

Proteome signature difference between respiratory viruses is associated with severity of bronchiolitis

To the editor

Bronchiolitis is the leading cause of hospitalization in American infants.¹ Most cases of severe bronchiolitis (ie, bronchiolitis requiring hospitalization) are caused by two distinct viruses: respiratory syncytial virus (RSV) and rhinovirus (RV).² Although bronchiolitis has been traditionally considered a single disease entity and guideline-recommended management does not differ by virus, emerging evidence suggests between-virus heterogeneity in the pathobiology and clinical severity.³ Indeed, recent studies suggest that RSV and RV bronchiolitis have different upper airway microbiome (eg, *Streptococcus* dominance in RSV infection vs. *Moraxella* and *Haemophilus* dominance in RV infection)⁴ and transcriptome⁵ profiles.

Despite the clinical and research relevance, it remains uncertain how different viruses systemically contribute to downstream molecules that influence outcomes of bronchiolitis. Proteomic approach addresses this knowledge gap through comprehensively profiling functional proteins that are the main effectors of cellular physiology.⁶ However, no study has examined the proteome signature of bronchiolitis, let alone their differences by major virus. Limited understanding of these issues has hindered efforts to develop targeted treatments for bronchiolitis, for which no effective treatment strategies are currently available.²

In this context, by comprehensively profiling the serum proteome of infants enrolled into a multicenter bronchiolitis cohort, we examined the relationships between respiratory viruses (RSV and RV), circulating proteome, and acute severity of bronchiolitis.

We analyzed data from the 35th Multicenter Airway Research Collaboration (MARC-35) cohort study—a multicenter prospective cohort study. Details of the study design, participants, data collection, testing, and statistical analysis may be found in the Online Repository. Briefly, investigators enrolled infants (age <1 year) hospitalized with bronchiolitis at 17 sites across 14 U.S. states (Table S1) in 2011–2014.⁴ Bronchiolitis was defined by the American Academy of Pediatrics guidelines. We excluded infants with a known heart-lung disease, immunodeficiency, immunosuppression, or gestational age of <32 weeks. Of infants enrolled into the MARC-35 cohort, the current analysis investigated 117 infants with RSV or RV infection who underwent serum proteomic testing (Table S2). By design, infants with RV infection were overrepresented in this substudy. The institutional review board at each of the participating hospitals approved the study.

Clinical data were collected via structured interview and chart reviews using a standardized protocol.⁴ Nasopharyngeal and serum specimens were also collected by trained site investigators within 24 hours of hospitalization. Nasopharyngeal specimens were tested for 17 respiratory viruses (including RSV and RV) by using real-time polymerase chain reaction assays. Serum specimens were tested for proteomic profiling, as described in the Supplemental Methods. Briefly, serum proteins were measured with the use of the Olink multiplex platform (13 panels; Olink Bioscience, Uppsala, Sweden). The expression value of each protein was normalized to a unit on a log₂ scale, proportional to its concentration.

The primary outcome was positive pressure ventilation (PPV) use, defined as the use of continuous positive airway pressure and/or intubation during the index hospitalization. The secondary outcome was intensive care use defined as a composite of intensive care unit admission and/or the use of PPV.

First, to identify serum proteins that best discriminate between the two virus groups (RSV vs. RV), we constructed an elastic net regression model with cross-validation. Second, we visualized the discriminatory ability of 30 selected proteins by using t-distributed stochastic neighbor embedding (t-sne) and computed Spearman's correlations between the viral genomic load and the 30 most discriminatory proteins. Third, we tested for the association of RV infection (RSV infection as the reference group) with the 30 proteins by constructing multivariable random-effects linear regression models adjusting for age, sex, and clustering within hospitals. We corrected for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) method. Fourth, to detect biologically meaningful pathways, we performed a protein-protein interaction network functional enrichment analysis. Lastly, to determine the association of selected proteins with severity outcomes, we constructed multivariable random-effects logistic regression models that adjust for age, sex, and virus and account for patient clustering. In the sensitivity analysis, we computed E-values to examine the robustness of causal inference to potential unmeasured confounding. The E-value is a method to gauge the evidence for causality, which indicates the minimum magnitude of association that an unmeasured confounder would need to have with both the exposure and outcome to fully explain away the association of interest.

As part of a 17-center prospective cohort study, 140 infants with bronchiolitis underwent serum proteomic testing. Of these,

we excluded 16 infants with coinfection and 7 with neither RSV nor RV infection, leaving 117 infants (66 RSV and 51 RV infection) eligible for the current analysis. The analytic and non-analytic cohorts did not differ in patient characteristics, except for the overrepresentation of RV in the analytic cohort (Table S2). Overall, the median age was 3 months at hospitalization and 61% were male. The RSV and RV groups did not differ in patient characteristics, except the RV group was older and had a larger weight at presentation (Table S3).

Proteomic profiling assayed a total of 1,109 circulating proteins. In the RSV-RV comparison (RSV infection as the reference group), the elastic net regression selected the 30 most discriminatory

proteins (Figure 1A). Of these 30 proteins, 17 proteins (eg, VSIG4) are upregulated in the RSV group while 13 proteins (eg, IL-17A, CCL25, MCP-4/CCL13) are upregulated in the RV group (Table S4). These differences between qualitative virus status (RSV vs RV) were consistent with the correlations between quantitative virus status (viral genomic load) and protein levels (Figure 1B). In the multivariable models, the between-virus differences remained significant for 23 proteins (FDR < 0.05; Table S4). The protein-protein interaction network functional enrichment analysis identified 80 differentially enriched pathways—for example, response to interleukin 1 (IL-1) pathways and regulation of mitogen-activated protein kinase (MAPK) cascade (FDR < 0.05; Figure 1C).

