nature research

Peer Review File

Manuscript Title: SARS-CoV-2-reactive T cells in COVID-19 patients and healthy donors

Editorial Notes:

Redactions – Mention of other journals

This document only contains reviewer comments, rebuttal and decision letters for versions considered at *Nature*. Mentions of the other journal have been redacted.

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee #1 (Remarks to the Author):

In this manuscript, Braun et al identify SARS-CoV-2 T cell responses in patients with COVID-19 and controls. The results are surprising because previous studies have not identified virus-specific T cell positivity in control patients, when patients infected with SARS-CoV or MERS-CoV were studied, making it important to validate the results. Most importantly, the data, if validated, would suggest that measurement of SARS-CoV-2 T cell responses would not be useful in prevalence studies or to measure vaccine responses.

Specific comments.
1. The gold standard to identify virus-specific CD4 T cells is peptide stimulation and detection of intracellular cytokines (IFN-g, TNF, IL-2). Using CD40L and 4-1BB could result in false positive results, especially in an inflammatory setting. Similarly, HLA-DR and CD38 are not specific activation markers. Most T cells will upregulate these two markers in an inflammatory environment. Given the importance of the conclusions, the results must be validated with an intracellular cytokine assay.
2. The assays use 1 ug of each peptide (approximately 150 ug) in each well. The assays should be repeated with smaller pools of overlapping peptides, perhaps in a checkerboard format with ICS staining, to further validate the results. Some of the pools should be negative.
3. Negative controls are shown in the supplementary figures. Positive controls should also be shown.
4. The authors conclude that there is likely cross-reactivity with cells primed by the 4 common cold CoV. PBMCs should be stimulated with peptide pools from these viruses to examine this possibility. Supplementary Figure 1 shows substantial homology of SARS-CoV-2 sequences with SARS-CoV sequences but much less so with the common cold CoV. It is unknown whether any of the latter epitopes would cross-react with SARS-CoV-2 epitopes.
5. Patients with severe disease have lower CD4 T cell responses. Were cells harvested from patients with severe disease as viable as those from those with mild disease or healthy donors?
6. Line 168-Were CD40L and 4-1BB also detected on the cell surface? Why was intracellular rather than surface staining used to detect these molecules?

Referee #2 (Remarks to the Author):

April 27, 2020

Dear Editors,

This manuscript by Braun et al. describes SARS-CoV-2 spike protein-reactive CD4+ T cells in peripheral blood of healthy individuals and COVID-19 patients. As CD4+ T cells are critical for clearance of coronavirus infection and lasting immunity, understanding how they react to SARS-CoV-2 is important for vaccine design and furthering our knowledge of disease pathology. This article uses a cohort of healthy donors (68 total) and COVID-19 patients with a spectrum of disease severity (18 total). The vast majority of COVID patients are men (72%) compared to 31% in the healthy population. Using a mix of peptide pools, they analyzed the peripheral blood of these patients to determine which populations of CD4+ T cells specific to the SARS-CoV-2 spike protein expand in COVID-19 patients. The authors found that SARS-CoV-2 naive patients do not possess a population of activated CD4+ T cells reactive to the N-terminus of the spike protein, while COVID-19 have an expanded activated pool of these cells. However, both the "reactive healthy donors" and COVID-19 patients have a population of "activated" CD4+ T cells in peripheral blood that are specific for the C-terminus of the SARS-CoV-2 spike glycoprotein. The authors go on to show that the C-terminus of the spike glycoprotein contains the most homology to the spike glycoprotein of other common endemic coronavirus strains. To the best of our knowledge, this is the first demonstration of this data and may inform future design of CD4+ T cell dominant vaccines.

However, there are several issues that we noticed with the manuscript:

One of the most novel and interesting findings reported is that healthy donors have cross-reactive CD4+ T cells that can respond to the C-terminus of the Spike protein, but not the N terminus. This is a very exciting finding but there are some caveats surrounding this interpretation:

1.The activation status of the healthy donor cells looks very different from the COVID19 individuals- they express lower MFIs of CD40L and 41BB (which should be graphed) and are not double positive for other activation markers seen in COV1D-19 individuals. What is the evidence that these are not just naive cells being activated during the 16 hour stim? Are there any other markers that would clearly designate these cells as memory in healthy donors? Or evidence of the kinetics of activation of naive versus memory cells with populations of PBMCs containing known antigen specific CD4 T cells (i.e. small pox vaccine +/-). While there is significance in the frequency of cells shown in 2D, the differences are very small.

2. Why are there not any virus-specific cells that you highlight as being activated in vivo in COVID-19 individuals (in the abstract) not HLA-DR+CD38+ in the unstimulated condition?

3. Line 160-162 says that the CD38+HLADR+ cells in COVID-19 individuals could not be reactivated. What does this mean and why do we not see those cells in the unstimulated samples? Please explain. Additionally, why are they assumed to be refractory as opposed to the cells being specific for other viral-derived (or really any other) antigens?.

4. Given the breadth of COVID-19 severity in the patient cohort, it would be edifying to stratify the data in Figure 3b, 3c, and 3d with patient outcome or severity. Is there a particular phenotype seen in S-reactive CD4+ T cells in patients with critical COVID-19 disease vs. mild disease? This could be essential in understanding why some patients have drastically distinct outcomes.

5. Similarly, Supp. Figure 2 shows the gating of the CD4+ T cells in all of the patients analyzed. It can be seen that in most of the "critical" COVID-19 patients (Patients 8, 12, 14, 15, 20 21), almost none of the S-reactive CD4+ T cells express high levels of Ki67 expression, but this is not discussed. Further analysis of this, combined with the previous point considering stratifying the data with disease severity or outcome, would add much importance to the findings. This lack of KI67 after restimulation could be an interesting biomarker for severe disease.

Minor points

Line 62: The authors claim that COVID-19 patients possess CD38+ and HLA-DR+ cells S-reactive cells that healthy donors do not, "indicating their recent in vivo activation". However, the data in Figure 3a which supports this claim was taken from cells stimulated with the S-protein, which would indicate that they were activated ex vivo. The authors need to remove this claim of "recent in vivo activation" and specify after in vitro restimulation.

Figure 2e: No statistics are present on this graph.

Supp. Figure 2: Inconsistent gating of 4-1BB x CD40L between patients. Please make gating consistent and update any associated data.

Supp. Figure 4: No positive or negative control present, which makes the data difficult to interpret.

Figure 2d: What does the data look like in the non-reactive healthy donors?

Referee #3 (Remarks to the Author):

In this study, through the use of a flow cytometry-based assay, the authors identified antigen-specific CD4 T cells that are directed against the spike glycoprotein protein (S) of SARS-CoV2. S-specific CD4 T cells from COVID-19 patients equally target both N- and C- portion of the protein, are highly prevalent among infected patients (67% and 83% against N- and C- portions, respectively), and harbor surface phenotypes indicative of recent in vivo TCR engagement (CD38, HLA-DR, and Ki-67). Importantly, 34% of SARS-CoV-2 seronegative healthy donors (HD) also harbor S-reactive CD4 T cells, which are found at lower frequency, preferentially target C-portion of the S protein, and do not express markers of recent TCR activation. Overall, this study is well written, well executed, and presents interesting insights. However, this study falls short of correlating the T cell response to disease outcome.

1) Were there any correlation with disease symptoms and the T cell responses seen among the

infected? Any correlation with age, sex, comorbidities?
2) Why was only one antigen examined? What is the rational for focusing on the S protein? Why were CD4 but not CD8 response examined?

3) In order to identify S-specific CD4 T cells, the authors designed an assay that used intracellular CD40L and 4-1BB as a readout to measure reactivity to S protein epitope mixtures. A main concern for the use of this assay, however, is the identification of background or false positive signals as bona fide CD40L and 4-1BB expression on CD4 T cells. As the identification of S-cross-reactive CD4 T cells in HD (which are at lower frequency compared to those in COVID-19 patients) is critically influenced by the sensitivity and specificity of the measurement, the following issues will need to be addressed to strengthen the validity of this approach.

3.1) First, based on data presented in Supplementary Fig. 2 it appeared that individual patient samples were subject to different gating strategy for the identification of CD40L+4-1BB+ cells. What was the rationale for doing so, instead of using the same gating across all samples? How was the gating determined? How robust were the main findings (i.e. frequency of S-reactive cells out of total CD4 T cells per sample as well as prevalence of S-reactive CD4 T cells among patients and HD) to changes in the gating strategy?

3.2) Second, dot plots showing CD40L, 4-1BB, CD38, HLA-DR, and Ki67 staining of stimulated HD samples, along with no-peptide-stimulation controls, are missing and should be provided (similar to Supplementary Fig. 2). Positive controls for the peptide stimulation assay (stimulated with SEB/TSST1 or PepMix HCMVA) should be provided as well.

3.3) Lastly, the authors ruled out the possibility that the presence of S-reactive CD4 T cells in HD could be due to early SARS-CoV-2 infection. However, only 10 donors (43.5% out of all subjects) were tested (negative) for viruses with nasopharyngeal swabs. The observation that none of the healthy donors developed antibody responses also did not suffice to rule out infection as antibodies take weeks to develop and reach detectable levels. Were there any other measurements or indications to confirm that the recruited subjects were uninfected at the time of sampling?

4) Through the combinatorial use of CD38, HLA-DR, and Ki-67, the authors convincingly demonstrated that CD4 T cells from the majority of COVID-19 patients that were reactive to ex vivo peptide stimulation also had recently undergone TCR activation in vivo. CD38 and HLA-DR (co)expression also well distinguished recently activated SARS-CoV-2-specific CD4 T cells found in COVID-19 patients from pre-existing S-cross-reactive CD4 T cells in HD that were presumably induced by previous endemic coronavirus strains.

While these efforts at characterizing the antigen-experienced state are commendable and highly informative, a thorough assessment of the nature of these antigen-specific CD4 T cell responses is needed. In particular, surface (CCR7, CD45RA) and intracellular markers (IFNγ, IL-4, IL-17, IL-21) indicative of the differentiation state and functional capacity of these CD4 T cells should be examined. Similarly, surface and intracellular phenotypes of S-specific CD4 T cells identified in reactive healthy donors (RHD) should be measured and compared with those found in COVID-19 patients.

