx of inflammatory cells (usually observed as many purple nuclei). Normally, one would also observe perivascular and peribronchiolar edema in case of bronchointerstitial pneumonia, but this cannot be observed in the slides from nCoV-infected ACE2 mice. The authors have included staining specific for 2019-nCoV, but when defining a new model it is essential to show that the expression of viral antigen overlaps with specific lesions. So nonfluorescent antibody stain in a counterstained slide needs to be included instead of the histology panels.

Author response: Thank you for your critical comment. We apologize for our insufficient quality data due to the emergency. We have supplemented new experiments and data according to your insightful advice. The organs were re-collected, appropriately fixed in 10% buffered formalin solution, and prepared in paraffin sections (3-4 μm in thickness) routinely. All the major tissues sections were stained with Hematoxylin and Eosin (H&E) and immunohistochemistry (IHC).

The detail and typical morphologic changes of histopathological lesions in the lungs can be clearly observed by the microscope. Perivascular infiltrates inflammatory cells, including lymphocytes and monocytes, were observed multifocally within and adjacent to affected areas of the lungs on 3 dpi and 5 dpi (Figure 2a). The types of specific inflammatory cells were also confirmed by IHC revealing that $MAC2^+$ macrophages, $CD3^+$ T lymphocytes, and $CD19^+$ B lymphocytes scattered dispersed or occasionally aggregated in the alveolar interstitium in the ACE2-HB-01 mice (Supp. Fig. 2).

Affected bronchioles were also filled with occasionally small amounts of periodic acid schiff (PAS) positive-exudation or denatured and detached bronchiolar epithelium at 3 dpi and 5 dpi (Supp. Fig. 1b). Consistently, the bronchi of mice have very few mucus- producing cells. The proximal intrapulmonary bronchi in mice are lined with columnar to cuboidal epithelium composed predominantly of Clara (59%) which is the principal secretory cell, and ciliated (28%) epithelial cells. The intrapulmonary bronchi in mice lack cartilaginous rings. In contrast,

humans have no bronchial Clara cells but have considerably more mucus-producing goblet cells (10%) and serous cells (3%). Therefore, less exudation can be observed in the infected hACE2 mice compared to the clinical cases in patients with SARS-CoV-2.

IHC staining of sequential section were detected to observe the expression of viral antigens overlaps with specific lesions. IHC staining revealed that viral antigens were found in alveolar macrophages, alveolar epithelia, swell and degenerative bronchial epithelial cells (Figure 2b).

[1] Treuting PM, Dintzis SM, Liggitt D, Frevert CW. (2011) Comparative anatomy and histology: a mouse and human atlas (expert consult). Academic Press.

9) The language needs attention. E.g. the term 'infection' is used inappropriately in several instances.

Author response: We have corrected this in the revised manuscript according to your suggestion.

Referee #3 (Remarks to the Author):

Key results: In this manuscript, the authors experimentally inoculate genetically modified mice expressing the receptor for SARS-CoV (human ACE2) on their cell surfaces, in order to determine the pathogenicity of SARS-CoV-2 (the formal name of the new coronavirus) and to fulfill Koch's postulates for COVID-19 (the formal name of the disease caused by infection with SARS-CoV-2). The most outstanding features of this work are the demonstration of body weight loss and pneumonia associated with SARS-CoV-2 infection.

Author response: Thank you for your insightful comment. We have fully considered your advice and answered your questions point by point.

Validity: The manuscript has important flaws.

1) The authors do not reach their aims, which are to clarify the pathogenicity and to determine whether Koch's postulates are fulfilled. They make no statement about the pathogenicity (that is, the virulence) of SARS-CoV-2. They conclude that the results of their study fulfill Koch's postulates, but they do neither specify what these postulates are nor indicate how the results of their study fulfill them. To elaborate on these Koch's postulates, they were originally used for bacteria, not viruses. Thomas Rivers adapted them for viruses in a 1937 paper in J. Bact. Although not very clearly indicated in the paper, the adapted postulates could be summarized as:

1. isolation of virus from diseased hosts

2. cultivation in host cells

3. proof of filterability

4. production of comparable disease in the original host species or a related one

5. re-isolation of the virus

6. detection of a specific immune response to the virus

Although the house mouse is not very closely related to the human being (postulate 4), it is a mammal, and the authors do show a comparable disease (pneumonia) as seen in affected people (ground-glass opacities in the lung on CT scan, consistent with interstitial pneumonia). The authors also re-isolate the virus from lung tissue of the inoculated mice. Therefore, if the mouse is considered sufficiently related to the human being for the purposes of these postulates, they have fulfilled postulates 4 and 5 of above list. Postulates 1 to 3 have been reached by others prior to this study, and postulate 6 is not yet fulfilled.

Author response: Thank you for your detailed comments about Koch's postulates. We have supplemented the detection of anti-Spike (SARS-CoV-2) specific antibody at 0 dpi and 21 dpi in WT-HB-01 mice and ACE2-HB-01 mice (Figure 1e). The postulates 4 (Figure 2), 5 (Figure 1) and 6 (Figure 1e) have been fulfilled in the revised manuscript.

2) The results of pathological analysis, which are the main results of this manuscript, are inadequately performed and described. The use of frozen sections, embedding in Cryo-Gel, and sectioning at 10 micrometers results in sections that are too thick to see enough histological details. A better method would be formalin fixation, paraffin embedding, and sectioning at 3 micrometers. In addition, panels d to i of Figure 2 indicate that the haematoxylin and eosin staining is not correct. Specifically, the eosin staining is too weak. Correct eosin staining would show the erythrocytes, which must be present in these sections, as bright red circles, but they are not visible at all in these photographs. Because of the thick sections and the poor staining, it is not possible to describe the details of the pathological changes at sufficient detail. For example, it is not possible to determine whether there is oedema fluid in the alveolar lumina (which would be expected in such a lesion), nor the identity of the inflammatory cells that are presumably infiltrating the affected tissue.

Author response: Thank you for your critical comment. We apologize for our insufficient quality data due to the emergency. We have supplemented all the methods and data according to your insightful advice. The organs were re-collected, fixed in 10% buffered formalin solution, and prepared in paraffin sections (3-4 μm in thickness) routinely. All the tissues sections were stained with Hematoxylin and Eosin (H&E) and immunohistochemistry (IHC). The detail and typical morphologic changes of histopathological lesions in the lungs can be clearly observed by the microscope.