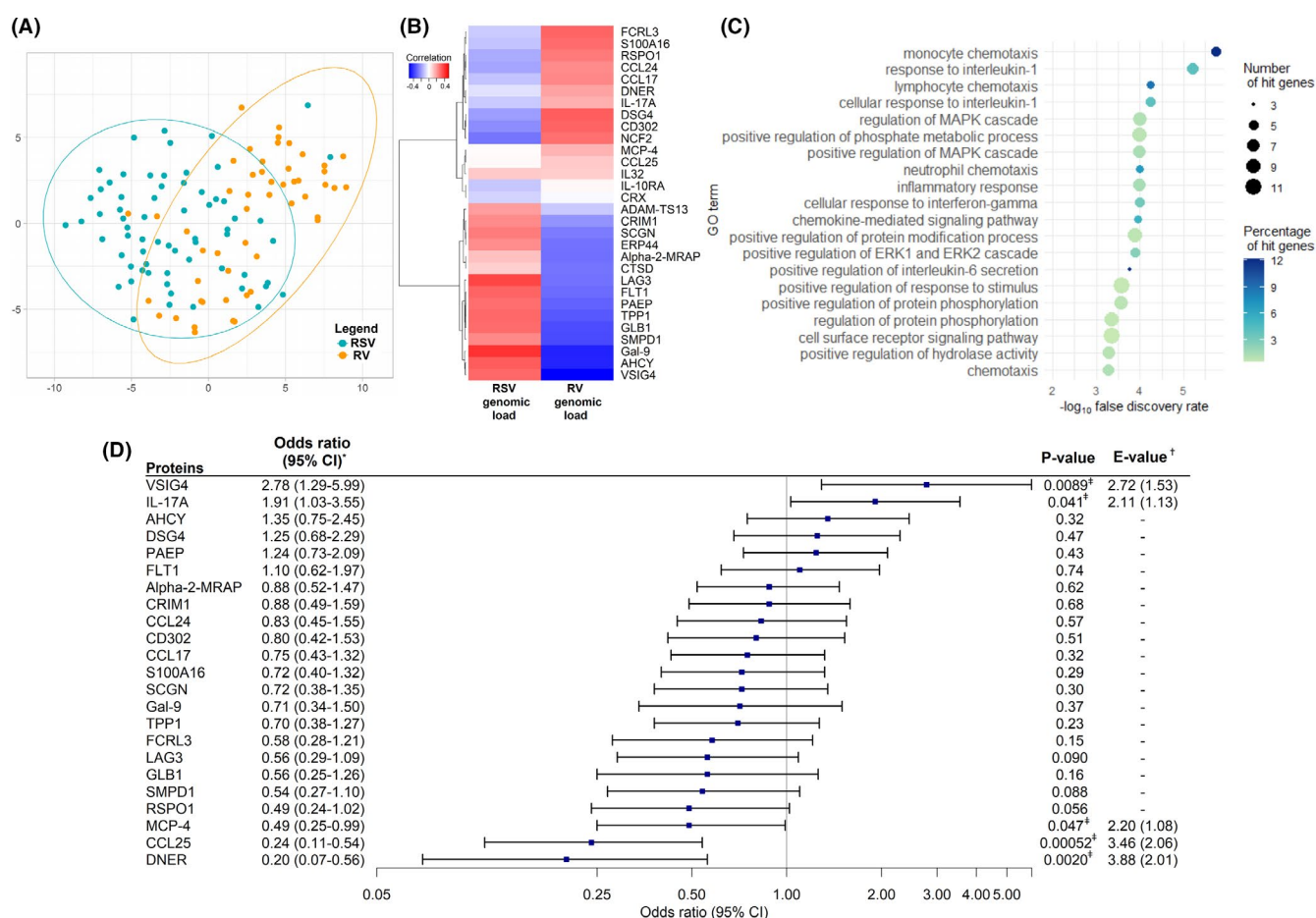


FIGURE 1 Associations of respiratory syncytial virus and rhinovirus species with serum proteins in infants with severe bronchiolitis. (A) The t-distributed stochastic neighbor embedding (t-SNE) plot—using the 30 most discriminatory proteins that are selected by an elastic net regression model—according to RSV (green) vs. RV (orange) infection for 117 infants. Each dot represents the serum proteome profile of a single infant. The ellipses are 95% confidence intervals. (B) Heatmap of Spearman's correlations between the genomic load (measured as the inverse cycle threshold) of RSV and RV and the 30 most discriminatory serum proteins in the elastic net regression model (Table S4). The color bar indicates the correlation coefficients: red color indicates a positive correlation while blue color indicates a negative correlation. (C) The protein-protein interaction network functional enrichment analysis with Gene Ontology terms using the selected proteins significantly associated with RV infection. The horizontal axis represents the negative $\log_{10}$ -transformed false discovery rate (FDR). Plot size represents the number of hit genes in the corresponding pathway. Plot color represents the percentage of hit genes in the corresponding pathway. (D) Forest plot for the associations of serum proteins with the risk of positive pressure ventilation use. Abbreviations: RSV, respiratory syncytial virus; RV, rhinovirus. *For 23 proteins that have a FDR of <0.05 in the virus-protein models (Table S4), multivariable random-effect logistic models adjusted for age, sex, respiratory virus, and clustering within hospitals were fit. The odds ratios were estimated per one standard deviation increment in $\log_2$ -transformed protein level. †The E-value (with its lower 95% confidence interval bound) represents how strongly an unmeasured confounder would have to be associated with the exposure and outcome for the observed association to be independent.

Of these 23 discriminatory proteins, five proteins were significantly associated with the risk of PPV use, indicating the clinical significance of these proteins (Figure 1D). For example, two proteins were associated with a higher risk of PPV use (Figure 1D): VSIG4 (adjOR, 2.78; 95% CI, 1.29–5.99; E-value = 2.72) and IL-17A (adjOR, 1.91; 95% CI, 1.03–3.55; E-value = 2.11). In contrast, three proteins were associated with a lower risk—for example, CCL25 (adjOR, 0.24; 95% CI, 0.11–0.54; E-value = 3.46) and MCP-4/CCL13 (adjOR, 0.49; 95% CI, 0.25–0.99; E-value = 2.20). The results did not materially change in the sensitivity analyses, with eight proteins significantly associated with the risk of intensive care use (Figure S1).