5) S1-specific IgG titers well segregate COVID-19 patients from RHD, suggesting these antibodies were developed as a result of exposure to the ongoing pandemic. Is there any correlation between virus-specific CD4 T cells and IgG titers? Could the absence of virus-specific CD4 T cells explain low antibody titers found in the several COVID-19 patients (Fig. 2f). If so, which parameter(s) on CD4 T cells (CD40L, 4-1BB, CD38, HLA-DR, and Ki-67) best predict the development of antibody responses?

6) Correlations (and related statistical analysis) between CD38 and HLA-DR and between CD38 and Ki-67 in Fig. 3e-f are missing.

7) Statistical analysis for comparisons of S-I-reactive or S-II-reactive CD4 T cell frequencies between mild, severe, and critical patients in Supplementary Fig. 3 is missing. It is also unclear how the authors reached the conclusion that most non-S-II-reactive COVID-19 patients have critical disease stage as there does not seem to be a noticeable difference of S-II-reactive CD4 T cell frequencies between different disease stages.

Author Rebuttals to Initial Comments:

Response to the Reviewers
We would like to thank all reviewers for the time and effort spent on evaluating our work. Their critical comments and suggestions have helped us to substantially improve our manuscript and we are very grateful for this valuable input.

Point-to-Point responses:

> Referee #1:
>
> General comment.
> In this manuscript, Braun et al identify SARS-CoV-2 T cell responses in patients with COVID-19 and controls. The results are surprising because previous studies have not identified virus-specific T cell positivity in control patients, when patients infected with SARS-CoV or MERS-CoV were studied, making it important to validate the results. Most importantly, the data, if validated, would suggest that measurement of SARS-CoV-2 T cell responses would not be useful in prevalence studies or to measure vaccine responses.

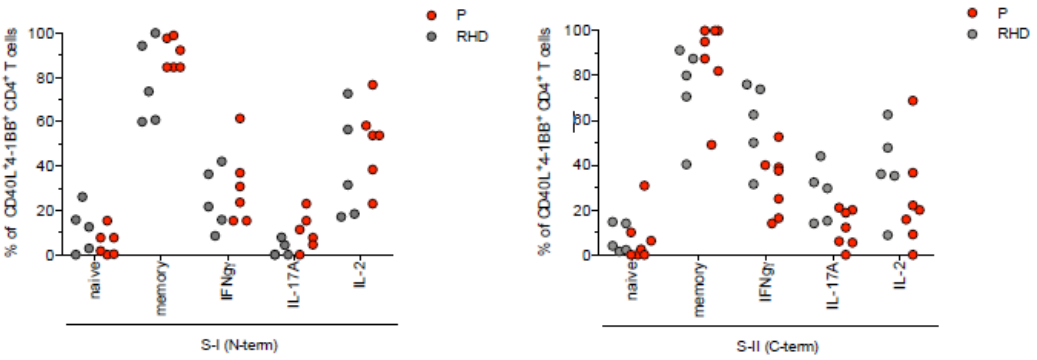
We thank the reviewers for acknowledging the novelty of our findings. Indeed, T cell cross-reactivity of SARS- and MERS-CoV-reactive T cells with endemics hCoVs has not been reported so far (Yang et al., 2006, [REDACTED]; Yang et al., 2007, [REDACTED]; Li et al., 2008, JI; Fan et al., 2009, [REDACTED]). In our opinion, the apparent discrepancy is mainly of technical nature. The previous studies analyzed only a very limited number of healthy control donors. In addition, lower cell numbers were stimulated due to culturing restrictions of an IFNγ ELISPOT assay. The latter results in an approximately log10 higher detection threshold compared to our assay. Importantly,
J. Zhao et al. (Zhao et al., 2016, [REDACTED]) reported that an HLA-II-restricted epitope of the SARS-CoV N protein mediates cross-protection to MERS-CoV in an HLA-DR transgenic mouse strain. This finding indicates that T cell cross-reactivity between different CoVs but can provide cross-protection against other CoV strains in experimental models. We respectfully disagree with the notion that our findings question the usefulness of measuring SARS-CoV-2 specific T cell responses as a read out for

vaccine responses. Our findings indicate that markers of recent activation, such as HLA-DR and CD38, discriminate resting antigen-specific (memory) T cells from activated effector/memory T cells. In contrast, IFN□ ELISPOT assays for detection of antigen-specific T cells cannot distinguish between these important types of responses.

Referee #1:
Specific comments.

1. The gold standard to identify virus-specific CD4 T cells is peptide stimulation and detection of intracellular cytokines (IFN-g, TNF, IL-2). Using CD40L and 4-1BB could result in false positive results, especially in an inflammatory setting. Similarly, HLA-DR and CD38 are not specific activation markers. Most T cells will upregulate these two markers in an inflammatory environment. Given the importance of the conclusions, the results must be validated with an intracellular cytokine assay.

The reviewer is concerned about background activation and false positive results. However, we have applied very stringent gating and background correction based on positive and negative controls (CMVpp65, SEB, unstimulated controls) to exclude false positive signals from the analysis. Moreover, we used CD40L and 4-1BB on purpose to identify the entire population of antigen-reactive CD4+ T cells, rather than just cytokine producers, and we have previously found these markers to be highly specific and suitable for detection of even very low frequencies of antigen-reactive T cells directly (Frentsch et al., 2005, [REDACTED].; Meier et al., 2008, [REDACTED]). Moreover, for 4-1BB we could previously demonstrate that its upregulation in the course of short-term activation in vitro is highly dependent on TCR-engagement, excluding the possibility of bystander activation through cytokines (Sattler et al., 2009, Blood). Both markers have been widely used since their introduction, also in clinical diagnostic analyses (Bacher et al., 2015, [REDACTED]; Bacher et al., 2016, [REDACTED]; Bacher et al., 2019, [REDACTED]).
Nonetheless, we agree with the reviewer that the analysis of intracellular cytokine expression after short-term activation in vitro (e.g. IFN-□, IL-2) is highly informative with respect to the functional status of antigen-specific T cells. Accordingly, we performed additional analyses, which we have now added the new Figure 2h, comparing frequencies of S-I and S-II-reactive IFN-□+ CD40L+ CD4+ T cells in reactive healthy donors (RHD) and COVID-19 patients (P). The data demonstrate not significant difference in IFN□ production between RHD and COVID-19 patients, which demonstrates the suitability of our assay to detect specific T cells.
In addition, we also compared expression patterns of multiple other cytokines such as IFNγ, IL-2 and IL-17A in 7 additional patients (P) and 5 reactive healthy donors (RHD) (Supplementary Fig. 3). Similar cytokine expression patterns were demonstrated in patients and RHD.

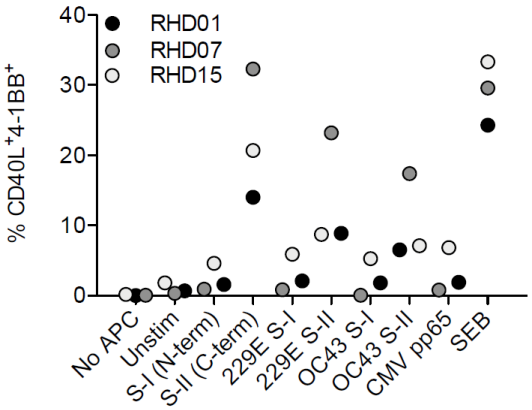


Referee #1: Specific comments.

2. The assays use 1 ug of each peptide (approximately 150 ug) in each well. The assays

should be repeated with smaller pools of overlapping peptides, perhaps in a checkerboard format with ICS staining, to further validate the results. Some of the pools should be negative.

We thank the reviewer for this important remark. We agree that high total amounts of peptides within peptide pools have to be used with care. However, comparable pools as we used here have been used and tested extensively in various studies demonstrating highly specific activation (Kern et al., 2000, EJI; Maecker et al., 2001, [REDACTED]; Kiecker et al., [REDACTED]).
To further underscore the specificity of the response, we employed peptide pool stimulation for direct enrichment and expansion of 4-1BB+CD40L+ CD4+ T cells from RHD. The resulting SARS-CoV2 S-II-specific T cell lines reacted specifically to SARS- CoV2 S-II, but not to SARS-CoV-2 S-I, or CMVpp65. Moreover, the same SARS-CoV2 S-II-specific T cell lines reacted to similarly well to SHCoV-II peptide pools of OC43 and 229E hCoVs (Fig. 2j). These additional experiments provide clear evidence that the measured activation is antigen-specific and not an artefact of high peptide concentration (Fig. 2j)



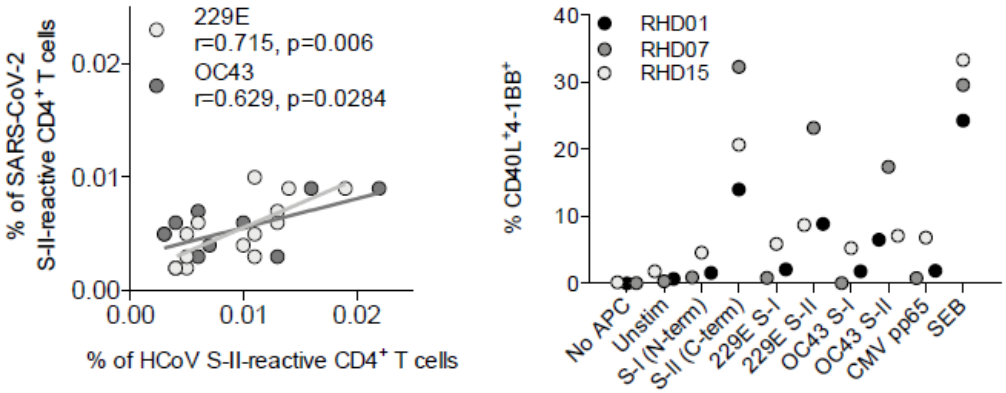
Referee #1:
Specific comments.
3. Negative controls are shown in the supplementary figures. Positive controls should also be shown.

As suggested, we have now included the related positive controls. Since referee #3 requested also to show gating of HD and RHD, we have revised Supplementary Fig. 2 and now present three examples of each donor group. We have further added a new Extended data file, in which we have compiled the gating of all samples for each donor group.

Referee #1:
Specific comments.
4. The authors conclude that there is likely cross-reactivity with cells primed by the 4 common cold CoV. PBMCs should be stimulated with peptide pools from these viruses to examine this possibility.

