


RESEARCH LETTER



## The influence of early life exposures on the infant gut virome

Yichang Zhang <sup>a\*</sup>, Josué L. Castro-Mejía <sup>a\*</sup>, Ling Deng <sup>a\*</sup>, Shiraz A. Shah <sup>b</sup>, Jonathan Thorsen <sup>b,c</sup>, Cristina Leal Rodríguez <sup>b</sup>, Leon E. Jessen <sup>b,d</sup>, Moira B. Dion <sup>e,f</sup>, Bo Chawes <sup>b</sup>, Klaus Bønnelykke <sup>b</sup>, Søren J. Sørensen <sup>g</sup>, Hans Bisgaard <sup>b</sup>, Sylvain Moineau <sup>e,f,h</sup>, Marie-Agnès Petit <sup>i</sup>, Jakob Stokholm <sup>a,b</sup>, and Dennis S. Nielsen <sup>a</sup>

<sup>a</sup>Department of Food Science, University of Copenhagen, Frederiksberg, Denmark; <sup>b</sup>Copenhagen Prospective Studies on Asthma in Childhood, Copenhagen University Hospital, Herlev-Gentofte, Gentofte, Denmark; <sup>c</sup>Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; <sup>d</sup>Department of Health Technology, Section for Bioinformatics, Technical University of Denmark, DTU, Lyngby, Denmark; <sup>e</sup>Département de Biochimie, de Microbiologie, et de Bio-Informatique, Faculté des Sciences et de Génie, Université Laval, Québec, QC, Canada; <sup>f</sup>Groupe de Recherche en Écologie Buccale, Faculté de Médecine Dentaire, Université Laval, Québec, QC, Canada; <sup>g</sup>Department of Biology, University of Copenhagen, Frederiksberg, Denmark; <sup>h</sup>Félix d'Hérelle Reference Center for Bacterial Viruses, Faculté de médecine dentaire, Université Laval, Québec, QC, Canada; <sup>i</sup>Université Paris-Saclay, INRAE, AgroParis Tech, Micalis Institute, Jouy-en-Josas, France

### ABSTRACT

The factors influencing the establishment of the gut bacterial community in early life are fairly well studied. However, the factors shaping the infant gut virome remain elusive. Interestingly, early life gut virome imbalances have recently been linked with increased risk of developing diseases like type 1 diabetes and asthma. We utilized the deeply phenotyped COPSAC2010 cohort to investigate how environmental factors influence the gut virome at one year age. We demonstrate that the presence of older siblings as well as residential location (urban or rural) had the strongest impact on gut virome composition at 1 year of age. A total of 16,118 species-level clustered viral representative contigs (here termed viral Operational Taxonomic Units – vOTUs) were identified and of these 2105 vOTUs varied in abundance with environmental exposures. Of these vOTUs 94.1% were phages mainly predicted to infect *Bacteroidaceae*, *Prevotellaceae*, and *Ruminococcaceae*. Strong co-abundance of phages and their bacterial hosts was confirmed underlining the predicted phage-host connections. Furthermore, we found some gut viruses affected by environmental factors encode enzymes involved in the utilization and degradation of major dietary components, potentially affecting infant health by influencing the bacterial host metabolic capacity. These findings provide a valuable insights for understanding the early life factors that predispose to autoimmune and metabolic disorders.

### ARTICLE HISTORY

Received 23 December 2024  
Revised 21 April 2025  
Accepted 29 April 2025

### KEYWORDS

Environmental exposures;  
infant; gut microbiota; virus;  
phage; bacterial host;  
metabolism

## Introduction


Early life gut microbiome (GM) establishment plays a fundamental role in shaping host physiology and health<sup>1,2</sup> with early life GM imbalances being linked to onset and progression of chronic diseases later in life, such as obesity,<sup>3</sup> diabetes,<sup>4,5</sup> and asthma.<sup>2</sup>

To date, GM research has generally focused on understanding the importance of the bacterial GM component, but recent findings indicate that the vast and diverse population of viruses found in the gut (collectively called the “gut virome”) also play a prominent role in gut microbial ecology<sup>6–9</sup>. Amidst

these biological entities, bacterial viruses, also termed bacteriophages (phages), are the most diverse and abundant particles of the GM<sup>9–11</sup> representing a major reservoir of genetic diversity influencing not only GM composition but also the GM metabolic potential.<sup>12,13</sup> Disease-specific alterations in the gut virome have been reported in several chronic conditions<sup>14</sup> such as inflammatory bowel disease,<sup>15</sup> colorectal cancer,<sup>16</sup> necrotizing enterocolitis in preterm infants,<sup>17</sup> severe acute malnutrition,<sup>18</sup> type-1 diabetes<sup>19,20</sup> and other autoimmune diseases such as rheumatoid arthritis.<sup>21</sup> Very recently, we have demonstrated that infant

**CONTACT** Ling Deng  [lingdeng@food.ku.dk](mailto:lingdeng@food.ku.dk); Dennis S. Nielsen  [dn@food.ku.dk](mailto:dn@food.ku.dk)  Department of Food Science, University of Copenhagen, Rolighedsvej 26, Frederiksberg C 1958, Denmark

\*These authors contributed equally.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/19490976.2025.2501194>

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

gut virome imbalance in the temperate phage pool at the age of 1 year is associated with increased risk of developing asthma before school age. Interestingly, it was found that this effect was additive to the increased risk of developing asthma due to imbalanced gut bacteriome at the age of 1 year, meaning that the influence of gut virome composition on asthma is not only mediated via influencing the bacterial gut component but also independent of the bacteriome.<sup>22</sup> The role of the gut virome in shaping the GM is further underlined by the observation that fecal virome transfer from healthy donors to recipients with a dysbiotic GM prevent or ameliorate symptoms associated with metabolic<sup>23</sup> and gastrointestinal<sup>24,25</sup> disorders.

While various early-life factors such as birth mode, siblings, diet and exposure to antibiotics have been found to influence the development of the gut bacterial community,<sup>1,26</sup> little is known about which factors shape the gut virome. The few attempts that have characterized the gut virome early in life have revealed that its composition is highly dynamic,<sup>27–29</sup> affected by delivery mode,<sup>6,30</sup> maternal factors such as gestational diabetes<sup>30</sup> and the first bacterial colonizers<sup>31</sup> as well as being enriched in phages belonging to the *Microviridae* family.<sup>10,28</sup> Moreover, its transmission-dynamics after birth follows a stepwise assembly, with breastfeeding playing a protective role against eukaryotic viral infections.<sup>32,33</sup> Understanding how environmental exposures and phenotypes intertwine the vector space conformed by viruses, bacteria, host, and their functional attributes remains hitherto an unsolved task.

In a recent detailed investigation of the infant gut virome, we showed a massive diversity of hitherto undescribed phages.<sup>9</sup> In this cross-sectional study of the gut virome of 645 infants at 1 year of age enrolled in the COPSAC<sub>2010</sub> cohort<sup>34</sup> more than 10000 viral species distributed over 248 viral families and 17 viral order-level clades were detected.<sup>9</sup> Recently, gut virome imbalance was as mentioned above linked to increased risk of developing asthma before school age in this cohort.<sup>22</sup>

Early life environmental exposures have influenced gut bacteriome development early in life.<sup>2,26,35,36</sup> Here we expand these findings by investigating how social, pre-, peri- and postnatal

factors influence the gut virome composition at 1 year of age. Our findings demonstrate how early life exposures are linked to the abundance of specific viruses, as well as their co-abundance and concordance with their predicted bacterial hosts. Metabolic functions encoded in the genomes of these viruses displayed enrichment of genes important for bacterial physiology in response to exposures, some of which are likely associated with dietary elements (e.g., degradation of complex carbohydrates) and others that may influence infant growth and health.

