Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Early prenatal antibiotic usage has limited impact on the microbiomes of infants

Hanna Theodora Noordzij
University of Oslo
Eric Jacques de Muinck
University of Oslo
Morten Nilsen
Norwegian University of Life Sciences
Knut Rudi
Norwegian University of Life Sciences
Riyas Vettukattil
Oslo University Hospital
Håvard Skjerven
Oslo University Hospital
Karin Lødrup Carlsen
Oslo University Hospital
Pål Trosvik
pal.trosvik@ibv.uio.no

University of Oslo

Article

Keywords:

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11351947/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Early prenatal antibiotic usage has limited impact on the microbiomes of infants

Hanna Theodora Noordzij[1], Eric Jacques de Muinck[1], Morten Nilsen[2], Knut Rudi[2], Riyas Vettukattil[3], Håvard Skjerven[3], Karin Lødrup Carlsen[3], Pål Trosvik[1*]

1. Centre for Ecological and Evolutionary Synthesis, University of Oslo, Oslo, Norway. 2. Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway. 3. Department of Paediatric Medicine, Oslo University Hospital, Oslo, Norway.

*Corresponding author: Pål Trosvik (pal.trosvik@ibv.uio.no)

Abstract

Background

Antibiotic resistance poses an increasingly significant challenge to public health. The human gut microbiome is a reservoir for antibiotic resistance genes (ARGs), and ARG carriage is particularly high in infants compared with adults. Here, we investigate the gastrointestinal microbiome in a cohort of mothers and their born-to-term infants sampled during pregnancy and at 3- 6- and 12-months of age using deep shotgun metagenome sequencing of fecal DNA.

Results

Despite a high level of taxonomic individuality, the infant ARG complement showed strong temporal structure in terms of composition, with decreasing dispersion and increasing diversity up to the first birthday, while the lowest levels of dispersion and highest diversity were observed in the mothers. ARG abundance decreased over time to levels comparable with adults at 12 months. These trends were also observed for genes encoding known virulence factors. We also assembled plasmid sequences and found similar temporal structure, but without any significant change in abundances over time. Surprisingly, at

no time point did the infants microbiomes resemble those of their own mothers more than those of unrelated women with respect to taxonomic composition, ARGs and VF gene content. We found minimal impact of pre-partum antibiotic use on the infant microbiome. Furthermore, antibiotic treatment during early infancy only had a minor impact on the microbiome at later time points.

Conclusions

This study adds to a growing body of literature demonstrating the temporal structure of key elements of the gut microbiome during infancy and offers some positive insights with respect to moderate exposure to antibiotics.

Introduction

The colonization period of the microbial community residing in the gastrointestinal (GI) tract is instrumental to infant development, as well as influencing health and disease later in life [1]. However, the same microbial community also acts as host to a sizeable complement of antibiotic resistance genes (ARGs), the so-called resistome, and infants tend to carry higher ARG abundances than their mothers [2]. Early colonizers of the infant GI, like *Escherichia coli* (*E.coli*), have been associated with high carriage of ARGs [3, 4]. As the gut and its microbiome matures, taxa more specialized to the gut environment, such as species of the phyla Bacteroidetes and Firmicutes, displace facultative anaerobes and become dominant. The period from early colonization to a more mature state is characterized by unstable community dynamics [5], which leads to high inter-individual variation in taxonomic composition during this critical developmental phase. Thus, the effects of antibiotic exposure on the infant microbiome can potentially vary a great deal depending on the specific context. Furthermore, use of antibiotics during pregnancy can also influence development of the infant GI microbiome [6, 7], including potential effects on the resistome.

Bacterial resistance to antibiotic substances is a natural phenomenon of ancient origin [8], and not necessarily problematic when it is not associated with pathogenic properties. However, studies have shown that high ARG carriage tends to correlate with high virulence potential [9] and accumulate in certain lineages as exemplified by Proteobacteria [10], and by *E. coli* in particular [4]. Proteobacteria are common early colonizers of the human gut, and commensal species and strains are particularly abundant in the infant GI microbiota [5].