Bronchiolitis has been considered a single disease entity.² However, recent studies have demonstrated differences in clinical characteristics (eg, severity)³ and molecular signatures^{5–7} by infecting virus. For example, among infants with severe bronchiolitis, the upper airway microbiota differed between RSV and RV.^{4,8} The current study builds on these previous reports indicating the heterogeneity of bronchiolitis, and extends them by comprehensively investigating the circulating proteome and demonstrating their between-virus differences in a prospective cohort of infants with severe bronchiolitis.

The mechanisms underlying the interrelationship between the infecting virus, host systemic response, and bronchiolitis outcome remain to be elucidated. In the current study, we found several notable pathways differentially enriched between RSV and RV—for example, IL-1 and MAPK pathways. In agreement with our findings, previous studies have suggested that RV infection triggers innate immunity including IL-1 pathways, upregulating pro-inflammatory signals, and recruiting immune cells.⁹ Prior research has also found that IL-1 upregulates IL-17 production from T_H17 cells, which contributes to the immune response to virus infection.⁷ Polymorphisms in IL-17 have been linked to severe bronchiolitis¹⁰ and its level was also associated with the disease severity in the current study. In contrast, the MAPK cascades—in conjunction with related chemokines (CCL25 and MCP-4/CCL13)—are thought to play important roles in the pathobiology of RSV infection, such as the production of reactive oxygen species and NETosis.¹¹ Additionally, RSV infection has been associated with reduced levels of these chemokines,¹² which also were down-regulated in RSV infection and negatively associated with the severity in the current study. Further clarification of the underlying mechanisms could lead to the development of treatments for severe bronchiolitis.

The current study has several potential limitations. First, the study did not have “healthy controls.” However, the study objective was *not* to assess the proteome related to the presence or absence of bronchiolitis *but* to determine the virus-specific proteome signature among infants with bronchiolitis. Second, as with any observational study, causal inference may be confounded by unmeasured factors, such as nutrition and between-hospital differences in intensive care resources. However, we statistically accounted for patient clustering at the hospital level. The large E-values also support causality. Fourth, while we applied approaches to minimize overfitting (eg, elastic net model with cross-validations), external validation is warranted. Lastly, although the cohort is a racially/ethnically diverse

multicenter sample, our inference may not be generalizable beyond infants with severe bronchiolitis. Regardless, our data are directly relevant to 110,000 infants hospitalized for bronchiolitis in the United States each year.¹

In conclusion, based on the multicenter prospective data of infants with severe bronchiolitis, we found virus-specific (RSV vs. RV) circulating proteome signatures with biologically relevant inflammatory pathways. Discriminatory proteins were associated with the severity of bronchiolitis. Our data provide additional evidence that the pathobiology of bronchiolitis is heterogeneous and varies by respiratory virus. Our findings should advance research into the complex interrelations between respiratory viruses and host systemic response, as well as their integrated contributions to clinical outcomes in bronchiolitis. Further, our data provide a foundation for the development of bronchiolitis management strategies (eg, modification of specific immune pathways according to the infecting pathogen and identification of patients at risk for higher severity) in this vulnerable population with a large morbidity burden.

KEYWORDS

bronchiolitis, proteome, respiratory syncytial virus, rhinovirus, severity

FUNDING INFORMATION

This study was supported by grants (UG3/UH3 OD-023253, R01 AI-127507, R01 AI-134940, and R01 AI-137091) from the National Institutes of Health (Bethesda, MD). The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funding organization was not involved in the collection, management, or analysis of the data; preparation or approval of the manuscript; or decision to submit the manuscript for publication.

CONFLICT OF INTEREST

C. A. C. and K. H. have received grants from the National Institutes of Health. All remaining authors have no conflicts of interest.

AUTHOR CONTRIBUTIONS

Tadao Ooka: Conceptualization (equal); Formal analysis (lead); Methodology (equal); Visualization (lead); Writing-original draft (lead). **Yoshihiko Raita:** Conceptualization (equal); Methodology (equal); Project administration (supporting); Writing-original draft (supporting); Writing-review & editing (equal). **Debby Ngo:** Investigation (equal); Project administration (equal); Writing-review & editing (equal). **Jonathan Mansbach:** Investigation (equal); Project administration (equal); Writing-review & editing (equal). **Zhaozhong Zhu:** Methodology (equal); Writing-review & editing (equal). **Carlos A Camargo Jr:** Funding acquisition (equal); Investigation (equal); Project administration (supporting); Writing-review & editing (equal). **Kohei Hasegawa:** Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Supervision (equal); Writing-original draft (supporting); Writing-review & editing (equal).

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/pai.13606>.

Tadao Ooka^{1,2} 

Yoshihiko Raita¹

Debby Ngo³

Jonathan M. Mansbach⁴

Zhaozhong Zhu¹

Carlos A. Camargo Jr.¹

Kohei Hasegawa¹

¹Department of Emergency Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA

²Department of Health Science, University of Yamanashi, Chuo, Japan

³Division of Cardiology, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

⁴Department of Pediatrics, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA

Correspondence

Yoshihiko Raita, Department of Emergency Medicine, Massachusetts General Hospital, 125 Nashua Street, Suite 920, Boston, MA 02114-1101, USA.
Email: yraita1@alumni.jh.edu

Editor: Ömer Kalayci

Tadao Ooka and Yoshihiko Raita contributed equally to this manuscript.

