Lifestyle associates with unique resistome and microbiome signatures in children

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

Katariina Pärnänen[2], 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[2], https://orcid.org/0000-0001-5537-637X

Harri Alenius[1,3], https://orcid.org/0000-0003-0106-8923

Anna Bergström[1,4] , https://orcid.org/0000-0002-7981-6314

Matti O. Ruuskanen[2], https://orcid.org/0000-0003-4221-2880

Nanna Fyhrquist[1,5], 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 Computing, University of Turku, Turku, Finland

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

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

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

Abstract

Antibiotic resistance is a global health crisis that is not solely explained by antibiotics consumption. However, environmental and lifestyle contributions to antimicrobial resistance (AMR) in humans and especially children are not fully understood. 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 study. 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, independent of lifestyle. Despite limited statistical power, our findings suggest that lifestyle and environment both shape the microbiome and resistome. This study highlights the need for further research on species and ARG transmission from the environment to humans.

Introduction

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 of billions of dollars, with both numbers expected to rise drastically in the coming decades[2]. Young children are an especially vulnerable group, as they are susceptible to infections while at the same time suffering from

potential declines in drug efficacy due to antimicrobial resistance (AMR)[3]. In this context, the human gut microbiome, i.e., the collection of microbes (bacteria, archaea, fungi and protists) 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 can evade the effects of different antibiotic classes through various mechanisms, some of which target the antibiotic itself, while others indirectly lead to evasion of the antimicrobial agent[6]. These resistance mechanisms can arise through mutations or can be acquired through horizontal gene transfer (HGT)[7]. HGT can occur through four different mechanisms: transformation (direct uptake of DNA from the environment), membrane vesicle fusion (lipid vesicles that transport DNA between bacteria), transduction (bacteriophage-mediated transfer of genetic material between donor and recipient bacteria) and conjugation (cell-cell contact dependent transfer of DNA)[8]. Elements that facilitate transfer of genetic material within or between bacterial species are commonly referred to as mobile genetic elements (MGEs), and include insertion sequences, transposons, gene cassettes/integrons, plasmids and integrative conjugative elements[9]. Thus, the human gut microbiome can function as a reservoir for antimicrobial resistance, as ARGs are known to be transported via MGEs, and can be passed on from commensals to pathogens, potentially leading to antimicrobial resistant infections[10].

Even though ARGs occur naturally in the environment and can be traced back to prehistoric ages[11], 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[12]. However, the drivers of ARG abundance in the environment remain uncertain, as antibiotics consumption alone does not explain the occurrence of antimicrobial resistance worldwide[13]. Furthermore, the transmission dynamics of ARGs between animals, humans and environments are far from being quantified and understood[14].

Environmental exposures and lifestyle factors are known to strongly influence the gut microbiome composition[15]. Yet most studies on childhood microbiomes and resistomes focus on early infancy[16], 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[16]. 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[17]. However, the exact mechanisms behind this protective effect remain unclear, though current studies suggest a complex interplay between the bacteria and the host immune system[18]. 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[19,20].

Given the growing global threat of AMR and the role of the gut microbiome as a potential reservoir for ARGs, it is essential to understand how lifestyle factors influence the resistome during childhood. In this study, we analyzed metagenomic data from the Swedish participants of the PARSIFAL study to investigate how anthroposophic and farm lifestyles affect the gut microbiome and resistome in school-aged children. Therefore, we conducted presence/absence, alpha and beta diversity, as well as differential abundance analyses, 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.

Results

Study population

The PARSIFAL cross-sectional study has previously been described in detail by Alfvén et al.[19]. 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). The variables used in this study are listed in Table 1.

Table 1 | Data for the PARSIFAL study population by lifestyle group. Percentages were calculated on a group basis. The values for some variables do not add to 100 % due to non-answers in the questionnaire.

	Farm *n* = 37	Farm ref *n* = 11	Steiner *n* = 19	Steiner ref *n* = 7
Sex (% male)	59	64	47	57
Age in years, median (range)	9 (6-13)	9 (6-12)	7 (5-11)	7 (6-9)
BMI[1] (% overweight and obese)	32	9	11	14
Never breastfed (%)	5	0	0	0
Breastfed exclusively < 5 months of age (%)	62	45	32	43
Breastfed exclusively ≥ 5 months of age (%)	30	45	63	57
Never used antibiotics (%)	3	9	32	0
First use of antibiotics ≤ 12 months of age (%)	41	45	32	71
First use of antibiotics > 12 months of age (%)	49	27	37	29
Never used antipyretics (%)	0	0	53	14
First use of antipyretics ≤ 12 months of age (%)	41	64	42	57
First use of antipyretics > 12 months of age (%)	49	27	5	29
Measles (% diseased)	0	0	37	0
Current environmental smoking at home (%)	11	18	16	0
Maternal smoking during pregnancy (%)	14	0	5	14
Use of organic and biodynamic food (%)	11	9	79	0
Household pets during the first year of life (%)	70	18	21	29
No older siblings (%)	30	27	16	14
Parent's education (% elementary school and lower)	14	9	0	0

[1]Child overweight and obesity as classified by Cole et al.[21]

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 taxa belonged to the class Clostridia, followed by Actinomycetia and Bacteroidia. On the family level, most species belonged to three families (*Lachnospiraceae*, *Bifidobacteriaceae* and *Oscillospiraceae*). The overwhelming majority of ARGs were tetracycline resistance genes. In addition to tetracycline resistance, resistance to four other classes of antibiotics were common (aminoglycosides, amphenicols, beta-lactams, and macrolide–lincosamide–streptogramin B; Fig. 1c)

