Peptide microarray based analysis of antibody responses to SARS-CoV-2 identifies unique epitopes with potential for diagnostic test development

Holenya, Pavlo, Lange, Paul Joris, Reimer, Ulf, Woltersdorf, Wolfram, Panterodt, Thomas, Glas, Michael, Wasner, Mark, Eckey, Maren, Drosch, Michael, Hollidt, Jörg-Michael, Naumann, Michael, Kern, Florian, Wenschuh, Holger, Lange, Robert, Schnatbaum, Karsten et al. (2021) Peptide microarray based analysis of antibody responses to SARS-CoV-2 identifies unique epitopes with potential for diagnostic test development. European Journal of Immunology. ISSN 0014-2980

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/98151/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
Peptide microarray based analysis of antibody responses to SARS-CoV-2 identifies unique epitopes with potential for diagnostic test development

Pavlo Holenya1*, Paul Joris Lange2,3,5,7*, Ulf Reimer1, Wolfram Woltersdorf3, Thomas Panter-odt2, Michael Glas3, Mark Wasner3, Maren Eckey1, Michael Drosch1, Jörg-Michael Hollied2,4, Michael Naumann5, Florian Kern1,6, Holger Wenschuh1, Robert Lange2,3, Karsten Schnatbaum1,9, Frank F. Bier2,7,9

1 JPT Peptide Technologies GmbH, Berlin, Germany
2 Institute for Molecular Diagnostics and Bioanalysis - IMDB gGmbH, Hennigsdorf, Germany
3 KH Labor GmbH, Bernburg, Germany
4 in.vent GmbH, Hennigsdorf, Germany
5 Institute of Experimental Internal Medicine, Medical Faculty, Otto von Guericke University, Magdeburg, Germany
6 Brighton and Sussex Medical School, Department of Clinical and Experimental Medicine, Brighton, UK
7 University of Potsdam, Institute for Biochemistry and Biology, Potsdam, Germany

*corresponding authors

Joint first authors*, „The authors wish it to be known that, in their opinion, the first 2 authors should be regarded as joint First Authors”

“The peer review history for this article is available at https://publons.com/publon/10.1002/eji.eji202049101”

Summary

Humoral immunity to the Severe Adult Respiratory Syndrome (SARS) Coronavirus (CoV)-2 is not fully understood yet but is a crucial factor of immune protection. The possibility of antibody cross-reactivity between SARS-CoV-2 and other human coronaviruses (HCoVs) would have important implications for immune protection but also for the development of specific diagnostic ELISA tests. Using peptide microarrays, n=24 patient samples and n=12 control samples were screened for antibodies against the entire SARS-CoV-2 proteome as well as...
the Spike (S), Nucleocapsid (N), VME1 (V), R1ab, and Protein 3a (AP3A) of the HCoV strains SARS, MERS, OC43 and 229E. While widespread cross-reactivity was revealed across several immunodominant regions of S and N, IgG binding to several SARS-CoV-2-derived peptides provided statistically significant discrimination between COVID-19 patients and controls. Selected target peptides may serve as capture antigens for future, highly COVID-19-specific diagnostic antibody tests.

Keywords: COVID-19, SARS-CoV-2, antibody, peptide microarray, diagnostics.

Introduction

The COVID-19 pandemic hit the world unprepared. At the time of writing global cases were in the millions and still rising [WHO, 2020]. Apart from social distancing, rapid testing of suspected cases and contacts is the key measure helping to contain the infection. The experience from many countries has shown that imprecise or incomplete knowledge of infection prevalence and viral transmission rates may result in a wrong assessment of the epidemiological situation, wrong predictions, and inadequate guidance and action with potentially grave consequences for societies and national economies. These observations underline the global need for reliable and highly specific tests.

SARS-CoV-2 infection is generally confirmed by RT-PCR on swabs taken from the nose and throat, however, virus load quickly becomes unmeasurable following the acute infection. Reliable serological monitoring will therefore be instrumental in determining the SARS-CoV-2 infection status of larger populations, determine existing immunity, and track local outbreaks [Pollán et al., 2020]. In the near future it may also greatly contribute to assessing vaccine-induced immunity at the population level. However, SARS-CoV-2 antibody testing is much more complex than appears at first glance. Currently available antibody tests are not very specific and may produce conflicting results [Krammer and Simón, 2020; Lisboa Bastos et al., 2020; Bond et al., 2020]. A plausible reason for that might be that antibody tests have generally been validated by samples from clinically symptomatic rather than mild or even asymptomatic cases. However, the reason for the overall unreliable performance of antibody tests is not fully understood. Interestingly, it has been suggested that, particularly in mild and asymptomatic cases, IgG responses may not be detectable until 3-4 weeks after infection [Fafi-Kremer et al., 2020] which may compound the issue.

A large number of studies have examined antibody responses to SARS-CoV-2 [Siracusano et al., 2020]. Most of these used full-length proteins as capture antigens and thus have not identified reactivity at the single epitope level. Only few references [Amrun et al., 2020; Far-
rera-Soler et al. 2020; Ladner et al., 2020; Poh et al., 2020; Wang et al., 2020; Zhang1 et al., 2020] including preprints [Zhang2 et al., 2020; Smith et al., 2020; Dahlke et al., 2020] provide more detailed information on antibody specificity. Most of these studies comprise small sample numbers (n ≤ 10) and only two of them [Ladner et al., 2020; Shrock et al., 2020] examined cross-reactivity with other HCoVs despite initial studies with full length antigens [Ng et al., 2020] and in silico analysis indicating that significant cross-reactivity is to be expected [Ahmed et al., 2020; Grifoni et al., 2020].

Here we present a first study of the humoral immune response against SARS-CoV-2 at the single linear peptide epitope level covering all SARS-CoV-2 antigens as well as the S, N, E and M proteins of the pandemic/epidemic-related human Severe Adult Respiratory Syndrome and Middle East Respiratory Syndrome-related human coronaviruses (SARS-CoV and MERS-CoV), and, in addition, the seasonal cold-related strains human coronavirus HCoV-OC43 and HCoV-229E. Our results provide deeper insight into the immune response to HCoVs than previously published work which could pave the way to SARS-CoV-2 specific and reliable COVID-19 diagnostic test development.

