RESEARCH

The dynamic changes and sex differences of 147 immune-related proteins during acute COVID-19 in 580 individuals

Guillaume Butler-Laporte[1,2], Edgar Gonzalez-Kozlova[3], Chen-Yang Su[1,4], Sirui Zhou[1,2], Tomoko Nakanishi[1,5,6,7], Elsa Brunet-Ratnasingham[8], David Morrison[1], Laetitia Laurent[1], Jonathan Afilalo[1,2], Marc Afilalo[9], Danielle Henry[1], Yiheng Chen[1,5], Julia Carrasco-Zanini[10], Yossi Farjoun[1], Maik Pietzner[10,11], Nofar Kimchi[1], Zaman Afrasiabi[1], Nardin Rezk[1], Meriem Bouab[1], Louis Petitjean[1], Charlotte Guzman[1], Xiaoqing Xue[1], Chris Tselios[1], Branka Vulesevic[1], Olumide Adeleye[1], Tala Abdullah[1], Noor Almamlouk[1], Yara Moussa[1], Chantal DeLuca[1], Naomi Duggan[1], Erwin Schurr[12], Nathalie Brassard[8], Madeleine Durand[8], Diane Marie Del Valle[13], Ryan Thompson[14], Mario A. Cedillo[15], Eric Schadt[14], Kai Nie[16], Nicole W. Simons[14], Konstantinos Mouskas[14], Nicolas Zaki[16], Manishkumar Patel[13], Hui Xie[16], Jocelyn Harris[16], Robert Marvin[16], Esther Cheng[14], Kevin Tuballes[13], Kimberly Argueta[16], Ieisha Scott[16], The Mount Sinai COVID-19 Biobank Team, Celia M. T. Greenwood[1,2], Clare Paterson[17], Michael Hinterberg[17], Claudia Langenberg[10,17], Vincenzo Forgetta[1], Vincent Mooser[5], Thomas Marron[3,13,18], Noam D. Beckmann[14], Ephraim Kenigsberg[14], Alexander W. Charney[19], Seunghee Kim-schulze[16], Miriam Merad[13], Daniel E. Kaufmann[8,20], Sacha Gnjatic[16] and J Brent Richards[1,2,5,21,22,23]*

Abstract

Introduction: Severe COVID-19 leads to important changes in circulating immune-related proteins. To date it has been difficult to understand their temporal relationship and identify cytokines that are drivers of severe COVID-19 outcomes and underlie differences in outcomes between sexes. Here, we measured 147 immune-related proteins during acute COVID-19 to investigate these questions.

Methods: We measured circulating protein abundances using the SOMAscan nucleic acid aptamer panel in two large independent hospital-based COVID-19 cohorts in Canada and the United States. We fit generalized additive models with cubic splines from the start of symptom onset to identify protein levels over the first 14 days of infection which were different between severe cases and controls, adjusting for age and sex. Severe cases were defined as individuals with COVID-19 requiring invasive or non-invasive mechanical respiratory support.

Results: 580 individuals were included in the analysis. Mean subject age was 64.3 (sd 18.1), and 47% were male. Of the 147 proteins, 69 showed a significant difference between cases and controls (p < 3.4×10^{-4}). Three clusters were

---

*Correspondence: brent.richards@mcgill.ca
[23] McGill University, King's College London (Honorary), Jewish General Hospital, Pavilion H-4133755 Côte-Ste-Catherine, Montréal, Québec H3T 1E2, Canada
Full list of author information is available at the end of the article

Butler-Laporte *et al. Clinical Proteomics* (2022) 19:34

Page 2 of 11

formed by 108 highly correlated proteins that replicated in both cohorts, making it difficult to determine which proteins have a true causal effect on severe COVID-19. Six proteins showed sex differences in levels over time, of which 3 were also associated with severe COVID-19: CCL26, IL1RL2, and IL3RA, providing insights to better understand the marked differences in outcomes by sex.

Conclusions: Severe COVID-19 is associated with large changes in 69 immune-related proteins. Further, five proteins were associated with sex differences in outcomes. These results provide direct insights into immune-related proteins that are strongly influenced by severe COVID-19 infection.

Keypoints

- COVID-19 is associated with changes in cytokines, interleukins, and other immune-related proteins. However, previous research has failed to account for the dynamic nature of these changes over the course of infection, leading to often contradictory results.
- We measured 147 immune-related protein in 580 individuals in three large academic centers to precisely map the evolution of these proteins during acute COVID-19.
- COVID-19 was associated with a clear change in 69 proteins. More importantly, 3 of them may also help explain sex-differences in COVID-19 outcomes.
- These results provide greater insight into the COVID-19 immune response, and how it leads to severe illness.

Keywords: COVID-19, Proteomics, SOMAscan, Immunity

Introduction

COVID-19 is characterized by a complex immune response which explains some of the observed variation in patient outcomes. In patients with a severe clinical course, some may develop a "COVID-19 cytokine storm" [1], though this term has been challenged due to a poor understanding of this response [2]. While previous publications found multiple cytokines and other immune-related proteins associated with COVID-19 outcomes [3–8], these associations were measured in small sample sizes, assessed a limited set of proteins, or did not provide a temporal analysis of the changes in cytokines during severe versus mild disease. These limitations may have contributed to contradicting results [9–11].

Similarly, it remains unclear if the host immune response explains a large proportion of differences in COVID-19 outcomes between males and females. While reports previously suggested that these differences were correlated with differential levels of cytokines (e.g. IL-8 and IL-18 [12]), these also suffered from small sample sizes and limited adjustment for temporal changes. These studies also likely contained many false positive associations due to multiple comparisons, as most of the sex-related differences were not replicated in other larger cohorts [13].

One way to address some of these limitations is by using high-throughput oligonucleotide-aptamer protein measurement technology [14]. These panels reliably measure thousands of blood circulating protein simultaneously, allowing for comprehensive measurements on larger number of subjects. The increase in sample size allows for better adjustments for time dependent changes in protein levels, providing a more granular understanding of their dynamics during infection. Here, we use the SOMAscan aptamer panel [15] (SomaLogic, Boulder, USA) in two prospectively enrolled cohorts from Canada and the United States (n = 580) to measure 147 proteins associated with the immune response over the first 14 days of COVID-19. This allowed us to clearly describe the temporal pattern of cytokines during COVID-19 disease progression.