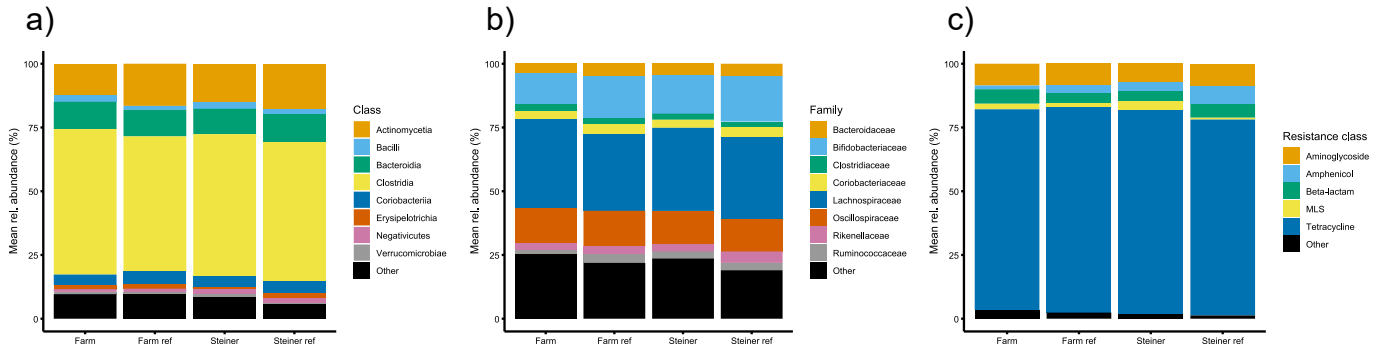

Fig. 1 | Microbiome and resistome compositions in the gut of the lifestyle groups. Mean relative abundances (%) are shown on the y-axis. **a** Most abundant bacterial classes. **b** Most abundant bacterial families. **c** Most abundant antibiotic resistance classes. Farm: Farm group (n = 37), Farm ref: 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 supplements (Supplementary Data 1, Supplementary Figure 2, 3). Furthermore, the proportions 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 proportions 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, $p > 0.05$).

