Auto-Immunoproteomics Analysis of COVID-19 ICU Patients Revealed Increased Levels of Autoantibodies Related to Male Reproductive System

Frank Schmidt[1*#], Houari B. Abdesselem[2,11#], Karsten Suhre[3], Muhammad U. Sohail[1], Maryam Al-Nesf[4,5], Ilham Bensmail[2], Fathima Mashod[1], Hina Sarwath[1], Joerg Bernhardt[6], Ti-Myen Tan[7,8,9], Priscilla E Morris[7,8,9], Edward J. Schenck[10], David Price[10], Nishant N. Vaikath[11], Vidya Mohamed-Ali[5,12], Mohammed Al-Maadheed[5,12], Abdelilah Arredouani[13], Julie Decock[14], Jonathan M. Blackburn[7,8,9], Augustine M.K. Choi[10], Omar M. El-Agnaf[11*]

[1]Proteomics Core, Weill Cornell Medicine - Qatar, Doha, Qatar.
[2]Proteomics Core Facility, Qatar Biomedical Research Institute (QBRI), Hamad Bin Khalifa University (HBKU), Qatar Foundation, Doha, Qatar.
[3]Bioinformatics Core, Weill Cornell Medicine - Qatar, Doha, Qatar.
[4]Hamad General Hospital, Hamad Medical Corporation, Doha, Qatar.
[5]Center of Metabolism and Inflammation, Division of Medicine, Royal Free Campus, University College London, Rowland Hill Road, London NW3 2PF, UK.
[6]Institute for Microbiology, University of Greifswald, Greifswald, Germany.
[7]Department of Integrative Biomedical Sciences, Faculty of Health Sciences, University of Cape Town, South Africa.
[8]Sengenics Corporation, Level M, Plaza Zurich, Damansara Heights, Kuala Lumpur 50490, Malaysia.
[9]Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, South Africa.
[10]Department of Medicine, Division of Pulmonary and Critical Care Medicine, New York Presbyterian Hospital – Weill Cornell Medical Center, Weill Cornell Medicine, New York, NY, USA.
[11]Neurological Disorders Research Center, QBRI, HBKU, Qatar Foundation, Doha, Qatar.
[12]Anti-Doping Laboratory Qatar, Doha, Qatar.
[13]Diabetes Research Center, QBRI, HBKU, Qatar Foundation, Doha, Qatar.
[14]Cancer Research Center, QBRI, HBKU, Qatar Foundation, Doha, Qatar.

[#] These authors contributed equally to the work.

[*] Corresponding Authors:

Frank Schmidt
Address: Proteomics Core, Weill Cornell Medicine - Qatar, Doha, Qatar
Tel: 0097450446815, E-Mail: frs4001@qatar-med.cornell.edu

Omar El-Agnaf
Address: Neurological Disorders Research Center, QBRI, HBKU, Qatar Foundation, Doha, Qatar.
Tel: 0097455935568, E-Mail: oelagnaf@hbku.edu.qa

Conflict of Interest Disclosures: no conflict of interest

Running Head: KREX-autoimmune COVID-19 replication study

Manuscript word count: 4871 words

Number of Figures: 9

Number of Supplemental Tables: 1, Number of Supplemental data file: 1

Abstract

The role of autoantibodies in coronavirus disease (COVID-19) complications is not yet fully understood. The current investigation screened two independent cohorts of 97 COVID-19 patients (Discovery (Disc) cohort from Qatar (n = 49) and Replication (Rep) cohort from New York (n = 48)) utilizing high-throughput KoRectly Expressed (KREX) immunome protein-array technology. Autoantibody responses to 57 proteins were significantly altered in the COVID-19 Disc cohort compared to healthy controls ($P \leq 0.05$). The Rep cohort had altered autoantibody responses against 26 proteins compared to non-COVID-19 ICU patients that served as controls. Both cohorts showed substantial similarities ($r^2 = 0.73$) and exhibited higher autoantibodies responses to numerous transcription factors, immunomodulatory proteins, and human disease markers. Analysis of the combined cohorts revealed elevated autoantibody responses against SPANXN4, STK25, ATF4, PRKD2, and CHMP3 proteins in COVID-19 patients. KREX analysis of the specific IgG autoantibody responses indicates that the targeted host proteins are supposedly increased in COVID-19 patients. The autoantigen-autoantibody response was cross-validated for SPANXN4 and STK25 proteins using Uniprot BLASTP and sequence alignment tools. SPANXN4 is essential for spermiogenesis and male fertility, which may predict a potential role for this protein in COVID-19 associated male reproductive tract complications and warrants further research.

Keywords: COVID-19, Autoantibodies, Immunoproteomics, SPANXN4, STK25, and Male reproductive system

77 **Significance Statement**

78 Coronavirus disease (COVID-19), caused by the SARS-CoV-2 virus, has emerged as a global
79 pandemic with a high morbidity rate and multiorgan complications. It is observed that the host
80 immune system contributes to the varied responses to COVID-19 pathogenesis. Autoantibodies,
81 immune system proteins that mistakenly target the body's own tissue, may underlie some of this
82 variation. We screened total IgG autoantibody responses against 1,318 human proteins in two
83 COVID-19 patient cohorts. We observed several novel markers in COVID-19 patients that are
84 associated with male fertility, such as sperm protein SPANXN4, STK25, and the apoptotic factor
85 ATF4. Particularly, elevated levels of autoantibodies against the testicular tissue-specific protein
86 SPANXN4 offer significant evidence of anticipating the protein role in COVID-19 associated male
87 reproductive complications.
88

Introduction

Coronavirus disease (COVID-19), caused by novel SARS-CoV-2 virus, has emerged as global pandemic with severe complications and high morbidity rate. The disease manifests a wide range of clinical symptoms, which are exacerbated by overactive and malfunctioning immune system of the host. Despite extensive research on innate and adaptive immune responses in COVID-19, little is known about the role of autoantibodies on disease progression and severe complications.

Infection with the SARS-CoV-2 causes a variety of symptoms, with most cases being moderate or asymptomatic, and only a smaller proportion advancing to more severe state of COVID-19 disease[1]. Many questions about the COVID-19 pathophysiology remain open, particularly why some people develop severe disease symptoms while others remain asymptomatic.

Acute respiratory distress syndrome (ARDS) affects a small percentage of patients, whereas others experience persistent lung damage and multi-organ illness that lasts months, even after the virus has been eliminated from the body[2]. High expression of angiotensin-converting enzyme 2 (ACE2) receptors in several organs of the body extends infection beyond respiratory tract, resulting in complex multiorgan complications[3]. ACE2 receptors are highly expressed in the male reproductive system, demonstrating the involvement of SARS-CoV-2 in male fertility, which is one of the unexplained manifestations of COVID-19[4].

Autoantibodies have been identified in significant proportion of COVID-19 hospitalized patients with positive correlation with immune responses to SARS-CoV-2 proteins[5]. Several studies observed significant rise in a diverse range of autoantibodies against immunomodulatory proteins, a- and w-interferons, cardiolipin and prothrombin during antiviral responses in severely ill COVID-19 patients[6, 7, 8, 9, 10, 11]. Particularly, autoantibodies against immune-related signaling proteins were found to contribute to COVID-19 pathogenesis by antagonizing the function of the innate immune system[12]. Although there have been some reports on disease-modifying autoantibody responses, the immunological and clinical consequences of autoantibodies in COVID-19 are yet to be fully understood. Here, we therefore screened total IgG autoantibody responses against 1,318 human proteins in COVID-19 patients using KREX immunome protein-array technology. Sengenics KREX technology employs full-length, naturally folded proteins that allow maximum epitopes binding to discover autoantibody biomarker proteins[13]. The quantitative signal measured on the arrays for each autoantibody-autoantigen pair is directly proportional to the autoantibody concentration in the blood with higher autoantibody titres to these proteins simplistically implying higher autoantigen concentrations in COVID-19 patients compared to controls, albeit the correlation is non-linear.

Autoantibody-based precision immuno-profiling has previously been shown to aid discovery of biomarkers of immune-related adverse events, as well as therapeutic prediction of drug response[14]. In the present study, by utilizing a broad array-based immunoproteomics strategy that simultaneously quantifies autoantibody responses across multiple organ systems in ICU COVID-19 patients and post recovery cohort, we aimed to better identify novel markers of comorbidities in COVID-19 patients. We identified a number of novel markers in COVID-19 patients that are also associated with male fertility, such as the sperm protein SPANXN4[15], the androgenic kinase

129 STK25[16, 17], the apoptotic factor ATF4[18], the calcium channel regulator protein kinase PRKD2[19],
130 and the multivesicular protein CHMP3[20].
131
132

Methods

Study design, samples collection and processing and ethics

We used blood samples and clinical data of patients from two independent COVID-19 cohorts to conduct a comprehenisve anlaysis of autoantibodies using novel KREX technology.

