Journal of proteome research

pubs.acs.org/jpr

Article

# NMR Spectroscopic Windows on the Systemic Effects of SARS-CoV-2 Infection on Plasma Lipoproteins and Metabolites in Relation to Circulating Cytokines

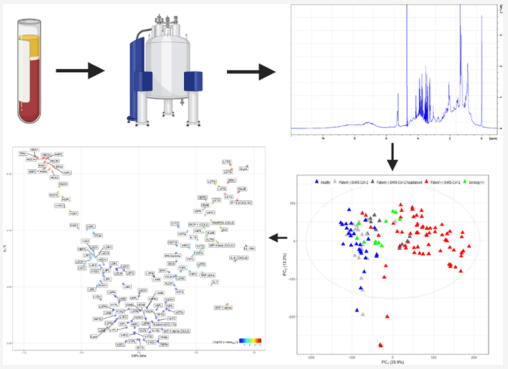
Samantha Lodge, Philipp Nitschke, Torben Kimhofer, Jerome D. Coudert, Sofina Begum, Sze-How Bong, Toby Richards, Dale Edgar, Edward Raby, Manfred Spraul, Hartmut Schaefer, John C. Lindon, Ruey Leng Loo, Elaine Holmes,* and Jeremy K. Nicholson*

Cite This: *J. Proteome Res.* 2021, 20, 1382−1396

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** To investigate the systemic metabolic effects of SARS-CoV-2 infection, we analyzed $^{1}$H NMR spectroscopic data on human blood plasma and co-modeled with multiple plasma cytokines and chemokines (measured in parallel). Thus, 600 MHz $^{1}$H solvent-suppressed single-pulse, spin-echo, and 2D J-resolved spectra were collected on plasma recorded from SARS-CoV-2 rRT-PCR-positive patients ($n$ = 15, with multiple sampling timepoints) and age-matched healthy controls ($n$ = 34, confirmed rRT-PCR negative), together with patients with COVID-19/influenza-like clinical symptoms who tested SARS-CoV-2 negative ($n$ = 35). We compared the single-pulse NMR spectral data with *in vitro* diagnostic research (IVDr) information on quantitative lipoprotein profiles (112 parameters) extracted from the raw 1D NMR data. All NMR methods gave highly significant discrimination of SARS-CoV-2 positive patients from controls and SARS-CoV-2 negative patients with individual NMR methods, giving different diagnostic information windows on disease-induced phenoconversion. Longitudinal trajectory analysis in selected patients indicated that metabolic recovery was incomplete in individuals without detectable virus in the recovery phase. We observed four plasma cytokine clusters that expressed complex differential statistical relationships with multiple lipoproteins and metabolites. These included the following: cluster 1, comprising MIP-1$\beta$, SDF-1$\alpha$, IL-22, and IL-1$\alpha$, which correlated with multiple increased LDL and VLDL subfractions; cluster 2, including IL-10 and IL-17A, which was only weakly linked to the lipoprotein profile; cluster 3, which included IL-8 and MCP-1 and were inversely correlated with multiple lipoproteins. IL-18, IL-6, and IFN-$\gamma$ together with IP-10 and RANTES exhibited strong positive correlations with LDL1−4 subfractions and negative correlations with multiple HDL subfractions. Collectively, these data show a distinct pattern indicative of a multilevel cellular immune response to SARS CoV-2 infection interacting with the plasma lipoproteome giving a strong and characteristic immunometabolic phenotype of the disease. We observed that some patients in the respiratory recovery phase and testing virus-free were still metabolically highly abnormal, which indicates a new role for these technologies in assessing full systemic recovery.



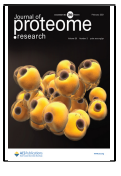
**KEYWORDS:** *plasma, COVID-19, SARS-CoV-2, NMR spectroscopy, single-pulse, spin-echo, IVDr, metabolic phenotyping, biomarkers, diagnostic modeling, lipoproteins*



## INTRODUCTION

COVID-19 is a complex respiratory and systemic disease caused by infection with the SARS-CoV-2 coronavirus. The virus, both directly and indirectly, causes systemic damage to multiple organs resulting in several new-onset pathologies including diabetes,[1] liver dysfunction,[2] neurological damage, and renal damage,[3,4] as well as possible pulmonary[5] and cardiovascular damage.[6] We have shown that SARS-CoV-2 infection induces a metabolic "phenoconversion" process that gives a characteristic complex blood plasma signature exhibiting multiple pathological effects of the disease and in a complex mosaic of biochemical change.[7] This COVID-19 phenoconversion signature is readily detectable by both nuclear magnetic resonance (NMR) and mass spectrometric techniques, with the spectral profiles providing multiple windows onto the disease sequelae in different organ systems.[7] The general term, phenoconversion, is defined here as follows:

**Received:** November 2, 2020
**Published:** January 11, 2021



"The systemic shift in multiple molecular phenotypic properties from the healthy to a characteristic activated or pathophysiological state in response to an infectious or a pathophysiological challenge". The term has previously been used to describe accurately, but more narrowly, the process of cellular metabolic activation by drugs to alter the biochemical response to subsequent drug treatments *via* enzyme induction.[8] The broader definition applied here extends the concept to cover disease processes and the resultant systemic metabolic state changes. This is particularly appropriate for COVID-19 that expresses itself in multiple clinical subphenotypes involving several major organ systems. Thus, the exact pattern of phenoconversion may give insights into the underlying pathophysiological processes and their individual variations.

We have suggested previously that the use of orthogonal phenoconversion tests for COVID-19 detection might be considered to augment polymerase chain reaction (PCR)-based diagnostics based on systemic disease effects.[7] Such diagnostics can also be extended to clinical trial monitoring[9,10] and long-term disease recovery studies. There are numerous possible spectroscopic approaches to such augmented COVID-19 diagnostics including NMR spectroscopy[7,11] and mass spectrometry,[12,13] and all these carry different and complementary information sets describing a particular condition. Multiple studies have shown the potential of metabolic phenotyping with respect to the diagnosis of SARS-CoV-2 infection, and a comparison with healthy controls has indicated a distinctive pattern of underlying pathologies.[7,14] Mass spectrometry has been used to identify characteristic patterns of amino acids, biogenic amines, and lipids,[7] whereas $^1$H NMR spectroscopic studies have identified perturbations in multiple lipoprotein subfraction patterns and an elevation in $\alpha$-1-acid glycoprotein [observed through characteristic *N*-acetyl proton signals from the oligosaccharide side-chains designated glycoprotein A (GlycA) and glycoprotein B (GlycB)], an acute phase reactive protein that is elevated in systemic inflammation.[7,15]

High-resolution $^1$H NMR spectroscopy of blood plasma and other biofluids has been shown to have a wide variety of diagnostic and biochemical applications.[16−18] A number of 1D and 2D experiments can be used to probe plasma biochemistry,[19] and recently, *in vitro* diagnostic research (IVDr) methods have been developed to extract lipoprotein information from plasma samples[20−22]. Information recovery has been shown to be reliable with respect to sample collection and storage, and spectral data are highly reproducible irrespective of sample handling, collection tube type, or freeze−thaw cycles.[23] However, sample preparation methods are key to obtaining reliable results, and prior heat treatment for virus inactivation is highly disruptive of NMR-based COVID-19 diagnostic biomarker information,[23] causing significant metabolic and lipoprotein disruption.

We have shown that a combination of IVDr-based lipoprotein measurements with mass spectrometry-based amino acid and biogenic amine analysis can discriminate between SARS-CoV-2 positive patients and healthy controls showing a distinctive pattern of underlying pathologies.[7] Here, we explore the potential of different $^1$H NMR methods to classify SARS-CoV-2 positive samples in relation to both healthy individuals and those with respiratory symptoms that were serologically confirmed to be SARS-CoV-2 negative. Three of the main NMR methods used in *in vitro* diagnostic suites are evaluated here, exploiting the relative strengths of each experiment and comparing their diagnostic performance and information recovery. There has been much speculation that acute acceleration of COVID-19-related morbidity is due to accelerated immunological reactions, and that the so-called "cytokine storm" events are significant contributors to mortality. Thus, we have also examined the relationships between these NMR data sets and a set of 34 circulating cytokines measured on plasma samples from the same patients taken at the same time in order to evaluate the possible metabolic connections with systemic immunological events that might relate to the underpinning pathological mechanisms.

## MATERIALS AND METHODS

### Patient Enrollment and Sample Collection

Blood plasma samples were collected from a cohort of adult individuals in a study initiated at Fiona Stanley Hospital in the Western Australian South Metropolitan Health Service catchment as part of the International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC)/World Health Organization (WHO) pandemic trial framework (SMHS Research Governance Office PRN:3976 and Murdoch University Ethics no. 2020/052). Healthy control participants were enrolled as volunteers, provided with study details, and written consent was obtained prior to data collection in accordance with ethical governance (Murdoch University Ethics no. 2020/053). Three groups of participants were recruited from the Fiona Stanley and Royal Perth Hospitals: (i) patients who presented COVID-19 disease symptoms and subsequently tested positive for SARS-CoV-2 infection from upper and/or lower respiratory tract swabs by reverse transcription PCR (rRT-PCR; $n$ = 15 participants sampled at various times resulting in $n$ = 68 specimens); (ii) healthy controls who had not exhibited COVID-19 disease symptoms ($n$ = 34 participants with 39 specimens); and (iii) patients with COVID-19 disease symptoms and who tested negative ($n$ = 35 participants with $n$ = 35 specimens). The SARS-CoV-2 negative participants were further classified as those who required hospitalization ($n$ = 11) versus those who were recruited from a COVID clinic but did not require hospitalization. The demographic data together with the clinical symptoms are shown in Supporting Information Tables S1 and S2.

As SARS-CoV2 has been reported to induce lingering metabolic effects, in order to ensure that the control group used to train the diagnostic model contained true healthy controls, rather than individuals that may have been exposed to the virus but were asymptomatic, additional serology testing was performed for a subset of 44 individuals using a commercial point-of-care serological COVID-19 IgA/IgG test. Serology testing was conducted at the by PathWest Laboratory Medicine WA, using 10 $\mu$L of plasma sample for 36 specimens (25 specimens from eight SARS-CoV-2 positive patients, 3 specimens from 3 SARS-CoV-2 negative patients, and 8 specimens from 7 healthy controls). Samples were considered as SARS-CoV-2 positive if IgA > 1.0 or equivocal where IgA = 0.8−1.0. IgG levels are reported in Supporting Information Table S3). Plasma samples were stored at −80 °C. Sample processing was performed according to Bruker IVDr protocols for the small molecule and lipoprotein data[24] and according to the protocols for multiplexed Human Procarta-


https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

Plex Panel 1A (Life Technologies, USA) for the cytokine and chemokine analyses.