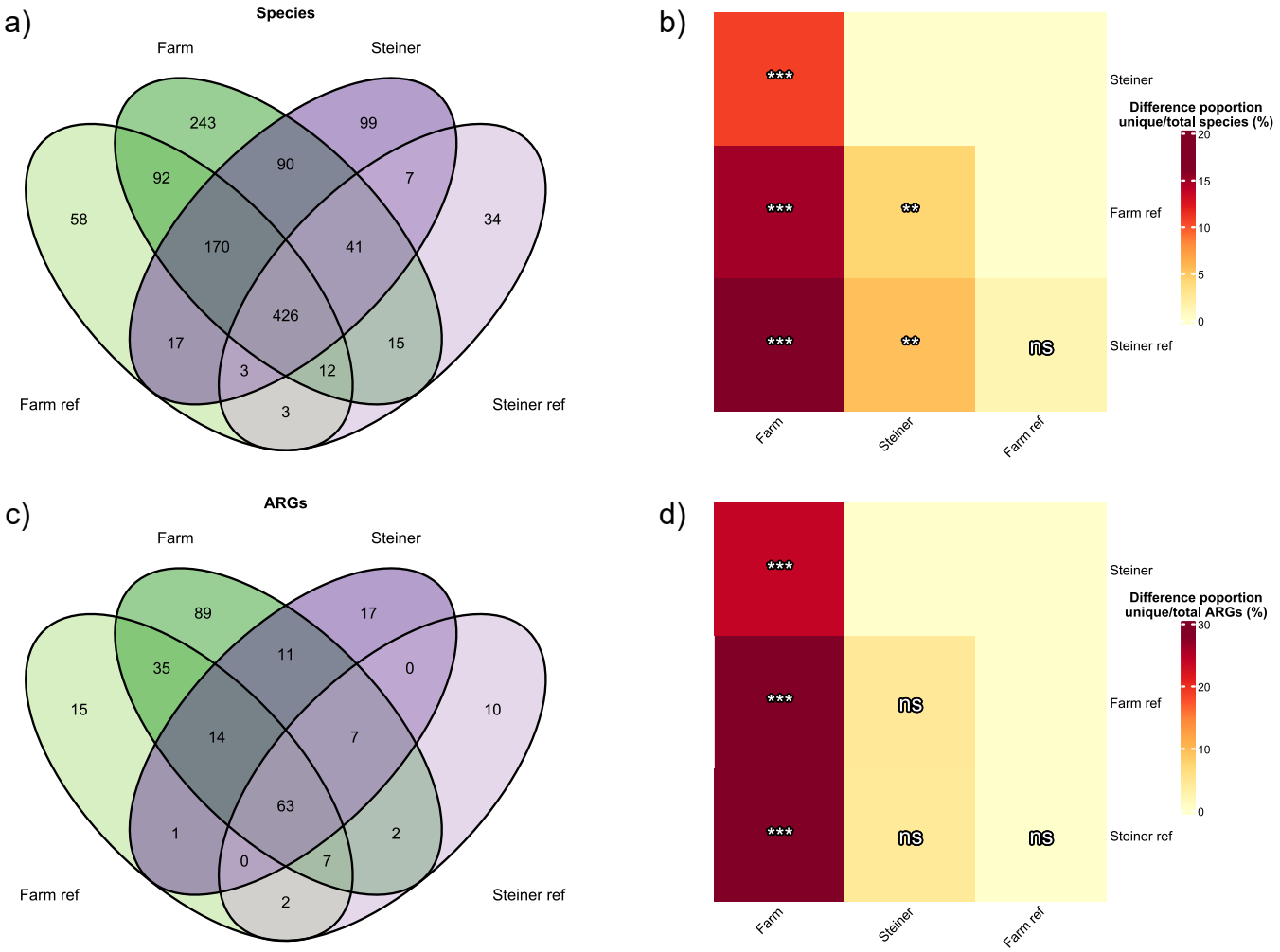

Fig. 2 | Venn diagrams and pairwise comparisons of proportions. Venn diagrams depict the shared species (n = 1310) and ARGs (n = 273) between lifestyle and reference groups. **a** presence/absence of species in a lifestyle group, **b** pairwise comparison of proportions of the ratios of unique to non-unique 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).

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. 7, $p \leq 0.05$), 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, $p \leq 0.05$).

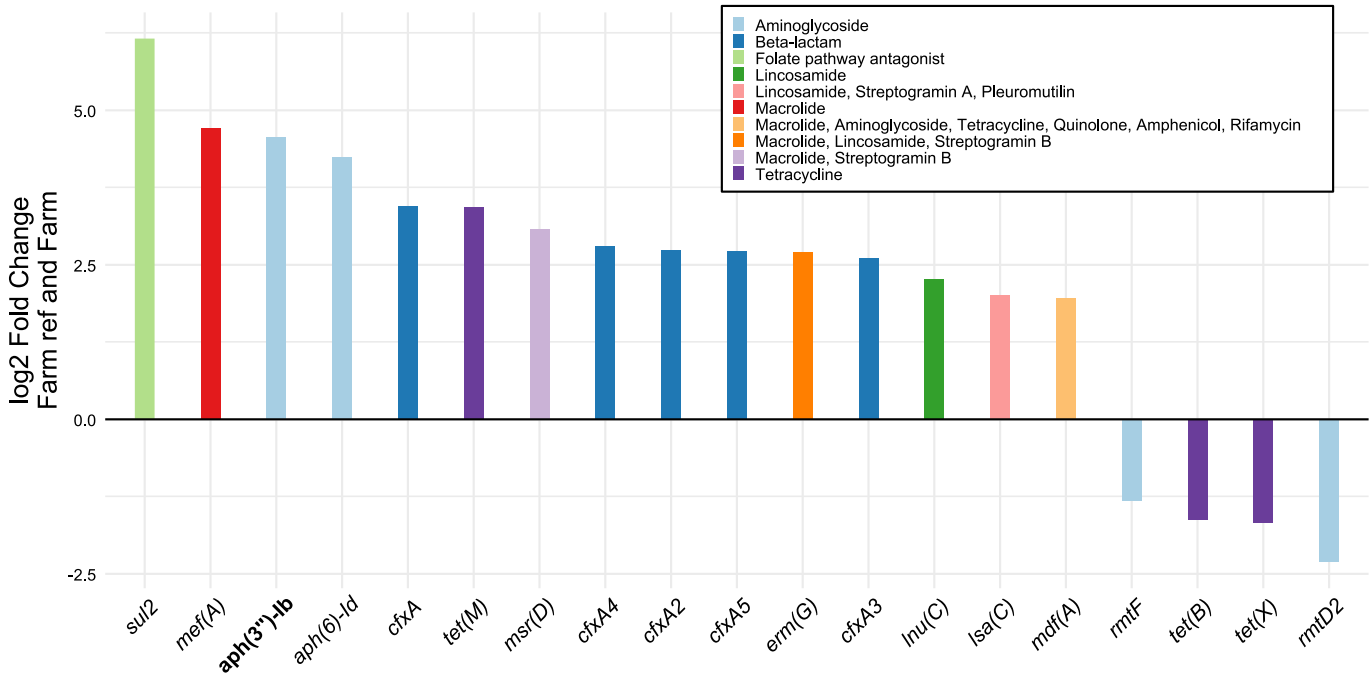

Fig. 3 | Differentially abundant ARGs and their corresponding resistance classes between farm children and their reference group. All genes shown were found to be significant in DESeq2 ($p \leq 0.05$). The y-axis shows log2 fold changes, the x-axis denotes gene names, positive log2 fold changes indicate differential abundance in the farm group, negative changes differential abundance in the farm reference group. *aph(3")-lb* was found to be differentially abundant in Steiner references compared to the Steiner group (log2 fold = -6.58). 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 group variables with the overall ARG load, a GLM was fitted using stepwise regression, using only complete case records. The body mass index (BMI) category (Fig. 4a), length of 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 % CI: 0.60-0.94). Children breastfed for more than 5 months had a 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).

