medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Lifestyle associates with unique resistome and microbiome signatures in children

Hendrik Dühr[1,2], https://orcid.org/0009-0007-6197-3235, hendrik.dhr@ki.se, 1st corresponding author

Katariina Pärnänen[3,4], https://orcid.org/0000-0003-2220-1738

Nina Kucháriková[1], https://orcid.org/0000-0002-1921-774X

Paulina Werner[1], https://orcid.org/0000-0002-8112-0252

Göran Pershagen[1], https://orcid.org/0000-0002-9701-1130

Leo Lahti[3], https://orcid.org/0000-0001-5537-637X

Harri Alenius[1,5], https://orcid.org/0000-0003-0106-8923

Anna Bergström[1,6] , https://orcid.org/0000-0002-7981-6314

Matti O. Ruuskanen[3,7], https://orcid.org/0000-0003-4221-2880

Nanna Fyhrquist[1,2], https://orcid.org/0000-0002-5408-0005, nanna.fyhrquist@ki.se, 2nd corresponding author

[1]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

[2]Department of Health Sciences, Karlstad University, Karlstad, Sweden

[3]Department of Computing, University of Turku, Turku, Finland

[4]Department of Microbiology, University of Helsinki, Helsinki, Finland

[5]Human Microbiome Research Program (HUMI), University of Helsinki, Helsinki, Finland

[6]Centre for Occupational and Environmental Medicine, Region Stockholm, Stockholm, Sweden

[7]Department of Life Technologies, University of Turku, Turku, Finland

Abstract

Background

Antibiotic resistance is a global health crisis that is not solely explained by antibiotics usage. However, environmental and lifestyle contributions to antimicrobial resistance (AMR) in children are not well understood, especially compared to adults. As the gut functions as a reservoir for antibiotic resistance genes (ARGs), the aim of this study was to better understand the influence of lifestyle on the gut microbiome and resistome using shotgun-metagenomic sequencing data of Swedish children from the PARSIFAL (Prevention of Allergy Risk factors for Sensitization In children related to Farming and Anthroposophic Lifestyle) study.

Results

Farm children exhibited high proportions of unique bacterial species and differentially abundant ARGs linked to the farm environment, and similar differences were found in anthroposophic children. Age, breastfeeding duration, and obesity significantly influenced the overall resistance load, independently of lifestyle. Despite limited statistical power, our findings suggest that lifestyle and environment both shape the microbiome and resistome of children.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Conclusions

This study corroborates the possible influence of the farm environment on the gut microbiome and resistome, revealing a highly individualized repertoire of low-abundance microbes and ARGs in farm children. Additionally, associations of age, obesity and the duration of exclusive breastfeeding with ARG load were found in a currently understudied age range. Overall, this study raises the need for further research on rare species and ARGs as well as their transmission dynamics in relation to the environment.

Background

In recent decades antibiotic resistance has become a major global problem affecting human health, currently being estimated to directly cause more than 1 million deaths per year(1), as well as economic costs in the order of billions of dollars, with both numbers expected to rise drastically in the coming decades(2). Especially young children could suffer from potential declines in drug efficacy due to antimicrobial resistance (AMR), as antibiotics currently provide an effective means to prevent child death (3). In this context, the human gut microbiome, i.e., the collection of microbes in the human gut(4), has been increasingly studied as it harbors many antibiotic resistance genes (ARGs), whose entirety make up the gut resistome(5).

Bacteria acquire resistance either through mutations or by horizontal gene transfer (HGT)(6), often mediated by mobile genetic elements (MGEs) such as plasmids, transposons and integrons(7). HGT may facilitate the spread of ARGs between commensal and pathogenic bacteria, making the human gut microbiome an important reservoir for antimicrobial resistance(8).

Even though ARGs occur naturally in the environment and can be traced back to prehistoric ages(9), it has been shown that human activity significantly increases the number of ARGs found in different environments, including the human gut and skin, with 23.78 % of the detected ARGs in all environments being considered a health risk(10). Although antibiotic consumption is a major contributing factor, it does not fully account for the global spread of AMR(11). Furthermore, the transmission dynamics of ARGs between animals, humans and environments are far from being quantified and remain poorly understood(12).

Environmental exposures and lifestyle factors are known to strongly influence the gut microbiome composition(13). Yet most studies on childhood microbiomes and resistomes focus on early infancy(14), leaving a knowledge gap with regards to school-aged children. Lifestyles such as the anthroposophic and farm lifestyle are of particular interest due to distinct characteristics and exposures. The anthroposophic lifestyle is characterized by limited use of antibiotics, antipyretics, and vaccinations, alongside a diet rich in fermented vegetables(15). In contrast, the farm lifestyle has been shown to be protective against allergies, which is potentially caused by a high exposure to microorganisms inhabiting the farm environment, leading to high microbial diversity in the gut(16). However, the exact mechanisms behind this protective effect remain unclear, though current studies suggest a complex interplay between the bacteria and the host

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

immune system(17). In this context, the PARSIFAL study population has been used to investigate protective lifestyle factors for the occurrence of allergic diseases, with results showing that both anthroposophic and farm lifestyles were protective against the occurrence of allergies in children(18,19).

Given AMR's global threat and the gut microbiome's role as an ARG reservoir, understanding how childhood lifestyle shapes the resistome is essential. We analyzed metagenomic data from the Swedish participants of the PARSIFAL (Prevention of Allergy Risk factors for Sensitization In children related to Farming and Anthroposophic Lifestyle) study to to assess how anthroposophic and farm lifestyles associate with the gut microbiome and resistome in school-aged children. Our analyses included presence/absence, alpha and beta diversity, as well as differential abundance analyses (DAA), focusing on comparisons between lifestyle and reference groups. Furthermore, we investigated the relation of ARG load (normalized ARG abundance per sample) to lifestyle variables using a generalized linear model (GLM). Additionally, we investigated the relation between ARG load and alpha diversity indices to bacterial class abundances using correlation analysis.

Materials and methods

Sample collection

To better understand the lifestyle and environmental drivers of the microbiome and resistome we analyzed 74 stool samples from the cross-sectional PARSIFAL study(18) using paired-end shotgun metagenomic sequencing. The study population consisted of children aged 5-13 years living in the Uppsala and Järna regions (Sweden) and entailed participants living on farms, children with anthroposophic lifestyles and two reference groups of children living in the same areas but not following these lifestyles. For this study, all 74 fecal samples of the Swedish part of the PARSIFAL study were obtained. Samples were taken by the parents, first stored at -20 °C, and finally stored at -80 °C before DNA extraction, as described by Dicksved et al(20). The participant data and samples that were used for the analyses in this study were collected between October 2000 and May 2002. A questionnaire was completed by the parents and has previously been reported in detail by Alfvén et al. (18).

DNA isolation and quantification

DNA was isolated using the QIAamp® PowerFecal® Pro DNA Kit (Qiagen, Germany, cat. no. 51804) according to manufacturer's instructions, using the Tissue Lyser II (Qiagen, cat. no. 85300) for lysis. Genomic DNA was eluted in 50 µL of 10 mM Tris. Subsequent quantification and quality control was performed by Novogene (UK) Co., Ltd. using Qubit Fluorometric Quantification (Thermo Fisher, USA).

Metagenomic sequencing

Library construction was performed according to the following workflow: Genomic DNA was randomly sheared into short fragments that were then end repaired, A-tailed and ligated with Illumina adapters. Fragments with adapters were PCR amplified, size selected, and purified. Paired-end sequencing (read length = 150 base pairs) was performed using the NovaSeq 6000 system (Illumina, USA, cat. no. 20012850)

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Metagenomic analysis

Microbial community profiling of metagenomic sequence reads was performed using MetaPhlAn 4 version 4.0.3(21). We mapped the R1 and R2 reads with Bowtie2 v 2.4.4(22) against the ResFinder database version 2.1.1(23) with the following options: "-D 20 -R 3 -N 1 -L 20 -i S,1,0 5" to identify ARGs. ResFinder was chosen as a database due to its high sensitivity with regards to clinically relevant ARGs such as beta-lactamase encoding genes, compared to other databases(24). The default quality scores in Bowtie were used to ensure high-quality matches are included.

Data normalization

For taxonomic data, MetaPhlAn relative abundance (%) were transformed into counts through division by 100 and multiplication by library size. Counts were then scaled so that the lowest count > 0 equaled 1. For ARG data, counts were first normalized to gene lengths, and then to library size. As with taxonomic data, counts were scaled so that the lowest count > 0 equaled 1. For the DESeq2 analysis of ARGs, a pseudo-count of 1 was added to all counts of both types of normalized data to allow comparisons between detected and undetected species/genes.