### $^1$H NMR Spectroscopy Data Acquisition and Processing Parameters

Plasma samples were thawed at 20 °C for 30 min and then centrifuged for 10 min at 13,000*g* at 4 °C. Plasma samples were prepared in 5 mm outer diameter SampleJet NMR tubes, following the recommended procedures for *in vitro* analytical and diagnostic procedures[22] using 300 μL of plasma mixed with 300 μL of phosphate buffer (75 mM $Na_2HPO_4$, 2 mM $NaN_3$, and 4.6 mM sodium trimethylsilyl propionate-[2,2,3,3-$2H_4$] (TSP) in $D_2O$, pH 7.4 ± 0.1). NMR spectroscopic analyses were performed on a 600 MHz Bruker Avance III HD spectrometer equipped with a 5mm BBI probe and fitted with the Bruker SampleJetTM robot cooling system set to 5 °C. A full quantitative calibration was completed prior to the analysis using a previously described protocol.[24] All experiments were completed using the Bruker *In Vitro* Diagnostics research (IVDr) methods.[21] For each sample, three experiments were completed in automation mode, amounting to a total of 12.5 min of acquisition time per sample: a standard 1D experiment with solvent presaturation (32 scans, 96 K data points, and a spectral width of 30 ppm), a Carr−Purcell−Meiboom−Gill (CPMG) spin-echo experiment (32 scans, 72 K data points, and a spectral width of 20 ppm) which filters the spectrum by differential T2 relaxation removing the peaks from the large molecules, and a 2D J-resolved experiment (40 $t_1$ increments with 2 scans each). Data were processed in automation using Bruker TopspinTM 3.6.2 and ICONTM NMR to achieve phasing and baseline correction. In addition to the information extracted from the standard 1D and CPMG spectra regarding the low molecular weight plasma components, a total of 112 lipoprotein parameters for each sample were generated using the Bruker IVDr Lipoprotein Subclass Analysis (B.I.-LISA) method, whereby the $-(CH_2)_n$ at $\delta$ = 1.25 and $-CH_3$ at $\delta$ = 0.80 peaks of the 1D spectrum after normalization to the Bruker QuantRef manager within Topspin were quantified using a PLS-2 regression model.[22] B.I.-LISA data consist of total plasma lipid analyte cholesterol, free cholesterol, phospholipids, triglycerides, apolipoproteins A1/A2/B100 and the B100/A1 ratio, and analyte distributions in different density classes of plasma-lipoproteins: high-density lipoprotein (HDL, density 1.063−1.210 kg/L), intermediate-density lipoprotein (density 1.006−1.019 kg/L), low-density lipoprotein (LDL, density 1.09−1.63 kg/L), and very low-density lipoprotein (VLDL, 0.950−1.006 kg/L). The main lipoprotein classes HDL, LDL, and VLDL were subdivided into different density subclasses (LDL-1: 1.019−1.031 kg/L, LDL-2: 1.031−1.034 kg/L, LDL-3: 1.034−1.037 kg/L, LDL-4: 1.037−1.040 kg/L, LDL-5: 1.040−1.044 kg/L, and LDL-6: 1.044−1.063 kg/L), HDL subfractions into four density classes (HDL-1 1.063−1.100 kg/L, HDL-2 1.100−1.125 kg/L, HDL-3 1.125−1.175 kg/L, and HDL-4 1.175−1.210 kg/L), and VLDL subfractions into five density classes. A list of all the 112 lipoprotein parameters are shown in Supporting Information Table S4. In addition to the 112 lipoprotein parameter, α-1-acid glycoprotein (*N*-acetyl-glucosamino *N*-acetyl) signal integrals were calculated as GlycA from the superimposition of terminal *N*-acetyl signals ($\delta$ 2.03[a]) and GlycB calculated from the branched chain *N*-acetyl signals ($\delta$ 2.07) from the CPMG experiment.

### NMR Data Modeling

Both the standard 1D and CPMG NMR spectral datasets were calibrated to the α-anomeric proton signal of glucose at $\delta$ 5.23. Each spectrum was baseline-corrected using an asymmetric least squares routine; spectral regions corresponding to the residual water resonance signal ($\delta$ 4.60−4.85) or predominantly noise ($\delta$ < 0.5 and $\delta$ > 9.5) were excluded from analyses. Spectra were normalized *via* a probabilistic quotient method using the median spectrum as reference.[25] Data were mean-centered and scaled to unit-variance prior to multivariate modeling. Principal component analysis (PCA) was used to assess the main sources of structured variation within each dataset (lipoprotein classes and subclasses, 1D and CPMG spectra). In order to identify the lipoproteins and metabolites differentiating SARS-CoV-2 positive participants from healthy controls, an orthogonal projection to latent structures discriminant analysis (OPLS-DA) model[26] was constructed using a training sample set comprising a single timepoint from PCR-confirmed SARS-CoV-2 patients ($n$ = 7) and age- and sex-matched healthy control subjects ($n$ = 7) confirmed to be serologically negative based on IgA and Supporting Information Table S3. The remaining dataset ($n$ = 61 SARS-CoV-2 positive, $n$ = 35 SARS-CoV-2 negative, $n$ = 28 healthy control samples, and $n$ = 7 control samples, subsequently identified as serology positive) was used as a test set and projected into the OPLS-DA training set model. This was performed for each spectral dataset (1D, CPMG, and quantitative lipoprotein) separately and the optimal number of orthogonal components was determined using the area under the receiver operator characteristic curve (AUROC) calculated from predictive component scores, generated using an internal leave-one-out cross-validation procedure. The SARS-CoV-2 negative samples were color-coded according to whether the samples were obtained from patients hospitalized for respiratory conditions (dark gray, $n$ = 11) versus those reporting to a "walk-in" COVID testing clinic (light gray, $n$ = 28).

### Cytokine and Chemokine Analyses

Where sample volume was sufficient ($n$ = 23 SARS-CoV-2 positive samples and $n$ = 15 healthy) xMAP cytokine assays were carried out in 96-well polystyrene microplates using a MagPix detection system (Luminex, USA). A total of 34 cytokines and chemokines were quantified using a multiplexed Human ProcartaPlex Panel 1A (Life Technologies, USA) according to manufacturer's instructions. Briefly, capture bead mixes were incubated overnight with 25 μL of serum sample at 4 °C and then washed three times in the wash buffer. Samples and standards were then incubated with detection antibodies for 30 min at room temperature, followed by three washes with the wash buffer; all washes were performed using a Bio-Plex Pro Wash station (Bio-Rad Laboratories). Streptavidin-PE was then added and the plate incubated for 30 min at room temperature (20 °C). Plates were subsequently washed three times before the reading buffer was added. All buffers and reagents were supplied with the Luminex detection kit (Life Technologies, USA). The plate was incubated for further 5 min at room temperature before analysis. All incubation steps were carried out using an orbital shaker at 500× rpm. Data were analyzed using Luminex xPonent software.

Cliff's delta statistic, a nonparametric effect size measure that quantifies the group differences of a variable, was calculated for both the quantitative lipoprotein data and the log-transformed cytokine data, separately and in combination, in order to assess


https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

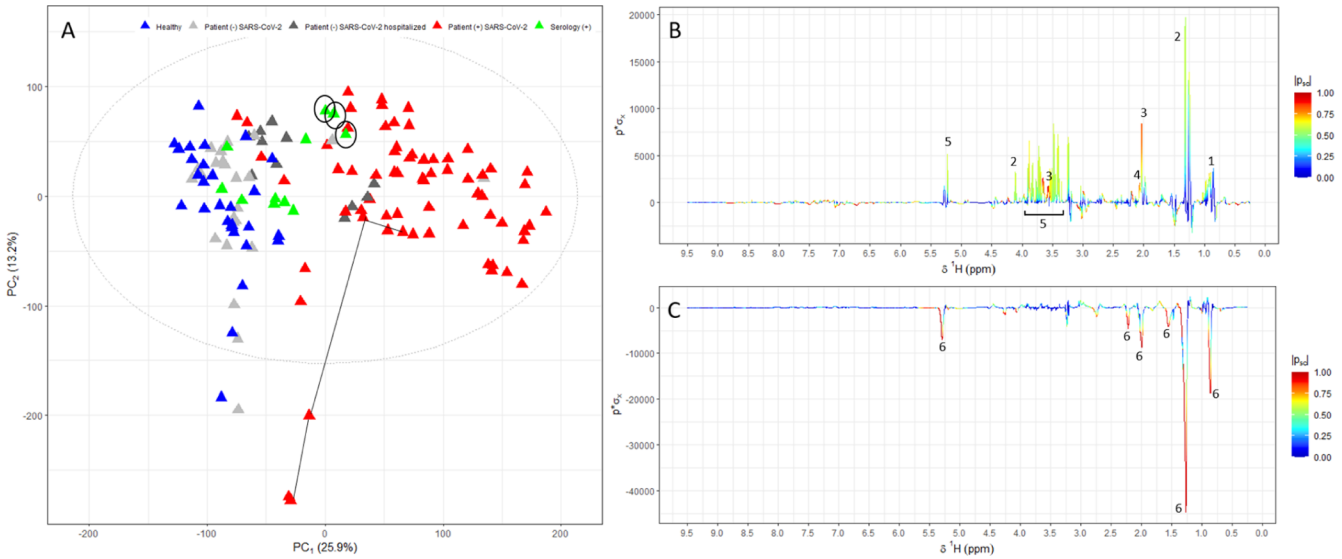



**Figure 1.** PCA of 1D 600 MHz $^{1}$H NMR spectral data. (A) PCA score plot showing clustering patterns for the healthy controls (blue), SARS-CoV-2 positive (red), nonhospitalized SARS-CoV-2 negative (light gray), hospitalized SARS-CoV-2 negative (dark gray) samples, and control participants who tested positive for IgA (green). The ellipse indicates Hotelling's T2 statistic ($\alpha$ = 0.95), which can be interpreted as a multivariate confidence interval; (B) PCA loading plots for PC 1 and (C) PC2. Key: 1 = valine; 2 = lactate; 3 = GlycA; 4 = GlycB; 5 = glucose; and 6 = triglycerides.

group differences. An absolute Cliff's delta of 1 denotes the maximum difference, while a value of zero denotes no differences between two groups. The arithmetic sign indicates the elevation or decrease in reference to a reference group (in this case, the control group). To test for statistical significance, a two-tailed paired Wilcoxon rank sum was applied. Correction for multiple hypothesis testing was not appropriate as many of the measured lipoprotein parameters are interdependent. Here, a statistical significance level was fixed at $\alpha$ = 0.05. All computation and data visualization were performed using R and RStudio IDE with the open-source R package metabom8 (version 0.2), available from GitHub (github.com/tkimhofer/metabom8).

## RESULTS AND DISCUSSION

### SARS-CoV-2 Positive Metabotypes are Distinct from Both Healthy Controls and SARS-CoV-2 Negative Respiratory Patients

The plasma from SARS-CoV-2 positive participants was clearly biochemically distinct from that of either the healthy controls or the SARS-CoV-2 negative participants based on information drawn from all the spectral datasets (1D, CPMG, and lipoproteins). Considering the PCA model of the 1D spectral datasets (Figure 1A), the healthy controls clustered on the left side of the Score Plot (plotted in blue) with the SARS-CoV-2 positive samples (plotted in red) clustering to the right side of the diagram. Patients that tested negative for SARS-CoV-2 were mainly nonhospitalized (plotted in light gray) and tend to cluster with healthy controls, while patients hospitalized for other respiratory conditions (plotted in dark gray) showed a greater overlap with those of SARS-CoV-2 positive patients. A similar pattern was evident in the PCA score plots generated from the CPMG data (Figure S1). Patient data and other metadata are provided in Supporting Information Tables S1 and S2.

Antibody testing for IgA and IgG levels was performed on a subset of samples in order to ensure that selected spectra used for training the OPLS-DA model were correctly classified. It is of note that 7 out of 34 participants originally recruited as healthy controls (denoted by green symbols in the PCA score plot; Figure 1) were found to have high IgA and/or IgG levels, suggesting that either these individuals had previously been unknowingly infected but were asymptomatic/mildly affected or that there may have been some cross-reactivity with other corona viruses.[27] These individuals map in-between the SARS-CoV-2 positive and healthy controls in the Principal Components (PCs) space, with one individual mapping within the SARS-CoV-2 positive cluster (identified by the circled co-ordinates in Figure 1A), with an IgA value of 2.9 (positive) and an IgG value of 0.8 (equivocal) mapping within the SARS-CoV-2 positive cluster (Figure 1). Of the seven individuals in the seropositive cluster, all but one had reported mild to severe cold or influenza-like symptoms ranging from fever and nasal decongestion to extreme fatigue in the 2 months prior to sampling but were not tested at the time. Further, five of the seven had traveled to countries with high levels of SARS-CoV-2 infection within the 2 months prior to testing. The fact that the seropositive but not PCR-confirmed samples lie at the boundary between the SARS-CoV-2 positive and healthy groups indicates that there are some metabolic similarities with the SARS-CoV-2 positive group and may represent participants who had been exposed to the virus.

