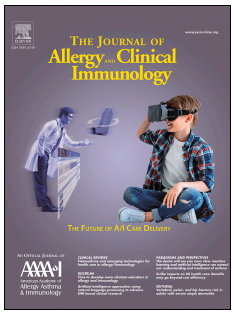
Journal Pre-proof



Integrated-omics endotyping of infants with rhinovirus bronchiolitis and risk of childhood asthma

Yoshihiko Raita, MD, MPH, MMSc, Carlos A. Camargo, Jr., MD, DrPH, FAAAAI, Yury A. Bochkov, PhD, Juan C. Celedón, MD, DrPH, FAAAAI, James E. Gern, MD, FAAAAI, Jonathan M. Mansbach, MD, MPH, Eugene P. Rhee, MD, Robert J. Freishtat, MD, MPH, Kohei Hasegawa, MD, MPH

PII: S0091-6749(20)31585-2

DOI: https://doi.org/10.1016/j.jaci.2020.11.002

Reference: YMAI 14829

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 20 July 2020

Revised Date: 24 October 2020

Accepted Date: 2 November 2020

Please cite this article as: Raita Y, Camargo Jr. CA, Bochkov YA, Celedón JC, Gern JE, Mansbach JM, Rhee EP, Freishtat RJ, Hasegawa K, Integrated-omics endotyping of infants with rhinovirus bronchiolitis and risk of childhood asthma, *Journal of Allergy and Clinical Immunology* (2020), doi: https://doi.org/10.1016/j.jaci.2020.11.002.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier Inc. on behalf of the American Academy of Allergy, Asthma & Immunology.

Journal Pre-proof

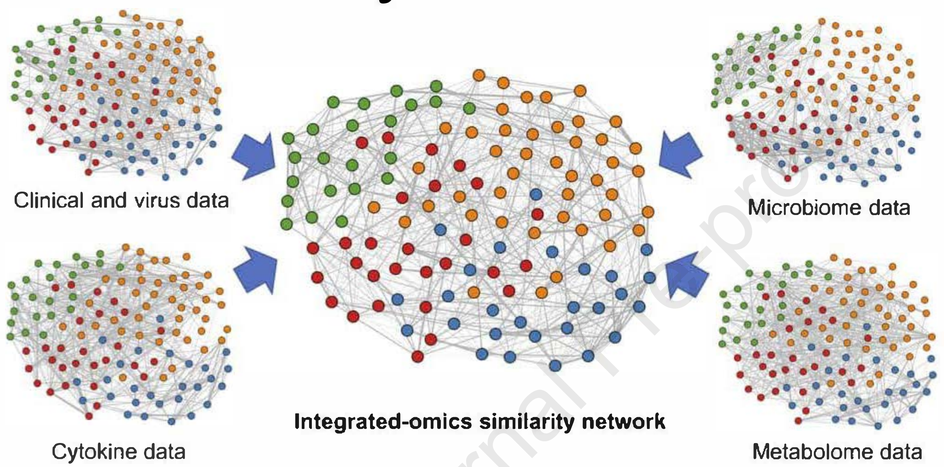
Integrated-omics endotyping of infants with rhinovirus bronchiolitis and risk of childhood asthma

Integrated-omics endotyping

122 infants hospitalized for rhinovirus bronchiolitis

- Clinical and virus data
- Microbiome data
- Cytokine data
- Metabolome data

Similarity network fusion



Derived endotypes

- A: virus$^{RV\text{-}C}$microbiome^{mixed}T2^{low}
- B: virus$^{RV\text{-}A}$microbiome$^{\textit{Haemophilus}}$T2^{low}
- C: virus$^{RSV/RV}$microbiome$^{\textit{Streptococcus}}$T2^{low}
- D: virus$^{RV\text{-}C}$microbiome$^{\textit{Moraxella}}$T2^{high}

Association of endotypes with clinical outcomes

Highest risk

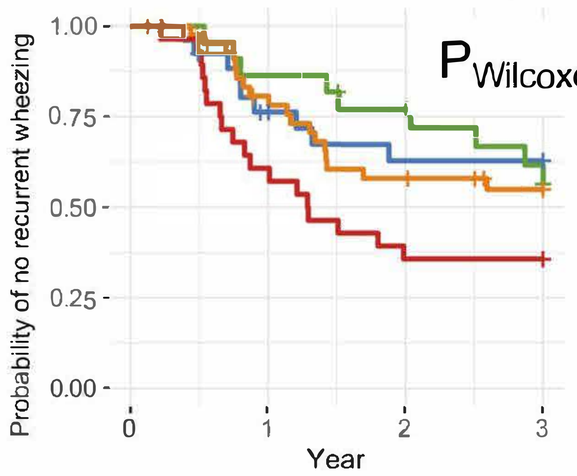
D: virus$^{RV\text{-}C}$microbiome$^{\textit{Moraxella}}$T2^{high}

vs.

Lowest risk

A: virus$^{RV\text{-}C}$microbiome^{mixed}T2^{low}

Recurrent wheeze by age 3 years



$P_{Wilcoxon}$=0.04

Hazard ratio
2.23 (1.00 – 4.96)
P-value=0.049

Asthma at age 5 years

Odds ratio
3.74 (1.21 – 12.6)
P-value=0.03

Abbreviations: RV-C, rhinovirus C; RV-A, rhinovirus A; RSV, respiratory syncytial virus; T2, type 2 cytokine

Journal Pre-proof

Integrated-omics endotyping of infants with rhinovirus bronchiolitis and risk of childhood asthma

Authors: Yoshihiko Raita, MD, MPH, MMSc[1]; Carlos A. Camargo, Jr., MD, DrPH, FAAAAI[1]; Yury A. Bochkov, PhD[2]; Juan C. Celedón, MD, DrPH, FAAAAI[3]; James E. Gern, MD, FAAAAI[2,4]; Jonathan M. Mansbach, MD, MPH[5]; Eugene P. Rhee, MD[6]; Robert J. Freishtat, MD, MPH[7,8,9]; and Kohei Hasegawa, MD, MPH[1]

Affiliations:

1. Department of Emergency Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
2. Department of Pediatrics, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA
3. Division of Pulmonary Medicine, Department of Pediatrics, UPMC Children's Hospital of Pittsburgh, University of Pittsburgh, Pittsburgh, PA
4. Department of Medicine, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, USA
5. Department of Pediatrics, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA
6. Nephrology Division and Endocrine Unit, Department of Medicine, Massachusetts General Hospital, Boston, Harvard Medical School, Boston, MA, USA
7. Division of Emergency Medicine, Children's National Hospital, Washington, DC, USA

Journal Pre-proof

Journal Pre-proof

8. Department of Pediatrics, George Washington University School of Medicine and Health Sciences, Washington, DC, USA
9. Department of Genomics and Precision Medicine, George Washington University School of Medicine and Health Sciences, Washington, DC, USA

Corresponding Author: Dr. Yoshihiko Raita, Department of Emergency Medicine, Massachusetts General Hospital, Harvard Medical School, 125 Nashua Street, Suite 920, Boston, MA 02114-1101. Tel: 617-726-5276, FAX: 617-724-4050.
E-mail: yraita1@mgh.harvard.edu

Word count: 3388

Authors Contribution: Y.R. carried out the main statistical analysis, drafted the initial manuscript, and approved the final manuscript as submitted. C.A.C. conceptualized and designed the study, obtained funding, collected the data, supervised the conduct of study and the analysis, critically reviewed and revised the initial manuscript, and approved the final manuscript as submitted. Y.A.B. and J.E.G. conducted virus testing, reviewed and revised the manuscript, and approved the final manuscript as submitted. J.C.C and J.M.M collected the data, reviewed and revised the manuscript, and approved the final manuscript as submitted. E.P.R assisted in the study design and data interpretation, reviewed and revised the manuscript, and approved the final manuscript as submitted. R.J.F conceptualized the study, conducted cytokine testing, reviewed and revised the manuscript, and approved the final manuscript as submitted. K.H. conceptualized the study, obtained funding, supervised the statistical analysis, reviewed and revised the initial manuscript, and approved the final manuscript as submitted.

Journal Pre-proof

Funding/Support: This study was supported by grants from the National Institutes of Health (Bethesda, MD): U01 AI-087881, R01 AI-114552, R01 AI-108588, R01 AI-134940, R21 HL-129909, and UG3/UH3 OD-023253. 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 organizations were 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: Dr. Celedón has received research materials from Merck and GSK (inhaled corticosteroids) and Pharmavite (vitamin D and placebo capsules) in order to provide medications free of costs to participants in NIH-funded studies, unrelated to the current work. Dr. Bochkov has patents on production methods of rhinoviruses. Dr. Gern is a paid consultant to AstraZeneca and Meissa Vaccines Inc., has stock options in Meissa Vaccines Inc., and has patents on production methods of rhinoviruses. The other authors have no financial relationships relevant to this article to disclose.

Journal Pre-proof

Journal Pre-proof

ABSTRACT

Background: Young children with rhinovirus (RV) infection—particularly bronchiolitis—are at high risk for developing childhood asthma. Emerging evidence suggests clinical heterogeneity within RV bronchiolitis. However, little is known about these biologically-distinct subgroups (endotypes) and their relations with asthma risk.

Objective: We aimed to identify RV bronchiolitis endotypes and examine their longitudinal relations with asthma risk.

Methods: As part of multicenter prospective cohort study of infants (age <12 months) hospitalized for bronchiolitis, we integrated clinical, RV species, nasopharyngeal microbiome (16S rRNA gene sequencing), cytokine, and metabolome (liquid chromatography–tandem mass spectrometry) data collected at hospitalization. We then applied network and clustering approaches to identify bronchiolitis endotypes. We also examined their longitudinal association with risks of developing recurrent wheeze by age 3 years and asthma by age 5 years.