To calculate the ARG load, the raw ARG counts were first normalized to the respective ARG gene length. Then, the sample-wise sums of gene-length normalized ARGs were calculated and normalized to library size. These values were then scaled to represent reads per kilobase of transcript per million mapped reads(25) (RPKM). Library size was chosen for data normalization as it yields similar results compared to 16S rRNA or single marker copy gene normalization(26).

Rarefaction, alpha and beta diversity

To ensure that the presence/absence analysis was not distorted by differences in groupor library size, we performed additional normalization stepsFirst, MetaPhlAn and ARG counts were normalized to group size before the scaling. Then, rarefaction as defined by Schloss(27), was performed with 100 iterations using rarefy_even_depth from phyloseq 1.50.0(28). The minimum library size was set equal to the smallest sample library size, and subsampling was performed without replacement. The rarefied count tables were then combined to control whether library size affected the presence/absence of species and ARGs. To obtain robust estimates for species alpha and beta diversity measures(27), rarefaction was performed with 100 iterations, using the rarefy_even_depth and getDissimilarity functions of the phyloseq(28) 1.50.0 and mia(29) 1.15.19 packages respectively. As before,the minimum library size was set equal to the smallest sample library size, and subsampling was performed without replacement. Alpha diversities (Shannon(30), inverse Simpson(31), Chao1(32)) were calculated using the estimate_richness function from phyloseq. Beta diversity indices (Bray-Curtis(33), Jaccard(34)) were calculated using the vegdist function (called with getDissimilarity) from the vegan(35) 2.6-8 package. ARG counts were not rarefied before alpha and beta diversity analyses.

Statistical analysis

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Associations between the participant data and the lifestyle groups were analyzed using Fisher's exact test(36) and the Kruskal-Wallis rank sum test(37), for categorical and continuous variables, respectively. The test statistics were calculated using gtsummary(38) 2.4.0, which was also used to create the final tables. Relative abundances of bacterial classes, families and ARG resistance classes were calculated using the transformAssay function (method = "relabundance") from the mia(29) 1.15.19 package. The differences in the ratio of the number of unique to shared species and ARGs between the different lifestyle groups were analyzed using a pairwise comparison of proportions implemented in pairwise.prop.test from the stats(39) 4.4.2 package. The venn diagrams were plotted using ggvenn(40) 0.1.10. To test for differentially abundant species between the lifestyle groups and their respective references both MaAsLin2 1.20.0(41) and the Wilcoxon test(42) were used, as previously suggested(43). For both methods, a 10 % prevalence filter was applied. In the Maaslin2 function, log-transformed, TSS-normalized counts were used. As no covariables except the lifestyle groups were used in our analysis, the MAasLin2 results were equal to those of a t-test. As suggested for the identification of differentially abundant genes in metagenomics(44), DESeq2(45) 1.46.0 was used to analyze differentially abundant ARGs between the lifestyle and respective reference groups, using the default settings of the DESeq function. To explore potential relations between the ARG load and the PARSIFAL study variables, a generalized linear model (GLM) was employed, and only participants with complete case records were analyzed (N = 57). The model was built using the variables listed in Table 1. The glm function of the stats(39) package was used to fit the model using the option family "gamma" (link = "log") as done previously(26). The step function of the stats(39) package was then used to fit the final model, using backwards stepwise regression. Variables were dropped based on the Akaike information criterion. The glht function of the multcomp(46) 1.4-26 was used to perform Tukey's post hoc test to obtain adjusted p-values for pairwise comparisons. Since our goal was to model ARG load rather than to test associations for specific exposures, we did not define covariates *a priori* but instead started from a full model including all predictors and applied stepwise reduction to obtain the final model. To test potential interplay between bacterial class abundance and ARG load, Shannon(30), inverse Simpson(31) and Chao1(32) indices, correlation analysis was performed. Therefore, classes with less than 50 % prevalence were aggregated to "Other". Then, transformAssay of the mia(29) package (method = "clr") was used to CLR transform the counts after adding a pseudocount (half of the minimum positive value of the class counts; 1.5 for bacterial classes, 0.5 for ARG classes). Kendall's rank correlation was then performed using getCrossAssociation from the mia(29) package. Differences in relative abundances of bacterial classes between the lifestyle and reference groups were calculated and graphed using metacoder(47) 0.3.8. To test differences between lifestyle and reference groups regarding alpha diversities, the Mann-Whitney U Test(48) was performed, as alpha diversities were not normally distributed, which was tested using the Shapiro-Wilk test(49). Principal coordinates analysis(50) (PCoA) was performed using the cmdscale function in the stats(39) package. Unless otherwise stated, p-values were corrected for

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

multiple testing using the Benjamini–Hochberg procedure(51). Results were considered significant for $p \leq 0.05$. Plots were created using ggplot2 3.5.1(52). All analyses were performed in R 4.4.2(39).

Results

Study population

The PARSIFAL cross-sectional study has previously been described in detail by Alfvén et al.(18). The Swedish portion of the study population was recruited in Uppsala and Järna (Sweden) and consists of four distinct lifestyle groups: 1. A farm group of children living on farms (N = 37). 2. A farm reference group, consisting of children living in the farmers' area, but not growing up on farms (N = 11). 3. A group of anthroposophs (N = 19; recruited at Steiner schools and called 'Steiner' throughout the text). 4. A Steiner reference group, consisting of children living in the same area but not following the anthroposophic lifestyle (N = 7). There were statistical differences across the lifestyle groups regarding the parental education level, household pet ownership during the first year of life, the use of antipyretics, previous measle infections, the consumption of organic and biodynamic food, and child age ($p \leq 0.05$ after correction).

Table 1 | Data for the PARSIFAL study population by lifestyle group. The BMI classification as defined by Cole et al.(53) was used. BMI: Body mass index. Farm: Farm group, Farm ref: Farm reference group, Steiner: Steiner group, Steiner ref: Steiner reference group.

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Variable	Overall N = 74[1]	Lifestyle group: Farm N = 37[1]	Farm ref N = 11[1]	Steiner N = 19[1]	Steiner ref N = 7[1]	p-value	q-value[2]
Sex (male)	42 [57%]	22 [59%]	7 [64%]	9 [47%]	4 [57%]	0.81[3]	0.81
Maternal smoking during pregnancy	7 [9.5%]	5 [14%]	0 [0%]	1 [5.3%]	1 [14%]	0.52[3]	0.62
Parental education						<0.001[3]	<0.001
Elementary school or lower	6 [8.3%]	5 [14%]	1 [9.1%]	0 [0%]	0 [0%]		
Gymnasium	34 [47%]	22 [63%]	8 [73%]	3 [16%]	1 [14%]		
University	32 [44%]	8 [23%]	2 [18%]	16 [84%]	6 [86%]		
Unknown	2	2	0	0	0		
Number of older siblings						0.45[3]	0.62
0	18 [24%]	11 [30%]	3 [27%]	3 [16%]	1 [14%]		
1	28 [38%]	12 [32%]	6 [55%]	6 [32%]	4 [57%]		
2	14 [19%]	6 [16%]	2 [18%]	4 [21%]	2 [29%]		
≥3	14 [19%]	8 [22%]	0 [0%]	6 [32%]	0 [0%]		
Environmental smoking at home	9 [12%]	4 [11%]	2 [18%]	3 [16%]	0 [0%]	0.75[3]	0.81
Housholds pets during the first year of life	34 [46%]	26 [70%]	2 [18%]	4 [21%]	2 [29%]	<0.001[3]	<0.001
Use of antibiotics						0.064[3]	0.12
Never use	8 [12%]	1 [2.9%]	1 [11%]	6 [32%]	0 [0%]		
First use >12 months of life	30 [43%]	18 [53%]	3 [33%]	7 [37%]	2 [29%]		
First use 0-12 months of life	31 [45%]	15 [44%]	5 [56%]	6 [32%]	5 [71%]		
Unknown	5	3	2	0	0		
Use of antipyretics						<0.001[3]	<0.001
Never use	11 [16%]	0 [0%]	0 [0%]	10 [53%]	1 [14%]		
First use >12 months of life	24 [35%]	18 [55%]	3 [30%]	1 [5.3%]	2 [29%]		
First use 0-12 months of life	34 [49%]	15 [45%]	7 [70%]	8 [42%]	4 [57%]		
Unknown	5	4	1	0	0		
Child had measles	7 [10%]	0 [0%]	0 [0%]	7 [37%]	0 [0%]	<0.001[3]	<0.001
Unknown	6	5	1	0	0		
Consumes organic and biodynamic food	20 [27%]	4 [11%]	1 [9.1%]	15 [79%]	0 [0%]	<0.001[3]	<0.001
Unknown	1	1	0	0	0		
Age (years)	8.69 (1.89)	9.49 (1.85)	8.73 (1.79)	7.63 (1.46)	7.29 (1.11)	<0.001[4]	0.002
Length of (exclusive) breastfeeding						0.27[3]	0.44
Never	5 [6.8%]	3 [8.1%]	1 [9.1%]	1 [5.3%]	0 [0%]		
0-4 months	37 [50%]	23 [62%]	5 [45%]	6 [32%]	3 [43%]		
≥5 months	32 [43%]	11 [30%]	5 [45%]	12 [63%]	4 [57%]		
BMI classification						0.50[3]	0.62
Not overweight	56 [78%]	23 [66%]	10 [91%]	17 [89%]	6 [86%]		
Overweight	13 [18%]	9 [26%]	1 [9.1%]	2 [11%]	1 [14%]		
Obese	3 [4.2%]	3 [8.6%]	0 [0%]	0 [0%]	0 [0%]		
Unknown	2	2	0	0	0		