Discovery (Disc) cohort

The Disc cohort included forty-nine COVID-19 patients from Qatar who were admitted to Hamad Medical Corporation hospitals. All recruited patients had confirmed SARS-CoV-2 positive RT-PCR results of sputum and throat swabs. All patients had severe COVID-19 disease (WHO guideline)[21] and were admitted to intensive care unit (ICU). Peripheral blood was collected within five to seven days of admission and processed into plasma and serum, which were stored at –80^{o}C, until further analysis. Ethical approval for this cohort was obtained from the Hamad Medical Corporation Institutional Review Board Research Ethics Committee (reference MRC-05-003), and Qatar Biomedical Research Institute- Institutional Review Board (Reference QBRI-IRB 2020-06-19).

Healthy controls

Age and gender matched healthy volunteers (n=48) with no prior COVID-19 infection history and with normal oxygen saturation and vital signs were used as controls. The Anti-Doping Laboratory-Qatar recruited them for blood collection. Individuals with medical history or with cognitive disabilities were excluded. All participants (patients and controls) provided written informed consent prior to enrolment in the study.

Replication (Rep) cohort

The replication cohort consisted of forty-eight adult patients who were admitted to the ICU of New York-Presbyterian Hospital (NYP)/Weill Cornell Medical Center (WCMC) from March to April 2020. All patients were RT-PCR confirmed SARS-CoV-2 positive and displayed ARDS or pneumonia symptoms. The cohort is part of the Weill Cornell Biobank of Critical Illness, a registry which attempts to recruit and enroll all patients being admitted to WCMC ICU for clinical investigations. The WCMC COVID Institutional Data Repository (COVID-IDR), a manually abstracted registry of COVID-19 patients that was developed to record patient demographics and allied health parameters. Laboratory parameters, ventilation records, respiratory variables and vital signs were recorded and documented at Weill Cornell-Critical Care Database for Advanced Research (WC-CEDAR)[22]. The processes for recruiting patients, collecting data, and processing samples had all been previously documented[23, 24]. Only patients that gave informed consent were included. IRB approvals for this cohort were obtained from NYP/WCMC with reference number 20-05022072 and 1405015116.

Non-COVID-19 ICU controls

Twenty-eight patients admitted to NYP hospital ICU between 2014 to 2019 were included as non-COVID-19 ICU controls for the Rep cohort. These patients were suffering from bacterial sepsis ARDS (N = 15), influenza ARDS (N = 4), and influenza pneumonia (N = 9). The patient recruitment, medical history, and sampling procedures for the non-COVID ICU control cohort are the same as those described for the Rep cohort.

Sengenics assay description and data pre-processing

The Disc cohort samples were processed at Qatar Biomedical Research Institution (QBRI) for KREX immunoproteomics. The Rep cohort samples were processed at the Sengenics facility in Kuala Lumpur, Malaysia. Samples of Disc cohort and controls were analyzed for antigen-specific autoantibodies using Immunome protein arrays (Sengenics), developed using KoRectly Expressed (KREX) technology to provide a high-throughput immunoassay based on correctly folded, full length and functional recombinant human proteins expressed in insect cells, thereby displaying a full repertoire continuous and discontinuous epitopes for autoantibody binding[25, 26]. The Immunome arrays contain more than 1,600 human antigens, enriched for kinases, signaling molecules, cytokines, interleukins, chemokines, as well as known autoimmune- and cancer antigens. Plasma samples of Rep cohort and non-COVID-19 ICU control patients were then processed for autoantibodies on a custom array containing a subset of 1,318 human proteins (Sengenics).

Samples were viral-inactivated in 10% Triton X-100 for 2 hours at room temperature. Samples were then diluted in Serum Albumin Buffer (SAB) at optimized dilution (50-fold dilution). Microarray slides were prepared in four-well plates slide. Samples including controls were randomized and applied to the microarray slides for 2 hours and samples' IgGs were then detected by secondary Cy3-labeled IgG antibodies. Slides were scanned at a fixed gain setting using the Agilent G4600AD fluorescence microarray scanner generating a 16-bit TIFF file. A visual quality control check was conducted and any array showing spot merging or other artefacts were re-assayed. A GAL (GenePix Array List) file containing information regarding the location and identity of all probed spots was used to aid with image analysis. Automatic extraction and quantification of each spot was performed using GenePix Pro 7 software (Molecular Devices) yielding the median foreground and local background pixel intensities for each spot.

Biotinylated human IgG (detected by fluorescently labelled secondary antibody) and biotinylated human anti-IgG (detected only when plasma or serum is added to the slide) were used as positive controls to assess assay integrity. Extrapolated data was then filtered, normalized and transformed as follows: Briefly, the median background pixel intensity (RFU) was subtracted from the median foreground pixel intensity (RFU) for each antigen to give the median net intensity per spot (RFU); CVs were calculated for each antigen based on the quadruplicate technical replica spots for each antigen on a given array, any antigens with CV above 20% were flagged and outlier spots removed, providing that at least two valid values remained; net intensity values for each antigen in a given sample were calculated as the mean of the net intensity values for technical replica spots on that array; and data was normalised across replica arrays based on the Cy3-BSA controls as previously

described[27]. Z-scores were then calculated by subtracting the overall mean antigen intensity (within a single sample) from the net intensity data for each antigen in that sample, and dividing that result by the standard deviation of all of the measured net intensities in that sample, according to the formula: $\mathbf{z = (x - \mu) / \sigma}$ where $\mathbf{x}$ is the net intensity of an antigen in a given sample, m is the mean net intensity calculated across all antigens in that sample, and s is the standard deviation of the net intensities for all antigens in that sample. All downstream statistical analysis was done based on the calculated z-scores.

Sequence identity and antigen specificity analysis for selected proteins

We needed to be cautious in directly comparing the results across different antigens on the arrays because the autoantigen-autoantibody response is not always linear and is an indirect way of prediction of protein concentrations. Since it can depend amongst others on both B cell activation and sequence identity among proteins that express similar antigen epitopes. To check this latter possibility, we selected two proteins (SPANXN4 and STK25) that showed the highest autoantibody alterations to perform their sequence alignment and antigen specificity analysis. Uniprot BLASTP program was used to compare proteins sequences. All human and viral protein sequences with more than 50% sequence similarity were aligned for epitope mapping to determine whether the evaluated RFU values were specific to the protein of interest or could be derived from highly homologous epitopes on other proteins.

Protein pathways prediction

The assingment of KREX array proteins to functional KEGG categories and their hierarchical organisation was displayed by using Paver, a software for the visualization of Voronoi Treemaps[28]. Any main category is displayed in different colors. The cell sizes were calculated according the signal intensity of the proteins immunofluorescence (highly fluorescent signals give larger cells). Functional Enrichment Analysis was performed to identify biological functions that were over-represented in differentially expressed proteins with a p-value less than 0.05. Differentially expressed proteins, both up-regulated and down-regulated, were used separately as proteins of interest and the proteins detected from all probes were used as the background set. The proteins were further annotated using KEGG- and WIKI-Pathways data prior to performing Fisher's exact test to determine pathways in which the proteins of interest were significantly over-represented. This analysis was performed on R 3.6.2 using clusterProfiler 3.14.3. GOSemSim was used to eliminate redundant GO-BP results. Only significantly over-represented pathways with a p-value less than 0.05 (-log10 p-value cut-off 1.3) are shown.

Statistical analysis

Proteins are reported using the symbols of the genes that encode them to offer a clear and uniform nomenclature. Autoantibody response, measured as relative fluorescence units (RFU), was normalized to calculate z-score. Statistical analysis was performed using R (version 4.1.0) and

rstudio (version 1.4.1717). Two kinds of inferal statistical tests were performed to test the hypotheses of whether a given autoantibody was differentially expressed in COVID cases versus controls. First, the means between cases and controls were compared using a linear model, using the z-scored autoantibody responses as dependant variables and the COVID state as independent variable (coded as 0=controls and 1=cases). Note that this approach is equivalent to conducting an unrelated T-test and that the effect size of the linear model matches the estimated difference of the means in a T-test. Second, binarized autoantibody responses were tested against cases versus controls using Fisher's exact test. The cutoff for binarization of the autoimmune response was set to one. As the response is z-scored, this means that all samples with an RFU score above one standard deviation from the mean were considered as being positive for the respective autoantibody whereas all other were considered negative. The Disc and Rep cohorts were analyzed separedly and then merged. Therefore three sets of p-values were obtained for each of the continous and binarized trait analyses.

Following comparisons were made:

- Differential autoantibody response analysis: COVID-19 cases versus controls for
    - Forty-nine Disc COVID-19 patients vs. fourty-eight controls.
    - Forty-eight Rep COVID-19 patients vs. twenty-eight non-COVID-19 ICU controls.
    - Combined ninety-seven COVID-19 vs. seventy-six controls from both Disc and Rep cohorts.
- Pearson's correlation analysis of fifteen Disc cohort patients sampled at the time of ICU admission (T1) and six weeks follow-up (T2).
- Principal component analysis and pearson's correlation analysis to compare two cohorts.