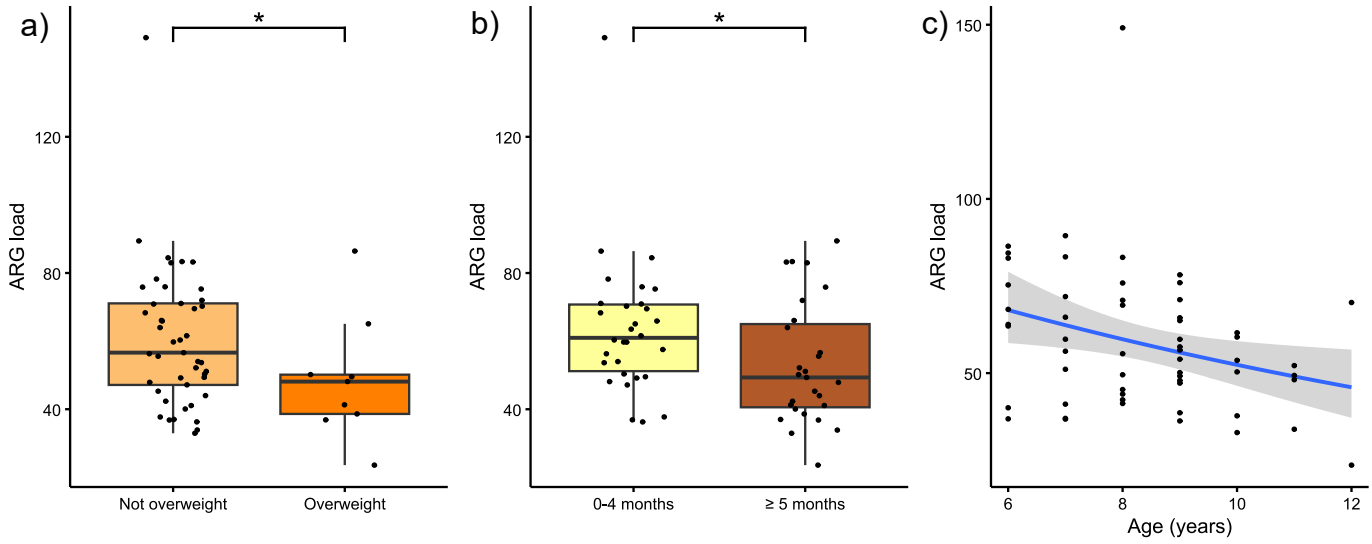

Fig. 4 | ARG load in relation to PARSIFAL study variables. To investigate the effects of variables on the ARG load a gamma-distributed GLM model was fitted using backward stepwise-regression. ARG load (n = 57) by: **a** BMI (not overweight and overweight), **b** length of breastfeeding (non-breastfed 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[21]. 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. 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).

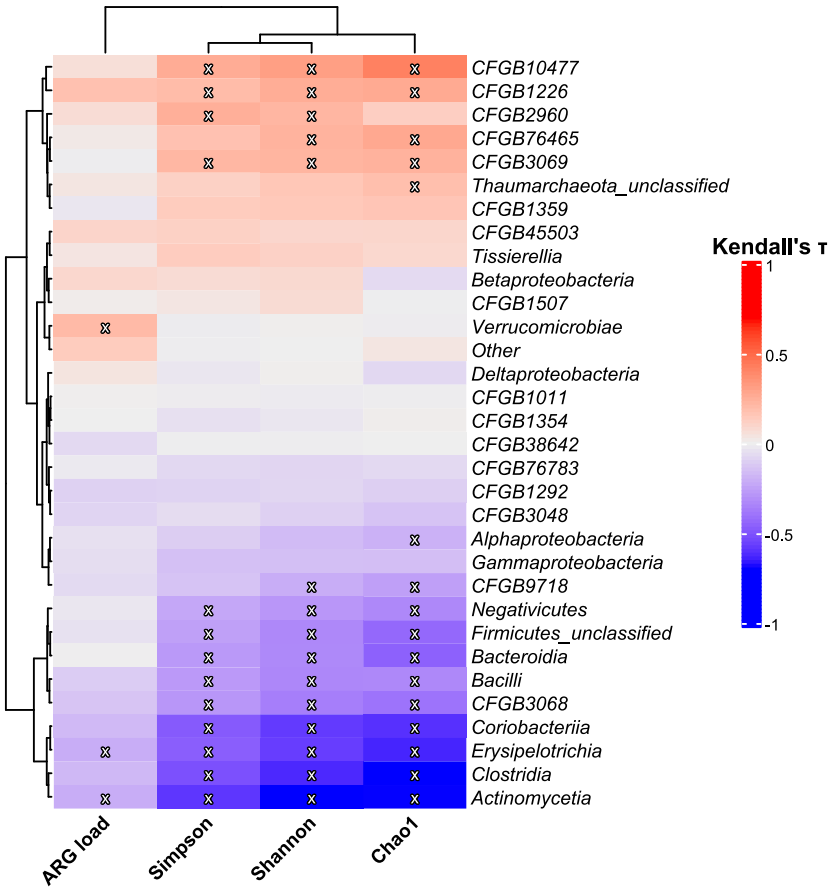

Fig. 5 | Correlation of bacterial taxa with ARG load and inverse Simpson, Shannon and Chao1 diversities. CLR transformed bacterial class abundances were correlated with diversity measures using the Kendall rank correlation coefficient (τ). Significant correlations (BH corrected $p \leq 0.05$) are denoted 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".

Discussion

Our study identified differences regarding species and ARG presence and absence between the group, as well as differentially abundant ARGs in the farm group. Furthermore, we were able to identify and confirm select lifestyle factors that influence the ARG load in school-aged children.