ORCID

Tadao Ooka  <https://orcid.org/0000-0002-1343-3986>

REFERENCES

1. Fujiogi M, Goto T, Yasunaga H, et al. Trends in bronchiolitis hospitalizations in the United States: 2000–2016. *Pediatrics*. 2019;144(6):2000–2016.

2. Hasegawa K, Dumas O, Hartert TV, et al. Advancing our understanding of infant bronchiolitis through phenotyping and endotyping: clinical and molecular approaches. *Expert Rev Respir Med*. 2016;10(8):891–899.
3. Hasegawa K, Mansbach JM, Camargo CA Jr. Infectious pathogens and bronchiolitis outcomes. *Expert Rev Anti Infect Ther*. 2014;12(7):817–828.
4. Mansbach JM, Hasegawa K, Henke DM, et al. Respiratory syncytial virus and rhinovirus severe bronchiolitis are associated with distinct nasopharyngeal microbiota. *J Allergy Clin Immunol*. 2016;137(6):1909–1913.
5. Mejias A, Dimo B, Suarez NM, et al. Whole blood gene expression profiles to assess pathogenesis and disease severity in infants with respiratory syncytial virus infection. *PLoS Medicine*. 2013;10(11):e1001549.
6. Bowler RP, Wendt CH, Fessler MB, et al. New strategies and challenges in lung proteomics and metabolomics. An official American thoracic society workshop report. *Ann Am Thorac Soc*. 2017;14(12):1721–1743.
7. Sutton CE, Lalor SJ, Sweeney CM, Brereton CF, Lavelle EC, Mills KH. Interleukin-1 and IL-23 induce innate IL-17 production from $\gamma\delta$ T cells, amplifying Th17 responses and autoimmunity. *Immunity*. 2009;31(2):331–341.
8. Fujiogi M, Camargo CA Jr, Bernot JP, et al. In infants with severe bronchiolitis: dual-transcriptomic profiling of nasopharyngeal microbiome and host response. *Pediatr Res*. 2020;88:144–146.
9. Piper SC, Ferguson J, Kay L, et al. The role of interleukin-1 and interleukin-18 in pro-inflammatory and anti-viral responses to rhinovirus in primary bronchial epithelial cells. *PLoS One*. 2013;8(5):e63365.
10. Janssen R, Bont L, Siezen CL, et al. Genetic susceptibility to respiratory syncytial virus bronchiolitis is predominantly associated with innate immune genes. *J Infect Dis*. 2007;196(6):826–834.
11. Muraro SP, Souza GF, Gallo SW, et al. Respiratory syncytial virus induces the classical ROS-dependent NETosis through PAD-4 and necroptosis pathways activation. *Sci Rep*. 2018;8:14166.
12. Wetzke M, Haid S, Decking LM, Pietschmann T, Hansen G, Happle C. Reduced chemokine responses in human neonatal cells upon infection with RSV. *Eur Respir J*. 2017;50(61):PA541.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

4Tadao Ooka, MD, PhD*; Yoshihiko Raita, MD, MPH, MMSc*; Debby Ngo, MD; Jonathan M.

5Mansbach, MD, MPH; Zhaozhong Zhu, ScD; Carlos A. Camargo Jr., MD, DrPH; and Kohei

6Hasegawa, MD, MPH, MS

8**Supplemental Methods**.....3

9**Supplemental References**.....11

10**Table S1.** Principal investigators at the 17 participating sites in MARC-35.....14

11**Table S2.** Comparisons between analytic and non-analytic cohorts in MARC-35.....15

12**Table S3.** Patient characteristics according to respiratory syncytial virus and rhinovirus

13infection.....17

14**Table S4.** Multivariable associations of respiratory syncytial virus and rhinovirus infection with

15serum proteins in infants with severe bronchiolitis.....18

16**Figure S1.** Multivariable associations of serum proteins with intensive treatment in infants with

17severe bronchiolitis.....	20
-----------------------------	----

18SUPPLEMENTAL METHODS

19Study Design, Setting, and Participants

20 We analyzed data from the 35th Multicenter Airway Research Collaboration (MARC-
2135)—a multicenter prospective cohort study of infants hospitalized with bronchiolitis [1].
22MARC-35 is coordinated by the Emergency Medicine Network (EMNet), a collaboration of 247
23participating hospitals [2]. Investigators enrolled infants (age <1 year) hospitalized with
24bronchiolitis at 17 sites across 14 U.S. states (**Table S1**) using a standardized protocol during
25three consecutive bronchiolitis seasons (from November 1 through April 30) during 2011-2014.
26Bronchiolitis was defined by the American Academy of Pediatrics guidelines as an acute
27respiratory illness with some combination of rhinitis, cough, tachypnea, wheezing, crackles, and
28retractions [3] and was diagnosed by an attending physician. We excluded infants who were
29transferred to a participating hospital >24 hours after the original hospitalization, those who were
30consented >24 hours after hospitalization, or those with known heart-lung disease,
31immunodeficiency, immunosuppression, or gestational age of <32 weeks. All patients were
32treated at the discretion of the treating physicians.

33 The institutional review board at each of the participating hospitals approved the study.
34Written informed consent was obtained from the parent or guardian. Of 1,016 infants enrolled
35into the MARC-35 cohort, the current analysis investigated infants with respiratory syncytial

36virus (RSV) or rhinovirus (RV) infection who underwent serum proteomic testing (**Table S2**).

37By design, infants with RV infection were overrepresented in this substudy.

38

39**Data Collection**

40 Clinical data (patients' demographic characteristics, and family, environmental, and
41medical history, and details of the acute illness) were collected via structured interview and chart
42reviews [1]. All data were reviewed at the EMNet Coordinating Center (Boston, Massachusetts,
43USA), and site investigators were queried about missing data and discrepancies identified by
44manual data checks.