Results

Study design and cohort specific information

In the present work, two ethnically independent cohorts of COVID-19 patients were evaluated for their autoimmune response (total IgG-response) against 1,318 naturally folded human proteins (antigens). The Disc cohort was recruited at ICU of HMC in Doha, Qatar and included 49 COVID-19 cases and 48 healthy controls, majority of whom were male. A second cohort was recruited from the ICU of NYP Hospital, USA, which included 48 COVID-19 cases and 28 control patients and served as a Rep study. In addition, patients who were admitted to the NYP ICU and had infectious diseases other than COVID-19, such as bacterial sepsis ARDS or H1N1 pneumonia, were included as controls for the Rep cohort. Because of the special composition of the cohorts, we were able to specifically look for COVID-19-related autoantibody signals compared with healthy-baseline- and general infection-baseline-titres. A combined analysis (discovery and replication) allowed stringent COVID-19-specific autoimmune responses to be monitored. Table 1 summarises the demographic and status-specific information of the study cohorts.

General autoantibody response in healthy and COVID-19 patients

To discover functional IgG-related autoantibodies that could influence COVID-19 predictions and/or outcomes, we used the KREX high-throughput autoantibody assay technology that includes a variety of known human-autoantigens such as cancer-, kinase-, interleukins-, cytokine, ribonuclear transcription and signaling-proteins[29]. Total IgG autoantibody responses were quantified for 1,600 proteins in the Discovery Cohort and for a subset of 1,318 proteins in the Replication Cohort. However, to increase stringency and reduce complexity, only the 1,318 overlapping proteins were subsequently used in the analysis pipeline. The majority of antigens on the array are found in the cytoplasm, nucleus, or cell membrane, but there are also proteins from the mitochondria, endoplasmic reticulum, and cytoskeleton.

The KREX assay reports RFU values for autoantigen-specific autoantibody binding, with linearity over 6 orders of magnitude and with a detection limit in the pg/ml range. These measured RFU values correlate directly with the antigen-specific IgG autoantibody titres, since ligand binding theory shows that the measured signal on-array is linearly proportional to autoantibody concentration. Thus, a higher RFU value for a specific autoantibody-autoantigen interaction indicates a higher autoantibody titre, whilst a higher antibody titer in turn implies a higher autoantigen concentration (or repeated exposure to the autoantigen), accepting that this latter correlation is non-linear. In a first overview, the general intensity distributions were calculated based on the mean autoantibody-antigen titers across all samples and further examined using KEGG-Brite-based Voronoi treemaps using the replication cohort as an example (Figure 1). Approximately 1,150 of the 1,318 proteins could be assigned to the annotation, with the relative size of each cell on the Voronoi treemaps reflecting the observed autoantibody response against that protein (Figure 1 left). Nearly all proteins showed a total IgG AB-signal in the cases and the corresponding controls, the latter represents the natural autoimmunity or the healthy repertoire of autoantibodies. In Figure 1 right, the corresponding pathways are summarized in different colors,

with most proteins belonging to the MAPK pathway (light blue), followed by transcription factors (green), chromosomal proteins (green), ribosomes (all blue), and metabolic proteins (yellow). A few proteins belong to the cell cycle (red), chemokines (cyan) or cancer (black). The 10 highest autoantibody titers were found against RBPJ, TPM1, TACC1, KRT19, PTPN20, TBCB, KRT15, AFF4, HSPD1, and CBFA2T3, many of these are structure related proteins. The 10 proteins with the lowest titers were AIF1, IL18, NCK1, COMMD3, NEK11, TGFBR2, SLA, PKM, MAPK6, and MLKL, many of which are cytoplasmic proteins involved in phosphorylation.

Relative autoantibody response in the Disc cohort revealed significantly higher level of SPANXN4 and ATF4

To examine the effects of SARS-CoV-2 infection on the autoantibody response, we first performed a differential expression analysis in Disc cohort between COVID-19 cases and healthy controls using T-test. Autoantibody responses of fifty-seven proteins were altered significantly (T-test p-value ≤ 0.05) (Supplementary file sheet 3). Autoantibody responses in COVID-19 patients were increased for forty proteins, while decreased for seventeen proteins (Figure 2A). The most elevated autoantibody responses in COVID-19 patients were against ATF4 (effect size (beta) = 3.32 SD; T-test p-value ≤ 0.001) and the sperm protein associated with the nucleus on the X chromosome N4 (SPANXN4) (effect size (beta) = 3.32 SD; T-test p-value ≤ 0.001). The latter is also known as spermiogenesis-related protein and belongs to the family of cancer/testis-associated proteins (CTAs)[30].

We then conducted an analysis using binarized autoimmune response, assuming that all samples with an autoimmune response that exceeds the mean by one s.d. as positive and all others as negative (Supplementary file sheet 4). Using Fisher's exact test, we found twenty-five COVID-19 patients had higher RFU values for SPANXN4 compared to only five in controls (Fisher's test p-value ≤ 0.0001) (Figure 2B). Autoantibodies against ATF4, recombining signal binding protein J (RBPJ), and programmed cell death 5 (PDCD5) were also significantly elevated (Fisher's test p-value ≤ 0.05) in the COVID-19 patients. Only the binarized SPANXN4 association reaches the most stringent Bonferroni significance level, that is $p < 0.05$ / number of proteins = 1,318. In the control group EAPP, SSNA1, and LDHB proteins showed higher autoantibody responses than the cases.

Autoantibody response in the Disc Follow-up cohort confirmed high levels against SPANXN4 and other proteins in COVID-19 patients

Following the initial blood sample collection at the time of ICU admission, follow-up samples were collected from fifteen patients at six weeks after recovery from COVID-19. For several proteins, a strong correlation (Pearson's $r^2 \geq 0.69$) was observed between the autoantibody responses at the two sampling time points (Figure 3A). Autoantibody responses against several proteins, including SPANXN4, STK25, TRAF3IP1, AMOTL2, PSMD4, and PPP1R2P9 remained highly elevated ($p \leq 0.05$) at 6 weeks post-recovery follow-up. Particularly, autoantibody responses against SPANXN4 (Figure 3B) stayed elevated at both initial (T1) and follow-up (T2)

time points. These observations reveal that SPANXN4 autoantibody responses remain elevated for extended periods, suggesting potential association with chronic health issues.

Relative autoantibody response in the Rep cohort confirms the trend in the Disc cohort

Autoantibody response for the Rep cohort (n = 48) was compared with the non-COVID-19 ICU control patients (N = 28) (Figure 4A). Autoantibody responses of twenty-six proteins altered significantly (T-test p-value ≤ 0.05) in the Rep cohort. Based on T-test analysis, the most elevated autoantibody response in Rep COVID-19 cohort was found for PRKD2 and BACH1 proteins, which are known for their roles in male reproductive tract development (PRKD2)[31] and spermatogenesis (BACH1)[32]. Autoantibody response for SPANXN4 was also higher (effect size (beta) = 1.61) in the Rep COVID-19 patients, albeit p-value was slightly higher than 0.05.
However, Fisher's exact test indicated that autoantibody response to SPANXN4 remained the highest (Fisher's test p-value = 0.0036) (Figure 4B). At sigma 1, fifteen COVID-19 patients had higher RFU values for SPANXN4 compared to only one in controls. SPANXN4 can therefore be considered fully replicated under the highest standards of a discovery-replication design. PDCD2L, PRKD2, and STK25 showed also higher autoantibody responses in COVID-19 patients (Fisher's test p-value ≤ 0.05).

Analysis of combined cohorts supports that SPANXN4 and STK25 are significantly elevated in COVID-19 patients independent of sampling martix or patients ethicity

Principal components analysis (PCA) of protein RFU data from the two cohorts demonstrated strong overlap between COVID-19 samples and the two cohorts did not separate into discrete clusters (Figure 5). Pearson's correlation analysis revealed that the autoantibody responses of the two cohorts have high correlation ($r^2 = 0.73$).
At third stage, we combine data from both the Disc and Rep cohorts (n = 97) and compared them with combined controls (n = 76). Case vs. control analysis revealed that autoantibody responses against fifty-six proteins were significantly altered: 35 autoantibodies with increased and 21 autoantibodies with decreased responses (T-test $p \leq 0.05$) (Figure 6A). SPANXN4, ATF4, STK25, and PRKD2 were the proteins with the highest effect size (beta). In total forty patients had SPANXN4 RFU higher than 1 sigma value (Fisher's exact test p-value ≤ 0.0001) in the combined COVID-19 cohorts compared with the six patients only in controls (Figure 6B).
Furthermore, the autoantibody responses, expressed as RFU z-score for fifty-six proteins that differed significantly between the study groups are shown in Figure 7A. The heatmap shows that most of the proteins display similar pattern of autoantibody ratios across the study cohorts. These analyses demonstrate that our autoantibody response data are highly reproducible despite differences in population ethnicity, different laboratories, and sampling materials (serum vs. plasma in Disc vs. Rep cohorts, respectively).

Protein pathways analysis uncovered up-regulated immune pathways in COVID-19 patients

KEGG and WIKI pathways analysis was performed to identify the functional contribution of autoantibodies targeted proteins in cellular processes and immune-inflammatory systems. Pathways associated with T helper cells (Th1, Th2, and Th17) differentiation, bacterial/viral infections, stress hormones release, and prostate cancer were upregulated in COVID-19 patients (Figure 7B). WIKI pathways were also activated for host immunity and interferon signaling, including T cell activation for SARS-CoV-2 and *Staphylococcus aureus* infections (Figure 7B).