By using large-scale protein measurement and accounting for temporal changes over the course of infection, we describe which proteins are likely associated with severe COVID-19, and which ones also underlie sex differences in outcomes.

Methods

Overview of study design

We used the SomaScan assay to measure 147 cytokines and other immune-related proteins in cases and controls in the Biobanque Québécoise de la COVID-19 [16] (BQC19) in Montreal, and in the Mount Sinai Biobank (MSB) at Icahn School of Medicine in New York City. We then combined those results using generalized additive models to identify proteins temporally associated with severe COVID-19.

Butler-Laporte *et al. Clinical Proteomics* (2022) 19:34

Page 3 of 11

Population

The BQC19 and MSB are hospital-based prospective cohorts enrolling subjects with PCR proven SARS-CoV-2 infections, as well as individuals who presented with signs or symptoms consistent with COVID-19, but without a microbiological diagnosis of COVID-19. For this study, the BQC19 cohort was limited to subjects enrolled at the Jewish General Hospital and Centre Hospitalier de l'Université de Montréal, both university affiliated hospitals. Demographic characteristics and clinical risk factors were obtained by medical chart review or subject interview performed by clinicians or trained research coordinators in all cohorts. Specifically, time from onset of symptoms used for all analyses were recorded by trained clinical assistants or physicians based on medical records review or patient or relatives interview.

COVID-19 case/control outcome definitions

Severe COVID-19 cases were defined as subjects with a positive SARS-CoV-2 PCR test result who either died or required invasive or non-invasive mechanical respiratory support. Mechanical respiratory support was defined as any one of the following: intubation, new positive airway pressure (CPAP) or bilevel positive airway pressure (BiPAP) ventilation, or high-flow nasal cannula. Controls were defined as any subjects with a positive PCR test who did not require invasive ventilation, or any subject with signs or symptoms consistent with COVID-19, but who had negative PCR tests for the virus. However, we also excluded participants who with severe non-COVID-19 disease (i.e. participants with respiratory support as defined above, but not due to COVID-19). This control definition was chosen to emphasize severe COVID-19 specific immune responses, as compared to a general hospital population.

Protein measurements

We used the SOMAscan (v4) platform to measure 5284 circulating proteins from each participant, and then prioritized 147 immune-related proteins for the analysis. These proteins were selected to include all available interleukins (n=38), CC motif chemokines (n=23), CXC motif chemokines (n=14), interferons (n=17), toll-like receptors (n=6), and immunoglobulins (n=5) available from the SOMAscan panel, as well as 6 other proteins (G-CSF, GM-CSF, M-CSF, MIF, TNF-α, TNF-β) known to be involved in viral immune responses [17–19]. We also included all 38 soluble interleukin receptors measured by SOMAscan. These soluble receptors act as decoy receptors for their respective interleukins. Biologically, they bind to their interleukins in the circulation, preventing them from binding membrane-bound receptors, and having their usual biological effect. Their action may

predict the effect of pharmacologic interleukin receptor blocking agents [20]. Owing to differences in the choice of aptamers in each SOMAscan panel, of the 147 proteins available in the BQC19 cohort, 15 were not available in the MSB cohort (IL-2, IL-7, IL-9, IL-34, IL-37, IL-12RB2, CCL1, CCL3, CXCL2, IFNB1, TLR2, IgD, IgE, IgG, and IgM). The full protein list in each cohort is available in Additional file 1.

To reflect acute illness, we limited this study to samples collected within 14 days of symptom onset (i.e. one sample per participant). To better control for the effect of COVID-19 treatment on circulating protein levels, we limited our analysis to only the first measurement of circulating proteins per subject, since these samples were less likely to be collected from individuals already starting therapy for severe COVID-19.

Samples were obtained and processed as per the manufacturer's instructions. Briefly, blood samples were collected in acid-citrate-dextrose tubes (to prevent coagulation) and frozen at -80 °C until analysis. Protein levels were measured using resonance fluorescence units, and further normalized and calibrated by SomaLogic to remove any systematic bias (e.g. batch effects). For the statistical analysis, we further standardized protein levels by subtracting their mean and dividing by their standard deviation to allow for easier interpretation and analysis.

Statistical analysis

To find clusters of proteins that varied together, we first drew Spearman's correlation heatmaps within cases and within controls separately. This was also done in both cohorts separately (i.e. 4 times in total). To better visualise Spearman correlation clusters, the proteins were ordered using a hierarchical clustering algorithm with the "complete linkage" method (implemented with the hclust base function in R [21], with default settings).

Second, to adjust for the time of onset of symptoms, which is expected to affect protein levels, we fit generalized additive models [22] (GAMs) on each protein levels during the first 14 days since symptom onset in cases and controls. In short, this analysis aims to model the natural history of protein levels by using measurements done at clinical presentation on different subjects (which have different time from onset of symptoms on presentation). GAMs fit spline between different immunity related proteins from days 1 to 14, it is therefore uniformly more powerful than dichotomizing protein levels in two time periods and comparing their levels. The GAMs were fit using cubic regression splines, the restricted maximum likelihood (REML) method, and with up to 15 knots allowed (the model chooses the optimal number of knots). All models were also adjusted for subjects' age and sex. The GAMs

Butler-Laporte *et al. Clinical Proteomics* (2022) 19:34

Page 4 of 11

Table 1 Subject characteristics in the two participating cohorts. Numbers presented as count (percentage) except where otherwise 570 noted. Hypertension information was not available for the Mount Sinai Biobank cohort.