There were a few edema fluids in the alveolar lumina which were confirmed by periodic acid schiff (PAS) stain with occasionally small amounts of periodic acid schiff (PAS) positive-exudation or denatured and detached bronchiolar epithelium (Supplementary Figure 2b). From anatomical histology, the proximal intrapulmonary bronchi in mice are lined with columnar to cuboidal epithelium composed predominantly of Clara (59%) which is the principal secretory cell, and ciliated (28%) epithelial cells [1].The intrapulmonary bronchi in mice lack cartilaginous rings. In contrast, humans have no bronchial Clara cells but have considerably more mucus-producing goblet cells (10%) and serous cells (3%). Therefore, less exudation can

be observed in the infected hACE2 mice compared to the clinical cases in patients with SARS-CoV-2.

Perivascular infiltrates inflammatory cells, including lymphocytes and monocytes, were observed multifocally within and adjacent to affected areas of the lungs (Figure 2a). On the 5 dpi, the lung progressed into coalescing interstitial pneumonia with diffuse lesions. Thickened alveolar septa were filled with lymphocytes and monocytes (Figure 2a). The types of specific inflammatory cells were also confirmed by IHC revealing $MAC2^+$ macrophages, $CD3^+$ T lymphocytes, and $CD19^+$ B lymphocytes were scattered dispersed or occasionally aggregated in the alveolar interstitium in the ACE2-HB-01 mice (Supp. Fig. 2).

[1] Treuting PM, Dintzis SM, Liggitt D, Frevert CW. (2011) Comparative anatomy and histology: a mouse and human atlas (expert consult). Academic Press.

3) There is inadequate support from pathological analyses for SARS-CoV-2 as the cause of the observed pneumonia. The key evidence for this is colocalisation of virus antigen expression and histopathological changes. The authors do not show this. All they show, in Figure 3q is colocalisation of virus antigen expression and hACE2 expression.

Author response: Thank you for your critical comment. We have supplemented the immunohistochemical staining of sequential section with the HE staining. The viral antigens were colocalized with the histopathological changes, mainly in the bronchial epithelial cells, alveolar macrophages, and alveolar epithelial cells (Figure 2b).

4) There is inadequate support from the photos of the gross lesions (panels a to c in Figure 2) for the presence of pneumonia. The arrow in panel c is pointing at a supposed area of pneumonia, but this cannot be identified as such. It appears that the lungs in panels a to c are collapsed, a normal phenomenon that occurs after the lungs are taken out of the thoracic cavity (which is under vacuum during life) at autopsy. This makes identification of subtle gross lesions difficult. A possible solution for this, in future autopsies, is to close the trachea by use of a small haemostat clamp before opening the thoracic cavity and removing the lungs. In this way, the lungs remain filled with air and it is much easier to distinguish normal from abnormal lung tissue. The description of the gross lesions is inadequate, both qualitatively and quantitatively. Qualitatively, description of the gross lesion should include colour and whether lesions were raised, level, or depressed compared to surrounding tissue. Quantitatively, the affected area

of lung and the weight of the lungs (compared to the negative controls) would be suitable measures to estimate the proportion of affected lung tissue.

Author response: Thank you for your critical comment. We have supplemented new data to support our results. The mice were autopsied at different intervals after infection. Compared to WT-HB-01 mice or ACE2-mock mice with homogeneously pink and slightly deflated, after the animals were euthanized, postmortem examinations, all the ACE2-HB-01 mice at 3 dpi displayed relatively comparable gross lesions, respectively. ACE2-HB-01 mice were necropsied at 3 dpi, all lungs showed focal to multifocal dark red discoloration (Figure 2a). These lesions progressed into multifocal to coalescent scattered-dark reddish purple areas and focal palpable nodules throughout the lung lobes at 5 dpi (Figure 2a). The lungs enlarged in size in all the SARS-CoV-2-infected hACE2 mice.

5) Originality and significance: The results are original. SARS-CoV-2 is a recently identified virus, and to my knowledge no studies on experimentally infected laboratory animals have been published. The results of such studies, if performed adequately, would be of interest to people from several disciplines.

Author response: Thank you for your critical comment. We apologize that we have submitted preliminary data with insufficient abundance as the we attempt to provide valuable information to an emergent epidemic as soon as we could. During the stage of under-review, we have replenished the relative work and highly improved the abundance of the experiments. We highly appreciated the insightful comments you have addressed and we have revised the manuscript.

6) Data & methodology: There is inadequate support from the results that "viral RNA was positive by RT-PCR (Figure 1b) and identified by sequencing only in the lung of infected-hACE2 mice at 3 dpi and 5 dpi". The cited panel (b of Figure 1) shows bands at the same level as the positive control not only for the lung sample, but also for the spleen and liver samples. Whether there are bands at the same level as the positive control for intestine, kidney, and brain samples is not clear because the resolution of the photo is too low. The authors state "identified by sequencing", which presumably indicates that they sequenced the PCR products in the bands cut out of the gel, but they do not state the sequencing method in the methods section. Because these mice presumably express hACE2 on cells throughout the body, it is

possible that SARS-CoV-2 replicates in other tissues than the lungs, and therefore it would not be unexpected to detect viral RNA in other tissues. In this regard, it is important to also specifically state what the results of histopathology and immunofluorescence in all the organs tested by RT-PCR. The authors state that "the major organs of mice were examined by histopathology and immunofluorescence", but only show results for the lung.

Author response: Thank you for your insightful comment. Firstly, we have sequenced all PCR products and confirmed that the viral PCR could only been found in tissue from lungs. Remained PCR products were believed as non-specific bands during experiments as it could not validated by sequencing. Secondly, to identify and clarify the viral replication outside the respiratory tract, we have supplemented the time-dependent viral loads in all primary tissues along detecting timeline (1 dpi, 3 dpi, 5 dpi and 7 dpi, Figure 1b). We also particularly compared viral titers in lungs among WT-HB-01, ACE2-MOCK and ACE2-HB-01 mice along the timeline (Figure 1c). Thirdly, we have systematically supplemented the histopathology and immunohistochemistry in all primary tissues (Figure 2 and Supp. Fig. 4). These findings were highly consistent. It was clearly shown that the virus actively replicates in lung, and led to the pneumonia associated with SARS-CoV-2 infection. A transient low expression in intestine at 1 dpi could be supposed as the residue of the virus input inoculum.