SPANXN4 and STK25 share sequence identity with SPANX- and STK-family proteins but showed unique AB-titers in COVID-19 patients

In order to check cross-reactivities, sequence homology and antigen specificity analysis were performed for SPANXN4 and STK25 against human and viral protein databases. Only few proteins appeared to have more than 50% sequence identity with our target proteins ((SPANXN4 with SPANXN1, 2, 3, and 5) and (STK25 with STK 3, 4, 24, and 26)) (Figure 8 and 9). However, many of these homologous proteins were also part of our KREX immunome panel but did not show any significant changes, which means that the observed RFUs are highly specific against the targeted proteins.

Discussion

In the current COVID-19 pandemic, there is increasing interest globally in understanding the underlying immunology of COVID-19, as well as revealing new health issues arising from COVID-19 complications. Several papers have described the existence and cross-reactivity of SARS-CoV-2 specific T-cell responses[33, 34, 35, 36], as well as correlations with male reproductive system and infertility[4, 37]. The present study identified and validated several autoantibody responses by screening two independent cohorts of COVID-19 patients with the KREX immunome protein array. The proteins identified with higher autoantibody responses serve important physiological functions and are strongly associated with various immunological and pathological parameters associated with COVID-19 disease.

The KREX immunome array contains proteins involved in physiological processes such as MAPK signaling, metabolism, transcription, cell cycle, immunity, and cancer-related pathways. Few of the proteins with the highest mean autoantibody response in COVID-19 patients were RBPJ, TPM1, TACC1, KRT19, and PTPN20. These proteins perform a variety of physiological functions in the human body, with many of them being structural proteins involved in tissue damage and repair mechanisms[38]. The presence of a high autoantibody response to these proteins suggests that they are overproduced during a pathological condition, such as cancer or a cardiovascular complication[39, 40]. For example, notch signally protein RBPJ has been associated with COVID-19 pathophysiology and cardiovascular complication[41]. Similarly, keratin family proteins (KRT19 and KRT15) that are responsible for epithelial cell structural integrity are linked to COVID-19 pathogenesis and disease severity[42]. Furthermore, many of these proteins are also involved in male

reproductive system physiology and fertility, yet there has been no previous report in COVID-19 patients.

Our Disc cohort reported higher autoantibodies against SPANXN4, ATF4, RBPJ, and PDCD5 proteins compared to the controls. Comparison between COVID-19 baseline (T1) vs. follow-up (T2) samples indicated that SPANXN4 autoantibodies remained elevated at post-recovery stage. Prolonged autoantibody responses may highlight COVID-19 post-acute sequelae by stimulating the humoral immune response in a way that leads to long-term autoantibody production[43]. The diverse variety of proteins linked to a prolonged autoantibody response suggest that SARS-CoV-2 may stimulate autoantibody formation by molecular mimicry[44], targeting cardiolipin, cardiolipin-binding proteins, platelet factor 4, prothrombin, and coagulation factors, suggesting their role in coagulopathies, chronic comorbidities and post-infection recovery[45, 46, 47]. We hypothesize that elevated autoimmune antibodies against SPANXN4, STK25, TRAF3IP1, AMOTL2, PSMD4, and PPP1R2P9 might suggest a similar role. However, Dotan et al.[48] investigated *in-silico* sequence homology of all human proteins with the virus but could not find evidence that any of the proteins mentioned here are part of such a mimicry process. Vice versa, we cannot exclude that the titers might be elevated before the exposure to SARS-CoV-2, due to pre-existing diseases such as cancer or prolonged inflammation.

In contrast, the Rep cohort had higher levels of autoantibody responses to SPANXN4, PDCD2L, PRKD2, and STK25 proteins than the controls. Except for SPANXN4, all other proteins with the high autoantibody response were not significantly elevated between the two cohorts but often showed similar trends. These differences could be attributed to the fact that the control group in the Disc cohort was comprised of healthy volunteers, whereas the control group in the Rep study was comprised of ICU patients suffering from bacterial or viral ARDS, or pneumonia.

When all COVID-19 patients (N = 97) from both cohorts were merged and compared to all controls (N = 76) from both cohorts, the most significant autoantibody responses were observed against SPANXN4, ATF4, STK25, and PRKD2. ATF4 regulates metabolic and redox processes in the human body, and an increased ATF4 response has been observed in previous coronavirus disease[49, 50]. Fischer et al.[18] suggested that ATF4 also plays role in differentiation of the vas deferens lamina propria layer that helps improve spermatozoa fertilization rate. STK25 and PRKD2 are two important kinases with several physiological roles in our body. However, their role in male reproductive tract physiology is least discussed. A few studies highlight STK25 as androgenic kinase[16, 17] and PRKD2 role in male reproductive tract development[19].

SPANXN4 belongs to a protein family called “sperm protein associated with nucleus in the X chromosome” (SPANX) that are essential for motility and fertilization capacity of male-ejaculated spermatozoa[15]. SPANXNs are also known as cancer testis antigens (CTAs) because of their overexpression in tumor tissues, in addition to their normal physiological role in the testis and spermatozoa of healthy males[51]. SPANX proteins are expressed in various regions of sperms, and

research has shown that the protein family has relevance to male fertility. Particularly, the presence of ACE2 receptors in testicular tissues suggests that SARS-CoV-2 influences male fertility, but the pathogenesis is not clear. No previous COVID-19 study has mentioned SPANXN4 involvement in male infertility, neither the pathogenesis is explained. Therefore, the autoantibody response measured in the current study may suggest a novel diagnostic and treatment marker for male fertility. Previously, one hepatitis C virus study has shown that SPANXN4 interacts with the virus, potentially increases virus infectivity, albeit no reproductive performance was discussed[52]. Increased levels of autoantibodies against testis-related proteins suggest their role in affecting male reproductive system, and thus declining male fertility in COVID-19. Although several investigations have found that COVID-19 patients have altered seminal parameters and decreased reproductive hormone levels[53], histological or functional abnormalities in male genital system[54], damaged blood-testis barrier[55], and impaired spermatogenesis[56], the cause of this comorbidity has not yet been investigated, and remains unknown.

Despite their ethnic diversity, which included Middle Eastern, Africans, Caucasians, Asians, and South Asians populations, correlation analysis, hierarchical, and PCA clustering demonstrated that both cohorts shared similarities in autoantibody responses. Therefore, strong correlation ($r^2 = 0.73$) between these cohorts demonstrate that autoantibody response data of COVID-19 patients are highly reproducible among different ethnic populations. These findings are consistent with our prior COVID-19 proteomics study that looked at immune-inflammatory markers in five different demographic cohorts (manuscript accepted).

Several previous COVID-19 studies have reported elevated immune-inflammatory responses, including cytokine-storm in COVID-19 patients. Perhaps, we observed relatively elevated albeit non-significant autoantibody responses to immune cytokines like IL1A and IL1B proteins. KEGG and WIKI pathways analysis showed that autoantibody responses to immune proteins activated T cell responses against infection and T helper cell differentiation. Li et al.,[57] observed that Th17 differentiation and cytokine response pathways play a key role in pathogenesis of COVID-19 and autoimmune diseases. Pathways analysis suggests that many immune cell responses specific to SARS-CoV-2 or bacterial infections may precede chronic inflammatory disorders and the respiratory failure[58]. Furthermore, an abnormal T helper cell response, combined with overactive interferon signaling, promote the differentiation of B cells, which produce autoantibodies and cause autoimmune diseases[59].

In conclusion, these findings reveal unique autoantibody response against several proteins that play diverse though important function in COVID-19 complications. These observations also highlight the importance of the humoral immune response, as well as numerous other previously unknown immunological pathways in COVID-19 pathogenesis. Particularly, elevated levels of autoantibodies against the testicular tissue specific protein SPANXN4 in both cohorts offer significant evidence of anticipating the protein's role in COVID-19 associated male reproductive complications. Overall, these finding not only revalidate autoantibody responses against SPANXN4 in COVID-19 but also predict novel pathological associations that may contribute to COVID-19 post-recovery comorbidities. SPANXN family proteins are known as CTAs[60] that play

essential role in spermatogenesis, however, their role in male fertility in the COVID-19 patients is previously unknown.

Author Contributions:

F.S., H.B.A. and O.M.E. designed, conceived, and led the study. M.A.Y.A., V.M.A., and A.M. led the Disc Cohort sample collection, processing, and ethical approvals. H.B.A. and N.V. optimized the assays on Disc Cohort. I.B. and H.B.A. run the assays on Disc samples. K.S. and M.U.S. performed statistics. E.J.S., D.P., H.S., and A.M.K.C. organized and collected samples for Rep Cohort. T-M.T., P.E.M., and J.M.B. run the assays on Rep samples J.B. and F.M. performed annotation and treemap analyses. K.S., F.S., F.M., M.U.S., H.B.A., O.M.E., J.D., J.B., and A.A. interpreted the data. M.U.S., K.S., H.B.A., and F.S. wrote the manuscript. All authors reviewed the manuscript and have read and agreed to the published version of the manuscript.

Funding: Research conducted on Qatar cohort was funded by project funding from Infectious Diseases Interdisciplinary Research Program (ID-IDRP) at QBRI and a grant fund from Hamad Medical Corporation (fund number MRC-05-003).