[1] n [%]; Mean (SD)

[2] Benjamini & Hochberg correction for multiple testing

[3] Fisher's exact test

[4] Kruskal-Wallis rank sum test

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Community composition

Overall, differences in relative abundances of the microbial taxa and ARGs between the lifestyle groups and references were comparatively small (Fig. 1, Supplementary Fig. 1). There was a high number of low-abundance, non-characterized bacterial classes, especially within Firmicutes (Supplementary Figure 1). Most of the bacterial classes (Fig. 1a) belonged to the class *Clostridia*, followed by *Actinomycetia* and *Bacteroidia*. On the family level (Fig. 1b), most species belonged to three families (*Lachnospiraceae*, *Bifidobacteriaceae* and *Oscillospiraceae*). A total of 8 bacterial genera made up around 50 % of the total genera in the study group, with *Bifidobacterium* being the most abundant (Fig. 1c). The overwhelming majority of ARGs were tetracycline resistance genes (Fig. 1d). In addition to tetracycline resistance, resistance to four other classes of antibiotics were common (aminoglycosides, amphenicols, beta-lactams, and macrolide–lincosamide–streptogramin B; Fig. 1d).

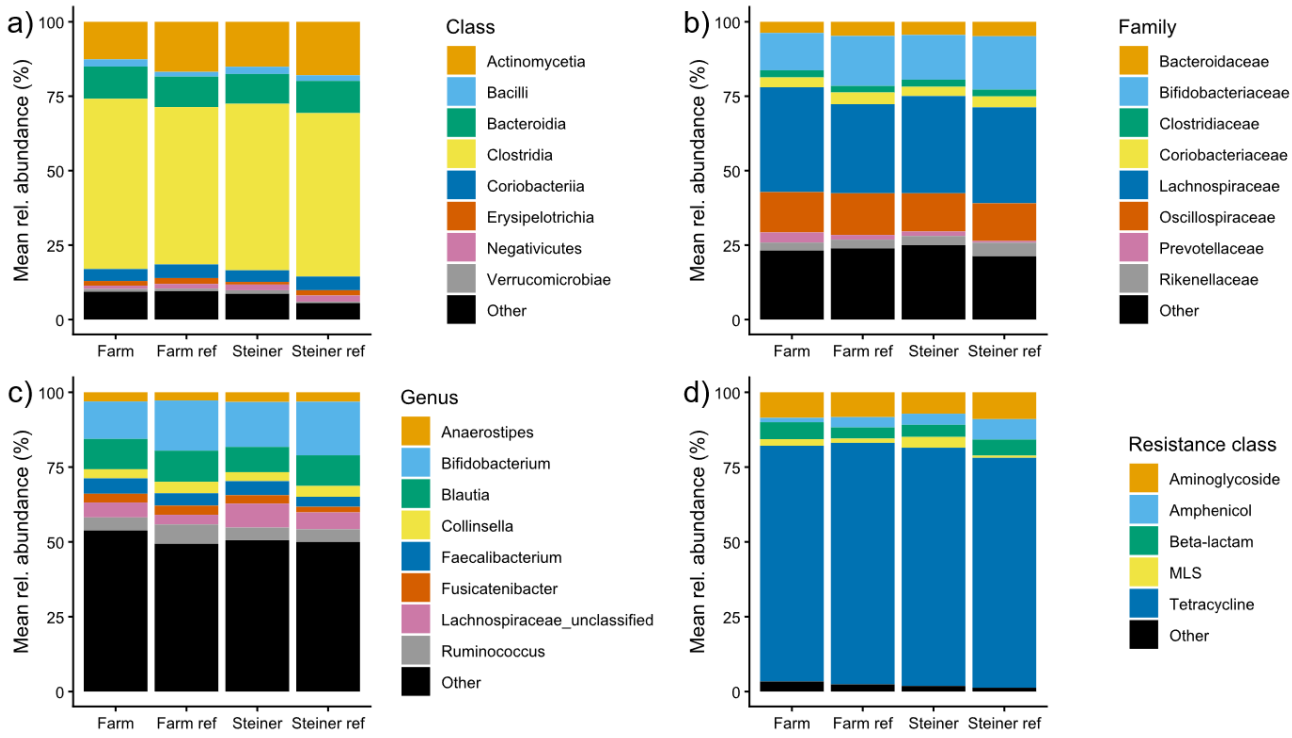

Fig. 1 | Microbiome and resistome compositions in the gut of the lifestyle groups. Mean relative abundances (%) are shown on the y-axis. **a** Most abundant bacterial classes. **b** Most abundant bacterial families. **c** Most abundant bacterial genera. **d** Most abundant antibiotic resistance classes. Farm: Farm group (N = 37), Farm ref: Farm reference group (N = 11), Steiner: Steiner group (N = 19), Steiner ref: Steiner reference group (N = 7). MLS: Macrolide, Lincosamide, Streptogramin B.

Presence/absence analysis

Venn diagrams depicting shared species and ARGs between the lifestyle groups indicated a shared core microbiome and resistome, i.e., shared species and ARGs across all lifestyle groups (Figure 2). Each group had several unique species and ARGs, with the farm and farm reference group containing the highest number of unique species and ARGs. More detailed information regarding the composition of unique species and ARGs is displayed in the supplementary data (Supplementary Figure 2, 3). Additionally,

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

complete lists of all species and ARGs and information on their mean prevalence and relative abundance within the lifestyle groups are provided in Supplementary data 1 and 2.The ratios of unique to shared species were significantly different ($p \leq 0.05$) between farmers and all other lifestyle groups, as well as between Steiner children and both reference groups (Fig. 2b). The ratio of unique/shared ARGs were significantly different ($p \leq 0.05$) between the farm group and all other lifestyle groups, though other groups did not significantly differ from each other (Fig. 2d, $p > 0.05$). The lifestyle exclusive species and ARGs were present mostly in single individuals, as shown in the beta diversity analysis, where no group-dependent groupings were observed (Supplementary Figure 4). Furthermore, more than 90 % of these group-exclusive species and ARGs were of low relative abundance (< 1%) within a given individual. Additionally, we compared the lifestyle and reference groups using alpha diversity indices of species and ARGs but found no significant differences (Supplementary Figure 5, $p > 0.05$).