7) Appropriate use of statistics and treatment of uncertainties: P-values are indicated clearly in Figure 1A. Nature requires requires detailed statements on statistical tests, some of which have not been fulfilled. It has not been stated whether the t-test used is one- or two-sided. It also has not been stated that it is assumed that both populations have normal distributions, which is a requirement for Welch's t-test. Also, the test statistic (t) with degrees of freedom and P value (exact value) are not noted.

Author response: Thank you for your valuable advice. We have modified the appropriate statistics according to your suggestion (Figure 1, Supp. Fig. 5).

8) References:

Since this paper refers to Koch's postulates—and more precisely, River's modifications for viral diseases—a reference to these postulates should be included: Rivers, T. M. J. Bacteriol. 33, 1–12 (1937).

Author response: We have cited this publication in the revised manuscript (line 154). We highly appreciate the valuable advice that you have offered once again.

9) Minor comments:

Line 50: It is not clear to which group of mice the statement "n = 7" refers.

Author response: We have corrected this in the revised manuscript (line 56-61).

Line 50: Please state the inoculum volume, whether and how mice were anaesthetised for inoculation, and how they were held.

Author response: We have corrected this in the revised manuscript (line 58-59).

Line 52: Please clarify what is meant by "for 10 dpi".

Author response: We have corrected this in the revised manuscript (line 65-66).

Line 53: Please clarify whether you used a score card to record clinical signs, and which signs were specifically checked.

Author response: We have corrected this in the revised manuscript (line 61-66). We established the observation system for mice infected with virus, and the degree of disease progression was comprehensively determined such as emaciation, arched back, bristles, and decreased projection of external stimuli. In this study, we observed emaciation and slight bristles in infected mice.

Line 54: Please list which major organs were assessed for distribution of SARS-CoV-2.

Author response: We have corrected this in the revised manuscript (line 67-69).

Line 56: Please clarify what is meant by "by sequencing", and provide methodological details.

Author response: We have corrected this in the revised manuscript (line 211-215).

Line 57: Please state from which tissue virus was isolated. Please indicate whether it was confirmed, e.g. by RT-PCR, that the cytopathic effect and the observed coronavirus particles corresponded to SARS-CoV-2.

Author response: We have corrected this in the revised manuscript (line 67-81).

Line 67: Please clarify what is meant by "interstitial hyperplasia". Specifically, which cell types were involved in this hyperplasia?

Author response: Thank you for your valuable advice. Microscopically, lung tissues from ACE2-HB-01 mice at 3 dpi had moderate, multifocal, interstitial pneumonia. Alveolar septa became widened with infiltration of mainly lymphocytes and monocytes. The types of specific inflammatory cells were also confirmed by IHC. The results revealed that $MAC2^+$ macrophages, $CD3^+$ T lymphocytes, and $CD19^+$ B lymphocytes were scattered dispersed or occasionally aggregated in the alveolar interstitium in the ACE2-HB-01 mice (Supp. Fig. 2).

Additionally, the increased collagen fiber in the thickened alveolar interstitium in ACE2-HB-01 mice was more than that in WT-HB-01 mice confirmed by Modified Masson's Trichrome stain, revealing that there may occur the proliferation of fibrous connective tissue (Supp. Fig. 1a). The bronchiolar epithelial cells showed swelling, degeneration, and some dissolved. Perivascular infiltrates inflammatory cells, including lymphocytes and monocytes, were observed multifocally within and adjacent to affected areas of the lungs (Figure 2a).

Line 67: Please specify which cell types comprised these inflammatory cells.

Author response: Thank you for your valuable advice. Microscopically, the inflammatory cells in the lung tissues of ACE2-HB-01 mice were mainly lymphocytes, monocytes and alveolar macrophage.

Line 69: Were no pathological changes observed in alveolar epithelial cells?

Author response: Thank you for your valuable advice. After infection, the alveolar epithelial cells displayed swelling, degeneration, dissolved, necrosis, and mild proliferation.

Line 70: Please specify which cell types comprised these leukocytes.

Author response: Thank you for your valuable advice. Microscopically, the inflammatory cells in the lung tissues of ACE2-HB-01 mice were mainly lymphocytes and monocytes.

Line 72: Was is possible to distinguish between type 1 and type 2 alveolar epithelial cells?

Author response: Thank you for your valuable comment. We thought it is possible to distinguish between typical type 1 and type 2 alveolar epithelial cells in the lesion areas where not too many distorted mixed-cells aggregated together. Type I (membranous) pneumonocytes are remarkably thin and cover most of the alveolar wall. Type II (granular) pneumonocytes tend to bulge into the air space, appearing as large cuboidal cells with lamellar bodies (surfactant) in the cytoplasm. In some areas, the alveolar walls collapsed, a mixture of desquamated epithelial cells and inflammatory cells accumulated in the alveolar cavities. Under this circumstances, it is not easy to distinguish type 1 and type 2 alveolar epithelial cells.

Line 161: Please indicate whether negative controls, positive controls, and omission controls were used, and provide details about these controls.

Author response: We have corrected this in the revised manuscript (line 270-273).

Supplementary figure 1: An explanation for the asterisk is missing.

Author response: We have corrected this in the revised manuscript (line 407-408).

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

In this revised manuscript, the authors have done an excellent job addressing the technical concerns raised in the previous reviews and increased the numbers of mice that were studied. The results show that hACE2 Tg mice, with hACE2 expression driven by the mouse ACE2 promoter, are susceptible to infection with SARS-CoV-2. They develop very minor clinical disease.
While the authors conclude that the infected mice fulfill Koch's postulates, the results really just show that hACE2 confers SARS-CoV-2 susceptibility to mice, without mimicking human disease to a substantial extent. Notably, this proof of susceptibility is an important advance for the COVID-19 field since a major research goal is develop a mouse model for this disease. However, the transient disease in these mice will make them only minimally useful for studies of vaccines and therapeutic interventions.

