Title

Immunoprofiles of COVID-19 uniquely differentiated from other viruses: A machine learning investigation of multiplex immunoassay data.

Journal

PNAS Nexus, 3(8)

Authors

Kaur, Ashneet

Krishnan, Viswanathan

Publication Date

2024-08-01

DOI

10.1093/pnasnexus/pgae327

PNAS Nexus, 2024, 3, pgae327
https://doi.org/10.1093/pnasnexus/pgae327
Advance access publication 7 August 2024
Research Report

Immunoprofiles of COVID-19 uniquely differentiated from other viruses: A machine learning investigation of multiplex immunoassay data

Ashneet Kaur[a] and Viswanathan V. Krishnan[b,c,*]

[a]Department of Biology, California State University, Fresno, CA 93740, USA
[b]Department of Chemistry and Biochemistry, California State University, Fresno, CA 93740, USA
[c]Department of Pathology and Laboratory Medicine, University of California Davis School of Medicine, Sacramento, CA 95817, USA
*To whom correspondence should be addressed: Email: krish@csufresno.edu, vvkrishnan@ucdavis.edu
Edited By Bruce Levine

Abstract

Cytokines and chemokines are vital in maintaining a healthy state by efficiently controlling invading microbes. In addition, the dysregulation of these immune mediators can contribute to viral infection pathology. We comprehensively analyzed the profiles of host immunomodulators in response to infections with members of several virus families, particularly if the SARS-CoV-2 infection produces a unique immune profile compared with other viral infections. Multiplex microbead immunoassay results from 219 datasets with a range of viruses were curated systematically. The curated immunoassay data, obtained using Luminex technology, include 35 different viruses in 18 different viral families; this analysis also incorporated data from studies performed in 7 different cell model systems with 28 different sample types. A multivariate statistical analysis was performed with a specific focus involving many investigations (>10), which include the viral families of Coronaviridae, Orthomyxoviridae, Retroviridae, Flaviviridae, and Hantaviridae. A random forest-based classification of the profiles indicates that IL1-RA, C-X-C motif chemokine ligand 9, C-C motif chemokine ligand 4, interferon (IFN)-λ1, IFN-γ-inducing protein 10, and interleukin (IL)-27 are the top immunomodulators among human samples. Similar approaches only between Coronaviridae (COVID-19) and Orthomyxoviridae (influenza A/B) indicated that transforming growth factor-β, IFN-λ1, IL-9, and eotaxin-1 are important features. In particular, the IFN-λ1 protein was implicated as one of the significant immunomodulators differentiating viral family infection. It is evident that Coronaviridae infection, including SARS-CoV-2, induces a unique cytokine–chemokine profile and can lead to specific immunoassays for diagnosing and prognosis of viral diseases based on host immune responses. Alternatively, we can use diagnosing and prognosing. It is also essential to note that meta-analysis-based predictions must be appropriately validated before clinical implementation.

Significance Statement

The research aimed to compare the immune response to SARS-CoV-2 infection with other viruses. After analyzing almost 200 studies and using data mining techniques on multiplex microbead immunoassay data, the results indicate that the Coronaviridae viral family, including SARS-CoV-2, exhibits a cytokine–chemokine profile that distinguishes it from other viruses. This distinct immunological signature can potentially develop targeted immunoassays, which could help diagnose and predict the progression of COVID-19 by assessing immune system activity.

Introduction

The immune response to a viral infection triggers a cascade of complex cytokines, chemokines, and other immunomodulating molecules. These circulating molecules, which have proinflammatory or antiinflammatory activities, are classified into interleukins (ILs), lymphokines, chemokines, and other mediators of cell signaling (1). Immunomodulators are integral in many functions, including innate immunity, antigen presentation, cellular recruitment, and adhesion molecule expression. Therefore, a comprehensive understanding of cytokine activity has been crucial in immunology and inflammation research on host responses to infection and other diseases.

Viruses, such as coronaviruses and influenza viruses, trigger the secretion of specific immune modulators, cytokines/chemokines in response to infection. The immune system's secretion of these proteins is essential in directing innate and adaptive immune responses to infections. These molecules play a crucial role in recruiting other immune cells to the site of infection and can thus promote the clearance of infected cells. The immune response mediated by cytokines is determined by a highly regulated and efficient cytokine

Competing Interest: The authors declare no competing interests.
Received: March 7, 2024. **Accepted:** July 22, 2024

production and a balance between proinflammatory and antiinflammatory cytokines. However, an excessive or dysregulated production of proinflammatory cytokines—designated "cytokine storm"—has also been observed in individuals with severe disease and has captured the public and scientific community's attention, particularly concerning COVID-19. At the same time, the concept of pathogenesis due to immune system dysregulation is well established (2–6). Overall, the complex interplay among cytokines, chemokines, and other immunomodulating agents is critical in determining the outcome of viral infections and the effectiveness of the immune response against them.

Recent literature provides many studies that focus on how the profiles of the immunomodulators are particularly effective in relating specific manifestations of COVID-19 in patients. The knowledge acquired on the pathophysiology of the immunomodulators in recent years is vast, and a "title" search of PubMed utilizing the terms "cytokine" and "COVID-19" yielded >200 citations. While most studies focus mainly on the immunoprofiles of patients with COVID-19, our study compared the immune profiles of COVID-19 infection with those of other viral infections. The cytokine–chemokine profiles related to a particular infection can indicate how the host immune system responds to a viral insult. Accordingly, we hypothesize that these differential profiles might lead to a better understanding of the various degrees of acuteness, severity, and persistence essential to understanding the pathophysiology of COVID-19. Furthermore, the results have the potential to form a basis for newer approaches to disease diagnosis and prognosis, as well as treatments.

Multiplex microbead immunoassay data obtained using Luminex technology were systematically collected from peer-reviewed publications. Fold changes of the immunomarkers calculated with reference to a healthy control group from each dataset were used to construct the profiles from 219 independent experiments. With a focus on studies predominantly in human samples and aiming to increase the statistical power of the analysis, data from five virus families (Coronaviridae, Orthomyxoviridae, Retroviridae, Flaviviridae, and Hantaviridae) were selected. A data mining investigation using the Random Forest classification was performed to identify the cytokines and chemokines that differentiate host responses to specific viral families. Random forest classification and predictive analytics were applied to all the sample types and then narrowed to comparing infection with Coronaviridae and Orthomyxoviridae. The results of the prospective study indicate that cytokine–chemokine profiles provide an understanding of the differential host response to COVID-19 compared with other viral infections.

Results

Description of the data

Sixty-eight publications were included in the meta-analysis of multiple immunomodulators from studies on 35 viruses from 18 viral families. The number of studies that measured these analytes is shown in a bar plot as a function of the viral families (Fig. S1a). Table 1 lists the viral families with the corresponding list of viruses in each family, indicating that the most significant number of studies were conducted on infection with the viral families Orthomyxoviridae (influenza A virus and influenza B virus), Retroviridae (human immunodeficiency virus, human T-cell leukemia virus type 1, simian immunodeficiency virus, xenotropic murine, and xenotropic murine leukemia virus–related virus), and Coronaviridae (SARS-CoV-2, coronavirus NL63/229E, coronavirus OC43, and feline coronavirus), resulting in 42, 42, and 41 datasets, respectively. The virus families, Arteriviridae, Flaviviridae, and Herpesviridae, are represented by 20, 20, and 12 datasets, respectively, whereas the remainder contain <10 datasets each (Fig. S1a). Most of the studies are related to human infections (120 datasets), followed by infections in mice (50 datasets) and cell culture systems (22 sets; Fig. S1b). Immunomodulator data were also obtained from several studies on Rhesus monkeys, African green monkeys, cats, and pigs. Within each system, a wide variety of samples were studied, with serum (64 datasets) and plasma (56 datasets) topping the list; whole PBMCs and other cell types followed (Fig. S1c). Interestingly, tear samples have also been used (7).

Table 1. Number of studies on each virus and viral families.

Viral family	Virus species	Number of studies
Coronaviridae	SARS-CoV-2	14
	Coronavirus ML63/229E	2
	Coronavirus OC43	1
	Feline coronavirus	1
Retroviridae	Human immunodeficiency virus	12
	Human T-cell leukemia virus type 1	1
	Simian immunodeficiency virus	2
	Xenotropic murine leukemia virus–related virus	1
Orthomyxoviridae	Influenza A virus	8
	Influenza B virus	2
Flaviviridae	Dengue virus	2
	Hepatitis C virus	2
	West Nile virus	1
	Zika virus	3
Hantaviridae	Hantavirus	2
Adenoviridae	Adenovirus	1
Anelloviridae	Torque Tino mini viruses	1
Arteriviridae	Porcine reproductive and respiratory syndrome virus	1
Filoviridae	Ebola virus	1
Hepadnaviridae	Hepatitis B virus	1
Herpesviridae	Cytomegalovirus	3
	Herpes simplex virus	4
	Kaposi's sarcoma-associated herpesvirus	1
	Varicella-Zoster virus	2
Matonaviridae	Rubella virus	1
Papillomaviridae	Human papillomavirus	2
Paramyxoviridae	Parainfluenza virus	3
Parvoviridae	Parvovirus B19	1
Picornaviridae	Coxsackie B virus	1
	Enterovirus	1
	Human parechovirus	1
	Rhinovirus	4
Pneumoviridae	Human metapneumovirus	1
	Respiratory syncytial virus	4
Togaviridae	Chikungunya virus	1

The curated experimental data contains a range of immunomodulators, encompassing 165 cytokines and chemokines (Table 2). Multiple types of analytes in the data were ILs, interferons (IFNs), tumor necrosis factors (TNFs), colony-stimulating factors (CSFs), and chemokines—C-C motif, X-C motif, C-X-C motif, and C-X3-C motif chemokines. The data also included growth factors, inhibitory factors, and other proteins with cytokine activities. The choice of the set of immunomodulators measured varies widely among the publications due to the selection of the panel of analytes chosen for the study.