A major concern in current public health is the spread of antibiotic resistance, including the potential for horizontal gene transfer (HGT) of ARGs from commensal to pathogenic bacteria [11]. As such, the infant gut microbiome could act as a reservoir and hub for the dissemination of ARGs across taxonomic boundaries. Plasmids are major vehicles of HGT and may disseminate among species, or even genera [12, 13]. The human GI microbiome harbors a substantial collection of plasmids [14], sometimes referred to as the plasmidome. These molecules can have complex dynamics that do not necessarily follow the population dynamics of carrier species [15]. Furthermore, some plasmids appear to transmit from mother to infant [16].

Here, we examine the microbiome in a cohort of 50 mother/infants (MI) pairs, with feces collected during pregnancy and from infants at 3, 6 and 12 months of age. We characterize changes in community structure during infancy towards the mature configuration seen in the maternal microbiomes, with decreasing β-diversity and increasing α-diversity. In addition, we describe a resistome that follows similar trends as seen for the taxonomic composition, but with declining overall carriage with increasing age. These results were mirrored in the metagenomic content of virulence factor genes (VFGs), while for plasmids we saw similar trends in diversity, but no trend for overall abundance. Importantly, we investigated potential impacts of early prenatal antibiotic usage on the microbiomes of infants. Overall, we found minimal impacts on taxonomic composition, overall ARG carriage or resistome composition in any of the three infant age groups. Furthermore, exposure to antibiotics during early infancy did not

have significant effects on ARG carriage at subsequent time points. These results suggest that early antibiotic exposure may not always lead to significant long-term impacts on the microbiome.

Methods

Study population

The study participants were a subset of a larger, general population-based mother-child cohort, recruited at 18-week pregnancy, and originally recruited for studying Atopic Dermatitis and allergies in children (The PreventADALL study; [17]). This study enrolled 2701 pregnant women in Norway and Sweden. The sub-population investigated here consisted of 50 MI pairs (Table 1), including one group where the mothers had received antibiotic treatment during pregnancy, one group where the infants received antibiotics during the first year of life, and one group where neither mother nor infant had received antibiotics. For the infants we analyzed fecal samples collected at 3, 6 and 12 months of age, while for the mothers we analyzed a single sample taken at 11-18 weeks gestational age. 23 mothers had received at least one treatment with antibiotics at some point during pregnancy (median 208 days before birth, range 95-279), taken to be within 280 days before giving birth, the mean gestational age. Each of these women received penicillin during that period. 12 infants received antibiotics during the first year of life (median 164 days after birth, range 43-337), with infants receiving beta-lactam (n=8), macrolide (n=4), diaminopyrimidine (n=2) or sulfomamide (n=2) types of drugs, either alone or in combination. 8 infants were delivered by C-section and 42 vaginally. 31 of the infants were boys while 19 were girls. Only 3 infants did not receive at least some breast milk until 6 months of age (data missing for two infants).

Sample type	# samples	Sex		Delivery mode	
MI pairs (50)	168	Girl	Boy	Vaginal	C-section
Maternal samples	44				
3 month samples	42	17	25	35	7
6 month samples	37	10	27	30	7
12 month samples	45	17	28	39	6

Table 1. Metadata of the sample set.

DNA sequencing and bioinformatic processing

DNA from fecal samples was extracted with the MagAttract PowerSoil DNA EP kit (QIAGEN), and sequencing library preparation was carried out using a tagmentation procedure [18]. Sequencing was done on an Illumina NovaSeq X apparatus using three lanes on a 25B flowcell.

Raw sequence reads were trimmed and filtered with Trimmomatic v0.39 [19]. Reads mapping to the human genome sequence CRCh38.p13 were identified and removed with Bowtie2 v2.5.1 [20]. Taxonomic assignment was carried out with MetaPhlAN 4 [21]. Reads were mapped to the Comprehensive Antibiotic Resistance Database (CARD) v3.3.0 [22] and the Virulence Factor Database (VFDB; downloaded September 27 2024) [23] with BWA v0.7.18 [24]. Before mapping of reads to the databases redundancy in the databases was removed with the cluster_fast function in VSEARCH v2.25.0 [25] using the –centroids argument with –id set to 0.99. Alignment coverage and the percent mapping rate were computed with the 'coverage' and 'flagstat' functions in SAMTools v.1.18 [26], respectively. Sourmash analysis was carried out with sourmash v4.8.14 [27], using the R1 files for each sequencing library, k-mer size set to 31 and the scaled argument set to 1000. Coverage data were normalized to reads per kilobase per million mapped reads (RPKM), i.e. (mapped reads * 10^9) / (gene length * total reads), prior to downstream analysis. All downstream analyses were done using R [28].