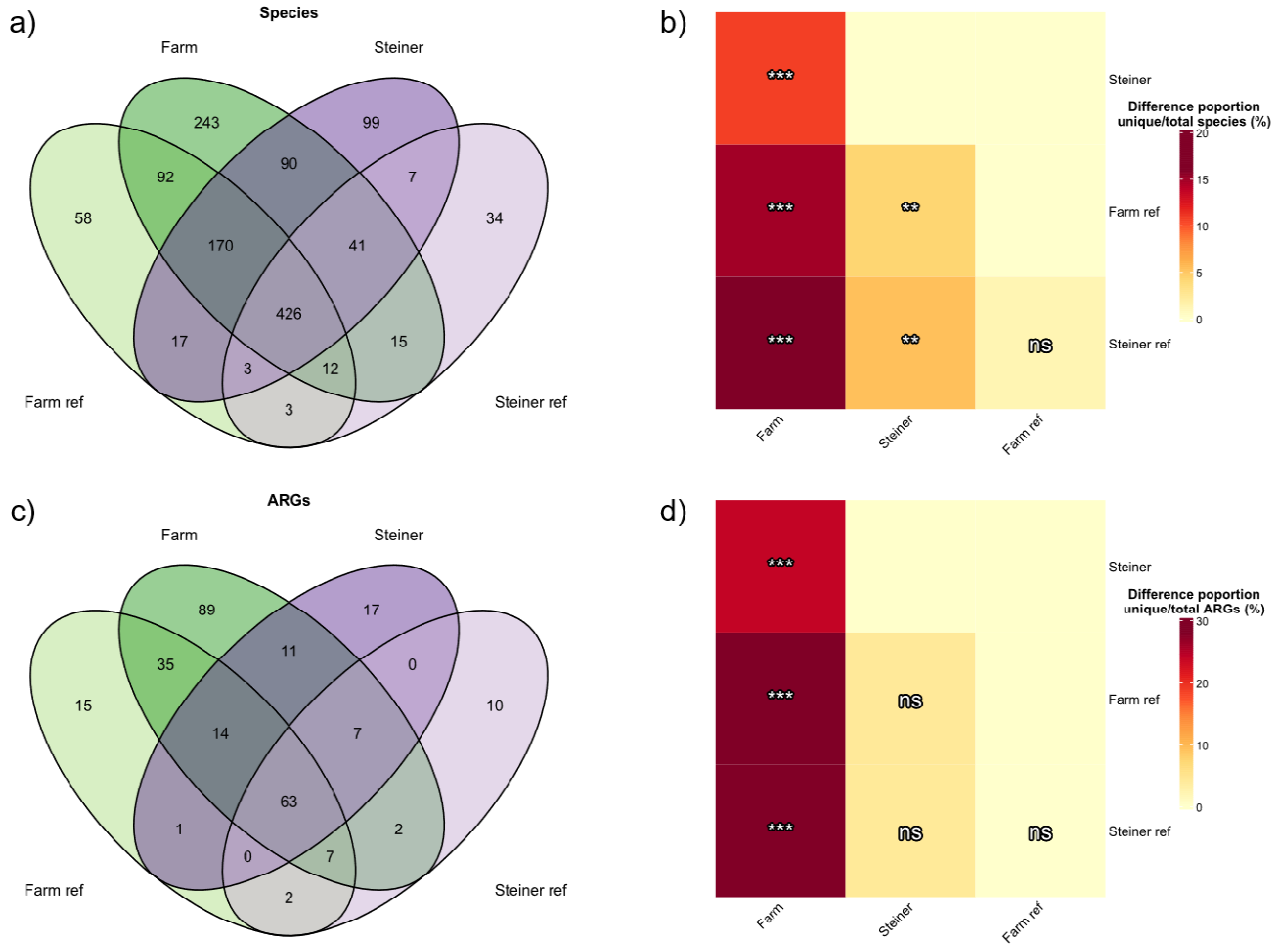

Fig. 2 | Venn diagrams and pairwise comparisons of proportions. Venn diagrams depict the shared species (N = 1310) and ARGs (N = 273) between lifestyle and reference groups. **a** presence/absence of species in a lifestyle group, **b** pairwise comparison of proportions of the ratios of unique to non-unique species, **c** presence/absence of ARGs in a lifestyle group, **d** pairwise comparison of proportions of the ratios of unique to non-unique ARGs. Significance of pairwise comparison of proportions is denoted as follows; ns: $p > 0.05$; *$p \leq 0.05$ **$p \leq 0.01$; ***$p \leq 0.001$. Farm: Farm group (N = 37), Farm ref: Farm reference group (N = 11), Steiner: Steiner group (N = 19), Steiner ref: Steiner reference group (N = 7).

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Differential abundance analyses

To test if any species or ARGs associated with a lifestyle group compared to its reference, differential abundance analysis was performed using MaAsLin2 and the Wilcoxon test for species, and DESeq2 for ARGs. Only one species, *Clostridium sp AM49 4BH* was found to be differentially abundant in farmers compared to the reference group using MaAsLin2 ($p \leq 0.05$) but was not found with the Wilcoxon test ($p > 0.05$). No differential species were detected between Steiner children and their references with either tool ($p > 0.05$).

Between farmers and references, 19 ARGs were found to be differentially abundant (Fig. 3, $p \leq 0.05$), representing 10 of the 22 resistance classes that were detected in the whole study population. Of these, 15 ARGs were differentially abundant in farm children, four in the reference group. *aph(3'')-Ib* was found to be differentially abundant in Steiner references compared to Steiner children (log2 fold = -6.58, $p \leq 0.05$).

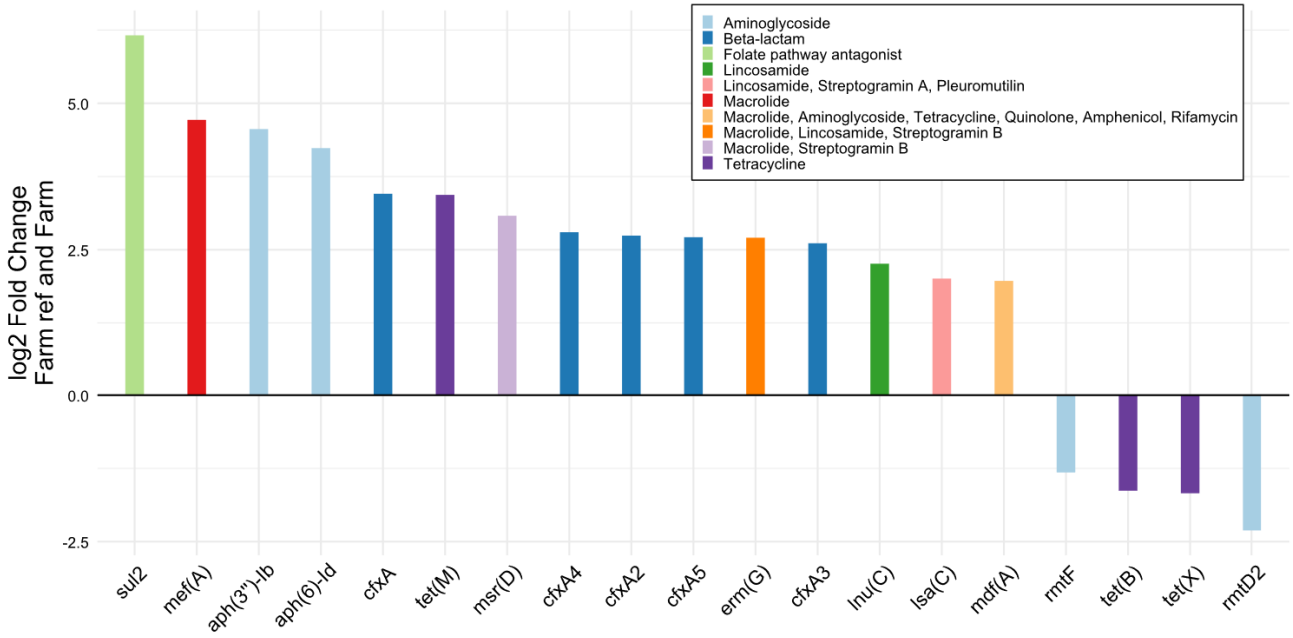

Fig. 3 | Differentially abundant ARGs and their corresponding resistance classes between farm children and their reference group. All genes shown were found to be significant in DESeq2 ($p \leq 0.05$). The y-axis shows log2 fold changes, the x-axis denotes gene names, positive log2 fold changes indicate differential abundance in the farm group, negative changes indicate differential abundance in the farm reference group. Farm: Farm group (N = 37), Farm ref: Farm reference group (N = 11), Steiner: Steiner group (N = 19), Steiner ref: Steiner reference group (N = 7).

ARG load

To study the associations of the study variables with the ARG load, a GLM was fitted using stepwise regression, using only complete case records for the variables listed in Table 1 (N = 57, see Supplementary Table 1 for a full description). The body mass index (BMI) category (Fig. 4a), length of exclusive breastfeeding (Fig. 4b) and age (Fig. 4c) were significantly associated with ARG load ($p \leq 0.05$). Overweight children had a significantly lower ARG load than non-overweight children (fold change: 0.75, 95 % Confidence interval [CI]: 0.60-0.94). Children breastfed exclusively for more than 5 months had a

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

significantly lower ARG load than children breastfed for shorter periods (fold change: 0.83, 95 % CI: 0.70-0.98). Increasing age (in years) was significantly associated with a lower ARG load (fold change: 0.92, 95 % CI: 0.88-0.97).