A comprehensive view of the immunomodulators in the top 10 viral families (Orthomyxoviridae, Retroviridae,

Table 2. List of the classes of immunomodulators.

Analyte type	Analyte name
Interleukins (ILs)	IL-1, IL-1α, IL-1β, IL-1RA, IL-1R1, IL-1R2, IL-2, IL-2RA, IL-3, IL-4, IL-4RA, IL-5, IL-6, IL-6RA, IL-6RB, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-12(p40), IL-12(p70), IL-13, IL-15, IL-16, IL-17, IL-17B, IL-17F, IL-18, IL-18BP, IL-19, IL-20, IL-21, IL-22, IL-23, IL-25, IL-26, IL-27, IL-28A, IL-31, IL-32, IL-33, IL-34, IL-35
Interferons (IFNs)	IFN-α, IFN-α2, IFN-β, IFN-γ, IFN-λ1, IFN-λ2
Tumor necrosis factors (TNFs)	TNF-α, TNF-β, TNFSF5, TNFSF6, TNFSF10, TNFSF12, TNFSF13, TNFSF13B, TNFSF14, TNFRSF1A, TNFRSF1B, TNFRSF7, TNFRSF8, TNFRSF11B, Trail-R3
Colony-stimulating factors (CSFs)	GMCSF, GCSF, MCSF, CSF
Chemokines	C-C motif chemokines: eotaxin-1, eotaxin-2, eotaxin-3, CCL1, CCL2, CCL3, CCL4, CCL7, CCL8, CCL12, CCL13, CCL15, CCL16, CCL17, RANTES, CCL18, CCL19, CCL20, CCL21, CCL22, CCL23, CCL25, CCL27, CCL28 X-C motif chemokines: CXCL1 C-X-C motif chemokines: IP-10, CXCL1, CXCL2, CXCL5, CXCL6, CXCL9, CXCL11, CXCL12, CXCL13, CXCL16, stromal cell–derived factor-1alpha + beta C-X3-C motif chemokines: CX3CL1
Growth Factors	bFGF, aFGF, HGF, PDGF, PDGF-AA, PDGF-BB, endothelin 1, VEGF-A, VEGF-C, VEGF-D, Flk-1, TGF-α, TGF-β, TGF-β1, epidermal growth factor (EGF), nerve growth factor, FLT-1, FLT-4, serum stem-cell growth factor-beta, bone morphogenetic protein 9, heparin-binding EGF-like growth factor, EGFR, PGF
Inhibitory Factors	Leukemia inhibitory factor, macrophage migration inhibitory factor, plasminogen activator inhibitor 1, tissue inhibitor of metalloproteinases (TIMP) 1, TIMP-2, TIMP-3, TIMP-4
Others	Serum soluble Fas, C-reactive protein (CRP), monomeric-CRP, Fms-like tyrosine kinase 3 ligand, cluster of differentiation (CD)-14, CD40L, CD-62P, CD163, activated leukocyte cell adhesion molecule, Pentraxin 3, stem-cell factor, Pan GRO, thymic stromal lymphopoietin, thrombopoietin, OPN, CHI3L protein, OCN, TIE2, angiopoietin (ANGPT) 1, ANGPT2, leptin, lipocalin 2, resistin, adipsin, adiponectin, RAGE, DCGF-beta, follistatin

Coronaviridae, Arteriviridae, Flaviviridae, Herpesviridae, Hantaviridae, Picornaviridae, Filoviridae, and Pneumoviridae) is shown in the upset interaction plot (Fig. S3). The bar plot (Fig. S2a) indicates that the data are highly heterogeneous due to the choice of the model system (human, mouse, porcine, rhesus monkey, African green monkey, feline, and various cell lines) and different sample types. Human samples constitute the majority, followed by those from cell-line model experiments. Twenty-eight different types of samples were collected to measure analytes across all model systems (Fig. S2b).

The data indicate that all COVID-19 studies have exclusively focused on human subjects, utilizing plasma or serum samples. Thus, comparing the immunoprofiles among these human samples holds significant relevance. In addition, the survey, based on the number of independent datasets (Fig. S3), shows a bias for the viral families Coronaviridae, Retroviridae, Flaviviridae, and Hantaviridae. However, considering the importance of how influenza infection might need to be differentiated, particularly with respect to the coronavirus family (8, 9), a detailed analysis of the profiles of these five viral families, including Orthomyxoviridae, is necessary. Therefore, Fig. 1 shows the upset plot of the five different viral families (Coronaviridae, Orthomyxoviridae, Retroviridae, Flaviviridae, and Hantaviridae) and the corresponding number of immunomodulators studied. Though many immunomodulators (158) have been reported, the number of analytes studied for all five viral families (in at least one study) is 27. The immunomodulators are defined as the "common set": basic fibroblast growth factor (bFGF), C-C motif chemokine ligand 2 (CCL2), CCL3, CCL4, X-C motif chemokines ligand 1 (CXCL1), CXCL9, eotaxin-1, GCSF, granulocyte-macrophage CSF (GMCSF), IFN-α, IFN-γ, IL-10, IL-12p40, IL-12p70, IL-13, IL-15, IL-17, IL-1b, IL-1RA, IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IFN-γ-inducing protein 10 (IP-10), platelet-derived growth factor (PDGF)-BB, RANTES, TNF-α, and vascular endothelial growth factor A (VEGF-A). Figure 1 also indicates that 48 immunomodulators were not investigated for the Coronaviridae and Orthomyxoviridae viral families but in all 3 other families.

Immunoprofiles of the top five viral families

The cytokine profiles of responses to infection with the Coronaviridae, which includes SARS-CoV-2, comprised a total of 122 analytes when all of the studies were taken together (Fig. 2a). For the Coronaviridae viral family, 12 immunomodulators are highly up-regulated (elevated by >4-fold), namely—acidic-FGF (aFGF), CCL8, CRP, CXCL11, Flt-1, hepatocyte growth factor (HGF), IL-17F, IL-2RA, IP-10, MCSF, SCDF-β, and transforming growth factor (TGF)-α. IFN-λ2, a type III IFN, was the only cytokine down-regulated >4-fold (Fig. 2a). The SDs for some of the analytes are large due to differences among the experimental system, sample conditions, and sources of the reagents. The horizontal dotted line shows an increase or decrease in the levels of these immunomodulators by a $\log_2$-fold change value of two or more with reference to healthy control.

For comparison, the immune profiles of the other four viral families (Orthomyxoviridae, Retroviridae, Flaviviridae, and Hantaviridae) are also shown in Fig. 2b–e. The immune profile represented by influenza A/B, another respiratory virus belonging to the Orthomyxoviridae virus family, has a total of 29 analytes (Fig. 2b). Out of these 29 analytes, 5 were elevated (by >4-fold), namely, IL-6, IL-8, IL-9, IP-10, and TGF-β, to the healthy controls. Similar to the Coronaviridae, a type III IFN, IFN-λ1, was down-regulated in Orthomyxoviridae-infected humans compared with healthy human controls.

The third immune profile in Fig. 2 illustrates the fold change of immunomodulators released in response to the Retroviridae, predominantly HIV. For the *Retroviridae*, different analytes were measured; nine were elevated or down-regulated across all model systems—CCL4, IL-1β, IL-9, placental growth factor (PGF), RANTES, TNF superfamily factor 5 (TNFSF5), IL-12p40, IL-7, and PDGF (data not shown). As for human cell models, 3 cytokines (IL-9, PGF, and TNFSF5) were elevated >4-fold, and 2 analytes (IL-12p40 and PDGF) were down-regulated >4-fold compared with healthy human controls (Fig. 2c).