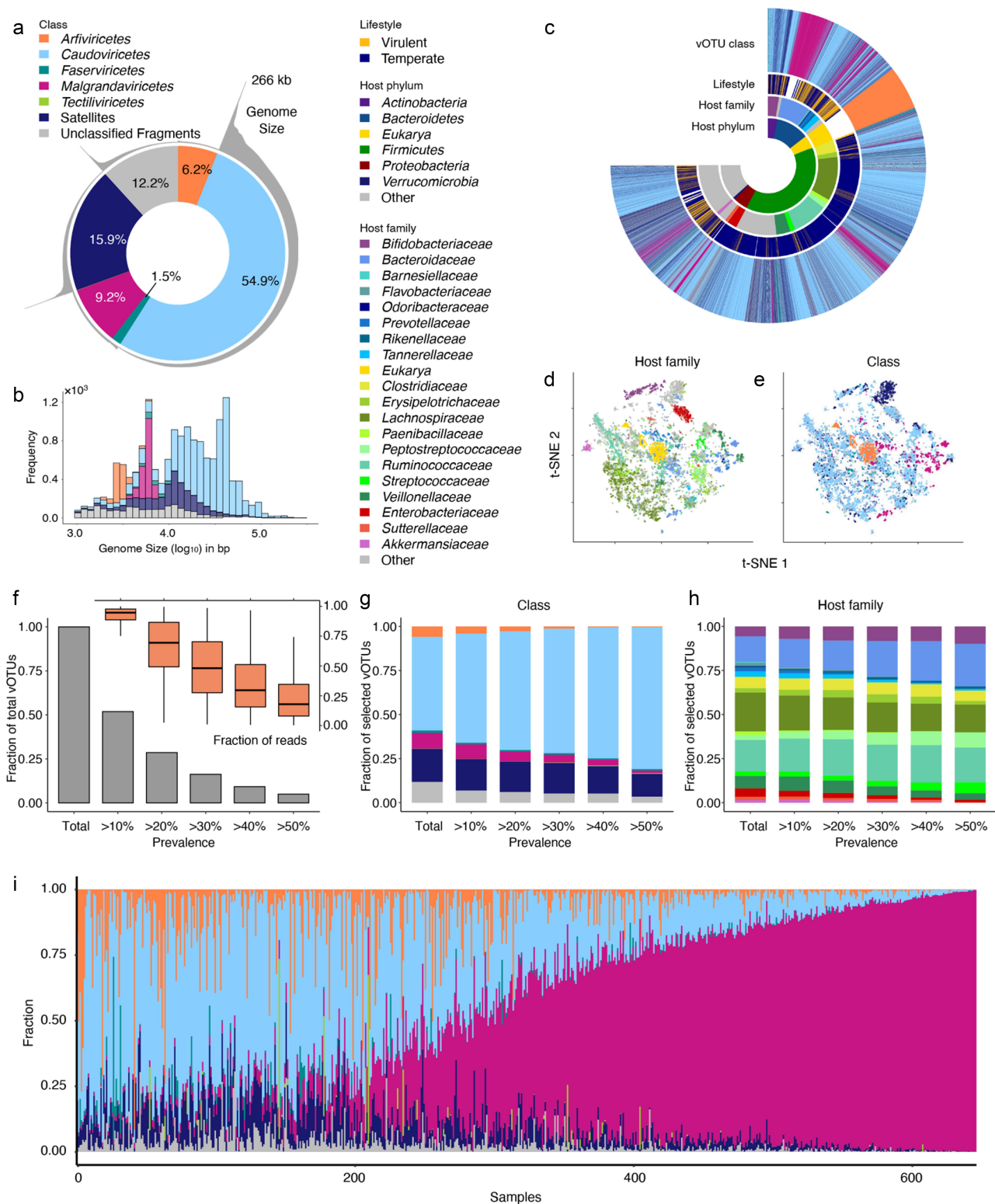
## Results

### Composition of DNA viruses in the gut of Danish infants

A total of 645 stool samples from 1-year-old infants in the COPSAC<sub>2010</sub> cohort<sup>34</sup> were obtained and analyzed.<sup>9</sup> Virions were isolated, concentrated and their genome was sequenced using a shotgun metagenome strategy.<sup>9,37</sup> Following assembly, a total of 16,118 species-level clustered viral representative contigs (here termed viral Operational Taxonomic Units – vOTUs) were obtained. Around 70% of the vOTUs were affiliated to five viral classes (*Arfiviricetes*, *Caudoviricetes*, *Faserviricetes*, *Malgrandaviricetes* and *Tectiliviricetes*) (Figure 1(a, l)). Almost 18.8% of the vOTUs ( $n = 3,029$ ) were considered putative satellite phages as contigs lacked genes coding for structural proteins but encoded other viral proteins (e.g., integrases or replicases) and were conserved in size and gene content across multiple samples. In addition, 11.8% of the vOTUs ( $n = 1,895$ ) were categorized as unclassified viral fragments (Figure 1(a)).

The largest genomes (>10 kb) were observed among *Caudoviricetes*, which constituted the vast majority of vOTUs (Figure 1(b)). The genomes, dominated by *Caudoviricetes* (tailed, double-stranded DNA phages) and *Malgrandaviricetes* (non-tailed, single-stranded DNA phages), followed a bi-/multi-modal distribution (Hartigan's Dip test,  $p < 0.0001$ ) based on their genome sizes (Figure 1(b)).

Bacterial hosts as well as lifestyle (temperate/virulent) of the vOTUs were predicted using CRISPR spacers and the presence of integrases,<sup>9</sup> respectively



**Figure 1.** Virome structure of the infants enrolled in the COPSAC<sub>2010</sub> cohort. (a) Distribution of the 16,118 vOTUs identified colored by their taxonomic class annotation. (b) Cumulative frequency of viral genomes (kb) identified by their taxonomic class annotation. (c) Circular diagram showing the distribution of vOTUs colored by their targeted bacterial hosts (at phylum and family levels), viral class and lifestyle. (d,e) *t*-Stochastic Neighbor Embedding (*t*-SNE) plots clustering *tetra*-mer vOTUs profiles identified according to host family (d) and viral class (e). (f–h) Percentage of vOTUs that appear at a specific prevalence (f), and vOTUs' distribution colored by their taxonomic class (g) and host family (h). (i) Relative abundance of vOTUs across all samples at the class level. Samples were sorted by *Malgrandaviricetes* abundance.

(Figure 1(c)). Because phages tend to have comparable  $k$ -mer frequencies to those of their hosts,<sup>38,39</sup> we also performed dimensionality reduction on tetramer vectors to confirm global host associations as a complement to our viral taxonomy.<sup>9</sup> Using unsupervised stochastic neighbor embedding (t-SNE) dimensionality reduction, vOTUs targeting the same hosts as determined by CRISPR spacers (Figure 1(d)) or belonging to the same viral classes (Figure 1(e)) were found to clearly cluster together. Previously only *Enterobacteriaceae* and *Bacteroidetes* have been shown to be the hosts of non-tailed *Malgrandaviricetes*,<sup>40</sup> but when examining the bacterial hosts, we observed that in addition to *Bacteroidetes*, also *Ruminococcaceae*, *Clostridiaceae*, *Erysipelotrichaceae* and *Sutterellaceae* are predicted as hosts of *Malgrandaviridetes* viruses (Figure 1(c) and S1). With respect to lifestyle, vOTUs predicted to infect *Streptococcaceae* and most families of *Bacteroidetes* exhibit a higher proportion of virulent phages compared vOTUs predicted to infect other families (Figure 1(c) and Table S2).

The distribution of vOTUs was very individual-specific, with less than 5% of vOTUs appearing in more than 50% of the samples (Figure 1(f)). However, this still adds up to around 800 vOTUs that are shared among a majority of infants and representing, on average, more than 20% of the sequencing reads (Figure 1(f)). The proportion of vOTUs classified as *Caudoviricetes* (Figure 1(g)) as well as those infecting *Bacteroidaceae* and *Bifidobacteriaceae* (Figure 1(h)) increased as a function of prevalence.

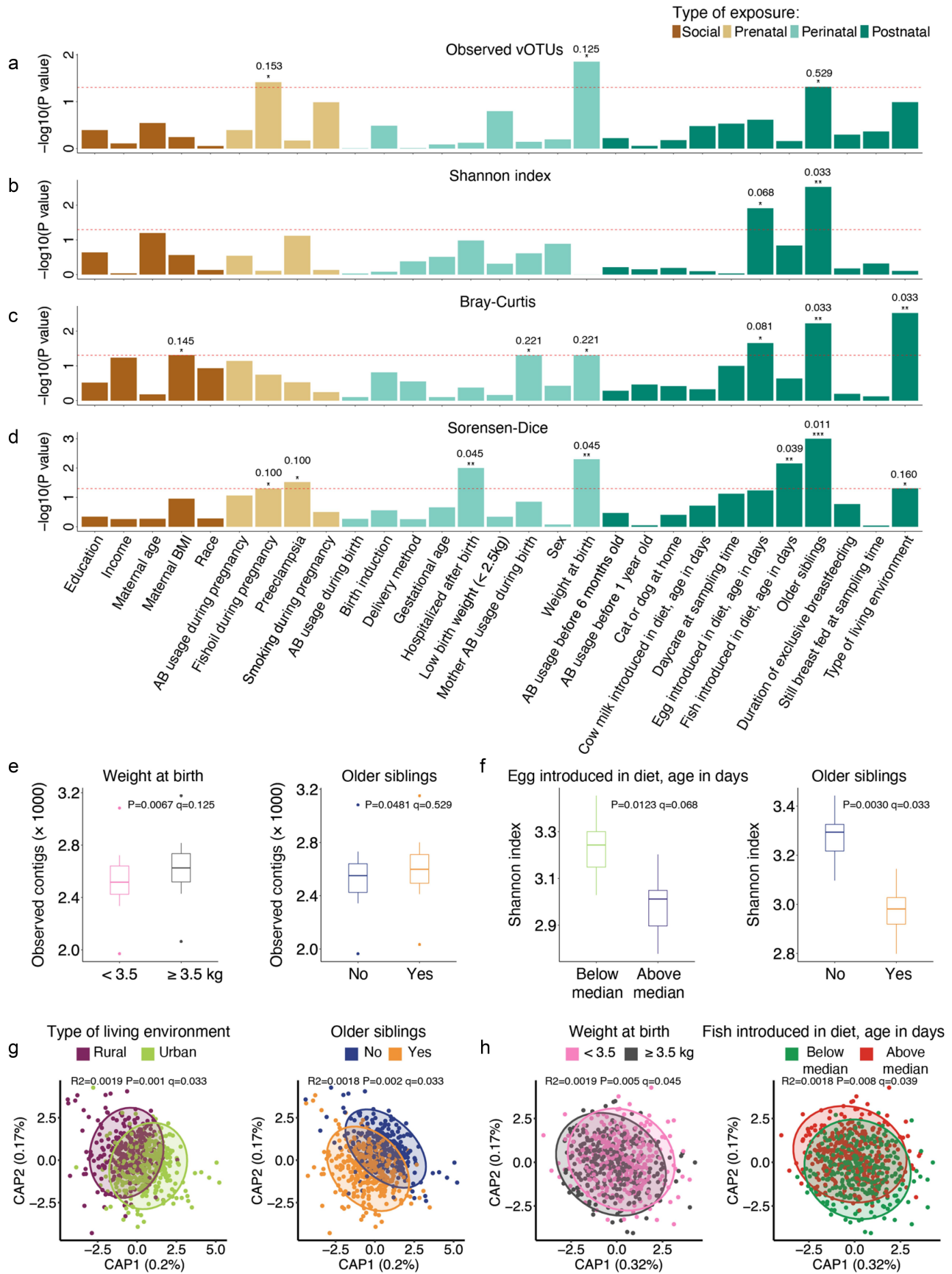
### Environmental exposures influence viral diversity

A range of pre-, peri-, and postnatal as well as social factors were recorded for the enrolled infants and their families (Supplementary Table S1). Having older siblings was associated with higher vOTU richness (linear mixed model,  $p = 0.048$ ,  $q = 0.529$ , estimate = 69.14, 95% CI = [0.58, 137.52]) and lower evenness (Shannon  $H'$ ) (linear mixed model,  $p = 0.003$ ,  $q = 0.033$ , estimate = -0.30, 95% CI = [-0.50, -0.10]) (Figures 2(a-b, e-f)) at 1 year of age. Likewise, a higher birth weight was linked to higher vOTU richness (linear mixed model,  $p = 0.007$ ,  $q = 0.125$ , estimate = -85.76, 95% CI = [-153.98, -17.56]) (Figure 2(a,e)). Dietary factors

were also found to influence the gut virome at 1 year of age, with late introduction of eggs in the diet being associated with lower viral evenness (Shannon  $H'$ ) (linear mixed model,  $p = 0.012$ ,  $q = 0.068$ , estimate = 0.25, 95% CI = [0.05, 0.45]) (Figure 2(a,f)). The mothers were enrolled in a nested randomized placebo-controlled trial of fish oil to the mothers during the third trimester of pregnancy.<sup>41,42</sup> Receiving fish oil during pregnancy was associated with increased gut vOTU richness (linear mixed model,  $p = 0.038$ ,  $q = 0.153$ , estimate = 71.60, 95% CI = [3.90, 139.22]) of the infants at 1 year of age (Figure 2(a)). The design also examined the difference in vitamin D between high and standard doses,<sup>43</sup> which had no effect on the viral community in our analysis. Interestingly, other factors that have been found to influence the bacterial GM component during infancy such as birth mode, use of antibiotics, and duration of exclusive breastfeeding<sup>1,26</sup> did not seem to influence gut virome alpha-diversity measures at 1 year of age in this cohort (Figure 2(a-b)).