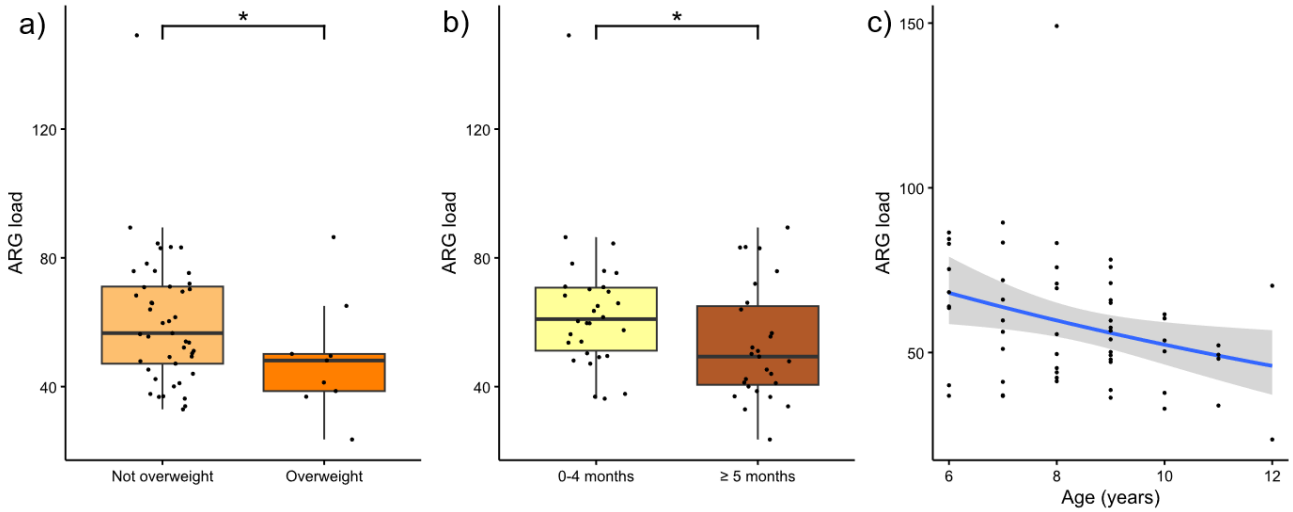

Fig. 4 | ARG load in relation to PARSIFAL study variables. To investigate the effects of variables on the ARG load a gamma-distributed GLM model was fitted using backward stepwise-regression. ARG load (N = 57) by: **a** BMI (not overweight and overweight), **b** length of breastfeeding (non-breastfed category not shown) and **c** age ($p \leq 0.05$). The y axis denotes ARG load (ARG counts normalized to gene lengths and library size, scaled to RPKM). Significance levels are denoted as follows (Tukey's post-hoc test): *p ≤ 0.05. Boxplot hinges represent 25 % and 75 % percentiles; the center line represents the median. The length of whiskers denotes 1.5 × IQR. The x-axis in c represents age in years. A regression line was fitted using gamma-distributed GLMs. Not overweight: Non-overweight children (N = 45), Overweight: Overweight children(53). 0-4 months: exclusive breastfeeding for less than 5 months (N = 30), ≥ 5 months: exclusive breastfeeding for 5 or more months (N = 27).

Correlation analysis

We investigated associations between ARG load and alpha diversity with bacterial and ARG class abundance using correlation analysis, as both ARG load and alpha diversity have been linked to human health. The abundances were center log ratio (CLR) transformed to remove compositionality bias. The ARG load and alpha diversity were not significantly correlated with the CLR transformed abundances of ARG classes (all p > 0.05; Supplementary Figure 6). Conversely, we found significant correlations between ARG load and alpha diversity indices and the CLR transformed abundances of bacterial classes (Fig. 5). Overall, the different alpha diversity indices showed similar correlation patterns regardless of being based on relative abundance of species/ARGs (Shannon, inverse Simpson), or estimating the total number of species based on presence/absence data and rare taxa (Chao1). Several unclassified bacterial classes correlated positively with alpha diversity (p ≤ 0.05); *Actinomycetia*, *Clostridia*, *Erysipelotrichia*, and *Coriobacteria* were strongly negatively correlated with all alpha diversity indices. In contrast, *Verrucomicrobiae*, *Eryisipelotrichia* and *Actinomycetia* were the only classes to significantly, though weakly, correlate with the ARG load, with the latter two correlating negatively and *Verrucomicrobiae* correlating positively with the ARG load (Fig. 5).

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

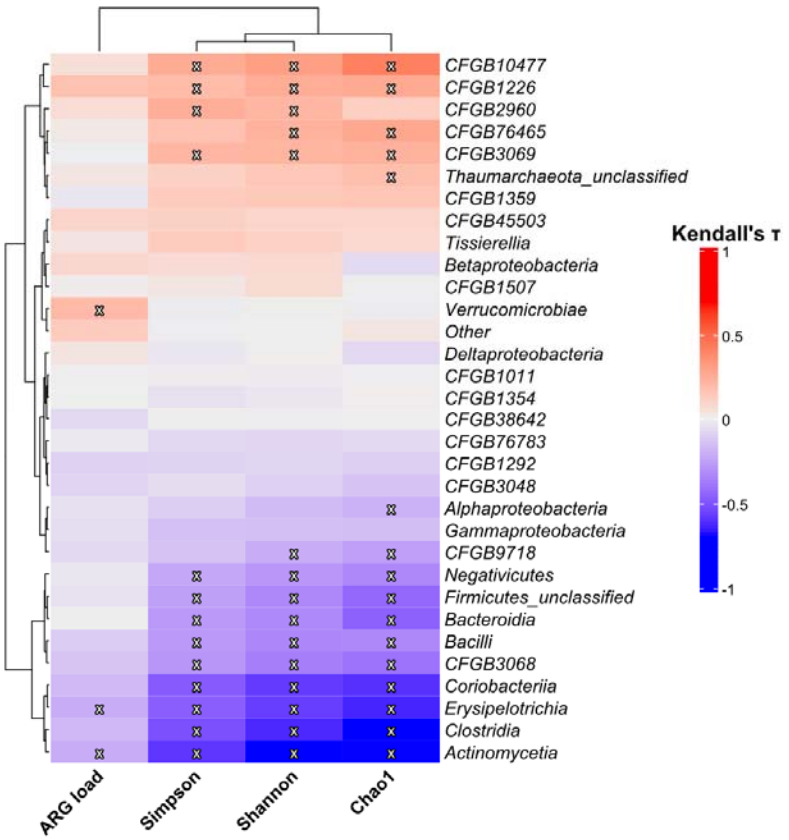

Fig. 5 | Correlation of bacterial taxa with ARG load and inverse Simpson, Shannon and Chao1 diversities. CLR transformed bacterial class abundances were correlated with diversity measures using the Kendall rank correlation coefficient (τ). Significant correlations (BH corrected $p \leq 0.05$) are denoted with an x. Dendrograms are the result of row- and column-wise hierarchical clustering (complete-linkage) using Euclidian distances. Classes with a prevalence < 50% were aggregated to "Other". Bacterial class names according to Genome Taxonomy Database (GTDB)(54).

Discussion

Our study revealed differences in species and ARG presence and absence between the groups, and identified differentially abundant ARGs in the farm group. Furthermore, we were able to identify and corroborate specific lifestyle factors that were associated with variation in ARG load in school-aged children. Nonetheless, these results should be read carefully in the light of the comparatively small sample size and unequal group sizes.

The presence/absence analysis showed that differences in the gut microbiome of both farm and anthroposophic children were primarily driven by the presence of rare, low-abundance species, with both farm and Steiner children having significantly higher proportions of group-exclusive species compared to the reference groups. Interestingly, farm children also had a significantly higher proportion of group-exclusive ARGs compared to all other lifestyle groups. As shown by the non-existent grouping in our beta diversity analysis, which included the presence/absence based Jaccard index, these group-exclusive species and ARGs are highly individual. These findings align with previous evidence of a strong impact of the farm environment on the human gut microbiome and resistome(55), supported by the presence of multiple farm-

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

exclusive species that were strongly associated with farm animals, such as *Mogibacterium kristiansenii*, *Baileyella intestinalis*(56) (both pigs), *Candidatus Neoanaerotignum tabaqchaliae*(57) (chicken) and *Clostridium cuniculi*(58) (rabbits). In anthroposophic children group-exclusive species included bacteria associated with fermentation of foods (*Leuconostoc citreum*(59), *Lactobacillus farciminis*(59)) and/or are lactic acid bacteria (*Pediococcus pentosaceus*(60), *Ligilactobacillus salivarius*(61)). Lactic acid bacteria have long been discussed to have a variety of beneficial health effects(62), including on the immune system, and may be present in this lifestyle group due to their higher consumption of fermented foods. However, additional studies are needed to determine the persistence of these species. Previous studies focused on the association between the anthroposophic lifestyle and the microbiome identified certain factors (e.g., low antibiotic use, opting for home birth) related to this lifestyle that might affect the gut microbial composition in infants and children(15,63). However, these studies cannot be directly compared to our study due to differences in age (subjects below 2 years of age), methodology and study design. Given the small sample sizes in the Steiner and respective reference groups in this study, differences associated with the anthroposophic lifestyle that were found in the aforementioned studies, e.g., higher abundances of *Bifidobacterium* and *Lactobacilli* might become discernible with increased sample sizes.