For the Hantaviradae and Flaviviridae, profiles are also presented in Fig. 2d and e, respectively, with 103 and 79

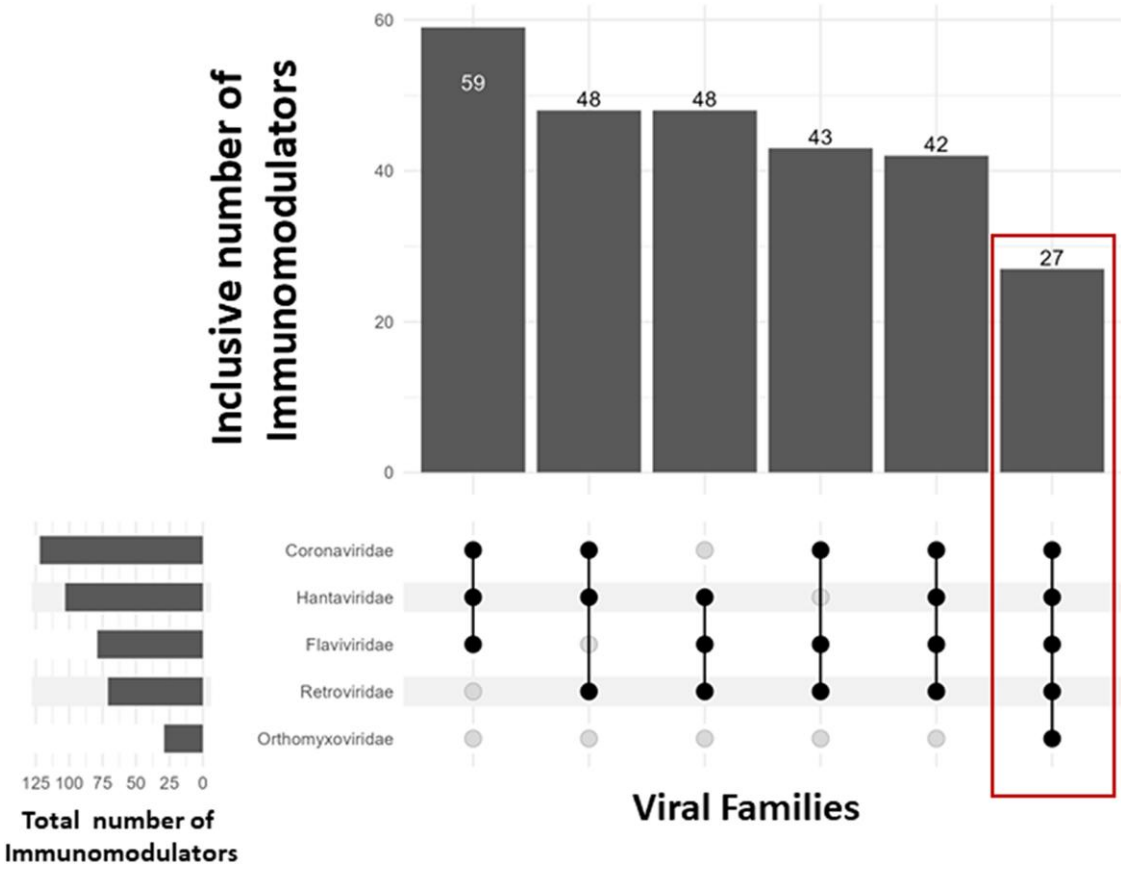

Fig. 1. Interaction (UpSet) plot selective viral families. The UpSet plot describes the common, inclusive number of immunomodulators in the dataset. The number of immunomodulators in each of the five viral families is plotted as horizontal bars, and the number of inclusively common analytes is shown as vertical bars with corresponding interactions in the matrix below. For example, the boxed red region indicated that 27 analytes are inclusively shared in all the studies with the 5 viral families listed.

immunomodulators. In Hantaviradae-infected humans, 17 analytes were highly elevated when compared with healthy humans namely, CCL2, CCL3, CCL4, CXCL9, C-X3-C motif chemokine ligand 1 (CX3CL1), IFN-λ1, IFN-λ2, IL-11, IL-15, IL-22, IL-26, IL-27, IL-1β, IL-32, IL-34, IL-6RB, and TNF-α. The analytes that were highly down-regulated (>4-fold) when compared with normal controls were CCL17, CCL22, chitinase 3 like 1 (CHI3L1), IL-6RA, osteocalcin (OCN), osteopontin (OPN), TNF receptor superfamily factor 1A (TNFRSF1A), TNFRSF1B, TNFRSF8, TNFSF12, TNFSF5, and TNFSF13B. The cytokine profile of Flaviviridae (Fig. 2e) consisted of 79 cytokines and chemokines, out of which 11 were elevated (>4-fold)—CCL4, CXCL9, eotaxin-3, IFN-β, IL-10, IL-20, IL-6, IL-8, IP-10, TNFRSF1B, and TNFRSF8—and 1 analyte, CXCL12, was highly down-regulated in Flaviviridae infection compared with healthy controls.

Classification of immunoprofiles

To understand the broader role of the individual immunomodulator profiles in infection with the viral families, a partial least squares-discriminant analysis (PLS-DA) was performed (Fig. 3). Immunomodulators studied only in the human samples were considered for the classification. The greatest number of experiments is represented by the Coronaviridae, consisting of 39 studies encompassing 122 immunomodulators, followed by Retroviridae and Flaviviridae, respectively, with 27 studies (70 analytes) and 20 studies (78 analytes). Human studies on the Orthomyxoviridae were low, with 4 studies comprising 22 analytes, and Hantaviridae studied with a limited number of 9 studies but with a variety of 103 analytes. The discriminatory power of the immunomodulators of the Coronaviridae and Hantaviridae may also be due to the number of unique sets of analytes measured in these respective experiments. Coronaviridae experiments measured 35 unique analytes, while Hantaviridae only had 16 unique cytokines and chemokines observed in Hantavirus infections compared with other viruses in the study. The number of components to be used in the discrimination between the immunoprofiles of the different viral infections dictates the performance of the PLS-DA model to correct classification (10). Cross-validation (5-fold × 10 repeats) was performed, and the overall error rate and balanced error rate for the three different distance metrics (Max, Centroid, and Mahalanobis) across the first 10 components were calculated (Fig. S4). Figure S4 indicates that the discriminatory power of the data is diminished significantly after three components. Both Coronaviridae and Hantaviridae show the largest discrimination in the levels of the immunomodulators (Fig. 3), while the other viral families (Orthomyxoviridae, Retroviridae, and Flaviviridae) exhibit the slightest discrimination. The overall classification of the differential profiles of the immunomodulators is modest and thus encourages a detailed analysis of the individual changes in the analytes between the viral infections.

Differential fold changes of immunomodulators

The curated data (see Experimental design and methodology) consist of fold changes calculated in comparison with the corresponding healthy group from the same dataset; all the analysis was aimed to estimate the "differential fold change" (ratio of the