Regarding virome composition, Bray-Curtis dissimilarity analysis (weighted measure, which is therefore mainly influenced by more abundant vOTUs) indicate a link (ANOVA-like permutation test,  $p = 0.049$ ,  $q = 0.145$ ,  $R^2 = 0.0016$ ) between maternal body mass index (BMI) and virome composition at 1 year of age (Figure 2(c) and S2A); while Sørensen-Dice distance (unweighted binary metric and therefore mainly influenced by more rare vOTUs) revealed that a number of pre- and perinatal exposures were linked with virome composition differences (ANOVA-like permutation test,  $p \leq 0.05$ ), namely weight at birth, fish oil supplementation during pregnancy, hospitalization after birth, and preeclampsia (Figure 2(d,h) and S2B). Both Bray-Curtis and Sørensen-Dice metrics showed significant differences in virome composition for children having older siblings (ANOVA-like permutation test,  $p = 0.006$ ,  $q = 0.033$ ,  $R^2 = 0.0018$  and  $p = 0.001$ ,  $q = 0.011$ ,  $R^2 = 0.0029$  for Bray-Curtis and Sørensen-Dice dissimilarity metrics, respectively), and whether the family was living in an urban or a rural area (ANOVA-like permutation test,  $p = 0.003$ ,  $q = 0.033$ ,  $R^2 = 0.0019$  and  $p = 0.049$ , FDR = 0.160,  $R^2 = 0.0016$  for Bray-Curtis and Sørensen-Dice dissimilarity metrics, respectively) (Figure 2(c-d,h) and S2A-D).





**Figure 2.** Virome diversity and composition covariates with early life exposures. (a–d) Barplot showing the strength of associations ( $-\log_{10}$  p-value) of the alpha diversity metrics observed vOTUs (a) and Shannon index (b) across different exposures (linear mixed

### Environmental exposure variables influence the abundance of specific vira

Subsequently, we determined how the distribution of vOTUs differed between the nine exposures (Figure 2c–d) found to significantly influence overall gut virome composition (preeclampsia was not included due to highly unbalanced sample size, see Supplementary Table S1). A total of 2,105 differentially abundant vOTUs affiliated to 173 viral families and 19 families of bacterial hosts were identified by DESeq2, with having older siblings being associated with 822 differential abundant vOTUs, while being hospitalized after birth being associated with 212 differential abundant vOTUs (Figure 3). For perinatal covariates, vOTUs differing in abundance were predicted to infect a range of different hosts, but interestingly revealed a pronounced lower abundance of vOTUs predicted as having *Bacteroidaceae*, *Ruminococcaceae* and *Streptococcaceae* as hosts with maternal antibiotic usage and hospitalization after birth (Figure 3). Postnatal factors like specific dietary patterns (late introduction of eggs in the diet), presence of older siblings in the house and living in a rural environment, were associated with a higher abundance of vOTUs infecting *Bifidobacteriaceae*, *Bacteroidaceae*, *Prevotellaceae*, *Tannerellaceae*, *Ruminococcaceae* and *Sutterellaceae*.

To further integrate these findings in the context of the gut bacterial component, we used 16S rRNA gene (V4 region) amplicon sequencing (bacterial OTUs – bOTUs) data previously published for this cohort<sup>2</sup> to determine virus-host co-abundances. Spearman correlation coefficients ( $\rho$ ) were calculated between the abundance of the above identified differentially different abundant vOTUs and bOTUs across samples. Only bOTUs that were strongly associated ( $\rho \geq 0.3$  and  $p$ -values  $\leq 0.05$ ) with at least one vOTU were retained. In most cases, when a vOTU was correlated with a bOTU, the bOTU family aligned with the predicted host family of the vOTU. However, some exceptions were observed, such as vOTU infecting

*Clostridiaceae* and *Lactobacillaceae* bOTU, reflecting the complexity of phage-host dynamics in the gut microbiome (Figure S3A). These virus-host co-abundances indicate there is a high degree of inter-relatedness between phages and their host in response to environmental exposures. This was supported by the fact that the same perinatal and postnatal covariates were also significantly associated with bOTU diversity and composition (Figure S4A–D). Overall, the 91 co-abundant vOTUs ( $\rho \geq 0.3$  and  $p$ -values  $\leq 0.05$ ), vOTUs that infect the same bacterial host family were in most cases closely related genetically, indicating a high degree of co-evolution between bacterial hosts and the phages that infect them (Figure S3B). To confirm the above-mentioned findings, we repeated the analysis of virus-host co-abundances using shotgun metagenomic data from the same cohort.<sup>44</sup> We found again that viruses and their bacterial hosts were highly correlated supporting the same conclusion as above (Figure S4E).

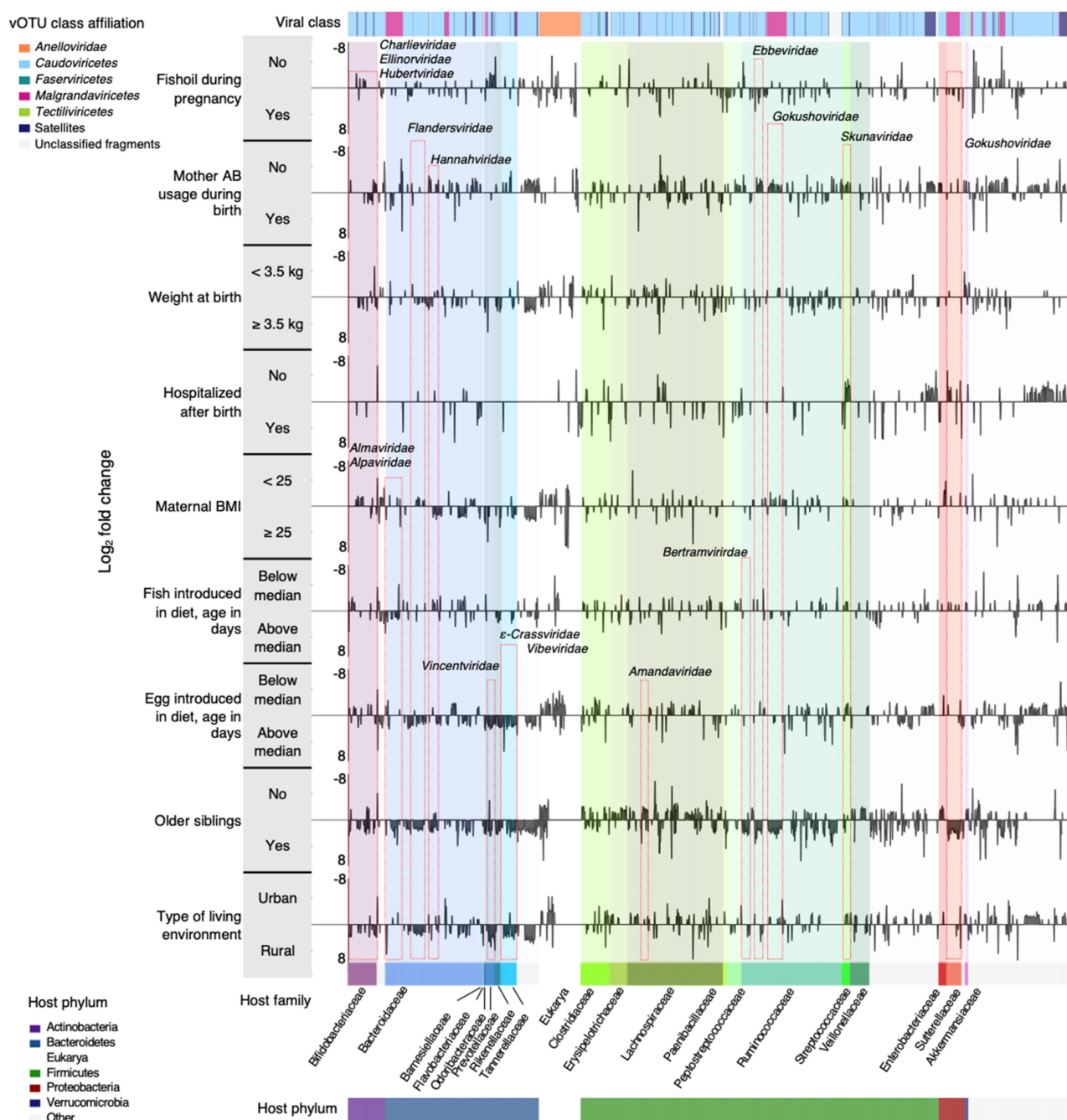
Because of the close relationship between phages and their bacterial hosts, we sought to determine whether bacterial communities mediate the effect of environmental exposures on the phages. Through mediation analysis, our findings indicate that, for certain exposures – such as having older siblings – bacterial communities partially mediate their influence on phage composition (Figure S5A). In contrast, other factors, including living environment and birth weight, operate more direct effects on the phageome (Figure S5B–C). Finally, some exposures such as age for introduction of eggs in the diet did not exhibit significant direct or indirect effects (Figure S5D).

### Functional profiles of gut viruses are linked with environmental exposures

Differentially abundant vOTUs were subjected to gene (open reading frame, ORF) prediction, and annotated based on KEGG Orthology (KO) using

---

model) as well as beta diversity using distance-based redundancy analysis (db-RDA) on bray-curtis dissimilarity (c) and sorenson-dice distance (d) matrices. Statistical significance levels are represented as  $p < 0.001$  (\*\*\*),  $p < 0.01$  (\*\*) and  $p \leq 0.05$  (\*). P-values were adjusted by Benjamini – Hochberg method and shown as q-values (only shown for significant exposures). (e,f) distribution of observed vOTUs for weight at birth and siblings (e) and Shannon index for siblings and dietary introduction of egg (f). Dietary introduction of egg is indicated in days. (g,h) db-RDA constrained-components based on bray-curtis distances for location and siblings (g), and sorenson-dice distances for weight at birth and dietary introduction of fish (h).



**Figure 3.** Viral host family, relative abundance and lifestyle associate with environmental exposures at one year of age. Visualization of differential abundance analysis of 2105 vOTUs across the nine exposures significantly associated with virome diversity and composition. Log<sub>2</sub> Fold change panel displays the change in abundance between the two groups for each exposure. The viral families to which vOTU belongs, surrounded by red boxes, are labeled. Adjusted  $p \leq 0.001$  and Log<sub>2</sub>-Fold changes  $\geq |1|$  were used to select differentially abundant vOTUs.