PC1 loadings (Figure 1B), which define the direction of metabolic separation of SARS-CoV-2 and healthy control classes, were dominated by elevated signals from glucose and $\alpha$-1-acid glycoprotein[15,28] consistent with previous observations of diabetes and acute inflammation associated with COVID-19 disease.[28] The NMR signals for GlycA at $\delta$ 2.03 and GlycB at $\delta$ 2.07 were integrated and significant intensity differences were found between groups, with SARS-CoV-2 positive samples containing the highest median concentrations of both GlycA and GlycB, followed by SARS-CoV-2 negative samples (Figure S2). GlycA has been shown to be a robust marker of inflammatory and proinflammatory conditions including inflammatory bowel disease,[29] gestational diabetes,[30]


https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

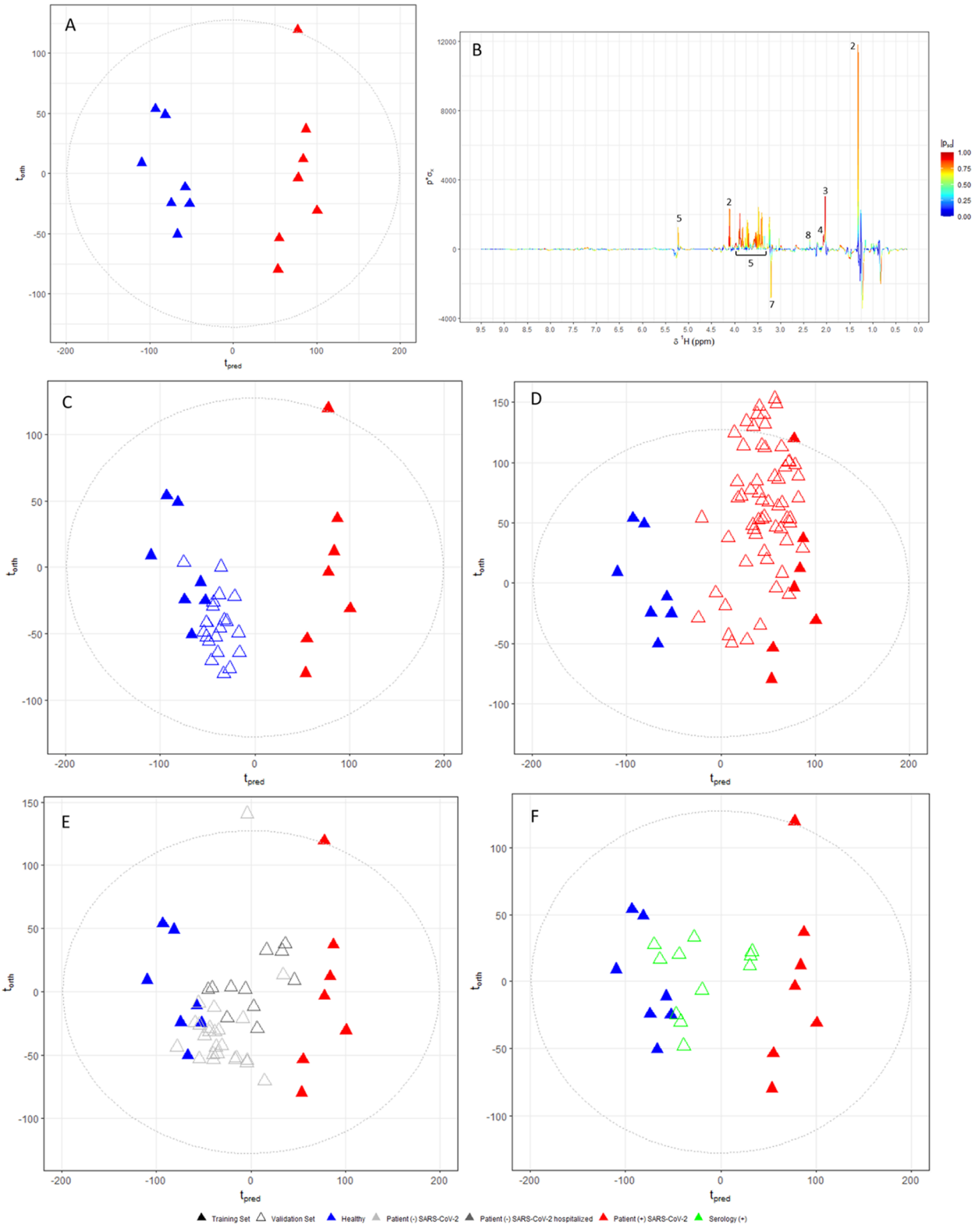



**Figure 2.** OPLS-DA for the training (closed triangle) and test sets (open triangle) of the 1D spectral dataset. (A) Score plot for OPLS-DA of the training set using the 1D spectra of the healthy controls and COVID-19 positive ($n$ = 7 per group) patients. (B) Coefficient plot of the OPLS-DA model showing that the SARS-CoV-2 patients were dominated by signals from GlycA, GlycB, glucose, and lactate; whereas the control group was driven by higher concentrations of a phosphocholine molecule. Projection of test set into the OPLS-DA training set model for (C) healthy individuals (blue open triangle); (D) SARS-CoV-2 positive patients (red open triangle); and (E) SARS-CoV-2 negative patients for hospitalized (dark gray open triangle), nonhospitalized patients (light gray open triangle), and (F) controls who tested positive according to the IgA serology results but who had not been formally diagnosed as having SARS-CoV-2. Key: 2 = lactate; 3 = GlycA; 4 = GlycB; 5 = glucose; 6 = triglycerides; 7 = lysophosphatidylcholine; and 8 = pyruvate.



https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

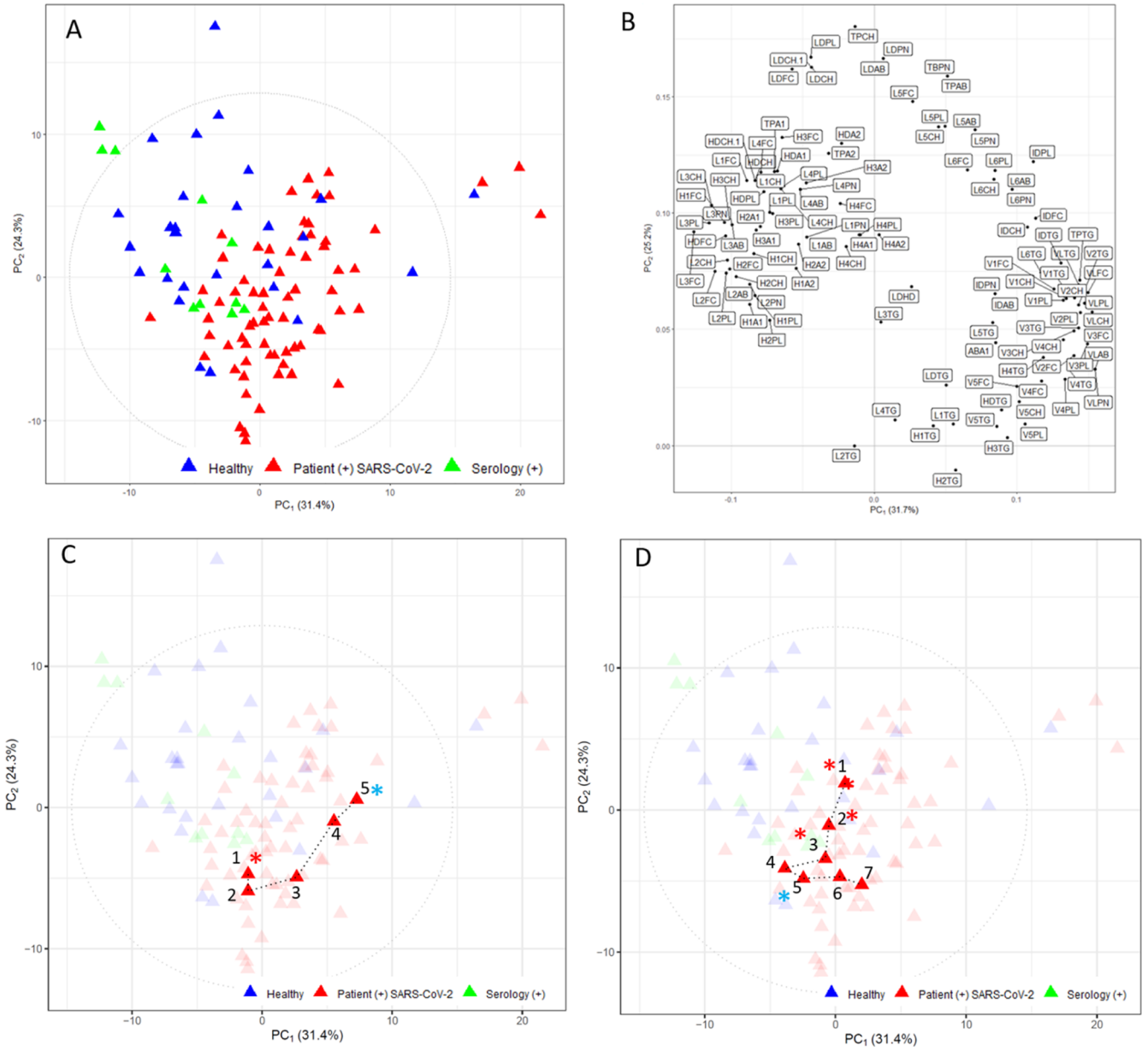



**Figure 3.** PCA of lipoprotein parameters. (A) PCA of lipoprotein parameters of the healthy (blue); individuals from the healthy cohort with a seropositive IgA result (green) and SARS-CoV-2 positive patients (red); (B) PCA lipoprotein loading plot; PCA of the lipoprotein parameters with an individual SARS-CoV-2 positive patient trajectory (C) showing change in the metabolic lipoprotein profile over five collection timepoints; and (D) another individual with seven collection timepoints. Coordinates marked with red * represent timepoints where the participant tested PCR positive and cyan * denotes a negative PCR result.

alcoholic liver disease,[31] cardiovascular risk,[32] rheumatoid arthritis[33] and certain cancers,[28] although GlycA is confounded by hemolysis and is not an appropriate marker of inflammatory conditions such as sickle cell disease.[34] In general, GlycA has been found to be a better predictor of inflammatory conditions than either C reactive protein (CRP) or interleukin-6 (IL-6), and is independent of both, although GlycA and CRP have been found to be strongly correlated in many studies.[28] Kettunen *et al.* reported GlycA to be predictive of mortality risk in angiography patients[31] and has also been shown to be an early predictor of subclinical cardiovascular disease and treatment response in the early stage of rheumatoid arthritis[33] and in the development of gestational diabetes.[30]

The signal defining GlycA at $\delta$ 2.03 is mainly attributed to $\alpha$-1-acid glycoprotein but represents a conglomerate of five circulating glycoproteins, with alpha-1 antitrypsin, haptoglobin, transferrin, and $\alpha$-1-antichymotrypsin additionally contributing to the signal intensity.[35] Ritchie *et al.* calculated the concentrations of each of the five components and assessed their correlation with the GlycA signal and with mortality risks and found that while $\alpha$-1-acid glycoprotein demonstrated the strongest correlation with GlycA, $\alpha$-1 antitrypsin levels were the most predictive of morbidity and mortality for a range of diseases including heart failure, influenza, and liver diseases. GlycA has also been associated with microbial diversity, underscoring the relationship between the gut microbiome and inflammation.[36] While there are fewer studies relating to GlycB, associations between GlycB and inflammatory conditions have also been reported.[33]


https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

The second PC loading vector (Figure 1C) is dominated by triglyceride signals consistent with obesity in several of the participants shown in Figure 1A, with all of the samples mapping outside the Hotelling's T ellipse derived from overweight or obese individuals. A PLS regression of the $^1$H NMR spectral data against the body mass index (BMI) for all the SARS-CoV-2 positive participants and all the healthy control participants yielded similar loadings to those identified in the PCA model of the dataset (Figure S3), indicating that BMI was the major driver for the variance observed in the second PC. This observation is consistent with other studies showing the association between BMI and high levels of serum triglycerides.[37] The second PC was also weighted by one SARS-CoV-2 positive patient, represented by several samples obtained at different times postinfection, with a pre-existing type 2 diabetes diagnosis and a BMI of 37.4 (indicated by the coordinates connected by a gray line in Figure 1A).