Results

Discrimination between SARS-CoV-2 infected and control group by detection of immunodominant epitopes

Serum samples from COVID-19 patients analyzed by peptide microarray technology exhibited high levels of antibody reactivity against peptides derived from SARS-CoV2. In contrast, almost no reactivity with only few scattered signals was detected in the control group. Several immunodominant regions were identified in color-coded heatmaps reflecting obtained signal intensities. These were located primarily in the SARS-CoV-2 S and N proteins (Figure 1, Supplementary Figures 1-6). For four exemplarily chosen immunodominant regions (denoted A, B, C and D, Figure 1) the respective signal distributions were exemplified by box plots. These showed a clear discrimination between the SARS-CoV-2 infected and the control group, thus supporting the visually observed immunodominance. Besides S and N, there was also strong IgG reactivity to several peptides derived from the SARS-CoV-2 M, AP3A and R1AB proteins, however, immune dominance was less pronounced in these proteins (Supplementary Figures 3, 5 and 6). All remaining SARS-CoV-2 antigens demonstrated no or only minor immunogenicity in terms of IgG response.

For quantitative characterization of the identified immunodominant regions, responses in SARS-CoV-2 infected individuals were compared to those in the control group for each single peptide. Table 1 shows a selection of identified epitopes represented either by single
peptides or by common core sequences derived from two or more overlapping peptides. To narrow down the identified immunodominant peptides to the potential immunodominant epitope sequences, the peptides with the highest accuracy scores (calculated as the sum of true positive and true negative divided by the total number of observations) were listed in the order of their starting position in the corresponding antigens. When applicable, sequence overlaps derived from two or more recognized overlapping peptides were marked in red. A complete list of peptides showing the significant discrimination between the SARS-CoV-2-infected and the control group, as defined by an accuracy score of >0.4, can be found in Supplementary Table 2.

**Figure 2** shows results for three selected peptides that statistically significantly discriminate between SARS-CoV-2 infected individuals and the control group by the non-parametric Wilcoxon Rank Sum test. Also the best combinations of three, four and five peptides are presented. Clearly, the combination of peptides further increases the potential to discriminate between the groups.

**Microarray assay results strongly correlate with a commonly used diagnostic test**

To evaluate the diagnostic potential of a peptide-based assay, a comparison with a commonly used diagnostic test utilizing a full-length viral antigen was performed. **Figure 3** shows the quantitative results obtained for the SARS-CoV-2 infected group with a) the commercial ELISA test (EUROIMMUN) and b) the peptide microarray. The assays demonstrated a strong correlation with a Spearman's rank correlation coefficient of 0.88.

**Cross-reactivity with seasonal common cold coronaviruses experimentally confirmed**

Antibody cross-reactivity between SARS-CoV-2 and seasonal common cold coronaviruses may be highly important with respect to the clinical course of COVID-19 and at the same time represents a challenge for the development of a specific test. We, therefore, analyzed IgG reactivity to peptides derived from the S, N, E, and M proteins of SARS-CoV, MERS-CoV, HCoV-OC43 and HCoV-229E (Supplementary Table 1) in both groups. This revealed highly similar IgG binding profiles against the N and M proteins of SARS-CoV-2 and SARS-CoV, whereas the corresponding proteins of the remaining viral strains showed only weak and scattered signals (Supplementary Figures 1 and 3). Most notably, peptides derived from the S protein of all viral strains showed clear signal patterns among SARS-CoV-2 positive patient samples, suggesting that this antigen contained cross-reactive determinants (Supplementary Figure 2). IgG reactivity was directed at two sequence regions in particular (Figure 4), both located in the S2 domain of the spike protein. Each of these regions was represented by overlapping peptides containing highly conserved cross-reactive epitopes (Figure 4, right panel). One consensus motif, R_{9815}S-IED-LF_{0823} (numbers referring to the location of the epitope in the SARS-CoV-2 antigen) was present in all five viruses, and a
second consensus motif, $F_{1148}^{\rightarrow}E LD^{\rightarrow} F^{\rightarrow} K N_{1158}^{\rightarrow}$ was found in the viruses SARS-CoV-2, SARS-CoV, MERS-CoV and HCoV-OC43, but not HCoV-229E.

To further analyze serologic cross-reactivity of SARS-CoV-2 positive patients against common cold coronaviruses we performed a gapless alignment of the N and S sequences from all coronaviruses (Figure 5). For each strain, this yielded the sequence regions that were complementary to the previously identified SARS-CoV-2 epitopes. Finally, for each of the amino acid residues of the thus identified sequences, the median signal of all overlapping peptides containing that residue was calculated and visualized by the same color coding as in the previously presented heatmaps (Figure 5). The immunodominant epitopes #6 and #7 of the N protein and #1 and #3 of the S protein of SARS-CoV-2 (see Table 1 for a list of the sequences) displayed strong signals with samples from SARS-CoV-2 infected individuals compared to the control samples. At the same time, no signals were found for the homologous sequences of the other coronavirus strains, indicating that IgG responses to these epitopes may be truly specific for SARS-CoV-2 (Figure 5, Top). SARS-CoV-2 epitopes #1-#5 of the N and #2 and #6 of the S protein showed a strong homology to SARS-CoV which resulted in serologic cross-reactivity to this virus (Figure 5, Middle). Finally, epitopes #4 and #5 of the S protein were the least SARS-CoV-2 specific and exhibited considerable cross-reactivity between all sequence variants (Figure 5, Bottom).

Discussion

The work presented here used a systematic approach that provides deeper insight into the humoral immune response to HCoVs and lays the foundation for the development of specific serological tests for SARS-CoV-2 infection. The approach is based on four pillars: (i) peptide microarray based analysis of the humoral immune response against the entire SARS-CoV-2 proteome and other HCoV proteomes, (ii) identification of all immunodominant epitopes at high resolution, (iii) selection of SARS-CoV-2 peptides that are recognized by serum antibodies of COVID-19 patients but not by serum antibodies from controls and (iv) selection of SARS-CoV-2 peptides whose homologous peptides in other HCoVs are not recognized by serum antibodies from COVID-19 patients.