Plasmid assembly and binning

Clean and trimmed reads were assembled for circular plasmids using metaplasmidSPAdes.py from SPAdes v3.15.3 [29]. Assembled circular contigs were assigned as plasmids using end-to-end geNomad 1.7.4 with default settings [30]. Using the default settings of MobMess v1 [31], plasmid contigs were clustered to a non-redundant subset of plasmid representatives per infant. In addition, MobMess placed similar plasmids containing the same plasmid maintenance genes but different cargo genes in evolutionary plasmid systems.

Protein annotation and mobility class of plasmids

Resulting plasmid references were annotated for ARGs using Resistance Gene Identifier (RGI) software v6.0.1 and CARD v3.3.0 [22]. The option --low_quality was used for potential partial genes at the start and end of the sequence, and --keep to keep found ORFs. Annotations with at least 80% fragment coverage and identity were considered hits. ORFs were further annotated with the non-redundant UniRef90 database from Universal Protein Resource (UniProt) [32] using DIAMOND v11.3.0 [33]. Annotations with at least 80% fragment coverage and an e-value < 0.001 were considered hits. In addition, ORFs were annotated with the online tool of eggNOG-mapper v5 with default settings for metagenomic contigs.

Nucleotide sequences were screened for origins of replication (DoriC) and origins of transmission (OriT) using blastn in BLAST+ v2.13.0 [34] and the DoriC and OriT databases respectively [35, 36], hits with at least 95% identity and 90% coverage were kept. In addition, CONJScan from MacSyFinder v2 [37] was used to classify plasmids that can conjugate (CONJ), are mobile with the help of another residing CONJ plasmid (MOB) or are not mobile (MOBless). Plasmids assigned MOBless by MacSyFinder but that have an OriT were reassigned to MOB.

Plasmid abundance

Reads were mapped to the MI pairs' own plasmid reference using BWA mem v0.7.17 [24]. Mapped reads were filtered, sorted, indexed, and coverage was calculated using SAMtools v1.17 [26], proper paired reads with at least 120 bp alignment (-f 2 -e 'rlen>120') were kept. Plasmid presence was decided for plasmids with at least 80% coverage and a minimum read depth of 10. Abundance was calculated as RPKM as above (with gene length substituted with plasmid size).

Taxonomy of putative plasmid host

Putative plasmid hosts were found as recently described [38]. Shortly, taxonomic assignment was based on a combination of the following methods. Contained plasmids within the plasmid database v2021_06_23_v2 (PLSDB) [39] were found using mash screen v2.3 [40], with the options -v 0.1 -i 0.99 -w. Taxa from annotated proteins from UniRef90 were counted per plasmid. Highest number of ORFs belonging to a certain genus was used as plasmid taxa, no taxa was assigned in the case of an equal count. Plasmid sequences were classified with Kraken v2 [41], using the Genome Taxonomy Database (GTDB) [42], specifically, the previously separated databases for chromosomal and plasmid sequences (GTDBp and GTDBc) [43]. Plasmids from PLSDB that were found in more than one taxonomic genus were assigned as putative Broad Host Range (pBHR) plasmids.

Phylogeny of plasmids

Nucleotide sequences of similar plasmid bins were aligned to start with the same protein. Multiple sequence alignment of the linearized plasmids or protein coding sequences was done with kalign v3.3.5 [44]. A phylogenetic tree was calculated with FastTree v2.1.11 and the options -gtr -gamma (and -nt for nucleotide sequences) [45]. The trees were visualized with the R library ape [46].