Institutional Review Board Statement: The study on Qatar Cohort was conducted according to the Ministry of Public Health (MOPH) guidelines and approved by the Institutional Review Board Research Ethics Committee of the Hamad Medical Corporation (reference MRC-05-003).

Acknowledgments: We would like to thank all the patients, volunteers, and healthcare co-workers from HMC and NYP/WCMC hospitals, and Rudolph Engelke for providing the R-code for functional enrichment. We also acknowledge the help provided by the Anti-Doping Lab-Qatar (ADLQ) and Qatar Red Crescent (QRC) for recruiting control individuals. The Qatar cohort research was supported by the Proteomics core facility and ID-IDRP project at QBRI. This work was also strongly supported by the Biomedical Research Program at Weill Cornell Medicine in Qatar, a program funded by the Qatar Foundation.

Conflicts of Interest: There is no conflict of interest.

References

1. Chen G*, et al.* Clinical and immunological features of severe and moderate coronavirus disease 2019. *J Clin Invest* **130**, 2620-2629 (2020).

2. Pfortmueller CA, Spinetti T, Urman RD, Luedi MM, Schefold JC. COVID-19-associated acute respiratory distress syndrome (CARDS): Current knowledge on pathophysiology and ICU treatment – A narrative review. *Best Practice & Research Clinical Anaesthesiology*, (2020).

3. Ashraf UM*, et al.* SARS-CoV-2, ACE2 expression, and systemic organ invasion. *Physiological Genomics* **53**, 51-60 (2021).

4. Sun J. The hypothesis that SARS-CoV-2 affects male reproductive ability by regulating autophagy. *Med Hypotheses* **143**, 110083 (2020).

5. Chang SE*, et al.* New-onset IgG autoantibodies in hospitalized patients with COVID-19. *Nature Communications* **12**, 5417 (2021).

6. Bastard P*, et al.* Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science* **370**, (2020).

7. Garcia-Beltran WF*, et al.* COVID-19-neutralizing antibodies predict disease severity and survival. *Cell* **184**, 476-488.e411 (2021).

8. Wang EY*, et al.* Diverse functional autoantibodies in patients with COVID-19. *Nature* **595**, 283-288 (2021).

9. Woodruff MC, Ramonell RP, Lee FE-H, Sanz I. Clinically identifiable autoreactivity is common in severe SARS-CoV-2 Infection. *medRxiv*, 2020.2010.2021.20216192 (2020).

10. Woodruff MC*, et al.* Extrafollicular B cell responses correlate with neutralizing antibodies and morbidity in COVID-19. *Nature Immunology* **21**, 1506-1516 (2020).

11. Zuo Y*, et al.* Prothrombotic autoantibodies in serum from patients hospitalized with COVID-19. *Science Translational Medicine* **12**, eabd3876 (2020).

12. Tay MZ, Poh CM, Rénia L, MacAry PA, Ng LFP. The trinity of COVID-19: immunity, inflammation and intervention. *Nature Reviews Immunology* **20**, 363-374 (2020).

13. Applications of Functional Protein Arrays for Detection and Development of Autoantibody-based Diagnostics and Therapeutics.) (2018).

14. Fritzler MJ, Martinez-Prat L, Choi MY, Mahler M. The Utilization of Autoantibodies in Approaches to Precision Health. *Frontiers in Immunology* **9**, (2018).

15. Urizar-Arenaza I*, et al.* SPANX-A/D protein subfamily plays a key role in nuclear organisation, metabolism and flagellar motility of human spermatozoa. *Scientific Reports* **10**, 5625 (2020).

16. Gill-Sharma MK. Testosterone Retention Mechanism in Sertoli Cells: A Biochemical Perspective. *Open Biochem J* **12**, 103-112 (2018).

17. Hammes A*, et al.* Role of Endocytosis in Cellular Uptake of Sex Steroids. *Cell* **122**, 751-762 (2005).

18. Fischer C*, et al.* Activating transcription factor 4 is required for the differentiation of the lamina propria layer of the vas deferens. *Biol Reprod* **70**, 371-378 (2004).

19. Dai D-H*, et al.* Exploration of miRNA and mRNA Profiles in Fresh and Frozen-Thawed Boar Sperm by Transcriptome and Small RNA Sequencing. *International Journal of Molecular Sciences* **20**, 802 (2019).

20. Razavi SM, Sabbaghian M, Jalili M, Divsalar A, Wolkenhauer O, Salehzadeh-Yazdi A. Comprehensive functional enrichment analysis of male infertility. *Scientific reports* **7**, 15778-15778 (2017).

21. Organization WH. Clinical management of severe acute respiratory infection when novel coronavirus (nCoV) infection is suspected: interim guidance, 25 January 2020.). World Health Organization (2020).

22. Su C*, et al.* Identifying organ dysfunction trajectory-based subphenotypes in critically ill patients with COVID-19. *medRxiv*, 2020.2007.2016.20155382 (2020).

23. Finkelsztein EJ*, et al.* Comparison of qSOFA and SIRS for predicting adverse outcomes of patients with suspicion of sepsis outside the intensive care unit. *Crit Care* **21**, 73 (2017).

24. Schenck EJ*, et al.* Circulating cell death biomarker TRAIL is associated with increased organ dysfunction in sepsis. *JCI Insight* **4**, (2019).

25. Adeola HA, Smith M, Kaestner L, Blackburn JM, Zerbini LF. Novel potential serological prostate cancer biomarkers using CT100+ cancer antigen microarray platform in a multi-cultural South African cohort. *Oncotarget* **7**, 13945-13964 (2016).

26. Blackburn JM, Shoko A, Beeton-Kempen N. Miniaturized, microarray-based assays for chemical proteomic studies of protein function. *Methods Mol Biol* **800**, 133-162 (2012).

27. Da Gama Duarte J, Goosen RW, Lawry PJ, Blackburn JM. PMA: Protein Microarray Analyser, a user-friendly tool for data processing and normalization. *BMC research notes* **11**, 156 (2018).

28. Bernhardt J, Funke S, Hecker M, Siebourg J. Visualizing Gene Expression Data via Voronoi Treemaps. In: *2009 Sixth International Symposium on Voronoi Diagrams*) (2009).

29. Blackburn J*, et al.* Quantitative, Epitope-specific, Serological Screening of COVID-19 Patients Using a Novel Multiplexed Array-based Immunoassay Platform. *medRxiv*, 2020.2009.2025.20201269 (2020).

30. Kouprina N*, et al.* The SPANX gene family of cancer/testis-specific antigens: rapid evolution and amplification in African great apes and hominids. *Proc Natl Acad Sci U S A* **101**, 3077-3082 (2004).

31. Nie X, Arend LJ. Novel roles of Pkd2 in male reproductive system development. *Differentiation* **87**, 161-171 (2014).

32. Yu B*, et al.* Cigarette Smoking Is Associated with Human Semen Quality in Synergy with Functional NRF2 Polymorphisms1. *Biology of Reproduction* **89**, (2013).

33. Grifoni A*, et al.* Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **181**, 1489-1501.e1415 (2020).

34. Peng Y*, et al.* Broad and strong memory CD4(+) and CD8(+) T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nat Immunol* **21**, 1336-1345 (2020).

35. Reynolds CJ*, et al.* Discordant neutralizing antibody and T cell responses in asymptomatic and mild SARS-CoV-2 infection. *Sci Immunol* **5**, (2020).

36. Sekine T*, et al.* Robust T Cell Immunity in Convalescent Individuals with Asymptomatic or Mild COVID-19. *Cell* **183**, 158-168.e114 (2020).

37. Delle Fave RF, Polisini G, Giglioni G, Parlavecchio A, Dell'Atti L, Galosi AB. COVID-19 and male fertility: Taking stock of one year after the outbreak began. *Arch Ital Urol Androl* **93**, 115-119 (2021).

38. Zhang X, Yin M, Zhang L-J. Keratin 6, 16 and 17-Critical Barrier Alarmin Molecules in Skin Wounds and Psoriasis. *Cells* **8**, 807 (2019).

39. Lv Q, Shen R, Wang J. RBPJ inhibition impairs the growth of lung cancer. *Tumour Biol* **36**, 3751-3756 (2015).

40. Saha SK*, et al.* KRT19 directly interacts with β-catenin/RAC1 complex to regulate NUMB-dependent NOTCH signaling pathway and breast cancer properties. *Oncogene* **36**, 332-349 (2017).

41. Breikaa RM, Lilly B. The Notch Pathway: A Link Between COVID-19 Pathophysiology and Its Cardiovascular Complications. *Front Cardiovasc Med* **8**, 681948-681948 (2021).

42. Gisby J*, et al.* Longitudinal proteomic profiling of dialysis patients with COVID-19 reveals markers of severity and predictors of death. *Elife* **10**, e64827 (2021).

43. Proal AD, VanElzakker MB. Long COVID or Post-acute Sequelae of COVID-19 (PASC): An Overview of Biological Factors That May Contribute to Persistent Symptoms. *Frontiers in Microbiology* **12**, (2021).

44. Liu Y, Sawalha AH, Lu Q. COVID-19 and autoimmune diseases. *Curr Opin Rheumatol* **33**, 155-162 (2021).

45. Leng L*, et al.* Potential microenvironment of SARS-CoV-2 infection in airway epithelial cells revealed by Human Protein Atlas database analysis. *bioRxiv*, 2020.2004.2016.045799 (2020).