KofamScan.<sup>45</sup> As seen from figure S6A, 0.82% of genes matched known metabolism-related orthologs, while the remaining genes with KO assignments (8.48% of predicted genes) encoded genes related to genetic information processing and signaling and cellular processes, representing typical

viral-associated traits required to accomplish replication.<sup>46</sup> The remaining 90.7% of the predicted genes were not annotated by the database.

Next, we focused on determining genes with metabolic functions having the potential to enhance host fitness and drive metabolic

reprogramming of the bacterial host.<sup>47</sup> The gut virome of infants with older siblings were enriched in genes related to O-antigen nucleotide sugar biosynthesis and seleno-compound metabolism, while infants without siblings were enriched in genes related to carbon fixation in photosynthetic organisms (Fisher's exact test,  $p < 0.05$ ; Figure 4(a)) (the link to photosynthetic microorganisms may be caused by the KEGG database not being optimized for vira). The gut of infants living in rural areas or that were introduced to eggs in their diet later in life (above the median age when eggs were introduced in the diet) were enriched in viral encoded genes associated with glycolysis/gluconeogenesis and O-antigen nucleotide sugar biosynthesis, whereas the gut of infants living in urban areas or that were introduced to eggs relatively early in life were enriched in viral genes associated with thiamine metabolism. Infants with birth weight above the median or whose mothers who had a pre-pregnancy BMI  $< 25$  also encoded genes involved in diverse pathways involved in, e.g., vitamin synthesis. Further, the gut virome of infants whose mothers received fish oil during pregnancy or were prescribed antibiotics during delivery encoded genes related to purine metabolism (Figure 4(a)).

To determine how virally encoded gene functions associate with the microbial composition, we linked back enriched genes to the vOTU of origin (Figure 4(b)). 94% of lifestyle predicted vOTUs ( $n = 17$ ) were temperate. Genes associated with two classes of amino acid metabolisms (i.e. alanine, aspartate and glutamate metabolism and lysine biosynthesis) were conserved across *Alistipes* and *Faecalibacterium* targeting vOTUs, respectively. In addition, multiple carbohydrate metabolism enzyme encoding genes were found to be widely encoded by *Blautia*, *Prevotella*, *Ruminococcus* and *Faecalibacterium* targeting vOTUs. These encoded enzymes including L-lactate dehydrogenase, ribose-phosphate pyrophosphokinase and aldose 1-epimerase (Figure S6D). Energy metabolism genes were found in *Prevotella* and *Faecalibacterium* targeting vOTUs, while nicotine and nicotinamide metabolism genes were mapped in *Ruminococcus* and *Escherichia* targeting vOTUs (Figure 4(b)).

Phage-host co-abundance (Figure S3A), was further confirmed by Procrustes analysis. The linking of the virome and bacteriome compositions revealed a strong correlation one to another ( $p < 0.001$ ,  $r = 0.52$ ) (Figure S6B-C). The cumulative abundance of all bOTUs belonging to the bacterial genera *Ruminococcus*, *Prevotella* and *Faecalibacterium*, which were found to be the main bacterial host of vOTUs carrying the above metabolic genes (Figure 4(b)) and having previously been reported to be highly associated with stable viral communities,<sup>48</sup> was highly correlated with rural vs. urban living and having older siblings (Figure 4(c-e)). These results emphasize the potential role of phage-host association in metabolic regulation.

## Discussion

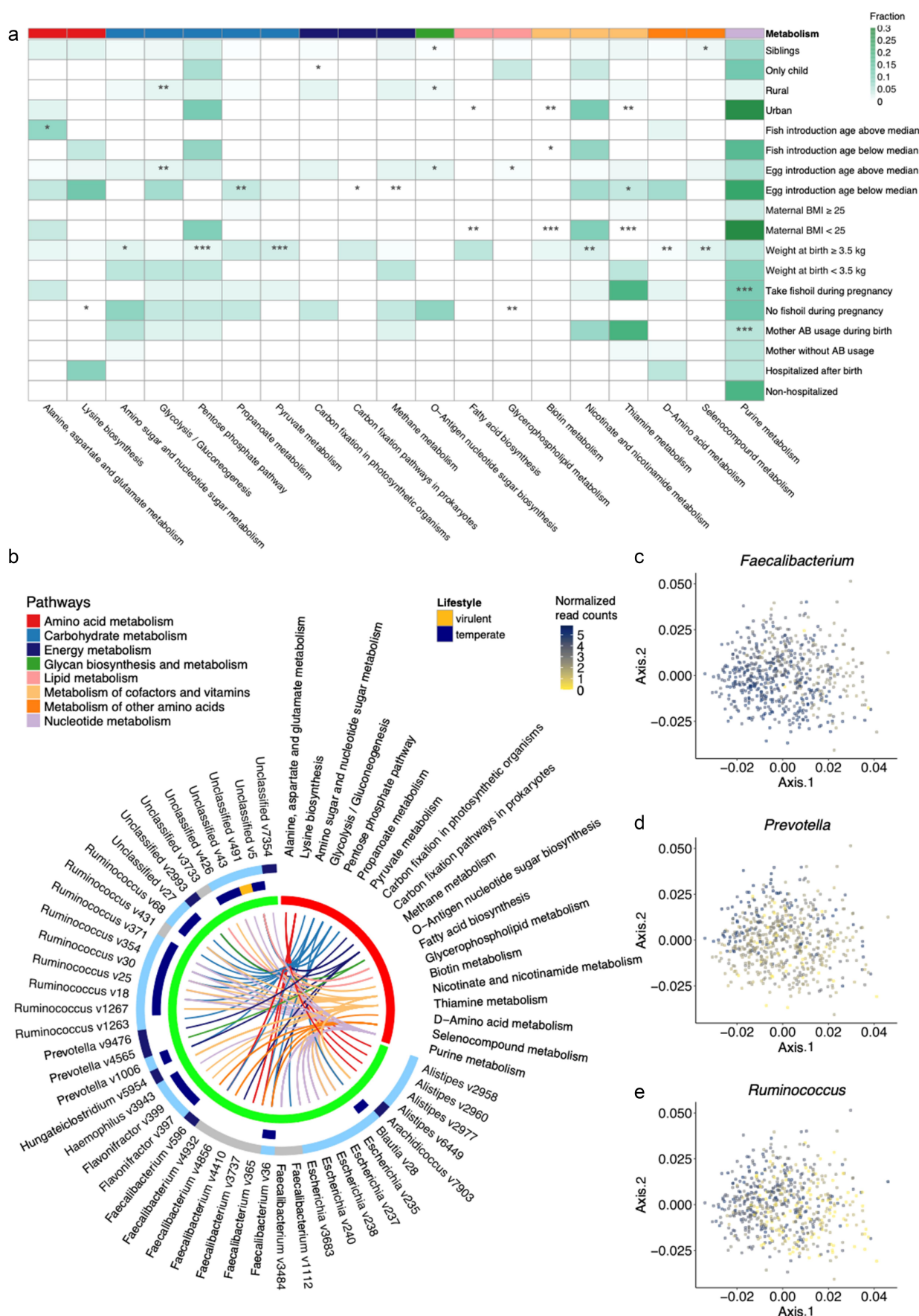
The gut of healthy newborns is usually devoid of viruses at birth, but it is rapidly colonized afterward.<sup>28,32</sup> Still relatively few studies have focused on the assembly of the gut virome within the first year of life and the factors that influence it<sup>6,27,28,32</sup> and even less is known about the environmental exposures that shape the gut virome.

Here, we leveraged a massive gut virome dataset from healthy infants at 1-year of age, and integrated measures of viral diversity such as sequence composition, viral hosts, and phage lifestyles,<sup>9</sup> (see Figure 1) with social, pre-, peri- and postnatal environmental exposures. We revealed the effects of these exposures on viral community and the possible effects on metabolism.

In previous reports, *Crassvirales* (class *Caudoviricetes*) and *Microviridae* (class *Malgrandaviricetes*) phages were found to be the two most abundant viral groups in the adult human gut, with their relative abundance being negatively correlated.<sup>19,48-51</sup> Here, in one-year-old infants, a similar observation was made, members of the *Caudoviricetes* and *Malgrandaviricetes* classes were the most abundant phages.

Interestingly, ongoing exposures such as having older siblings and residential location, as well as past exposures (e.g., birth weight, preeclampsia) were linked with gut virome composition at one year of age. However, it is still possible that the





**Figure 4.** Abundance of phage accessory genes differ in connection with exposures. (a) Abundance of genes (3<sup>rd</sup> level KEGG pathway) in the virome of infants with significant ( $p \leq 0.05$ ) enzymatic enrichments that are associated with the presence of siblings and residential location. (b) Viral host families that contribute to metabolism pathways. the outermost circle represents the vOTU class. The

prenatal and perinatal exposures still influenced the immune education earlier in life and remnants of the interplay are still tangible at 1-year of age.<sup>5</sup> In line with this, previous studies have shown that delivery mode and gestational diabetes influence the infant gut virome composition in the period immediately after birth and with gestational diabetes even being associated with higher abundance of pathogenic herpesvirus and poxvirus in meconium samples.<sup>30,52</sup> Among the exposures significantly influencing the gut virome composition, the largest effect sizes were from residential location (rural vs. urban) and having older siblings (see Figure 2(c,d)). Interestingly, urbanization has been reported to have a significant impact on the composition of the adult viral community, with individuals living in urban areas having higher abundance of *Lactococcus* (family *Streptococcaceae*) phages.<sup>53</sup> The latter is presumably associated with the consumption of dairy products. We show that the living environment also affects the gut virome of infants, and that *Streptococcaceae* targeting phages are also more abundant in infants living in urban areas, possibly reflecting differences in dietary habits rather than residence *per se* (Figure 3).