Our analysis of rare taxa showed that differences in the gut microbiome of both farm and anthroposophic children were mainly defined 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 rare species and ARGs are highly individual. This aligns with previous findings that suggest a strong impact of the farm environment on the human gut microbiome and resistome[22], especially since we observed multiple farm-exclusive species that were strongly associated with farm animals, such as *Mogibacterium kristiansenii, Baileyella intestinalis*[23] (both pigs), *Candidatus Neoanaerotignum tabaqchaliae*[24] (chicken) and *Clostridium cuniculi*[25] (rabbits). With regards to the anthroposophic children, some of the species found exclusively in this group were associated with

fermentation of foods (*Leuconostoc citreum*[26], *Lactobacillus farciminis*[26]) and/or are lactic acid bacteria (*Pediococcus pentosaceus*[27], *Ligilactobacillus salivarius*[28]). Lactic acid bacteria have long been discussed to have a variety of beneficial health effects[29], 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[16,30]. However, these studies cannot be directly compared to our study as both methodology and study design (subjects below 2 years of age) differ strongly from ours. Given the small sample sizes for both 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[31]. 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[32]. 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. However, it should be mentioned that the original study used a larger subsample from several European countries, which may influence comparisons.

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[19]. Nonetheless, our findings are in line with the biodiversity hypothesis[33], 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[34,35] the presence of a high number of taxa unique to this group suggests a potential link to immune resilience, which merits 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[36].Given the strong protective influence of soil[37] and dust[38] exposure on the microbiome and immune function, the inclusion environmental measurements will be essential in future research.

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[39], kidney stone formation[40] and postural orthostatic tachycardia syndrome[41]. This result is not surprising, as we employed the commonly used prevalence filter cutoff of 10 % that removes rare and often low-abundance

taxa from DAA analyses[42]. However, focusing solely on DAA to analyze differences between treatment and lifestyle groups omits differences regarding rare species that might be part of the core microbiome[43]. 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 seeding sludges or sewage samples[44]. This indicates the farm environment as a potential risk factor for the acquisition of antibiotic resistance genes. Transmission of species that carry ARGs is possibly even airborne, as Bai et al.[45] found that this transmission route is relevant even at a distance from the actual farm working environment.

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)[46]. Furthermore, the results of the current study indicate that extended breastfeeding may be beneficial in reducing the overall ARG load. In general, breastfeeding has been identified as a factor that lowers the ARG load in infants compared to formula-based diets[47]. However, it remains unclear to what extent this effect is due to reduced exposure to antibiotic resistant bacteria, such as through dietary sources, or molecular mechanisms influencing the composition of the gut microbiome[48]. Additionally, our study found a lower ARG load in overweight children compared to non-overweight children. This result aligns with the finding in a large cohort from Finland[49] 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 the high-fat and high-processed food diets. However, given the association of high phylogenetic diversity of foods with a low ARG load in adults[50] these findings warrant further investigation, as a phylogenetically diverse diet is associated with a healthy lifestyle[51]. The GLM employed in our study would maybe have confirmed well known ARG load influencing factors such as antibiotic treatment if it was not for the small sample size of this study.

Verrucomicrobiae was the only bacterial class which abundance was positively correlated with the ARG load. This may be due to a high competitiveness of the Verrucomicrobia phylum after antibiotic treatment[52], 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[53]. 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 germ-free mice[54], contrasting with our findings in humans. The negative correlation between Actinomycetia abundance and ARG load is unexpected, given its known production of antimicrobial compounds[55]. However, our analysis cannot determine causality. The significant correlations between bacterial class abundances and alpha

diversity suggest a decoupling of species diversity and ARG load. Since no significant links were found between ARG class abundances and overall ARG load or alpha diversity, our results indicate that children's resistomes might differ from those of infants, which are strongly influenced by *Bacteroides* and *Bifidobacterium*[56].

Many current studies examining the microbiome and resistome in the gut of children focus on the first few weeks to months after birth, taking into account select factors such as nutrition and antibiotic treatment that allow well-defined study groups[47,57]. 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[58]. 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 and longitudinal studies that focus on the farm environment are needed to explore connections between the environment, microbiome and resistome, due to the cross-sectional design and comparatively small number of participants in this study

To conclude, our study showed clear influences of the farm environment on the gut microbiome and resistome of children, while indicating a highly individualized exposure to/repertoire of low-abundance microbes and ARGs. Further research on rare species and ARGs with a focus on environmental transmission may therefore be warranted. Additionally, we were able to confirm (lifestyle) variables such as age, obesity and the duration of breastfeeding that affect the ARG load in a currently understudied age-range.

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 (Prevention of Allergy Risk factors for Sensitization In children related to Farming and Anthroposophic Lifestyle) study[19] 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[32].

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)

Metagenomic analysis

Paired-end reads were aligned using Bowtie 2 version 2.4.4[59]. Aligned reads were then filtered using SAMtools 1.16.1 and converted into BAM format[60]. Microbial community profiling of filtered metagenomic sequence reads was performed using MetaPhlAn 4 version 4.0.3[61]. We mapped the R1 and R2 reads with Bowtie2 v 2.4.4 against the ResFinder database version 2.1.1[62] with the following options: "-D 20 -R 3 -N 1 -L 20 -i S,1,0 5" to identify ARGs. ResFinder was chosen since it is an ARG database that exclusively contains acquired and clinically relevant ARGs, as our focus was on these types of ARGs, not on efflux pumps or other genes with less clear roles in clinical AMR. The default quality scores in Bowtie were used to ensure high-quality matches are included.