46. Mishra A, Chanchal S, Ashraf MZ. Host-Viral Interactions Revealed among Shared Transcriptomics Signatures of ARDS and Thrombosis: A Clue into COVID-19 Pathogenesis. *TH Open* **4**, e403-e412 (2020).

47. Root-Bernstein R. COVID-19 coagulopathies: Human blood proteins mimic SARS-CoV-2 virus, vaccine proteins and bacterial co-infections inducing autoimmunity. *BioEssays* **n/a**, 2100158.

48. Dotan A, Muller S, Kanduc D, David P, Halpert G, Shoenfeld Y. The SARS-CoV-2 as an instrumental trigger of autoimmunity. *Autoimmun Rev* **20**, 102792 (2021).

49. Köditz J*, et al.* Oxygen-dependent ATF-4 stability is mediated by the PHD3 oxygen sensor. *Blood* **110**, 3610-3617 (2007).

50. Liao Y, Fung TS, Huang M, Fang SG, Zhong Y, Liu DX. Upregulation of CHOP/GADD153 during coronavirus infectious bronchitis virus infection modulates apoptosis by restricting activation of the extracellular signal-regulated kinase pathway. *Journal of virology* **87**, 8124-8134 (2013).

51. Zendman AJ*, et al.* The human SPANX multigene family: genomic organization, alignment and expression in male germ cells and tumor cell lines. *Gene* **309**, 125-133 (2003).

52. Ngo HTT, Pham LV, Kim J-W, Lim Y-S, Hwang SB. Modulation of mitogen-activated protein kinase-activated protein kinase 3 by hepatitis C virus core protein. *Journal of virology* **87**, 5718-5731 (2013).

53. Khalili MA*, et al.* Male Fertility and the COVID-19 Pandemic: Systematic Review of the Literature. *World J Mens Health* **38**, 506-520 (2020).

54. Youssef K, Abdelhak K. Male genital damage in COVID-19 patients: Are available data relevant? *Asian J Urol*, (2020).

55. Peirouvi T*, et al.* COVID-19 disrupts the blood-testis barrier through the induction of inflammatory cytokines and disruption of junctional proteins. *Inflamm Res* **70**, 1165-1175 (2021).

56. Li H*, et al.* Impaired spermatogenesis in COVID-19 patients. *EClinicalMedicine* **28**, 100604 (2020).

57. Li Z*, et al.* Bioinformatic analyses hinted at augmented T helper 17 cell differentiation and cytokine response as the central mechanism of COVID-19–associated Guillain-Barré syndrome. *Cell Proliferation* **54**, e13024 (2021).

58. Ouyang Y*, et al.* Downregulated Gene Expression Spectrum and Immune Responses Changed During the Disease Progression in Patients With COVID-19. *Clinical Infectious Diseases* **71**, 2052-2060 (2020).

59. Zhang X-M, Liu C-Y, Shao Z-H. Advances in the role of helper T cells in autoimmune diseases. *Chin Med J (Engl)* **133**, 968-974 (2020).

60. Chiriva-Internati M, Cobos E, Da Silva DM, Kast WM. Sperm fibrous sheath proteins: a potential new class of target antigens for use in human therapeutic cancer vaccines. *Cancer Immun* **8**, 8 (2008).

733 **Table 1:** Summary metadata of COVID-19 case and control cohorts

734

Study Specification	Condition	Discovery Cohort (Disc)	Replication Cohort (Rep)
Cohort Size	Control	48	28
	COVID-19	49	48
	COVID-19 Follow-up	15	
Gender (Male, Female)	Control	46 M + 2 F	19 M + 9 F
	COVID-19	48 M + 1 F	40 M + 8 F
	COVID-19 Follow-up	15 M	
Age (IQR)	Control	38 (9)	66 (26)
	COVID-19	47 (20)	57 (22)
	COVID-19 Follow-up	49 (11)	
BMI (mean (IQR))	Control	25.3 (9.2)	25.9 (9.4)
	COVID-19	29.9 (6.2)	28.9 (7.3)
	COVID-19 Follow-up	28.7 (6.4)	
Sampling Time (media days (IQR))	COVID-19	5 (3) days after ICU admission	6 (6) days after ICU admission
	COVID-19 Follow-up	Six weeks after recovery	
Status Controls		Healthy	ICU Bacterial ARDS & Pneumonia, ICU H1N1 ARDS & Pneumonia
Status COVID-19		Severe	Severe, ARDS, Pneumonia
Hospital		ICU, HMC, Qatar	ICU, New York, USA
Ethnicity (%)		South Asian (69), Middle East and North Africa (MENA) (25). Other (6)	White (31), Asian (10), African (9), Other/Unspecified (50)
Matrix, Tubes, and Virus Inactivation		Serum, non-EDTA coated, viral inactivation using 10% Triton X100	Plasma, EDTA coated, viral inactivation using 10% Triton X100

735

Figure 1: Mapping of KREX Array proteins to KEGG categories (KEGG Pathway and KEGG Brite): Protein symbols and median fluorescent antibody signals (treemap cell size) are represented according their KEGG category assignment (www.kegg.jp; accessed on 14.Nov.2021). The other main categories are defined as cellular processes (top left - red), Human diseases (top middle - greyish purple), Organismal systems (to right - magenta), Genetic information processing (left - blue), Brite protein families (center - dark green), metabolism (right - orange), environmental information processing (bottom right - cyan). Unmapped proteins are considered as "Not included in Pathway or Brite" (bottom left - grey).

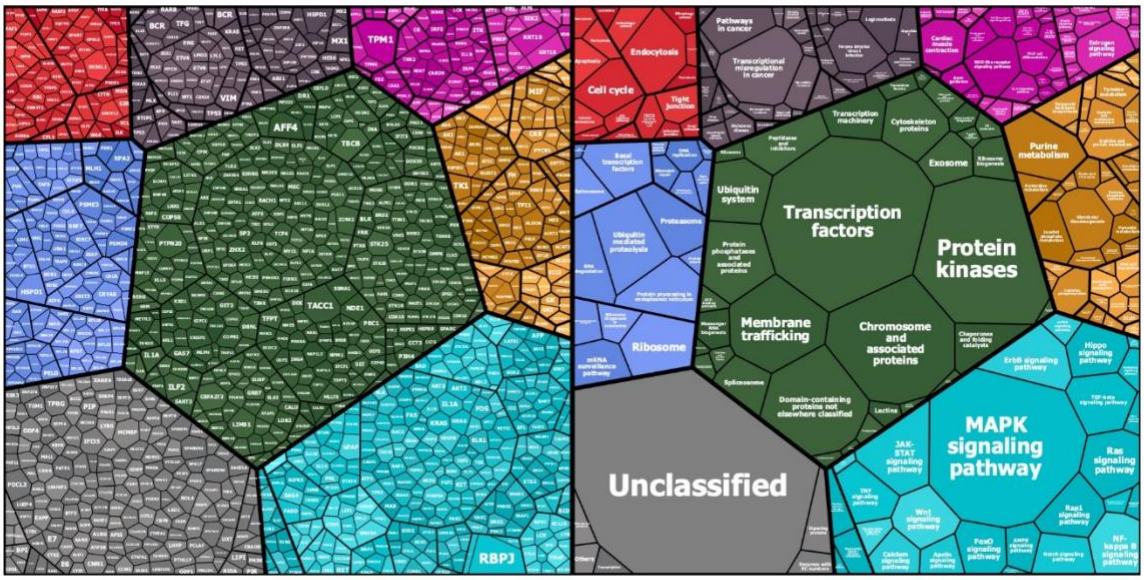
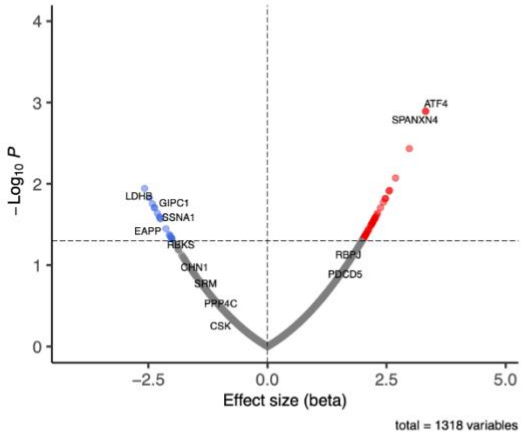


Figure 2: Differential protein autoantibody response analysis of COVID-19 Discovery cohort performed using T-test (A) and Fisher's exact test (B). A) Volcano graph of 1,318 proteins compares COVID-19 case (n = 49) vs. healthy controls (n = 48). Red dots represent proteins with an elevated autoantibody response, while blue dots represent proteins with a lower autoantibody response in COVID-19 patients. Proteins with Fisher's test p-value ≤ 0.05 are labelled in the volcano graph. B) Table on Fisher's exact statistics comparing subjects (numbers) of COVID-19 (n = 49) and the control (n = 48) groups for only thirteen proteins that showed significantly altered (p-value ≤ 0.05) autoantibody responses at sigma > 1.