Results: Of 122 infants hospitalized for RV bronchiolitis (median age, 4 months), we identified four distinct endotypes—mainly characterized by RV species, microbiome, and type 2 cytokine (T2) response: A) virus$^{\text{RV-C}}$microbiome$^{\text{mixed}}$T2$^{\text{low}}$, B) virus$^{\text{RV-A}}$microbiome$^{\textit{Haemophilus}}$T2$^{\text{low}}$, C) virus$^{\text{RSV/RV}}$microbiome$^{\textit{Streptococcus}}$T2$^{\text{low}}$, and D) virus$^{\text{RV-C}}$microbiome$^{\textit{Moraxella}}$T2$^{\text{high}}$. Compared with endotype A infants, endotype D infants had a significantly higher rate of recurrent wheeze (33% vs. 64%; HR, 2.23; 95%CI, 1.00-4.96; P=0.049) and higher risk for developing asthma (28% vs. 59%; OR, 3.74: 95%CI, 1.21-12.6; P=0.03)

Conclusion: Integrated-omics analysis identified biologically-meaningful RV bronchiolitis endotypes in infants, such as one characterized by RV-C infection, *Moraxella*-dominant

85 microbiota, and high T2 cytokine response, at higher risk for developing recurrent wheeze and
86 asthma. This study should facilitate further research to validating our inferences.
87 **Word count**: 250
88
89 **Clinical Implications:** Our findings lend support to the concept that "bronchiolitis" represents
90 several diseases that have distinct biological mechanisms, and provide an evidence base for the
91 early identification of high-risk children
92
93 **Capsule Summary:** By integrating clinical, virus, nasopharyngeal microbiome, cytokine, and
94 metabolome data, we identified four endotypes of RV bronchiolitis. Endotype characterized by
95 RV-C infection, *Moraxella*-dominant microbiota, high type 2 cytokine response had the highest
96 risk for asthma.
97
98 **Keywords:** integrated-omics, microbiome, metabolome, cytokines, endotyping, bronchiolitis,
99 asthma
100
101 **Abbreviations:** MARC-35, the 35th Multicenter Airway Research Collaboration; RV, rhinovirus;
102 RV-A, rhinovirus A; RV-C, rhinovirus C; RSV, respiratory syncytial virus; T2, type 2 cytokines;
103 IQR, interquartile range; ICU, intensive care unit; sIgE, specific immunoglobulin E; CI,
104 confidential interval; PCR, polymerase chain reaction; FDR, false discovery rate; IL, interleukin;
105 TSLP, thymic stromal lymphopoietin; IFN-beta, interferon-beta; MIP-1, macrophage
106 inflammatory protein-1; RANTES, regulated upon activation, normal T cell expressed and

107 presumably secreted; DHA, docosahexaenoic acid; GPS, glycerophosphatidylserine; GPC,
108 glycerophosphorylcholine; GPE, glycerophosphoethanolamine; GPG, glycerophosphoglycerol
109
110

Journal Pre-proof

Journal Pre-proof

INTRODUCTION

Viral respiratory infection in infancy is an important asthma risk factor, with the largest population attributable fraction for childhood asthma. [1] One such infection, bronchiolitis, is the leading cause of hospitalization in US infants, accounting for ~110,000 hospitalizations annually.[2] Approximately 20%-40% of infants hospitalized for bronchiolitis develop recurrent wheeze and asthma in childhood.[3–8] Of the major causative pathogens, infection with rhinovirus (RV)—consisting of three species (A, B and C)—has a stronger association with these chronic airway morbidities (e.g., recurrent wheeze and childhood asthma).[9,10]

While bronchiolitis has been considered a single disease with similar mechanisms, recent studies have shown heterogeneity in clinical presentations,[11,12] chronic morbidities (e.g., subsequent risk of asthma),[9,13] and upper airway microbiome,[14–16] cytokine,[17,18] and metabolome[19] profiles among infants with bronchiolitis. However, these studies have relied *solely* on single-level data (e.g., only clinical or microbiome data). To our knowledge, no study has used clinical, virus, and multi-level omics data to investigate the endotypes of bronchiolitis—let alone RV bronchiolitis—and their longitudinal relation with recurrent wheeze and asthma. The inadequate understanding of bronchiolitis endotypes during infancy—an important period of lung development—has hindered efforts to develop endotype-specific bronchiolitis treatment and asthma prevention strategies in this high-risk population.

To address this knowledge gap, we aimed to identify biologically- and clinically-meaningful endotypes of RV bronchiolitis through integrating clinical, virus, nasopharyngeal microbiome, cytokine and metabolome data from a multicenter prospective cohort study. We hypothesized that these endotypes have differential risks of developing recurrent wheeze and asthma in later childhood.

Journal Pre-proof

METHODS

Study Design, Setting, and Participants

This is an analysis of the data from a multicenter prospective cohort study of infants hospitalized for bronchiolitis—the 35th Multicenter Airway Research Collaboration (MARC-35) study.[15] MARC-35 is coordinated by the Emergency Medicine Network (EMNet, www.emnet-usa.org), an international research collaboration with 247 participating hospitals. Details of the study design, setting and participants, data collection, testing, and statistical analysis may be found in the **Supplementary Methods** in **Online Supplement**. Briefly, at 17 sites across 14 U.S. states (**Table E1**), MARC-35 enrolled infants (aged <1 year) who were hospitalized with an attending physician diagnosis of bronchiolitis during three consecutive bronchiolitis seasons (November 1 to April 30) from 2011 to 2014. The diagnosis of bronchiolitis was made according to the American Academy of Pediatrics bronchiolitis guidelines, defined as acute respiratory illness with a combination of rhinitis, cough, tachypnea, wheezing, crackles, or retractions.[20] We excluded infants who were transferred to a participating hospital >24 hours after initial hospitalization or with a preexisting heart and lung disease, immunodeficiency, immunosuppression or gestational age of <32 weeks. All patients were treated at the discretion of the treating physicians. The institutional review board at each participating hospital approved the study with a written informed consent obtained from parent or guardian. Of 921 infants enrolled into the longitudinal cohort, the current analysis investigated 122 infants with RV infection who underwent nasopharyngeal microbiome, cytokine, and metabolomics testing (**Figure E1**).

Data Collection, Virus, Microbiome, Cytokine, and Metabolomics Testing

Clinical data (patients' demographic characteristics, and family, environmental, and medical history, and details of the acute illness) were collected via structured interview and chart reviews using a standardized protocol.[15,21] Additionally, nasopharyngeal samples were collected within 24 hours of hospitalization using a standardized protocol.[15] The details of the data collection methods are described in the **Supplementary Methods**. We conducted A) real-time PCR to test for respiratory viruses, with subsequent RV genotyping,[15] B) 16S rRNA gene sequencing to profile microbiome,[15] C) electrochemiluminescence immunoassays and enzyme-linked immunosorbent assay to measure cytokines,[18] and D) liquid chromatography–tandem mass spectrometry (LC-MS/MS) to profile the metabolome.[19]

Real-time PCR to Test for Respiratory Viruses

The details of real-time PCR to test for respiratory viruses are described in the **Supplementary Methods**. For RV detection, complementary DNA was generated using virus-specific primers for RV, and singleplex real-time PCR was used. Next, to identify the RV species (A, B, and C), RV-positive samples were partially sequenced to determine the RV species and type[22] at the University of Wisconsin-Madison (Madison, WI). To quantify the relative RV genomic load, cycle threshold (CT) values were used.

Nasopharyngeal Airway Microbiome Profiling

The details of microbiome profiling and quality control measures are described in **Supplementary Methods**. In brief, 16S rRNA gene sequencing methods were adapted from the methods developed for the National Institutes of Health (NIH) Human Microbiome Project.[23] Bacterial genomic DNA was extracted using a MO BIO PowerSoil DNA Isolation Kit (Mo Bio

Laboratories, Carlsbad, CA, USA).[24] The 16S rDNA V4 region was amplified by PCR and sequenced in the MiSeq platform (Illumina, SanDiego, CA, USA) using a 2×250 bp paired-end protocol that yields paired-end reads which overlap almost completely. Sequencing read pairs were demultiplexed based on the unique molecular barcodes, and reads were merged using USEARCH v7.0.1090.[25] Rarefaction curves of bacterial operational taxonomic units (OTUs) were constructed using sequence data for each sample to ensure coverage of the bacterial diversity present. Samples with suboptimal amounts of sequencing reads were re-sequenced to ensure that the majority of bacterial taxa were encompassed in our analyses. 16S rRNA gene sequences were clustered into OTUs at a similarity cut-off value of 97% using the UPARSE algorithm.[26] OTUs were determined by mapping the centroids to the SILVA database.[27]

Nasopharyngeal Cytokines Measurement

The details of cytokine measurement are described in **Supplementary Methods**. Investigators collected nasopharyngeal airway samples within 24 hours of hospitalization by using a standardized protocol. Ten nasopharyngeal cytokines (interleukin-4 [IL□4], IL□5, IL□10, IL□12, IL□13, IL□25, IL□33, interferon [IFN]□β, macrophage inflammatory protein [MIP]□1α, thymic stromal lymphopoietin [TSLP]) were quantified using electrochemiluminescence immunoassays. Additionally, RANTES concentration was quantified by using an enzyme-linked immunosorbent assay.

Nasopharyngeal Metabolome Testing

The details of metabolomic profiling are described in **Supplementary Methods**. In brief, metabolome profiling used LC-MS/MS by Metabolon (Durham, NC, USA). Instrument variability

Journal Pre-proof

was 4%, as determined by calculating the median relative standard deviation for the internal standards that were added to each sample before the sample was injected into the mass spectrometer.

Outcome Measures

The clinical outcomes were 1) recurrent wheeze by age 3 years, and 2) asthma at age 5 years. The definition of recurrent wheeze was based on the 2007 NIH asthma guidelines as having at least two corticosteroid-requiring exacerbations in six months or at least four wheezing episodes in one year that last at least one day and affect sleep. The definition of asthma was based on commonly-used epidemiologic definition of asthma : physician-diagnosis of asthma at age 5 years, plus either asthma medication use (e.g., albuterol inhaler, inhaled corticosteroids, montelukast) or asthma-related symptoms in the preceding year.