Our analysis of IgG responses of COVID-19 patients against all peptides of the SARS-CoV-2 proteome revealed that IgG binding was mainly focused on both the S and N proteins, indicating that these two proteins are indeed immunodominant (Figure 1, Supplementary Figures 1-6). The fact that the M, AP3A and R1ab proteins were also recognized, however, shows that responses are not exclusive to S and N. It remains to be resolved in future studies if the most dominant immune responses against certain antigens are also the most effec-
tive responses concerning disease protection. With respect to the location of linear epitopes within the S protein sequence, we made several interesting observations. First, linear epitopes were found in the cytoplasmic tail portion of the S protein, although such would not normally be accessible to antibody binding, as they are hidden away from the viral surface. However, following the destruction of SARS-CoV-2-infected cells or the destruction of SARS-CoV-2 virions otherwise, these portions of the sequence are likely to get exposed and may be recognized by B-cells. Secondly, our analysis did not detect much binding of antibodies to potential linear epitopes in the RBD region of the S protein (Figure 1), which stands in contrast to the detection of antibody binding to the correctly folded domains reported by others [Premkumar et al., 2020]. Both observations appear to confirm that linear epitopes may occur largely independently of conformational epitopes [Baraniak et al., 2019].

It should be noted that some signals of reactivity were also observed in a few of the control group samples, i.e. in the N but also in the S protein (Figure 1). While these reactivities have been included in all statistical analyses throughout the manuscript, due to the relatively small control group size we refrained from further elaborating on the proportion of possible cross reactive samples in non-infected individuals.

Applying stringent bioinformatic criteria we identified several linear epitopes that were immunodominant among SARS-CoV-2 positive patients (Table 1). The majority of the epitopes were located within the N protein as well as the S1 and S2 domain of the S protein (Table 1). Previous analyses of antibody epitopes in SARS-CoV revealed that many such epitopes were indeed linear [Guo et al., 2004; Shichijo et al., 2004]. Interestingly, some immunodominant regions in the N and S proteins that have been previously reported in the literature [Poh et al., 2020; Amrun et al., 2020] correspond closely to epitopes #3 from N and #1 and #4 from S (Table 1), respectively. This confirms our findings and illustrates the potential of the peptide microarrays for epitope identification. Longer sequence stretches encompassing some of the epitopes identified here have also been described in published manuscripts [Wang et al., 2020; Farrera-Soler et al., 2020; Zhang1 et al., 2020] or in non-peer-reviewed preprint publications, however, without narrowing them down [Zhang2 et al., 2020; Smith et al., 2020; Dahlke et al., 2020]. With our comprehensive peptide library we further completed the list of the previously described immunodominant SARS-CoV-2 regions and specified the corresponding epitopes.

The most important question with respect to developing a reliable diagnostic test was, however, whether the IgG response to the identified SARS-CoV-2-specific peptides can effectively distinguish COVID-19 patients and controls. This was indeed the case, as a compari-
son of serum reactivity in COVID-19 patients versus the control group, peptide by peptide, identified many peptides to which responses were significantly different between the groups (Table 1, Figure 2A). The combination of selected peptides derived from different SARS-CoV-2 antigens further improved the discrimination between the sample groups as shown in Figure 2B, and also might increase sensitivity when applied in a diagnostic test.

As a further step to assess the diagnostic potential of the obtained data, the quantitative results obtained with the peptide microarrays were compared with those generated by a commercial ELISA test (Figure 3). We observed a very strong correlation between the ELISA signals and the summed microarray signals for the SARS-CoV-2 S and N peptides. This suggests that a combination of short linear synthetic peptides may be used for the development of a diagnostic test instead of full-length antigens. The strong correlation was not expected per se, as the ELISA test is based on the S1 domain of the S protein and not the complete antigen. As it was found that only a small proportion of all peptides forms B cell epitopes, the selection of peptides for a diagnostic test must not necessarily span the whole antigen, but might be limited to immunodominant regions identified by a peptide microarray screening.

Multiple studies utilizing in silico prediction of SARS-CoV-2 immunogenicity [Wang et al., 2020; Yuan et al., 2020; Palm et al., 2020; Smith et al., 2020] postulated the existence of cross-reactive immunity between different coronaviruses, in particular between SARS-CoV-2 and SARS-CoV, because several known B-cell and T-cell target peptides in SARS-CoV exist almost identically in SARS-CoV-2. Cross-reactive T cell epitopes were experimentally confirmed by several authors [Braun et al., 2020; Mateus et al., 2020] and some degree of B cell cross reactivity to HCoVs has been identified by western blot or flow cytometry analysis using full length antigens [Bonifacius et al., 2020; Ng et al., 2020]. However, there is still a shortage of experimental data identifying cross-reactive B cell epitopes. We found two immunodominant regions in the S2 domain of the SARS-CoV-2 S protein (1148-FKEELDKYFKN-1158 and 0815-RSFIEDLLF-0823) which were highly cross-reactive with the prevalent seasonal "common cold" coronaviruses HCoV-OC43 and HCoV-229E (Figure 4). A similar result obtained with different assays based on 30-mer peptide cDNA-fusion pools or 20-mer phage-display technology confirms our findings [Ladner et al., 2020; Shrock et al., 2020]. Although a few of the control samples shown in Figure 1 display some reactivity to SARS-CoV-2 derived peptides, given the small number of control sera tested it is not possible to draw any conclusions regarding the frequency of cross-reactive samples in SARS-CoV-2 negative individuals.
The identified cross-reactivity is of particular interest, as it suggests that exposure to common cold HCoVs may provide a degree of immunity to SARS-CoV-2. The presence of such preexisting immunity may have a positive impact on the course and outcome of COVID-19. Arguments against this hypothesis are, however, that it has been reported that the common cold coronaviruses generate only a weak and transient humoral immune response [Huang et al., 2020] and that convalescents from SARS-CoV – a virus that is structurally more closely related to SARS-CoV-2 than the common cold coronaviruses – did not have SARS-CoV-2 neutralization titers in plasma [Lv et al., 2020]. Of note, preexisting immunity could also be detrimental, as it may inhibit production of antibodies with certain specificity, i.e. a phenomenon called ‘original antigenic sin’. This phenomenon may be responsible for ineffective vaccine response to a recently developed Cytomegalovirus vaccine [Baraniak et al., 2019]. Even worse, existing antibodies could promote infection of target cells by facilitating viral uptake, a phenomenon called antibody-dependent enhancement [Fierz et al., 2020; Arvin et al., 2020].