In order to extract a clearer signal of SARS-CoV-2 infection that was less confounded by BMI and other factors such as comorbidities, OPLS-DA was applied to the 1D $^1$H NMR spectral data and yielded a strong model (AUROC = 1.0, based on one predictive and one orthogonal component) separating the control and SARS-CoV-2 positive patients. The scores, together with the corresponding loading plots for the training set, based on seven healthy control and seven SARS-CoV-2 patients, are shown in (Figure 2A,B). The test set projections of the healthy control, SARS-CoV-2 positive patients, and SARS-CoV-2 negative patients into the OPLS-DA model show good prediction of healthy control participants (Figure 2C) and SARS-CoV-2 positives (Figure 2D), whereas the SARS-CoV-2 negative patients cluster in-between the healthy and SARS-CoV-2 positive groups but map more closely to the healthy group (Figure 2E). Samples from individuals originally included in the healthy control set but who were seropositive for IgA (Figure 2F; green triangles) were projected between the healthy and SARS-CoV-2 positive groups with three samples from the same individual identified in Figure 1 mapping closer to the positive group.

PCA and OPLS-DA analyses were also performed using CPMG spin-echo NMR data (Figures S1 and S4) in addition to the 1D spectral dataset. In addition to elevated $\alpha$-1-acid glycoprotein signals (GlycA and GlycB), key discriminatory features of SARS-CoV-2 positive patients were detected in the CPMG dataset, and these included higher concentrations of alanine, lactate, and pyruvate compared to healthy controls, which were highly statistically correlated indicative of a shift in the energy metabolism. The lactate/pyruvate ratio was significantly greater ($p < 0.0002$) for the SARS-CoV-2 positive group (28.20 ± 11.95) compared to that of the healthy group (21.31 ± 4.57) and has been shown to be a good test of impaired mitochondrial respiration as it indirectly reflects the NADH:NAD+ redox state in the cytoplasm[38] and is high in poorly ventilated patients. The discriminatory powers of CPMG spin-echo spectra for identifying SARS-CoV-2 positive individuals which edits out most macromolecular signals except from those from GlycA and GlycB acetyl signals from *N*-glycans ($\delta$ 2.03) and the attached oligosaccharide backbone sugar ring protons. Even though these moieties are carried on a macromolecule, there is significant segmental motion in the oligomeric chain which means that the $T_2$ relaxation times are relatively long and so are not attenuated.[39,40]

The loadings for the aromatic region were difficult to interpret due to the partial overlap of signals from various metabolites. However, as alluded to in previous publications, a general disturbance in amino acid metabolism was manifested in the SARS-CoV-2 positive group with increased phenylalanine [$\delta$ 7.33 (doublet); ($\delta$ 7.43 (multiplet)] being in the SARS-CoV-2 positive group in contrast to 1-methylhistidine [$\delta$ 7.05 (singlet); $\delta$ 7.77 (singlet)], which was decreased in comparison with the healthy control group (Figure S4). NMR signals from the ether glucuronide conjugate of the widely used analgesic drug acetaminophen (paracetamol, Tylenol) were also noted in some of the SARS-CoV-2 positive patient group (8 out of 15), consistent with disease-related analgesia, and with a third of SARS-CoV-2 negative group (12 out of 39) showing acetaminophen signals but none in the healthy control ($n$ = 28) as shown in the comparison of an exemplar 2-D J-resolved of a healthy control sample (blue) overlaid with a SARS-CoV-2 sample (Figure S5). Detailed NMR signal analysis of acetaminophen metabolites has been reported previously[41] and their presence in normal human populations from widespread over-the-counter self-administration.[42]

## Differential Effects of SARS-CoV-2 on Lipoprotein Profiles

The standard Bruker IVDr quantification was applied to the 1D NMR data in order to extract the impact of SARS-CoV-2 infection on the lipoprotein profile. The PCA score plot (Figure 3A) of the 112 lipoprotein concentrations and their ratios, together with the corresponding loadings (Figure 3B), shows a clear differentiation between SARS-CoV-2 positive patients and healthy controls. The quantitative lipoprotein model produced a slightly inferior classification of SARS-CoV-2 positivity to the simple 1D NMR spectra shown previously, which is consistent with the loss of the extra discriminatory power conveyed by the GlycA and GlycB signals that carry extra information on the inflammatory processes. Nonetheless, most of the samples from healthy participants (blue), including those from the healthy group who had tested positive for IgA (green), are clustered on the left hand side of the plot with SARS-CoV-2 positive patients tending to cluster to the right hand side. Within the SARS-CoV-2 positive group, several patients were represented by multiple samples, which showed a metabolic trajectory, or progression of infection, within the SARS-CoV-2 cluster, indicating that the metabolic effects of the virus were relatively stable through the time in which they were studied. Such metabolic time trajectories have previously been demonstrated in a variety of experimental toxicity states and provide useful insights into the progression of disease[43−45] showing onset, progression, and resolution of disease.

The individual patient journey trajectories are illustrated for two patients where longitudinal data were available (Figure 3C,D). For both of these patient trajectories, the biochemical profiles mapped within the SARS-CoV-2 positive space, indicating that there was no biochemical recovery toward "normal" during that time period even though the nasal swab confirmed the absence of virus by the end of their hospital stay. This raises an interesting and important clinical point in that "recovery" for COVID-19 is normally assessed on respiratory symptoms. However, a patient cannot be said to be truly recovered unless they are also biochemically normal and free of some of the more complex, but less visually obvious systemic effects of the disease. For example, the trajectory depicted in Figure 3C shows a patient who was diagnosed as SARS-CoV-2 positive at timepoints 1, 2, and 3 (indicated by a red*) and confirmed as negative by timepoint 5 (cyan*). Similarly, the second trajectory Figure 3D tested positive at timepoint 1


https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

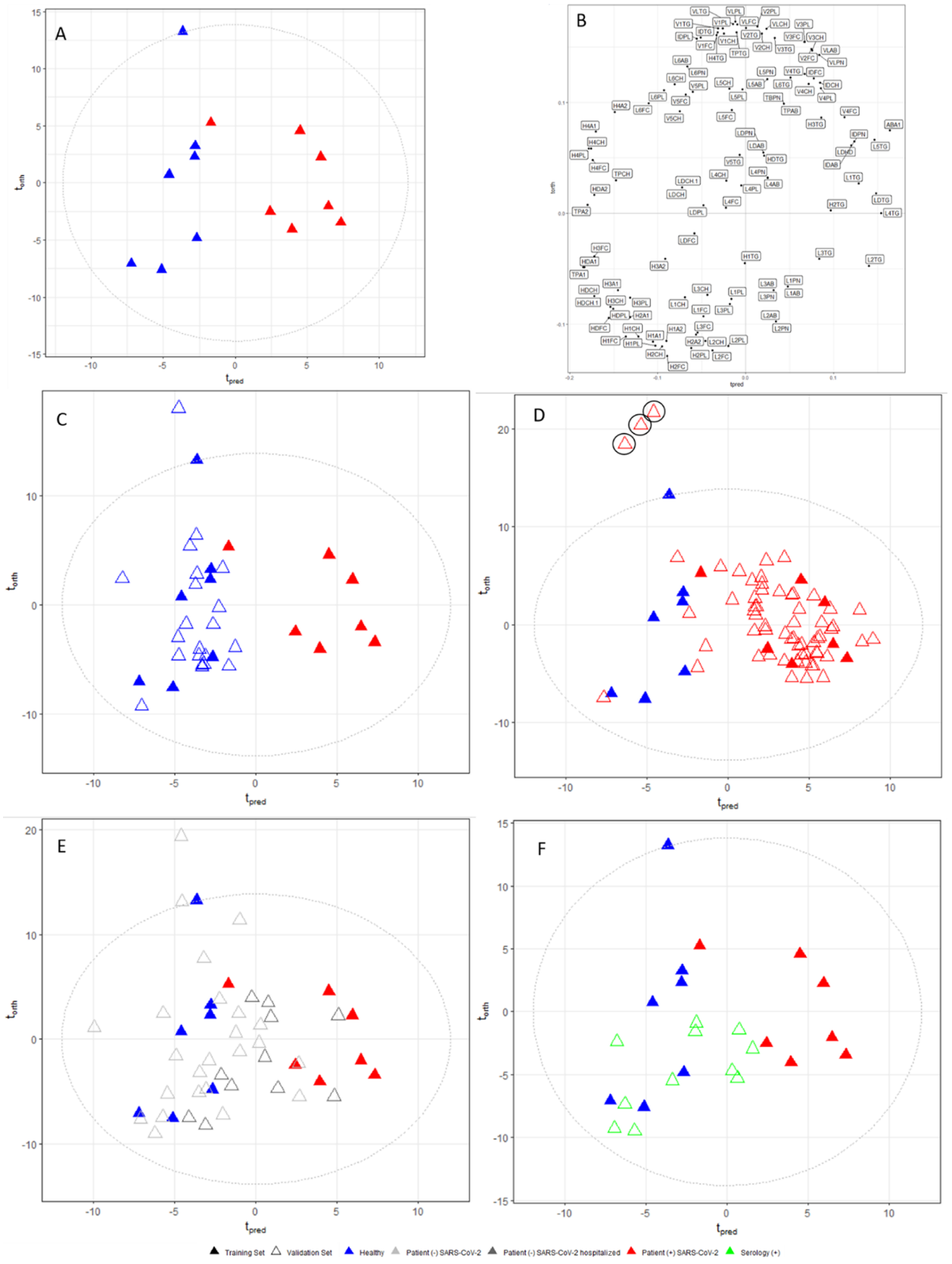



**Figure 4.** OPLS-DA of lipoprotein training set. (A) OPLS-DA score plot for healthy controls and SARS-CoV-2 positive patients; (B) corresponding loadings of the OPLS-DA; projection of the test set into the OPLS-DA training model for (C) healthy individuals; (D) SARS-CoV-2 positives patients; (E) SARS-CoV-2 negative patients, and (F) serology IgA positive.[53]