	BQC19 (n = 333)		Mount Sinai Biobank (n = 247)	
	Cases (n = 91)	Controls (n = 242)	Cases (n = 119)	Controls (n = 128)
Age in years (mean)	67.2	66.2	64.8	59.2
Female sex	35 (38.5%)	133 (55.0%)	49 (41.2%)	57 (44.5%)
Hospital site	–	–	–	–
Centre Hospitalier de l'Université de Montréal	32 (35.2%)	22 (9.1%)	–	–
Jewish General Hospital	59 (64.8%)	220 (90.1%)	–	–
Mount Sinai Hospital	–	–		
COVID-19 positive	91 (100%)	202 (83.5%)	119 (100%)	128 (100%)
Diabetes	38 (41.8%)	71 (29.3%)	30 (25.2%)	32 (25.0%)
Chronic obstructive pulmonary disease	16 (17.6%)	27 (11.2%)	14 (11.8%)	8 (6.3%)
Chronic kidney disease	16 (17.6%)	24 (9.9%)	15 (12.6%)	26 (20.3%)
Congestive heart failure	13 (14.3%)	28 (11.6%)	14 (11.8%)	11 (8.6%)
Hypertension	60 (65.9%)	134 (55.4%)	–	–
Liver disease	2 (2.2%)	4 (1.7%)	6 (5.0%)	4 (3.1%)
Smoking status				
Current smoker	5 (5.5%)	6 (2.5%)	10 (8.4%)	9 (7.0%)
Ex-smoker	11 (12.1%)	30 (12.4%)	32 (26.9%)	35 (27.3%)
Never smoker	39 (42.9%)	172 (71.1%)	47 (39.5%)	65 (50.8)
Don't know	36 (39.6%)	34 (14.0%)	30 (25.2%)	19 (14.8%)

obtained in the two cohorts were de-identified and meta-analyzed (if measured in both cohorts) using the metagam package [23] (v0.2.0). The resulting meta-analyzed models were then plotted for a 65-year-old male and female (65 was chosen because it was the mean age in the BQC19 cohort).

To check if the protein levels were different between cases and controls, we used GAM ANOVA using a model without case/control status as predictor of protein level as the nested model. Similarly, we used GAM ANOVA with nested models with and without sex variables to check for difference in cytokine levels between sexes. Approximate p-values for this null hypothesis of no difference between cases and controls were obtained using GAM ANOVA. GAMs were fitted using the mgcv package [22] (v1.8–33). Sample code is available in Additional file 2. Finally, GAM ANOVA p-values were meta-analyzed across cohorts using the logistic method with the metap package [24] (v1.4). We considered that protein levels differed between cases and controls if the resulting p-value was below Bonferroni correction (alpha = 0.05/147 = 0.0003). We acknowledge that this correction is overly conservative due to the correlatedness of protein levels.

All analyses were done using R [21] (v4.0.3).

Results
Population

Table 1 shows basic characteristic of the participants in each cohort. Mean age was similar between cases and controls in the BQC19 (67.2 vs 66.2 year-old), but cases were slightly older in the MSB (64.8 vs 59.2 year-old). In both cohort, there were less females amongst cases compared to controls: 38.5% vs 55.0% in the BQC19, and 41.2% vs 44.5% in the MSB. There were more diabetic cases than controls in the BQC19 (41.8% vs 29.3%) but a similar proportion in the MSB. In both cohorts, there were more cases with chronic obstructive pulmonary disease: 17.6% vs 11.2% in the BQC-19, and 11.8% vs 6.3% in the MSB. There were also more heart failure diagnoses in cases in both cohorts: 14.3% vs 11.6% in the BQC-19, and 11.8% vs 8.6% in the MSB. Finally, there were less never-smokers amongst cases: 42.9% vs 71.1% in the BQC-19, and 39.5% vs 50.8% in the MSB. These values are comparable to other reported large COVID-19 cohorts [25].

Immune-related protein levels dynamics over time

Many cytokines and related proteins showed statistically significant time-dependent differences between cases and controls (Bonferroni threshold 0.05/147 = 0.00034): 17 of the 38 interleukins, 24 of the 38 soluble interleukin receptors, 11 of the 23 CC chemokines, 6 of the 14 CXC chemokines, 8 of the 17 interferons related proteins, and 3 of 17 other immune-related proteins (Table 2).

Butler-Laporte *et al. Clinical Proteomics* (2022) 19:34

Page 5 of 11

Table 2 Immune-related proteins with differences between severe COVID-19 cases and controls in our meta-analysis of the BQC-19 and MSB results (Bonferroni adjusted threshold $0.05/147 = 0.00034$)

Proteins	P-values	Proteins	P-values
Interleukins		**Soluble interleukin receptors**	
IL1A	9.04×10^{-6}	IL1R1	1.50×10^{-8}
IL1B	1.21×10^{-6}	IL1R2	1.91×10^{-7}
IL3	2.34×10^{-4}	IL1RAPL2	4.23×10^{-11}
IL4	1.35×10^{-11}	IL1RL1	2.90×10^{-12}
IL6	9.50×10^{-10}	IL1RL2	1.45×10^{-4}
IL11	1.14×10^{-10}	IL1RN	2.89×10^{-6}
IL12	1.26×10^{-4}	IL2RB	1.04×10^{-4}
IL13	3.04×10^{-7}	IL3RA	4.37×10^{-7}
IL17D	1.66×10^{-8}	IL4R	1.67×10^{-8}
IL17F	2.14×10^{-10}	IL7R	5.12×10^{-11}
IL18	8.26×10^{-7}	[b]IL10RB	6.18×10^{-5}
IL19	2.32×10^{-4}	IL10RA.soma2	5.17×10^{-8}
IL24	1.52×10^{-11}	IL11RA	1.64×10^{-6}
IL25	2.76×10^{-7}	IL12RB1	5.46×10^{-13}
IL36B	1.52×10^{-9}	IL13RA1	5.15×10^{-7}
IL36G	1.28×10^{-6}	[b]IL15RA.soma2	1.10×10^{-7}
[a]IL37	3.17×10^{-4}	IL17RB	1.22×10^{-4}
Interferons		IL17RC	1.02×10^{-4}
IFNA4	9.47×10^{-8}	IL18RAP	1.50×10^{-5}
IFNA6	1.09×10^{-7}	IL21R	1.28×10^{-11}
IFNA8	2.43×10^{-5}	IL22RA1	1.66×10^{-14}
IFNA10	4.74×10^{-4}	IL22RA2	1.10×10^{-5}
[a]IFNB1	1.80×10^{-4}	IL23R	5.80×10^{-5}
IFNL1	6.53×10^{-5}	IL27RA	2.11×10^{-6}
IFNL2	3.01×10^{-11}	**CXCL chemokines**	
IFNL3	3.60×10^{-5}	CXCL5	4.53×10^{-8}
CCL chemokines		CXCL10	3.22×10^{-9}
CCL7	5.21×10^{-10}	CXCL12	1.21×10^{-4}
CCL8	4.14×10^{-5}	CXCL13	1.01×10^{-5}
CCL11	5.23×10^{-6}	CXCL14	5.42×10^{-10}
CCL13	4.01×10^{-11}	CXCL16	2.26×10^{-5}
CCL19	1.93×10^{-6}	**Others**	
CCL20	1.30×10^{-6}	M-CSF	1.60×10^{-13}
CCL22	8.84×10^{-10}	TLR1.soma1	2.62×10^{-4}
CCL23	8.15×10^{-6}	LT-α / TNF-β	2.08×10^{-10}
CCL24	3.50×10^{-10}		
CCL26	2.78×10^{-8}		
CCL27	3.27×10^{-13}		