Our results demonstrate that some of the patient sera, e.g. numbers SARS_20 to SARS_23, show stronger reactivity towards the HCoV-OC43 variant of the cross-reactive epitope II than to the respective SARS-CoV-2 epitope (Figure 4). This is an interesting observation as there is no other logical explanation for this than that the antibody was made in response to the HCoV-OC43 virus in a previous HCoV-OC43 infection. In other words this is suggesting that these antibodies may be bona-fide anti-HCoV-OC43 antibodies re-boosted by exposure to SARS-CoV-2 infection. A number of additional peptides that led to strong IgG responses (Figure 4) support the notion of preexisting B cell immunity to SARS-CoV-2 in those exposed to other coronaviruses. It will be up to future studies with larger cohorts to determine whether preexisting immunity has an effect on disease severity.

While the physiological effects of serologic cross-reactivity require further studies, it is obvious that it makes the development of reliable antibody tests more difficult. Antibody based ELISA tests that use full-length conformational antigens or parts thereof as capture antigens might give false positive results in individuals with previous exposure to common cold coronaviruses. In this regard, the usage of the S protein antigen in a diagnostic test might be problematic due to several highly homologous regions between SARS-CoV-2 and other coronaviruses. The detailed analysis of IgG reactivity against SARS-CoV, MERS-CoV, HCoV-OC43 and HCoV-229E S and N proteins indeed revealed widespread cross-reactivity in several immunodominant regions (Figure 4), which is most probably a consequence of sequence homology. However, the observed cross-reactivity could not be predicted by simple sequence alignment unless 100% homology was present. With a view to develop a diagnostic test, our finding that epitopes #6 and #7 of N and #1 and #3 of S appear to provide a
high specificity for SARS-CoV-2 (Figure 5) is of major interest. A combination of such epitopes in a peptide ELISA may increase sensitivity without increasing the risk of false positives.

Because the prevalence of mutated versions of SARS-CoV-2 became eminent during the revision of this manuscript [CDC, 2021], we analyzed if the discovered specific peptides were located within mutated regions of the respective variants. In fact none of the sequences in Figure 4 were located in regions where the currently most discussed emerging SARS-CoV-2 variants (B.1.1.7 – originating from the UK, B.1.351 – originating from South Africa, P.1 – originating from Brazil) are mutated (Figure 1). This implies that a diagnostic test based on the epitopes discovered in this study should also allow the detection of the three mutated viruses.

In conclusion, our study experimentally confirms antibody cross-reactivity between SARS-CoV-2 and other HCoVs, but at the same time reveals antibody reactivity to several epitopes that are unique to SARS-CoV-2. These epitopes are of particular interest for the development of diagnostic tests. Additional, larger studies are required to confirm these results and accelerate the development of a sensitive and specific test for SARS-CoV-2 infection that we urgently need in order to deal with the current and possible future waves of this pandemic.

MATERIALS AND METHODS

**Blood sampling.** Following written informed consent peripheral blood samples were obtained by venepuncture / from peripheral venous catheters. Multiple samples were collected into a Vacutainer Tube (Serum-separating tube) with a volume of 8 ml. Serum was collected after centrifugation, samples were allowed to clot for 1 hour, and the clot was removed by centrifugation. The samples were divided and stored in aliquots at −80°C until use.

**RT-PCR RNA collection according to Guidelines of the CDC.** Samples were collected from the upper respiratory tract using a nasopharyngeal (NP) with plastic by trained medical staff. In the case of the NP the swab was inserted deeply into the nostril until resistance was encountered and the depth was equal to the distance from the nose to the ears of the patients. The swab was gently rubbed against the nasopharynx and left there in order to absorb enough secret. Swabs were then immediately placed into a sterile transport tube. Samples were cooled and tested no more than 48 hours after collection.

**Participants.**

Controls: Controls consist of excess serum samples collected in accordance to the regulations for the use of rest samples for scientific purposes at the KH Labor. All samples have been collected from form patients hospitalized in April 2020. Sera in the control group tested
negative for COVID-19 infection in the EUROIMMUN IgG Elisa Test. We used samples from different specialties, including cardiology, gastroenterology, nephrology and the intensive care unit.

SARS-CoV-2 infected group: COVID-19 patients (n=13 female, age range 23 to 70 years; n=11 male, age range 35 to 70 years) were recruited across the spectrum of clinical severity, ranging from asymptomatic/mild to severe/SARS. COVID-19 diagnosis was confirmed either by diagnostic PCR and/or commercial ELISA carried out 14 days or later after the onset of symptoms. Detailed information on all patients and samples is provided in Supplementary Table 3.

Inclusion/exclusion criteria were: Fully conversant in the German or English language, able and willing to attend to the study requirements, written informed consent to participate in this study / history of four or more infections with antibiotic treatment in the year preceding the blood-draw, severe liver or kidney disease, patients undergoing immunotherapy, and cancer patients.

A detailed medical history was obtained from all patients. To validate COVID-19 cases, RT-PCR, clinical diagnosis and the result of the Euroimmune IgG Assay were used in conjunction. Patients were questioned about the time of onset of symptoms. All Patients included in this study could precisely specify at which date symptoms occurred. After the onset of symptoms all but one Patient underwent RT-PCR testing and tested positive for COVID-19. We ensured that no patient took any medicine or suffered from a disease that could potentially affect the development of antibodies. The severity of COVID-19 was determined based on WHO guidelines into asymptomatic, mild (common cold-like symptoms), moderate (dyspnea, clinical signs of pulmonary distress) and severe (requiring oxygen supplementation, intensive care with or without ventilation) [WHO Global, 2020].

**Peptide microarrays.** Antibody response profiling was performed with a peptide microarray (JPT Peptide Technologies GmbH) containing 5513 peptides covering the full proteome of SARS-CoV-2 and, in addition, the S, N, E, and M proteins of the following HCoVs: SARS-CoV-1, MERS-CoV, HCoV-229E, HCoV-OC43. All proteins were represented by overlapping peptides covering the entire protein sequence (15 aa length, 11 aa overlap). All proteins covered by the peptide library are listed in Supplementary Table 1.