Fig. 2

Discovery Cohort: Case *vs.* Control - 1318 Protein AB-titers

A **Volcano plot T-test vs. ES (beta)**

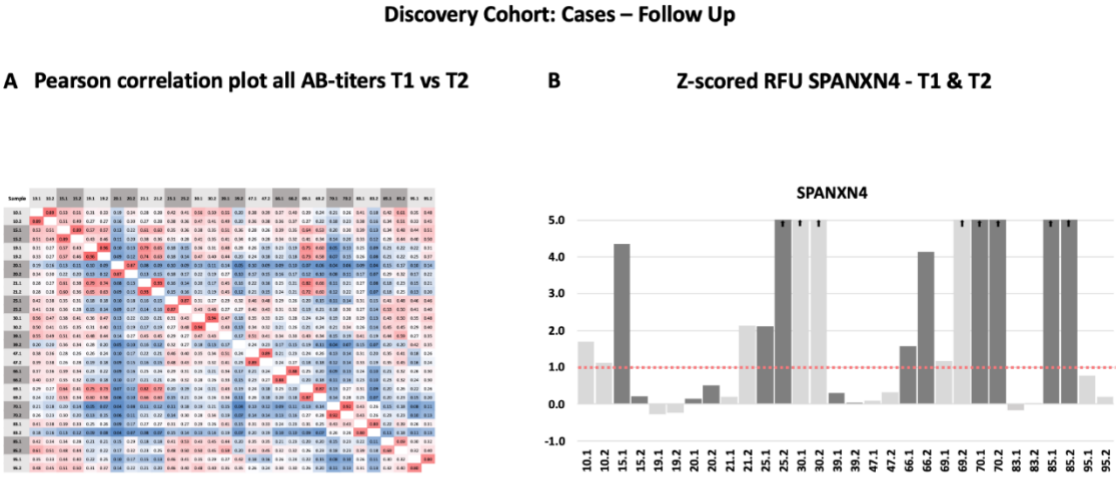


B **Fisher's exact test p < 0.05**

		Controls (N = 48)		COVID-19 (N = 49)	
Protein	Fisher p-value	Sigma ≤ 1	Sigma > 1	Sigma ≤ 1	Sigma > 1
SPANXN4	1.87E-05	43	5	24	25
ATF4	5.70E-03	48	0	41	8
GIPC1	1.24E-02	42	6	49	0
LDHB	1.24E-02	42	6	49	0
SSNA1	1.41E-02	21	27	34	15
RBPJ	1.41E-02	33	15	21	28
PPP4C	2.66E-02	43	5	49	0
SRM	2.66E-02	43	5	49	0
CHN1	2.66E-02	43	5	49	0
RBKS	3.07E-02	36	12	45	4
CSK	3.07E-02	41	7	48	1
EAPP	4.31E-02	30	18	40	9
PDCD5	4.76E-02	38	10	29	20

Figure 3: Autoantibody response in the Discovery cohort (T1 = sampling during ICU admission) and follow-up patients (T2 = sampling after recovery). A) Spearman rank correlation analysis of fifteen Discovery cohort samples collected at two different time points show strong correlation ($r \geq 0.69$) of autoantibody responses for the proteins. B) The histogram of Discovery cohort samples shows that the z-score RFU of SPANXN4 protein remains elevated in many patients even after COVID-19 recovery.

Fig. 3

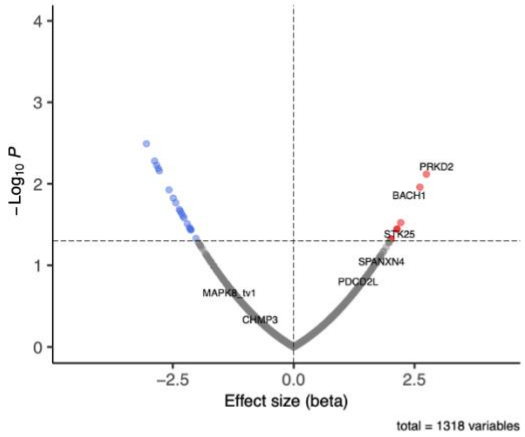


764 Figure 4: Differential protein autoantibody response analysis of COVID-19 Replication cohort
765 performed using T-test (A) and Fisher's exact test (B). A) Volcano graph of 1,318 proteins
766 compares COVID-19 cases (n = 48) vs. non-COVID-19 ICU controls (n = 28). Red dots represent
767 proteins with a high autoantibody response, while blue dots represent proteins with a low
768 autoantibody response in COVID-19 positive patients. Only proteins with Fisher's test p-value ≤
769 0.05 are labelled in the volcano graph. B) Table on Fisher's exact statistics comparing subjects
770 (numbers) of COVID-19 (n = 49) and the control (n = 48) groups for only thirteen proteins that
771 showed significantly altered (p-value ≤ 0.05) autoantibody responses at sigma > 1.

Fig. 4

Replication Cohort: Case *vs.* Control - 1318 Protein AB-titers

A **Volcano plot T-test vs. ES (beta)**



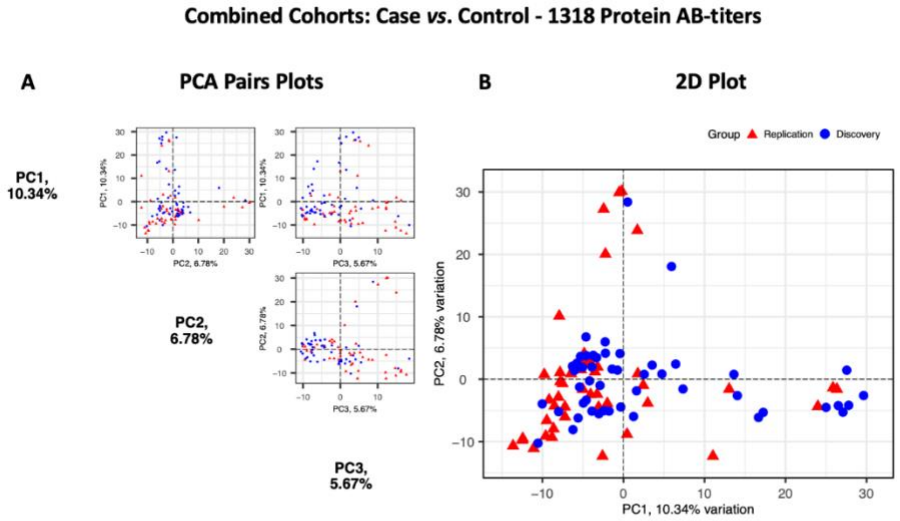
B **Fisher's exact test p < 0.05**

		Controls (N = 28)		COVID-19 (N = 48)	
Protein	Fisher p-value	Sigma ≤ 1	Sigma > 1	Sigma ≤ 1	Sigma > 1
SPANXN4	3.60E-03	27	1	33	15
PDCD2L	1.09E-02	28	0	38	10
PRKD2	1.29E-02	27	1	35	13
BACH1	2.32E-02	28	0	40	8
STK25	3.87E-02	26	2	34	14
CHMP3	4.66E-02	25	3	48	0
MAPK8_tv1	4.66E-02	25	3	48	0

772
773

774 Figure 5: Principal Components Analysis of the Discovery (n = 49, blue circles) and the
775 Replication cohorts (n = 48, red triangles). Each point represents a sample. A) PCA pair plot
776 compares PC1 to PC3. The proportion of variance explained in our cohorts by each PC is shown
777 in parentheses on the axis labels. B) PCA 2D plot with PC1 and PC2, which together describe
778 17.12% diversity between the cohorts.

Fig. 5



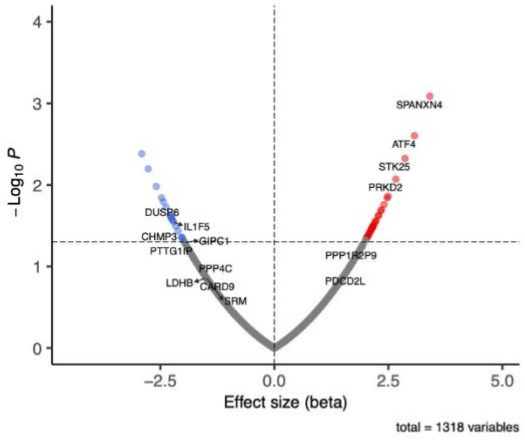
779
780

Figure 6: Differential protein autoantibody response analysis of Combined Cohorts performed using T-test (A) and Fisher's exact test (B). A) Volcano graph of 1,318 AB-protein titers compares COVID-19 cases (n = 97) vs. controls (n = 76). Red dots represent proteins with a high autoantibody response, while blue dots represent proteins with a low autoantibody response in COVID-19 positive patients. Only proteins with Fisher's test p-value ≤ 0.05 are labelled in the volcano graph. B) Table on Fisher's exact statistics comparing subjects (numbers) of COVID-19 (n = 49) and the control (n = 48) groups for only thirteen proteins that showed significantly altered (p-value ≤ 0.05) autoantibody responses at sigma > 1.