(red*) and tested negative at timepoint 5 (cyan*) with no testing carried out between timepoints 1 and 5. This observation suggests that SARS-CoV-2 infection leaves a residual metabolic signature with an, as yet, undetermined time



https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

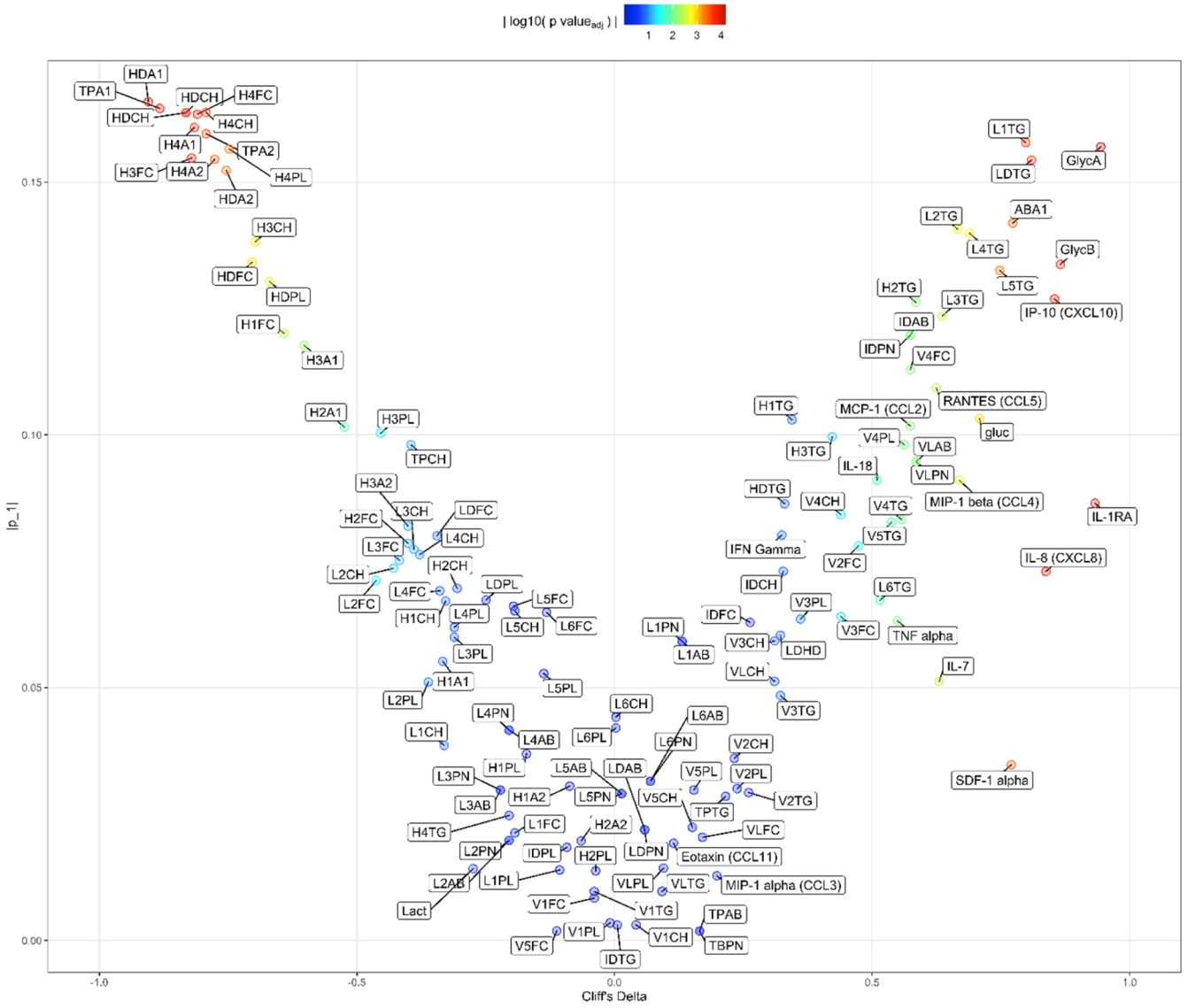



**Figure 5.** Eruption plot of the combined lipoprotein data, cytokine data, and GlycA and GlycB ratios for the SARS-CoV-2 positive samples formed from Cliff's delta (abscissa) and O-PLS-DA loadings (ordinate). Variables are color-coded for statistical significance.

of resolution. This type of information could be clinically useful in determining the rates of recovery of patients or indeed incomplete recovery with respect to new onset co-morbidities described previously.[7]

The PCA loadings demonstrate relatively higher concentrations of LDL cholesterol (LDCH), LDL phospholipids (LDPL), LDL-free cholesterol (LDFC), LDL apolipoprotein B, HDL cholesterol (HDCH), and phospholipids (HDPL) in the healthy control group, whereas the SARS-CoV-2 positive group was associated with higher concentrations of VLDL cholesterol (*e.g.*, V1CH, V2CH, and so forth) and triglycerides (*e.g.*, V1TG, V2TG, and so forth) subclasses in general. In order to interrogate further the impact of the virus on the serum lipoprotein composition, an OPLS-DA model was constructed (Figure 4) based on the same training set as the 1D and CPMG spectral dataset. The model yielded an AUROC of 1.0 attesting to the strength of the model and echoing the observations from the PCA model. The strong influence of high concentrations of LDL triglycerides and its subclasses (*e.g.*, L1 to L4TG) in driving the SARS-CoV-2 signature was reinforced along with VLDL cholesterol (V3CH) and phospholipids (V4PL). In addition, the ratio of apolipoprotein A1 to apolipoprotein B100 (ABA1) was strongly positively associated with the SARS-CoV-2 positive group (Figure 4B and Table S5). Healthy participants demonstrated higher concentrations of HDL and LDL cholesterol and phospholipids (HDCH, HDPL, LDCH, and LDPL) and total plasma apolipoproteins A1 and A2 (TPA1 and TPA2). The lipoprotein signature attributed to the SARS-CoV-2 positive group shares similarities with the lipoprotein signature associated with type 1 diabetes reported by Llauradó *et al.*,[46] who showed that diabetes was associated with high concentrations of LDL triglycerides, VLDL particle number and LDL, HDL, and total cholesterol. A lower LDL particle number was also a key feature of both SARS-CoV-2 infection and type 1 diabetes. Coronary artery disease (CAD) in diabetics has previously been associated with higher concentrations of small dense LDL (LDL 5 and 6), which demonstrated an association with increased risk of CAD in a diabetic population.[47] Other studies have confirmed the observation that small and medium LDLs are predictive of CAD risk in type 1 diabetics and further have shown that the


https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

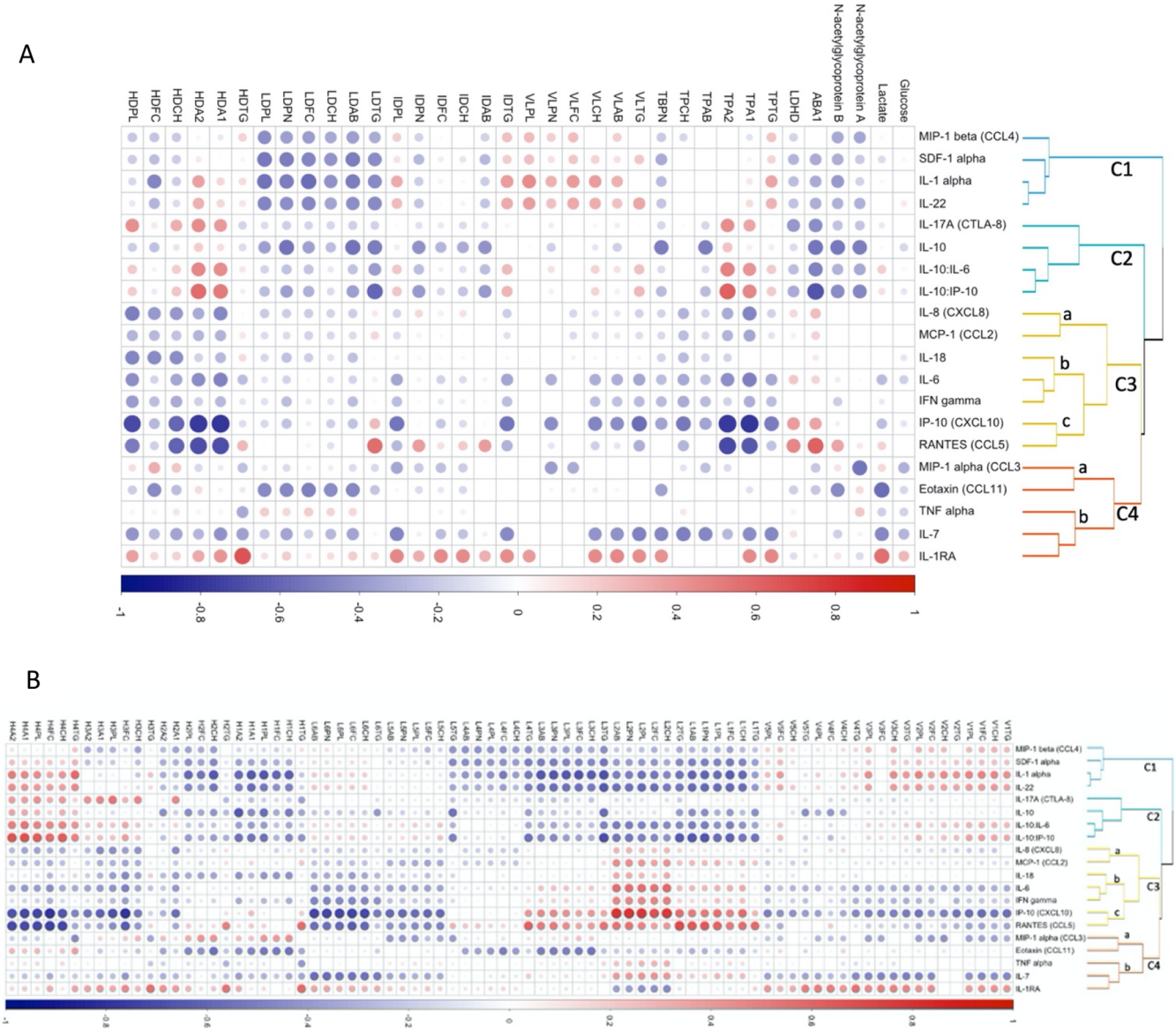



**Figure 6.** Immunometabolic correlation plot between hierarchically clustered cytokine subpatterns of significantly expressed proteins in COVID-19 positive patients versus their quantified lipoprotein patterns for (A) major fraction components only together with glucose, lactate, and glycoprotein levels and (B) lipoprotein subfractions. Cluster subpatterns C1−C4 are discussed in the main text.

LDL particle concentrations were associated with carotid intima media thickness.[48] This is of particular relevance to the consideration of the longer term effects of the disease which include abnormal cardiac MRI patterns in 78% of COVID-19 patients up to 3 months after so-called recovery,[49] which could have potentially serious longer term changes to patients' disease risk profiles. This would point to the potential value of lipoprotein monitoring to help assess long-term recovery in COVID-19 patients. It is of note that type 1 diabetes is typically characterized by the increased liver fat content, which is also associated with CAD. Soedamah-Muthu *et al.* showed a strong correlation between the liver fat content and higher VLDL and VLDL levels[48] consistent with our observations in SARS-CoV-2 positive patients, which is consistent with liver dysfunction identified as characteristic of SARS-CoV-2 infection, as indicated by the elevated concentrations of tyrosine, phenylalanine, and glutamine/glutamate and Fischer's ratios.[7]

Infection with SARS-CoV-2 has been associated with adverse cardiac events.[6] A comparison of perturbation in the lipoprotein profile associated with patients with coronary artery calcification, carotid intima thickening, and SARS-CoV-2 positive patients of the current study shows an overlap in higher LDL triglycerides under the disease conditions compared to healthy controls.[50] In contrast, the lipoprotein signature associated with intima thickening showed distinct differences compared to that of SARS-CoV-2 positive concentrations. For example, where total LDL, LDL-5, and LDL-6 cholesterol and phospholipid apolipoprotein B are positively associated with intima thickening, they do exert a strong influence on the SARS-CoV-2 model. Although there is a clear trend in the literature toward a distinctive lipoprotein signature for diabetes, CAD, and liver disease, it should be borne in mind that the lipoprotein profiles are significantly influenced by gender[51] and ethnicity.[52] It is yet to emerge if SARS-CoV-2 infection produces a unique phenotypic pattern of dyslipoproteinemia.


https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

The remaining study participants (test set) were predicted in the SARS-CoV training model and showed, as expected, the healthy controls were classified with the controls in the training set (Figure 4C). For the SARS-CoV-2 positive validation set, most of the validation set projected correctly with only one sample mapping with the control group and a further four mapping on the border between the groups in the healthy control (Figure 4D). Three further samples from the SARS-CoV-2 validation group (denoted by symbols within black circles) mapped outside the Hotelling's T ellipse and were attributed to a patient with type 2 diabetes. The patients presenting with respiratory symptoms but who were identified as SARS-CoV-2 negative from upper and/or lower respiratory tract swabs by rRT-PCR mainly mapped to the control group with six samples falling clearly within the SARS-CoV-2 positive sample grouping and SARS-CoV-2 negative groups (Figure 4E). Unlike the standard 1D spectral dataset, the control samples later shown to be seropositive for IgA largely co-mapped with the control group (Figure 4F).