Results

Taxonomic composition

DNA was successfully extracted from 168 samples that were then prepared and sequenced, with 44 samples from mothers and 42, 37 and 45 from infants at 3, 6 and 12 months, respectively. The sequencing output, after filtering of low quality or host derived reads, had a mean of 82,327,281 reads (± 60,131,446 s.d; median 76,756,560; range 3,084,777-442,004,019).

On the phylum level, Firmicutes was the overall most prevalent taxon. This phylum steadily increased from a mean of 18.0% at 3 months to 62.6% in the mothers, with a similar, albeit less pronounced trend seen for Bacteroidetes (Fig. 1). Actinobacteria showed the opposite trend, decreasing from 56.5% at 3 months to 8.3% in the mothers, with a similar but weaker trend for Proteobacteria.

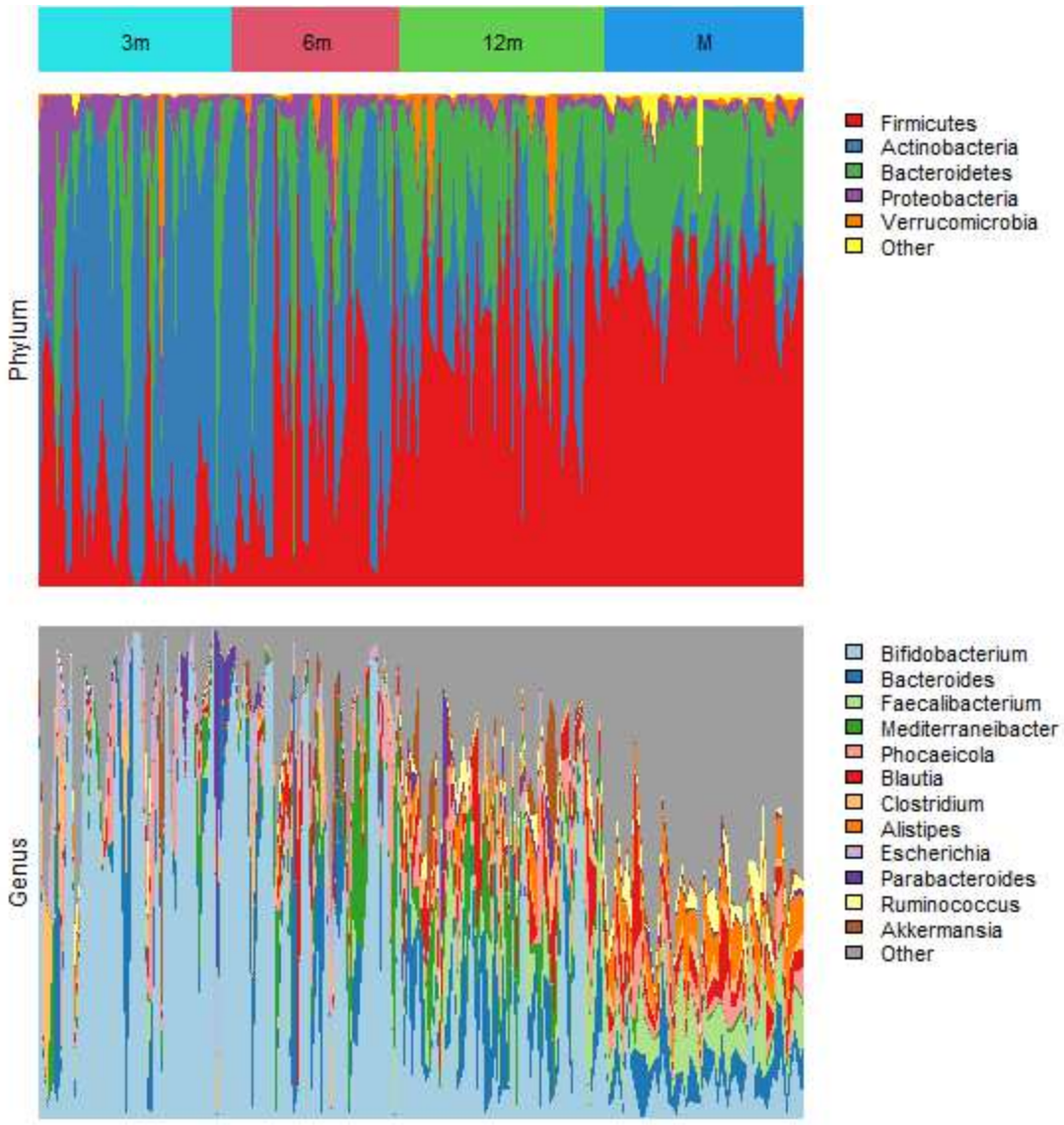


Fig. 1. Phylum and genus level composition in all 168 samples in each of the four sample groups, indicated by the colored rectangles at the top of the figure.