[a] For IL37 and IFNB1, the p-values from the BQC-19 p-value are shown (protein not available in the MSB panel)

[b] For IL10RB and IL15RA.soma2, the p-value from the MSB p-value is shown (GAM ANOVA approximation failure)
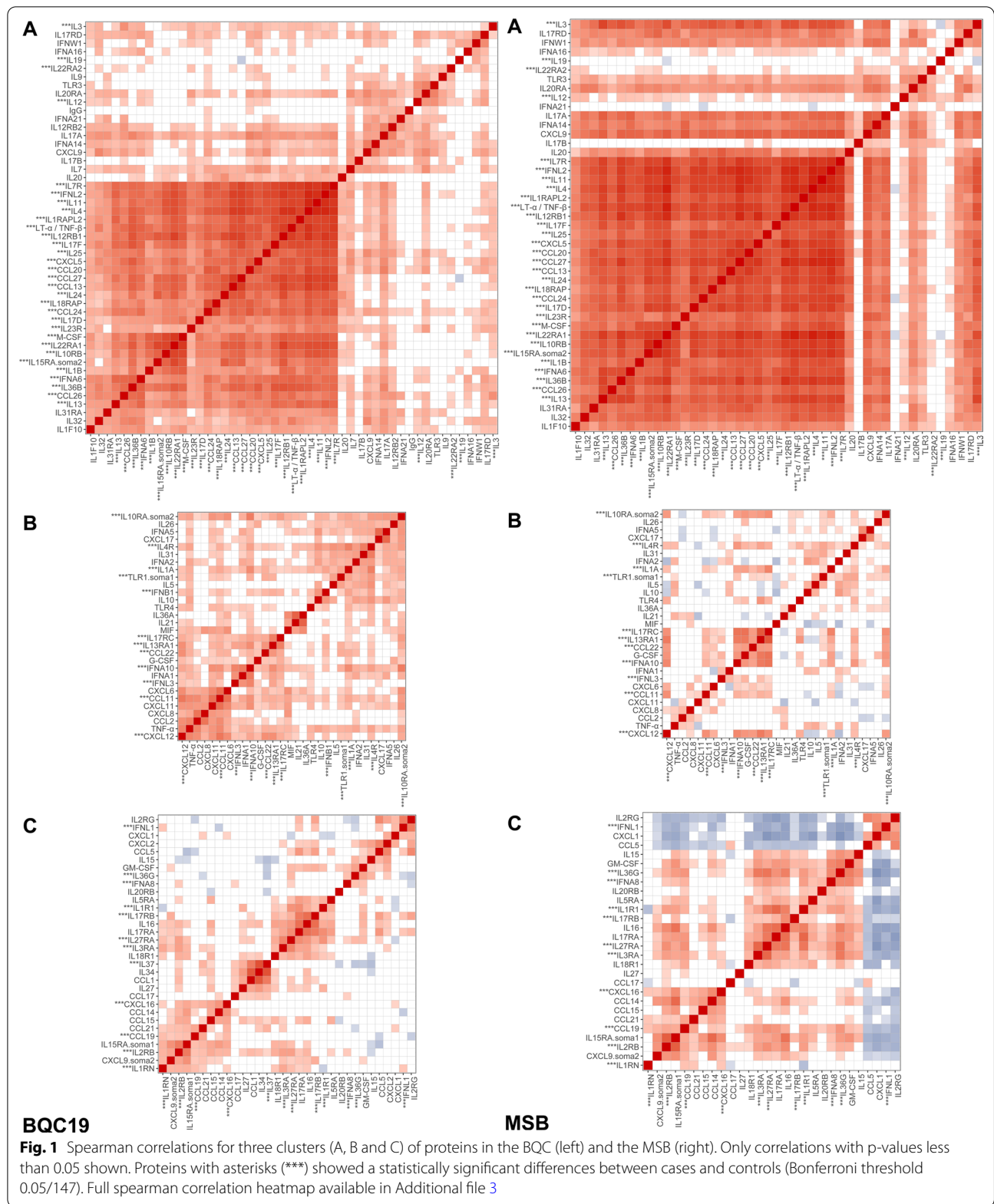
Hierarchical clustering and Spearman correlation delineated clear clusters of proteins that varied together over the course of infection (Fig. 1 and Additional file 3).
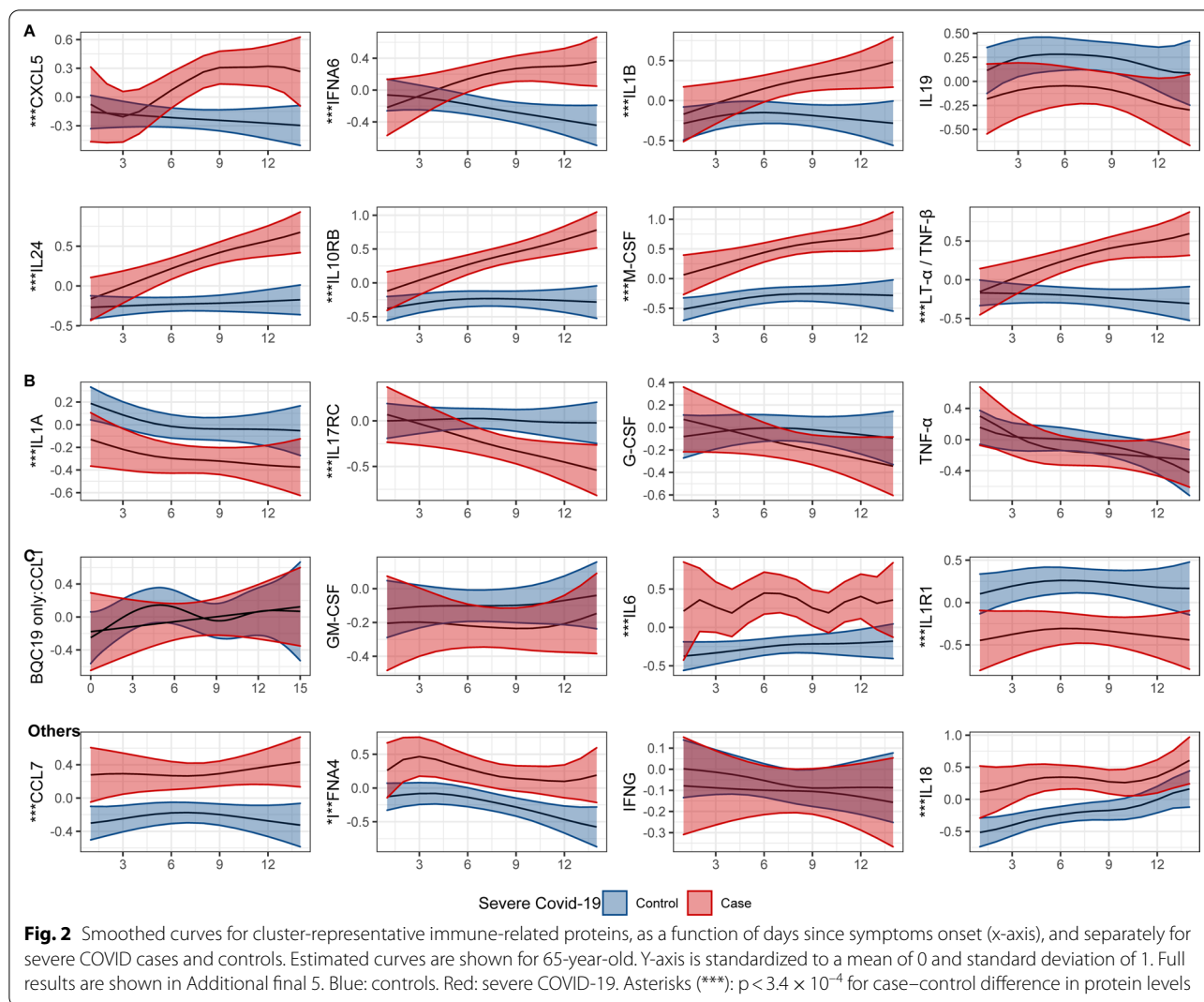
Visual inspection of the BQC19 with the MSB heatmaps reveals three large protein clusters whose members show similar changes in levels. Importantly, the cluster with the highest proportion number of proteins showing an association with case and control status (cluster A; 31 out of 49 proteins) is also the one with the highest mean absolute Spearman correlation in both cohorts. Cluster A also showed a clear increase in Spearman correlations between cases and controls (mean increase of 0.163 across cohorts), supporting the fact that an immune overactivation underlies severe COVID-19 (Additional file 3). Proteins in this cluster show increasing levels in cases over the first 14 days of symptoms, whereas they remain stable in controls. In contrast, clusters B and C showed a negative correlation with cluster A, with higher but slowly tapering protein levels in controls. These clusters did not replicate as clearly between cohorts and have a lower proportion of proteins associated with case and control status.