Having older siblings influences the development of the bacterial community in early life<sup>35,54,55</sup> and here we show that having older siblings is also associated with gut virome composition at 1 year of age. Both low as well as high birth weight has been associated with increased asthma risk during childhood.<sup>56,57</sup> Interestingly, in the present study, we also find that infants with a birth weight <2.5 kg differ in gut virome composition at age 1 year, relative to infants born with a weight >2.5 kg. Importantly, focusing on the curated DNA phage community (i.e. not the total vOTU pool, but only the fraction predicted as being bacteriophages) we recently demonstrated that phageome imbalances are associated with increased risk of developing before school age, but also that having older siblings was negatively associated with higher virome asthma signature score, implying a lower asthma

rate, showing how the influence of environmental exposures on virome composition also have implications for child health.<sup>22</sup> Importantly, from a translational angle, early-life exposures may affect the establishment of health phenotypes, such as the protective role of breastfeeding against eukaryotic-viral infections in the neonatal period.<sup>32</sup> Combining gut bacterial compositional data with gut virome composition (Figure S3A and S6B-C) in our cohort elucidates the co-abundance of phages and their hosts, underlying the role of phage-host interactions in shaping the GM. Most of these viruses (Figure S3A) have temperate lifestyles, as evidenced by the presence of genes coding for integrases. Thus, these temperate phages appear to have the ability to integrate their genome into the bacterial hosts and become prophages at some point. In a recent study<sup>22</sup> investigating the influence of early-life gut virome imbalances on the risk of developing asthma before school-age we used mediation analysis to show that gut viruses influence asthma partly mediated via gut bacteria partly independent of gut bacteria. Using a similar approach we in the present study show that for some exposures like the having older siblings, the influence on the gut phageome is partly mediated via gut bacteria. However, for other exposures such birth weight and living environment, the influence on the gut phageome is largely mediated independently of gut bacteria. How phages can influence the human host without being mediated via bacteria is far from fully understood but interactions with immune system via, e.g., TLR9 receptors likely play a role.<sup>22,58</sup>

Gut virome members have the potential to modulate biochemical processes.<sup>12,13,59</sup> The functional prediction of the genes derived from vOTUs co-varying with exposures, revealed up to 90% of genes with unknown functions. It emphasizes that proteins with yet uncharacterized functions are potentially playing a role in the regulation of human host phenotypes. Certain predicted gene functions linked to metabolic activities, such as alanine, aspartate and glutamate metabolism,

---

second circle represents the viral lifestyle (virulent vs. temperate). In the innermost circle, red represents metabolic pathways and green represents the vOTUs. The pathway and the family of the vOTU host are labeled. (c–e) Bacterial points extracted from the procrustes analysis (e). The points are colored according to the abundance of the specific genus in each sample.

amino sugar and nucleotide sugar metabolism and glycolysis/gluconeogenesis, which are likely associated with dietary intake and degradation of macronutrients, were associated with fish in the diet, birth weight, residence location and egg in the diet (Figure 4(a)). Birth weight also show negative correlation with higher virome asthma signature score (asthma rate),<sup>22</sup> and time-to-asthma analysis proved that the effects of phages on asthma were additive and statistically independent of the bacterial gut microbiome component.<sup>22</sup> Maternal obesity alters fatty acid metabolism and changes in gene expression of lipid metabolism in infants, which cause a higher risk of developing obesity and its complications, neuropsychiatric disorders and asthma.<sup>60,61</sup> We find here that viral genes associated with normal weight mothers were predominantly enriched in fatty acid biosynthesis compared to obese mothers, which may be an intermediate pathway by which maternal obesity affects child health. In addition, for biotin metabolism, which is known to be impaired by severe obesity,<sup>62</sup> many phage genes are also observed to be enriched in infants from mothers with BMI below 25 in our data. The mothers enrolled in the cohort participated in a randomized clinical trial where they were randomized to receiving fish oil or a placebo from week 24 of pregnancy to 1 week after birth.<sup>42,63</sup> The design also examined the difference in vitamin D between high and standard doses,<sup>43</sup> which had no effect on the viral community in our analysis. Of note, the supplementation of fish oil during pregnancy was not found to influence the gut bacterial component at age 1 year. Here, we report that the same intervention has some influence on the gut virome at age 1 year, but the effect is only borderline significant. The infants of mothers that received fish oil had viral genes involved in lysine biosynthesis, glycerophospholipid metabolism, and purine metabolism – metabolic activities that have been associated to fish oil supplementation,<sup>64,65</sup> but never attributed to gut virome composition. Interestingly, most of these metabolism-related genes were conserved across temperate vOTUs targeting *Ruminococcus*, *Faecalibacterium* and *Prevotella* spp. (Figure 4(b)). These genera have been consistently reported to be enriched in Danish and American subjects with a diet rich in carbohydrates, resistant starch, and

fibers, and being determinants of the so-called *Prevotella*-enterotype.<sup>66,67</sup> The *Prevotella*-enterotype is established early in life (between 9 and 36 months of age)<sup>68–70</sup> and have been previously suggested as markers of GM maturity at age 1 year.<sup>2,71,72</sup> Stockholm et al. (2018) reported delayed GM maturation as a risk factor for later development of asthma indicating the importance of these microbes for immune maturation.

Although our study is currently unable to assess how these gut virome-associated genes are actively involved in either enhancing either phage or host fitness, or both, our data as well as our recent finding, that early life gut virome imbalance is associated with increased asthma before school age risk<sup>22</sup> underlines the potential importance of bacteriophage-encoded metabolic genes and delivers an initial insight of the type of metabolic content conveyed by the gut virome in association to environmental variables.

In summary, our data provide detailed insight into the influence of common environmental factors that shape the gut virome during early life. We also uncover that key gut metabolic functions can be encoded by viral genes, which suggest that, in addition of shaping gut bacteriome composition, phages may directly play a role in metabolic activities.

## Methods

### Study participants

Participants belong to the COPSAC2010 cohort.<sup>34</sup> Fecal samples for virome extraction sequencing and analysis were collected from all infants at age 1 year.

### Sample collection, sequencing, virome assembly

Preparation of fecal samples, and extraction and sequencing of virions was carried out using a previously described protocol.<sup>37</sup> Briefly, viral-associated DNA was subjected to short MDA amplification and libraries were prepared following using manufacturer's procedures for the Nextera XT kit (FC-131-1096 Illumina, California). Libraries were single-end high-throughput sequenced on the Illumina HiSeq X platform.

Details of the pipeline for data processing, de-novo assembly, quality control, bacterial-host and life-style predictions, abundance-mapping (vOTU table), and taxonomy of complete and partial viral genomes (here termed vOTUs) can be found in Shah et al.<sup>9</sup> 16S rRNA gene amplicon data (bOTU table) from the same cohort's participants were retrieved from Stokholm et al. (2018).

### Environmental exposures

Briefly, during scheduled visits to the COPSAC clinic, information on a wide range of exposures was collected. We included a subset of variables, which were typical host, social and environmental factors suspected to influence the early life microbiome/virome. To avoid potential redundancy, we applied vifcor (Variance Inflation Factor and correlation-based filtering) in the R package “usdm” to identify and remove collinear variables, leading to the current variable list. A total of 30 environmental exposures were investigated and were grouped into social ( $n = 6$ ), pre- ( $n = 4$ ), peri- ( $n = 9$ ) and postnatal ( $n = 11$ ) exposures based on whether they occurred or existed before birth. See Supplementary Table S1 for a complete list of the exposures. For numeric variables like birth weight, we wanted the variable to capture being born with a low birth weight, and in pediatrics, this is defined by a weight  $< 2.5$  kg. These classifications of living areas are based on an objective measure of the living environment. This living environment variable included the land cover in a 3 km radius around the home address at birth using the European CORINE satellite-based land cover database.<sup>73</sup>

### Statistics and data analysis

Analyses on diversity were carried out on contingency tables gathering vOTUs abundance. Abundance data was normalized by reads per kilobase per million (RPKM). Alpha-diversity (Observed vOTUs and Shannon Index) indices and Beta-diversity (Bray-Curtis and Sørensen-Dice distances) matrices were generated using the package phyloseq (version 1.42.0).<sup>74</sup> The contribution of each covariate to explain vOTUs community structure (as determined by Sørensen-Dice similarity and Bray-Curtis dissimilarity metrics) was calculated using distance-based redundancy analysis (db-

RDA) models coupled to ANOVA-like permutation test ( $n$  permutations = 999) in package vegan (version 2.6–2),<sup>75</sup> while the effect size of the same covariates on alpha-diversity was calculated with linear mixed models from the package lmerTest (version 3.1–3).<sup>76</sup> All linear mixed models accounted for technical variation between runs using sequencing lane as the random effect. None of the exposures exhibited significant differences in dispersion.

Different differential abundance analysis methods were evaluated by DAtest.<sup>77</sup> DESeq2 (version 1.36.0) performed well with a low false-positive rate and a high ability to detect differential vOTUs for our data.<sup>78</sup> The sequencing lane was considered as a factor-covariate. The raw reads count table of each sample for vOTUs were prepared as input. All parameters are default except for sftype which is set to poscount. The Benjamin-Hochberg method was adapted to correct the p-values. vOTUs with adjusted p-value  $\leq 0.001$  and  $\log_2$  fold change  $\geq |1|$  were selected for downstream analyses.

Spearman's rank correlations were used to test univariate associations of continuous data, and results were visualized in a heatmap. MAFFT<sup>79</sup> was used to generate the phylogenetic tree file for those highly correlated vOTUs. The phylogenetic tree was visualized using the R package ggtree (version 3.4.0).<sup>80</sup> Procrustes analysis (R package vegan) was performed on vOTUs as target block and 16S rRNA gene data as rotatory block ( $n$  permutations = 999), while using the first two constrained components (CAP1 and CAP2) of db-RDA models for each data block. Bacterial and viral communities were each ordinated via principal coordinates analysis (PCoA) or db-RDA, and their unconstrained coordinate axes were used to represent overall community composition. The “mediation” R package was applied to assess whether bacterial composition partially mediates the early-life exposure effects on viral composition.

ORF calling on selected vOTUs was executed with Prodigal.<sup>81</sup> To determine metabolic function, genes were annotated based on KEGG Orthology using KofamScan<sup>45</sup> and filtered by default thresholds. Enricher function in clusterProfiler package (version 4.6.0) was applied to detect whether genes in differently abundant vOTUs were enriched in the metabolic pathway.<sup>82</sup>



All analyses were carried out in R (version 4.0.2) and results were visualized with the package ggplot2 (version 3.3.6).<sup>83</sup>

## Acknowledgments

We express our deepest gratitude to the children and families of the COPSAC<sub>2010</sub> cohort study for all their support and commitment. We acknowledge and appreciate the unique efforts of the COPSAC research team. We acknowledge the related preprints and abstract book was published somewhere else.<sup>78–80</sup>

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work is supported by the Joint Programming Initiative ‘Healthy Diet for a Healthy Life’, specifically here, the Danish Agency for Science and Higher Education, Institut National de la Recherche Agronomique (INRA), and the Canadian Institutes of Health Research (Team grant on Intestinal Microbiomics, Institute of Nutrition, Metabolism, and Diabetes, grant number [143924]). JT is supported by the BRIDGE Translational Excellence Program (bridge.ku.dk) at the Faculty of Health and Medical Sciences, University of Copenhagen, funded by the Novo Nordisk Foundation [grant no. NNF18SA0034956]. S.M. holds the Tier 1 Canada Research Chair in Bacteriophages [950-232136]. JS and DSN are recipients of Novo Nordisk Foundation grant [NNF20OC0061029].