Overall, relative abundances of bacterial classes and families observed in the PARSIFAL study closely resemble those found in adult fecal microbiomes(64). While the microbial composition on class and family levels indicated slight variations at higher taxonomic levels between the groups, these differences were non-significant and could not be found with differential abundance analysis. Furthermore, our analysis of alpha diversities did not confirm the differences in alpha diversity that were previously detected in the PARSIFAL study population(20). The discrepancies between the original study and our study may be due to differences in methodology, as our metagenomic analysis was untargeted compared to the targeted T-RLFP method employed in the original study. Furthermore, it should be mentioned that the original 2007 study used a larger subsample from several European countries, which may influence comparisons. The whole PARSIFAL study group will need to be analyzed to resolve the different findings of the studies.

Our study lacks the statistical power to draw definitive conclusions about the protective effects of farm and anthroposophic lifestyles on respiratory allergies which have previously been observed in the original study(18). Nonetheless, our findings are in line with the biodiversity hypothesis(65), particularly given the high number of unique species in the farm group. Considering the well-established role of the (gut) microbiome in shaping immune function(66,67) the presence of a high number of group-exclusive taxa could merit further investigation. Expanding the analysis to include the full study group and/or future studies might help clarify the associations between the environmental exposure to microbes, the gut microbiome and protection against allergic diseases in humans(68).Given the strong protective influence of soil(69) and dust(70) exposure on the microbiome and immune function, the inclusion environmental measurements will be essential in future research.

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

In contrast to our presence/absence analysis of bacterial species, we only found one species, *Clostridium sp AM49 4BH*, to be differentially abundant in farm children compared to their references. This species has not been associated with the farm lifestyle yet, but has been associated with various diseases, including IBD(71), kidney stone formation(72) and postural orthostatic tachycardia syndrome(73). This result is not surprising, as we employed the commonly used prevalence filter cutoff of 10 %, which removes rare and often low-abundance taxa from DAA analyses(74). However, focusing solely on DAA to analyze differences between groups omits differences regarding rare species that might be part of the core microbiome(75). Though environment specific origins of these rare taxa are likely, we are unable to infer from our data whether the group-specific species are persistent or not, raising the need for studies that include multiple points of sampling.

The differentially abundant ARGs we identified in the farm group compared to the reference group contain numerous ARGs that have been identified in the farm environment, including *sul2*, *mef(A)*, *tetM* and *lnu(C)*, whilst others such as *aph(3'')-IIb* or *msr(D)* have been identified in farm-related seeding sludges (76). This may indicate the farm environment as a potential risk factor for the acquisition of antibiotic resistance genes. The transmission of species that carry ARGs could occur even airborne, as Bai et al.(77) found that this transmission route could be relevant even at a distance from the actual farm environment and thus may not require direct contact with ARG carriers such as farm animals. However, our study cannot quantify the specific contribution of the farm environment to ARG acquisition, as it lacks environmental samples.

In this study, higher age was associated with decreased ARG loads. Increased age has previously been associated with a reduced ARG load in the human gut, demonstrated in a comparative study between infants (age 11-24 months) and young adults (age 17-21)(78). Furthermore, the results of the current study could indicate that extended and exclusive breastfeeding may be beneficial in reducing the overall ARG load, though longitudinal studies are needed to explore this association. In general, breastfeeding has been identified as a factor that lowers the ARG load in infants compared to formula-based diets(26). However, it remains unclear to what extent this effect is due to a reduced exposure to antibiotic resistant bacteria in dietary components other than breastmilk, or due to molecular effects driven by breastmilk consumption(79). Additionally, our study found a lower ARG load to be associated with overweight children compared to non-overweight children. This result aligns with observations in a large cohort from Finland(80) that found an inverse association of BMI and ARG load in adults and pointed to the potential lack of ARG containing foods (fresh vegetables and poultry) in high-fat and high-processed foods and diets. However, these findings are purely observational, and due to the lack of detailed dietary data, our study is unable to explore such associations further. Nonetheless, given the association of high phylogenetic diversity of foods with a low ARG load in adults(81) further investigations are warranted, especially since a phylogenetically diverse diet is generally associated with a healthy lifestyle(82).

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Verrucomicrobiae was the only bacterial class whose abundance was positively correlated with the ARG load. This may be due to a high competitiveness of the *Verrucomicrobia* phylum after antibiotic treatment(83), indicating a potential competitive advantage of this class in high ARG load environments. However, this may also be due to the fact that the host range of ARGs is taxonomically limited(84). The negative correlations between ARG load and bacterial classes (*Erysipelotrichia* and *Actinomycetia*) were few and comparatively weak in the current study. To the best of our knowledge, literature on these associations is currently sparse.

A lowered abundance of *Erysipelotrichia* after antibiotic treatment has been reported in gnotobiotic mice(85), contrasting with our findings in humans. The negative correlation between *Actinomycetia* abundance and ARG load is unexpected, given its known production of antimicrobial compounds(86). However, our analysis cannot determine causality. The many significant correlations between bacterial class abundances and alpha diversity compared to ARG load suggest a decoupling of species diversity and ARG load. However, further analyses with larger study groups are needed to identify the bacterial drivers of ARG load in children as demonstrated previously in infants (87).

Many current studies examining the microbiome and resistome in the gut of children focus on the first few weeks to months after birth, considering specific factors such as nutrition and antibiotic treatment that enable well-defined study groups(26,88). Additionally, not all these studies take both the microbiome and resistome into account in their analyses, and only cover a maximum age of 5 years(14). In contrast, this study examined both the microbiome and resistome in distinct lifestyle groups (farmers, anthroposophs) with an age range (5-13) that is not adequately represented in the current literature. Nonetheless, larger sample sizes and longitudinal studies that focus on the farm environment are needed to explore connections between the environment, microbiome and resistome, as our study lacks statistical power and a balanced design with similar group sizes.

Conclusions

In conclusion, high-abundance taxa and ARGs were largely similar across groups, while the farm environment appears to influence the gut microbiome and resistome primarily through low-abundance, highly individualized species and ARGs. We also observed associations between lifestyle variables (age, obesity, length of exclusive breastfeeding) and ARG load in a currently understudied age range. These findings highlight the potential role of environmental and lifestyle factors in shaping the resistome and support further research focusing on rare taxa, ARGs and their transmission dynamics.

Declarations

Ethics approval and consent to participate

This study was approved by Karolinska Institutet's regional research ethics committee. Informed consent was obtained from the parents of each child for the questionnaire and the stool sample (Dnr 02-475 and 00-140).

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Consent for publication

Not applicable.

Availability of data and materials

The datasets analyzed in this study are not publicly available due to the private and sensitive nature of the research data, which could compromise participant privacy. However, the data can be made available by the corresponding authors upon reasonable and qualified request.

Availability of computer code and software

The analysis code for this study is available from the GitHub repository Duehr_et_al_2025_supplementary_code and can be accessed via the following link: https://doi.org/10.5281/zenodo.15534317.

Competing Interests

All authors declare no financial or non-financial competing interests.

Funding

This project was funded by the Institute of Environmental Medicine (Karolinska Institutet) strategic grants for pilot collaboration projects. The PARSIFAL study was supported by a research grant from the European Union (QLRT 1999 01391). M.O.R was supported by grants from the Research Council of Finland (No. 338818) and the Finnish Cultural foundation.

Authors' contributions

H.D.: Data analysis, Visualization, Writing – Original Draft, Review, Editing. K.P. & M.O.R.: Conceptualization, Supervision, Metagenomic data analysis, Writing – Review, Editing. N.K. & P.W.: DNA isolation, Writing – Review, Editing. G.P. & A.B.: Cohort data curation, Resources, Writing – Review, Editing. L.L. & H.A.: Resources, Funding, Writing – Review, Editing. N.F.: Conceptualization, Supervision, Resources, Funding, Writing – Review, Editing. All authors approved the final version.