Statistical Analysis

In this study, the inferential goals are to identify biologically-distinct RV bronchiolitis endotypes (description [clustering]) and to relate them to the outcome risks (association). The analytic flow is summarized in **Figure 1**; the details of the statistical analysis may be found in the **Supplementary Methods**. Briefly, we first preprocessed variables to compute a distance matrix for each of datasets (i.e., clinical and virus, microbiome, cytokines, and metabolome). For the clinical data, to generate a distance matrix from the mixed-type data (i.e., data with binary, categorical and continuous variables), we computed Gower distance using variables that are selected based on *a priori* knowledge—age, sex, birth weight, history of breathing problem(s), lifetime antibiotic use, parental asthma, blood eosinophilia of $\geq$4%, specific immunoglobulin E

(sIgE) sensitization, and viruses (RV species, RV genomic load, coinfection with respiratory syncytial virus [RSV]).[1,9] For the microbiome data, we computed Bray-Curtis distance using the relative abundance data (i.e., total sum scaling) of 20 most abundant genera, which accounted for 97% of total abundance. For the cytokine data, we included all measured cytokines. To derive a distance matrix for the cytokine data, we computed Euclidian distance using all measured cytokines with normalization with RNase P (host housekeeping gene) cycling threshold value, log2 transformation, and Pareto scaling. For metabolome data, we computed the Euclidian distance using 100 metabolites with highest variances with adjustment for batch effects using R *sva* package.[31] Second, to conduct integrated-omics endotyping, we computed an affinity matrix of each dataset separately, and computed a fused affinity matrix by similarity network fusion using R *SNFtool* package.[32] Lastly, to identify mutually-exclusive endotypes, we conducted spectral clustering using the fused affinity matrix.[32] To choose the optimal number of endotypes, we used both eigengaps and clinical plausibility. To determine the magnitude of variables' contribution to endotypes, we also computed the ranking of normalized mutual information score of each variable.[32]

To examine the longitudinal relation of derived endotypes with recurrent wheeze, we modeled the time to recurrent wheeze by fitting a Cox proportional hazards model. Patients who did not have an outcome were censored at their last follow-up interview or at the time of withdrawal during the 36-month follow-up period. The proportionality of hazards assumption was verified through evaluating Schoenfeld residuals. To determine the association of endotypes with the risk of childhood asthma, we fitted a logistic regression model. To detect biologically-meaningful pathways, we performed metabolite set enrichment analysis using R *MetaboAnalystR* package.[33]

Journal Pre-proof

In the sensitivity analysis, to address potential heterogeneity of asthma, we constructed the model with stratifying the asthma outcome by recurrent wheeze. Next, we also examined the robustness of endotype-outcome associations using different numbers of endotypes. To visualize consistencies across the different numbers chosen (3, 4, or 5 endotypes), we used alluvial plot. Then, we examined the between-endotype differences in clinical, virus, microbiome, cytokine, and metabolome data, and estimated the association of endotypes with asthma using a logistic regression model. We analyzed the data using R version 3.5.1 (R foundation for Statistical Computing, Vienna, Austria). All P-values were two-tailed, with $P<0.05$ considered statistically significant. We corrected for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) method.

Journal Pre-proof

RESULTS

In this multicenter prospective cohort study, 122 infants with RV bronchiolitis who underwent nasopharyngeal microbiome, cytokine, and metabolome testing were eligible for the current analysis (**Figure E1**). The analytic cohort and non-analytic cohort did not differ in most patient characteristics, except for lifetime antibiotic use, IgE sensitization, and RSV coinfection ($P \geq 0.05$; **Table E2**). Among the analytic cohort, the median age was 4 (IQR, 2-6) months, 34% were female, and 41% were non-Hispanic white.

Integrated-omics Endotyping

By integrating clinical, virus, nasopharyngeal microbiome, cytokine, and metabolome data, we identified 4 distinct endotypes among infants with RV bronchiolitis (**Figure 2**), with the number of endotypes chosen according to eigengap values and silhouette scores (**Figure E2**) as well as clinical plausibility.[1] These endotypes were *chiefly* characterized by RV species, major bacteria genera in the nasopharyngeal airway microbiome, and type 2 cytokine (T2) response: A) virus$^{\text{RV-C}}$microbiome$^{\text{mixed}}$T2$^{\text{low}}$ (22%), B) virus$^{\text{RV-A}}$microbiome$^{\textit{Haemophilus}}$T2$^{\text{low}}$ (19%), C) virus$^{\text{RSV/RV}}$microbiome$^{\textit{Streptococcus}}$T2$^{\text{low}}$ (36%), and D) virus$^{\text{RV-C}}$microbiome$^{\textit{Moraxella}}$T2$^{\text{high}}$ (23%) (**Figure 2**).

Across the endotypes, several clinical characteristics (e.g., lifetime antibiotic use, parental history of asthma, blood eosinophilia) were different (**Tables 1**). Additionally, there were significant differences in all measured cytokines—including type 2 (IL-4, IL-5, and IL-13) and epithelial-derived (IL-17E/IL-25, IL-33, and TSLP) cytokines (all $P<0.05$; **Table E3**). Similarly, there were significant differences in metabolites, predominantly lipid metabolites such as sphingolipids and arachidonate ($P<0.05$; **Table E3**). These characteristics with significant

between-endotype difference also had a high-ranked normalized mutual information score—i.e., a large contribution to the similarity network (**Figure E3**).

Descriptively, endotype A was characterized by younger age, low proportion of lifetime antibiotics use and parent history of asthma, low proportion of positive pressure ventilation use, high proportion of RV-C, mixed microbiome profile, and generally-low cytokine profile. Endotype B was characterized by RV-A, *Haemophilus* dominance, and low T2 cytokine profile. Endotype C was characterized by high proportion of positive pressure ventilation use, RV-RSV coinfection, *Streptococcus* dominance, and low T2 but high IFN-β and RANTES levels. Lastly, endotype D was characterized by older age, lifetime antibiotic use, blood eosinophilia, IgE sensitization, RV-C, *Moraxella* dominance, and high T2 cytokine profile (**Figure 3** and **Table E3**).

Association with Risks of Developing Recurrent Wheeze and Asthma

Overall, the Kaplan-Meier curves for recurrent wheeze outcome significantly differed between the endotypes ($P_{Wilcoxon}$=0.04; **Figure 4**). Specifically, the Cox proportional hazards model demonstrated that, compared with endotype A infants (virus$^{RV\text{-}C}$microbiome^{mixed}T2^{low}), the rate of developing recurrent wheeze was not significantly different in endotype B or C infants. In contrast, endotype D infants (virus$^{RV\text{-}C}$microbiome$^{\textit{Moraxella}}$T2^{high}) had a significantly higher rate of recurrent wheeze (33% vs. 64%; HR 2.23; 95%CI 1.00-4.96; P=0.049; **Table 2**). Likewise, compared with endotype A infants, only endotype D infants had a significantly higher risk for asthma at age 5 years (28% vs. 59%; OR 3.74; 95%CI 1.21-12.6; P=0.03; **Table 2**). To address potential heterogeneity of asthma, we further stratified the outcome by recurrent wheeze status. For example, despite the reduced sample size, compared with endotype A infants, endotype D

infants had a significantly higher risk of developing asthma after recurrent wheeze outcome (26% vs. 67%; OR 5.60; 95%CI 1.43–25.1; P=0.02; **Table E4**). In the examination of asthma risk without recurrent wheeze, endotype D infants had a non-significantly higher risk of outcome (13% vs. 40%; OR 4.67; 95%CI 0.71–41.0; P=0.12).

Comparison Between Endotypes A vs. D

To better understand the difference between the lowest-risk (endotype A) and highest-risk (endotype D) groups—both of which were also characterized by RV-C infection, we compared clinical, nasopharyngeal microbiome, cytokine, and metabolome data between these endotypes (**Tables E5-8**). Compared with endotype A infants, endotype D infants had a higher abundance of *Moraxella* (**Table E6**) and higher levels of type 2 (IL-4, IL-5, and IL-13) and epithelial-derived (IL-17E/IL-25, IL-33, TSLP) cytokines (all FDR<0.05; **Table E7**). Furthermore, endotype D infants had a higher intensity of specific lipid metabolites (e.g., sphingolipids, arachidonate, hydroxyoctadecadienoic acids [HODE]; FDR<0.05; **Table E8**). To examine biologically-meaningful pathways, we also performed metabolite set enrichment analysis. The analysis identified differentially-enriched pathways—e.g., arginine and proline, sphingolipid, and arachidonic acid metabolism pathways in the endotype A vs. D comparison (all FDR<0.05; **Table E9** and **Figure E4**).

Sensitivity Analysis

To examine the robustness of the findings, we examined different numbers of endotypes. Alluvial plot (**Figure E5**) demonstrated consistency of the original endotypes across the different numbers chosen (3, 4, or 5 endotypes). With the use of 5 endotypes, the first one had 100%

concordance with the original endotype A and the fifth one had 96% concordance with the original endotype D (**Table E10**). Also similar to the primary analysis, these 5 endotypes were also characterized by RV species, major nasopharyngeal bacteria genera (e.g., *Moraxella, Haemophilus, Streptococcus*), and type 2 cytokines (**Table E10**). Lastly, compared to endotype 1 (which is consistent with endotype A), endotype 5 (consistent with endotype D) infants had a non-significant but consistently higher risk of asthma at age 5 years (33% vs. 62%; OR 3.25; 95%CI 0.94-12.1; P=0.07).

Journal Pre-proof

Journal Pre-proof

DISCUSSION

By integrating clinical, RV species, nasopharyngeal microbiome, cytokine, and metabolome data from a multicenter prospective cohort study of 122 infants with RV bronchiolitis, we identified four biologically-distinct endotypes. These endotypes were mainly characterized by RV species, major bacteria of the microbiome, and cytokine responses. In particular, the endotype D—which is characterized by prior antibiotic use, eosinophilia, IgE sensitization, RV-C infection, *Moraxella*-dominant microbiota, high T2 cytokine response, and specific lipid profiles (e.g., enriched sphingolipids, arachidonate, HODE)—had the highest risk for developing recurrent wheeze and asthma in later childhood. To our knowledge, this is the first study that has identified biologically-meaningful endotypes in infants, let alone in infants with RV bronchiolitis, and demonstrated their longitudinal relations with the risk of developing recurrent wheeze and asthma.