The Cliff's delta eruption plot was modeled based on a combination of lipoprotein and cytokine concentrations in addition to the GlycA/GlycB ratios to compare the strength of all the parameters in the discrimination between the SARS-CoV-2 positive patients and the healthy controls (Figure 5). The comparative effect size of the difference between the lipoprotein profile, cytokine dataset, and GlycA and GlycB ratios associated with the healthy versus SARS-CoV-2 positive groups (abscissa) and the absolute weight of parameters in the OPLS-DA model (ordinate) is shown in the eruption plot.

Thus, those parameters that differentiate the two groups on the basis of both the multivariate and univariate metric are at the top corners of the plot with those that are elevated in the SARS-CoV-2 positive group mapping to the right-hand side of the plot and those that are elevated in the control group mapping to the top left. The parameters are colored according to the absolute log-transformed $p$-value corrected for the false discovery rate. Thus, the "hotter" the color, the higher the discriminatory power of the lipoprotein, cytokine, or GlycA/GlycB ratio variables. The plot emphasizes the higher levels of the HDL lipoprotein class in the plasma of the control group along with the total apolipoprotein particles, whereas the triglycerides, particularly those associated with LDL and the ratio of apolipoprotein A1 to apolipoprotein B100 ratio (ABA1) drive the signature associated with the COVID-19 disease. The ABA1 ratio has been associated with the atherogenic risk and represents the balance between atherogenic particles that are rich in apolipoprotein B and particles that are rich in apolipoprotein-1, which is generally considered to be antiatherogenic.[53] IP-10 mapped together with ABA1 and several of the LDL triglyceride components with high weightage in both Cliff's delta (abscissa) and O-PLS-DA loadings (ordinate) and was key in defining the SARS-CoV-2 positive class. IL-1RA, IL-8 and SDF-1$\alpha$ all ranked highly using the Cliff's delta statistic but did not exert such a strong influence on the O-PLS-DA loadings. As expected, none of the cytokines were associated with metabolic patterns in the healthy control group.

The SARS-CoV-2 positive and healthy control serum samples were clearly differentiated based on their cytokine composition (as illustrated in the OPLS-DA score plot; Figure S6). Ten of the panels of inflammatory cytokines and chemokines were found to be significantly different between healthy controls and SARS-CoV-2 positive patients (Table S6 and Figure S7), with most of these also differentiating between the SARS-CoV-2 positive and SARS-CoV-2 negative groups. IL-1RA, IP-10, and IL-8 were the strongest differential cytokines.

Many of the lipoprotein parameters showed a strong correlation with multiple cytokines, and four main clusters of cytokine−lipoprotein correlation patterns were observed when the correlation matrix was ranked by hierarchically clustering the cytokines. Cluster C1 consisted of MIP-1$\beta$, SDF-1$\alpha$, Il-1$\alpha$, and IL-22, which were associated with decreased levels of LDL particles and increased VLDL particles (Figure 6). MIP-1$\beta$ was strongly discriminatory between the control and SARS-CoV-2 positive groups, and together with the rest of the cluster, most likely reflects the inflammation mediated by innate immune cells and is consistent with the activation of monocytes and/or macrophages in response to the virus. The second cluster (C2), consisting of IL-17 and IL-10 and ratios of IL-10 to IL-6 and IP-10 demonstrated a much weaker general correlation with the lipoproteins. IL-8 and MCP-1 (cluster 3a) were characteristic of the virus infection and were weakly inversely correlated with most of the lipoproteins, with the exception of L1 and L2 LDL subfractions of cholesterol and triglycerides. Confirming the observation with cluster 1, this group reflects inflammation mediated essentially by innate immune cells and implies a central role played by activated neutrophils and monocytes/macrophages. IP-10 and RANTES (subcluster 3c) show the strongest association of all of the cytokines with the lipoprotein panel (correlated with L1−L4 LDL subfractions and anticorrelated with HDL, VLDL, and L6 LDL parameters). Additionally, RANTES is directly correlated with GlycA, a marker of systemic inflammation and has been proposed as a more robust marker of inflammation than CRP or IL-6.[54]

Both RANTES and IP-10 (cluster 3C) are associated with inflammatory and cytotoxic T cell responses that help eliminate viruses, and their correlation with GlycA may suggest that the increase in these circulating cytokines is reflecting the severity of the respiratory infection[55,56] Neotarangelo showed that, during the course of disease, certain inflammatory markers, such as IL-6, did not change significantly, whereas others, such as soluble IL-33R (sIL-33R) and CXCL10, decreased in patients who eventually recovered but remained persistently elevated in those who succumbed to COVID-19.[57]

Cluster 4b is defined by IL-7 and IL-1RA, which are significantly elevated in SARS-CoV-2 patients and TNF-$\alpha$. The functions of TNF-$\alpha$ and IL-1RA would suggest a good prognosis with reduced inflammation and increased elimination of the infected cells. However, their levels may be insufficient to counter the level of inflammatory cytokines released in response to the virus. IL-1RA is strongly correlated with HDL and VLDL lipoprotein particles. HDL is generally classed as an anti-inflammatory lipoprotein with protective effects against oxidized lipids[58] and has been shown to negatively regulate T cell activation,[59] thus the correlation of HDL with IL-1RA supports the hypothesis that IL-1RA reduces the inflammatory cascade.

## CONCLUSIONS

We have demonstrated that single-pulse, spin-echo, and quantitative NMR methods all provide strong differential diagnostic models for SARS CoV-2 positive patients versus controls and SARS CoV-2 negative patients based on models of their characteristic lipoprotein, glycoprotein, and metabolite


https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

profiles. Each NMR method provides valuable and complementary windows on the complex COVID-19 systemic pattern. There were strong indications of the underlying systemic disease in the NMR data indicative or diabetes, liver dysfunction, and cardiovascular abnormalities. Long-term cardiovascular problems appear to be common for COVID-19 patients, and these may be reflected in abnormal lipoproteins after the acute respiratory symptoms have subsided, indicating a potential role for these measurements in the assessment of systemic patient recovery. There are strong and characteristic statistical relationships between multiple cytokines and chemokines and the lipoproteins in COVID-19, suggesting a major role in the overall immune response to the disease. However, we await more detailed longitudinal studies to evaluate the exact roles of each cytokine cluster−lipoprotein interaction in the overall immunometabolic mechanism and signature of the disease progression. We noted that some patients who were in the recovery process and also testing negative for the virus were still highly metabolically abnormal, indicating that the clinical recovery should not be assessed by reduction of respiratory symptoms alone.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00876.

Full cohort demographic data and the demographic data for the training set used in the modeling of 1D, CPMG, and lipoprotein data; symptom presentation in SARS-CoV-2 positive patients; serology IgA and IgG concentrations; annotation of keys used by the B.I.LISA method; lipoprotein OPLS-DA model loadings, Cliff's delta values, and *p* values; ranked importance of cytokines; loadings of OPLS-DA and Cliff's delta of healthy controls *vs* SARS-CoV-2 positive patients; PCA of the CPMG experiment and loading plots of PC 1 and 2; GlycA and GlycB NMR peak integrals for the healthy controls, SARS-CoV-2 positive patients, and SARS-CoV-2 negative participants; BMI regressed against the full $^{1}$H 1D spectrum; OPLS-DA analysis of the training set CPMG data and the validation set projections of healthy, SARS-CoV-2 negative, SARS-CoV-2 positive patients, and serology (+) participants; overlaid JRES spectra of a SARS-CoV-2 positive patient and a healthy control; the presence of paracetamol in a SARS-CoV-2 positive patient; OPLS-DA analysis of lipoprotein and cytokine/chemokine data for the healthy and SARS-CoV-2 positive patients; and box plots showing the cytokine and chemokine concentrations of healthy controls, SARS-CoV-2 negative, and SARS-CoV-2 positive patients (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Authors

**Elaine Holmes** − *Australian National Phenome Centre, Health Futures Institute and Centre for Computational and Systems Medicine, Health Futures Institute, Murdoch University, Perth, Western Australia 6150, Australia; Section of Nutrition Research , Department of Metabolism, Nutrition and Reproduction, Faculty of Medicine, Sir Alexander Fleming Building, Imperial College London, London SW7 2AZ, U.K.;* orcid.org/0000-0002-0556-8389; Email: elaine.holmes@murdoch.edu.au, elaine.holmes@imperial.ac.uk

**Jeremy K. Nicholson** − *Australian National Phenome Centre, Health Futures Institute and Centre for Computational and Systems Medicine, Health Futures Institute, Murdoch University, Perth, Western Australia 6150, Australia; Division of Surgery, Medical School, Faculty of Health and Medical Sciences, University of Western Australia, Perth, Western Australia 6150, Australia; Institute of Global Health Innovation, Imperial College London, London SW7 2NA, U.K.*; Email: Jeremy.nicholson@murdoch.edu.au, j.nicholson@imperial.ac.uk

### Authors

**Samantha Lodge** − *Australian National Phenome Centre, Health Futures Institute and Centre for Computational and Systems Medicine, Health Futures Institute, Murdoch University, Perth, Western Australia 6150, Australia;* orcid.org/0000-0001-9193-0462

**Philipp Nitschke** − *Australian National Phenome Centre, Health Futures Institute, Murdoch University, Perth, Western Australia 6150, Australia*

**Torben Kimhofer** − *Australian National Phenome Centre, Health Futures Institute and Centre for Computational and Systems Medicine, Health Futures Institute, Murdoch University, Perth, Western Australia 6150, Australia;* orcid.org/0000-0001-7158-9930

**Jerome D. Coudert** − *Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, Western Australia 6150, Australia; Perron Institute for Neurological and Translational Science, Nedlands, Western Australia 6009, Australia; School of Medicine, University of Notre Dame, Fremantle, Western Australia 6160, Australia*

**Sofina Begum** − *Australian National Phenome Centre, Health Futures Institute, Murdoch University, Perth, Western Australia 6150, Australia; Section of Nutrition Research , Department of Metabolism, Nutrition and Reproduction, Faculty of Medicine, Sir Alexander Fleming Building, Imperial College London, London SW7 2AZ, U.K.*

**Sze-How Bong** − *Australian National Phenome Centre, Health Futures Institute, Murdoch University, Perth, Western Australia 6150, Australia*

**Toby Richards** − *Division of Surgery, Medical School, Faculty of Health and Medical Sciences, University of Western Australia, Perth, Western Australia 6150, Australia*

**Dale Edgar** − *Faculty of Health and Medical Sciences, University of Western Australia, Perth, Western Australia 6150, Australia*

**Edward Raby** − *Department of Clinical Microbiology, PathWest Laboratory Medicine WA, Perth, Western Australia 6150, Australia*

**Manfred Spraul** − *Biospin GmbH, Ettlingen 76275, Germany*

**Hartmut Schaefer** − *Biospin GmbH, Ettlingen 76275, Germany*

**John C. Lindon** − *Division of Systems Medicine, Department of Metabolism, Nutrition and Reproduction, Faculty of Medicine, Sir Alexander Fleming Building, Imperial College London, London SW7 2AZ, U.K.;* orcid.org/0000-0002-0916-6360

**Ruey Leng Loo** − *Australian National Phenome Centre, Health Futures Institute and Centre for Computational and Systems Medicine, Health Futures Institute, Murdoch*




https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

*University, Perth, Western Australia 6150, Australia;*
orcid.org/0000-0001-5307-5709

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jproteome.0c00876

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank the Spinnaker Health Foundation, WA, The McCusker Foundation, WA, the Western Australian State Government, and the MRFF for funding the Australian National Phenome Centre for this and related work. We thank the UK MRC for funding (S.B.), and the Department of Jobs, Tourism, Science and Innovation, Government of Western Australian Premier's Fellowship for funding R.-L.L. and E.H.; and ARC Laureate Fellowship funding for E.H. We would also like to acknowledge the Western Australian COVID Research Response team (https://research-au.net/covid-research-response/), Giuliana D'Aulerio, Kelly Beer, Rolee Kumar, Doug Robb, Joseph Miocevich, Dominic Mallon, Michael Epis, Merrilee Needham, Daniel Fatovich, Aron Chakera, Thomas Gilbert, Nathanael Foo, @STRIVE WA, Candice Peel, Sheeraz Mohd, and Ali Alishum for the coordination, sampling, and biobanking of patient samples and clinical metadata. We would also like to acknowledge Darren McKee and David Morrison for support in the establishment of infrastructure for cytokine analysis. We would like to thank Anuradha Sooda and Emily McLeish for their technical assistance with cytokine analysis.