Other comments:
The histological changes in Figure 2 should be quantified, especially since it is likely that the pathological changes are patchy.

Referee #3 (Remarks to the Author):

Overall, the authors have made substantial improvements in the revised manuscript, but in pathological and virological analyses. However, I still have significant concerns about the manuscript in its current form, as detailed below. Although I think that these hACE2 transgenic mice may be useful componentns of animal models for COVID-19, e.g. to test antiviral therapeutics and vaccines, there is no factual basis for the authors' statements that both clinical features and pathological changes of SARS-CoV-2 infection in mice resemble those in SARS-CoV-2-infected humans.

Key results:
-The authors have changed their aims. The stated aim now is "to study the pathogenicity of the virus." Although more a statement of method than an aim, they do indeed study the pathogenicity of SARS-CoV-2 in mice and so reach their aim. However, their conclusion (line 148) is not correct: they have only described the pathogenicity of SARS-CoV-2 infection, they have not clarified it. For example, they have not provided any explanation for the higher pathogenicity of SARS-CoV infection than SARS-CoV-2 infection in mice, although both apparently use the same receptor.

Although no longer stated as an aim, they do state that they have fulfilled Koch's postulates in the abstract and at the end of the discussion. They now cite Rivers (1937) as the source of Koch's postulates amended for virus infecitons, but still do not state what the postulates are that have been fulfilled by this study.

-The authors have made a great effort to improve the methodology of pathological analysis, by repeating the experimental infections, collecting tissues in formalin, embedding them in paraffin, and making much thinner (3 to 4 micron-thick) tissue sections that allow detailed histological analysis. Because of this improved methodology, it is now possible to determine to which degree the histopathological changes in the lungs of these mice correspond to those observed in fatal human cases of SARS-CoV-2 infection.

-The authors now partly provide evidence of colocalisation of virus antigen expression and histopathological changes. They show that virus antigen expression is present in respiratory epithelial cells and mononuclear cells (which they assume to be alveolar macrophages) within lesional areas in the lungs. What is still missing is information whether virus antigen expression in these cell types is absent in non-lesional areas of the lungs.

-In the revised manuscript, the photos of the gross lesions support the described gross lesions, and are a great improvement. Also, there is now an adequate qualitative written description of the gross lesions in the lungs.

Data & methodology:

-It appears that the authors no longer make use of the standard RT-PCR methods and results they presented in the original manuscript, which led to my questions about the interpretation of the bands in the sequencing gel in other tissues than lungs. They now have used a quantitative real-time RT-PCR (qRT-PCR) to determine presence and quantity of viral RNA in lungs and extra-respiratory tissues. The expected result from qRT-PCR is a cycle threshold (CT) value, yet the authors express their results in log*10 copies/ml in Figure 1b. Apparently they are using a

conversion factor, perhaps from a dilution series of viral RNA of known concentration. However, the method of conversion is not stated in the methodology. Also, qRT-PCR does not result in PCR products that can be used to verify that the test is specific for SARS-CoV-2, yet this is what the authors state in lines 281-285. Both the conversion method and verification method need to be clarified.

-In their rebuttal, the authors state that viral PCR (I assume they mean viral RNA) could only be found in tissue from lungs, yet in Figure 1b they also show a viral load (I assume they mean viral RNA) in the intestine at 1 dpi. Please clarify.

-The authors now state that no significant histopathological changes or SARS-CoV-2 antigen were found in other organs, as requested

Appropriate use of statistics and treatment of uncertainties:

-The authors now provide indicate that the t-test used is two-sided. However, they do not provide statements about the assumption of normal distributions of the populations, and do not note the test statistic t, the degrees of freedom, or the exact P value, although this is requested by the Nature journal.

Conclusions:

-The authors state that the clinical results in these mice are comparable to clinical reports of human SARS-CoV-2 cases (lines 128-131). Clinical signs in people, according to reference #7, include fever, cough, myalgia, fatigue, sputum production, headache, diarrhoea, and dyspnoea. In comparison, the mice showed body weight loss and 'bristles' (which I interpret as a raised, unkempt hair coat). I conclude that these clinical results are not comparable.

-The authors state that the lesions in these mice correspond to the lesions in a fatal human case of SARS-CoV-2 infection (lines 132-138). The human fatal case (reference #8) showed diffuse alveolar damage (the pathological correlate to ARDS), characterized by pulmonary oedema, hyaline membranes, fibromyxoid exudate (meaning exudate containing fibrin and mucus), syncytial cells in alveolar lumina, enlarged pneumocytes, and accumulation of inflammatory cells, mostly lymphocytes. In comparison, the mice had multifocal or coalescing interstitial pneumonia, characterized mainly be accumulation of macrophages and lymphocytes. Important features that were missing in mice are alveolar oedema and hyaline membranes, which need to be present to make the pathological diagnosis of diffuse alveolar damage (e.g., J. E. Craighead, Pathology and pathogenesis of human viral disease, Academic Press 2000). I conclude that these autopsy results do not correspond so well to each other.

-I have two other major concerns about the pathological analysis of the mice. First, the authors now (as I requested) specify which cell types comprised the leukocyte infiltration (lines 88-110). However, they mention only monocytes/macrophages and lymphocytes. I miss the description of neutrophils in the lung lesions of mice at 3 and 5 dpi. At such an early stage of virus infection, this inflammatory cell type (small round cells with characteristic segmented nuclei) should be a prominent component of the inflammatory infiltrate (see, e.g. Camp & Jonsson, 2017, A Role for Neutrophils in Viral Respiratory Disease. Front Immunol. 2017 May 12;8:550), yet the authors do not mention neutrophils. Second, the authors suggest that increased collagen fibre is an important component of the thickened alveolar walls at 3 dpi (lines 89 to 92). This implies increased fibrous connective tissue (that is, fibrosis), which is a chronic response to a severe virus infection that one would not expect to be the main component of the lesion so soon (3 and 5 days) after inoculation, although it may be starting then.

Minor comments:
-As requested, the authors have now provided information on the anaesthesia of the mice at the

time of inoculation. Although they indicate which anaesthetic they used (2.5% avertin), I missed other key data: route of anesthesia, dose of anesthetic in mg per kg body weight. I also missed how (i.e. in which position) mice were held at the time of inoculation. This is important to know whether the inoculum would have spread directly to the lungs.