**Peptide synthesis and microarray production.** Peptides were synthesized and immobilized on peptide microarray slides as described previously [Stephenson et al., 2015]. In brief, the peptides were synthesized using SPOT synthesis [Wenschuh et al., 2000], cleaved from
the solid support and chemoselectively immobilized on functionalized glass slides. Each peptide was deposited on the microarray in triplicates.

**Microarray assay and data analysis.** The peptide microarrays were incubated with sera (applied dilution 1:200) in an HS 4800 microarray processing station (Tecan) for two hours at 30°C, followed by incubation with 0.1 μg/mL fluorescently labelled anti human IgG detection antibody (Jackson Immunoresearch, 109-605-098). Washing steps were performed prior to every incubation step with 0.1% Tween-20 in 1x TBS. After the final incubation step the microarrays were washed (0.05% Tween-20 in 0.1x SSC) and dried in a stream of nitrogen. Each microarray was scanned using a GenePix Autoloader 4300 SL50 (Molecular Devices, Pixel size: 10 μm). Signal intensities were evaluated using GenePix Pro 7.0 analysis software (Molecular Devices). For each peptide, the MMC2 value of the three triplicates was calculated. The MMC2 value was equal to the mean value of all three instances on the microarray except when the coefficient of variation (CV) – standard-deviation divided by the mean value – was larger than 0.5. In this case the mean of the two values closest to each other (MC2) was assigned to MMC2. Further data analysis and generation of the heatmaps was performed using the statistical computing and graphics software R (Version 4.0.2, www.r-project.org).

**RT-PCR test.**

We used the protocol described by the centers for disease control and prevention for the CDC 2019 Novel Coronavirus (2019-nCoV) Real-Time Reverse Transcriptase (RT)-PCR Diagnostic Panel.

**ELISA test.** To detect antibodies, we used the Euroimmun Anti-SARS-CoV-2 ELISA (IgG), which was manufactured by Euroimmun Medizinische Labordiagnostika AG, Lübeck, Germany. The break-off microplate wells were covered with the antigen, a recombinant structural spike 1 (S1) protein of SARS-CoV-2 expressed in HEK293 [FDA, 2020]. First, the diluted patient samples were incubated in the wells which lead to specific IgG antibodies binding to the antigens. In order to detect the bound antibodies, an enzyme-labelled anti-human IgG antibody detected antigen-antibody complexes and catalysed a color reaction. By calculating the extinction of the sample over the extinction of the calibrator the results could be evaluated semi-quantitatively. A ratio below 0.8 was considered a negative result. A ratio between 0.8 and 1.1 was considered a borderline result. A ratio above 1.1 was considered a positive result [Euroimmun Medizinische Labordiagnostika AG, 2020].
QUANTIFICATION AND STATISTICAL ANALYSIS

Data analysis and heatmap generation was performed using the statistical computing and graphics software R (Version 4.0.2, www.r-project.org).

The data set used for statistical analysis of the microarray results and for generation of all heatmap presentations was based on MMC2 values calculated from signals of the triplicate spots of each individual peptide. The MMC2 value calculation is described in detail in the Microarray assay and data analysis section.

The principle of the color coding of all heatmaps is described in the corresponding figure legends. The colored fields in the heatmaps reflect the calculated MMC2 values of each individual peptide.

The boxplots in Figure 1 and Figure 2 show the median of the corresponding sample groups. The central rectangles span the first quartile to the third quartile (the interquartile range or IQR). The whiskers represent the following values: upper whisker = min(max(x), Quartile 3 + 1.5 * IQR); lower whisker = max(min(x), Quartile 1 – 1.5 * IQR). The points in the boxplots show the MMC2 values for the individual samples.

Statistical comparison of two sample groups was done by the Wilcoxon Rank Sum test (R package stats, n=24 patient samples and n=12 control samples). ROC analysis was performed using the R package ROCR. The selection of the peptide hits with significant separation between the sample groups was based on the accuracy value calculated as '(true positive + true negative)/total number of observations'. An accuracy value of ≥0.4 was considered to indicate immunogenicity (Supplementary Table 2).

The Spearman's rank correlation coefficient reflecting the relationship between the commercial ELISA and the peptide microarray (Figure 3) was calculated as described in the figure legend using the R package ggpubr.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics approval statement

The study was conducted in agreement with the declaration of Helsinki. Ethical approval was obtained from the Ärztekammer Sachsen-Anhalt (approval reference number 64/20).

This article is protected by copyright. All rights reserved.
Acknowledgements

The authors are grateful for the assistance of the following local health authorities (Gesundheitsämter) of Landkreis Börde (Frau Dr. Kontzog and Team), Landkreis Salzlandkreis (Frau Leonhardt and Team), and Landkreis Harz (Frau Dr. Christiansen and Team) for their support in identifying appropriate blood donors. The project got financial support from the German Ministry of Education and Research (BMBF) grant 03COV24A and was part of the COVID19-initiative of the Diagnostics Network Berlin-Brandenburg.

Author contributions


Conflict of interest disclosure

P.H., U.R., M.E., M.D., F.K., H.W. and K.S. are employees of JPT Peptide Technologies GmbH which commercialized the peptide microarrays used in this study. P.J.L, W.W., T.P. M.G., M.W., R.L. are employees of KH Labor, and J.-M. H. is CEO of in.vent GmbH, who supported the project via the Diagnostics Network. All other authors have no commercial or financial conflict of interest.