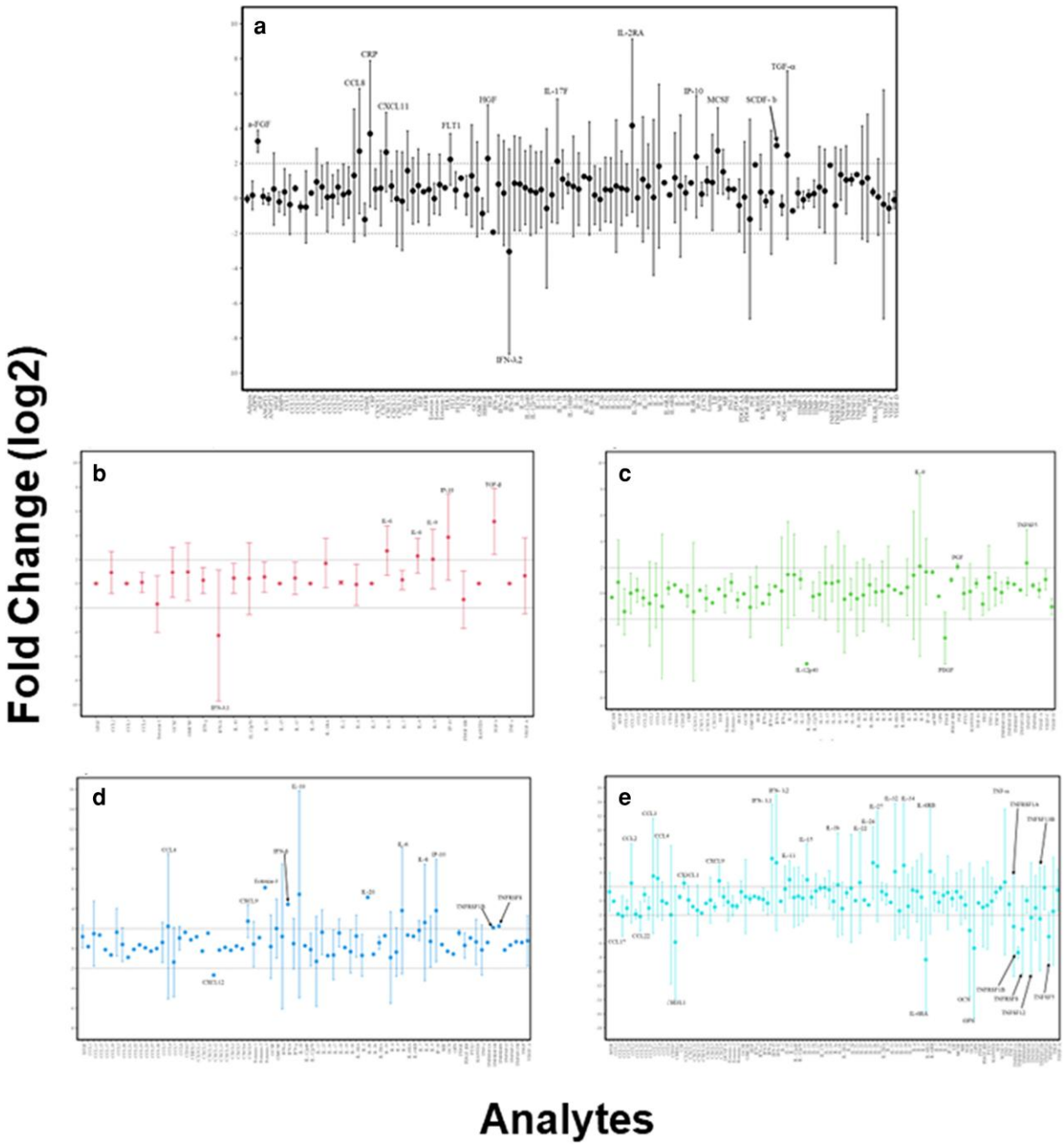

Fig. 2. Immunomodulator profiles of viral families: fold changes of the viral modulator profiles measured in human samples for a) Coronaviridae, b) Orthomyxoviridae, c) Retroviridae, d) Flaviviridae, and e) Hantaviridae. The mean values of each analyte are plotted with the error bars representing the SDs across all the measurements. The fold changes (log_2) were calculated with respect to healthy controls under the same experimental conditions. The horizontal dashed lines show positive or negative 2-fold changes (log_2), with some of the analytes above the 2-fold changes indicated in the respective plots. It is important to note that the error bars are large as the data come from multiple studies in the meta-analysis.

fold changes) and corresponding statistical significance. A multivariate analysis of the experimental data of the five viral families was performed to identify the analytes with significant differential fold changes (fold change differences between two treatments). Twenty-eight immunomodulators were found to be significantly altered for at least one of the four viral families (Retroviridae, Flaviviridae, Hantaviradae, and Orthomyxoviridae) in comparison with the Coronaviridae (Table 3). Violin plots of the fold change distributions of a select set of immunomodulators are shown in Fig. 4 (a: IL1-RA; b: CXCL9; c: CCL4; d: IFN-λ1; e: IP-10; f: IL-8; g: IL-10; and h: IL-27). Interestingly, infections with Retroviridae did not show many differentially regulated analytes; Hantaviradae infections showed the greatest number of analytes with a high differentiation. For example, IL-6RA (Table 3) exhibits the largest positive differentiation for Hantaviradae in comparison with Coronaviridae (log_2(5.6) ~47-fold), while IFN-λ1 (Fig. 4d) shows the most significant negative differentiation (–log_2(4) ~16-fold reduction). IL-8 (Table 3) has a negative differential change of 4 (–log_2(1.9)) in Coronaviridae, while IFN-λ1 shows an almost 16 times increase in Orthomyxoviridae (log_2(4.3)) (Table 3) family. The immunomodulators IL-10, IL-8, IP-10, and CCL4 (Table 3 and Fig. 4) show negative differentiation in Flaviviridae to Coronaviridae by more than three times. The results indicate that the immunomodulator responses to various viral

families tend to be selective, except for a few exceptions. IFN-λ1 indicates a robust differential response between Orthomyxoviridae (positive) and Hantaviradae (negative) (Fig. 4d). In contrast, IL-8 shows a similar response for Orthomyxoviridae and Flaviviridae (Fig. 4f), whereas CCL4 exhibits a similar response in Flaviviridae and Hantaviridae compared with the Coronaviridae (Fig. 4c). Coronaviridae studies measured 35 unique analytes, while Hantaviridae examined 16 unique cytokines and chemokines observed in Hantavirus infections compared with other viruses in the study (Fig. 3). The overall classification of the differential profiles of the immunomodulators supported a detailed analysis of the individual changes in the analytes.

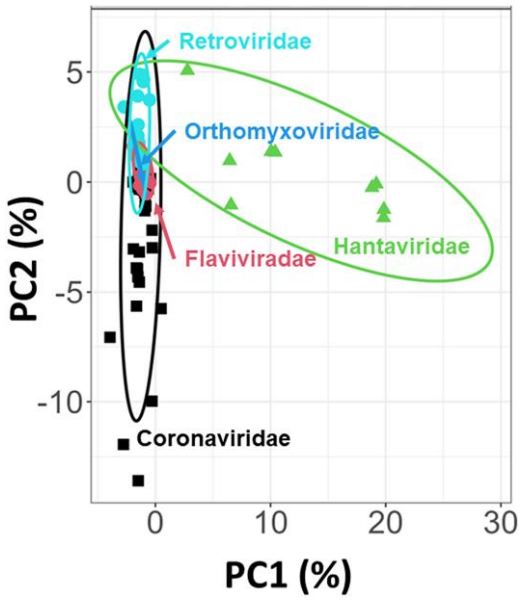

Fig. 3. PLS-DA of the immunomodulators. The first two components are plotted with the five viral families Coronaviridae (black), Orthomyxoviridae (blue), Retroviridae (teal), Flaviviridae (red), and Hantaviridae (green), identified with different color symbols. The ellipses show the 90% confidence in the classification.

Predictive immunomodulators by machine learning

A random forest-based feature selection algorithm is used to identify features (immunomodulators) that are significant for differentiating between the five viral families (11, 12). The Boruta approach generates shadow datasets from the original dataset to identify specific features from the immunoprofiles of the viral families as defined by "importance" by the algorithm. The feature selection process was performed with several starting datasets. When all 28 immunomodulators showed differential fold changes in at least 1 of 4 viral families in comparison with Coronaviridae (Table 3 and Fig. 4), IL-1RA emerged as the most statistically significant immunomodulator that could be used to differentiate among the viral families with the largest importance score (Fig. 5). A second set of analytes, CXCL9, CCL4, IFN-λ1, IP-10, and IL-27, shows a decreasing order of importance in differentiation between the viral families. CCL4, IFN-λ1, and IP-10 tend to have significant differential fold changes (Fig. 5 and Table 3).

As noticed previously (Figs. 4 and 5, Table 3), the immunomodulator response to Hantaviridae, in relation to Coronaviridae, tends to be strong because many analytes are highly modulated, and the experimental data from Hantaviridae could potentially skew

Table 3. Differential fold changes of viral families compared with the Coronaviridae family.[a,b]

	Immunomodulators	vs. Orthomyxoviridae	vs. Retroviridae	vs. Flaviviridae	vs. Hantaviradae
1	IL-10	0.20	**(0.18)**	**(3.16)**	**(1.02)**
2	IL-8	**(1.87)**	**(0.56)**	**(1.79)**	0.57
3	**IP-10**	**(2.02)**	0.97	**(1.79)**	0.67
4	**CCL4**	0.16	0.36	**(1.65)**	**(1.87)**
5	TNFRSF1B	**(0.08)**	**(0.17)**	**(0.19)**	**4.79**
6	IL-6RA	0.02	0.01	**(0.04)**	**5.57**
7	IL-11	0.00	0.00	0.00	**(2.00)**
8	IL-26	0.00	0.00	0.03	**(3.59)**
9	CHI3L1	0.00	0.00	**(0.04)**	**3.90**
10	OPN	0.00	0.01	0.03	**4.46**
11	TNFRSF8	0.10	0.10	**(0.01)**	**2.80**
12	IL-34	0.00	0.00	0.00	**(3.34)**
13	**IL-27**	0.20	0.20	0.20	**(3.06)**
14	TNFRSF1A	0.05	0.05	0.02	**2.48**
15	IFN-l2	**(0.23)**	**(0.23)**	**(0.23)**	**(3.83)**
16	**IFN-λ1**	**4.29**	0.00	0.00	**(3.99)**
17	IL-22	0.07	0.07	0.07	**(2.50)**
18	IL-6RB	0.01	0.01	(0.06)	**(2.77)**
19	OCN	0.00	0.00	0.01	**2.80**
20	IL-32	0.00	0.00	0.00	**(2.76)**
21	TNFSF13B	0.00	**(0.01)**	**(0.03)**	**1.69**
22	TNFSF12	0.06	0.06	0.06	**1.64**
23	TNFSF5	0.14	**(0.03)**	0.14	**1.83**
24	IL-15	0.30	0.00	**(0.69)**	**(2.68)**
25	CCL3	0.11	0.50	**(0.35)**	**(2.62)**
26	**IL-1RA**	**(0.91)**	**0.88**	**(0.33)**	**1.64**
27	VEGF-A	**(0.83)**	**(0.30)**	**(0.65)**	**(1.71)**
28	**CXCL9**	0.37	0.39	**(1.15)**	**(2.46)**

[a]Immunomodulators given in block letters and underlined were identified in the Boruta feature selection algorithm. [b]Differential fold changes are defined as the difference in fold change between the Coronaviridae family and any other viral family. Significance in the differential fold changes is defined by multiple comparison-adjusted P-value <0.05 (bold and underlined). Differential fold changes that are negative in the Coronaviridae viral family are given in within parenthesis and bold fonts.