Recent studies—which are mostly limited to clinical data or single data element (e.g., microbiome)—have examined the relationships between infant bronchiolitis and the development of its chronic morbidities. For example, in agreement with the current observations, we previously reported that infants with RV-C bronchiolitis with IgE sensitization, compared with those with RSV bronchiolitis, had a higher risk of developing recurrent wheeze.[9] Similarly, and also consistent with the current study, prior studies have demonstrated the association of unique airway microbiome profiles with the risk of recurrent wheeze and childhood asthma. For example, studies have reported associations between dominance/presence by *Moraxella* in the upper airway and a high risk of recurrent wheeze,[34,35] incident asthma,[36] and asthma exacerbation.[37,38] Additionally, recent population-based cohort study has demonstrated that, in agreement with our data, antibiotic exposures during early infancy are associated with a higher risk of developing asthma and its effect is mediated by the changes in nasal microbiome.[39] Studies

have also shown that enhanced T2 inflammatory response in the airway of infants with RV[18] and RSV[40] infection is associated with a higher risk of developing recurrent wheeze and asthma. Moreover, recent studies have also underscored the significance of the metabolome, which represents the downstream functional products of the microbiome, child's genetic make-up, and immune response. Indeed, studies have associated specific lipid profiles—e.g., sphingolipids, arachidonate, HODE (12/15–lipoxygenase metabolite of linoleic acid [an ω-6 polyunsaturated fatty acid])—in the airway with the pathobiology of bronchiolitis[19] and asthma. [41,42] The current study corroborates and extends these earlier findings, not only by identifying biologically-meaningful bronchiolitis endotypes in infants through integrated-omics endotyping but also by demonstrating their longitudinal relations with recurrent wheeze and asthma.

The exact mechanisms underlying the observed endotypes—particularly endotype D (virus$^{\text{RV-C}}$microbiome$^{\textit{Moraxella}}$T2$^{\text{high}}$ with a unique lipid profile) with the highest risk for recurrent wheeze and asthma—warrant clarification. Emerging evidence has demonstrated that, compared to RV-A and -B, RV-C infection has a distinct pathobiology, including its cellular receptor (human cadherin-related family member 3 [CDHR3]),[43] inflammatory profiles,[44] and its interaction with IgE sensitization.[9] Indeed, in hospitalized children (aged 24-72 months) with RV-C infection, nasal cytokine profile was characterized by high levels of T2 cytokines (e.g., IL-4, IL-9, IL-13).[44] Furthermore, RV infection —particularly RV-C infection—among infants[14] and school-age children[45] has been associated with *Moraxella*-dominant microbiome profile in the airway. Studies have also suggested that interrelations between *Moraxella* and host immune response in the airway play an important role in the airway inflammation. For example, *in vitro*, *Moraxella catarrhalis* upregulated gene expression of pro-inflammatory cytokines (e.g., IL-33) and led to greater magnitude of epithelial damage.[37] Lastly, recent studies have reported the role of lipids—major

structural components of cell membrane and molecular signaling—in the pathobiology of airway diseases. For example, sphingolipids play important roles in immune response, inflammation, apoptosis, and cell proliferation.[46] Sphingolipids have also attracted attention because *ORMDL3* gene within the 17q21 locus—one of the most important genomic regions in asthma pathogenesis—regulates *de novo* sphingolipid synthesis.[42,47] Study of human lung epithelial cell model has demonstrated that siRNA knockdown of *ORMDL3* modifies the rate-limiting step of sphingolipid production.[48] Additionally, ω-6 polyunsaturated fatty acids play important roles in asthma pathobiology. For example, arachidonate is a precursor of the arachidonic cascade that produces inflammatory eicosanoids—such as leukotrienes and prostaglandins. Likewise, research has also reported that HODE (another metabolite of ω-6 polyunsaturated fatty acids) dysregulates NF-κB and decreases GR-α transcript in human bronchial epithelial cell and allergic asthma mouse models.[49] Notwithstanding the complexity of these mechanisms, the identification of endotypes in infants with RV bronchiolitis, and their longitudinal relations with chronic outcomes, are important findings that can facilitate future research focusing on the potential target for the primary prevention of childhood asthma.

Limitations

Our study has several potential limitations. First, we focused on infants with RV bronchiolitis as RV infection during early childhood is known to have a stronger association with chronic airway morbidities (compared to non-RV pathogens, such as RSV).[9,10] Regardless, examining endotypes of non-RV bronchiolitis will be the focus of our future investigations. Second, bronchiolitis involves inflammation of the lower airways in addition to the upper airways. While our study is based on the nasopharyngeal samples, studies have shown that upper airway

sampling provides reliable representation of the lung microbiome[50] and transcriptomic[51] profiles. Furthermore, the use of upper airway specimens is unavoidable as bronchoscopy or other methods of lower airway sampling would be too invasive in young infants. Third, we used an untargeted metabolomics approach which is subject to measurement errors, and may have missed small differences in individual metabolites. While the untargeted metabolomic approach enabled us to comprehensively measure hundreds of metabolites (e.g., sugars, amino acids, lipids, organic acids), for future diagnostic goals, the use of targeted approaches is needed. Fourth, the sample size of this study is relatively smaller to the number of variables examined. This study should facilitate further validation research. Fifth, the nasopharyngeal samples were obtained <24 hours of hospitalization at a single time-point. While longitudinal molecular data (e.g., succession of the airway microbiome over childhood) are also important, the study objective was to identify endotypes of RV bronchiolitis *during* infancy. Additionally, even with single-time point data, we successfully identified biologically-distinct endotypes that are longitudinally associated with the development of recurrent wheeze and asthma. Sixth, the current study did not have healthy "controls". Yet, we note that the study objective is not to assess molecular signatures of bronchiolitis development (yes/no) but to identify endotypes within infants with RV bronchiolitis. Lastly, the study sample consisted of a racially/ethnically- and geographically-diverse US infants hospitalized for bronchiolitis. While our sample had a large severity contrast, our inferences may not be generalizable to infants in ambulatory settings with mild-to-moderate bronchiolitis. Nonetheless, our data remain relevant for tens of thousands of infants with RV bronchiolitis hospitalized yearly in the US,[2] a vulnerable population with high chronic morbidity.

Conclusion

Journal Pre-proof

By integrating the clinical, RV species, nasopharyngeal microbiome, cytokine, and metabolome data from a multicenter prospective cohort study of 122 infants hospitalized for RV bronchiolitis, we identified four biologically-distinct and clinically-meaningful endotypes. Specifically, the bronchiolitis endotype characterized by previous antibiotic use, eosinophilia, IgE sensitization, RV-C infection, *Moraxella*-dominant nasopharyngeal microbiota, high T2 cytokine response, and specific lipid profiles (e.g., sphingolipids, arachidonate, HODE) had the highest risk for developing recurrent wheeze and asthma. Considering the relatively small sample size of infants hospitalized for bronchiolitis in the current study, we need an external validation. However, our data lend support to the emerging concept that "bronchiolitis" represents several diseases that share biological mechanisms (endotypes). For clinicians, our findings could provide an evidence base for the early identification of high-risk children during an important period of airway development—infancy. Our data should also facilitate further investigation (e.g., modification of microbiome and host response) into the development of endotype-specific asthma prevention strategies.

Journal Pre-proof

Journal Pre-proof

Acknowledgments: We thank the MARC-35 study hospitals and research personnel for their ongoing dedication to bronchiolitis and asthma research (**Table E1** in the Online Supplement), and Janice A. Espinola, MPH and Ashley F. Sullivan, MS, MPH (Massachusetts General Hospital, Boston, MA) for their many contributions to the MARC-35 study. We also thank Alkis Togias, MD, at the National Institutes of Health (Bethesda, MD) for helpful comments about the study results.

Journal Pre-proof

REFERENCES

1. Abreo A, Gebretsadik T, Stone CA, Hartert TV. The impact of modifiable risk factor reduction on childhood asthma development. Clin Transl Med. 2018;7:15.
2. Fujiogi M, Goto T, Yasunaga H, Fujishiro J, Mansbach JM, Camargo CA, et al. Trends in bronchiolitis hospitalizations in the United States: 2000-2016. Pediatrics. 2019;144.
3. Bacharier LB, Cohen R, Schweiger T, Yin-Declue H, Christie C, Zheng J, et al. Determinants of asthma after severe respiratory syncytial virus bronchiolitis. J Allergy Clin Immunol. 2012;130:91-100.e3.
4. Mikalsen IB, Halvorsen T, Øymar K. The outcome after severe bronchiolitis is related to gender and virus. Pediatr Allergy Immunol. 2012;23:391–8.
5. Sigurs N, Aljassim F, Kjellman B, Robinson PD, Sigurbergsson F, Bjarnason R, et al. Asthma and allergy patterns over 18 years after severe RSV bronchiolitis in the first year of life. Thorax. 2010;65:1045–52.
6. Castro M, Schweiger T, Yin-DeClue H, Ramkumar TP, Christie C, Zheng J, et al. Cytokine response after severe respiratory syncytial virus bronchiolitis in early life. J Allergy Clin Immunol. 2008;122:726-733.e3.
7. Kotaniemi-Syrjänen A, Vainionpää R, Reijonen TM, Waris M, Korhonen K, Korppi M. Rhinovirus-induced wheezing in infancy--the first sign of childhood asthma? J Allergy Clin Immunol. 2003;111:66–71.
8. Reijonen TM, Kotaniemi-Syrjänen A, Korhonen K, Korppi M. Predictors of asthma three years after hospital admission for wheezing in infancy. Pediatrics. 2000;106:1406–12.

Journal Pre-proof

9. Hasegawa K, Mansbach JM, Bochkov YA, Gern JE, Piedra PA, Bauer CS, et al. Association of rhinovirus C bronchiolitis and immunoglobulin E sensitization during infancy with development of recurrent wheeze. JAMA Pediatr. 2019;173:544–52.