45 In addition to the clinical data, investigators also collected nasopharyngeal airway and
46serum specimens within 24 hours of hospitalization by using standardized protocols [1,4]. All
47sites used the same collection equipment and collected the samples within 24 hours of a child's
48arrival on the medical ward or intensive care unit. Frozen nasopharyngeal specimens were
49shipped in batches to Baylor College of Medicine (Houston, Texas, USA) where they were tested
50for respiratory viruses using real-time polymerase chain reaction (RT-PCR) [1,4,5]. Serum
51specimens were shipped to Beth Israel Deaconess Medical Center (Boston, Massachusetts, USA)
52for proteomic profiling [8].

53 **Respiratory Virus Testing**

55 Nasopharyngeal samples were tested for 17 respiratory viruses (including RSV and RV)
56 by using RT-PCR assays [5]. For RV detection, complementary DNA was generated using virus-
57 specific primers for RV and singleplex RT-PCR was used. The details of the RV primers and
58 probes have been described elsewhere [9].

59 **Serum Proteomic Profiling**

60 *Protein Measurement*

62 Serum samples were analyzed for circulating protein markers using Olink multiplex
63 arrays (Proseek multiplex arrays; Olink Bioscience, Uppsala, Sweden) [10]. Olink multiplex
64 panels were run on 96-well plates. Each panel assayed 92 proteins. In total, 13 panels of
65 biomarkers were run to assay 1,196 proteins—human cardiometabolic, cardiovascular disease
66 (CARDIOVASCULAR II / III), cell regulation, development, immune response, inflammation,
67 metabolism, neuro exploratory, neurology, oncology (ONCOLOGY II / III), and organ damage
68 [10]. Briefly, an incubation mix was prepared containing pairs of oligonucleotide-labelled
69 antibodies (Proseek probes) to each protein. Presence of the analyte in the sample brings the
70 partner probes in close proximity, allowing the complementary oligonucleotides to form a PCR

71target sequence. Addition of an extension mix then allows the specific target sequences to be
72detected and quantified using real-time PCR. The reaction was run in 96-well plate format, in a
73thermal cycler at 50°C for 20 minutes, 95°C for 5 minutes, followed by 17 cycles of 95°C for 30
74seconds, 54°C for 1 minutes, and 60°C for 1 minutes. Detection was carried out using a 96
75Dynamic Array™ Integrated Fluidic Circuit (IFC) on the Biomark HD (Fluidigm, South San
76Francisco, California, USA), at 50°C for 120 seconds, 70°C for 1800 seconds, 25°C for 600
77seconds, 95°C for 300 seconds, followed by 40 cycles of 95°C for 15 seconds and 60°C for 60
78seconds [11].

79*Quality Control*

80 Four internal controls were added to each sample to monitor the quality of assay
81performance, as well as the quality of individual samples. The quality control (QC) was
82performed in two steps. First, each sample plate was examined on the standard deviation of the
83internal controls. This should be below 0.2 normalized protein expression (NPX) units. Only data
84from sample plate that pass this QC were reported. Second, the quality of each sample was
85determined by evaluating the deviation from the median value of the controls for each individual
86sample. Samples that deviate less than 0.3 NPX from the median passed the quality control.

87*Intra- and Inter-Assay Coefficient of Variance (%CV)*

88 Intra- and inter-Assay %CVs were based on external control samples (replicated pooled
89 plasma samples) included on each plate. Calculations were made using linear generate NPX-
90 values. Intra-assay and inter-assay coefficient variance of proteins across all panels ranged from
91 were 3%-15% and 7%-17%, respectively. .

92 *Protein Detection Results, Protein Concentration Calculation, and Output*

93 On average, 94% of tested proteins were detected in >75% of the specimens. Protein
94 concentration data were analyzed using Real-Time PCR Analysis Software (Fluidigm) and a
95 macro program provided by Olink in which the output of quantification cycles from the PCR
96 reaction was adjusted, using controls and a correction factor to generate NPX. NPX value using a
97 unit of a $\log_2$ scale, was proportional to concentration. Each serum sample was analyzed for total
98 protein content using a BCA kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA),
99 according to manufacturer's instructions, and values used to normalize the NPX quantitation [9].

100 Of 1,196 proteins that are originally measured, we removed overlapping proteins
101 between the 13 test panels and proteins with missingness of >40%, leaving 1,109 proteins for the
102 following analysis. We imputed the missing values of 4 proteins with missingness of <1% by
103 using R *missForest* package [12].

104

105 **Outcomes**

106 The primary outcome was positive pressure ventilation (PPV) use, defined as the use of
107continuous positive airway pressure and/or intubation at any time during the index
108hospitalization [11]. The secondary outcome was intensive care use defined as a composite of
109intensive care unit (ICU) admission and/or the use of PPV during the hospitalization [12].

110

111**Statistical Analysis**

112 In the current analysis, we categorized the infants into two mutually exclusive groups:
113RSV (reference) and RV groups. First, to identify serum proteins that best discriminate the virus
114groups, we constructed an elastic net regression model with 4-fold cross-validation (repeated 20
115times) to minimize potential overfitting, by using R *caret* package [13]. Second, we visualized
116the discriminatory ability of these selected proteins by using t-distributed stochastic neighbor
117embedding (t-SNE), by using *Rtsne* package [14]. Third, we computed Spearman's correlations
118between the viral genomic load (measured as the inverse cycle threshold) and the levels of the
119selected 30 most discriminatory proteins. Fourth, we then tested for the association of RV (RSV
120as the reference group) infection with each of the proteins by constructing a multivariable
121random-effects linear regression model adjusting for age and sex and accounting for clustering
122within hospitals. To detect biologically meaningful pathways, we performed protein-protein

123interaction networks functional enrichment analysis using the selected proteins significantly
124associated with RV infection. We used STRING 11.0 [15] for the functional enrichment analysis.