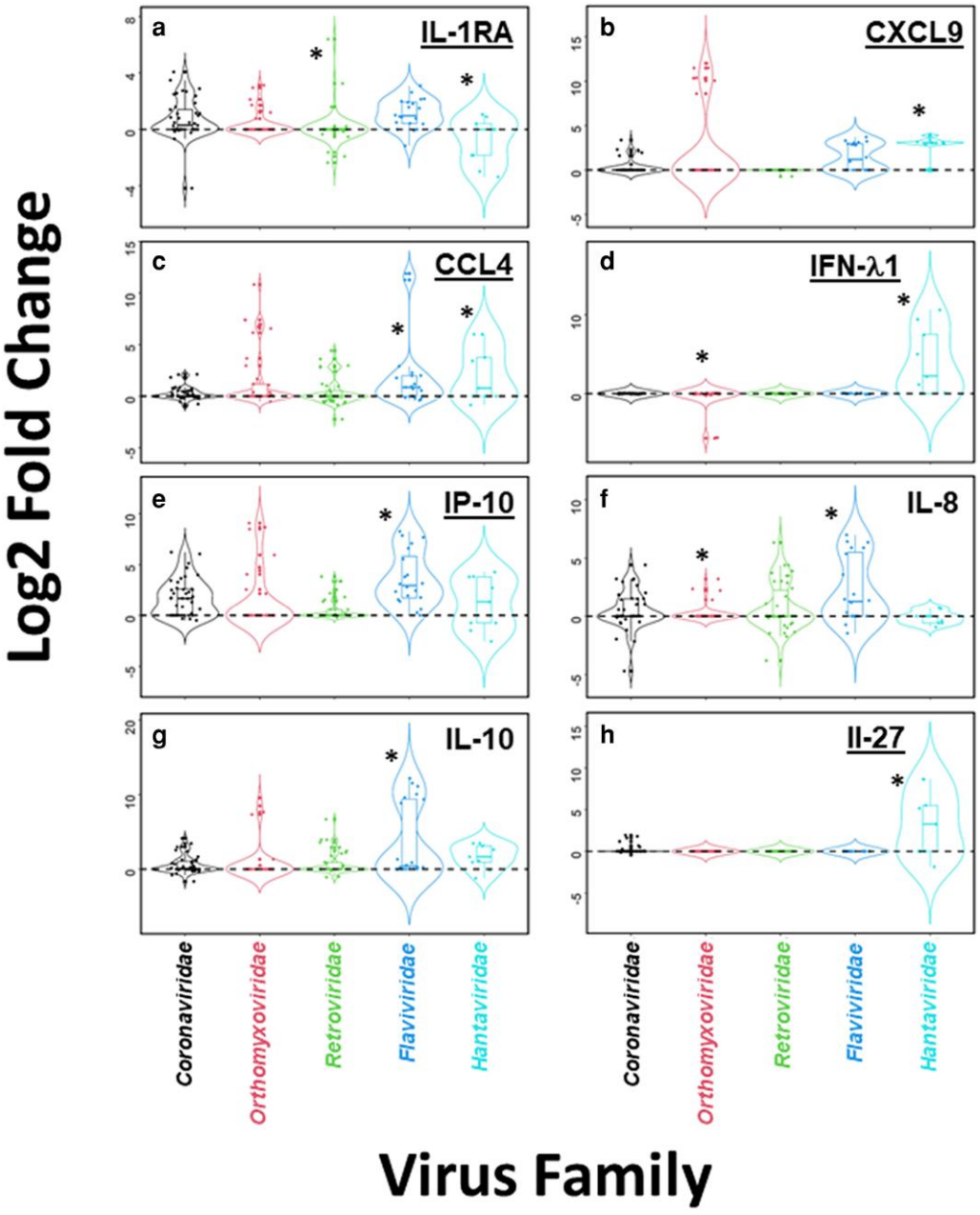

Fig. 4. Immunomodulator distributions: violin plots of the select set of immunomodulators for the five virus families, Coronaviridae (black), Orthomyxoviridae (blue), Retroviridae (teal), Flaviviridae (red), and Hantaviridae (green). a) IL1-RA, b) CXCL9, c) CCL4, d) IFN-λ1, e) IP-10, f) IL-8, g) IL-10, and h) IL-27. The immunomodulators highlighted are one of the important features of the Boruta algorithm. The viral group that has significant differential fold changes (>1.5 with $P < 0.05$) with respect to Coronaviridae is shown with an asterisk.

the feature selection process. Therefore, the Boruta algorithm analysis was repeated on the same set of 28 analytes (Table 3) without including the contribution from Hantaviridae. The analysis reveals that IL-1RA tends to be the most important immunomodulator, with the second set analytes (CXCL9, CCL4, IFN-λ1, IP-10, and IL-27) remaining listed with a lesser importance score (Fig. S5).

The co-occurrence of SARS-CoV-2 and influenza viral infections and diagnostic tools to differentiate them are public health interests. Thus, a third Boruta algorithm is repeated for Coronaviridae and Orthomyxoviridae. This data mining algorithm selected the immunomodulators TFG-β, IFN-λ1, IL-9, and eotaxin-1 (with an importance score > 80, Fig. 6a). The violin plots for these four analytes are shown in Fig. 6b–e. Differential fold changes (Fig. 6) indicate that specific immunomodulator IFN-λ1 is a significant analyte across all the viral families compared (Figs. 5 and 6). At the same time, IL1-RA and IP-10, to a lesser extent, are notable when only Coronaviridae and the Orthomyxoviridae families are compared (Table 3).

Discussion

The first goal of this study used a meta-analysis approach to determine whether the immunomodulator profile of infection with SARS-CoV-2, the causative agent of COVID-19, was distinctly different from other viral infections. Experimental multiplex microbead immunoassay data of an extensive collection of proteins with immunomodulatory activities were collected over 219 distinct datasets from 68 peer-reviewed publications. The combined

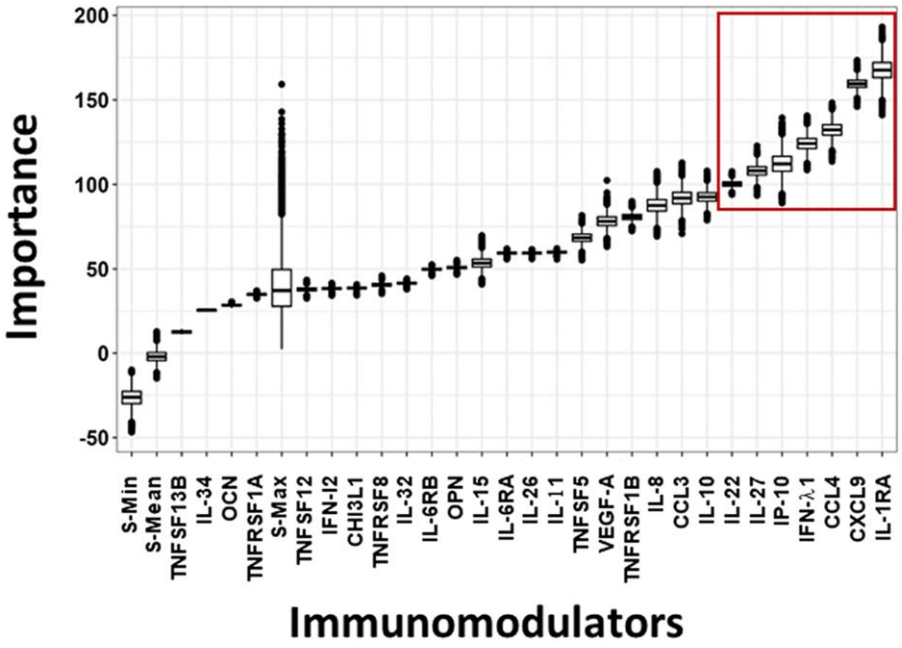

Fig. 5. Feature selection using Boruta. The common set of immunomodulators among the five viral families (Coronaviridae, Orthomyxoviridae, Retroviridae, Flaviviridae, and Hantaviridae) vs. the importance determined using the Boruta algorithm is plotted. The plot also includes the algorithm-generated shadow minimum (S-Min), maximum (S-Max), and mean (S-Mean). Immunomodulators with a score of >100 are highlighted in the red box.

analysis indicated that the immunoprofile of the Coronaviridae, which includes the SARS-CoV-2 virus, is distinct from most other viruses (Figs. 1 and 2). The meta-analysis followed the PRISMA requirements for data collection, curation, and analysis based on the availability of these tools (13). The second goal used a data mining approach based on a random forest classification scheme, aimed to identify the immunomodulators in human samples most responsible for the differentiation of the top four largest datasets of several viral families (Coronaviridae, Retroviridae, Flaviviridae, and Hantaviridae) and Orthomyxoviridae (Figs. 3–5). This process supported the importance of immunomodulators (IL1-RA, CXCL9, CCL4, IFN-λ1, IP-10, and IL-27; Figs. 4 and 5) for differentiating infections. A multivariate analysis was also performed to estimate the differential fold changes (difference in the fold changes with respect to the healthy samples between Coronaviridae and other viral families); this approach identified additional analytes of significance (IL-8 and IL-10; Table 3). To further differentiate the role of Coronaviridae and Orthomyxoviridae in the absence of data from other virus families, additional data mining investigation led to four analytes of importance (TGF-β, IFN-λ1, IL-9, and eotaxin-1; Fig. 6). The data mining approach to the curated meta-data produced a select set of immunomodulators that could be used to distinguish viral families.