-As requested, the authors have listed the clinical signs they checked. However, I do not understand what they mean by "decreased projection of external stimuli". Also, the term "bristles" is perhaps better described as "bristled fur" (Pasler et al. 2009. Endpoints for Mouse Abdominal Tumor Models: Refinement of Current Criteria. Comp Med. 2009 Jun; 59(3): 234–241).

-As requested, the authors have listed the major organs they collected. However, they need to explain that they did not collect testis from all mice, since both male and female mice (50% of each?) were used. Also, were ovaries collected from female mice, instead?

-The authors have clarified many questions abou the character of the lung lesions in the mice. One term, " epithelial cells ... dissolved" is not used normally in pathological description. Instead, I would refer to 'loss' or 'fragmentation' of bronchiolar or alveolar epithelial cells.

-As requested, the authors have provided information on negative controls and positive controls (lines 340 to 344). However, this is for confocal microscopy only. The same information needs to be provided for immunohistochemistry (lines 314-327). Furthermore, for both techniques, details need to be provided about omission controls; also isotype controls, if monoclonal antibodies were used.

Line 72-73: Virus titres in tissues should be given per g tissue.

Line 95: "dispersed" This is not clear. Do you mean "distended"?

Line 100: It is stated that more B lymphocytes were seen in the lungs of ACE2-HB-01 mice. However, extended data figure 2b, lower panel, does not show this. Please correct.

Line 111: It is stated that viral antigen was found in alveolar macrophages. How was this determined? In my experience, it is not possible to distinguish alveolar macrophages and type 2 pneumocytes in such lesions just on morphological features. It requires immunohistochemical staining, e.g. CD68 for macrophages and keratin for epithelial cells.

Line 113, 398: 'Histopathological lesions' is a pleonasm (because all lesions are pathological). Either write 'histological lesions' or 'histopathological changes'.

Line 137: I would not attach much importance to the presence of PAS-positive material. Based on extended data figure 1b, there is hardly any evidence for an increase; furthermore, it is a nonspecific pathological change.

Line 138: 'brochiolar' > 'bronchiolar' Also, the legend for extended data figure 1b states 'bronchial'.

Line 144: casculitis > vasculitis

Fig. 1: The EM image of the virus is too small to be able to identify it as a coronavirus, e.g. the characteristic spikes on the surface of the envelope. Please provide an image at higher magnification.

Lines 193-195: The text speaks of 'virus replication','viral load', and 'virus titer' This needs to be corrected, because it is viral RNA that was measured.

Line 200: Please specify that IgG was measured in sera.

Line 211: The cells to which the red arrows point are too small to identify as lymphocytes. Please provide an image at higher magnification.

Line 212, 215: The cells to which the green arrows point cannot be identified as alveolar macrophages on morphological criteria alone. Please provide evidence by immunohistochemistry about cell type, or be more conservative in naming the cells (e.g. 'mononuclear cells').

Fig. 3: Please indicate at which day post inoculation these tissues were obtained.

Line 301: 'Figure 1' > 'Figure 5'

Extended Data Fig. 1: The images are too small in panels a to determine in which histological structure the blue staining is present; and in panels b to see any PAS-positive material. Please provide images at higher magnification.

Author Rebuttals to First Revision:

Referee #1 (Remarks to the Author):

In this revised manuscript, the authors have done an excellent job addressing the technical concerns raised in the previous reviews and increased the numbers of mice that were studied. The results show that hACE2 Tg mice, with hACE2 expression driven by the mouse ACE2 promoter, are susceptible to infection with SARS-CoV-2. They develop very minor clinical disease.

While the authors conclude that the infected mice fulfill Koch's postulates, the results really just show that hACE2 confers SARS-CoV-2 susceptibility to mice, without mimicking human disease to a substantial extent. Notably, this proof of susceptibility is an important advance for the COVID-19 field since a major research goal is develop a mouse model for this disease. However, the transient disease in these mice will make them only minimally useful for studies of vaccines and therapeutic interventions.

Author response: Thank you for your insightful comment and kindness. We have fully considered your advice and answered your question as following. This mouse model with SARS-CoV-2 infection has been continuously detected viral replication in the lungs for 7 days, and exhibited the moderate interstitial pneumonia, which is applicable for protective evaluation of antiviral drugs and vaccines. We have modified this in the abstract and at the end of the discussion (line 35-37; line 156-158) according to your suggestion.

Other comments:

The histological changes in Figure 2 should be quantified, especially since it is likely that the pathological changes are patchy.

Author response: Thank you for your insightful comment. We have revised our description to quantify the histological changes, "Microscopically, the lung tissues from ACE2-HB-01 mice at 3 dpi displayed moderate interstitial pneumonia characterized by thickened alveolar septum accompanied with infiltration of inflammatory cells in 70-80% areas of the lung tissues, and accumulation of inflammatory cells in partial alveolar cavities (20-30%)" (line 89-93), "At 5 dpi, the lung progressed into coalescing interstitial pneumonia with diffuse lesions characterized by more severe thickened alveolar septum accompanied with infiltration of inflammatory cells, and accumulation of inflammatory cells in more alveolar cavities (40-50%)" (line 95-98).

Referee #3 (Remarks to the Author):

Overall, the authors have made substantial improvements in the revised manuscript, but in pathological and virological analyses. However, I still have significant concerns about the manuscript in its current form, as detailed below. Although I think that these hACE2 transgenic mice may be useful componentns of animal models for COVID-19, e.g. to test antiviral therapeutics and vaccines, there is no factual basis for the authors' statements that both clinical features and pathological changes of SARS-CoV-2 infection in mice resemble those in SARS-CoV-2-infected humans.

Author response: Thank you for your valuable comment. We have fully considered your advice and answered your questions point by point.

Key results:

-The authors have changed their aims. The stated aim now is "to study the pathogenicity of the virus." Although more a statement of method than an aim, they do indeed study the pathogenicity of SARS-CoV-2 in mice and so reach their aim. However, their conclusion (line 148) is not correct: they have only described the pathogenicity of SARS-CoV-2 infection, they have not clarified it. For example, they have not provided any explanation for the higher

pathogenicity of SARS-CoV infection than SARS-CoV-2 infection in mice, although both apparently use the same receptor.