Each cluster contained a heterogeneous set of proteins (Additional file 4), with many replicating previously published findings [3–7], supporting the robustness of our results. Of note, many of the interleukin-10 family cytokines (IL19, IL20, IL24) or their soluble receptors (IL10RB, IL20A, IL22RA1 and IL22RA2) were present in cluster A. Other notable proteins found in cluster A include two members of the IL1 family (IL1B and IL36B), IL11 (a member of the IL6 family which both act on the same receptor [26]), multiple members of the IL17 family (IL17A, IL17B, IL17D, IL17F, and IL25), and IL4 and IL13 which both act on the same receptor to drive severe asthma [27]. Interestingly, clusters B and C contained some proteins related to those in cluster A which still showed significant differences between cases and controls. These included IL1A, IL4R (the soluble receptor for IL4), and IL10RA (a component of the soluble receptor for IL10) in cluster B, and IL1R1 (a soluble receptor for IL1) in cluster C. Figure 2 shows some representative proteins from each cluster and from the rest of the proteins, and all protein level plots are shown in Additional file 5.

Differences in cytokine levels between sexes in cases and controls over time

Of the 147 proteins, 5 showed a difference between males and females: TLR5, CXCL17, CCL28, CCL26, IL1RL2, and IL3RA. However, only the last three were also associated with case and control status (Fig. 3). CCL6 was found in the previously described cluster A of highly associated proteins, while IL3A was in cluster C. Both proteins had slightly different levels on the first day of symptoms (higher in females for CCL26, lower in females for IL3RA) but trended similarly afterwards. ILRL2
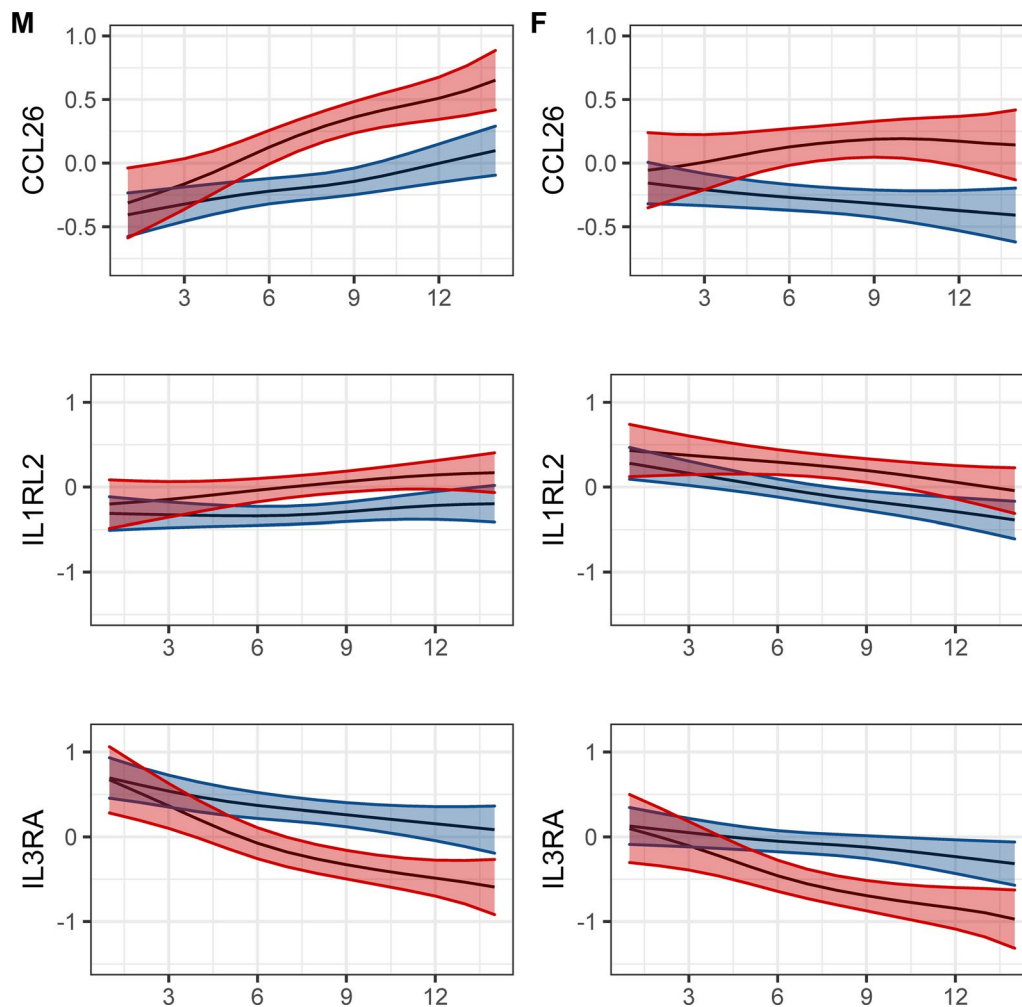
Butler-Laporte *et al. Clinical Proteomics* (2022) 19:34

Page 6 of 11

Fig. 1 Spearman correlations for three clusters (A, B and C) of proteins in the BQC (left) and the MSB (right). Only correlations with p-values less than 0.05 shown. Proteins with asterisks (***) showed a statistically significant differences between cases and controls (Bonferroni threshold 0.05/147). Full spearman correlation heatmap available in Additional file 3

Butler-Laporte *et al. Clinical Proteomics* (2022) 19:34

Page 7 of 11

Fig. 2 Smoothed curves for cluster-representative immune-related proteins, as a function of days since symptoms onset (x-axis), and separately for severe COVID cases and controls. Estimated curves are shown for 65-year-old. Y-axis is standardized to a mean of 0 and standard deviation of 1. Full results are shown in Additional final 5. Blue: controls. Red: severe COVID-19. Asterisks (***): $p < 3.4 \times 10^{-4}$ for case–control difference in protein levels

did not cluster well with other proteins and decreased towards normal levels more rapidly in females. Of note, IL3RA is the only one of these proteins for which the corresponding genes is located on a sex chromosome (chromosome X).

Discussion

In this study, we used two large prospective cohorts and a panel measuring 147 circulating immune-related proteins and found that severe COVID-19 was associated with a clear activation in many immune-related proteins, with most protein levels varying together closely overt time. These results also provided three proteins that were importantly different between the sexes: CCL26, IL1RL2 and IL3RA. These 3 sex-specific protein findings were not found in previous reports, which is partly explained by the fact that our panel included more proteins than other studies, but this could also suggest false positive