The results presented here do have some notable limitations. (i) It is challenging to find all the studies related to a meta-analysis study. One reason behind this is publication biases such as the "file-drawer effect," where the studies that produced a statistically nonsignificant result were not published (14, 15). In addition, financial, political, or ideological reasons might contribute to selective reporting of the data or studies (16). (ii) In some instances, lack of information and the possibility of error in reporting in experimental studies can influence the data analysis in a meta-analysis study (17). For example, suppose a particular manuscript does not provide the meta-data related to healthy controls. In that case, the calculation of fold changes may not be possible; hence, the data were not included. (iii) For infections with the Coronaviridae, the analysis did not differentiate between the various viral mutations or health effects induced by long or acute COVID-19. (iv) Finally, wide availability and use of the robust Luminex technology for multiplex microbead analysis, data obtained from other platforms of student and single-plex studies (e.g. ELISA) were not included. Despite these limitations, our foundational study can incorporate further research using automated meta-analysis data with emerging real-time tools and applications in data science.

One of the study's significant findings is that antiinflammatory type III IFN, IFN-λ1, emerges as one of the crucial analytes differentiating among the viral families. Recent studies further support the potential of IFN-λ as an inhibitor of SARS-CoV-2 in the infected host (18). Targeting such immunomodulatory agents could provide valuable tools in the battle against viral infections and offer new avenues for their treatment and management. In a recent study, Santer et al. (19) demonstrated that IFN-λ accelerates SARS-CoV-2 clearance without affecting virus-specific T-cell responses or antibody production in mild-to-moderate acute infection with this virus.

Profiles of host response immunomodulators have been reported in the literature for many viral infections as well as SARS-CoV-2. A prospective study in patients with COVID-19 by Tamayo-Velasco et al. (20), relating cytokine profiles to ABO blood groups, revealed that higher cytokine levels in the O blood group are associated with a better outcome than A/B/AB group. Several studies have focused on differences in cytokine profiles due to the degree of COVID-19 severity. Notably, differences in the cytokine profiles of COVID-19 severity with type 1 T helper (Th1) and acute respiratory distress syndrome were observed (21, 22). At the same time, Martonik et al. investigated lipid and cytokine profiles and noted that proinflammatory cytokines (IL-6, CXCL-10, and HGF) are elevated in cases of severe COVID-19 (23). A recent comprehensive study focused on the antibody and cytokine profiles of hospitalized patients in Belgium found that proinflammatory cytokines (IL-6, IP-10, and IL-8), the antiinflammatory cytokine IL-10, and the antiviral cytokines (IFN-α, IFN-β, and IFN-λ1) proportionally increase with disease

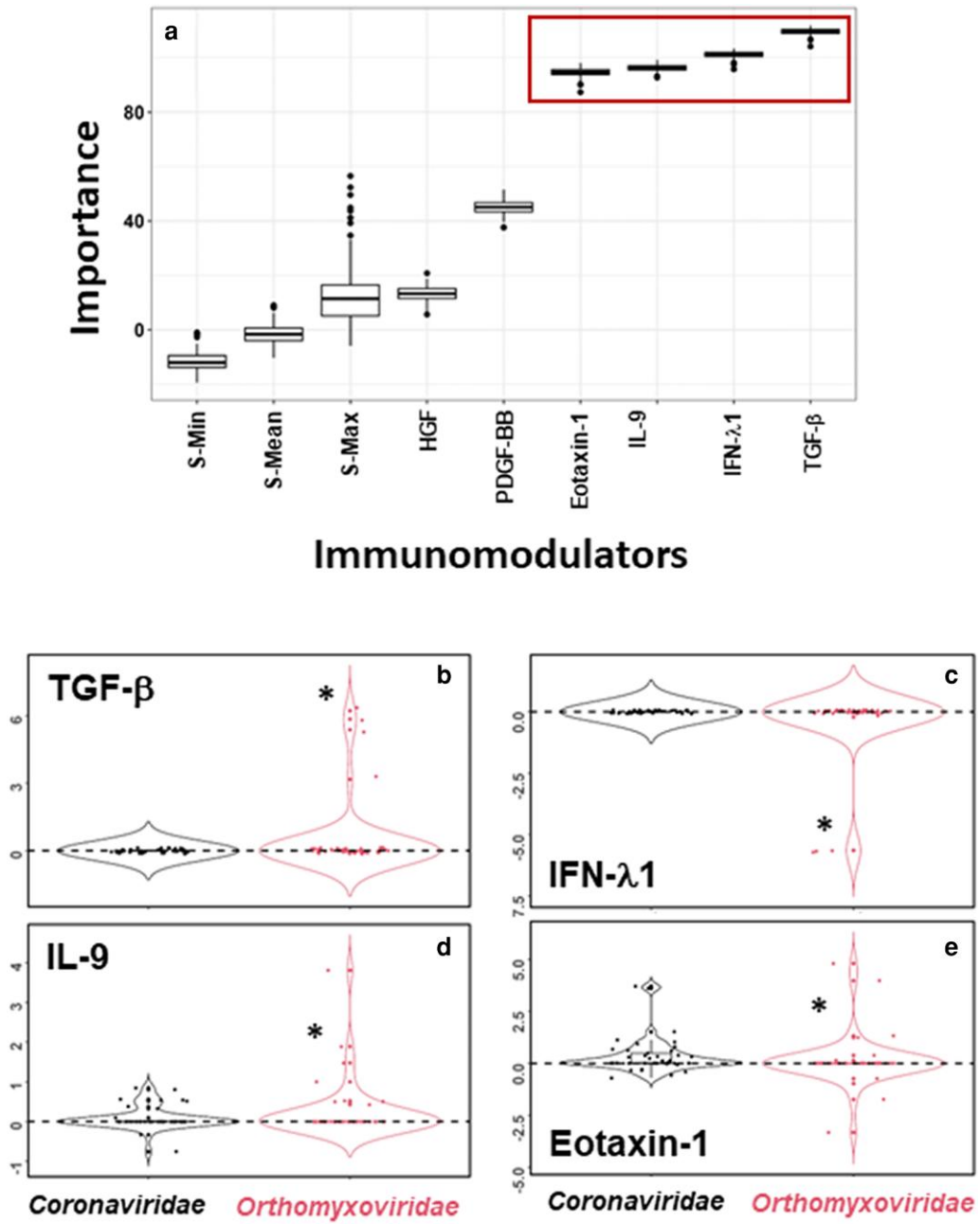

Fig. 6. a) Feature selection by Boruta between Coronaviridae (black) and Orthomyxoviridae (red). The plot also includes the algorithm-generated shadow minimum (S-Min), maximum (S-Max), and mean (S-Mean). The top four immunomodulators (red box) selected by the Boruta algorithm are shown as violin plots: b) TGF-β, c) IFN-λ1, d) IL-9, and e) eotaxin-1 with the asterisk representing a significant difference between the virus families.

severity. Remarkably, this study also found that IFN-λ1 positively associated with mortality. Another notable study examined the relationship of immunomodulators in patients with COVID-19 with chronic kidney disease (24).

Hong et al. (25) compared multiple data mining algorithms using cytokine and immune cell profiling data to predict disease severity for patients with COVID-19 with pneumonia; IL-10 predicted critically ill patients with COVID-19 pneumonia. In another study utilizing machine learning, Castro-Castro et al. (26) found that significant differences in the mortality rate were associated with the CXCL10/IL-38 ratio, and Huntington et al. (27) utilized a mutual information algorithm to correlate cytokine profiles with disease severity. However, the studies focus on a specific set of experimental conditions following a study design of choice. Combining meta-analysis with data mining provides a comprehensive yet complementary approach.

Insights from the data in our study present intriguing possibilities for several real-world biological applications in medicine of immunomodulator profiling by multiplex immunoassay analysis. In diagnostics, these findings open new avenues to capture a more holistic view of a particular infectious disease at an early stage. In addition to relying solely on methods that detect the causative virus or its components (i.e. genome or proteins) or antibody responses, a novel approach could be explored—multiplex profiling cytokines and chemokines—as a diagnostic aid for aiding disease identification. Based on clinical laboratory procedures, this approach would also offer cost-effective and time-efficient advantages (28). Moreover, such profiling methods, based on the

up-regulation and down-regulation of specific immunomodulators, would hold prospects for evaluating the stage of the disease at the time of patient presentation and thus assist in the early prediction of disease severity. However, realizing this diagnostic and prognostic potential of immunomodulator profiling would necessitate additional research, particularly through longitudinal studies involving patients with COVID-19 and other viral infections exploring analytes at different stages of the disease.