10. Altman MC, Beigelman A, Ciaccio C, Gern JE, Heymann PW, Jackson DJ, et al. Evolving concepts in how viruses impact asthma. J Allergy Clin Immunol. 2020;

11. Feldman AS, He Y, Moore ML, Hershenson MB, Hartert TV. Toward primary prevention of asthma. Reviewing the evidence for early-life respiratory viral infections as modifiable risk factors to prevent childhood asthma. Am J Respir Crit Care Med. 2015;191:34–44.

12. Dumas O, Mansbach JM, Jartti T, Hasegawa K, Sullivan AF, Piedra PA, et al. A clustering approach to identify severe bronchiolitis profiles in children. Thorax. 2016;71:712–8.

13. Dumas O, Hasegawa K, Mansbach JM, Sullivan AF, Piedra PA, Camargo CA. Severe bronchiolitis profiles and risk of recurrent wheeze by age 3 years. J Allergy Clin Immunol. 2019;143:1371-1379.e7.

14. Toivonen L, Camargo CA, Gern JE, Bochkov YA, Mansbach JM, Piedra PA, et al. Association between rhinovirus species and nasopharyngeal microbiota in infants with severe bronchiolitis. J Allergy Clin Immunol. 2019;143:1925-1928.e7.

15. Hasegawa K, Mansbach JM, Ajami NJ, Espinola JA, Henke DM, Petrosino JF, et al. Association of nasopharyngeal microbiota profiles with bronchiolitis severity in infants hospitalised for bronchiolitis. Eur Respir J. 2016;48:1329–39.

16. Rosas-Salazar C, Shilts MH, Tovchigrechko A, Schobel S, Chappell JD, Larkin EK, et al. Differences in the nasopharyngeal microbiome during acute respiratory tract infection with human rhinovirus and respiratory syncytial virus in infancy. J Infect Dis. 2016;214:1924–8.

Journal Pre-proof

17. Hasegawa K, Mansbach JM, Ajami NJ, Petrosino JF, Freishtat RJ, Teach SJ, et al. The relationship between nasopharyngeal CCL5 and microbiota on disease severity among infants with bronchiolitis. Allergy. 2017;72:1796–800.

18. Hasegawa K, Hoptay CE, Harmon B, Celedón JC, Mansbach JM, Piedra PA, 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.

19. Stewart CJ, Mansbach JM, Wong MC, Ajami NJ, Petrosino JF, Camargo CA, et al. Associations of nasopharyngeal metabolome and microbiome with severity among infants with bronchiolitis. A Multiomic Analysis. Am J Respir Crit Care Med. 2017;196:882–91.

20. Ralston SL, Lieberthal AS, Meissner HC, Alverson BK, Baley JE, Gadomski AM, et al. Clinical practice guideline: the diagnosis, management, and prevention of bronchiolitis. Pediatrics. 2014;134:e1474-1502.

21. Hasegawa K, Jartti T, Mansbach JM, Laham FR, Jewell AM, Espinola JA, 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.

22. Bochkov YA, Grindle K, Vang F, Evans MD, Gern JE. Improved molecular typing assay for rhinovirus species A, B, and C. Journal of Clinical Microbiology. 2014;52:2461–71.

23. Human Microbiome Project Consortium. A framework for human microbiome research. Nature. 2012;486:215–21.

24. PowerSoil® DNA Isolation Kit [Internet]. Available from: https://mobio.com/media/wysiwyg/pdfs/protocols/12888.pdf

25. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26:2460–1.

26. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods. 2013;10:996–8.

27. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:D590-596.

28. US Department of Health and Human Services. National asthma education and prevention program: Expert panel report 3 (EPR 3): Guidelines for the diagnosis and management of asthma (NIH publication 08– 4051). Bethesda, MD: National Institutes of Health. 2007;11.

29. Camargo CA, Ingham T, Wickens K, Thadhani R, Silvers KM, Epton MJ, et al. Cord-blood 25-hydroxyvitamin D levels and risk of respiratory infection, wheezing, and asthma. Pediatrics. 2011;127.

30. van den Berg RA, Hoefsloot HCJ, Westerhuis JA, Smilde AK, van der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics. 2006;7:142.

31. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28:882–3.

32. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M, et al. Similarity network fusion for aggregating data types on a genomic scale. Nat Methods. 2014;11:333–7.

33. MetaboAnalystR package | R Documentation [Internet]. [cited 2020 Mar 18]. Available from: https://www.rdocumentation.org/packages/MetaboAnalystR/versions/2.0.0

34. Teo SM, Tang HHF, Mok D, Judd LM, Watts SC, Pham K, et al. Airway microbiota dynamics uncover a critical window for interplay of pathogenic bacteria and allergy in childhood respiratory disease. Cell Host Microbe. 2018;24:341-352.e5.

35. Toivonen L, Hasegawa K, Waris M, Ajami NJ, Petrosino JF, Camargo CA, et al. Early nasal microbiota and acute respiratory infections during the first years of life. Thorax. 2019;74:592–9.

36. Bisgaard H, Hermansen MN, Buchvald F, Loland L, Halkjaer LB, Bønnelykke K, et al. Childhood asthma after bacterial colonization of the airway in neonates. N Engl J Med. 2007;357:1487–95.

37. McCauley K, Durack J, Valladares R, Fadrosh DW, Lin DL, Calatroni A, et al. Distinct nasal airway bacterial microbiotas differentially relate to exacerbation in pediatric patients with asthma. J Allergy Clin Immunol. 2019;144:1187–97.

38. Zhou Y, Jackson D, Bacharier LB, Mauger D, Boushey H, Castro M, et al. The upper-airway microbiota and loss of asthma control among asthmatic children. Nat Commun. 2019;10:5714.

39. Toivonen L, Schuez-Havupalo L, Karppinen S, Waris M, Hoffman KL, Camargo CA, et al. Antibiotic treatments during infancy, changes in nasal microbiota, and asthma development: Population-based cohort study. Clin Infect Dis. 2020;

40. Turi KN, Shankar J, Anderson LJ, Rajan D, Gaston K, Gebretsadik T, et al. Infant viral respiratory infection nasal immune-response patterns and their association with subsequent childhood recurrent wheeze. Am J Respir Crit Care Med. 2018;198:1064–73.

Journal Pre-proof

41. Kelly RS, Chawes BL, Blighe K, Virkud YV, Croteau-Chonka DC, McGeachie MJ, et al. An integrative transcriptomic and metabolomic study of lung function in children with asthma. Chest. 2018;154:335–48.

42. Kelly RS, Chawes BL, Guo F, Zhang L, Blighe K, Litonjua AA, et al. The role of the 17q21 genotype in the prevention of early childhood asthma and recurrent wheeze by vitamin D. Eur Respir J. 2019;54.

43. Bochkov YA, Gern JE. Rhinoviruses and their receptors: Implications for allergic disease. Curr Allergy Asthma Rep. 2016;16:30.

44. Sikazwe CT, Laing IA, Imrie A, Smith DW. Nasal cytokine profiles of patients hospitalised with respiratory wheeze associated with rhinovirus C. Viruses. 2019;11.

45. Bashir H, Grindle K, Vrtis R, Vang F, Kang T, Salazar L, et al. Association of rhinovirus species with common cold and asthma symptoms and bacterial pathogens. J Allergy Clin Immunol. 2018;141:822-824.e9.

46. Sturgill JL. Sphingolipids and their enigmatic role in asthma. Advances in biological regulation. 2018;70:74–81.

47. Zhu Z, Camargo CA, Hasegawa K. Metabolomics in the prevention and management of asthma. Expert Rev Respir Med. 2019;13:1135–8.

48. Zhang Y, Willis-Owen SAG, Spiegel S, Lloyd CM, Moffatt MF, Cookson WOCM. The ORMDL3 asthma gene regulates ICAM1 and has multiple effects on cellular inflammation. Am J Respir Crit Care Med. 2019;199:478–88.

49. Panda L, Gheware A, Rehman R, Yadav MK, Jayaraj BS, Madhunapantula SV, et al. Linoleic acid metabolite leads to steroid resistant asthma features partially through NF-κB. Sci Rep. 2017;7:9565.

50. Marsh RL, Kaestli M, Chang AB, Binks MJ, Pope CE, Hoffman LR, et al. The microbiota in bronchoalveolar lavage from young children with chronic lung disease includes taxa present in both the oropharynx and nasopharynx. Microbiome. 2016;4:37.

51. Poole A, Urbanek C, Eng C, Schageman J, Jacobson S, O'Connor BP, et al. Dissecting childhood asthma with nasal transcriptomics distinguishes subphenotypes of disease. J Allergy Clin Immunol. 2014;133:670-678.e12.

Journal Pre-proof

Journal Pre-proof

FIGURE LEGENDS

Figure 1. Analytic flow of integrated-omics endotyping

A. After an affinity matrix of each dataset was separately computed, and a fused affinity matrix was generated by similarity network fusion. Then, the fused affinity matrix was used to identify mutually-exclusive endotypes by spectral clustering.

B. Eigengaps and clinical plausibility were examined to choose the optimal number of endotypes. The concordance between different numbers of endotypes was also examined. After deriving endotypes, a similarity network was visualized.

C. The between-endotype differences in nasopharyngeal microbiome, cytokine, and metabolome were visualized using heatmap.

D. To examine the longitudinal relation of derived endotypes with recurrent wheeze, the time to recurrent wheeze was modeled by fitting a Cox proportional hazards model. The risk of childhood asthma was modeled by fitting a logistic regression model.

Abbreviations: RV-A, rhinovirus A; RV-C, rhinovirus C; RSV, respiratory syncytial virus; T2, type 2 cytokines

Figure 2. Similarity network visualization of rhinovirus bronchiolitis endotypes

Nodes (circles) with a same color represent infants with the corresponding endotype (A, B, C, or D). Network graphs in the integrated-omics similarity network are visualized using with Fruchterman-Reingold layout. Layout of the network from the four datasets are fixed with the position of integrated-omics similarity network. We selected 1000 edges with the highest similarity for the network visualization and the width of edge reflects the strength of similarity.