associations due to multiple testing. Hence, while most of the changes in immune-related proteins observed in severe COVID-19 are shared across sexes, it remains possible that disparity in outcomes between sexes may be mediated by differences in immune-related proteins levels.

Our study's main strengths include its large sample size with strong replication between the two independent cohorts, the large protein measurement panel and the fact that proteins were measured at different times during infection, a feature that was explicitly modelled into our analysis. These provided a granular depiction of time-dependent immune responses to COVID-19 and explain previously discordant reports on the association between different immune proteins and COVID-19. For example, visual inspection of the interferon level dynamics will often reveal that the differences in measurement timing can easily explain previously reported differences

Butler-Laporte *et al. Clinical Proteomics* (2022) 19:34

Page 8 of 11

Fig. 3 Smoothed protein level curves showing time-related and sex-related differences as a function of days since symptoms onset (x-axis) in a 65-year-old patient (p < 3.4 × 10⁻⁴ for sex differences in cytokine levels). Y-axis is standardized to a mean of 0 and standard deviation of 1 F: female. M: male. Blue: controls. Red: severe COVID-19

in direction of associations [11]. Our large sample and careful adjustments for multiple comparisons also likely avoided spurious associations.

Other studies have assessed at protein levels in acute COVID-19, and we share many of the same observations already made. For example, severe COVID-19 was linked to changes in IL-13 [3], IL6 and IL-1B [4], and multiple chemokines (e.g. CCL20, CCL27, CXCL10 [28]). Hence, while it is clear that immune-related proteins levels are associated with outcomes, differences in methodology led to varying observations. For example, Lucas et al [3] observed differences in IL-1B, IL-6, IL-18, and TNF-α between severe and non-severe individuals using the Eve Technologies (Calgary, Alberta, Canada) Luminex based HD71 assay. However, using a different Luminex-based assay, Wilson et al [29] found no such difference in these 4 cytokines between severe COVID-19 cases and

non-COVID-19 sepsis controls. Different results were again obtained from Filbin et al [30] who used the Olink (Uppsala, Sweden) multiplex antibody-oligonucleotide assay to highlight IL6, IL-1RL1, and IL-1RN's role in severe COVID-19. As mentioned above, these differences can likely be explained by either small sample sizes, insufficient control for time of onset of symptoms, and different choices of cases and controls. Indeed, we replicated the IL6, IL-1RL1, and IL-1RN results from Filbin et al. which is to our knowledge the previously largest proteomics study on acute Covid-19. This study used similar methods to ours but adjusted for day of hospitalization rather than onset of symptoms. Comparisons to other studies should therefore also keep these methodological differences in mind.