We thank the reviewer for this important suggestion. We stimulated PBMC from 12 RHD with spike glycoprotein I and II peptide pools from HCoV 229E and OC43 in comparison to stimulation with SARS-CoV-2 S II pools to assess T cell reactivities. Bivariate plotting of SARS-CoV2-S-II-reactive CD4+ T cells and HCoV-S-II CD4+ T cells in RHD revealed a strong positive correlation of T cell responses against the three different pools. The results are now included in the new Fig. 2i. Moreover, we enriched and expanded SARS-CoV2-S-II-reactive CD4+ T cells and restimulated the resulting T cell

lines with S-I and SII peptide pools from SARS-CoV-2 and the two endemic HCoVs 229E and OC43. The results are shown in the new Fig. 2j. SARS-CoV-2 S-II-specific T cell lines reasponded specifically to SARS-CoV2 S-II, but not to SARS-CoV2 S-I. Importantly, resulting SARS-CoV2 S-II-specific T cell lines reacted also to S-II peptide pools of OC43 and 229E strains, formally demonstrating their cross-reactivity and suggesting that previous infections with HCoVs may have induced SARS-CoV-2 S-cross-reactive CD4+ T cells.



Referee #1:
Specific comments.
Supplementary Figure 1 shows substantial homology of SARS-CoV-2 sequences with SARS-CoV sequences but much less so with the common cold CoV. It is unknown whether any of the latter epitopes would cross-react with SARS-CoV-2 epitopes.

We agree with the reviewer that the homology between SARS-CoV-2 sequences with SARS-CoV sequences is much higher, but clearly cross-reactive T cell responses may depend only on single MHC-II restricted epitopes. Hence a short stretch of homology in the amino acid sequence can be sufficient to elicit cross-reactivity.

Additionally, our new data discussed above in Fig. 2i and 2j emphasize that the presence of SARS-CoV-2-reactive CD4+ T cells in seronegative healthy donors is based on the induction of such cells during previous encounters with HCoVs.

Referee #1:
Specific comments.
5. Patients with severe disease have lower CD4 T cell responses. Were cells harvested from patients with severe disease as viable as those from those with mild disease or healthy donors?

We regularly checked cell viability with a viability dye marker and detected no differences comparing cells from severe disease patients with less severely affected patients. One explanation for the observed lower CD4+ T cell response could be the strong lymphopenia in severe COVID-19, which has been observed in other studies (Huang et al., 2020, [REDACTED]; Wang et al., 2020, [REDACTED]). Another reason could be as well exhaustion in particular of SARS-CoV-2-specific T cells, as recently demonstrated (Diao et al., 2020, [REDACTED].).

Referee #1:
Specific comments.

6. Line 168-Were CD40L and 4-1BB also detected on the cell surface? Why was intracellular rather than surface staining used to detect these molecules?

The reviewer is right that both, CD40L and 4-1BB, can be detected and analysed on the cell surface. However, we have previously demonstrated a higher sensitivity in detecting antigen-reactive CD40L+ CD4+ T cells when stained intracellularly (Meier et al., 2008, [REDACTED]).

Referee #2:

General comment.
This manuscript by Braun et al. describes SARS-CoV-2 spike protein-reactive CD4+ T cells in peripheral blood of healthy individuals and COVID-19 patients. As CD4+ T cells are critical for clearance of coronavirus infection and lasting immunity, understanding how they react to SARS-CoV-2 is important for vaccine design and furthering our knowledge of disease pathology. This article uses a cohort of healthy donors (68 total) and COVID-19 patients with a spectrum of disease severity (18 total). The vast majority of COVID patients are men (72%) compared to 31% in the healthy population. Using a mix of peptide pools, they analyzed the peripheral blood of these patients to determine which populations of CD4+ T cells specific to the SARS-CoV-2 spike protein expand in COVID-19 patients. The authors found that SARS-CoV-2 naive patients do not possess a population of activated CD4+ T cells reactive to the N-terminus of the spike protein, while COVID-19 have an expanded activated pool of these cells. However, both the "reactive healthy donors" and COVID-19 patients have a population of "activated" CD4+ T cells in peripheral blood that are specific for the C-terminus of the SARS-CoV-2 spike glycoprotein. The authors go on to show that the C-terminus of the spike glycoprotein contains the most homology to the spike glycoprotein of other common endemic coronavirus strains. To the best of our knowledge, this is the first demonstration of this data and may inform future design of CD4+ T cell dominant vaccines.

However, there are several issues that we noticed with the manuscript:

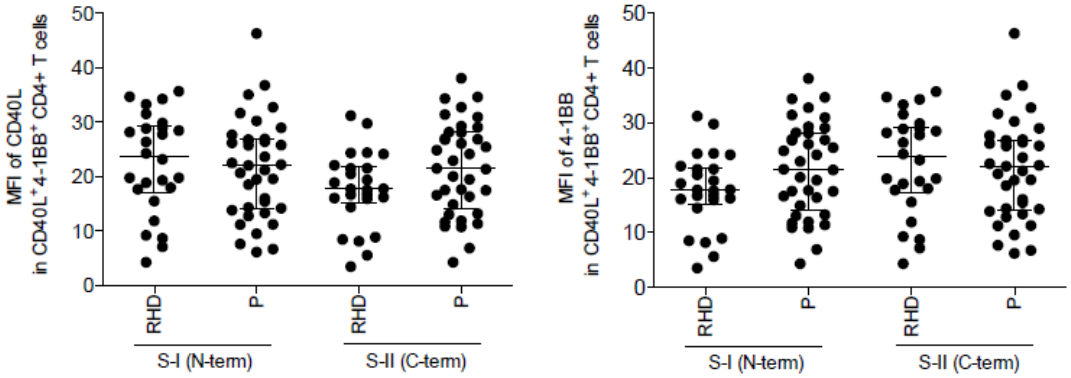
One of the most novel and interesting findings reported is that healthy donors have cross-reactive CD4+ T cells that can respond to the C-terminus of the Spike protein, but not the N terminus. This is a very exciting finding but there are some caveats surrounding this interpretation:

We thank the reviewer for these encouraging comments and the positive assessment of our study.

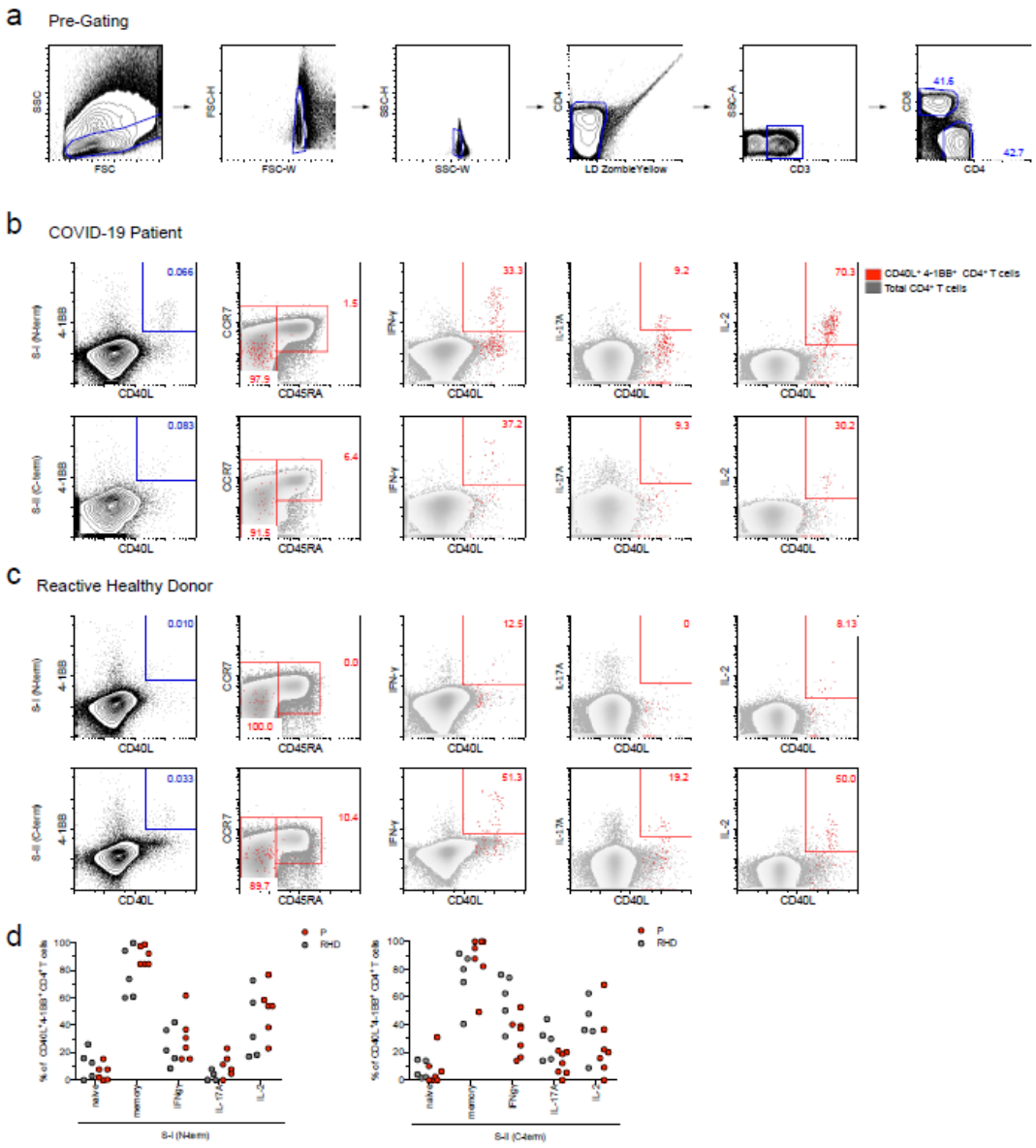
Referee #2:
1. The activation status of the healthy donor cells looks very different from the COVID19 individuals- they express lower MFIs of CD40L and 41BB (which should be graphed) and are not double positive for other activation markers seen in COVID-19 individuals. What is the evidence that these are not just naive cells being activated during the 16 hour stim? Are there any other markers that would clearly designate these cells as memory in healthy donors? Or evidence of the kinetics of activation of naive versus memory cells with populations of PBMCs containing known antigen specific CD4 T cells (i.e. small pox vaccine +/-). While there is significance in the frequency of cells shown in 2D, the differences are very small.