Fig. 6

Combined Cohorts: Case *vs.* Control - 1318 Protein AB-titers

A **Volcano plot T-test vs. ES (beta)**



B **Fisher's exact test p < 0.05**

		Controls (N = 76)		COVID-19 (N = 97)	
Protein	Fisher p-value	Sigma ≤ 1	Sigma > 1	Sigma ≤ 1	Sigma > 1
SPANXN4	6.10E-07	70	6	57	40
STK25	1.30E-03	74	2	79	18
ATF4	2.70E-03	76	0	87	10
PRKD2	6.40E-03	72	4	78	19
CHMP3	9.60E-03	65	11	94	3
PPP4C	1.52E-02	71	5	97	0
GIPC1	2.22E-02	69	7	96	1
CARD9	2.91E-02	38	38	65	32
PPP1R2P9	3.54E-02	76	0	91	6
DUSP6	3.56E-02	72	4	97	0
IL1F5	3.56E-02	72	4	97	0
PDCD2L	3.68E-02	73	3	84	13
PTTG1IP	4.39E-02	69	7	95	2
SRM	4.39E-02	69	7	95	2
LDHB	4.39E-02	69	7	95	2

Figure 7: Heatmap of autoantibody response pattern among different cohorts (A) and bar plots for pathways analysis. A) Heatmap of relative estimates (case vs. control) of autoantibody responses to fifty-six proteins in the Discovery, Replication, and Combined cohorts. Only proteins with significantly altered autoantibody responses were selected. Red color indicates higher and blue color indicates lower autoantibody responses against the proteins. B) KEGG and WIKI pathways analysis presented as bar-plot shows overactivated pathways in COVID-19 patients. Only pathways with T-test p-value ≤ 0.05 are presented in the bar-plot.

Fig. 7

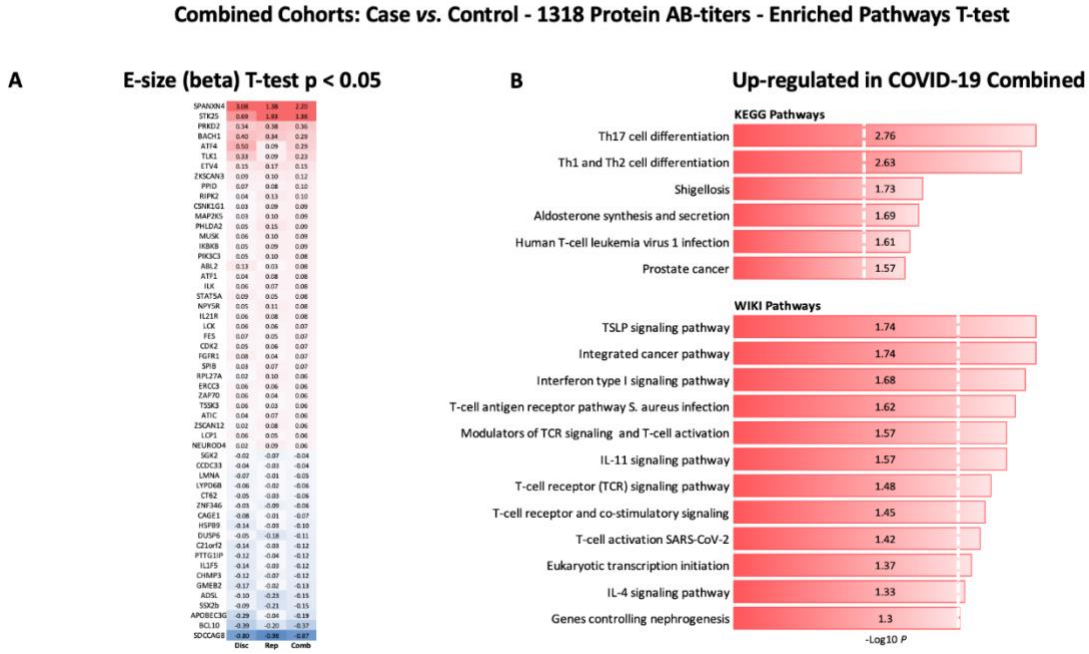
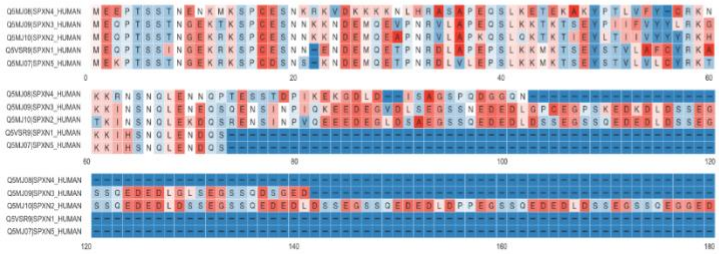


Figure 8: SPANXN4 protein sequence identity and antigen specificity analysis. A) Sequence alignment of proteins showing ≥ 50% identity with SPANXN4. B) Summary of properties of SPANXN4 protein from Human protein atlas. C) Autoantibody responses to SPANXN family proteins with high sequence identity to SPANXN4 that were part of the KREX immunome panel.

Fig. 8 **Overview SPANX4 Properties**

A BLAST Result vs. all human proteins [Identity > 50% Human]



B Human Protein Altlas Properties

Property	Description
Predicted location	plasma membrane & vesicles, nuclear membrane
Number of transcripts	2
RNA detected	testis, epididymis, glandular cells
RNA tissue specificity	testis enriched
Cell line specificity	HEL, HeLa, hTERT-HME1
Blood cell specificity	cell type enriched: neutrophil, total PBMC
Antibody specificity PrEsts	interaction only with its own antigen

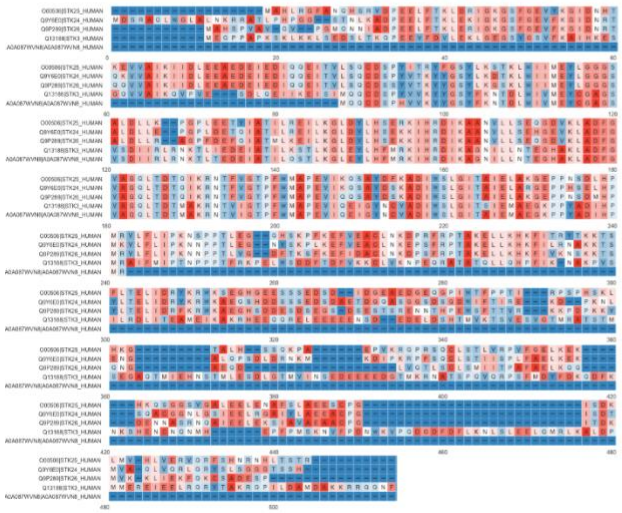
C Combined Cohorts: Case *vs*. Control - SPANX Family AB-titers

				Controls (N = 76)		COVID-19 (N = 97)	
Protein	T-test p-value	T-test Ratio	Fisher p-value	Sigma ≤ 1	Sigma > 1	Sigma ≤ 1	Sigma > 1
SPANXN4	8.17E-04	-2.20E+00	6.06E-07	70	6	57	40
SPANXN2	9.53E-01	2.13E-02	5.04E-01	52	24	71	26
SPANXN1	4.28E-01	-3.18E-01	1.00E+00	70	6	90	7

Figure 9: STK25 protein sequence identity and antigen specificity analysis. A) Sequence alignment of proteins showing ≥ 50% identity with STK25. B) Summary of properties of STK25 protein from Human Protein Atlas. C) Autoantibody responses to STK family proteins with high sequence identity to SKT25 that were part of the KREX immunome panel.

Fig. 9

A BLAST Result vs. all human proteins [Identity > 50% Human]



B Human Protein Altlas Properties

Property	Description
Predicted location	intracellular
Number of transcripts	19
RNA detected	all tissues
RNA tissue specificity	skeletal muscles
Cell line specificity	low cell line specificity
Blood cell specificity	detected in all without cell-type specificity

C Combined Cohorts: Case *vs.* Control - STK Family AB-titers

				Controls (N = 76)		COVID-19 (N = 97)	
Protein	T-test p-value	T-test Ratio	Fisher p-value	Sigma ≤ 1	Sigma > 1	Sigma ≤ 1	Sigma >
STK25	4.7E-03	1.36E+00	1.29E-03	74	2	79	18
STK10	8.2E-01	9.91E-03	5.05E-01	76	0	95	2
STK11	2.2E-01	2.83E-02	1.00E+00	76	0	97	0
STK16	3.3E-01	-2.66E-02	1.00E+00	76	0	97	0
STK17B	6.2E-01	1.78E-02	1.00E+00	76	0	96	1
STK24	9.8E-01	7.62E-04	1.00E+00	76	0	97	0
STK26	3.5E-01	4.61E-02	4.68E-01	74	2	91	6
STK3	9.3E-01	4.82E-03	1.00E+00	76	0	97	0
STK32A	7.4E-01	-1.32E-02	1.00E+00	76	0	97	0
STK32C	1.5E-01	-7.63E-02	1.00E+00	75	1	96	1
STK33	6.6E-01	-9.29E-02	1.00E+00	73	3	92	5
STK38	7.9E-01	9.64E-03	1.00E+00	76	0	97	0
STK38L	9.0E-01	4.48E-03	1.00E+00	76	0	97	0
STK4	7.8E-01	9.23E-03	1.00E+00	76	0	97	0