On the genus level, Bifidobacterium showed the highest overall prevalence, steadily decreasing with increasing age from a mean of 54.1% at 3 months to 6.5% in the mothers (Fig. 1). Several prominent genera of Firmicutes, like *Faecalibacterium*, *Roseburia*, *Blautia* and *Ruminococcus* were increased in 12-month samples and mothers relative to 3- and 6-month samples.

Unsurprisingly, there was a strong effect of infant age, and maternal provenance, on the taxonomic composition in the fecal samples ($p<0.001$, $R^2=0.18$, PERMANOVA; Fig. 2a). Between-sample distances were highest at 3 months (Fig. 2b), gradually declining towards a relatively tight cluster in maternal samples. Taxonomic diversity was slightly higher at 6 compared to 3 months ($p=0.018$, Wilcoxon rank sum test with Benjamini-Hochberg p-value correction), only to increase drastically at 12 months, and in maternal samples Shannon entropy was increased by a factor of 2.7 relative to 3 months and 1.4 relative to 12 months (Fig. 2c). Gender or delivery mode did not have significant effects on species composition in any age bracket ($p>0.3$ for all tests, PERMANOVA).

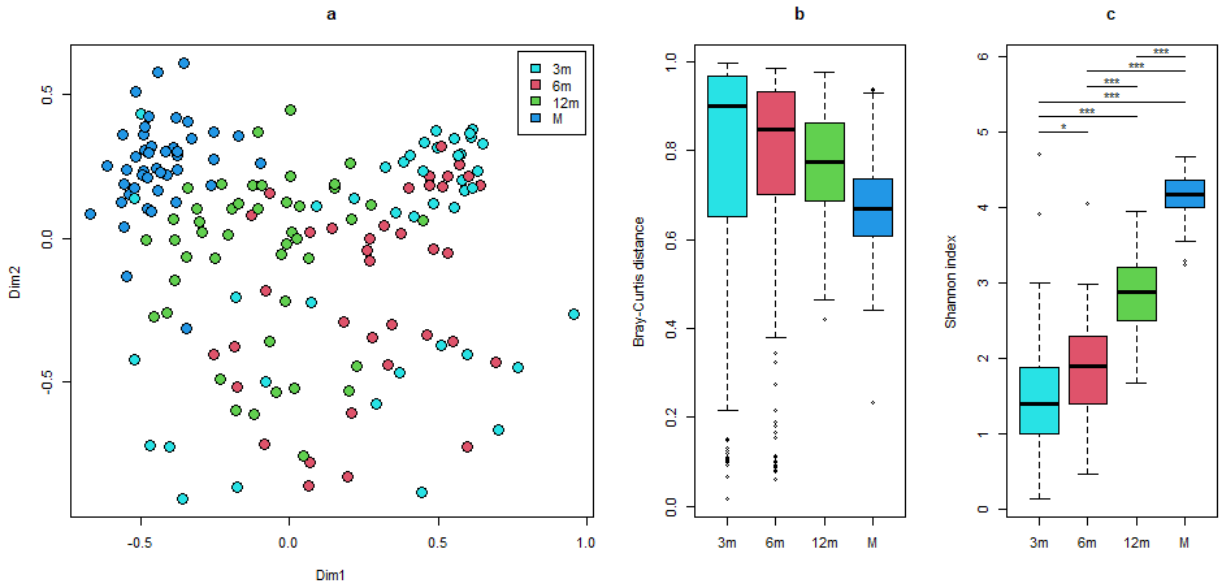


Fig. 2. a). NMDS plot based on Bray-Curtis distances of species abundance profiles for all 168 samples. b) Between-samples Bray-Curtis distances of species assignments. Each comparison is highly significant ($p<0.001$, Wilcoxon rank sum test with Benjamini-Hochberg p-value correction). c) Species diversity as measured by Shannon entropy. Asterisks above the boxes indicate the level of significance (*<0.05, **<0.01, ***<0.001, n.s.=non-significant. Statistical results are from Wilcoxon rank sum tests with Benjamini-Hochberg p-value adjustment). In the color key and x-axes labels M=mothers. The box plots