References


This article is protected by copyright. All rights reserved.
Figure 1. Peptide microarray based analysis of IgG responses to SARS-CoV-2 Nucleoprotein (N, NCAP_SARS2) and Spike glycoprotein (S, SPIKE_SARS2) measured in human sera. Antibody response profiling was performed with peptide microarrays containing 5513 peptides covering a scan through the full proteome of SARS-CoV-2. The heatmaps represent two groups of serum samples: SARS-CoV-2 infected (n=24) and control (n=12) group. The two upper rows of the heatmaps reflect signals obtained with detection antibody only (no serum). Columns represent peptide sequences, rows represent samples. Colors indicate the signal values obtained from triplicate spots ranging from white (0 or low intensity) over yellow (middle intensity) to red (maximal intensity of 65535 light units). Box plots A-D show the signal distribution in each group for selected immunodominant regions. The positions of these regions are indicated with the corresponding letters A-D in the heatmap. Peptide sequences as well as start and end positions are indicated in the same order above each block of boxplots. The boxplots show the median of the corresponding sample groups. The central rectangles span the interquartile range (IQR). The whiskers represent the following values: upper whisker = min(max(x), Quartile 3 + 1.5 * IQR); lower whisker = max(min(x), Quartile 1 – 1.5 * IQR). The points in the boxplots show the MMC2 values for the individual samples. Data are pooled from two experiments.
Figure 2. Selected single peptides and combination of peptides providing the best discrimination between sample groups. The same dataset as in Figure 1 was used for the statistical analysis. A. For each peptide the signal intensity was compared between the SARS-CoV-2 infected ("SARS2"; n=24) and the control ("CTRL"; n=12) group by Wilcoxon Rank Sum test with the box plots showing the ability of selected peptides to discriminate between groups. B. Combinations of three, four and five peptides providing the best discrimination between the groups by Wilcoxon Rank Sum test. Each data point represents the sum of the signals of the combined peptides. Thresholds were calculated using the FPR and TNR values from the ROC-Analysis by selecting the optimal value for balanced accuracy ((TPR+TNR)/2). The boxplots show the median of the corresponding sample groups. The central rectangles span the interquartile range (IQR). The whiskers represent the following values: upper whisker = min(max(x), Quartile 3 + 1.5 * IQR); lower whisker = max(min(x), Quartile 1 – 1.5 * IQR). The points in the boxplots show the MMC2 values for the individual samples. Data are pooled from two experiments.
Figure 3. Correlation between the commercial ELISA and the peptide microarray signals. The Spearman's correlation coefficient ($R=0.88; p \text{ value}=2.5\times10^{-6}$) between the ELISA signal and the summed microarray signals for SARS-CoV-2 S and N peptides was calculated for sera from SARS-CoV-2 infected patients (n=22). In the ELISA assay, which used a recombinant S glycoprotein as capture antigen, the quotient of the extinction of the patient sample in comparison to the calibrator was used as net assay signal. The total microarray signals for each sample were calculated as a sum of all corresponding signal intensities above the upper $10^{-16}\%$ quantile of the noise distribution.
Figure 4. Serologic cross-reactivity between SARS-CoV-2 and other Coronaviruses analyzed by peptide microarray. Left: Conserved S protein regions of different coronaviruses showing a strong IgG seroreactivity in sera from SARS-CoV-2 infected patients (n=24). Each region is represented by two overlapping peptides (framed red). Heatmap rows represent peptide sequences, columns represent samples. Colors indicate the signal values obtained from triplicate spots ranging from white (0 or low intensity) over yellow (middle intensity) to red (maximal intensity of 65535 light units). Right: cross-reactive motifs resulting from sequence alignment of the peptides reflecting a high homology. Motif II was present in all five viruses, whereas motif I was found in the viruses SARS-CoV-2, SARS-CoV, MERS-CoV and HCoV-OC43, but not HCoV-229E.
Figure 5. Potential Nucleoprotein (N) and Spike glycoprotein (S) epitope combinations for a diagnostic test. Each epitope (depicted with the same numbering as in Table 1) is represented by a heatmap showing aligned sequences of the investigated coronaviruses. Blue frames indicate the inferred epitope length. The colored background of each amino acid residue reflects the median signal intensity (IgG reactivity) of the SARS-CoV-2 positive cohort calculated from signals of all peptides on the microarray which contained this residue. The color code ranges from white (0 or low intensity) over yellow (middle intensity) to red (maximal intensity of 65535 light units). Thus, red signals indicate strong binding of antibodies to the respective amino acid, while yellow signals indicate weak binding. Per cent sequence homology to SARS-CoV-2 is indicated in green. Based on the observed cross-reactivity of the SARS-CoV-2-induced IgG responses toward the other virus types, certain combinations of SARS-CoV-2 epitopes may be used to develop diagnostic tests with different specificities.
Table 1. Selected immunodominant SARS-CoV-2 epitopes identified by peptide microarrays. Selected peptides providing statistically significant separation between the groups are listed. The first column contains the protein name which in combination with "_SARS2" forms the Uniprot ID (e.g. AP3A_SARS2). The numbers indicate the start and end positions of each peptide. Immunodominant epitopes (highlighted red) are derived from the sequences of (overlapping) peptides. Start and end positions are also indicated preceding and following the epitope sequences in the 'immunodominant epitope' column. The right column (#) provides a numeration of the epitopes that is used within the manuscript and also indicates the S protein subunit (S1 or S2) in which the respective peptide is found.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
<th>Ratio</th>
<th>Median SARS2/CTRL</th>
<th>P value</th>
<th>AUC</th>
<th>Immunodominant epitope</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP3A_0253-0267</td>
<td>SGVVNPVMEPIYDEP</td>
<td>1.8</td>
<td>1.2E-02</td>
<td>0.76</td>
<td></td>
<td>0261-EPIYDEP-0267</td>
<td>1</td>
</tr>
<tr>
<td>AP3A_0257-0271</td>
<td>NPVMEPIYDEPTTTTT</td>
<td>1.9</td>
<td>5.6E-03</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP3A_0261-0275</td>
<td>EPIYDEPTTTTSVPL</td>
<td>3.0</td>
<td>5.0E-03</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0009-0023</td>
<td>QRNA PRITFGGPSDS</td>
<td>2.1</td>
<td>6.1E-04</td>
<td>0.84</td>
<td></td>
<td>0013-PRITFGGPSDS-0023</td>
<td>1</td>
</tr>
<tr>
<td>NCAP_0013-0027</td>
<td>PRITFGGPSDSTGSN</td>
<td>2.6</td>
<td>1.6E-03</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0033-0047</td>
<td>SGARSQRRPQGLPN</td>
<td>2.7</td>
<td>7.3E-03</td>
<td>0.78</td>
<td></td>
<td>0037-SKQRPPQGLPN-0047</td>
<td>2</td>
</tr>
<tr>
<td>NCAP_0037-0051</td>
<td>SKQRPPQGLPNNTAS</td>
<td>4.6</td>
<td>4.0E-03</td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0157-0171</td>
<td>IVLQLPQGTTLPKGF</td>
<td>7.1</td>
<td>5.6E-03</td>
<td>0.78</td>
<td></td>
<td>0161-LPQGTTLPKGF-0171</td>
<td>3</td>
</tr>
<tr>
<td>NCAP_0161-0175</td>
<td>LPQGTTLPKGFYAEG</td>
<td>5.0</td>
<td>8.4E-02</td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0217-0231</td>
<td>AALALLLDRNLQLE</td>
<td>2.7</td>
<td>8.1E-04</td>
<td>0.83</td>
<td></td>
<td>0221-LLDLRNLQLE-0231</td>
<td>4</td>
</tr>
<tr>
<td>NCAP_0221-0235</td>
<td>LLLDLRNLQLESKMS</td>
<td>12.7</td>
<td>4.3E-07</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0357-</td>
<td>IDAYKTFPPTEPKKD</td>
<td>3.2</td>
<td>1.7E-01</td>
<td>0.87</td>
<td></td>
<td>0357-IDAYKTFPPTEPKKD</td>
<td>5</td>
</tr>
<tr>
<td>Code</td>
<td>Sequence</td>
<td>Score</td>
<td>E-value</td>
<td>p-value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------</td>
<td>-------</td>
<td>---------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0373-0387</td>
<td>KKKADETQALPQRQK</td>
<td>3.1</td>
<td>2.1E-03</td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0377-0391</td>
<td>DETQALPQRQKQKQT</td>
<td>6.1</td>
<td>5.7E-05</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0385-0399</td>
<td>RQKKQQTTLPPPADD</td>
<td>6.1</td>
<td>3.8E-04</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0389-0403</td>
<td>QQTVTLPPPADD</td>
<td>5.6</td>
<td>1.1E-03</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0393-0407</td>
<td>TLPADDDDFSKQL</td>
<td>39.4</td>
<td>1.7E-05</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCAP_0397-0411</td>
<td>AADDDDFSKQLQSM</td>
<td>30.9</td>
<td>5.3E-06</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_0553-0567</td>
<td>TESNNKFLPFOQFGR</td>
<td>3.4</td>
<td>5.2E-04</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_0557-0571</td>
<td>KKKFLPFOQFRDIAD</td>
<td>5.1</td>
<td>9.9E-05</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_0567-0671</td>
<td>NNSYEDIPAGIC</td>
<td>1.6</td>
<td>2.6E-02</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_0661-0675</td>
<td>ECDIPAGICASYQ</td>
<td>1.8</td>
<td>1.9E-02</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_0785-0799</td>
<td>VQIYKTPPIKDFGG</td>
<td>2.8</td>
<td>1.1E-06</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_0789-0803</td>
<td>YKTPPIKDFGGFNFS</td>
<td>4.8</td>
<td>4.5E-04</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_0809-0823</td>
<td>FSKPSKRFSFEDLLF</td>
<td>5.5</td>
<td>1.6E-03</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_0813-0827</td>
<td>SKRSFIEDLLFNKT</td>
<td>3.0</td>
<td>1.2E-02</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_1145-1159</td>
<td>LDSFKEELDHYFKNH</td>
<td>11.2</td>
<td>1.1E-04</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_1253-1267</td>
<td>CCKFDEDSEPVLKG</td>
<td>2.8</td>
<td>6.3E-03</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_1257-1271</td>
<td>DEDSEPVLKGVLKH</td>
<td>3.3</td>
<td>4.7E-05</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPIKE_1259-1273</td>
<td>DEDSEPVLKGVLKHYT</td>
<td>3.1</td>
<td>1.7E-05</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This article is protected by copyright. All rights reserved.
<table>
<thead>
<tr>
<th></th>
<th>Sequence</th>
<th>E-value</th>
<th>p-value</th>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>VME1_0005-0019</td>
<td>NGTITVEELKKLLEQ</td>
<td>6.3</td>
<td>8.1E-04</td>
<td>0005-NGTITVEELKKLLEQ-0019</td>
<td>1</td>
</tr>
<tr>
<td>R1AB_1657-1671</td>
<td>KYPQVNGLTSIKWAD</td>
<td>1.7</td>
<td>2.6E-02</td>
<td>1657-KYPQVNGLTSIKWAD-1671</td>
<td>1</td>
</tr>
<tr>
<td>R1AB_2153-2167</td>
<td>TTNIVTRCLNRVCTN</td>
<td>1.6</td>
<td>1.9E-02</td>
<td>0005-TTNIVTRCLNRVCTN-0019</td>
<td>2</td>
</tr>
</tbody>
</table>
MATERIALS AND METHODS