Author response: Thank you for your precise suggestion and kindness. For more accurate expression, we have used the term “confirmed” (line 35) or “demonstrated” (line 154) instead of “clarified” in the revised manuscript. We have modified this in the abstract and at the end of the discussion (line 35-37; line 156-158) according to your suggestion.

Although no longer stated as an aim, they do state that they have fulfilled Koch’s postulates in the abstract and at the end of the discussion. They now cite Rivers (1937) as the source of Koch’s postulates amended for virus infections, but still do not state what the postulates are that have been fulfilled by this study.

Author response: Thank you for your valuable comment. In the revised manuscript (line 138-144), we have added the description of Koch's postulates that have been completed in our study.

-The authors have made a great effort to improve the methodology of pathological analysis, by repeating the experimental infections, collecting tissues in formalin, embedding them in paraffin, and making much thinner (3 to 4 micron-thick) tissue sections that allow detailed histological analysis. Because of this improved methodology, it is now possible to determine to which degree the histopathological changes in the lungs of these mice correspond to those observed in fatal human cases of SARS-CoV-2 infection.

Author response: Thank you for your insightful comment.

-The authors now partly provide evidence of colocalisation of virus antigen expression and histopathological changes. They show that virus antigen expression is present in respiratory epithelial cells and mononuclear cells (which they assume to be alveolar macrophages) within lesional areas in the lungs. What is still missing is information whether virus antigen expression in these cell types is absent in non-lesional areas of the lungs.

Author response: Thank you for your insightful advice. Viral antigen could be widely detectable from lung tissues and they are mainly colocalized with the lesion areas. In non-

lesional areas of the lungs, there were also small amounts of viral antigen in respiratory epithelial cells. We have supplemented this in the revised manuscript (line 121-122).

-In the revised manuscript, the photos of the gross lesions support the described gross lesions, and are a great improvement. Also, there is now an adequate qualitative written description of the gross lesions in the lungs.

Author response: Thank you for your comment.

Data & methodology:

-It appears that the authors no longer make use of the standard RT-PCR methods and results they presented in the original manuscript, which led to my questions about the interpretation of the bands in the sequencing gel in other tissues than lungs. They now have used a quantitative real-time RT-PCR (qRT-PCR) to determine presence and quantity of viral RNA in lungs and extra-respiratory tissues. The expected result from qRT-PCR is a cycle threshold (CT) value, yet the authors express their results in log*10 copies/ml in Figure 1b. Apparently, they are using a conversion factor, perhaps from a dilution series of viral RNA of known concentration. However, the method of conversion is not stated in the methodology. Also, qRT-PCR does not result in PCR products that can be used to verify that the test is specific for SARS-CoV-2, yet this is what the authors state in lines 281-285. Both the conversion method and verification method need to be clarified.

Author response: Thank you for your valuable comment. Actually, we have sequenced all PCR products for the first RT-PCR assay and confirmed that the viral RNA could have been found only in lung tissue. Remained PCR products were non-specific bands during experiments as it could not validated by sequencing. Since the results of the standard RT-PCR methods are qualitative and other reviewers have suggested us that the data should be quantitative, so the qRT-PCR was used in the revised manuscript. We have corrected this in the methods (line 297-302) according to your suggestions.

-In their rebuttal, the authors state that viral PCR (I assume they mean viral RNA) could only be found in tissue from lungs, yet in Figure 1b they also show a viral load (I assume they mean viral RNA) in the intestine at 1 dpi. Please clarify.

Author response: Thank you for your valuable comment. From the time-dependent changes in viral titers, compared to viral loads, it was clearly shown that the virus actively replicates in lung, but not residual input inoculum. In addition, although the viral loads were detected in the intestine at 1 dpi, no virus in intestine was isolated, which was speculated the residual input inoculum from the nasal mucosa to intestines by swallowing. We have supplemented this in the revised manuscript (line 70-73).

-The authors now state that no significant histopathological changes or SARS-CoV-2 antigen were found in other organs, as requested

Author response: Thank you for your insightful comment.

Appropriate use of statistics and treatment of uncertainties:
-The authors now provide indicate that the t-test used is two-sided. However, they do not provide statements about the assumption of normal distributions of the populations, and do not note the test statistic t, the degrees of freedom, or the exact P value, although this is requested by the Nature journal.

Author response: Thank you for your insightful advice. We have corrected the statistical technique at your suggestion (line 379-385). We used Student's technique when the variances are not significantly different. We employed Welch's technique, as the data harbored unequal variances. We compared the weight loss with Mann-Whitney U tests. We are glad to see that the significance was unchanged. The exact P value was supplemented in the revised manuscript (line 196-198, 203-205, 208-209, 444).

Conclusions:

-The authors state that the clinical results in these mice are comparable to clinical reports of human SARS-CoV-2 cases (lines 128-131). Clinical signs in people, according to reference #7, include fever, cough, myalgia, fatigue, sputum production, headache, diarrhoea, and dyspnoea. In comparison, the mice showed body weight loss and 'bristles' (which I interpret as a raised, unkempt hair coat). I conclude that these clinical results are not comparable.

Author response: Thank you for your accurate suggestion. We deleted the "Weight loss" and only described "the interstitial pneumonia in mice infected by SARS-CoV-2 is similar (not comparable) to initial clinical reports of pneumonia caused by SARS-CoV-2". We have modified in the revised manuscript (line 139-143).

-The authors state that the lesions in these mice correspond to the lesions in a fatal human case of SARS-CoV-2 infection (lines 132-138). The human fatal case (reference #8) showed diffuse alveolar damage (the pathological correlate to ARDS), characterized by pulmonary oedema, hyaline membranes, fibromyxoid exudate (meaning exudate containing fibrin and mucus), syncytial cells in alveolar lumina, enlarged pneumocytes, and accumulation of inflammatory cells, mostly lymphocytes. In comparison, the mice had multifocal or coalescing interstitial pneumonia, characterized mainly be accumulation of macrophages and lymphocytes. Important features that were missing in mice are alveolar oedema and hyaline membranes, which need to be present to make the pathological diagnosis of diffuse alveolar damage (e.g., J. E. Craighead, Pathology and pathogenesis of human viral disease, Academic Press 2000). I conclude that these autopsy results do not correspond so well to each other.