show the interquartile range (boxes), the median (lines within boxes), and values falling within 1.5x the interquartile range (whiskers). Dots outside the whiskers are outliers.

ARG content

The ARG profiles of individual samples were strongly structured by sampling age or maternal provenance ($p<0.001$, $R^2=0.27$; Fig. 3a), with progressively lower between-sample Bray-Curtis distances with age, and the tightest clustering in the maternal samples (Fig. 3b). Furthermore, Shannon entropy increased with age, with the highest diversity seen in the maternal samples (Fig. 3c). However, mapping rates to CARD decreased over time, only to reach a level comparable with maternal samples at 12 months (Fig. 3d). Mode of birth or gender did not have significant effects on CARD mapping rates in infants in any of the age groups ($p>0.1$ for all comparisons, Wilcoxon rank sum test), nor did they have significant effects on ARG composition ($p>0.2$ for all comparisons, PERMANOVA).

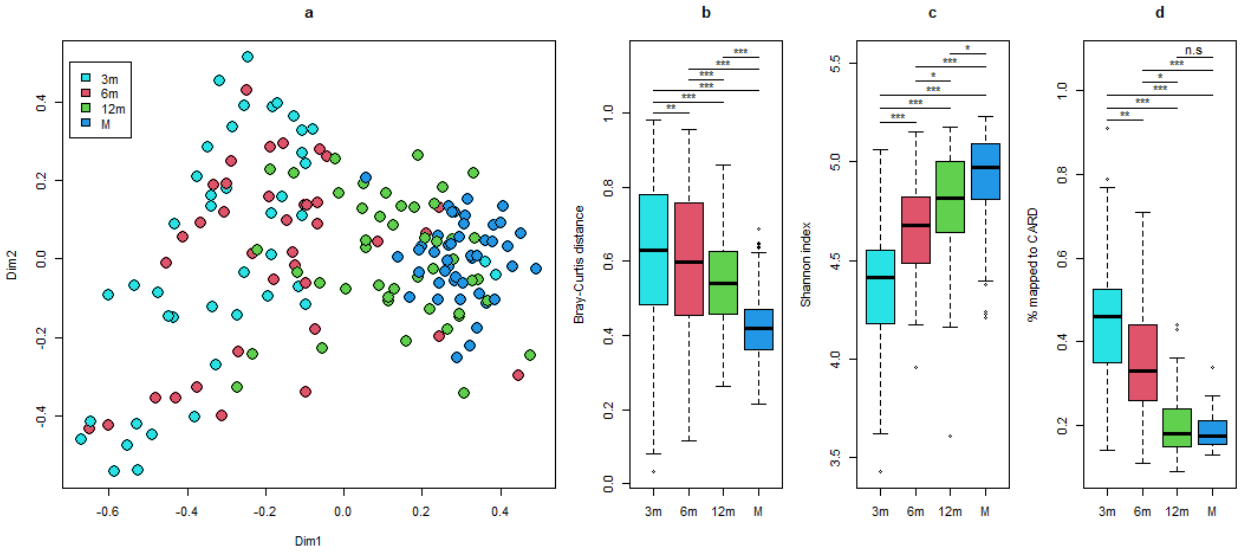


Fig. 3. a) NMDS plot based on Bray-Curtis distances of ARG abundance profiles. b) Between-samples Bray-Curtis distances of ARG profiles. c) ARG diversity as measured by Shannon entropy. d) Percentage of reads mapped to CARD. Significance testing and indication is as in figure 2.