Furthermore, our study presents promising avenues to inform and improve immunotherapies and treatment strategies for viral infections. Many approaches for viral infections primarily focus on targeting the virus. For instance, the mRNA vaccines for COVID-19 are designed around the virus's spike proteins (29, 30). However, such approaches would face challenges in the face of major virus mutations, potentially rendering the treatment less or ineffective (31). On the contrary, effective therapies can be developed regardless of viral mutations by mobilizing a host's immune response based on modulating cytokines and chemokines and targeting key cell receptors of these immune cell regulators. This, in turn, would enable the tailoring of more personalized immunotherapies and more nuanced immunomodulation rather than adopting a one-size-fits-all treatment approach.

Cytokines and chemokines may be helpful tools in deciphering the nature of a host's immunity in response and helpful as early predictors of the disease progression or remission in several diseases. A recent study found that the levels of cytokines like IL-6 can predict the remission of acute myeloid leukemia in patients receiving chemotherapy (32). The association of cytokines with autoimmune diseases and morbidity and mortality of various infections has led to several trials studying the effect of immunosuppressive therapies on acute and severe stages of disease. For example, clinical trials have shown that ruxolitinib is effective as an immune modulator in COVID-19 (33). Without carefully selecting patients for treatment with immunosuppressive agents, such therapies can be inconclusive and may do more harm than good. Thus, it is imperative to investigate immune profiles in different viral infections to understand the interconnection between infection and host immune response. The approaches presented here could augment recent advancements in high-throughput technologies, such as next-generation sequencing, computer modeling, and artificial intelligence. Multiplex immuneprofiling may be highly advantageous for identifying the roles of different determinants of the host immune system in disease progression and resolution.

One limitation of the data mining approach (broadly in the realm of meta-analysis) is that no additional experiments are presented to demonstrate the validity of the data-driven conclusions. The analysis presented here, however, follows the rigors of a scientific experiment with a well-defined protocol (multiplex microbead immune assay by Luminex) and a statistical approach that critically evaluates differential responses. At the same time, as the data derived from the published results, these approaches present the findings without a clinical trial's inclusion or exclusion criteria. This study aims to identify the unique cytokine–chemokine profile for SARS-CoV-2 infection based on a diverse set of differential effects (fold changes with healthy controls). Though results are not tested on a validating cohort, annual evidence suggests the specific immunomodulators identified here would form a solid basis for immunoassays to differentiate SARS-CoV-2 from other viral infections. These pieces of evidence include the patients with COVID-19 with pneumonia (25), mortality rate (26), disease severity (27), and impact of prior or concurrent COVID-19 infection on immune profiles of patients with tuberculosis and those with other respiratory diseases (34).

Although the research in our report focuses on cytokines/chemokines stimulated by viral infections, a broader approach to how these profiles differentiate from a particular pathogenic insult to the host immune system could be valuable. The purpose is to investigate whether the cytokine–chemokine profile, due to infection with a specific virus, is representative of that virus and whether a connection between viruses and their characteristic cytokine profiles, especially that the Coronaviridae infection in which SARS-CoV-2 is a member, can be beneficial for diagnostic and treatment purposes. It is important that predictions based on meta-analysis must be validated before implementing them in clinical practice, as it is equally crucial to ensure their accuracy and reliability. Immuneprofiling may aid in carefully selecting patients for immunosuppressive therapies and thereby inform studies identifying and experimenting with new drugs for immune modulation therapies of infectious diseases as well as other pathologies where the host immune system plays a role.

Experiment design and methodology

Collection of peer-reviewed experimental data

Peer-reviewed reporting experimental multiplex immunoassay data for our analyses were identified by systematically reviewing existing literature on cytokines and chemokines in several animal species infected by a wide range of viruses. This study initially focused only on publications that measured host response cytokine and chemokine levels using multiplex microbead assays (such as Luminex xMAP technology and related analyte detection systems). Initially, four keywords were used for the database searches: "cytokine," "chemokine," "virus," and "multiplex." The search started with the Luminex publications database using the search term "@(title, abstract) cytokine AND chemokine AND virus," which showed all the papers with these three terms in the title or abstract. A similar search was conducted in the PubMed database, apart from Luminex publications using "(((cytokine) AND (chemokine)) AND (virus)) AND (multiplex)." The primary search included papers published on both databases until 2022 March 31. The resulting articles were verified manually, and the data quality was checked according to inclusion and exclusion criteria.

Inclusion criteria for research papers

The following inclusion criteria were used: (i) papers studying the effect of viral infection on cytokine–chemokine profiles in a host where analyte measurements were conducted using multiplex microbead assays, for example, with the Luminex technology; (ii) no restrictions regarding the species, age, gender, and race of the host; (iii) no restriction regarding the date or country of the research (35); and (iv) reporting raw data on cytokine measurements in the publication.

Exclusion criteria for research papers

The following exclusion criteria were used: (i) any publication language other than English; (ii) papers not containing data in an accessible format; (iii) papers only showing data in graphical, correlation, or statistical form about the cytokine profiles (with no supplemental information on exact measurements of cytokine and chemokine levels); (iv) papers expressing data in units that could not be converted to pg/mL, e.g. pg/mg, IU/mL; (v) papers that measured cytokines and chemokines with techniques other than a multiplex assay, such as ELISA, PCR, flow cytometry, etc.; (vi) papers that did not present control data from normal/healthy/untreated groups; (vii) papers or data that focused on the effect of therapy or treatment

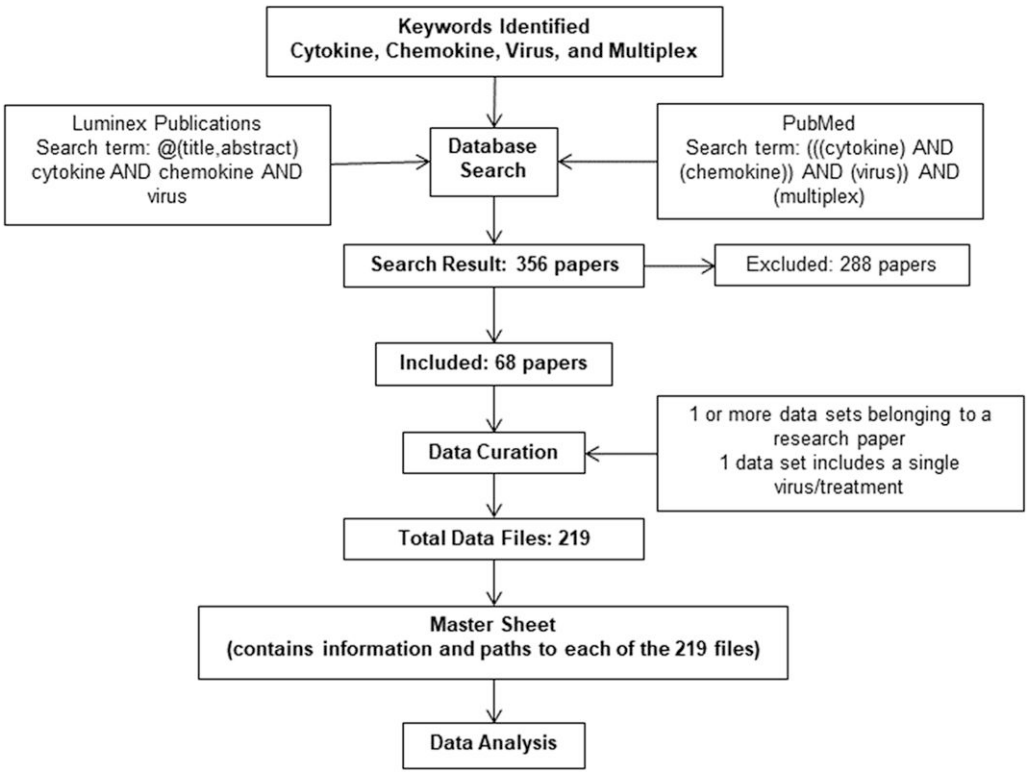

Fig. 7. Flow chart describing the summary of the various data collection and curation steps.

of an infection, e.g. data that were collected from HIV-positive patients on antiretroviral therapy; (vii) papers that studied other infections, e.g. bacterial infections, in conjunction with viral infections, for instance, COVID-19 and HIV co-infection, or *Mycobacterial tuberculosis*, and cryptococcal co-infection.

A second search was also conducted in August 2023 to include papers published after March 2022, producing 356 papers meeting the above criteria. These papers were subject to a manual data quality check, after which 288 reports were deemed ineligible and excluded from the final analysis according to the abovementioned exclusion. A data table that includes lists of PMID, Virus, Viral Family, Model System, Sample type, Sample Size, and Analytes is given in the supporting information. During the data quality inspection, 68 papers were included in the final analysis, and the supporting information was included.