Author response: We highly appreciate the valuable advice that you have offered once again. Your opinions inspired us and stimulated our thought process. Because the pathological results of human fatal case were correlated with ARDS, and it cannot be ruled out whether it is lung lesions caused by a combination of other chronic diseases. In contrast, the multifocal or coalescing interstitial pneumonia in a mouse model was directly caused by the SARS-CoV-2 infection, but not affected by other factors. Therefore, according to your insightful suggestion, we have decided to remove this discussion in the revised manuscript.

-I have two other major concerns about the pathological analysis of the mice. First, the authors now (as I requested) specify which cell types comprised the leukocyte infiltration (lines 88-110). However, they mention only monocytes/macrophages and lymphocytes. I miss the description of neutrophils in the lung lesions of mice at 3 and 5 dpi. At such an early stage of virus infection, this inflammatory cell type (small round cells with characteristic segmented nuclei) should be a prominent component of the inflammatory infiltrate (see, e.g. Camp & Jonsson, 2017, A Role for Neutrophils in Viral Respiratory Disease. Front Immunol. 2017 May 12;8:550), yet the authors do not mention neutrophils.

Author response: Thank you for your insightful suggestion. Your valuable suggestions have improved our manuscript. We have supplemented the description of neutrophils in the revised manuscript (line 94, 99, 106, and 117).

Second, the authors suggest that increased collagen fibre is an important component of the thickened alveolar walls at 3 dpi (lines 89 to 92). This implies increased fibrous connective tissue (that is, fibrosis), which is a chronic response to a severe virus infection that one would not expect to be the main component of the lesion so soon (3 and 5 days) after inoculation, although it may be starting then.

Author response: Thank you for your valuable comment. Although collagen fiber would not be expected as the main component of lung lesions at the early stage of infection, the increase in collagen fiber is an indicator of the host responses to lung repairment. In this study, we also observed that it appeared (a small amount) early in the infection process. We have corrected this description according to your suggestion (line 99-100).

Minor comments:

-As requested, the authors have now provided information on the anaesthesia of the mice at the time of inoculation. Although they indicate which anaesthetic they used (2.5% avertin), I missed other key data: route of anesthesia, dose of anesthetic in mg per kg body weight. I also missed how (i.e. in which position) mice were held at the time of inoculation. This is important to know whether the inoculum would have spread directly to the lungs.

Author response: Thank you for your accurate suggestion. 2.5% avertin at 0.2 ml/10g was intraperitoneally administered to mice. We have corrected this in the revised manuscript (line 55 and line 266-267).

-As requested, the authors have listed the clinical signs they checked. However, I do not understand what they mean by “decreased projection of external stimuli”. Also, the term “bristles” is perhaps better described as “bristled fur” (Pasler et al. 2009. Endpoints for Mouse Abdominal Tumor Models: Refinement of Current Criteria. Comp Med. 2009 Jun; 59(3): 234–241).

Author response: Thank you for your accurate suggestion. We have used the term “decreased response to external stimuli” instead of “decreased projection of external stimuli” in the revised manuscript (line 60). “bristled fur” have been corrected in the revised manuscript (line 58).

-As requested, the authors have listed the major organs they collected. However, they need to explain that they did not collect testis from all mice, since both male and female mice (50% of each?) were used. Also, were ovaries collected from female mice, instead?

Author response: Thank you for your suggestion. Testis was only used for male mice, and ovaries in female mice were not tested in this study. We have corrected this in the revised manuscript (line 65 and line 200).

-The authors have clarified many questions about the character of the lung lesions in the mice. One term, “epithelial cells ... dissolved” is not used normally in pathological description. Instead, I would refer to ‘loss’ or ‘fragmentation’ of bronchiolar or alveolar epithelial cells.

Author response: Your valuable suggestions have improved our manuscript. We have corrected this as your suggestion (line 102).

-As requested, the authors have provided information on negative controls and positive controls (lines 340 to 344). However, this is for confocal microscopy only. The same information needs to be provided for immunohistochemistry (lines 314-327). Furthermore, for both techniques, details need to be provided about omission controls; also isotype controls, if monoclonal antibodies were used.

Author response: We have corrected this in the revised manuscript (IHC, line 345-350 ; Confocal microscopy, line 363-365). We did not validate the IgG isotype of the anti-Spike monoclonal antibody (7D2) prepared in our laboratory, so the isotype control of the 7D2 monoclonal antibody was not used in this study.

Line 72-73: Virus titres in tissues should be given per g tissue.

Author response: We have supplemented wt/vol (0.1 g tissue in 100 μL DMEM) in preparation of tissues homogenates (line 282). After clarifying the methods of tissues homogenates (1g/mL), it can be expressed in volume units or weight units.

e.g. Gong S *et al.* Human-Derived A/Guangdong/Th005/2017 (H7N9) Exhibits Extremely High Replication in the Lungs of Ferrets and Is Highly Pathogenic in Chickens. *Viruses.* 2019 May 29;11(6). pii: E494. doi: 10.3390/v11060494.
Xue C *et al.* Vγ4+γδT Cells Aggravate Severe H1N1 Influenza Virus Infection-Induced Acute Pulmonary Immunopathological Injury via Secreting Interleukin-17A. *Front Immunol.* 2017 Aug 31;8:1054. doi: 10.3389/fimmu.2017.01054. eCollection 2017.

Line 95: "dispersed" This is not clear. Do you mean "distended"?

Author response: We have corrected this in the revised manuscript (line 105).

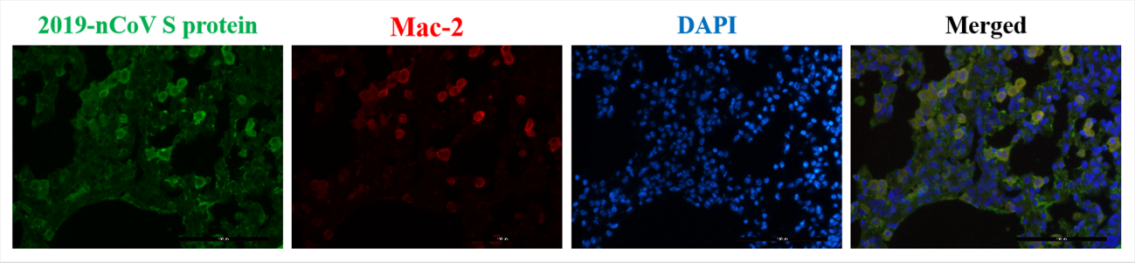
Line 100: It is stated that more B lymphocytes were seen in the lungs of ACE2-HB-01 mice. However, extended data figure 2b, lower panel, does not show this. Please correct.