609 Abbreviations: RV-A, rhinovirus A; RV-C, rhinovirus C; RSV, respiratory syncytial virus; T2,
610 type 2 cytokines
611

612 **Figure 3. Between-endotype differences in nasopharyngeal microbiome, cytokine, and**
613 **metabolome in infants with rhinovirus bronchiolitis**

614 To visualize the between-endotype differences, the microbiome data (five most abundant genera)
615 are processed by $\log_2$ transformation and Pareto scaling; the cytokine data are processed by $\log_2$
616 transformation and autoscaling; the metabolome data (30 metabolites with highest normalized
617 mutual information score, **Table E2**) are processed by $\log_2$ transformation with batch effect
618 adjustment. The color bar indicates the median value in the corresponding endotype—red color
619 indicates a higher median value while blue color indicates a lower median value. Super-pathway of
620 each metabolite is color-coded for each function.
621 Abbreviations: RV-A, rhinovirus A; RV-C, rhinovirus C; RSV, respiratory syncytial virus; IL,
622 interleukin; TSLP, thymic stromal lymphopoietin; IFN-beta, interferon-beta; MIP-1, macrophage
623 inflammatory protein-1; RANTES, regulated upon activation, normal T cell expressed and
624 presumably secreted; DHA, docosahexaenoic acid; GPS, glycerophosphatidylserine; GPC,
625 glycerophosphorylcholine; GPE, glycerophosphoethanolamine; GPG, glycerophosphoglycerol
626

627 **Figure 4. Kaplan-Meier curves for development of recurrent wheeze by age 3 years,**
628 **according to rhinovirus bronchiolitis endotypes**

629 The survival curves significantly differed across the endotypes ($P_{Wilcoxon}$=0.04). Compared with
630 endotype A (virus$^{RV\text{-}C}$microbiome^{mixed}T2^{low}) infants, the rate of developing recurrent wheeze by

631 age 3 years was not significantly different in endotype C or D. By contrast, the rate was
632 significantly higher in endotype D (virus$^{\text{RV-C}}$microbiome$^{\textit{Moraxella}}$T2$^{\text{high}}$) infants. Corresponding
633 hazards ratio estimates are presented in **Table 2**. We observed 15 censoring and assumed
634 non-informative censoring as the exposure (endotypes) contains the information of many
635 important covariates.
636 Abbreviations: RV-A, rhinovirus A; RV-C, rhinovirus C; RSV, respiratory syncytial virus
637

Journal Pre-proof

Journal Pre-proof

Table 1. Baseline characteristics and clinical course of infants, according to rhinovirus bronchiolitis endotypes

Characteristics	Overall (n=122; 100%)	Endotype A (n=27; 22%)	Endotype B (n=23; 19%)	Endotype C (n=44; 36%)	Endotype D (n=28; 23%)	P-value
Demographics						
Age (month), median (IQR)	4 (2–6)	3 (2–6)	4 (2–5)	4 (3–6)	5 (1–8)	0.44
Female sex	42 (34.4)	8 (29.6)	6 (26.1)	18 (40.9)	10 (35.7)	0.61
Race/ethnicity						0.89
Non-Hispanic white	50 (41.0)	11 (40.7)	8 (34.8)	21 (47.7)	10 (35.7)	
Non-Hispanic black	29 (23.8)	7 (25.9)	7 (30.4)	8 (18.2)	7 (25.0)	
Hispanic	39 (32.0)	7 (25.9)	8 (34.8)	14 (31.8)	10 (35.7)	
Other or unknown	4 (3.3)	2 (7.4)	0 (0.0)	1 (2.3)	1 (3.6)	
Prematurity (32-37 weeks)	27 (22.1)	5 (18.5)	4 (17.4)	11 (25.0)	7 (25.0)	0.87
Birth weight (kg)						0.33
0–1.3	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.0)	
1.4–2.2	11 (9.0)	0 (0)	3 (13.0)	6 (13.6)	2 (7.1)	
2.3–3.1	41 (33.6)	10 (37.0)	8 (34.8)	11 (25.0)	12 (42.9)	
≥3.2	70 (57.4)	17 (63.0)	12 (52.2)	27 (61.4)	14 (50.0)	
Mode of birth (caesarean delivery)	37 (30.3)	8 (29.6)	7 (30.4)	13 (29.5)	9 (32.1)	0.99
Previous breathing problems (count)						0.33
0	88 (72.1)	22 (81.5)	18 (78.3)	30 (68.2)	18 (64.3)	
1	28 (23.0)	5 (18.5)	5 (21.7)	9 (20.5)	9 (32.1)	
2	6 (4.9)	0 (0)	0 (0)	5 (11.4)	1 (3.6)	
Previous ICU admission	1 (0.8)	0 (0)	1 (4.3)	0 (0)	0 (0)	0.19
Lifetime antibiotic use	50 (41.0)	5 (18.5)	5 (21.7)	22 (50.0)	18 (64.3)	0.001
Ever attended daycare	34 (27.9)	9 (33.3)	4 (17.4)	15 (34.1)	6 (21.4)	0.38
Cigarette smoke exposure at home	13 (10.7)	2 (7.4)	4 (17.4)	5 (11.4)	2 (7.1)	0.66
Maternal smoking during pregnancy	12 (9.8)	2 (7.4)	5 (21.7)	4 (9.1)	1 (3.6)	0.22
Parental history of asthma	48 (39.3)	4 (14.8)	11 (47.8)	22 (50.0)	11 (39.3)	0.02
Parental history of eczema	27 (22.1)	3 (11.1)	7 (30.4)	8 (18.2)	9 (32.1)	0.18
Clinical presentation						
Weight (kg), median (IQR)	6.70 (5.18–8.00)	5.70 (4.85–7.60)	5.99 (5.05–	6.70 (5.56–	7.60 (4.72–	0.32
Respiratory rate (per minute), median	48 (40–60)	43 (32–54)	48 (40–60)	48 (40–60)	50 (42–56)	0.26

Journal Pre-proof

Oxygen saturation						0.71
<90%	11 (9.4)	1 (3.7)	2 (10.0)	5 (11.9)	3 (10.7)	
90-93%	14 (12.0)	4 (14.8)	4 (20.0)	3 (7.1)	3 (10.7)	
≥94%	92 (78.6)	22 (81.5)	14 (70.0)	34 (81.0)	22 (78.6)	
Blood eosinophilia (≥4%)	13 (10.7)	3 (11.1)	4 (17.4)	1 (2.3)	5 (17.9)	0.07
sIgE sensitization	15 (12.3)	3 (11.1)	3 (13.0)	4 (9.1)	5 (17.9)	0.74
Clinical course						
Positive pressure ventilation use [*]	4 (3.3)	0 (0)	0 (0)	3 (6.8)	1 (3.6)	0.55
Intensive treatment use[†]	13 (10.7)	1 (3.7)	3 (13.0)	4 (9.1)	5 (17.9)	0.37
Length of day (day), median (IQR)	2 (1-3)	1 (1–2)	2 (1–3.5)	2 (1–3)	2 (1–4)	0.08
Antibiotic use during hospitalization	39 (32.0)	6 (22.2)	7 (30.4)	18 (40.9)	8 (28.6)	0.40
Corticosteroid use during hospitalization	21 (17.2)	2 (7.4)	5 (21.7)	8 (18.2)	6 (21.4)	0.47
Respiratory virus						
RV-A	60 (49.2)	8 (29.6)	17 (73.9)	26 (59.1)	9 (32.1)	0.002
RV-B	6 (4.9)	2 (7.4)	1 (4.3)	1 (2.3)	2 (7.1)	0.76
RV-C	57 (46.7)	17 (63.0)	6 (26.1)	15 (34.1)	19 (67.9)	0.002
RSV coinfection	60 (49.2)	14 (51.9)	6 (26.1)	30 (68.2)	10 (35.7)	0.004

Abbreviations: IQR, interquartile range; ICU, intensive care unit; sIgE, specific immunoglobulin E; RV-A, rhinovirus A; RV-B, rhinovirus B; RV-C, rhinovirus C; RSV, respiratory syncytial virus

Data are no. (%) of infants unless otherwise indicated. Percentages may not equal 100, because of rounding and missingness.

[*] Infants with bronchiolitis who underwent continuous positive airway ventilation and/or mechanical ventilation.

[†] Infants with bronchiolitis who were admitted to ICU and/or who underwent positive pressure ventilation.

Journal Pre-proof

Table 2. Association of rhinovirus bronchiolitis endotypes in infants with development of recurrent wheeze and asthma outcomes

Endotypes	Recurrent wheeze by age 3 years*			Childhood asthma at age 5 years†		
	Number of outcomes	Hazard ratio (95% CI)	P-value	Number of outcomes	Odds ratio (95% CI)	P-value
Endotype A virus$^{RV\text{-}C}$microbiome^{mixed}T2^{low} (n=27)	9	1 [Reference]	-	7	1 [Reference]	-
Endotype B virus$^{RV\text{-}A}$microbiome$^{\textit{Haemophilus}}$T2^{low} (n=23)	9	0.98 (0.39 − 2.48)	0.97	8	1.47 (0.43 − 5.16)	0.54
Endotype C virus$^{RSV/RV}$microbiome$^{\textit{Streptococcus}}$T2^{low} (n=44)	18	1.20 (0.54 − 2.67)	0.66	15	1.48 (0.51 − 4.56)	0.47
Endotype D virus$^{RV\text{-}C}$microbiome$^{\textit{Moraxella}}$T2^{high} (n=28)	18	2.23 (1.00 − 4.96)	0.049	16	3.74 (1.21 − 12.6)	0.03

Abbreviation: CI, confidential interval

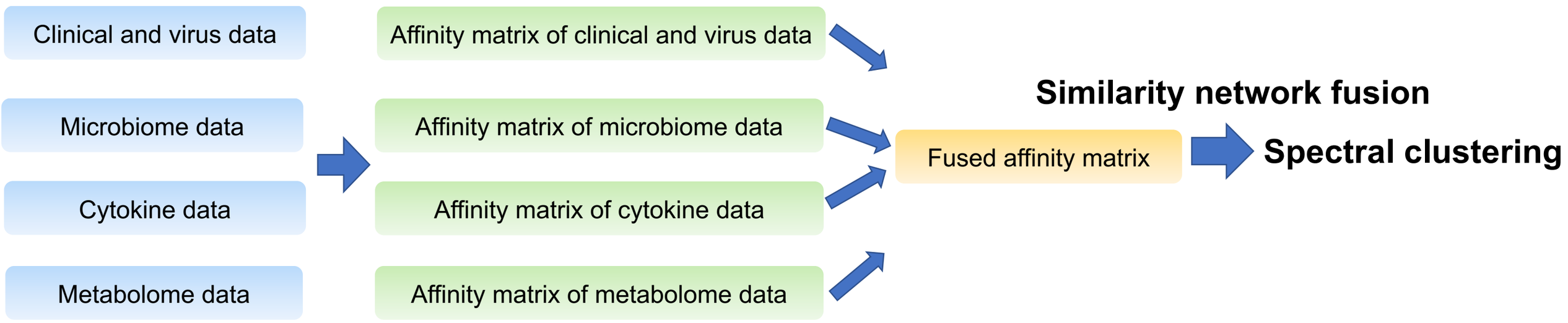
* Recurrent wheeze was defined as having at least two corticosteroid-requiring exacerbations in six months or at least four wheezing episodes in one year that last at least one day and affect sleep. To examine the association between RV bronchiolitis endotypes (endotype A as the reference) and the rate of recurrent wheeze, Cox proportional hazards model was fit. We observed 15 censoring and assumed non-informative censoring as the exposure (endotypes) contains the information of many important covariates.