Acknowledgments

We wish to acknowledge CSC – IT Center for Science, Finland, for computational resources. We thank the children and parents for participating in the PARSIFAL study. We also thank all the staff involved in the study.

References

1. Naghavi M, Vollset SE, Ikuta KS, Swetschinski LR, Gray AP, Wool EE, et al. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. The Lancet. 2024 Sep 28;404(10459):1199–226.

2. Aslam B, Wang W, Arshad MI, Khurshid M, Muzammil S, Rasool MH, et al. Antibiotic resistance: a rundown of a global crisis. Infection and Drug Resistance. 2018 Oct 10;11:1645–58.

3. Laxminarayan R, Matsoso P, Pant S, Brower C, Røttingen JA, Klugman K, et al. Access to effective antimicrobials: a worldwide challenge. The Lancet. 2016 Jan 9;387(10014):168–75.

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

4. Joos R, Boucher K, Lavelle A, Arumugam M, Blaser MJ, Claesson MJ, et al. Examining the healthy human microbiome concept. Nat Rev Microbiol. 2025 Mar;23(3):192–205.

5. van Schaik W. The human gut resistome. Philosophical Transactions of the Royal Society B: Biological Sciences. 2015 Jun 5;370(1670):20140087.

6. Huddleston JR. Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. IDR. 2014 Jun 20;7:167–76.

7. Darby EM, Trampari E, Siasat P, Gaya MS, Alav I, Webber MA, et al. Molecular mechanisms of antibiotic resistance revisited. Nat Rev Microbiol. 2023 May;21(5):280–95.

8. Brito IL. Examining horizontal gene transfer in microbial communities. Nat Rev Microbiol. 2021;19(7):442–53.

9. Aminov RI. A Brief History of the Antibiotic Era: Lessons Learned and Challenges for the Future. Front Microbiol. 2010 Dec 8;1.

10. Zhang Z, Zhang Q, Wang T, Xu N, Lu T, Hong W, et al. Assessment of global health risk of antibiotic resistance genes. Nat Commun. 2022 Mar 23;13(1):1553.

11. Collignon P, Beggs JJ, Walsh TR, Gandra S, Laxminarayan R. Anthropological and socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable analysis. The Lancet Planetary Health. 2018;2(9):e398–405.

12. Larsson DGJ, Flach CF. Antibiotic resistance in the environment. Nat Rev Microbiol. 2022;20(5):257–69.

13. Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, et al. Environment dominates over host genetics in shaping human gut microbiota. Nature. 2018 Mar;555(7695):210–5.

14. Lebeaux RM, Karalis DB, Lee J, Whitehouse HC, Madan JC, Karagas MR, et al. The association between early life antibiotic exposure and the gut resistome of young children: a systematic review. Gut Microbes. 2022;14(1):2120743.

15. Alm JS, Swartz J, Björkstén B, Engstrand L, Engström J, Kühn I, et al. An anthroposophic lifestyle and intestinal microflora in infancy. Pediatric Allergy and Immunology. 2002;13(6):402–11.

16. Ege MJ. The Hygiene Hypothesis in the Age of the Microbiome. Annals ATS. 2017 Nov;14(Supplement_5):S348–53.

17. Depner M, Taft DH, Kirjavainen PV, Kalanetra KM, Karvonen AM, Peschel S, et al. Maturation of the gut microbiome during the first year of life contributes to the protective farm effect on childhood asthma. Nat Med. 2020 Nov;26(11):1766–75.

18. Alfvén T, Braun-Fahrländer C, Brunekreef B, von Mutius E, Riedler J, Scheynius A, et al. Allergic diseases and atopic sensitization in children related to farming and anthroposophic lifestyle – the PARSIFAL study. Allergy. 2006;61(4):414–21.

19. Flöistrup H, Swartz J, Bergström A, Alm JS, Scheynius A, van Hage M, et al. Allergic disease and sensitization in Steiner school children. Journal of Allergy and Clinical Immunology. 2006 Jan 1;117(1):59–66.

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

20. Dicksved J, Flöistrup H, Bergström A, Rosenquist M, Pershagen G, Scheynius A, et al. Molecular Fingerprinting of the Fecal Microbiota of Children Raised According to Different Lifestyles. Applied and Environmental Microbiology. 2007 Apr;73(7):2284–9.

21. Blanco-Míguez A, Beghini F, Cumbo F, McIver LJ, Thompson KN, Zolfo M, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nat Biotechnol. 2023 Nov;41(11):1633–44.

22. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012 Apr;9(4):357–9.

23. Florensa AF, Kaas RS, Clausen PTLC, Aytan-Aktug D, Aarestrup FM. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. Microbial Genomics. 2022;8(1):000748.

24. Gschwind R, Ugarcina Perovic S, Weiss M, Petitjean M, Lao J, Coelho LP, et al. ResFinderFG v2.0: a database of antibiotic resistance genes obtained by functional metagenomics. Nucleic Acids Res. 2023 Jul 5;51(W1):W493–500.

25. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008;5(7):621–8.

26. Pärnänen KMM, Hultman J, Markkanen M, Satokari R, Rautava S, Lamendella R, et al. Early-life formula feeding is associated with infant gut microbiota alterations and an increased antibiotic resistance load. The American Journal of Clinical Nutrition. 2022;115(2):407–21.

27. Schloss PD. Waste not, want not: revisiting the analysis that called into question the practice of rarefaction. mSphere. 2023 Dec 6;9(1):e00355-23.

28. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE. 2013 Apr 22;8(4):e61217.

29. Ernst FGM, Shetty SA, Borman T, Lahti L. mia: Microbiome analysis [Internet]. 2025. Available from: https://microbiome.github.io/mia/

30. Shannon CE. A Mathematical Theory of Communication. Bell System Technical Journal. 1948;27(3):379–423.

31. Simpson EH. Measurement of Diversity. Nature. 1949 Apr;163(4148):688–688.

32. Chao A. Nonparametric Estimation of the Number of Classes in a Population. Scandinavian Journal of Statistics. 1984;11(4):265–70.

33. Bray JR, Curtis JT. An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecological Monographs. 1957;27(4):325–49.

34. Jaccard P. The Distribution of the Flora in the Alpine Zone. New Phytologist. 1912;11(2):37–50.

35. Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, et al. vegan: Community Ecology Package [Internet]. 2022. Available from: https://CRAN.R-project.org/package=vegan

36. Fisher RA. Statistical methods for research workers. 5th ed. Edinburgh: Oliver & Boyd; 1934. xiii + 319.

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

37. Kruskal WH, Wallis WA. Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association. 1952 Dec 1;47(260):583–621.

38. Sjoberg DD, Whiting K, Curry M, Lavery JA, Larmarange J. Reproducible Summary Tables with the gtsummary Package. The R Journal. 2021;13(1):570–80.

39. R Core Team. R: A Language and Environment for Statistical Computing. [Internet]. Vienna (Austria): R Foundation for Statistical Computing; 2024. Available from: https://www.R-project.org/

40. Yan L. ggvenn: Draw Venn Diagram by "ggplot2" [Internet]. 2023. Available from: https://CRAN.R-project.org/package=ggvenn

41. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, et al. Multivariable association discovery in population-scale meta-omics studies. PLOS Computational Biology. 2021 Nov 16;17(11):e1009442.

42. Wilcoxon F. Individual Comparisons by Ranking Methods. Biometrics Bulletin. 1945;1(6):80–3.

43. Pelto J, Auranen K, Kujala JV, Lahti L. Elementary methods provide more replicable results in microbial differential abundance analysis. Briefings in Bioinformatics. 2025 Mar 1;26(2):bbaf130.

44. Jonsson V, Österlund T, Nerman O, Kristiansson E. Statistical evaluation of methods for identification of differentially abundant genes in comparative metagenomics. BMC Genomics. 2016 Jan 25;17(1):78.

45. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014 Dec 5;15(12):550.

46. Hothorn T, Bretz F, Westfall P. Simultaneous Inference in General Parametric Models. Biometrical Journal. 2008;50(3):346–63.

47. Metacoder: An R package for visualization and manipulation of community taxonomic diversity data | PLOS Computational Biology [Internet]. [cited 2025 Apr 10]. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005404

48. Mann HB, Whitney DR. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics. 1947;18(1):50–60.

49. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika. 1965;52(3–4):591–611.

50. Gower JC. Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis. Biometrika. 1966;53(3/4):325–38.

51. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995 Jan 1;57(1):289–300.