Data normalization

The MetaPhlAn relative abundance (%) taxonomic data was transformed into counts through division by 100 and multiplication by library size. Counts were then scaled so that the lowest count > 0 equaled 1. ARG counts were first normalized to gene lengths, and then to library size. Again, counts were then 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, ARG counts were first normalized to gene length. Then, sample-wise sums of ARGs were calculated and normalized to library size. Normalized values were then scaled to represent reads per kilobase of transcript per million mapped reads[63] (RPKM). Library size was chosen for data normalization as it yields similar results compared to 16S rRNA or single marker copy gene normalization[47].

Rarefaction, alpha and beta diversity

To ensure the robustness of the presence/absence analysis, both MetaPhlAn and ARG counts were additionally normalized to group size before the scaling. Then, rarefaction as defined by Schloss[64], was performed with 100 iterations using rarefy_even_depth from phyloseq 1.50.0[65]. 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, rarefaction was performed with 100 iterations, using the rarefy_even_depth and getDissimilarity functions of the phyloseq[65] 1.50.0 and mia[66] 1.15.19 packages respectively. Again, the minimum library size was set equal to the smallest sample library size, and subsampling was performed without replacement. Alpha diversities (Shannon[67], inverse Simpson[68], Chao1[69]) were calculated using the estimate_richness function from phyloseq. Beta diversity indices (Bray-Curtis[70], Jaccard[71]) were calculated using the vegdist function (called with

getDissimilarity) from the vegan[72] 2.6-8 package. The ARG counts were not rarefied before alpha and beta diversity analyses.

Statistical analysis

Relative abundances of bacterial classes, families and ARG resistance classes were calculated using the transformAssay function (method = "relabundance") from the mia[66] 1.15.19 package. The differences in the ratio of the number of exclusive to non-exclusive species and ARGs between the different lifestyle groups were analyzed using a pairwise comparison of proportions implemented in pairwise.prop.test from the stats[73] 4.4.2 package. The venn diagrams were plotted using ggvenn[74] 0.1.10. To test for differentially abundant species between the lifestyle groups and their respective references both MaAsLin2 1.20.0[75] and the Wilcoxon test[76] were used, as previously suggested[77]. 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[78], DESeq2[79] 1.46.0 was used to analyze differentially abundant ARGs between the lifestyle 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 complete cases were analyzed (n = 57). The model was built using the variables listed in Table 1. The glm function of the stats[73] package was used to fit the model using the option family "gamma" (link = "log") as done previously[47]. The step function of the stats[73] 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[80] 1.4-26 was used to perform Tukey's post hoc test to obtain adjusted p-values for pairwise comparisons. To correlate the abundance of bacterial and ARG classes with ARG load, Shannon[67], inverse Simpson[68] and Chao1[69] indices, classes with less than 50 % prevalence were aggregated to "Other". Then, transformAssay of the mia[66] 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[66] package. Differences in relative abundances of bacterial classes between the lifestyle and reference groups were calculated and graphed using metacoder[81] 0.3.8. To test differences between lifestyle and reference groups regarding alpha diversities, the Mann-Whitney U Test[82] was performed, as alpha diversities were not normally distributed, which was tested using the Shapiro-Wilk test[83]. Principal coordinates analysis[84] (PCoA) was performed using the cmdscale function in the stats[73] package. Unless otherwise stated, p-values were corrected for multiple testing using the Benjamini–Hochberg procedure[85]. Results were considered significant for $p \leq 0.05$. Plots were created using ggplot2 3.5.1[86]. All analyses were performed in R 4.4.2[73].

Ethical permit

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).

Data availability

The datasets analyzed during the current study are not publicly available due to the sensitive nature of the research data but are available from the corresponding authors upon reasonable and qualified request.

Code availability

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

Acknowledgments

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. 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.

Author 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.

Competing Interests

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

References

1. Naghavi, M. *et al.* Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. *The Lancet* **404**, 1199–1226 (2024).
2. Aslam, B. *et al.* Antibiotic resistance: a rundown of a global crisis. *Infection and Drug Resistance* **11**, 1645–1658 (2018).
3. Laxminarayan, R. *et al.* Access to effective antimicrobials: a worldwide challenge. *The Lancet* **387**, 168–175 (2016).
4. Joos, R. *et al.* Examining the healthy human microbiome concept. *Nat Rev Microbiol* **23**, 192–205 (2025).
5. van Schaik, W. The human gut resistome. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**, 20140087 (2015).
6. Darby, E. M. *et al.* Molecular mechanisms of antibiotic resistance revisited. *Nat Rev Microbiol* **21**, 280–295 (2023).