Based on the reviewer´s observation, we analyzed the MFI levels of both markers. The results (figure below) demonstrate that COVID-19 individuals cannot be distinguished from SARS-CoV-2-reactive

HD (=RHD) by the MFI of CD40L or 4-1BB.



The question, whether we may measure naïve T cell responses in RHD is important . Accordingly, we assessed the phenotype of the S-II-reactive T cells by measuring CCR7 and CD45RA in five reactive healthy donors and seven patients. This analysis revealed that almost all S-II-reactive CD4+ T cells displayed a memory phenotype in COVID-19 patients as well as in RHD (Supplementary Fig. 3).
We furthermore analyzed the cytokine profile of CD40L+ 4-1BB+ CD4+ T cells, demonstrating high frequencies of IFNγ producing cells among S-II-reactive CD4+ T cells (Supplementary Fig. 3), consistent with a TH1 response.



Referee #2:

2. Why are there not any virus-specific cells that you highlight as being activated in vivo in COVID-19 individuals (in the abstract) not HLA-DR+CD38+ in the unstimulated condition?

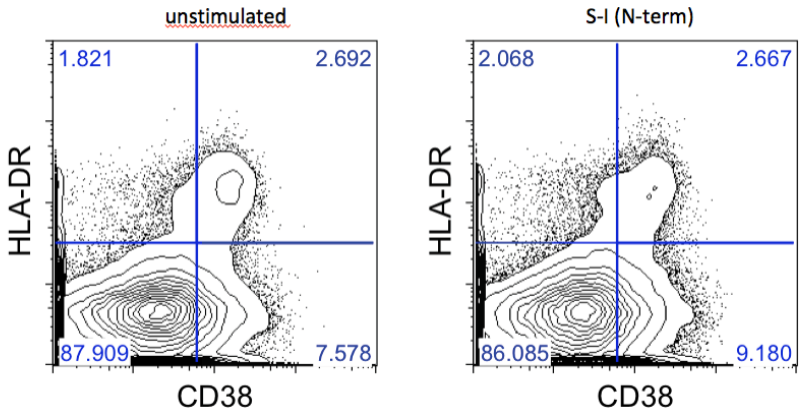
In our cohort of COVID-19 patients low frequencies of S-reactive T cells are the most likely explanation. Frequency alterations in the range of 0.002-0.2% may be not reliable detectable in the large populations of in vivo activated HLA-DR+ CD38+ CD4+ T cell population.

Referee #2:

3. Line 160-162 says that the CD38+HLADR+ cells in COVID-19 individuals could not be reactivated. What does this mean and why do we not see those cells in the unstimulated samples? Please explain. Additionally, why are they assumed to be refractory as opposed to the cells being specific for other viral- derived (or really any other) antigens?.

Supplementary Fig. 2 displays the expression of CD38 and HLA-DR among CD40L+4- 1BB+ CD4+ T cells. In some patients we observed high frequencies of CD38+HLA-DR+

CD4+ T cells as compared to HD, unchanged by stimulation (see figure below). Recent reports demonstrated an increase of exhaustion markers PD-1 and TIM3 on T cells in COVID-19 patients that may be indicative of a refractory functional state of T cells from COVID-19 patients (Diao et al., 2020, Front. Immunol.), which would result in a fail to respond and upregulate CD40L and 4-1BB in specific cells upon restimulation in vitro.



Referee #2:

4. Given the breadth of COVID-19 severity in the patient cohort, it would be edifying to stratify the data in Figure 3b, 3c, and 3d with patient outcome or severity. Is there a particular phenotype seen in S-reactive CD4+ T cells in patients with critical COVID-19 disease vs. mild disease? This could be essential in understanding why some patients have drastically distinct outcomes.

We agree with the referee and thank him/her for this important remark. This would indeed provide valuable information. However, given the limited group size of our COVID-19 patient cohort, 18 patients subdivided into three groups of seven patients with mild, five patients with severe, and six patients with critical disease onset (Table 1). For a reliable correlation of T cell responses with disease symptoms, higher numbers of patients would likely be required. Partially, this question has been addressed with an intermediate-sized patient cohort in another recent preprint (Anft et al., 2020, medRxiv). While this is a critically important topic to be examined by the scientific community, we would like to emphasize that the novelty of our work derives from the identification and characterisation of SARS-CoV-2-cross-reactive CD4+ T cells in seronegative healthy donors.

Referee #2:

5. Similarly, Supp. Figure 2 shows the gating of the CD4+ T cells in all of the patients analyzed. It can be seen that in most of the “critical” COVID-19 patients (Patients 8, 12, 14, 15, 20 21), almost none of the S-reactive CD4+ T cells express high levels of Ki67 expression, but this is not discussed. Further analysis of this, combined with the previous point considering stratifying the data with disease severity or outcome, would add much importance to the findings. This lack of KI67 after restimulation could be an interesting biomarker for severe disease.

We are grateful for this valuable comment. However, Ki67 expression underlies a timepoint sensitive kinetic and would require standardized sampling relative to the time of infection that cannot be determined accurately for most of the patients. Patients from our cohort were analyzed at different days after first symptoms were (Table 1). Larger cohort studies are required to assess this, which would be beyond the scope of this paper.

Referee #2:
Minor points
Line 62: The authors claim that COVID-19 patients possess CD38+ and HLA- DR+ cells S-reactive cells that healthy donors do not, “indicating their recent in vivo activation”. However, the data in Figure 3a which supports this claim was taken from cells stimulated with the S-protein, which would indicate that they were activated ex vivo. The authors need to remove this claim of “recent in vivo activation” and specify after in vitro restimulation.

Please refer to our response to comment #3, where we elaborate on CD38 and HLA- DR kinetics. Indeed, the cells are ex vivo activated by stimulation with spike peptide pools, but exhibit high CD38 and HLA-DR expression in patients, as compared to RHD, due to the long kinetic of these markers, which indicates “their recent in vivo activation”.

Referee #2:
Minor points
Figure 2e: No statistics are present on this graph.

Thank you for bringing this to our attention. We have added the statistic accordingly.

Referee #2:
Minor points
Supp. Figure 2: Inconsistent gating of 4-1BB x CD40L between patients. Please make gating consistent and update any associated data.

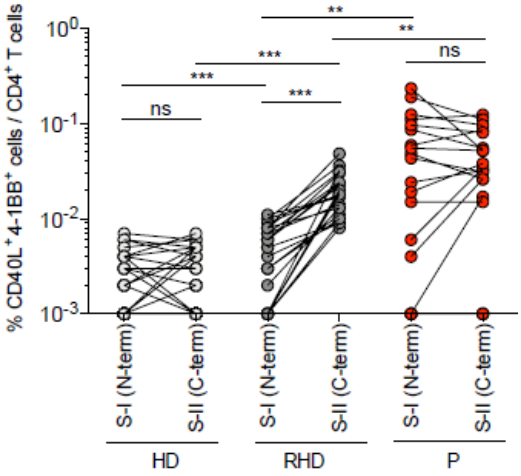
We included the negative and positive controls CMVpp65 peptide pool and SEB/TSST- 1 into the Supplementary Fig. 2 in order to highlight the observed inter-donor variability in marker expression intensity. The gating for 4-1BB/CD40L was donor-individually determined based on unstimulated background controls. Background levels vary from donor to donor, which is why it should always be determined and used as gating control.

Referee #2:
Minor points
Supp. Figure 4: No positive or negative control present, which makes the data difficult to interpret.

All humans have high antibody titers against the 4 common cold coronaviruses (HCoV 229E, NL63, OC43 and HKU1), which makes human positive/negative control sera largely unavailable. However, all applied coronavirus spike proteins harbor a C-terminal FLAG-tag. The FLAG-epitopes are detected by immunofluorescence test using a mouse anti-FLAG and a goat-anti mouse Alexa488. As transfection efficiency is approximately 30-50% there are always non-transfected VeroB4 cells on the glass slides that serve as reference for non-specific binding of antibodies. An example of the recombinant spike-based immunofluorescence test (here SARS-CoV-2) can be found in Wölfel, Corman et al., Nature 2020.

Referee #2:
Minor points
Figure 2d: What does the data look like in the non-reactive healthy donors?

We have now included the information about non-reactive healthy donors into the graph (Fig. 2d).



Referee #3:

General comments.
In this study, through the use of a flow cytometry-based assay, the authors identified antigen-specific CD4 T cells that are directed against the spike glycoprotein protein (S) of SARS-CoV2. S-specific CD4 T cells from COVID-19 patients equally target both N- and C-portion of the protein, are highly prevalent among infected patients (67% and 83% against N- and C- portions, respectively), and harbor surface phenotypes indicative of recent in vivo TCR engagement (CD38, HLA-DR, and Ki-67). Importantly, 34% of SARS-CoV-2 seronegative healthy donors (HD) also harbor S-reactive CD4 T cells, which are found at lower frequency, preferentially target C-portion of the S protein, and do not express markers of recent TCR activation. Overall, this study is well written, well executed, and presents interesting insights. However, this study falls short of correlating the T cell response to disease outcome.

We thank the reviewer for this acknowledging the importance of our study and for the encouraging comments.

Referee #3:
1) Were there any correlation with disease symptoms and the T cell responses seen among the infected? Any correlation with age, sex, comorbidities?

We thank the reviewer for this important comment. We agree that correlations of T cell responses and clinical data would be very interesting to investigate. However, given the limited sample size of our patient cohort, which was subdivides into seven COVID- 19 categories of patients with mild, five patients with severe and six patients with critical disease onset based on their symptoms, such stratification would not yield statistically valid conclusions. However, work of this nature has been recently published in an intermediate-sized patient cohort in a recent preprint (Anft et al., 2020, medRxiv). However large cohort studies would be required for in depth analyses in near future.

Referee #3:
2) Why was only one antigen examined? What is the rational for focusing on the S protein? Why were CD4 but not CD8 response examined?