## ORCID

Yichang Zhang  <http://orcid.org/0000-0001-5388-6376>  
 Josué L. Castro-Mejía  <http://orcid.org/0000-0002-1416-8205>  
 Ling Deng  <http://orcid.org/0000-0002-2190-9788>  
 Shiraz A. Shah  <http://orcid.org/0000-0002-4665-577X>  
 Jonathan Thorsen  <http://orcid.org/0000-0003-0200-0461>  
 Cristina Leal Rodríguez  <http://orcid.org/0000-0002-3133-0630>  
 Leon E. Jessen  <http://orcid.org/0000-0003-2879-2559>  
 Bo Chawes  <http://orcid.org/0000-0001-6846-6243>  
 Klaus Bønnelykke  <http://orcid.org/0000-0003-2003-1018>  
 Søren J. Sørensen  <http://orcid.org/0000-0001-6227-9906>  
 Hans Bisgaard  <http://orcid.org/0000-0003-4131-7592>  
 Sylvain Moineau  <http://orcid.org/0000-0002-2832-5101>  
 Marie-Agnès Petit  <http://orcid.org/0000-0003-2242-2768>  
 Jakob Stokholm  <http://orcid.org/0000-0003-4989-9769>  
 Dennis S. Nielsen  <http://orcid.org/0000-0001-8121-1114>

## Author contributions

L.D., S.M., M.A.P., J.S. and D.S.N. conceived the project and supervised all the research; B.C., K.B., J.S., S.J.S., L.J. and M.D. collected the samples and/or information; Y.Z., J.L.C.M., S.A. S. and D.S.N. analyzed the data; Y.Z., J.L.C.M. and D.S.N. wrote the manuscript with the assistance of L.D., S.A.S., J.T., C.L.R., S.M., M.A.P. and J.S.; L.D. prepared the virome and sequencing libraries; all authors contributed to, revised and approved the final manuscript.

## Data availability statement

Sequencing FASTQ files are available on ENA under project number PRJEB46943. All cohort participants’ individual-level data are protected by Danish and European law and are not publicly available. However, participant-level data can be made available under a data transfer agreement as part of a collaboration effort. Codes for data analyses are available on GitHub (<https://github.com/Yichang07/EarlyVir.git>).

## Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki and was approved by The National Committee on Health Research Ethics (H-B-2008-093) and the Danish Data Protection Agency (2015-41-3696). Both parents gave written informed consent before enrollment.

## References

1. Tamburini S, Shen N, Wu HC, Clemente JC. The microbiome in early life: implications for health outcomes. *Nat Med*. 2016;22(7):713–722. doi: [10.1038/nm.4142](https://doi.org/10.1038/nm.4142).
2. Stokholm J, Blaser MJ, Thorsen J, Rasmussen MA, Waage J, Vinding RK, Schoos AMM, Kunøe A, Fink NR, Chawes BL, et al. Maturation of the gut microbiome and risk of asthma in childhood. *Nat Commun*. 2018;9(1):141. doi: [10.1038/s41467-017-02573-2](https://doi.org/10.1038/s41467-017-02573-2).
3. Pihl AF, Fonvig CE, Stjernholm T, Hansen T, Pedersen O, Holm J-C. The role of the gut microbiota in childhood obesity. *Child Obes*. 2016;12(4):292–299. doi: [10.1089/chi.2015.0220](https://doi.org/10.1089/chi.2015.0220).
4. Kostic AD, Gevers D, Siljander H, Vatanen T, Hyötyläinen T, Hämäläinen A-M, Peet A, Tillmann V, Pöhö P, Mattila I, et al. The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. *Cell Host & Microbe*. 2015;17(2):260–273. doi: [10.1016/j.chom.2015.01.001](https://doi.org/10.1016/j.chom.2015.01.001).
5. Vatanen T, Kostic AD, d’Hennzel E, Siljander H, Franzosa E, Yassour M, Kolde R, Vlamakis H, Arthur T, Hämäläinen A-M, et al. Variation in microbiome LPS immunogenicity contributes to autoimmunity in

- humans. *Cell*. 2016;165(4):842–853. doi: [10.1016/j.cell.2016.04.007](https://doi.org/10.1016/j.cell.2016.04.007).
6. McCann A, Ryan FJ, Stockdale SR, Dalmaso M, Blake T, Ryan CA, Stanton C, Mills S, Ross PR, Hill C. Viromes of one year old infants reveal the impact of birth mode on microbiome diversity. *PeerJ*. 2018;6:e4694. doi: [10.7717/peerj.4694](https://doi.org/10.7717/peerj.4694).
7. Sutton TDS, Hill C. Gut bacteriophage: current understanding and challenges. *Front Endocrinol*. 2019;10:784. doi: [10.3389/fendo.2019.00784](https://doi.org/10.3389/fendo.2019.00784).
8. Shkoporov AN, Hill C. Bacteriophages of the human gut: the “known unknown” of the microbiome. *Cell Host Microbe*. 2019;25(2):195–209. doi: [10.1016/j.chom.2019.01.017](https://doi.org/10.1016/j.chom.2019.01.017).
9. Shah SA, Deng L, Thorsen J, Pedersen AG, Dion MB, Castro-Mejía JL, Silins R, Romme FO, Sausset R, Jessen LE, et al. Expanding known viral diversity in the healthy infant gut. *Nat Microbiol*. 2023;8(5):986–998. doi: [10.1038/s41564-023-01345-7](https://doi.org/10.1038/s41564-023-01345-7).
10. Gregory AC, Zablocki O, Zayed AA, Howell A, Bolduc B, Sullivan MB. The gut virome database reveals age-dependent patterns of virome diversity in the human gut. *Cell Host Microbe*. 2020;28(5):724–740.e8. doi: [10.1016/j.chom.2020.08.003](https://doi.org/10.1016/j.chom.2020.08.003).
11. Roux S, Paez-Espino D, Chen IMA, Palaniappan K, Ratner A, Chu K, Reddy TBK, Nayfach S, Schulz F, Call L, et al. IMG/VR v3: an integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses. *Nucleic Acids Research*. 2021;49(D1):D764–D775. doi: [10.1093/nar/gkaa946](https://doi.org/10.1093/nar/gkaa946).
12. Kieft K, Zhou ZC, Anderson RE, Buchan A, Campbell BJ, Hallam SJ, Hess M, Sullivan MB, Walsh DA, Roux S, et al. Ecology of inorganic sulfur auxiliary metabolism in widespread bacteriophages. *Nat Commun*. 2021;12(1):3503. doi: [10.1038/s41467-021-23698-5](https://doi.org/10.1038/s41467-021-23698-5).
13. Bi L, Yu DT, Du S, Zhang L-M, Zhang L-Y, Wu C-F, Xiong C, Han L-L, He J-Z. Diversity and potential biogeochemical impacts of viruses in bulk and rhizosphere soils. *Environmental Microbiology*. 2021;23(2):588–599. doi: [10.1111/1462-2920.15010](https://doi.org/10.1111/1462-2920.15010).
14. Tisza MJ, Buck CB. A catalog of tens of thousands of viruses from human metagenomes reveals hidden associations with chronic diseases. *Proc Natl Acad Sci*. 2021;118(23):e2023202118. doi: [10.1073/pnas.2023202118](https://doi.org/10.1073/pnas.2023202118).
15. Clooney AG, Sutton TDS, Shkoporov AN, Holohan RK, Daly KM, O'Regan O, Ryan FJ, Draper LA, Plevy SE, Ross RP, et al. Whole-virome analysis sheds light on viral dark matter in inflammatory bowel disease. *Cell Host Microbe*. 2019;26(6):764–778.e5. doi: [10.1016/j.chom.2019.10.009](https://doi.org/10.1016/j.chom.2019.10.009).
16. Zhao LY, Shi Y, Lau HC, Liu W, Luo G, Wang G, Liu C, Pan Y, Zhou Q, Ding Y, et al. 441: Uncovering 1,058 novel human enteric dna viruses through deep long-read third-generation sequencing and their clinical impact. *Gastroenterology*. 2022;162(7):S96–S96. doi: [10.1016/S0016-5085\(22\)60240-3](https://doi.org/10.1016/S0016-5085(22)60240-3).
17. Kaelin EA, Rodriguez C, Hall-Moore C, Hoffmann JA, Linneman LA, Ndao IM, Warner BB, Tarr PI, Holtz LR, Lim ES. Longitudinal gut virome analysis identifies specific viral signatures that precede necrotizing enterocolitis onset in preterm infants. *Nat Microbiol*. 2022;7(5):653–662. doi: [10.1038/s41564-022-01096-x](https://doi.org/10.1038/s41564-022-01096-x).
18. Reyes A, Blanton LV, Cao S, Zhao G, Manary M, Trehan I, Smith MI, Wang D, Virgin HW, Rohwer F, et al. Gut DNA viromes of Malawian twins discordant for severe acute malnutrition. *Proc Natl Acad Sci USA*. 2015;112(38):11941–11946. doi: [10.1073/pnas.1514285112](https://doi.org/10.1073/pnas.1514285112).
19. Zhao GY, Vatanen T, Droit L, Park A, Kostic AD, Poon TW, Vlamakis H, Siljander H, Härkönen T, Hämäläinen A-M, et al. Intestinal virome changes precede autoimmunity in type I diabetes-susceptible children. *Proc Natl Acad Sci USA*. 2017;114(30):E6166–E6175. doi: [10.1073/pnas.1706359114](https://doi.org/10.1073/pnas.1706359114).
20. Vehik K, Lynch KF, Wong MC, Tian X, Ross MC, Gibbs RA, Ajami NJ, Petrosino JF, Rewers M, Toppari J, et al. Prospective virome analyses in young children at increased genetic risk for type 1 diabetes. *Nat Med*. 2019;25(12):1865–1872. doi: [10.1038/s41591-019-0667-0](https://doi.org/10.1038/s41591-019-0667-0).
21. Tomofuji Y, Kishikawa T, Maeda Y, Ogawa K, Nii T, Okuno T, Oguro-Igashira E, Kinoshita M, Yamamoto K, Sonehara K, et al. Whole gut virome analysis of 476 Japanese revealed a link between phage and autoimmune disease. *Ann Rheum Dis*. 2021;81(2):278–288. doi: [10.1136/annrheumdis-2021-221267](https://doi.org/10.1136/annrheumdis-2021-221267).
22. Leal Rodriguez C, Shah SA, Rasmussen MA, Thorsen J, Boulund U, Pedersen CET, Castro-Mejía JL, Poulsen CE, Poulsen CS, Deng L, et al. The infant gut virome is associated with preschool asthma risk independently of bacteria. *Nat Med*. 2023;30(1):138–148. doi: [10.1038/s41591-023-02685-x](https://doi.org/10.1038/s41591-023-02685-x).
23. Rasmussen TS, Mentzel CMJ, Kot W, Castro-Mejía JL, Zuffa S, Swann JR, Hansen LH, Vogensen FK, Hansen AK, Nielsen DS. Faecal virome transplantation decreases symptoms of type 2 diabetes and obesity in a murine model. *Gut*. 2020;69(12):2122–2130. doi: [10.1136/gutjnl-2019-320005](https://doi.org/10.1136/gutjnl-2019-320005).
24. Ott SJ, Waetzig GH, Rehman A, Moltzau-Anderson J, Bharti R, Grasis JA, Cassidy L, Tholey A, Fickenscher H, Seegert D, et al. Efficacy of sterile fecal filtrate transfer for treating patients with clostridium difficile infection. *Gastroenterology*. 2017;152(4):799–811.e7. doi: [10.1053/j.gastro.2016.11.010](https://doi.org/10.1053/j.gastro.2016.11.010).
25. Brunse A, Deng L, Pan XY, Hui Y, Castro-Mejía JL, Kot W, Nguyen DN, Secher JBM, Nielsen DS, Thymann T. Fecal filtrate transplantation protects against necrotizing enterocolitis. *ISME J*. 2022;16(3):686–694. doi: [10.1038/s41396-021-01107-5](https://doi.org/10.1038/s41396-021-01107-5).
26. Stewart CJ, Ajami NJ, O'Brien JL, Hutchinson DS, Smith DP, Wong MC, Ross MC, Lloyd RE, Doddapaneni H, Metcalf GA, et al. Temporal development of the gut microbiome in early childhood from