Data curation

The raw multiplex immunoassay data were manually extracted and assessed to ensure adherence to the inclusion criteria. In case a paper studied more than one virus, multiple stages of infection (early phase/convalescent phase), or different types of samples (e.g. serum, plasma), data for each condition were tabulated and stored as a separate entry for analysis. Log_2-fold change values were calculated for the final analysis between healthy control/untreated and patient/treatment groups. At the end of the curation process, a total of 219 datasets were developed from 68 papers containing the following columns: PubMed ID of the source paper; the title of the paper; the form of data (mean/median and units); author(s); the year that the article was published; the journal in which the paper was published; the virus(s) that the report studied; viral family; cell model system(s) studied in the article; type(s) of samples that were collected to measure cytokines and chemokines; sample size, if available; the names of analytes measured; name of the multiplex assay; URL for the source paper; associated CSV file names (with linked computer paths to each file)—all files were named consistently, PMIDxxxxxxxx_T(number).csv, denoting the PMID number of the source paper along with a treatment number, T(number), given to each data file from the same source paper; and description of the file, briefly explaining what data are contained in the file. The flow chart describing the data collection and curation is shown in Fig. 7.

Multiplex immunoassay data analysis

All analyses were performed using R Statistical Software using the RStudio interface (36, 37). Starting with the master list file, the data were transformed into two data tables—a long-form data table and a wide-form matrix-like data table to allow for different avenues of data manipulation for the subsequent analysis. The final data analysis began with basic data exploration and a meta-analysis overview to understand the entire data using the R packages such as "tidyverse" and ggplot2 (12, 38). First, the data were subjected to a broader classification using a PLS-DA between the subgroups defined by the virus infection. Feature selection (selection of most relevant cytokine/chemokines) was performed using the Boruta algorithm based on the Random Forest classification scheme (11). The Boruta approach generates "shadow features" by shuffling/duplicating the original features as controls for defining an "importance score."

A linear model fit was determined using the estimated fold changes of each analyte to estimate the "differential fold change" between the virus groups (e.g. Coronaviridae vs. Hantaviridae). An *F*-test detected differential fold changes in samples across the groups. Using the Benjamini–Hochberg procedure, *P*-values for different measurements were transformed to compensate for multiple comparisons using the false discovery rate (39). The significance threshold was a *P*-value <0.05 for all tests with a fold change $(\log_2) > 2.0$.

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

This work was partly supported by the National Institute of Health (T32 GM 37948 and R25 GM131956). Dr Paul A. Luciw, Emeritus Professor at the University of California—Davis, provided critical comments and participated in editing the manuscript.

Author Contributions

Conceptualization, resources, supervision, and project administration: V.V.K.; data curation: A.K.; methodology, software, validation, formal analysis, investigation, visualization, and writing—original draft, review, and editing: A.K. and V.V.K.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary material.

References

1 Borish LC, Steinke JW. 2003. 2. Cytokines and chemokines. *J Allergy Clin Immunol*. 111(2):S460–S475.
2 Clark IA. 2007. The advent of the cytokine storm. *Immunol Cell Biol*. 85(4):271–273.
3 Tisoncik JR, *et al.* 2012. Into the eye of the cytokine storm. *Microbiol Mol Biol Rev*. 76(1):16–32.
4 Kimura H, Yoshizumi M, Ishii H, Oishi K, Ryo A. 2013. Cytokine production and signaling pathways in respiratory virus infection. *Front Microbiol*. 4:276.
5 Clark IA, Vissel B. 2017. The meteorology of cytokine storms, and the clinical usefulness of this knowledge. *Semin Immunopathol*. 39(5):505–516.
6 Teijaro JR, editor. 2017. *Cytokine storms in infectious diseases. Seminars in immunopathology*. Springer.
7 Gijs M, *et al.* 2021. In-depth investigation of conjunctival swabs and tear fluid of symptomatic COVID-19 patients, an observational cohort study. *Transl Vis Sci Technol*. 10(12):32.
8 Mudd PA, *et al.* 2020. Distinct inflammatory profiles distinguish COVID-19 from influenza with limited contributions from cytokine storm. *Sci Adv*. 6(50):eabe3024.
9 Swets MC, *et al.* 2022. SARS-CoV-2 co-infection with influenza viruses, respiratory syncytial virus, or adenoviruses. *Lancet*. 399(10334):1463–1464.
10 Lê Cao K-A, Welham ZM. 2021. *Multivariate data integration using R: methods and applications with the mixOmics package*. Chapman and Hall/CRC.
11 Kursa MB, Rudnicki WR. 2010. Feature selection with the Boruta package. *J Stat Softw*. 36:1–13.
12 Wickham H, Wickham H. 2016. *Data analysis*. Springer.
13 Page MJ, *et al.* 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev*. 10(1):89.
14 Rosenthal R. 1979. The file drawer problem and tolerance for null results. *Psychol Bull*. 86(3):638–641.
15 Bartolucci AA, Hillegass WB. 2010. Overview, strengths, and limitations of systematic reviews and meta-analyses. In: Chiappelli F, editor. *Evidence-based practice: toward optimizing clinical outcomes*. Berlin, Heidelberg: Springer. p. 17–33.
16 Rothstein HR. 2008. Publication bias as a threat to the validity of meta-analytic results. *J Exp Criminol*. 4(1):61–81.
17 Walker E, Hernandez AV, Kattan MW. 2008. Meta-analysis: its strengths and limitations. *Cleve Clin J Med*. 75(6):431–439.
18 Oleinik LA, Madonov PG, Pykhtina MB. 2023. Potential of interferon lambda as an inhibitor of SARS-CoV-2. *Mol Biol*. 57(2):291–298.
19 Santer DM, *et al.* 2022. Interferon-λ treatment accelerates SARS-CoV-2 clearance despite age-related delays in the induction of T cell immunity. *Nat Commun*. 13(1):6992.
20 Tamayo-Velasco Á, *et al.* 2021. Can the cytokine profile according to ABO blood groups be related to worse outcome in COVID-19 patients? Yes, they can. *Front Immunol*. 12:726283.
21 Ghazavi A, Ganji A, Keshavarzian N, Rabiemajd S, Mosayebi G. 2021. Cytokine profile and disease severity in patients with COVID-19. *Cytokine*. 137:155323.
22 Ling L, *et al.* 2021. Longitudinal cytokine profile in patients with mild to critical COVID-19. *Front Immunol*. 12:763292.
23 Martonik D, *et al.* 2023. Effect of antiviral and immunomodulatory treatment on a cytokine profile in patients with COVID-19. *Front Immunol*. 14:1222170.
24 Ciceri P, *et al.* 2022. Cytokine and chemokine retention profile in COVID-19 patients with chronic kidney disease. *Toxins (Basel)*. 14(10):673.
25 Hong W, *et al.* 2022. A comparison of XGBoost, random forest, and nomograph for the prediction of disease severity in patients with COVID-19 pneumonia: implications of cytokine and immune cell profile. *Front Cell Infect Microbiol*. 12:819267.
26 Castro-Castro AC, *et al.* 2022. Difference in mortality rates in hospitalized COVID-19 patients identified by cytokine profile clustering using a machine learning approach: an outcome prediction alternative. *Front Med (Lausanne)*. 9:987182.
27 Huntington KE, *et al.* 2021. Cytokine ranking via mutual information algorithm correlates cytokine profiles with presenting disease severity in patients infected with SARS-CoV-2. *eLife*. 10:e64958.
28 Munro SB, Kuypers J, Jerome KR. 2013. Comparison of a multiplex real-time PCR assay with a multiplex luminex assay for influenza virus detection. *J Clin Microbiol*. 51(4):1124–1129.
29 Polack FP, *et al.* 2020. Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. *N Engl J Med*. 383(27):2603–2615.
30 Wrapp D, *et al.* 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 367(6483):1260–1263.
31 Weisblum Y, *et al.* 2020. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *eLife*. 9:e61312.
32 Hu R, *et al.* 2023. Cytokine levels in patients with non-M3 myeloid leukemia are key indicators of how well the disease responds to chemotherapy. *Clin Exp Med*. 23(8):4623–4632.
33 Cao Y, *et al.* 2020. Ruxolitinib in treatment of severe coronavirus disease 2019 (COVID-19): a multicenter, single-blind, randomized controlled trial. *J Allergy Clin Immunol*. 146(1):137–46.e3.
34 Cottam A, *et al.* 2023. The impact of prior SARS-CoV-2 infection on host inflammatory cytokine profiles in patients with TB or other respiratory diseases. *Front Immunol*. 14:1292486.
35 Tawfik GM, *et al.* 2019. A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Trop Med Health*. 47:46.
36 RStudio Team. 2020. *RStudio: Integrated Development for R*. RStudio, PBC, Boston, MA. http://www.rstudio.com/.
37 R Core Team. 2023. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
38 Wickham H, *et al.* 2019. Welcome to the Tidyverse. *J Open Source Softw*. 4(43):1686.
39 Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc: Ser B (Methodol)*. 57(1):289–300.