125 Lastly, to determine the association of the selected proteins with the primary severity
126outcome, we constructed multivariable random-effects logistic regression models adjusting for
127age, sex, virus, and clustering within hospitals. In the sensitivity analysis, we repeated the
128analysis using the intensive care use outcome. To examine the robustness of causal inference to
129potential unmeasured confounding, we also computed E-values by using the R *Evalue* package
130[16]. The E-value is a method to gauge the evidence for causality, which indicates the minimum
131magnitude of association that an unmeasured confounder would need to have with both the
132exposure and outcome to fully explain away the association of interest, conditional on the
133measured covariates. For instance, an E-value of 2.0 means that the odds ratio (OR) for the
134association of an unmeasured confounder with both the exposure and outcome would have to be
135 ≥ 2.0 to fully explain away the observed exposure-outcome association.

136 All P-values were 2-tailed, with a P-value of <0.05 considered statistically significant.
137We corrected for multiple testing using the Benjamini-Hochberg false-discovery rate (FDR)
138method. We performed the analysis using R version 3.6.1.

SUPPLEMENTARY REFERENCES

- Hasegawa K, Mansbach JM, Ajami NJ, et al. Association of nasopharyngeal microbiota profiles with bronchiolitis severity in infants hospitalised for bronchiolitis. *Eur Respir J*, **2016**; 48: 1329–39.
- Emergency Medicine Network [Internet]. [cited 2021 Mar 26]. Available from: <http://www.emnet-usa.org/>
- Ralston SL, Lieberthal AS, Meissner HC, et al. Clinical practice guideline: the diagnosis, management, and prevention of bronchiolitis. *Pediatrics*, **2014**; 134: e1474-1502.
- Hasegawa K, Jartti T, Mansbach JM, et al. Respiratory syncytial virus genomic load and disease severity among children hospitalized with bronchiolitis: multicenter cohort studies in the United States and Finland. *J Infect Dis*, **2015**; 211: 1550–9.
- Mansbach JM, Piedra PA, Stevenson MD, et al. Prospective multicenter study of children with bronchiolitis requiring mechanical ventilation. *Pediatrics*, **2012**; 130: e492-500.
- Hasegawa K, Mansbach JM, Ajami NJ, et al. The relationship between nasopharyngeal CCL5 and microbiota on disease severity among infants with bronchiolitis. *Allergy*, **2017**; 72: 1796–800.
- Hasegawa K, Hoptay CE, Harmon B, et al. Association of type 2 cytokines in severe rhinovirus bronchiolitis during infancy with risk of developing asthma: A multicenter prospective study. *Allergy*, **2019**; 74: 1374–7.

1588. Stewart CJ, Mansbach JM, Wong MC, et al. Associations of nasopharyngeal metabolome
159 and microbiome with severity among infants with bronchiolitis. A multiomic analysis. *Am*
160 *J Respir Crit Care Med*, **2017**; 196: 882–91.
1619. Lu X, Holloway B, Dare RK, et al. Real-time reverse transcription-PCR assay for
162 comprehensive detection of human rhinoviruses. *J Clin Microbiol*, **2008**; 46: 533–9.
16310. Olink [Internet]. [cited 2021 Mar 25]. Available from: <https://www.olink.com/>
16411. Assarsson E, Lundberg M, Holmquist G, et al. Homogenous 96-plex PEA immunoassay
165 exhibiting high sensitivity, specificity, and excellent scalability. *PLoS ONE*, **2014**; 9(4):
166 e95192.
16712. Stekhoven DJ and Bühlmann P. MissForest—non-parametric missing value imputation for
168 mixed-type data, *Bioinformatics*, **2012**; 28(1): 112–118.
16913. Kuhn M. Building predictive models in R using the caret package. *J Stat Softw*, **2008**;
170 28(5): 1-26.
17114. Donaldson J, Donaldson MJ. Package ‘tsne’. *CRAN Repository*, **2010**.
17215. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association
173 networks with increased coverage, supporting functional discovery in genome-wide
174 experimental datasets. *Nucleic Acids Res*, **2019**; 47: D607-613.

17516. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-
176 value. *Ann Intern Med*, **2017**; 167: 268–74.
177

178**Table S1. Principal investigators at the 17 participating sites in MARC-35**

Amy D. Thompson, MD	Alfred I. duPont Hospital for Children, Wilmington, DE
Federico R. Laham, MD, MS	Arnold Palmer Hospital for Children, Orlando, FL
Jonathan M. Mansbach, MD, MPH	Boston Children's Hospital, Boston, MA
Vincent J. Wang, MD, MHA and Susan Wu, MD	Children's Hospital of Los Angeles, Los Angeles, CA
Michelle B. Dunn, MD and Jonathan M.Spergel, MD, PhD	Children's Hospital of Philadelphia, Philadelphia, PA
Juan C. Celedón, MD, DrPH	Children's Hospital of Pittsburgh, Pittsburgh, PA
Michael R. Gomez, MD, MS-HCA and Nancy Inhofe, MD	The Children's Hospital at St. Francis, Tulsa, OK
Brian M. Pate, MD and Henry T. Puls, MD	The Children's Mercy Hospital & Clinics, Kansas City, MO
Stephen J. Teach, MD, MPH	Children's National Medical Center, Washington, D.C.
Richard T. Strait, MD and Stephen C. Porter, MD, MSc, MPH	Cincinnati Children's Hospital and Medical Center, Cincinnati, OH
Ilana Y. Waynik, MD	Connecticut Children's Medical Center, Hartford, CT
Sujit Iyer, MD	Dell Children's Medical Center of Central Texas, Austin, TX
Michelle D. Stevenson, MD, MS	Norton Children's Hospital, Louisville, KY
Margaret Samuels-Kalow, MD, MPhil, Wayne G. Shreffler, MD, PhD, and Ari R. Cohen, MD	Massachusetts General Hospital, Boston, MA
Anne K. Beasley, MD and Cindy S. Bauer, MD	Phoenix Children's Hospital, Phoenix, AZ
Thida Ong, MD and Markus Boos, MD, PhD	Seattle Children's Hospital, Seattle, WA
Charles G. Macias, MD, MPH	Texas Children's Hospital, Houston, TX