## ■ ADDITIONAL NOTE

[a]When calibrated to the $\alpha$ anomeric proton signal of glucose, the chemical shifts for GlycA and GlycB are $\delta$ 2.03 and $\delta$ 2.07, respectively. When calibrated to the TSP signal using the IVDr method, GlycA and GlycB signals are at $\delta$ 2.06 and $\delta$ 2.10, respectively.

## ■ REFERENCES

(1) Rubino, F.; Amiel, S. A.; Zimmet, P.; Alberti, G.; Bornstein, S.; Eckel, R. H.; Mingrone, G.; Boehm, B.; Cooper, M. E.; Chai, Z.; Del Prato, S.; Ji, L.; Hopkins, D.; Herman, W. H.; Khunti, K.; Mbanya, J.-C.; Renard, E. New-Onset Diabetes in Covid-19. *N. Engl. J. Med.* **2020**, *383*, 789−790.

(2) Zhang, C.; Shi, L.; Wang, F.-S. Liver injury in COVID-19: management and challenges. *Lancet Gastroenterol. Hepatol.* **2020**, *5*, 428−430.

(3) Ellul, M. A.; Benjamin, L.; Singh, B.; Lant, S.; Michael, B. D.; Easton, A.; Kneen, R.; Defres, S.; Sejvar, J.; Solomon, T. Neurological associations of COVID-19. *Lancet Neurol.* **2020**, *19*, 767−783.

(4) Ronco, C.; Reis, T.; Husain-Syed, F. Management of acute kidney injury in patients with COVID-19. *Lancet Respir. Med.* **2020**, *8*, 738−742.

(5) Shi, P.; Ren, G.; Yang, J.; Li, Z.; Deng, S.; Li, M.; Wang, S.; Xu, X.; Chen, F.; Li, Y.; Li, C.; Yang, X.; Xie, Z.; Wu, Z.; Chen, M. Clinical characteristics of imported and second-generation coronavirus disease 2019 (COVID-19) cases in Shaanxi outside Wuhan, China: a multicentre retrospective study. *Epidemiol. Infect.* **2020**, *148*, No. e238.

(6) Zheng, Y.-Y.; Ma, Y.-T.; Zhang, J.-Y.; Xie, X. COVID-19 and the cardiovascular system. *Nat. Rev. Cardiol.* **2020**, *17*, 259−260.

(7) Kimhofer, T.; Lodge, S.; Whiley, L.; Gray, N.; Loo, R. L.; Lawler, N. G.; Nitschke, P.; Bong, S.-H.; Morrison, D. L.; Begum, S.; Richards, T.; Yeap, B. B.; Smith, C.; Smith, K. G. C.; Holmes, E.; Nicholson, J. K. Integrative Modelling of Quantitative Plasma Lipoprotein, Metabolic and Amino Acid Data Reveals a Multi-organ Pathological Signature of SARS-CoV-2 Infection. *J. Proteome Res.* **2020**, *19*, 4442−4454.

(8) Shah, R. R.; Smith, R. L. Inflammation-induced phenoconversion of polymorphic drug metabolizing enzymes: hypothesis with implications for personalized medicine. *Drug Metab. Dispos.* **2015**, *43*, 400−410.

(9) Nicholson, J. K.; Holmes, E.; Kinross, J. M.; Darzi, A. W.; Takats, Z.; Lindon, J. C. Metabolic phenotyping in clinical and surgical environments. *Nature* **2012**, *491*, 384−392.

(10) Garcia-Perez, I.; Posma, J. M.; Gibson, R.; Chambers, E. S.; Hansen, T. H.; Vestergaard, H.; Hansen, T.; Beckmann, M.; Pedersen, O.; Elliott, P.; Stamler, J.; Nicholson, J. K.; Draper, J.; Mathers, J. C.; Holmes, E.; Frost, G. Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. *Lancet Diabetes Endocrinol.* **2017**, *5*, 184−195.

(11) Capati, A.; Ijare, O. B.; Bezabeh, T. Diagnostic Applications of Nuclear Magnetic Resonance-Based Urinary Metabolomics. *Magn. Reson. Insights* **2017**, *10*, 1178623X1769434.

(12) Whiley, L.; Sen, A.; Heaton, J.; Proitsi, P.; García-Gómez, D.; Leung, R.; Smith, N.; Thambisetty, M.; Kloszewska, I.; Mecocci, P.; Soininen, H.; Tsolaki, M.; Vellas, B.; Lovestone, S.; Legido-Quigley, C.; AddNeuroMed, C. Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. *Neurobiol. Aging* **2014**, *35*, 271−278.

(13) Chuang, S.-C.; Fanidi, A.; Ueland, P. M.; Relton, C.; Midttun, O.; Vollset, S. E.; Gunter, M. J.; Seckl, M. J.; Travis, R. C.; Wareham, N.; Trichopoulou, A.; Lagiou, P.; Trichopoulos, D.; Peeters, P. H. M.; Bueno-de-Mesquita, H. B.; Boeing, H.; Wientzek, A.; Kuehn, T.; Kaaks, R.; Tumino, R.; Agnoli, C.; Palli, D.; Naccarati, A.; Aicua, E. A.; Sanchez, M.-J.; Quiros, J. R.; Chirlaque, M.-D.; Agudo, A.; Johansson, M.; Grankvist, K.; Boutron-Ruault, M.-C.; Clavel-Chapelon, F.; Fagherazzi, G.; Weiderpass, E.; Riboli, E.; Brennan, P. J.; Vineis, P.; Johansson, M. Circulating biomarkers of tryptophan and the kynurenine pathway and lung cancer risk. *Cancer Epidemiol., Biomarkers Prev.* **2014**, *23*, 461−468.

(14) Ayres, J. S. A metabolic handbook for the COVID-19 pandemic. *Nat. Metab.* **2020**, *2*, 572−585.

(15) Bell, J. D.; Brown, J. C. C.; Nicholson, J. K.; Sadler, P. J. Assignment of resonances for 'acute-phase' glycoproteins in high resolution proton NMR spectra of human blood plasma. *FEBS Lett.* **1987**, *215*, 311−315.

(16) Nicholson, J. K.; Buckingham, M. J.; Sadler, P. J. High resolution 1H n.m.r. studies of vertebrate blood and plasma. *Biochem. J.* **1983**, *211*, 605−615.

(17) Bales, J. R.; Higham, D. P.; Howe, I.; Nicholson, J. K.; Sadler, P. J. Use of high-resolution proton nuclear magnetic resonance spectroscopy for rapid multi-component analysis of urine. *Clin. Chem.* **1984**, *30*, 426−432.

(18) Nicholson, J. K.; O'Flynn, M. P.; Sadler, P. J.; Macleod, A. F.; Juul, S. M.; Sönksen, P. H. Proton-nuclear-magnetic-resonance studies of serum, plasma and urine from fasting normal and diabetic subjects. *Biochem. J.* **1984**, *217*, 365−375.

(19) Nicholson, J. K.; Foxall, P. J. D.; Spraul, M.; Farrant, R. D.; Lindon, J. C. 750 MHz 1H and 1H-13C NMR spectroscopy of human blood plasma. *Anal. Chem.* **1995**, *67*, 793−811.

(20) Otvos, J. D.; Jeyarajah, E. J.; Bennett, D. W. Quantification of plasma lipoproteins by proton nuclear magnetic resonance spectroscopy. *Clin. Chem.* **1991**, *37*, 377−386.

(21) Monsonis Centelles, S.; Hoefsloot, H. C. J.; Khakimov, B.; Ebrahimi, P.; Lind, M. V.; Kristensen, M.; de Roo, N.; Jacobs, D. M.; van Duynhoven, J.; Cannet, C.; Fang, F.; Humpfer, E.; Schäfer, H.; Spraul, M.; Engelsen, S. B.; Smilde, A. K. Toward Reliable Lipoprotein Particle Predictions from NMR Spectra of Human Blood: An Interlaboratory Ring Test. *Anal. Chem.* **2017**, *89*, 8004−8012.

(22) Jiménez, B.; Holmes, E.; Heude, C.; Tolson, R. F.; Harvey, N.; Lodge, S. L.; Chetwynd, A. J.; Cannet, C.; Fang, F.; Pearce, J. T. M.;