† Asthma was defined as physician-diagnosis of asthma by age 5 years, plus either asthma medication use (e.g., albuterol inhaler, inhaled corticosteroids, montelukast) or asthma-related symptoms in the preceding year. To examine the association between RV bronchiolitis endotypes (endotype A as the reference) and the risk of developing childhood asthma, logistic regression model was fit.

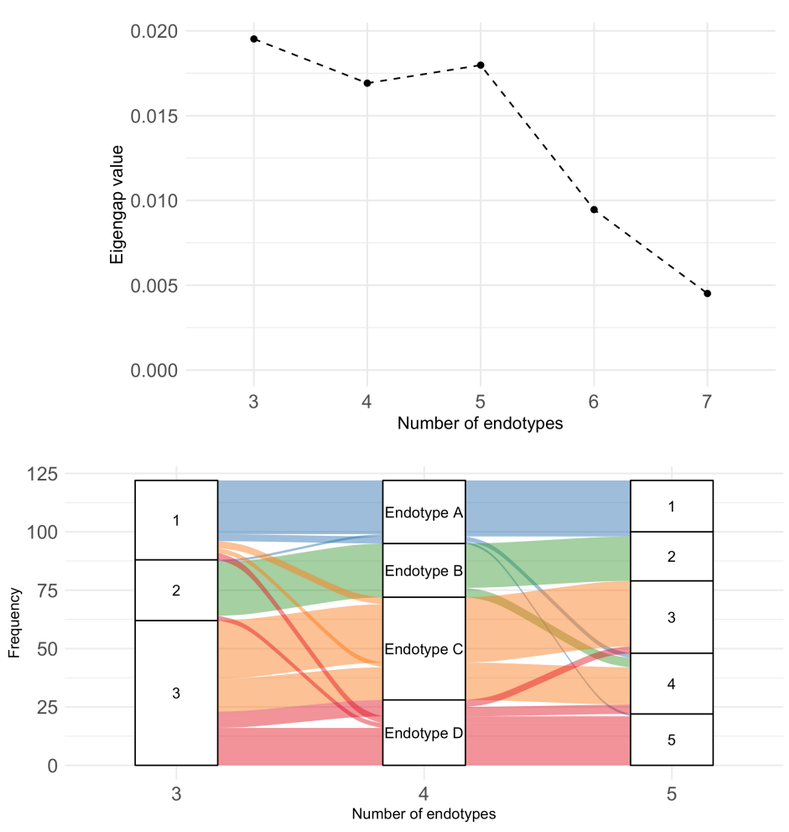
Journal Pre-proof

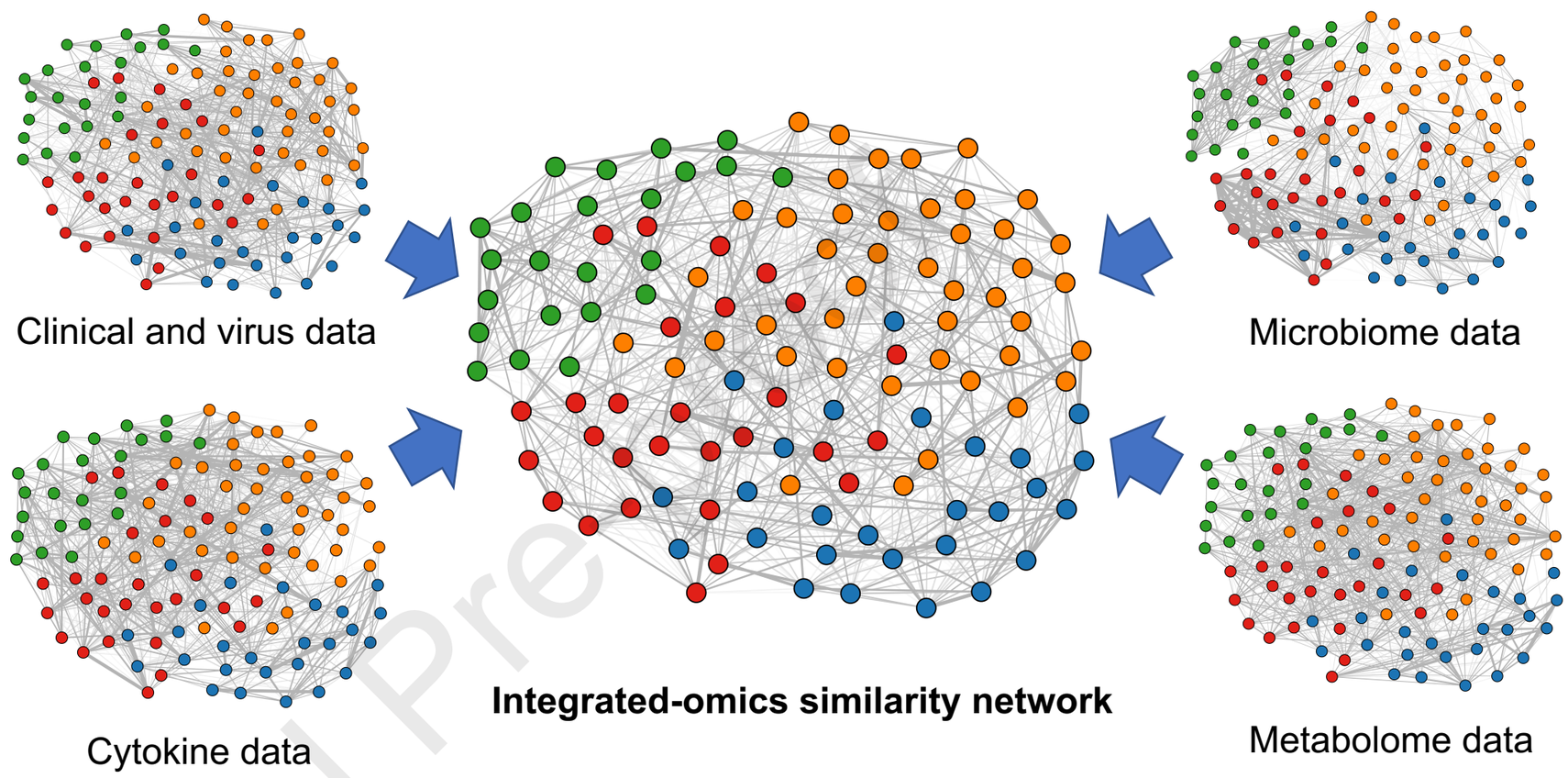
Journal Pre-proof

A. Integrated-omics endotyping

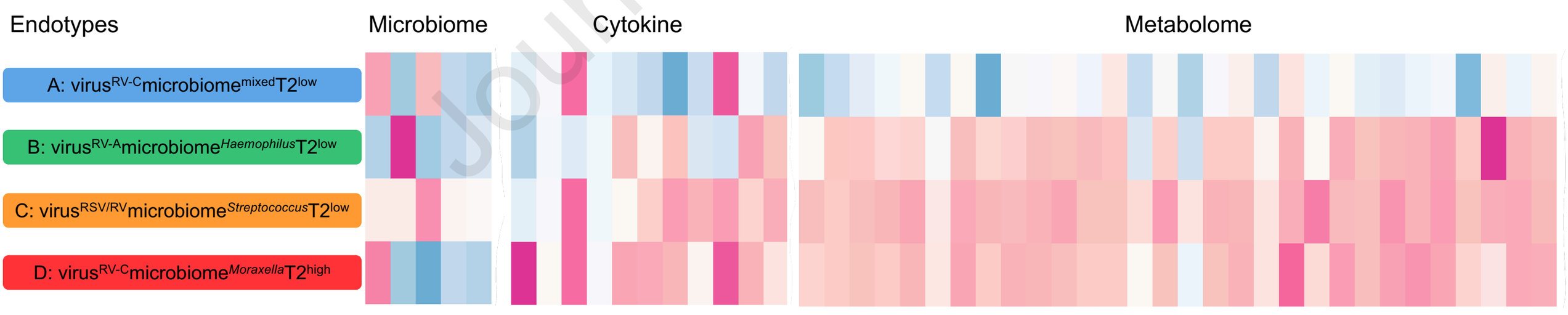