Less is known about the role of immune-related proteins in COVID-19 outcome sex differences. A previous

Butler-Laporte *et al. Clinical Proteomics* (2022) 19:34

Page 9 of 11

report [12] suggested a role for IL8 and IL18, but these were not replicated in other studies [13]. Our study is the first to report on difference in TLR5, CXCL17, CCL26, IL1RL2, or IL3RA levels in sexes during infection. While the mechanism by which they could influence outcomes is unclear, it is worth noting that the gene encoding for IL3RA is located on the X chromosome, providing a plausible explanation for the observed difference. Further CCL26 (also known as eotaxin-3) is known to induce eosinophils tissue infiltration [31], which could influence COVID-19 outcomes [32]. However, multiple studies have shown differences in cellular immune responses, COVID-19 specific antibody levels, and many commonly measured inflammatory markers in clinical practice (e.g. C-reactive protein) [12, 33]. Hence, it remains possible that another immune pathway that was not measured by our panel might be involved in the observed sex differences in outcomes. However, our observations on TLR5, CXCL17, CCL26, IL1RL2, and IL3RA provide clear proteins to explore to explain sex differences in COVID-19 outcomes.

Nevertheless, our study still has limitations. First, while we assayed proteins in the first collected samples, it remains possible that some subjects received immunomodulatory drugs (e.g. dexamethasone) which would have affected protein levels. However, this would likely attenuate the differences between the cases and the controls, and our results would therefore be biased towards the null hypothesis. Second, given that protein time trends were obtained using multiple different subjects, unmeasured confounders could explain some of our findings. While these cannot be easily measured, it is reassuring that our results replicated across two cohorts, arguing against the presence of confounders with large effect sizes. Third, the control group made up of non-severe COVID-19 participants as well as non-COVID-19 disease may have biased some of our results towards the null. However, this specific choice of control arm made our results more specific for severe COVID-19, rather than critical illness in general, and we still found clear associations with many immune-related proteins. Fourth, the use of SOMAscan may make comparisons difficult with other studies using different protein measurement technologies. Despite this, SOMAscan showed great sensitivity, specificity, and reproducibility when benchmarked against mass spectrometry [34], and our conclusions are unlikely to be greatly biased by the choice of protein measuring platform. Lastly, while this is one of the largest panels of immune-related proteins studied for COVID-19, there are multiple proteins that were not measured, and we cannot assess whether other unmeasured proteins may also have important effects on the outcomes.

In conclusion, using two large independent cohorts with broad protein measurements, we showed that severe COVID-19 was associated with clear time-dependent changes in multiple immune-related proteins, and that these may in part explain difference in COVID-19 outcomes between sexes.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12014-022-09371-z.

Additional file 1: List of immunity-related proteins measured.

Additional file 2: Sample code for the generalized additive models.

Additional file 3: Protein correlation heatmaps.

Additional file 4: Protein clusters.

Additional file 5: Inferred protein levels over time.

Additional file 6: Mount Sinai investigators.

Author contributions

Conception and design: GBL, JBR. Data analyses: GBL, EGK, CYS, SZ. Data acquisition: TN, GBL, DM, DEK, JA, MA, LL, EBR, DH, NK, ZA, NR, MB, LP, CG, XX, CT,

Butler-Laporte *et al. Clinical Proteomics* (2022) 19:34

Page 10 of 11

Author details
[1]Lady Davis Institute, Jewish General Hospital, McGill University, Montréal, Québec, Canada. [2]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, Québec, Canada. [3]Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [4]Department of Computer Science, McGill University, Montréal, Québec, Canada. [5]Department of Human Genetics, McGill University, Montréal, Québec, Canada. [6]Graduate School of Medicine, McGill International Collaborative School in Genomic Medicine, Kyoto University, KyotoKyoto, Japan. [7]Japan Society for the Promotion of Science, Tokyo, Japan. [8]Research Centre of the Centre Hospitalier de L'Université de Montréal, Montréal, Québec, Canada. [9]Department of Emergency Medicine, Jewish General Hospital, McGill University, Montréal, Québec, Canada. [10]MRC Epidemiology Unit, School of Clinical Medicine, University of Cambridge, Cambridge, UK. [11]Computational Medicine, Berlin Institute of Health at Charité—Universitätsmedizin Berlin, Berlin, Germany. [12]Infectious Diseases and Immunity in Global Health Program, Research Institute of the McGill University Health Centre, Montréal, Québec, Canada. [13]Precision Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [14]Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [15]Department of Radiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [16]Human Immune Monitoring Center, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [17]SomaLogic Inc, Boulder, CO, USA. [18]Early Phase Trials Unit, Mount Sinai Hospital, New York, NY, USA. [19]Mount Sinai Clinical Intelligence Center, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [20]Department of Medicine, Université de Montréal, Montréal, Québec, Canada. [21]Department of Twin Research, King's College London, London, UK. [22]5 Prime Sciences, Montreal, Québec, Canada. [23]McGill University, King's College London (Honorary), Jewish General Hospital, Pavilion H-4133755 Côte-Ste-Catherine, Montréal, Québec H3T 1E2, Canada.