7. Huddleston, J. R. Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. *IDR* **7**, 167–176 (2014).
8. McInnes, R. S., McCallum, G. E., Lamberte, L. E. & van Schaik, W. Horizontal transfer of antibiotic resistance genes in the human gut microbiome. *Current Opinion in Microbiology* **53**, 35–43 (2020).
9. Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clinical Microbiology Reviews* **31**, 10.1128/cmr.00088-17 (2018).
10. Brito, I. L. Examining horizontal gene transfer in microbial communities. *Nat Rev Microbiol* **19**, 442–453 (2021).
11. Aminov, R. I. A Brief History of the Antibiotic Era: Lessons Learned and Challenges for the Future. *Frontiers in Microbiology* **1**, (2010).
12. Zhang, Z. *et al.* Assessment of global health risk of antibiotic resistance genes. *Nat Commun* **13**, 1553 (2022).
13. Collignon, P., Beggs, J. J., Walsh, T. R., Gandra, S. & Laxminarayan, R. Anthropological and socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable analysis. *The Lancet Planetary Health* **2**, e398–e405 (2018).
14. Larsson, D. G. J. & Flach, C.-F. Antibiotic resistance in the environment. *Nat Rev Microbiol* **20**, 257–269 (2022).
15. Rothschild, D. *et al.* Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**, 210–215 (2018).
16. Alm, J. S. *et al.* An anthroposophic lifestyle and intestinal microflora in infancy. *Pediatric Allergy and Immunology* **13**, 402–411 (2002).
17. Ege, M. J. The Hygiene Hypothesis in the Age of the Microbiome. *Annals ATS* **14**, S348–S353 (2017).
18. Depner, M. *et al.* Maturation of the gut microbiome during the first year of life contributes to the protective farm effect on childhood asthma. *Nat Med* **26**, 1766–1775 (2020).
19. Alfvén, T. *et al.* Allergic diseases and atopic sensitization in children related to farming and anthroposophic lifestyle – the PARSIFAL study. *Allergy* **61**, 414–421 (2006).
20. Flöistrup, H. *et al.* Allergic disease and sensitization in Steiner school children. *Journal of Allergy and Clinical Immunology* **117**, 59–66 (2006).
21. Cole, T. J., Bellizzi, M. C., Flegal, K. M. & Dietz, W. H. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* **320**, 1240 (2000).
22. Sun, J. *et al.* Environmental remodeling of human gut microbiota and antibiotic resistome in livestock farms. *Nat Commun* **11**, 1427 (2020).
23. Wylensek, D. *et al.* A collection of bacterial isolates from the pig intestine reveals functional and taxonomic diversity. *Nat Commun* **11**, 6389 (2020).
24. Glendinning, L., Stewart, R. D., Pallen, M. J., Watson, K. A. & Watson, M. Assembly of hundreds of novel bacterial genomes from the chicken caecum. *Genome Biology* **21**, 34 (2020).
25. Djukovic, A. *et al.* Gut colonization by a novel Clostridium species is associated with the onset of epizootic rabbit enteropathy. *Veterinary Research* **49**, 123 (2018).
26. Nam, S.-H. *et al.* Genome Sequence of Lactobacillus farciminis KCTC 3681. *Journal of Bacteriology* **193**, 1790–1791 (2011).
27. Jiang, S., Cai, L., Lv, L. & Li, L. Pediococcus pentosaceus, a future additive or probiotic candidate. *Microb Cell Fact* **20**, 45 (2021).
28. Guerrero Sanchez, M., Passot, S., Campoy, S., Olivares, M. & Fonseca, F. Ligilactobacillus salivarius functionalities, applications, and manufacturing challenges. *Appl Microbiol Biotechnol* **106**, 57–80 (2022).
29. 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* **44**, 454–489 (2020).
30. Hesla, H. M. *et al.* Impact of lifestyle on the gut microbiota of healthy infants and their mothers – the ALADDIN birth cohort. *FEMS Microbiology Ecology* **90**, 791–801 (2014).
31. King, C. H. *et al.* Baseline human gut microbiota profile in healthy people and standard reporting template. *PLOS ONE* **14**, e0206484 (2019).
32. Dicksved, J. *et al.* Molecular Fingerprinting of the Fecal Microbiota of Children Raised According to Different Lifestyles. *Applied and Environmental Microbiology* **73**, 2284–2289 (2007).