- the TEDDY study. *Nature*. 2018;562(7728):583–588. doi: [10.1038/s41586-018-0617-x](https://doi.org/10.1038/s41586-018-0617-x).
27. Breitbart M, Haynes M, Kelley S, Angly F, Edwards RA, Felts B, Mahaffy JM, Mueller J, Nulton J, Rayhawk S, et al. Viral diversity and dynamics in an infant gut. *Research Microbiology*. 2008;159(5):367–373. doi: [10.1016/j.resmic.2008.04.006](https://doi.org/10.1016/j.resmic.2008.04.006).
  28. Lim ES, Zhou YJ, Zhao GY, Bauer IK, Droit L, Ndao IM, Warner BB, Tarr PI, Wang D, Holtz LR. Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nat Med*. 2015;21(10):1228–1234. doi: [10.1038/nm.3950](https://doi.org/10.1038/nm.3950).
  29. Taboada B, Moran P, Serrano-Vazquez A, Iša P, Rojas-Velázquez L, Pérez-Juárez H, López S, Torres J, Ximenez C, Arias CF, et al. The gut virome of healthy children during the first year of life is diverse and dynamic. *PLOS ONE*. 2021;16(4):e0240958. doi: [10.1371/journal.pone.0240958](https://doi.org/10.1371/journal.pone.0240958).
  30. Shi X, Liu Y, Ma T, Jin H, Zhao F, Sun Z. Delivery mode and maternal gestational diabetes are important factors in shaping the neonatal initial gut microbiota. *Front Cell Infect Microbiol*. 2024;14:1397675. doi: [10.3389/fcimb.2024.1397675](https://doi.org/10.3389/fcimb.2024.1397675).
  31. Liang G, Bushman FD. The human virome: assembly, composition and host interactions. *Nat Rev Microbiol*. 2021;19(8):514–527. doi: [10.1038/s41579-021-00536-5](https://doi.org/10.1038/s41579-021-00536-5).
  32. Liang GX, Zhao CY, Zhang HJ, Mattei L, Sherrill-Mix S, Bittinger K, Kessler LR, Wu GD, Baldassano RN, DeRusso P, et al. The stepwise assembly of the neonatal virome is modulated by breastfeeding. *Nature*. 2020;581(7809):470–474. doi: [10.1038/s41586-020-2192-1](https://doi.org/10.1038/s41586-020-2192-1).
  33. Walters WA, Granados AC, Ley C, Federman S, Stryke D, Santos Y, Haggerty T, Sotomayor-Gonzalez A, Servellita V, Ley RE, et al. Longitudinal comparison of the developing gut virome in infants and their mothers. *Cell Host & Microbe*. 2023;31(2):187–198.e3. doi: [10.1016/j.chom.2023.01.003](https://doi.org/10.1016/j.chom.2023.01.003).
  34. Bisgaard H, Vissing NH, Carson CG, Bischoff AL, Følsgaard NV, Kreiner-Møller E, Chawes BLK, Stokholm J, Pedersen L, Bjarnadóttir E, et al. Deep phenotyping of the unselected COPSAC2010 birth cohort study. *Clin Experimental Allergy*. 2013;43(12):1384–1394. doi: [10.1111/cea.12213](https://doi.org/10.1111/cea.12213).
  35. Christensen ED, Hjelmso MH, Thorsen J, Shah S, Redgwell T, Poulsen CE, Trivedi U, Russel J, Gupta S, Chawes BL, et al. The developing airway and gut microbiota in early life is influenced by age of older siblings. *Microbiome*. 2022;10(1):106. doi: [10.1186/s40168-022-01305-z](https://doi.org/10.1186/s40168-022-01305-z).
  36. Gasparrini AJ, Wang B, Sun X, Kennedy EA, Hernandez-Leyva A, Ndao IM, Tarr PI, Warner BB, Dantas G. Persistent metagenomic signatures of early-life hospitalization and antibiotic treatment in the infant gut microbiota and resistome. *Nat Microbiol*. 2019;4(12):2285–2297. doi: [10.1038/s41564-019-0550-2](https://doi.org/10.1038/s41564-019-0550-2).
  37. Deng L, Silins R, Castro-Mejia JL, Kot W, Jessen L, Thorsen J, Shah S, Stokholm J, Bisgaard H, Moineau S, et al. A protocol for extraction of infective viromes suitable for metagenomics sequencing from low volume fecal samples. *Viruses*. 2019;11(7):667. doi: [10.3390/v11070667](https://doi.org/10.3390/v11070667).
  38. Edwards RA, McNair K, Faust K, Raes J, Dutilh BE. Computational approaches to predict bacteriophage–host relationships. *FEMS Microbiology Reviews*. 2016;40(2):258–272. doi: [10.1093/femsre/fuv048](https://doi.org/10.1093/femsre/fuv048).
  39. Pride DT, Wassenaar TM, Ghose C, Blaser MJ. Evidence of host-virus co-evolution in tetranucleotide usage patterns of bacteriophages and eukaryotic viruses. *BMC Genom*. 2006;7(1):1–13. doi: [10.1186/1471-2164-7-8](https://doi.org/10.1186/1471-2164-7-8).
  40. Krupovic M, Forterre P, Qiu J. Microviridae goes temperate: microvirus-related proviruses reside in the genomes of bacteroidetes. *PLOS ONE*. 2011;6(5):e19893. doi: [10.1371/journal.pone.0019893](https://doi.org/10.1371/journal.pone.0019893).
  41. Vinding RK, Stokholm J, Sevelsted A, Chawes BL, Bønnelykke K, Barman M, Jacobsson B, Bisgaard H. Fish oil supplementation in pregnancy increases gestational age, size for gestational age, and birth weight in infants: a randomized controlled trial. *J Nutr*. 2019;149(4):628–634. doi: [10.1093/jn/nxy204](https://doi.org/10.1093/jn/nxy204).
  42. Bisgaard H, Stokholm J, Chawes BL, Vissing NH, Bjarnadóttir E, Schoos AMM, Wolsk HM, Pedersen TM, Vinding RK, Thorsteinsdóttir S, et al. Fish oil-derived fatty acids in pregnancy and wheeze and asthma in offspring. *N Engl J Med*. 2016;375(26):2530–2539. doi: [10.1056/NEJMoa1503734](https://doi.org/10.1056/NEJMoa1503734).
  43. Chawes BL, Bønnelykke K, Stokholm J, Vissing NH, Bjarnadóttir E, Schoos AMM, Wolsk HM, Pedersen TM, Vinding RK, Thorsteinsdóttir S, et al. Effect of vitamin D 3 supplementation during pregnancy on risk of persistent wheeze in the offspring. *JAMA*. 2016;315(4):353–361. doi: [10.1001/jama.2015.18318](https://doi.org/10.1001/jama.2015.18318).
  44. Li XJ, Stokholm J, Brejnrod A, Vestergaard GA, Russel J, Trivedi U, Thorsen J, Gupta S, Hjelmso MH, Shah SA, et al. The infant gut resistome associates with *E. coli*, environmental exposures, gut microbiome maturity, and asthma-associated bacterial composition. *Cell Host & Microbe*. 2021;29(6):975–987.e4. doi: [10.1016/j.chom.2021.03.017](https://doi.org/10.1016/j.chom.2021.03.017).
  45. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics*. 2020;36(7):2251–2252. doi: [10.1093/bioinformatics/btz859](https://doi.org/10.1093/bioinformatics/btz859).
  46. Sanz-Gaitero M, Seoane-Blanco M, van Raaij MJ. Structure and function of bacteriophages. *Bacteriophages Biol Technol Ther*. 2021;19–91.
  47. Warwick-Dugdale J, Buchholz HH, Allen MJ, Temperton B. Host-hijacking and planktonic piracy: how phages command the microbial high seas. *Virol J*. 2019;16(1):1–13. doi: [10.1186/s12985-019-1120-1](https://doi.org/10.1186/s12985-019-1120-1).
  48. Shkoporov AN, Clooney AG, Sutton TDS, Ryan FJ, Daly KM, Nolan JA, McDonnell SA, Khokhlova EV, Draper LA, Forde A, et al. The human gut virome is highly diverse, stable, and individual specific. *Cell Host*