References
1. Tang Y, Liu J, Zhang D, Xu Z, Ji J, Wen C. Cytokine storm in COVID-19: the current evidence and treatment strategies. Front Immunol. 2020;11:1708.
2. Sinha P, Matthay MA, Calfee CS. Is a "cytokine storm" relevant to COVID-19? JAMA Intern Med. 2020. https://doi.org/10.1001/jamainternmed.2020.3313.
3. Lucas C, Wong P, Klein J, et al. Longitudinal analyses reveal immunological misfiring in severe COVID-19. Nature. 2020;584:463–9.
4. Del Valle DM, Kim-Schulze S, Huang H-H, et al. An inflammatory cytokine signature predicts COVID-19 severity and survival. Nat Med. 2020;26:1636–43.
5. Blanco-Melo D, Nilsson-Payant BE, Liu W-C, et al. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. Cell. 2020;181:1036-1045.e9.
6. Contoli M, Papi A, Tomassetti L, et al. Blood interferon-α levels and severity, outcomes, and inflammatory profiles in hospitalized COVID-19 patients. Front Immunol. 2021;12:536.
7. Xiao N, Nie M, Pang H, et al. Integrated cytokine and metabolite analysis reveals immunometabolic reprogramming in COVID-19 patients with therapeutic implications. Nat Commun. 2021;12:1618.
8. Ebihara T, Matsumoto H, Matsubara T, et al. Cytokine elevation in severe COVID-19 from longitudinal proteomics analysis: comparison with sepsis. Front Immunol. 2021;12:798338.
9. Lee JS, Shin E-C. The type I interferon response in COVID-19: implications for treatment. Nat Rev Immunol. 2020;20:585–6.
10. Hadjadj J, Yatim N, Barnabei L, et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science. 2020;369:718–24.
11. da Silva RP, Gonçalves JIB, Zanin RF, Schuch FB, de Souza APD. Circulating type I interferon levels and COVID-19 severity: a systematic review and meta-analysis. Front Immunol. 2021;12:1717.
12. Takahashi T, Ellingson MK, Wong P, et al. Sex differences in immune responses that underlie COVID-19 disease outcomes. Nature. 2020;588:315–20.
13. Lau ES, McNeill JN, Paniagua SM, et al. Sex differences in inflammatory markers in patients hospitalized with COVID-19 infection: Insights from the MGH COVID-19 patient registry. PLoS ONE. 2021;16: e0250774.
14. Jing M, Bowser MT. Methods for measuring aptamer-protein equilibria: a review. Anal Chim Acta. 2011;686:9–18.

Butler-Laporte *et al. Clinical Proteomics* (2022) 19:34

Page 11 of 11

15. Kraemer S, Vaught JD, Bock C, et al. From SOMAmer-based biomarker discovery to diagnostic and clinical applications: a SOMAmer-based, streamlined multiplex proteomic assay. PLoS ONE. 2011;6:e26332–e26332.

16. Tremblay K, Rousseau S, Zawati MH, et al. The Biobanque québécoise de la COVID-19 (BQC19)—a cohort to prospectively study the clinical and biological determinants of COVID-19 clinical trajectories. PLoS ONE. 2021;16:e0245031.

17. Mogensen TH, Paludan SR. Molecular Pathways in Virus-Induced Cytokine Production. Microbiol Mol Biol Rev. 2001;65:131–50.

18. Koroleva EP, Fu Y-X, Tumanov AV. Lymphotoxin in physiology of lymphoid tissues—implication for antiviral defense. Cytokine. 2018;101:39–47.

19. Kutza J, Crim L, Feldman S, et al. Macrophage colony-stimulating factor antagonists inhibit replication of HIV-1 in human macrophages. J Immunol. 2000;164:4955–60.

20. Rosa M, Chignon A, Li Z, et al. A Mendelian randomization study of IL6 signaling in cardiovascular diseases, immune-related disorders and longevity. Npj Genomic Med. 2019;4:23.

21. R Core Team (2020). R: A language and environment for statistical computing. 2020. https://www.r-project.org/.

22. Wood SN. Generalized additive models: an introduction with R. 2nd ed. Boca raton: Chapman and Hall/CRC; 2017. https://doi.org/10.1201/97813 15370279.

23. Sørensen Ø, Brandmaier AM, Macià D, et al. Meta-analysis of generalized additive models in neuroimaging studies. Neuroimage. 2021;224: 117416.

24. Dewey M. metap: meta-analysis of significance values. R package version 1.4. 2020.

25. Yang J, Zheng Y, Gou X, et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. Int J Infect Dis. 2020;94:91–5.

26. Lamertz L, Rummel F, Polz R, et al. Soluble gp130 prevents interleukin-6 and interleukin-11 cluster signaling but not intracellular autocrine responses. Sci Signal. 2018;11:7388.

27. Gour N, Wills-Karp M. IL-4 and IL-13 signaling in allergic airway disease. Cytokine. 2015;75:68–78.

28. Quartuccio L, Fabris M, Sonaglia A, et al. Interleukin 6, soluble interleukin 2 receptor alpha (CD25), monocyte colony-stimulating factor, and hepatocyte growth factor linked with systemic hyperinflammation, innate immunity hyperactivation, and organ damage in COVID-19 pneumonia. Cytokine. 2021;140: 155438.

29. Wilson JG, Simpson LJ, Ferreira A-M, et al. Cytokine profile in plasma of severe COVID-19 does not differ from ARDS and sepsis. JCI Insight. 2020. https://doi.org/10.1172/jci.insight.140289.

30. Filbin MR, Mehta A, Schneider AM, et al. Longitudinal proteomic analysis of severe COVID-19 reveals survival-associated signatures, tissue-specific cell death, and cell-cell interactions. Cell Reports Med. 2021. https://doi.org/10.1016/j.xcrm.2021.100287.

31. Bhattacharya B, Carlsten J, Sabo E, et al. Increased expression of eotaxin-3 distinguishes between eosinophilic esophagitis and gastroesophageal reflux disease. Hum Pathol. 2007;38:1744–53.

32. Lindsley AW, Schwartz JT, Rothenberg ME. Eosinophil responses during COVID-19 infections and coronavirus vaccination. J Allergy Clin Immunol. 2020;146:1–7.

33. Huang B, Cai Y, Li N, et al. Sex-based clinical and immunological differences in COVID-19. BMC Infect Dis. 2021;21:647.

34. Gold L, Ayers D, Bertino J, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. PLoS ONE. 2010;5: e15004.

Publisher's Note