**Blood sampling.** Following written informed consent peripheral blood samples were obtained by venepuncture / from peripheral venous catheters. Multiple samples were collected into a Vacutainer Tube (Serum-separating tube) with a volume of 8 ml. Serum was collected after centrifugation, samples were allowed to clot for 1 hour, and the clot was removed by centrifugation. The samples were divided and stored in aliquots at −80°C until use.

**RT-PCR RNA collection according to Guidelines of the CDC.** Samples were collected from the upper respiratory tract using a nasopharyngeal (NP) with plastic by trained medical staff. In the case of the NP the swab was inserted deeply into the nostril until resistance was encountered and the depth was equal to the distance from the nose to the ears of the patients. The swab was gently rubbed against the nasopharynx and left there in order to absorb enough secret. Swabs were then immediately placed into a sterile transport tube. Samples were cooled and tested no more than 48 hours after collection.

**Participants.**

Controls: Controls consist of excess serum samples collected in accordance to the regulations for the use of rest samples for scientific purposes at the KH Labor. All samples have been collected from form patients hospitalized in April 2020. Sera in the control group tested negative for COVID-19 infection in the EUROIMMUN IgG Elisa Test. We used samples from different specialties, including cardiology, gastroenterology, nephrology and the intensive care unit.

SARS-CoV-2 infected group: COVID-19 patients (n=13 female, age range 23 to 70 years; n=11 male, age range 35 to 70 years) were recruited across the spectrum of clinical severity, ranging from asymptomatic/mild to severe/SARS. COVID-19 diagnosis was confirmed either by diagnostic PCR and/or commercial ELISA carried out 14 days or later after the onset of symptoms. Detailed information on all patients and samples is provided in Supplementary Table 3.