B. Endotypes and network visualization





C. Endotypes and molecular data

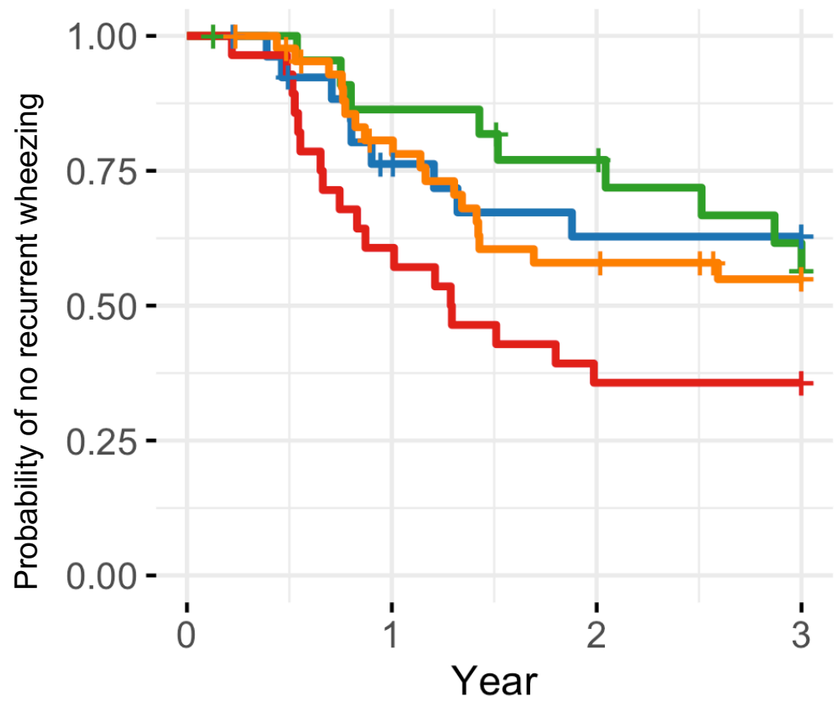


D. Association of endotypes with clinical outcomes

Recurrent wheeze by age 3 years

Kaplan-Meier curves for survival function estimation

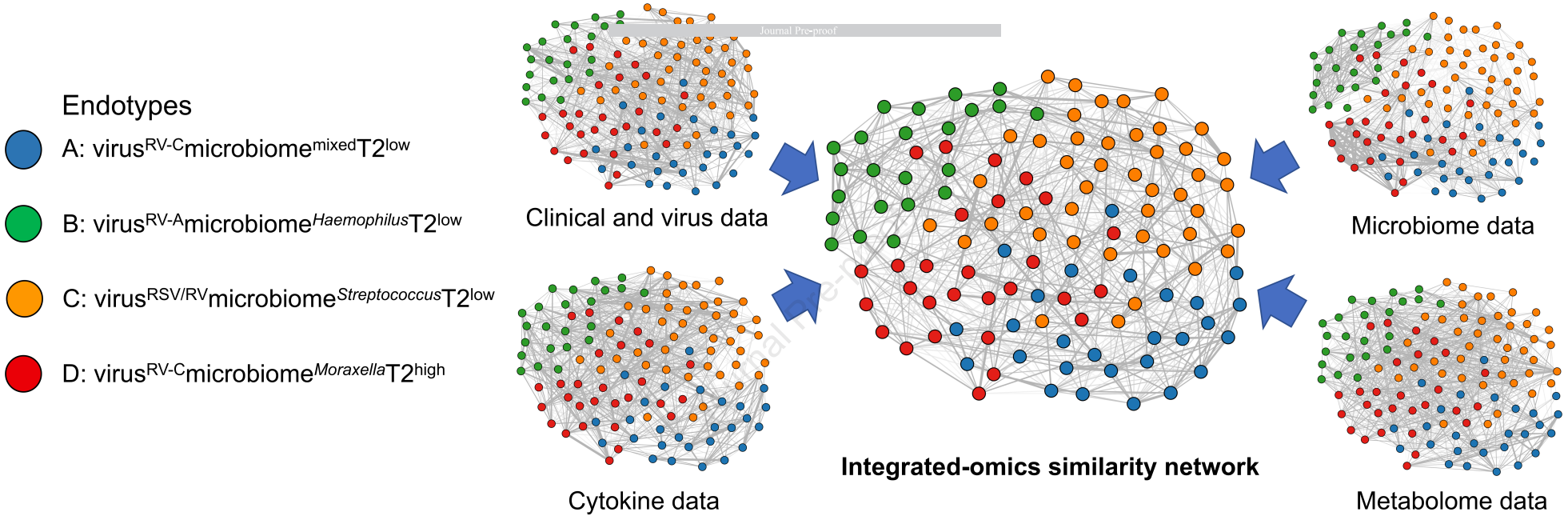
Hazard ratio estimation using Cox proportional hazards model

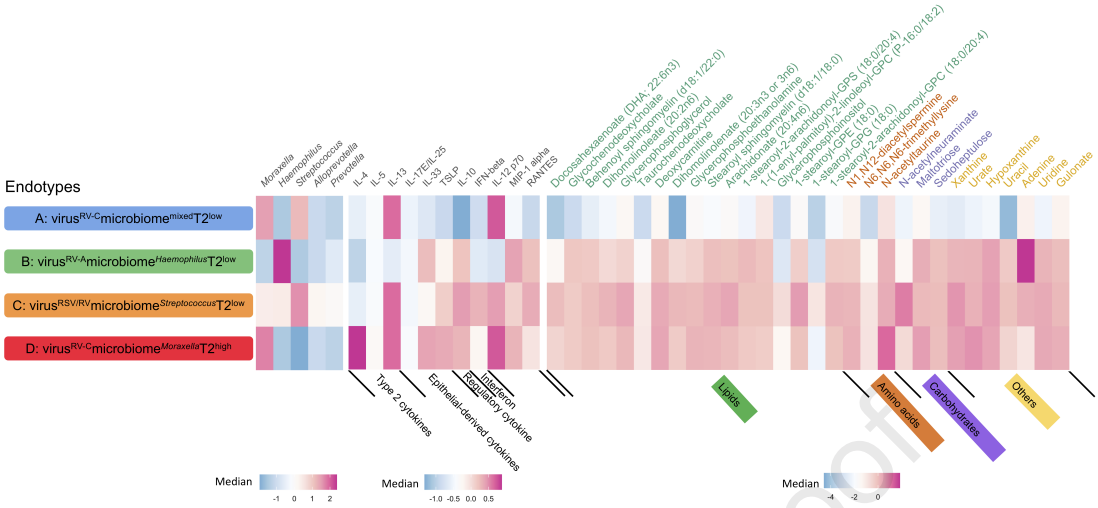


Asthma at age 5 years

Odds ratio estimation using logistic regression model

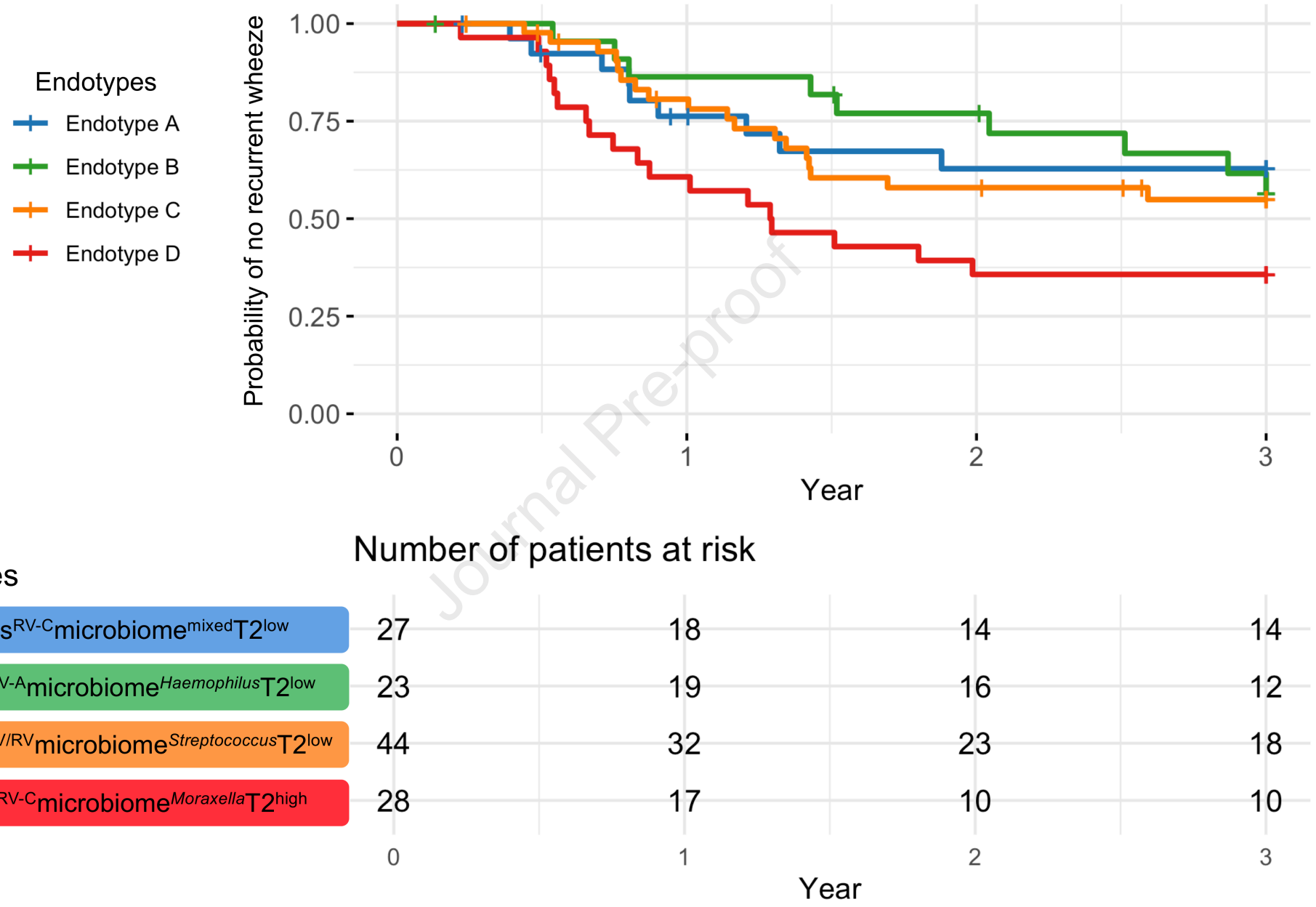
Journal Pre-proof





Journal Pre-proof

Journal Pre-proof



Journal Pre-proof