Lewis, M. R.; Viant, M. R.; Lindon, J. C.; Spraul, M.; Schäfer, H.; Nicholson, J. K. Quantitative Lipoprotein Subclass and Low Molecular Weight Metabolite Analysis in Human Serum and Plasma by (1)H NMR Spectroscopy in a Multilaboratory Trial. *Anal. Chem.* **2018**, *90*, 11962−11971.
(23) Loo, R. L.; Lodge, S.; Kimhofer, T.; Bong, S.-H.; Begum, S.; Whiley, L.; Gray, N.; Lindon, J. C.; Nitschke, P.; Lawler, N. G.; Schäfer, H.; Spraul, M.; Richards, T.; Nicholson, J. K.; Holmes, E. Quantitative In-Vitro Diagnostic NMR Spectroscopy for Lipoprotein and Metabolite Measurements in Plasma and Serum: Recommendations for Analytical Artefact Minimization with Special Reference to COVID-19/SARS-CoV-2 Samples. *J. Proteome Res.* **2020**, *19*, 4428−4441.
(24) Dona, A. C.; Jiménez, B.; Schäfer, H.; Humpfer, E.; Spraul, M.; Lewis, M. R.; Pearce, J. T. M.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Anal. Chem.* **2014**, *86*, 9887−9894.
(25) Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal. Chem.* **2006**, *78*, 4281−4290.
(26) Bylesjo, M.; Rantalainen, M.; Cloarec, O.; Nicholson, J.; Holmes, E.; Trygg, J. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J. Chemom.* **2007**, *20*, 341−351.
(27) Okba, N. M. A.; Müller, M. A.; Li, W.; Wang, C.; GeurtsvanKessel, C. H.; Corman, V. M.; Lamers, M. M.; Sikkema, R. S.; de Bruin, E.; Chandler, F. D.; Yazdanpanah, Y.; Le Hingrat, Q.; Descamps, D.; Houhou-Fidouh, N.; Reusken, C. B. E. M.; Bosch, B.-J.; Drosten, C.; Koopmans, M. P. G.; Haagmans, B. L. Severe Acute Respiratory Syndrome Coronavirus 2-Specific Antibody Responses in Coronavirus Disease Patients. *Emerging Infect. Dis.* **2020**, *26*, 1478−1488.
(28) Fuertes-Martín, R.; Correig, X.; Vallvé, J. C.; Amigó, N. Human Serum/Plasma Glycoprotein Analysis by (1)H-NMR, an Emerging Method of Inflammatory Assessment. *J. Clin. Med.* **2020**, *9*, 354.
(29) Dierckx, T.; Verstockt, B.; Vermeire, S.; van Weyenbergh, J. GlycA, a Nuclear Magnetic Resonance Spectroscopy Measure for Protein Glycosylation, is a Viable Biomarker for Disease Activity in IBD. *J. Crohns Colitis* **2019**, *13*, 389−394.
(30) Mokkala, K.; Vahlberg, T.; Pellonperä, O.; Houttu, N.; Koivuniemi, E.; Laitinen, K. Distinct Metabolic Profile in Early Pregnancy of Overweight and Obese Women Developing Gestational Diabetes. *J. Nutr.* **2020**, *150*, 31−37.
(31) Kettunen, J.; Ritchie, S. C.; Anufrieva, O.; Lyytikäinen, L.-P.; Hernesniemi, J.; Karhunen, P. J.; Kuukasjärvi, P.; Laurikka, J.; Kähönen, M.; Lehtimäki, T.; Havulinna, A. S.; Salomaa, V.; Männistö, S.; Ala-Korpela, M.; Perola, M.; Inouye, M.; Würtz, P. Biomarker Glycoprotein Acetyls Is Associated With the Risk of a Wide Spectrum of Incident Diseases and Stratifies Mortality Risk in Angiography Patients. *Circ.: Genomic Precis. Med.* **2018**, *11*, No. e002234.
(32) Ballout, R. A.; Remaley, A. T. GlycA: A New Biomarker for Systemic Inflammation and Cardiovascular Disease (CVD) Risk Assessment. *J. Lab. Precis. Med.* **2020**, *5*, 17.
(33) Rodríguez-Carrio, J.; Alperi-López, M.; López, P.; Pérez-Álvarez, Á. I.; Gil-Serret, M.; Amigó, N.; Ulloa, C.; Benavente, L.; Ballina-García, F. J.; Suárez, A. GlycA Levels during the Earliest Stages of Rheumatoid Arthritis: Potential Use as a Biomarker of Subclinical Cardiovascular Disease. *J. Clin. Med.* **2020**, *9*, 2472.
(34) Weisman, J. K.; Meeks, D.; Mendelsohn, L.; Remaley, A. T.; Sampson, M.; Allen, D. T.; Nichols, J.; Shet, A. S.; Thein, S. L. GlycA is not a useful biomarker of inflammation in sickle cell disease. *Int. J. Lab. Hematol.* **2018**, *40*, 704−709.
(35) Ritchie, S. C.; Kettunen, J.; Brozynska, M.; Nath, A. P.; Havulinna, A. S.; Männistö, S.; Perola, M.; Salomaa, V.; Ala-Korpela, M.; Abraham, G.; Würtz, P.; Inouye, M. Elevated serum alpha-1 antitrypsin is a major component of GlycA-associated risk for future morbidity and mortality. *PLoS One* **2019**, *14*, No. e0223692.
(36) Mokkala, K.; Houttu, N.; Koivuniemi, E.; Sorensen, N.; Nielsen, H. B.; Laitinen, K. GlycA, a novel marker for low grade inflammation, reflects gut microbiome diversity and is more accurate than high sensitive CRP in reflecting metabolomic profile. *Metabolomics* **2020**, *16*, 76.
(37) Ahmad, M. S.; Alsaleh, M.; Kimhofer, T.; Ahmad, S.; Jamal, W.; Wali, S. O.; Nicholson, J. K.; Damanhouri, Z. A.; Holmes, E. Metabolic Phenotype of Obesity in a Saudi Population. *J. Proteome Res.* **2017**, *16*, 635−644.
(38) Debray, F.-G.; Mitchell, G. A.; Allard, P.; Robinson, B. H.; Hanley, J. A.; Lambert, M. Diagnostic accuracy of blood lactate-to-pyruvate molar ratio in the differential diagnosis of congenital lactic acidosis. *Clin. Chem.* **2007**, *53*, 916−921.
(39) Liu, M.; Nicholson, J. K.; Lindon, J. C. High-resolution diffusion and relaxation edited one- and two-dimensional 1H NMR spectroscopy of biological fluids. *Anal. Chem.* **1996**, *68*, 3370−3376.
(40) Liu, M.; Nicholson, J. K.; Parkinson, J. A.; Lindon, J. C. Measurement of biomolecular diffusion coefficients in blood plasma using two-dimensional 1H-1H diffusion-edited total-correlation NMR spectroscopy. *Anal. Chem.* **1997**, *69*, 1504−1509.
(41) Bales, J. R.; Nicholson, J. K.; Sadler, P. J. Two-dimensional proton nuclear magnetic resonance "maps" of acetaminophen metabolites in human urine. *Clin. Chem.* **1985**, *31*, 757−762.
(42) Holmes, E.; Loo, R. L.; Cloarec, O.; Coen, M.; Tang, H.; Maibaum, E.; Bruce, S.; Chan, Q.; Elliott, P.; Stamler, J.; Wilson, I. D.; Lindon, J. C.; Nicholson, J. K. Detection of urinary drug metabolite (xenometabolome) signatures in molecular epidemiology studies via statistical total correlation (NMR) spectroscopy. *Anal. Chem.* **2007**, *79*, 2629−2640.
(43) Holmes, E.; Bonner, F. W.; Sweatman, B. C.; Lindon, J. C.; Beddell, C. R.; Rahr, E.; Nicholson, J. K. Nuclear magnetic resonance spectroscopy and pattern recognition analysis of the biochemical processes associated with the progression of and recovery from nephrotoxic lesions in the rat induced by mercury(II) chloride and 2-bromoethanamine. *Mol. Pharmacol.* **1992**, *42*, 922−930.
(44) Waters, N. J.; Waterfield, C. J.; Farrant, R. D.; Holmes, E.; Nicholson, J. K. Metabonomic deconvolution of embedded toxicity: application to thioacetamide hepato- and nephrotoxicity. *Chem. Res. Toxicol.* **2005**, *18*, 639−654.
(45) Keun, H. C.; Ebbels, T. M. D.; Bollard, M. E.; Beckonert, O.; Antti, H.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Geometric trajectory analysis of metabolic responses to toxicity can define treatment specific profiles. *Chem. Res. Toxicol.* **2004**, *17*, 579−587.
(46) Llauradó, G.; Amigó, N.; Cano, A.; Ballesta, S.; Albert, L.; Mazarico, I.; Fernández-Veledo, S.; Pedro-Botet, J.; Vendrell, J.; González-Clemente, J.-M. Specific Nuclear Magnetic Resonance Lipoprotein Subclass Profiles and Central Arterial Stiffness in Type 1 Diabetes Mellitus: A Case Control Study. *J. Clin. Med.* **2019**, *8*, 1875.
(47) Erbey, J. R.; Robbins, D.; Forrest, K. Y.-Z.; Orchard, T. J. Low-density lipoprotein particle size and coronary artery disease in a childhood-onset type 1 diabetes population. *Metabolism* **1999**, *48*, 531−534.
(48) Soedamah-Muthu, S. S.; Chang, Y.-F.; Otvos, J.; Evans, R. W.; Orchard, T. J. Lipoprotein subclass measurements by nuclear magnetic resonance spectroscopy improve the prediction of coronary artery disease in Type 1 diabetes. A prospective report from the Pittsburgh Epidemiology of Diabetes Complications Study. *Diabetologia* **2003**, *46*, 674−682.
(49) Puntmann, V. O.; Carerj, M. L.; Wieters, I.; Fahim, M.; Arendt, C.; Hoffmann, J.; Shchendrygina, A.; Escher, F.; Vasa-Nicotera, M.; Zeiher, A. M.; Vehreschild, M.; Nagel, E. Outcomes of Cardiovascular Magnetic Resonance Imaging in Patients Recently Recovered From Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol.* **2020**, *5*, 1265−1273.
(50) Tzoulaki, I.; Castagné, R.; Boulangé, C. L.; Karaman, I.; Chekmeneva, E.; Evangelou, E.; Ebbels, T. M. D.; Kaluarachchi, M. R.; Chadeau-Hyam, M.; Mosen, D.; Dehghan, A.; Moayyeri, A.; Ferreira, D. L. S.; Guo, X.; Rotter, J. I.; Taylor, K. D.; Kavousi, M.; de


https://dx.doi.org/10.1021/acs.jproteome.0c00876
J. Proteome Res. 2021, 20, 1382−1396

Vries, P. S.; Lehne, B.; Loh, M.; Hofman, A.; Nicholson, J. K.; Chambers, J.; Gieger, C.; Holmes, E.; Tracy, R.; Kooner, J.; Greenland, P.; Franco, O. H.; Herrington, D.; Lindon, J. C.; Elliott, P. Serum metabolic signatures of coronary and carotid atherosclerosis and subsequent cardiovascular disease. *Eur. Heart J.* **2019**, *40*, 2883−2896.

(51) Knopp, R. H.; Paramsothy, P.; Retzlaff, B. M.; Fish, B.; Walden, C.; Dowdy, A.; Tsunehara, C.; Aikawa, K.; Cheung, M. C. Gender differences in lipoprotein metabolism and dietary response: basis in hormonal differences and implications for cardiovascular disease. *Curr. Atheroscler. Rep.* **2005**, *7*, 472−479.

(52) Garcia, A. E.; Kasim, N.; Tamboli, R. A.; Gonzalez, R. S.; Antoun, J.; Eckert, E. A.; Marks-Shulman, P. A.; Dunn, J.; Wattacheril, J.; Wallen, T.; Abumrad, N. N.; Flynn, C. R. Lipoprotein Profiles in Class III Obese Caucasian and African American Women with Nonalcoholic Fatty Liver Disease. *PLoS One* **2015**, *10*, No. e0142676.

(53) Kaneva, A. M.; Potolitsyna, N. N.; Bojko, E. R.; Odland, J. O. The apolipoprotein B/apolipoprotein A-I ratio as a potential marker of plasma atherogenicity. *Dis. Markers* **2015**, *2015*, 591454.

(54) Connelly, M. A.; Otvos, J. D.; Shalaurova, I.; Playford, M. P.; Mehta, N. N. GlycA, a novel biomarker of systemic inflammation and cardiovascular disease risk. *J. Transl. Med.* **2017**, *15*, 219.

(55) Thompson, T. M.; Roddam, P. L.; Harrison, L. M.; Aitken, J. A.; DeVincenzo, J. P. Viral Specific Factors Contribute to Clinical Respiratory Syncytial Virus Disease Severity Differences in Infants. *Clin Microbiol.* **2015**, *4*, 206.

(56) Kevat, A. C.; Carzino, R.; Vidmar, S.; Ranganathan, S. Glycoprotein A as a biomarker of pulmonary infection and inflammation in children with cystic fibrosis. *Pediatr. Pulmonol.* **2020**, *55*, 401−406.

(57) Buszko, M.; Park, J.-H.; Verthelyi, D.; Sen, R.; Young, H. A.; Rosenberg, A. S. The dynamic changes in cytokine responses in COVID-19: a snapshot of the current state of knowledge. *Nat. Immunol.* **2020**, *21*, 1146−1151.

(58) Brites, F.; Martin, M.; Guillas, I.; Kontush, A. Antioxidative activity of high-density lipoprotein (HDL): Mechanistic insights into potential clinical benefit. *BBA Clin.* **2017**, *8*, 66−77.

(59) Kaji, H. High-density lipoproteins and the immune system. *J. Lipids* **2013**, *2013*, 684903.



https://dx.doi.org/10.1021/acs.jproteome.0c00876
*J. Proteome Res.* 2021, 20, 1382−1396