179

180 **Table S2. Comparisons between analytic and non-analytic cohorts in MARC-35**

	Analytic cohort*	Non-analytic cohort*	P-value
Patient characteristics	(n=117)	(n=899)	
Baseline characteristics			
Age (month), median (IQR)	3 (1–6)	3 (2–6)	0.40
Male sex	71 (60.7)	539 (60.0)	0.96
Race/ethnicity			0.70
Non-Hispanic white	46 (39.3)	384 (42.7)	
Non-Hispanic black	31 (26.5)	208 (23.1)	
Hispanic	37 (31.6)	271 (30.1)	
Other or unknown	3 (2.6)	36 (4.0)	
Maternal smoking during pregnancy	18 (15.4)	129 (14.6)	0.94
C-section delivery	43 (36.8)	305 (34.6)	0.73
Prematurity (32.0–36.9 weeks)	22 (18.8)	164 (18.2)	0.98
History of eczema	16 (13.7)	133 (14.8)	0.85
Mostly breastfed for the first 3 month of age	57 (51.8)	366 (47.2)	0.42
Cigarette smoke exposure at home	11 (9.4)	145 (16.1)	0.08
Corticosteroid use before the index hospitalization	23 (19.7)	124 (13.8)	0.12
Clinical presentation			
Weight at presentation (kg), median (IQR)	5.8 (4.4–7.4)	6.1 (4.7–7.7)	0.24
Respiratory rate at presentation (per minute), median (IQR)	48 (40–60)	48 (40–60)	0.31
Oxygen saturation at presentation			0.47

<90%	10 (8.8)	81(9.1)	
90-93%	23 (20.4)	132 (15.0)	
≥94%	80 (70.8)	669 (75.9)	
Detected pathogen			
RSV	81 (69.2)	740 (82.3)	<0.001
RSV-only	66 (56.4)	520 (57.8)	0.85
RV	51 (43.6)	163 (18.1)	<0.001
Other pathogen [†]	16 (13.7)	221 (24.6)	0.01

Abbreviations: IQR, interquartile range; RSV, respiratory syncytial virus; RV, rhinovirus

*Data are no. (%) of infants unless otherwise indicated.

[†]Adenovirus, bocavirus, Bordetella pertussis, enterovirus, human coronavirus NL63, OC43, 229E, or HKU1, human metapneumovirus, influenza A or B virus, Mycoplasma pneumoniae, parainfluenza virus 1-3.

182**Table S3. Patient characteristics according to respiratory syncytial virus and rhinovirus**
183**infection**

Patient characteristics	RSV* (n=66)	RV* (n=51)	P-value
Baseline characteristics			
Age (month), median (IQR)	2 (1–5)	4 (2–7)	0.003
Male sex	38 (57.6)	33 (64.7)	0.55
Race/ethnicity			0.49
Non-Hispanic white	28 (42.4)	18 (35.3)	
Non-Hispanic black	19 (28.8)	12 (23.5)	
Hispanic	17 (25.8)	20 (39.2)	
Other or unknown	2 (3.0)	1 (2.0)	
Maternal smoking during pregnancy	10 (15.2)	8 (15.7)	0.99
C-section delivery	25 (37.9)	18 (35.3)	0.93
Prematurity (32.0–36.9 weeks)	11 (16.7)	11 (21.6)	0.66
History of eczema	6 (9.1)	10 (19.6)	0.17
Mostly breastfed for the first 3 month of age	30 (48.4)	27 (56.2)	0.53
Cigarette smoke exposure at home	6 (9.1)	5 (9.8)	0.99
Corticosteroid use before the index hospitalization	9 (13.6)	14 (27.5)	0.10
Clinical presentation			
Positive pressure ventilation use	18 (27.3)	3 (5.9)	0.006
Intensive care use	33 (50.0)	13 (25.5)	0.01
Weight at presentation (kg), median (IQR)	5.2 (4.2–6.6)	6.7 (5.3–8.2)	0.01
Respiratory rate at presentation (per minute), median (IQR)	48 (38–60)	48 (40–60)	0.36
Oxygen saturation at presentation			0.91
<90%	6 (9.5)	4 (8.0)	
90–93%	14 (22.2)	9 (18.0)	
≥94%	43 (68.3)	37 (74.0)	

Abbreviations: IQR, interquartile range; RSV, respiratory syncytial virus; RV, rhinovirus

* Data are no. (%) of infants unless otherwise indicated.

184

185**Table S4. Multivariable associations of respiratory syncytial virus and rhinovirus infection**
186**with serum proteins in infants with severe bronchiolitis**