We thank the reviewer for commenting on this issue. Spike glycoprotein was selected for several reasons. First of all, because it has been demonstrated to be a major target of neutralizing antibodies (Wang et al., 2020, [REDACTED].; Nie et al., 2004, J. [REDACTED].; Hofmann et al., 2004, [REDACTED]), and CD4+ T cell responses are instrumental for high-affinity antibody responses. Moreover, S is the target of nearly all vaccine initiatives worldwide, yet there is no information on the T cell immunogenicity of S. We also analyzed CD8+ T cell responses to spike glycoprotein peptide pools, however detected only weak responses in HD and patients. One reason could be the suboptimal peptide length of 15mers present in our peptide pools, since MHC-I dependent presentation would requires shorter peptides. The finding of SARS-CoV-S-II reactive CD4+ T cells in HD further shifted our focus to a more comprehensive analysis of these cells.

Referee #3:

3) In order to identify S-specific CD4 T cells, the authors designed an assay that used intracellular CD40L and 4-1BB as a readout to measure reactivity to S protein epitope mixtures. A main concern for the use of this assay, however, is the identification of background or false positive signals as bona fide CD40L and 4-1BB expression on CD4 T cells. As the identification of S-cross-reactive CD4 T cells in HD (which are at lower frequency compared to those in COVID- 19 patients) is critically influenced by the sensitivity and specificity of the measurement, the following issues will need to be addressed to strengthen the validity of this approach.

Referee #3:

3.1) First, based on data presented in Supplementary Fig. 2 it appeared that individual patient samples were subject to different gating strategy for the identification of CD40L+4-1BB+ cells. What was the rationale for doing so, instead of using the same gating across all samples? How was the gating determined? How robust were the main findings (i.e. frequency of S-reactive cells out of total CD4 T cells per sample as well as prevalence of S-reactive CD4 T cells among patients and HD) to changes in the gating strategy?

Once again, we thank the reviewer for his/her helpful comment. Indeed, the gates were individually fitted according to CMVpp65 peptide pool and SEB/TSST-1 stimulations as well as unstimulated controls. The reason for the differences in the gatings may originate from batch effects since the samples were processed on different days. The gating for 4-1BB/CD40L was donor-individually determined based on these controls. For full transparency, we have now included the corresponding positive controls for each individual donor in Supplementary Fig. 2 as well an `all donor overview´ in the Extended data file.

Referee #3:

3.2) Second, dot plots showing CD40L, 4-1BB, CD38, HLA-DR, and Ki67 staining of stimulated HD samples, along with no-peptide-stimulation controls, are missing and should be provided (similar to Supplementary Fig. 2). Positive controls for the peptide stimulation assay (stimulated with SEB/TSST1 or PepMix HCMVA) should be provided as well.

We appreciate this comment and included the requested information of HD & RHD and the plots of the positive controls in Supplementary Fig. 2, which now presents three representative examples of each donor group. Furthermore, we added a new Extended data file, in which we have compiled the gating of all samples for each donor group.

Referee #3:

3.3) Lastly, the authors ruled out the possibility that the presence of S-reactive CD4 T

cells in HD could be due to early SARS-CoV-2 infection. However, only 10 donors (43.5% out of all subjects) were tested (negative) for viruses with nasopharyngeal swabs. The observation that none of the healthy donors developed antibody responses also did not suffice to rule out infection as antibodies take weeks to develop and reach detectable levels. Were there any other measurements or indications to confirm that the recruited subjects were uninfected at the time of sampling?

This is a very important issue. In order to address this, we were able to re-recruit 65 of the 68 HD and conducted a second round of ELISA IgG S1 testing for SARS-CoV-2 specific antibodies, including 24 out of 25 reactive healthy donors. All donors were still seronegative for SARS-CoV-2, confirming their status as SARS-CoV-2 naïve. Results of this second test are shown now in Fig. 2g.

Referee #3:

4) Through the combinatorial use of CD38, HLA-DR, and Ki-67, the authors convincingly demonstrated that CD4 T cells from the majority of COVID-19 patients that were reactive to ex vivo peptide stimulation also had recently undergone TCR activation in vivo. CD38 and HLA-DR (co)expression also well distinguished recently activated SARS-CoV-2-specific CD4 T cells found in COVID-19 patients from pre-existing S-cross-reactive CD4 T cells in HD that were presumably induced by previous endemic coronavirus strains.

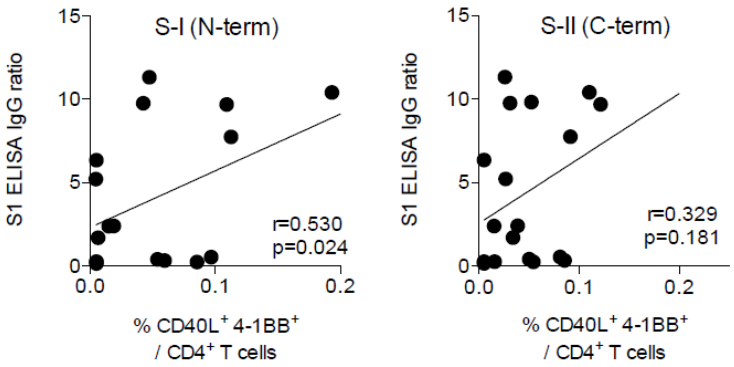
While these efforts at characterizing the antigen-experienced state are commendable and highly informative, a thorough assessment of the nature of these antigen-specific CD4 T cell responses is needed. In particular, surface (CCR7, CD45RA) and intracellular markers (IFNγ, IL-4, IL-17, IL-21) indicative of the differentiation state and functional capacity of these CD4 T cells should be examined. Similarly, surface and intracellular phenotypes of S-specific CD4 T cells identified in reactive healthy donors (RHD) should be measured and compared with those found in COVID-19 patients.

We thank the reviewer for this suggestion. We now performed intracellular cytokine measurement of IL-2, IL-17A and IFNγ as well as of CCR7 and CD45RA in SARS-CoV- 2-reactive CD4+ T cells from seven patients and five RDH (new Supplementary Fig. 3). We found that almost all SARS-CoV-2 reactive T cells displayed a memory phenotype in patients as well in the RHD. Furthermore, no differences in the cytokine profile between patients and RHD were observed, elucidating a comparable functionality and predominating Th1 cytokine profile.

Referee #3:

5) S1-specific IgG titers well segregate COVID-19 patients from RHD, suggesting these antibodies were developed as a result of exposure to the ongoing pandemic. Is there any correlation between virus-specific CD4 T cells and IgG titers? Could the absence of virus-specific CD4 T cells explain low antibody titers found in the several COVID-19 patients (Fig. 2f). If so, which parameter(s) on CD4 T cells (CD40L, 4-1BB, CD38, HLA-DR, and Ki-67) best predict the development of antibody responses?

We thank the reviewer for this interesting suggestion. Accordingly, we plotted the frequencies of S-I and S-II-specific CD4+ T cells against IgG titers and observed for S- I-reactive cells a weak positive correlation to anti-S1 IgG titers (Pearson correlation).



However, we have decided against including these figures in the manuscript, since sampling time points of the different severity level COVID-19 patients varied substantially. Moreover, there are many more factors influencing antibody titer levels which is why such a plot could easily be misleading without further investigations.

Referee #3:

6) Correlations (and related statistical analysis) between CD38 and HLA-DR and between CD38 and Ki-67 in Fig. 3e-f are missing.

We apologize for this oversight and have now added the Pearson correlation coefficient and the related p values into the diagram (Fig. 3 e,f).

Referee #3:

7) Statistical analysis for comparisons of S-I-reactive or S-II-reactive CD4 T cell frequencies between mild, severe, and critical patients in Supplementary Fig. 3 is missing. It is also unclear how the authors reached the conclusion that most non-S-II-reactive COVID-19 patients have critical disease stage as there does not seem to be a noticeable difference of S-II-reactive CD4 T cell frequencies between different disease stages.

We assume the referee is referring to this sentence:
“Most non-reactive COVID-19 patients were characterized by critical disease states“. Indeed, this statement as written is not correct, since we refer to the low S-I-reactivity in critical patients. We apologize this mistake and thus, change the sentence in the corrected manuscript to:
„Most COVID-19 patients with critical disease state exhibited no reactivity to S-I (N- term)“. We thank the reviewer for his/her careful reading!

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

In this revised manuscript, Braun et al respond appropriately to the comments of the reviewers. The results demonstrate that healthy donors responded to SARS-CoV-2 peptide pools, suggesting pre-existing immune responses to the virus. Initially, this was demonstrated by measuring two markers of cellular activation, raising the concern that this activation was non-specific. In the revised manuscript, the authors show that in some of these patients, these cells also expressed IFN-gamma.

These results support the conclusion that cross-reactivity is specific. The results still remain somewhat enigmatic because there is not much homology between the S2 sequences of the endemic CoV and SARS-CoV-2. However, these results are in general agreement with a previous publication from the Crotty/Sette groups ([REDACTED], online).

Specific comments:
1. While the IFN-g results are good, assessing cells double labeled for IFN-g and TNF is even more specific. Were most of the IFN-g cells also TNF-expressing? If so, some of these data should be included.
2. Figure 1C. The IFN-g positivity is 0.01% but in the extended data, the IFN-g positivity was marked as 0%. This should be rectified.
3. The title to Figure SF6 is not clear and should be modified.