- & Microbe. 2019;26(4):527–541.e5. doi: [10.1016/j.chom.2019.09.009](https://doi.org/10.1016/j.chom.2019.09.009).
49. Kim KH, Bae JW. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. *Appl Environ Microbiol.* 2011;77(21):7663–7668. doi: [10.1128/AEM.00289-11](https://doi.org/10.1128/AEM.00289-11).
  50. Roux S, Solonenko NE, Dang VT, Poulos BT, Schwenck SM, Goldsmith DB, Coleman ML, Breitbart M, Sullivan MB. Towards quantitative viromics for both double-stranded and single-stranded DNA viruses. *PeerJ.* 2016;4:e2777. doi: [10.7717/peerj.2777](https://doi.org/10.7717/peerj.2777).
  51. Yutin N, Makarova KS, Gussow AB, Krupovic M, Segall A, Edwards RA, Koonin EV. Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. *Nat Microbiol.* 2018;3(1):38–46. doi: [10.1038/s41564-017-0053-y](https://doi.org/10.1038/s41564-017-0053-y).
  52. Wang J, Zheng J, Shi W, Du N, Xu X, Zhang Y, Ji P, Zhang F, Jia Z, Wang Y, et al. Dysbiosis of maternal and neonatal microbiota associated with gestational diabetes mellitus. *Gut.* 2018;67(9):1614–1625. doi: [10.1136/gutjnl-2018-315988](https://doi.org/10.1136/gutjnl-2018-315988).
  53. Zuo T, Sun Y, Wan YT, Yeoh YK, Zhang F, Cheung CP, Chen N, Luo J, Wang W, Sung JJY, et al. Human-gut-DNA virome variations across geography, ethnicity, and urbanization. *Cell Host & Microbe.* 2020;28(5):741–751.e4. doi: [10.1016/j.chom.2020.08.005](https://doi.org/10.1016/j.chom.2020.08.005).
  54. Laursen MF, Zachariassen G, Bahl MI, Bergström A, Høst A, Michaelsen KF, Licht TR. Having older siblings is associated with gut microbiota development during early childhood. *BMC Microbiol.* 2015;15(1):1–9. doi: [10.1186/s12866-015-0477-6](https://doi.org/10.1186/s12866-015-0477-6).
  55. Hedin CR, van der Gast CJ, Stagg AJ, Lindsay JO, Whelan K. The gut microbiota of siblings offers insights into microbial pathogenesis of inflammatory bowel disease. *Gut Microbes.* 2017;8(4):359–365. doi: [10.1080/19490976.2017.1284733](https://doi.org/10.1080/19490976.2017.1284733).
  56. Liu X, Olsen J, Agerbo E, Yuan W, Cnattingius S, Gissler M, Li J. Birth weight, gestational age, fetal growth and childhood asthma hospitalization. *All Asth Clin Immun.* 2014;10(1):13. doi: [10.1186/1710-1492-10-13](https://doi.org/10.1186/1710-1492-10-13).
  57. Sin DD, Spier S, Svenson LW, Schopflocher DP, Senthilselvan A, Cowie RL, Man SFP. The relationship between birth weight and childhood asthma: a population-based cohort study. *Arch Pediatr Adolesc Med.* 2004;158(1):60–64. doi: [10.1001/archpedi.158.1.60](https://doi.org/10.1001/archpedi.158.1.60).
  58. Gogokhia L, Buhrke K, Bell R, Hoffman B, Brown DG, Hanke-Gogokhia C, Ajami NJ, Wong MC, Ghazaryan A, Valentine JF, et al. Expansion of bacteriophages is linked to aggravated intestinal inflammation and colitis. *Cell Host & Microbe.* 2019;25(2):285–299.e8. doi: [10.1016/j.chom.2019.01.008](https://doi.org/10.1016/j.chom.2019.01.008).
  59. Brown EM, Arellano-Santoyo H, Temple ER, Costliow ZA, Pichaud M, Hall AB, Liu K, Durney MA, Gu X, Plichta DR, et al. Gut microbiome ADP-ribosyltransferases are widespread phage-encoded fitness factors. *Cell Host & Microbe.* 2021;29(9):1351–1365.e11. doi: [10.1016/j.chom.2021.07.011](https://doi.org/10.1016/j.chom.2021.07.011).
  60. Alvarez D, Munoz Y, Ortiz M, Maliqueo M, Chouinard-Watkins R, Valenzuela R. Impact of maternal obesity on the metabolism and bioavailability of polyunsaturated fatty acids during pregnancy and breastfeeding. *Nutrients.* 2020;13(1):19. doi: [10.3390/nu13010019](https://doi.org/10.3390/nu13010019).
  61. Costa SM, Isganaitis E, Matthews TJ, Hughes K, Daher G, Dreyfuss JM, da Silva GAP, Patti M-E. Maternal obesity programs mitochondrial and lipid metabolism gene expression in infant umbilical vein endothelial cells. *Int J Obes (Lond).* 2016;40(11):1627–1634. doi: [10.1038/ijo.2016.142](https://doi.org/10.1038/ijo.2016.142).
  62. Belda E, Volland L, Tremaroli V, Falony G, Adriouch S, Assmann KE, Prifti E, Aron-Wisniewsky J, Debédat J, Le Roy T, et al. Impairment of gut microbial biotin metabolism and host biotin status in severe obesity: effect of biotin and prebiotic supplementation on improved metabolism. *Gut.* 2022;71(12):2463–2480. doi: [10.1136/gutjnl-2021-325753](https://doi.org/10.1136/gutjnl-2021-325753).
  63. Hjelmso MH, Shah SA, Thorsen J, Rasmussen M, Vestergaard G, Mortensen MS, Brejnrod A, Brix S, Chawes B, Bønnelykke K, et al. Prenatal dietary supplements influence the infant airway microbiota in a randomized factorial clinical trial. *Nat Commun.* 2020;11(1):426. doi: [10.1038/s41467-020-14308-x](https://doi.org/10.1038/s41467-020-14308-x).
  64. Hishikawa D, Valentine WJ, Iizuka-Hishikawa Y, Shindou H, Shimizu T. Metabolism and functions of docosahexaenoic acid-containing membrane glycerophospholipids. *FEBS Letters.* 2017;591(18):2730–2744. doi: [10.1002/1873-3468.12825](https://doi.org/10.1002/1873-3468.12825).
  65. Cao Y, Gao Q, Li X, Zhou Y, Dong S, Wang Y, Dai Z. Integrated analysis of metabolomics and transcriptomics for assessing effects of fish meal and fish oil replacement on the metabolism of rainbow trout (*Oncorhynchus mykiss*) [original research]. *Front Mar Sci.* 2022;9:208. doi: [10.3389/fmars.2022.843637](https://doi.org/10.3389/fmars.2022.843637).
  66. Roager HM, Vogt JK, Kristensen M, Hansen LBS, Ibrügger S, Mærkedahl RB, Bahl MI, Lind MV, Nielsen RL, Frøkiær H, et al. Whole grain-rich diet reduces body weight and systemic low-grade inflammation without inducing major changes of the gut microbiome: a randomised cross-over trial. *Gut.* 2019;68(1):83–93. doi: [10.1136/gutjnl-2017-314786](https://doi.org/10.1136/gutjnl-2017-314786).
  67. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R, et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science.* 2011;334(6052):105–108. doi: [10.1126/science.1208344](https://doi.org/10.1126/science.1208344).
  68. Bergstrom A, Skov TH, Bahl MI, Roager HM, Christensen LB, Ejlerskov KT, Mølgaard C, Michaelsen KF, Licht TR. Establishment of intestinal microbiota during early life: a longitudinal, explorative study of a large cohort of Danish infants. *Appl Environ Microbiol.* 2014;80(9):2889–2900. doi: [10.1128/AEM.00342-14](https://doi.org/10.1128/AEM.00342-14).



69. Backhed F, Roswall J, Peng YQ, Feng Q, Jia H, Kovatcheva-Datchary P, Li Y, Xia Y, Xie H, Zhong H, et al. Dynamics and stabilization of the human gut microbiome during the first year of life. *Cell Host & Microbe*. 2015;17(5):690–703. doi: [10.1016/j.chom.2015.04.004](https://doi.org/10.1016/j.chom.2015.04.004).
70. Roswall J, Olsson LM, Kovatcheva-Datchary P, Nilsson S, Tremaroli V, Simon M-C, Kiilerich P, Akrami R, Krämer M, Uhlén M, et al. Developmental trajectory of the healthy human gut microbiota during the first 5 years of life. *Cell Host & Microbe*. 2021;29(5):765–776.e3. doi: [10.1016/j.chom.2021.02.021](https://doi.org/10.1016/j.chom.2021.02.021).
71. Larsen JM. The immune response to *Prevotella* bacteria in chronic inflammatory disease. *Immunology*. 2017;151(4):363–374. doi: [10.1111/imm.12760](https://doi.org/10.1111/imm.12760).
72. Park H, Shin JW, Park SG, Kim W. Microbial communities in the upper respiratory tract of patients with asthma and chronic obstructive pulmonary disease. *PLOS ONE*. 2014;9(10):e109710. doi: [10.1371/journal.pone.0109710](https://doi.org/10.1371/journal.pone.0109710).
73. Lehtimäki J, Thorsen J, Rasmussen MA, Hjelmsø M, Shah S, Mortensen MS, Trivedi U, Vestergaard G, Bønnelykke K, Chawes BL, et al. Urbanized microbiota in infants, immune constitution, and later risk of atopic diseases. *J Allergy Clin Immunol*. 2021;148(1):234–243. doi: [10.1016/j.jaci.2020.12.621](https://doi.org/10.1016/j.jaci.2020.12.621).
74. McMurdie PJ, Holmes S, Watson M. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE*. 2013;8(4):e61217. doi: [10.1371/journal.pone.0061217](https://doi.org/10.1371/journal.pone.0061217).
75. Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci*. 2003;14(6):927–930. doi: [10.1111/j.1654-1103.2003.tb02228.x](https://doi.org/10.1111/j.1654-1103.2003.tb02228.x).
76. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Softw*. 2017;82(13):1–26. doi: [10.18637/jss.v082.i13](https://doi.org/10.18637/jss.v082.i13).
77. Russel J, Thorsen J, Brejnrod AD. Datest: a framework for choosing differential abundance or expression method. *bioRxiv*. 2018; 241802.
78. Love MI, Huber W, Anders S. Moderated estimation of Fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. 2014;15(12):550. doi: [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8).
79. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–780. doi: [10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010).
80. Yu GC, Smith DK, Zhu HC, Guan Y, Lam TTY. GGTREE: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol*. 2017;8(1):28–36. doi: [10.1111/2041-210X.12628](https://doi.org/10.1111/2041-210X.12628).
81. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform*. 2010;11(1):1–11. doi: [10.1186/1471-2105-11-119](https://doi.org/10.1186/1471-2105-11-119).
82. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb)*. 2021;2(3):100141. doi: [10.1016/j.xinn.2021.100141](https://doi.org/10.1016/j.xinn.2021.100141).
83. Hadley W. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.