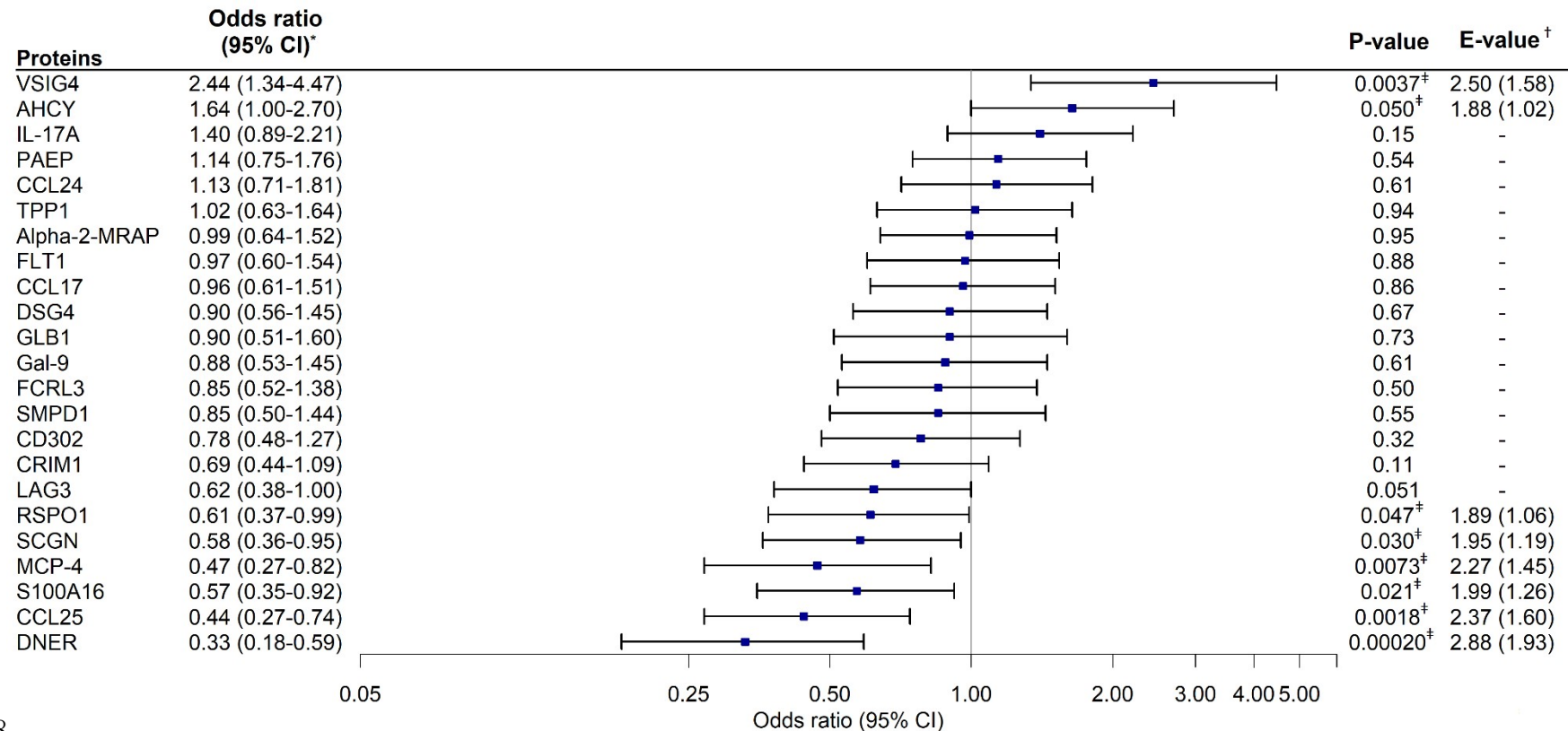
Protein[*]	Coefficient[†]	Standard error[†]	P value[†]	FDR[†]
AHCY	-0.56	0.13	<0.001	<0.001
VSIG4	-0.52	0.11	<0.001	<0.001
Gal-9	-0.49	0.12	<0.001	<0.001
LAG3	-0.39	0.13	0.0031	0.0094
FLT1	-0.36	0.12	0.0039	0.0099
PAEP	-0.33	0.13	0.014	0.027
SMPD1	-0.32	0.11	0.0049	0.011
CRIM1	-0.31	0.13	0.024	0.037
Alpha-2-MRAP	-0.3	0.13	0.022	0.037
SCGN	-0.3	0.13	0.021	0.036
TPP1	-0.28	0.12	0.026	0.039
CRX	-0.25	0.13	0.062	0.071
CTSD	-0.23	0.13	0.082	0.091
GLB1	-0.23	0.11	0.035	0.045
IL-10RA	-0.18	0.13	0.18	0.19
ERP44	-0.13	0.12	0.27	0.28
ADAM-TS13	-0.09	0.14	0.53	0.53
NCF2	0.24	0.12	0.054	0.065
DSG4	0.27	0.12	0.031	0.042
IL32	0.28	0.14	0.04	0.05
RSP01	0.29	0.13	0.029	0.041
IL-17A	0.33	0.13	0.014	0.027
CCL25	0.34	0.13	0.014	0.027
FCRL3	0.37	0.12	0.0039	0.0099
CCL24	0.39	0.12	0.0022	0.0074
CD302	0.41	0.12	0.0011	0.0056
CCL17	0.42	0.13	0.0017	0.0072
DNER	0.42	0.13	0.002	0.0074
S100A16	0.43	0.13	0.001	0.0056
MCP-4/CCL13	0.44	0.12	<0.001	0.0043

Abbreviations: CI, confidence interval; FDR, false discovery rate

* Of 1,109 measured serum proteins, the top 30 discriminating (respiratory syncytial virus vs. rhinovirus) proteins with a large regression coefficient in the elastic net model are examined. Of these 30 proteins, 23 proteins with bold letters have a significant association (FDR<0.05).

† Multivariable random-effect linear models adjusted for age, sex, and clustering within hospitals were fit. The coefficients represent between-virus differences in protein levels using autoscaling after log₂ transformation (the respiratory syncytial virus [RSV] group is the reference). A negative coefficient indicates a higher adjusted mean protein level in RSV infection compared to rhinovirus (RV) infection. In contrast, a positive coefficient indicates a higher adjusted mean protein level in the RV infection group compared to the RSV group.

187**Figure S1. Multivariable associations of serum proteins with intensive treatment in infants with severe bronchiolitis**



188

189Abbreviation: CI, confidence interval

190* For 23 proteins that have a FDR of <0.05 in the virus-protein models (**Table S4**), multivariable random-effect logistic models adjusted for age, sex, respiratory

191virus, and clustering within hospitals were fit. The odds ratios were estimated per one standard deviation increment in log2-transformed protein level.

192†The E-value (with its lower 95% confidence interval bound) represents how strongly an unmeasured confounder would have to be associated with the exposure
193and outcome in order for the observed association to be independent.

194†P-value <0.05