Referee #2 (Remarks to the Author):

We thank the authors for their efforts to answer our questions and concerns. The key novel findings initially were that COVID patients (P) and those with potentially cross-reactive CD4+ T cells, termed "reactive healthy donors" or RHD, both had responses to the C-terminal region of the spike protein, which is the region with the most homology with other coronaviruses. COVID patients, however, only had antigen-specific T cell responses to the Spike N-terminal. The difference in the percent of CD40L+41BB+ cells as shown in Fig 2d in the RHD group versus HD group (or even compared to the S-I specific cells in RHD patients) is significant, but extremely small (mean ~ 10-2 in RHD vs 10-2.5 in HD individuals- although the representative flow plots for HD show 0%). These data still raise concerns about the significance of these findings.
In our first review, we were very concerned that the assessment of 4-1BB and CD40L expression on T cells after a 16 hour peptide stimulation would reveal naive cells that could be activated by the culture conditions, and not truly identify previously activated memory cells that suggest cross reactivity. A major concern was the lack of quantification of the MFI of CD40L and 4-1BB, which appeared to be different. In the current data included in the rebuttal, which should also be included as a supplement in the manuscript, they have now quantified the MFI of these markers in the P and RHD groups in response to both regions of the spike, and there are no differences in either region. Surprisingly however, the number of dots presented in these data now suggest that there are equivalent numbers of CD40L 4-1BB+ cells across both groups and both regions of the protein, which is very confusing. This needs to be addressed.
To further demonstrate that these were not naive cells activated by culture conditions, we suggested looking for better indicators of memory status. The authors first looked at CCR7 and CD45RA on CD40L+41BB+ cells after peptide stimulation, but these too could be down-regulated by a 16 hour culture, which the data suggest. To truly determine memory status, demonstrating rapid cytokine production that could only be achieved by differentiated memory cells would be necessary. The authors agreed and therefore assessed the production of IFNg, IL-17, and IL-2 in the P versus RHD groups by ICS. They show beautiful data demonstrating N-term (S-I)-specific IFNg production in a representative individual from the P group (and a smaller population specific for the S-II region)

and lovely data in the representative plots in the RHD individual of both IFNg and IL2 demonstrating a N-terminal response (Supp. fig. 3b and c). However, when they go beyond these representative plots and graph the frequencies of IFNG, IL17a or IL2+ cells of the CD40L+41BB+ individuals (Supp Fig.3b and C), there are not significant differences in responses between either the N-terminus or the C-terminus-specific responses between COVID patients and RHDs for any of those parameters, which doesn't support their conclusion that only COVID patients have N-terminal responses and suggests instead that they are looking at an artifact of in vitro culture. What was the total frequency of IFNg or IL17 or IL-2 producers without first gating on the CD40L+ and how did any of these frequencies of cytokine producers compare to the HD group under the same conditions?
Most importantly, the authors generate data displaying the CD38+HLA-DR+ cells among CD40L+4-1BB+ cells in Supp. Figure 2 to demonstrate frequencies of S-1 vs S-II-specific CD4+ T cells in the various groups. We were very surprised to see that in this data set, unlike in the COVID-19 Patients (P), there were hardly any cells in the 3 RHD individuals that looked activated as determined by 4-1BB by CD40L or HLA-DR by CD38 or CD38 by KI67 in response to the C-term (in RHD 15, RHD 21, or RHD 43) whereas these cells were easily discernible in the COVID-19 patients. Again,unfortunately, greater interrogation of the data leads to less faith in the results. These data demonstrate that significant CD4 T cell responses are mounted in response to this novel pandemic virus that were not previously generated by other corona viruses. Interpreting these data as anything different is not accurate.

Referee #3 (Remarks to the Author):

The authors have addressed the concerns of my original review.

Author Rebuttals to First Revision:

Referee #1:

Specific comments:

1. "While the IFN results are good, assessing cells double labeled for IFN and TNF is even more specific. Were most of the IFN cells also TNF-expressing? If so, some of these data should be included."

We thank the reviewer for this valuable suggestion. We have reanalysed our data accordingly for a subset of donors. We have now included this dataset of five COVID-19 patients and five reactive healthy donors (RHD). We selected samples, in which we detected significant IFN-expression among S-II reactive $CD4^+$ T cells for further analysis of TNF co-expression (Figure S5). While S-II reactive IFN$^+$ $CD4^+$ T cells from RHD mostly co-expressed TNF, SII reactive IFN$^+$ $CD4^+$ T cells from COVID-19 patients displayed heterogenous TNF co-expression patterns. This is likely reflecting the different disease stages during the acute SARS-CoV-2 infection of the individual patients included in our study.

2. "Figure 1C. The IFN-g positivity is 0.01% but in the extended data, the IFN-g positivity was marked as 0%. This should be rectified."

We apologise for this labelling error. We have corrected accordingly in Fig 2c.

3. "The title to Figure SF6 is not clear and should be modified."

We agree with the reviewer and have changed the title of this figure now into:

Supplementary Figure 6 is now Supplementary Figure 7.

Supplementary Figure 7: Frequencies of SARS-CoV-2-S-I/S-II-reactive $CD4^+$ T cells in healthy donors do not correlate with frequencies of S_{HCoV}-I-reactive or CMV-reactive $CD4^+$ T cells.

Referee #2:

"The key novel findings initially were that COVID patients (P) and those with potentially cross-reactive CD4+ T cells, termed "reactive healthy donors" or RHD, both had responses to the C-terminal region of the spike protein, which is the region with the most homology with other coronaviruses. COVID patients, however, only had antigen-specific T cell responses to the Spike N-terminal."

We thank the referee for the summary of our findings; however we would like to point out an important misinterpretation.

Contrary to the statement "*COVID patients, however, **only** had antigen-specific T cell responses to the Spike **N-terminal***", we demonstrate that $CD4^+$ T cells in COVID-19 patients equally target **both the N- and the C-terminal** parts of S, while $CD4^+$ T cells of RHDs are mostly directed towards the C-terminal part (S-II) (Fig. 2d).

Furthermore, we had also added another key finding in the revised version: we demonstrate direct cross-reactivity of SARS CoV-2 S-II-reactive T cells from RHD to spike glycoprotein (S-II) of endemic coronaviruses (229E and OC43) (Fig 2j,k).

"The difference in the percent of CD40L+41BB+ cells as shown in Fig 2d in the RHD group versus HD group (or even compared to the S-I specific cells in RHD patients) is significant, but extremely small (mean ~ 10-2 in RHD vs 10-2.5 in HD individuals- although the representative flow plots for HD show 0%. These data still raise concerns about the significance of these findings."

We thank the reviewer for this comment. It is important to consider that frequencies of circulating peptide-specific memory T cells in humans are generally low. We agree that S-I- reactivity in RHD is negligibly low, similar to non-reactive HD. Importantly though - and this is a key point of our study - S-II reactivity in RHD is higher and in fact comparable to that in COVID-19 patients. Again, despite the

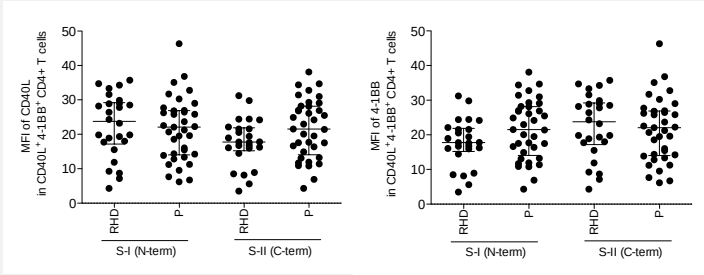
naturally low frequencies, the observed differences are highly specific, demonstrating activation of rare antigen-reactive memory T cells.

In further support of the specificity of the detection method, we demonstrate that SARS-CoV-2 S-II reactivity (even at low frequencies) correlates with reactivity against HCoV S-II (Fig. 2i), but not with CMV-reactivity (Fig. S7c).

"In the current data included in the rebuttal, which should also be included as a supplement in the manuscript, they have now quantified the MFI of these markers in the P and RHD groups in response to both regions of the spike, and there are no differences in either region. Surprisingly however, the number of dots presented in these data now suggest that there are equivalent numbers of CD40L 4-1BB+ cells across both groups and both regions of the protein, which is very confusing. This needs to be addressed."

We apologize if our data representation was misleading. In this figure from our first point-by-point reply, each dot represents the **mean fluorescence intensity (MFI) of the entire CD40L$^+$4-1BB$^+$ population** – **regardless of the number of cells present within this population**.

Based on the reviewer´s observation, we analyzed the MFI levels of both markers. The results (figure below) demonstrate that COVID-19 individuals cannot be distinguished from SARS-CoV-2-reactive HD (=RHD) by the MFI of CD40L or 4-1BB.



Thus, the dots represent MFIs of S-I (N-terminal) and S-II (C-terminal) reactive cells in patients (P) and RHD. However, S-I reactive T cells in RHD were detected at extremely low frequencies (see Fig. 2d). This information is not contained in this graph, since it shows a mean expression value of the *population*. As shown, there is **no difference in the MFIs** – in contrast to the **clear differences in percentage of reactive cells** (Fig. 2d,e). Since MFIs of CD40L or 4-1BB are not informative of the number of reactive cells, we have not included this information here again, but are happy to do so, if this is requested.

Thus, our data clearly demonstrate a **significant difference in the number** of S-I *vs.* S-II reactive T cells in RHD, but not in HD or P.

We hope that the misunderstanding could be clarified and we apologize for any confusion it may have caused.

"However, when they go beyond these representative plots and graph the frequencies of IFNG, IL17a or IL2+ cells of the CD40L+41BB+ individuals (Supp Fig.3b and C), there are not significant differences in responses between either the N-terminus or the C-terminus-specific responses between COVID patients and RHDs for any of those parameters, which doesn't support their conclusion that only COVID patients have N-terminal responses and suggests instead that they are looking at an artifact of in vitro culture."

Unfortunately, the data presentation may have caused another misunderstanding. We apologize for the confusion, and we would like to clarify the data.

Suppl. Fig. 4b and 4c show representative FACS plots of IFN / IL17 / IL-2 expression versus CD40L expression in $CD4^+$ T cells in one COVID-19 patient and one RHD.

Suppl. Fig. 4d shows a quantification of five newly enrolled COVID-19 patients and five RHD, depicting the percentage of cytokine-expressing cells (IFN$^+$ / IL-17^+ / IL-2^+) ***among*** $CD40L^+4\text{-}1BB^+$ T cells.

In order to compare cytokine expression among antigen-reactive T cells between RHD and COVID-19 patients, we **specifically selected** RHD, with known detectable S-I- reactive and S-II-reactive $CD4^+$ T cells. Again, the figure displays the **percentage** of cytokine expressing cells **among** the entire population of $CD40L^+$ T cells – it does **not depict the frequency of reactive T cells** (which in fact was different, see Fig. 2d,e). These data **do not contradict our key finding message** in any way: COVID-19 patients target both N- and C-terminal parts of SARS-CoV-2 S (S-I und S-II), while RHD mostly target only the C-terminal part (S-II) (Fig. 2d,e).

In fact, the finding that there is no significant difference in the cytokine profiles within antigen-reactive cells further demonstrates that $CD40L^+$ T cells in patients and as well in reactive healthy donors are indeed virus-induced memory T cells.