52. Wickham H. ggplot2: Elegant Graphics for Data Analysis [Internet]. New York: Springer-Verlag; 2016. Available from: https://ggplot2.tidyverse.org

53. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ. 2000 May 6;320(7244):1240.

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

54. Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil PA, Hugenholtz P. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nucleic Acids Res. 2022 Jan 7;50(D1):D785–94.

55. Sun J, Liao XP, D'Souza AW, Boolchandani M, Li SH, Cheng K, et al. Environmental remodeling of human gut microbiota and antibiotic resistome in livestock farms. Nat Commun. 2020 Mar 18;11(1):1427.

56. Wylensek D, Hitch TCA, Riedel T, Afrizal A, Kumar N, Wortmann E, et al. A collection of bacterial isolates from the pig intestine reveals functional and taxonomic diversity. Nat Commun. 2020 Dec 15;11(1):6389.

57. Glendinning L, Stewart RD, Pallen MJ, Watson KA, Watson M. Assembly of hundreds of novel bacterial genomes from the chicken caecum. Genome Biology. 2020 Feb 12;21(1):34.

58. Djukovic A, Garcia-Garcera M, Martínez-Paredes E, Isaac S, Artacho A, Martínez J, et al. Gut colonization by a novel Clostridium species is associated with the onset of epizootic rabbit enteropathy. Veterinary Research. 2018 Dec 20;49(1):123.

59. Nam SH, Choi SH, Kang A, Kim DW, Kim RN, Kim A, et al. Genome Sequence of Lactobacillus farciminis KCTC 3681. Journal of Bacteriology. 2011 Mar 11;193(7):1790–1.

60. Jiang S, Cai L, Lv L, Li L. Pediococcus pentosaceus, a future additive or probiotic candidate. Microb Cell Fact. 2021 Feb 16;20(1):45.

61. Guerrero Sanchez M, Passot S, Campoy S, Olivares M, Fonseca F. Ligilactobacillus salivarius functionalities, applications, and manufacturing challenges. Appl Microbiol Biotechnol. 2022 Jan 1;106(1):57–80.

62. De Filippis F, Pasolli E, Ercolini D. The food-gut axis: lactic acid bacteria and their link to food, the gut microbiome and human health. FEMS Microbiology Reviews. 2020 Jul 1;44(4):454–89.

63. Hesla HM, Stenius F, Jäderlund L, Nelson R, Engstrand L, Alm J, et al. Impact of lifestyle on the gut microbiota of healthy infants and their mothers – the ALADDIN birth cohort. FEMS Microbiology Ecology. 2014 Dec 1;90(3):791–801.

64. King CH, Desai H, Sylvetsky AC, LoTempio J, Ayanyan S, Carrie J, et al. Baseline human gut microbiota profile in healthy people and standard reporting template. PLOS ONE. 2019 Nov 9;14(9):e0206484.

65. Haahtela T. A biodiversity hypothesis. Allergy. 2019;74(8):1445–56.

66. McCoy KD, Burkhard R, Geuking MB. The microbiome and immune memory formation. Immunology & Cell Biology. 2019;97(7):625–35.

67. Negi S, Das DK, Pahari S, Nadeem S, Agrewala JN. Potential Role of Gut Microbiota in Induction and Regulation of Innate Immune Memory. Front Immunol. 2019 Oct 25;10.

68. Tischer C, Kirjavainen P, Matterne U, Tempes J, Willeke K, Keil T, et al. Interplay between natural environment, human microbiota and immune system: A scoping review of interventions and future perspectives towards allergy prevention. Science of The Total Environment. 2022 May 15;821:153422.

69. Ottman N, Ruokolainen L, Suomalainen A, Sinkko H, Karisola P, Lehtimäki J, et al. Soil exposure modifies the gut microbiota and supports immune tolerance in a mouse model. Journal of Allergy and Clinical Immunology. 2019 Mar 1;143(3):1198-1206.e12.

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

70. Deckers J, Lambrecht BN, Hammad H. How a farming environment protects from atopy. Current Opinion in Immunology. 2019 Oct 1;60:163–9.

71. Chatterjee S, Paul E, Liu Z, Gao J, Basak P, Bhattacharyya A, et al. Repurposing without Reinventing the Wheel - Ensemble Models for Differential Analysis [Internet]. bioRxiv; 2025 [cited 2025 Apr 16]. p. 2025.04.07.647549. Available from: https://www.biorxiv.org/content/10.1101/2025.04.07.647549v1

72. Al KF, Joris BR, Daisley BA, Chmiel JA, Bjazevic J, Reid G, et al. Multi-site microbiota alteration is a hallmark of kidney stone formation. Microbiome. 2023 Nov 25;11(1):263.

73. Hamrefors V, Kahn F, Holmqvist M, Carlson K, Varjus R, Gudjonsson A, et al. Gut microbiota composition is altered in postural orthostatic tachycardia syndrome and post-acute COVID-19 syndrome. Sci Rep. 2024 Feb 9;14(1):3389.

74. Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al. Microbiome differential abundance methods produce different results across 38 datasets. Nat Commun. 2022 Jan 17;13(1):342.

75. Chakraborty D, Jousset A, Wei Z, Banerjee S. Rare taxa in the core microbiome. Trends in Microbiology. 2025 Mar 27;0(0).

76. Zhuang M, Achmon Y, Cao Y, Liang X, Chen L, Wang H, et al. Distribution of antibiotic resistance genes in the environment. Environmental Pollution. 2021 Sep 15;285:117402.

77. Bai H, He LY, Wu DL, Gao FZ, Zhang M, Zou HY, et al. Spread of airborne antibiotic resistance from animal farms to the environment: Dispersal pattern and exposure risk. Environment International. 2022 Jan 1;158:106927.

78. Li X, Brejnrod A, Thorsen J, Zachariasen T, Trivedi U, Russel J, et al. Differential responses of the gut microbiome and resistome to antibiotic exposures in infants and adults. Nat Commun. 2023 Dec 22;14(1):8526.

79. Nadimpalli ML, Bourke CD, Robertson RC, Delarocque-Astagneau E, Manges AR, Pickering AJ. Can breastfeeding protect against antimicrobial resistance? BMC Med. 2020 Dec 15;18(1):392.

80. Pärnänen K, Ruuskanen MO, Sommeria-Klein G, Laitinen V, Kantanen P, Méric G, et al. Variation and prognostic potential of the gut antibiotic resistome in the FINRISK 2002 cohort. Nat Commun. 2025 Jul 1;16(1):5963.

81. Oliver A, Xue Z, Villanueva YT, Durbin-Johnson B, Alkan Z, Taft DH, et al. Association of Diet and Antimicrobial Resistance in Healthy U.S. Adults. mBio. 2022 May 10;13(3):e00101-22.

82. Asnicar F, Berry SE, Valdes AM, Nguyen LH, Piccinno G, Drew DA, et al. Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. Nat Med. 2021 Feb;27(2):321–32.

83. Dubourg G, Lagier JC, Armougom F, Robert C, Audoly G, Papazian L, et al. High-level colonisation of the human gut by Verrucomicrobia following broad-spectrum antibiotic treatment. International Journal of Antimicrobial Agents. 2013 Feb 1;41(2):149–55.

84. Diebold PJ, Rhee MW, Shi Q, Trung NV, Umrani F, Ahmed S, et al. Clinically relevant antibiotic resistance genes are linked to a limited set of taxa within gut microbiome worldwide. Nat Commun. 2023 Nov 14;14(1):7366.

medRxiv preprint doi: https://doi.org/10.1101/2025.06.25.25330258; this version posted October 13, 2025. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

85. Tiffany CR, Lee JY, Rogers AWL, Olsan EE, Morales P, Faber F, et al. The metabolic footprint of Clostridia and Erysipelotrichia reveals their role in depleting sugar alcohols in the cecum. Microbiome. 2021 Aug 19;9(1):174.

86. Genilloud O. Actinomycetes: still a source of novel antibiotics. Nat Prod Rep. 2017 Oct 18;34(10):1203–32.

87. Jokela R, Pärnänen ,Katariina MM, Ponsero ,Alise J, Lahti ,Leo, Kolho ,Kaija-Leena, de Vos ,Willem M, et al. A cohort study in family triads: impact of gut microbiota composition and early life exposures on intestinal resistome during the first two years of life. Gut Microbes. 2024 Dec 31;16(1):2383746.

88. Gibson MK, Wang B, Ahmadi S, Burnham CAD, Tarr PI, Warner BB, et al. Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. Nat Microbiol. 2016 Mar 7;1(4):1–10.