Author response: Your valuable suggestions have improved our manuscript. We have corrected our description as " some $CD19^+$ B lymphocytes were also observed at 5 dpi" according to your suggestion (line 115).

Line 111: It is stated that viral antigen was found in alveolar macrophages. How was this determined? In my experience, it is not possible to distinguish alveolar macrophages and type 2 pneumocytes in such lesions just on morphological features. It requires immunohistochemical staining, e.g. CD68 for macrophages and keratin for epithelial cells.

Author response: Thank you for your valuable comment. To distinguish alveolar macrophages and type 2 pneumocytes, MAC-2 staining for mouse macrophages [1] were detected. Here, we supplemented the immunofluorescence experiment to provide the additional evidence. The co-localization (yellow color) of SARS-CoV-2 S protein (green color,7D2 monoclonal antibody) and MAC2+ macrophages (red color, Cedarlane Laboratories, CL8942AP) demonstrated that the viral antigen was found in alveolar macrophages.

[1] Jerrold M. Ward. et al. Immunohistochemical Markers for the Rodent Immune System. Toxicologic Pathology, 34:616-630, 2006.



Line 113, 398: 'Histopathological lesions' is a pleonasm (because all lesions are pathological). Either write 'histological lesions' or 'histopathological changes'.

Author response: We have corrected this in the revised manuscript (line 123 and 431).

Line 137: I would not attach much importance to the presence of PAS-positive material. Based on extended data figure 1b, there is hardly any evidence for an increase; furthermore, it is a nonspecific pathological change.

Author response: We have removed this sentence in the revised manuscript.

Line 138: 'brochiolar' > 'bronchiolar' Also, the legend for extended data figure 1b states 'bronchial'.

Author response: We have corrected this in the revised manuscript (line 103 and line 420).

Line 144: casculitis > vasculitis

Author response: We have corrected this in the revised manuscript (line 150).

Fig. 1: The EM image of the virus is too small to be able to identify it as a coronavirus, e.g. the characteristic spikes on the surface of the envelope. Please provide an image at higher magnification.

Author response: The EM image have been shown at higher magnification (Figure 1e) according to your suggestion.

Lines 193-195: The text speaks of ‘virus replication’,’viral load’, and ‘virus titer’ This needs to be corrected, because it is viral RNA that was measured.

Author response: We have corrected this in the revised manuscript (line 198-199).

Line 200: Please specify that IgG was measured in sera.

Author response: We have corrected this in the revised manuscript (line 208).

Line 211: The cells to which the red arrows point are too small to identify as lymphocytes. Please provide an image at higher magnification.

Author response: Thank you for your valuable comment. We have magnified the lymphocytes to make their morphology clearer.

Line 212, 215: The cells to which the green arrows point cannot be identified as alveolar macrophages on morphological criteria alone. Please provide evidence by immunohistochemistry about cell type, or be more conservative in naming the cells (e.g. ‘mononuclear cells’).

Author response: We have corrected alveolar macrophages to mononuclear cells to make the description appropriate in the revised manuscript (line 221, 224).

Fig. 3: Please indicate at which day post inoculation these tissues were obtained.

Author response: We have added this information in the revised manuscript (line 231). These tissues were obtained at 3 dpi.

Line 301: ‘Figure 1’ > ‘Figure 5’

Author response: We have corrected this in the revised manuscript (line 318).

Extended Data Fig. 1: The images are too small in panels a to determine in which histological structure the blue staining is present; and in panels b to see any PAS-positive material. Please provide images at higher magnification.

Author response: Thank you for your valuable comment. We have magnified the details on each figure to make their morphology clearer.

Reviewer Reports on the Second Revision:

Referee #3 (Remarks to the Author):

The authors have responded very well to my comments on the revised manuscript, and I no longer have major concerns.

There are just a few minor comments, mainly on use of technical terms in pathology:

Lines 31, 32, 93, 98, 106, 117, 119: The technical term for these cells of the mononuclear phagocyte system is indeed "monocytes" when they is circulating in the blood. When they migrate from the blood to the tissue as part of the inflammatory response, however, they change their phenotype and are technically called "macrophages". Therefore, in these sentences, I think these cells should be called "macrophages" because they are part of the inflammatory infiltrate in the lung tissue.

Line 85: after 'deflated', insert 'lung lobes'?

Lines 91, 96, 113, 425: Plural of 'septum' is 'septa'. I think you mean 'septa' in these sentences.

Line 267: After mL/g, insert 'body weight'?

Author Rebuttals to Second Revision:

Referee #3 (Remarks to the Author):

The authors have responded very well to my comments on the revised manuscript, and I no longer have major concerns.

Author response: Thank you for your insightful comment.

There are just a few minor comments, mainly on use of technical terms in pathology:

Lines 31, 32, 93, 98, 106, 117, 119: The technical term for these cells of the mononuclear phagocyte system is indeed "monocytes" when they is circulating in the blood. When they migrate from the blood to the tissue as part of the inflammatory response, however, they

change their phenotype and are technically called "macrophages". Therefore, in these sentences, I think these cells should be called "macrophages" because they are part of the inflammatory infiltrate in the lung tissue.

Author response: We have corrected this in the revised manuscript (line 31, 32, 94, 99, 106, 117, 120).

Line 85: after ‘deflated’, insert ‘lung lobes’?

Author response: We have corrected this in the revised manuscript (line 85).

Lines 91, 96, 113, 425: Plural of ‘septum’ is ‘septa’. I think you mean ‘septa’ in these sentences.

Author response: We have corrected this in the revised manuscript (line 91, 96, 113, 431).

Line 267: After mL/g, insert ‘body weight’?

Author response: We have corrected this in the revised manuscript (line 272).