"We were very surprised to see that in this data set, unlike in the COVID-19 Patients (P), there were hardly any cells in the 3 RHD individuals that looked activated as determined by 4-1BB by CD40L or HLA-DR by CD38 or CD38 by KI67 in response to the C-term (in RHD 15, RHD 21, or RHD 43) whereas these cells were easily discernible in the COVID-19 patients."

Again, we would like to clarify the data.

Antigen-reactive CD40L & 4-1BB expression serves to identify **antigen-reactive** $CD4^+$ T cells (Frentsch et al. 2005; Schönbrunn, Frentsch et al., 2012). Fig. 2b, c and Supp. Fig. 2 show the presence of $CD40L^+$ $4\text{-}1BB^+$ (= SARS-CoV-2 reactive) $CD4^+$ T cells in both P and RHD.

HLADR, CD38 and Ki-67 expression indicates **recent *in vivo* activation** (Schulz et al. 2015 doi: 10.4049/[REDACTED].1500598). In RHD, $CD40L^+$ $4\text{-}1BB^+$ T cells do not co-express these markers – demonstrating that they have not been recently activated *in vivo*. This is perfectly in line with the finding, that these cells in RHD are in fact cross-reactive remnants of previous encounters with

endemic HCoVs (Fig. 2i-k), as opposed to CD40L$^+$ 4-1BB$^+$ T cells in COVID-19 patients, which were primed during acute SARS-CoV-2 infection.

"Again,unfortunately, greater interrogation of the data leads to less faith in the results. These data demonstrate that significant CD4 T cell responses are mounted in response to this novel pandemic virus that were not previously generated by other corona viruses. Interpreting these data as anything different is not accurate."

We respectfully disagree with this summary of our findings. We hope that the clarifications provided in reply to the comments will alleviate the concerns and lead the referee to a similar conclusion.

The origin of S-reactive CD4$^+$ T cells in COVID-19 patients remains unknown, but the obvious and most reasonable origin is their *de novo* priming during acute SARS-CoV-2 infection. We do not state otherwise. Regarding the origin of S-II-reactive CD4$^+$ T cells in RHD, which show a clear memory T cell phenotype (Suppl Fig. 3d), we provide data demonstrating their cross-reactivity to S-II pools from endemic coronaviruses and SARS-CoV-2, but not S-I pools or unrelated viruses (CMV) (Fig. 2i,k and Supplementary Fig. 7). Although cross-reactivity with other antigens cannot be completely ruled out, we believe that these data warrant the conclusion that these preformed memory T cells, detected in a subset of healthy donors, were generated in response to previous encounters with endemic coronaviruses, since SARS-CoV-2 is a newly emerged virus.

It is important to stress that the healthy donors in our cohort were repeatedly tested negative for both SARS-CoV-2 PCR and serology by the Department of Virology at Charité (lead by Christian Drosten), which is the German reference laboratory for coronaviruses.

In summary, our study is **the first to demonstrate direct cross-reactivity of SARS-CoV-2 spike reactive CD4$^+$ T cells with spike of 229E and OC43 coronaviruses** in a large cohort of unexposed, healthy individuals. These findings have potentially broad epidemiological implications and we believe therefore that they are of significant interest to the scientific community.

Reviewer Reports on the Second Revision:

Referee #2 (Remarks to the Author):

Response to Revision:

We are well aware of assays that measure expression of CD40L in response to peptide stimulation, especially those that stimulate cells for four to six hours. The concern here is that the lengthy 16 hour stimulation could also induce CD154 expression in naive antigen specific cells. As the authors state. "the strongest evidence…against artifacts of in vitro stimulation stem from the demonstration of direct cross reactivity of SARS-CoV2 reactive T cells with spike glycoproteins from OC43 and 229E." They further state that "if our technique was "inaccurate" it would not have allowed us to

derive three specifically SII-cross reactive T cell lines form three healthy donors." But of course it would if you have identified a previously naive pool through this stimulation. We are specifically stating that a long enough restimulation of naïve cells from a HD without any cross-reactive memory cells could identify spike cross reactive clones that could then be specific for the other SII proteins. We have never suggested that they were not antigen specific, only that they are NOT necessarily memory cells, which the authors highlight as a key finding of this paper. There are no stats performed on Figure 2K, nor is there any indication that 2 out of the 3 lines are responsive to the 229E S-2 or the OC43 S-II.

The more convincing data supporting the fact that RHD have S-II specific memory cells is the IFNg by TNFa data that you have now included in SF5. Make this into a real graph for all of the samples, and not just a few representative samples as explained below.

Point #1: We thank the referee for the summary of our findings; however we would like to point out an important misinterpretation.
Contrary to the statement "COVID patients, however, only had antigen-specific T cell responses to the Spike N-terminal", we demonstrate that CD4+ T cells in COVID-19 patients equally target both the N- and the C-terminal parts of S, while CD4+ T cells of RHDs are mostly directed towards the C-terminal part (S-II) (Fig. 2d).

-That was a grammatical mistake on our part and we apologize. We of course have understood throughout this review that the whole point of the paper was that HD individuals maintain memory T cells specific for the Spike C-terminus (S-II region) from prior cross-reactive exposures, but demonstrate no responsiveness to the N-term, while COVID patients have both. The sentence should have stated, "COVID19 patients only, however, had antigen-specific T cell responses to the Spike N-terminal."

Point #2:We thank the reviewer for this comment. It is important to consider that frequencies of circulating peptide-specific memory T cells in humans are generally low. We agree that S-I- reactivity in RHD is negligibly low, similar to non-reactive HD. Importantly though - and this is a key point of our study - S-II reactivity in RHD is higher and in fact comparable to that in COVID-19 patients.

-We also understand the scarcity of antigen-specific memory cells, and that S-II reactivity in RHD is higher than SI reactivity in the same patients, however there should also be a significant difference in the percent of S-II specific CD40L+41-BB+ cells in Fig. 2D in the RHD versus the HD if there is truly an expanded population of S-II specific cells in the RHD, but this is not marked as significant. Perhaps this is a mistake in the display of the data, but this needs to be explained if it is not.

Point #3: We apologize if our data representation was misleading. In this figure from our first point-by- point reply, each dot represents the mean fluorescence intensity (MFI) of the entire CD40L+4-1BB+ population – regardless of the number of cells present within this population.

-Thank you. This is much clearer.

Point #4: Again, the figure displays the percentage of cytokine expressing cells among the entire

population of CD40L+ T cells – it does not depict the frequency of reactive T cells (which in fact was different, see Fig. 2d,e).

-Thank you for the clarification.

Point #5: Antigen-reactive CD40L & 4-1BB expression serves to identify antigen-reactive CD4+ T cells (Frentsch et al. 2005; Schönbrunn, Frentsch et al., 2012). Fig. 2b, c and Supp. Fig. 2 show the presence of CD40L+ 4-1BB+ (= SARS-CoV-2 reactive) CD4+ T cells in both P and RHD.

-See above concerns regarding 2D

Point #6: HLADR, CD38 and Ki-67 expression indicates recent in vivo activation (Schulz et al. 2015 doi: 10.4049/[REDACTED].1500598). In RHD, CD40L+ 4-1BB+ T cells do not co-express these markers – demonstrating that they have not been recently activated in vivo. This is perfectly in line with the finding, that these cells in RHD are in fact cross-reactive remnants of previous encounters with endemic HCoVs (Fig. 2i-k), as opposed to CD40L+ 4-1BB+ T cells in COVID- 19 patients, which were primed during acute SARS-CoV-2 infection.

-We may somehow be misunderstanding this as it is not well explained in the text, however this is our interpretation of the data:
P7, P8 and P24 samples that are unstimulated demonstrate very little activation by any marker examined and only upon restimulation up-regulate HLA-DR, CD38 and Ki67 to varying degrees. RHD15, RHD21 and RH43 samples that are unstimulated also demonstrate very little activation by any markers examined, However they also do not increase the expression of these markers after 16 hours of restimulation.

In the Schulz JI paper cited, Thiel and colleagues track HLA-DR+ CD38+ cells on freshly drawn PBMCs and show very nicely that expression tracks with acutely stimulated cells directly ex vivo after vaccination. There is no data presented in that paper demonstrating how the expression of HLA-DR, CD38 and KI67 changes on memory cells before and after 16 hours of stimulation. Nor is there any explanation of why PBMC from both P and RHD samples do not express HLA-DR and CD38 in the unstimulated conditions (which should be closest to direct ex vivo), yet only the “memory population” stimulated with the S-II peptides in the P group expresses siginificant amounts of HLA-DR and CD38, while the RHD S-II stimulated samples do not? This needs to be addressed better.

-Lastly, as reviewer #1 also suggested, the cytokine expression data showing TNFa by IFNg should be shown as a graph displaying either the total percent and/or numbers of cytokine DP cells in the HD, RHD and P samples, not just the 5 representative plots shown in SF5 from the RHD and P. This is the real comparison that is needed to make their point that there is a significant S-II reactive population in 35% of the RHD that is different than the HD stimulated under the same conditions.

Referee #4 (Remarks to the Author):

As I understand, this manuscript went through two rounds of review, and was rejected based on the

comments from Referee#2 regarding the reliability of the findings. The authors have now appealed and are rebutting the statements from Referee #2. I was called to provide expert advice to arbitrate. I have reviewed the manuscript revision, and the reviews provided from the other three reviewers. I must also add that I was familiar with the paper as it appeared in BioRX, I believe around the time of submission. In short, this is a fantastic contribution. The data is very novel, and since the short time it became known in the scientific literature it has raised a lot of interest debate and spurred new investigations. The impact of this paper will be profound for years to come.
I believe that the authors have done a good job in answering the reviewer queries in a balanced and thorough manner. I believe that the concerns over the lack of reliability are unfounded, based on three different lines of reasoning. First, at the technical level, the authors have addressed the questions in what I believe is a convincing and credible manner. Second, as mentioned by one of the reviewer, a study from the Sette group reaches essentially the same conclusions, and furthermore a study from Bertoletti's group in Singapore, and one from R. De Vries from the Netherlands also appeared in BioRX, also report preexisting reactivity. Third, and most importantly, in this revision the authors provide direct evidence of cross-reactivity using SARS CoV2 T cell lines and common cold corona peptide pools. I think this evidence is particularly strong.
Based on all of the above, I have no hesitation in recommending the expedite acceptance of this beautiful piece of work.