Inclusion/exclusion criteria were: Fully conversant in the German or English language, able and willing to attend to the study requirements, written informed consent to participate in this
study / history of four or more infections with antibiotic treatment in the year preceding the blood-draw, severe liver or kidney disease, patients undergoing immunotherapy, and cancer patients.

A detailed medical history was obtained from all patients. To validate COVID-19 cases, RT-PCR, clinical diagnosis and the result of the Euroimmune IgG Assay were used in conjunction. Patients were questioned about the time of onset of symptoms. All Patients included in this study could precisely specify at which date symptoms occurred. After the onset of symptoms all but one Patient underwent RT-PCR testing and tested positive for COVID-19. We ensured that no patient took any medicine or suffered from a disease that could potentially affect the development of antibodies. The severity of COVID-19 was determined based on WHO guidelines into asymptomatic, mild (common cold-like symptoms), moderate (dyspnea, clinical signs of pulmonary distress) and severe (requiring oxygen supplementation, intensive care with or without ventilation) [WHO Global, 2020].

Peptide microarrays. Antibody response profiling was performed with a peptide microarray (JPT Peptide Technologies GmbH) containing 5513 peptides covering the full proteome of SARS-CoV-2 and, in addition, the S, N, E, and M proteins of the following HCoVs: SARS-CoV-1, MERS-CoV, HCoV-229E, HCoV-OC43. All proteins were represented by overlapping peptides covering the entire protein sequence (15 aa length, 11 aa overlap). All proteins covered by the peptide library are listed in Supplementary Table 1.

Peptide synthesis and microarray production. Peptides were synthesized and immobilized on peptide microarray slides as described previously [Stephenson et al., 2015]. In brief, the peptides were synthesized using SPOT synthesis [Wenschuh et al., 2000], cleaved from the solid support and chemoselectively immobilized on functionalized glass slides. Each peptide was deposited on the microarray in triplicates.

Microarray assay and data analysis. The peptide microarrays were incubated with sera (applied dilution 1:200) in an HS 4800 microarray processing station (Tecan) for two hours at 30°C, followed by incubation with 0.1 μg/mL fluorescently labelled anti human IgG detection antibody (Jackson Immunoresearch, 109-605-098). Washing steps were performed prior to every incubation step with 0.1% Tween-20 in 1x TBS. After the final incubation step the microarrays were washed (0.05% Tween-20 in 0.1x SSC) and dried in a stream of nitrogen.
Each microarray was scanned using a GenePix Autoloader 4300 SL50 (Molecular Devices, Pixel size: 10 μm). Signal intensities were evaluated using GenePix Pro 7.0 analysis software (Molecular Devices). For each peptide, the MMC2 value of the three triplicates was calculated. The MMC2 value was equal to the mean value of all three instances on the microarray except when the coefficient of variation (CV) – standard-deviation divided by the mean value – was larger than 0.5. In this case the mean of the two values closest to each other (MC2) was assigned to MMC2. Further data analysis and generation of the heatmaps was performed using the statistical computing and graphics software R (Version 4.0.2, www.r-project.org).

**RT-PCR test.**

We used the protocol described by the centers for disease control and prevention for the CDC 2019 Novel Coronavirus (2019-nCoV) Real-Time Reverse Transcriptase (RT)-PCR Diagnostic Panel.

**ELISA test.** To detect antibodies, we used the Euroimmun Anti-SARS-CoV-2 ELISA (IgG), which was manufactured by Euroimmun Medizinische Labordiagnostika AG, Lübeck, Germany. The break-off microplate wells were covered with the antigen, a recombinant structural spike 1 (S1) protein of SARS-CoV-2 expressed in HEK293 [FDA, 2020]. First, the diluted patient samples were incubated in the wells which lead to specific IgG antibodies binding to the antigens. In order to detect the bound antibodies, an enzyme-labelled anti-human IgG antibody detected antigen-antibody complexes and catalysed a color reaction. By calculating the extinction of the sample over the extinction of the calibrator the results could be evaluated semi-quantitatively. A ratio below 0.8 was considered a negative result. A ratio between 0.8 and 1.1 was considered a borderline result. A ratio above 1.1 was considered a positive result [Euroimmun Medizinische Labordiagnostika AG, 2020].

**QUANTIFICATION AND STATISTICAL ANALYSIS**

Data analysis and heatmap generation was performed using the statistical computing and graphics software R (Version 4.0.2, www.r-project.org).

The data set used for statistical analysis of the microarray results and for generation of all heatmap presentations was based on MMC2 values calculated from signals of the triplicate spots of each individual peptide. The MMC2 value calculation is described in detail in the
Microarray assay and data analysis section.

The principle of the colour coding of all heatmaps is described in the corresponding figure legends. The coloured fields in the heatmaps reflect the calculated MMC2 values of each individual peptide.

The boxplots in Figure 1 and Figure 2 show the median of the corresponding sample groups. The central rectangles span the first quartile to the third quartile (the interquartile range or IQR). The whiskers represent the following values: upper whisker = min(max(x), Quartile 3 + 1.5 * IQR); lower whisker = max(min(x), Quartile 1 – 1.5 * IQR). The points in the boxplots show the MMC2 values for the individual samples.

Statistical comparison of two sample groups was done by the Wilcoxon Rank Sum test (R package stats, n=24 patient samples and n=12 control samples). ROC analysis was performed using the R package ROCR. The selection of the peptide hits with significant separation between the sample groups was based on the accuracy value calculated as ‘(true positive + true negative)/total number of observations’. An accuracy value of ≥0.4 was considered to indicate immunogenicity (Supplementary Table 2).

The Spearman’s rank correlation coefficient reflecting the relationship between the commercial ELISA and the peptide microarray (Figure 3) was calculated as described in the figure legend using the R package ggpubr.

References


This article is protected by copyright. All rights reserved.


Graphical abstract

SARS-CoV-2 whole-proteome peptide microarray analysis of antibody reactivity in Covid-19 patient sera revealed several epitopes and confirmed the presence of humoral cross-reactivity between SARS-CoV-2 and ubiquitous "common cold" human coronaviruses. The identified epitopes may provide the basis for the development of a specific and sensitive diagnostic test.