33. Haahtela, T. A biodiversity hypothesis. *Allergy* **74**, 1445–1456 (2019).
34. McCoy, K. D., Burkhard, R. & Geuking, M. B. The microbiome and immune memory formation. *Immunology & Cell Biology* **97**, 625–635 (2019).
35. Negi, S., Das, D. K., Pahari, S., Nadeem, S. & Agrewala, J. N. Potential Role of Gut Microbiota in Induction and Regulation of Innate Immune Memory. *Front. Immunol.* **10**, (2019).
36. Tischer, C. *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* **821**, 153422 (2022).
37. Ottman, N. *et al.* Soil exposure modifies the gut microbiota and supports immune tolerance in a mouse model. *Journal of Allergy and Clinical Immunology* **143**, 1198-1206.e12 (2019).
38. Deckers, J., Lambrecht, B. N. & Hammad, H. How a farming environment protects from atopy. *Current Opinion in Immunology* **60**, 163–169 (2019).
39. Chatterjee, S. *et al.* Repurposing without Reinventing the Wheel - Ensemble Models for Differential Analysis. 2025.04.07.647549 Preprint at https://doi.org/10.1101/2025.04.07.647549 (2025).
40. Al, K. F. *et al.* Multi-site microbiota alteration is a hallmark of kidney stone formation. *Microbiome* **11**, 263 (2023).
41. Hamrefors, V. *et al.* Gut microbiota composition is altered in postural orthostatic tachycardia syndrome and post-acute COVID-19 syndrome. *Sci Rep* **14**, 3389 (2024).
42. Nearing, J. T. *et al.* Microbiome differential abundance methods produce different results across 38 datasets. *Nat Commun* **13**, 342 (2022).
43. Chakraborty, D., Jousset, A., Wei, Z. & Banerjee, S. Rare taxa in the core microbiome. *Trends in Microbiology* **0**, (2025).
44. Zhuang, M. *et al.* Distribution of antibiotic resistance genes in the environment. *Environmental Pollution* **285**, 117402 (2021).
45. Bai, H. *et al.* Spread of airborne antibiotic resistance from animal farms to the environment: Dispersal pattern and exposure risk. *Environment International* **158**, 106927 (2022).
46. Li, X. *et al.* Differential responses of the gut microbiome and resistome to antibiotic exposures in infants and adults. *Nat Commun* **14**, 8526 (2023).
47. Pärnänen, K. M. M. *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* **115**, 407–421 (2022).
48. Nadimpalli, M. L. *et al.* Can breastfeeding protect against antimicrobial resistance? *BMC Med* **18**, 392 (2020).
49. Pärnänen, K. *et al.* Population variation and prognostic potential of gut antibiotic resistome. 2024.08.08.24311663 Preprint at https://doi.org/10.1101/2024.08.08.24311663 (2024).
50. Oliver, A. *et al.* Association of Diet and Antimicrobial Resistance in Healthy U.S. Adults. *mBio* **13**, e00101-22 (2022).
51. Asnicar, F. *et al.* Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. *Nat Med* **27**, 321–332 (2021).
52. Dubourg, G. *et al.* High-level colonisation of the human gut by Verrucomicrobia following broad-spectrum antibiotic treatment. *International Journal of Antimicrobial Agents* **41**, 149–155 (2013).
53. Diebold, P. J. *et al.* Clinically relevant antibiotic resistance genes are linked to a limited set of taxa within gut microbiome worldwide. *Nat Commun* **14**, 7366 (2023).
54. Tiffany, C. R. *et al.* The metabolic footprint of Clostridia and Erysipelotrichia reveals their role in depleting sugar alcohols in the cecum. *Microbiome* **9**, 174 (2021).
55. Genilloud, O. Actinomycetes: still a source of novel antibiotics. *Nat. Prod. Rep.* **34**, 1203–1232 (2017).
56. Jokela, R. *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* **16**, 2383746 (2024).
57. Gibson, M. K. *et al.* Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat Microbiol* **1**, 1–10 (2016).
58. Lebeaux, R. M. *et al.* The association between early life antibiotic exposure and the gut resistome of young children: a systematic review. *Gut Microbes* **14**, 2120743 (2022).

59. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).
60. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
61. Blanco-Míguez, A. *et al.* Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat Biotechnol* **41**, 1633–1644 (2023).
62. Florensa, A. F., Kaas, R. S., Clausen, P. T. L. C., Aytan-Aktug, D. & Aarestrup, F. M. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microbial Genomics* **8**, 000748 (2022).
63. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**, 621–628 (2008).
64. Schloss, P. D. Waste not, want not: revisiting the analysis that called into question the practice of rarefaction. *mSphere* **9**, e00355-23 (2023).
65. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217 (2013).
66. Ernst, F. G. M., Shetty, S. A., Borman, T. & Lahti, L. mia: Microbiome analysis. (2025).
67. Shannon, C. E. A Mathematical Theory of Communication. *Bell System Technical Journal* **27**, 379–423 (1948).
68. Simpson, E. H. Measurement of Diversity. *Nature* **163**, 688–688 (1949).
69. Chao, A. Nonparametric Estimation of the Number of Classes in a Population. *Scandinavian Journal of Statistics* **11**, 265–270 (1984).
70. Bray, J. R. & Curtis, J. T. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* **27**, 325–349 (1957).
71. Jaccard, P. The Distribution of the Flora in the Alpine Zone. *New Phytologist* **11**, 37–50 (1912).
72. Oksanen, J. *et al.* vegan: Community Ecology Package. (2022).
73. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (2024).
74. Yan, L. ggvenn: Draw Venn Diagram by 'ggplot2'. (2023).
75. Mallick, H. *et al.* Multivariable association discovery in population-scale meta-omics studies. *PLOS Computational Biology* **17**, e1009442 (2021).
76. Wilcoxon, F. Individual Comparisons by Ranking Methods. *Biometrics Bulletin* **1**, 80–83 (1945).
77. Pelto, J., Auranen, K., Kujala, J. V. & Lahti, L. Elementary methods provide more replicable results in microbial differential abundance analysis. *Briefings in Bioinformatics* **26**, bbaf130 (2025).
78. Jonsson, V., Österlund, T., Nerman, O. & Kristiansson, E. Statistical evaluation of methods for identification of differentially abundant genes in comparative metagenomics. *BMC Genomics* **17**, 78 (2016).
79. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).
80. Hothorn, T., Bretz, F. & Westfall, P. Simultaneous Inference in General Parametric Models. *Biometrical Journal* **50**, 346–363 (2008).
81. Metacoder: An R package for visualization and manipulation of community taxonomic diversity data | PLOS Computational Biology. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005404.
82. Mann, H. B. & Whitney, D. R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics* **18**, 50–60 (1947).
83. Shapiro, S. S. & Wilk, M. B. An analysis of variance test for normality (complete samples). *Biometrika* **52**, 591–611 (1965).
84. Gower, J. C. Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis. *Biometrika* **53**, 325–338 (1966).
85. 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)* **57**, 289–300 (1995).
86. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag (2016).