Author Rebuttals to Second Revision:

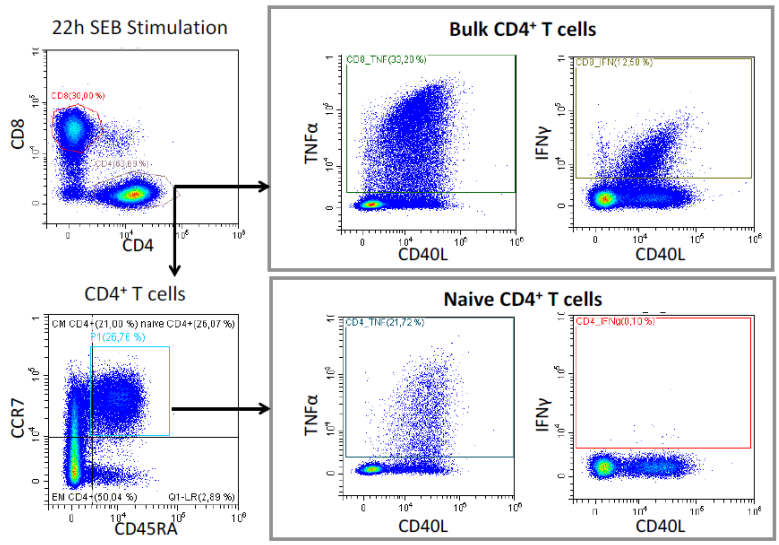
Remaining referees' comments:

Referee #2 (Remarks to the Author):

Response to Revision:
We are well aware of assays that measure expression of CD40L in response to peptide stimulation, especially those that stimulate cells for four to six hours. The concern here is that the lengthy 16-hour stimulation could also induce CD154 expression in naive antigen specific cells. As the authors state. "the strongest evidence…against artifacts of in vitro stimulation stem from the demonstration of direct cross reactivity of SARS-CoV2 reactive T cells with spike glycoproteins from OC43 and 229E." They further state that "if our technique was "inaccurate" it would not have allowed us to derive three specifically SII-cross reactive T cell lines form three healthy donors." But of course, it would if you have identified a previously naive pool through this stimulation. We are specifically stating that a long enough restimulation of naïve cells from a HD without any cross-reactive memory cells could identify spike cross reactive clones that could then be specific for the
other SII proteins. We have never suggested that they were not antigen specific, only that they are NOT necessarily memory cells, which the authors highlight as a key finding of this paper. There are no stats performed on Figure 2K, nor is there any indication that 2 out of the 3 lines are responsive to the 229E S-2 or the OC43 S-II.

The more convincing data supporting the fact that RHD have S-II specific memory cells is the IFNγ by TNFα data that you have now included in SF5. Make this into a real graph for all of the samples, and not just a few representative samples as explained below.

We agree with the referee that after lengthy stimulation, 4-1BB is upregulated also on antigen-specific naïve T cells. However, IFNγ is only produced by memory cells. To further clarify this point, we here share older experimental data with you, clearly demonstrating that even after 22 hours stimulation with SEB, naïve T cells do not express IFNγ, but can express TNFα (see Figure below).



Given the fact that even naive CD4+ T cells can express TNFα, we do not agree with the reviewer. Solely the expression of IFNγ together with CD40L and 4-1BB are sufficient to define memory T cells after short-term antigen-specific stimulation in vitro. In this respect we kindly ask that we don't follow the referee suggestion here to show a graph for all samples because of space restrictions.

Point #1: We thank the referee for the summary of our findings; however, we would like to point out an important misinterpretation.

Contrary to the statement "COVID patients, however, only had antigen-specific T cell responses to the Spike N-terminal", we demonstrate that CD4+ T cells in COVID-19 patients equally target both the N- and the C-terminal parts of S, while CD4+ T cells of RHDs are mostly directed towards the C-terminal part (S-II) (Fig. 2d).

-That was a grammatical mistake on our part, and we apologize. We of course have understood throughout this review that the whole point of the paper was that HD individuals maintain memory T cells specific for the Spike C-terminus (S-II region) from prior cross- reactive exposures, but demonstrate no responsiveness to the N-term, while COVID patients have both. The sentence should have stated, "COVID19 patients only, however, had antigen- specific T cell responses to the Spike N-terminal."

We would like to thank the referee for the clarification.

Point #2:We thank the reviewer for this comment. It is important to consider that frequencies of circulating peptide-specific memory T cells in humans are generally low. We agree that S- I- reactivity in RHD is negligibly low, similar to non-reactive HD. Importantly though - and this is a key point of our study - S-II reactivity in RHD is higher and in fact comparable to that in COVID-19 patients.

-We also understand the scarcity of antigen-specific memory cells, and that S-II reactivity in RHD is higher than SI reactivity in the same patients, however there should also be a significant difference in the percent of S-II specific CD40L+41-BB+ cells in Fig. 2D in the RHD versus the HD if there is truly an expanded population of S-II specific cells in the RHD, but this is not marked as significant. Perhaps this is a mistake in the display of the data, but this needs to be explained if it is not.

Indeed, we had applied a statistical test to compare S-I and S-II-reactivity in HD and RHD in previous versions of fig. 2D. However, after internal discussions, we have taken this out, since this difference in S-II-reactivity is the basis on which HD and RHD groups were formed.

Point #6: HLADR, CD38 and Ki-67 expression indicates recent in vivo activation (Schulz et al. 2015 doi: 10.4049/[REDACTED].1500598). In RHD, CD40L+ 4-1BB+ T cells do not co- express these markers – demonstrating that they have not been recently activated in vivo. This is perfectly in line with the finding, that these cells in RHD are in fact cross-reactive remnants of previous encounters with endemic HCoVs (Fig. 2i-k), as opposed to CD40L+ 4- 1BB+ T cells in COVID- 19 patients, which were primed during acute SARS-CoV-2 infection.

-We may somehow be misunderstanding this as it is not well explained in the text, however this is our interpretation of the data:

P7, P8 and P24 samples that are unstimulated demonstrate very little activation by any marker examined and only upon restimulation up-regulate HLA-DR, CD38 and Ki67 to varying degrees.

RHD15, RHD21 and RH43 samples that are unstimulated also demonstrate very little activation by any markers examined, However they also do not increase the expression of these markers after 16 hours of restimulation.

In the Schulz JI paper cited, Thiel and colleagues track HLA-DR+ CD38+ cells on freshly drawn PBMCs and show very nicely that expression tracks with acutely stimulated cells directly ex vivo after vaccination. There is no data presented in that paper demonstrating how the expression of HLA-DR, CD38 and KI67 changes on memory cells before and after 16 hours of stimulation. Nor is there any explanation of why PBMC from both P and RHD samples do not express HLA-DR and CD38 in the unstimulated conditions (which should be closest to direct ex vivo), yet only the "memory population" stimulated with the S-II peptides in the P group expresses significant amounts of HLA-DR and CD38, while the RHD S-II stimulated samples do not? This needs to be addressed better.

In the Extended Data Figure 2 (previously Supplementary Figure 2), the displayed dot plots with CD38, HLA-DR and Ki-67 are all gated on CD40L+ 4-1BB+ CD4+ T cells **(„the frequencies of S-I and S-II-reactive CD4+ T cells and the ratios of CD38+, HLA-DR+ and Ki-67+ among them").** Therefore, there are almost no cells included in the unstimulated control and these background cells falling into the gate do – by chance – not express CD38, HLA-DR or Ki-67. However, S-reactive cells from P

express these markers, due to their recent in vivo activation. In vivo activation of S-reactive cells in RHD is most probably not recent.

Kinetics of CD38 and HLA-DR expression in the course of T cell activation have been published. CD38 peaks later than 4 days post infection, while HLA-DR may be induced after 15 hours but expression lasts until more than 8 days. Please see table below.

Name / CD number	Function	Expression time	Background in blood	Ref.
CD38	cyclic ADP ribose hydrolase regulates intracellular Ca^{2+} &NAD^{+} levels	Peaked > 4d	0-2%	PMID:23576305 PMID: 2981640 PMID:18468462 PMID:11927944
HLA-DR	MHC class II cell surface receptor; antigen presentation	Peaked ~72h (15h->8d)	$CD4^{+}$ -> 3% $CD8^{+}$ ->8%	PMID 9000587 PMID:11180110

-Lastly, as reviewer #1 also suggested, the cytokine expression data showing TNFa by IFNg should be shown as a graph displaying either the total percent and/or numbers of cytokine DP cells in the HD, RHD and P samples, not just the 5 representative plots shown in SF5 from the RHD and P. This is the real comparison that is needed to make their point that there is a significant S-II reactive population in 35% of the RHD that is different than the HD stimulated under the same conditions.

We have discussed the disputable value of TNFα in defining memory versus naïve CD4+ T cells after in vitro stimulation before. Again, we cannot see a benefit or added value in providing such a summary figure.

Referee #4 (Remarks to the Author):

As I understand, this manuscript went through two rounds of review, and was rejected based on the comments from Referee#2 regarding the reliability of the findings. The authors have now appealed and are rebutting the statements from Referee #2. I was called to provide expert advice to arbitrate.

I have reviewed the manuscript revision, and the reviews provided from the other three reviewers. I must also add that I was familiar with the paper as it appeared in BioRX, I believe around the time of submission. In short, this is a fantastic contribution. The data is very novel, and since the short time it

became known in the scientific literature it has raised a lot of interest debate and spurred new investigations. The impact of this paper will be profound for years to come.

I believe that the authors have done a good job in answering the reviewer queries in a balanced and thorough manner. I believe that the concerns over the lack of reliability are unfounded, based on three different lines of reasoning. First, at the technical level, the authors have addressed the questions in what I believe is a convincing and credible manner. Second, as mentioned by one of the reviewers, a study from the Sette group reaches essentially the same conclusions, and furthermore a study from Bertoletti's group in Singapore, and one from R. De Vries from the Netherlands also appeared in BioRX, also report preexisting reactivity. Third, and most importantly, in this revision the authors provide direct evidence of cross-reactivity using SARS CoV2 T cell lines and common cold corona peptide pools. I think this evidence is particularly strong.

Based on all of the above, I have no hesitation in recommending the expedite acceptance of this beautiful piece of work.

We thank the referee #4 for his summary of the revision process and his especially impressive estimation of the manuscript's impact.