In the 3- and 6-month-old infants, beta-lactam ARGs were the most prevalent, followed by fluoroquinolone ARGs (Fig. S1a and b). By the age of 12 months tetracycline ARGs had overtaken fluoroquinolone ARGs as the second most prevalent class, with fluoroquinolone ARGs in fifth place (Fig. S1c). In the maternal samples tetracycline ARGs were the most prevalent class, followed by beta-lactam ARGs, with fluoroquinolone ARGs in seventh place (Fig. S1d). The beta-lactam class was composed of six subclasses, whereof penam was the most prevalent in infants of all ages, while cephalosporin was narrowly more prevalent in maternal samples (Fig. S2).

ARGs conferring resistance to most major classes of antibiotics followed the same pattern of decreasing abundances from 3 to 6 to 12 months, with even lower levels observed in the maternal samples (Table 2; Fig. S3). Going from 3 to 6 months, variations in carriage among individuals were large and decreases were mostly not significant. An exception was rifamycin resistance genes, whose abundance decreased significantly over this period. Going from 3 to 12, as well as 6 to 12 months, genes within most classes saw significant decreases in abundance. Similarly, most ARG classes were significantly decreased in maternal samples relative to infants at 3 and 6 months. This was the general trend, albeit weaker, when comparing infants at 12 months to maternal samples. ARGs conferring resistance to a few classes of antibiotics broke with the general trend. The relatively rare ARG classes conferring resistance to lincosamide, sulfonamide and glycopeptide antibiotics did not differ significantly in abundance in infants going from 3 to 6 to 12 months, or between infants and mothers. Interestingly, while ARGs conferring

resistance to tetracycline followed the general trend in the infants, maternal levels were significantly increased relative to infants at 12 months.

	Mean RPKM				P-value					
	3m	6m	12m	M	3m vs. 6m	3m vs. 12m	3m vs. M	6m vs. 12m	6m vs. M	12m vs. M
Beta-lactam	1296,14	823,77	258,26	141,37	0,300	<0.001	<0.001	<0.001	<0.001	<0.001
Tetracycline	368,71	263,62	158,90	206,92	0,241	<0.001	0,050	0,010	0,382	0,004
Fluoroquinolone	488,66	326,74	94,33	41,28	0,485	<0.001	<0.001	<0.001	<0.001	0,001
Rifamycin	368,60	272,91	104,16	56,48	0,010	<0.001	<0.001	<0.001	<0.001	<0.001
Macrolide	315,14	234,82	99,47	73,76	0,491	<0.001	<0.001	0,004	<0.001	0,104
Aminoglycoside	196,31	141,32	51,74	37,42	0,335	<0.001	<0.001	<0.001	<0.001	0,059
Chloramphenicol	162,36	111,74	37,33	23,76	0,431	<0.001	<0.001	<0.001	<0.001	0,020
Aminocoumarin	143,13	95,57	26,59	9,42	0,480	<0.001	<0.001	<0.001	<0.001	<0.001
Lincosamide	48,78	61,16	51,53	59,28	0,500	0,300	0,190	0,300	0,190	0,300
Nucleoside	99,22	77,19	22,82	8,72	0,981	0,002	<0.001	<0.001	<0.001	<0.001
Glycopeptide	15,60	18,80	21,24	19,98	0,177	0,052	0,052	0,459	0,459	0,841
Sulfonamide	15,59	16,69	3,84	2,81	0,980	0,980	0,980	0,980	0,980	0,980

Table 2. Mean normalized abundances of ARGs conferring resistance to major classes of antibiotics in infants at 3-, 6- and 12 months, and maternal samples (M), as well as pairwise comparisons of abundances in each sample category (Wilcoxon rank sum tests with Benjamini-Hochberg adjusted p-values).

In infants at 3- and 6- months, antibiotic efflux was the main mechanism of resistance, followed by target alteration and replacement, respectively (Fig. S4a and b). At 12 months target alteration was a marginally more prevalent mechanism than efflux (Fig. S4c), while in the maternal samples target protection was the main mechanism of resistance, followed by target alteration, drug inactivation and efflux (Fig. S4d).

Out of a total of 290 ARGs identified in the dataset, 49 were found in all 168 samples sequenced. In infants at 3 months, the core resistome consisted of 54 genes, 90 genes at 6 months, 97 genes at 12 months, and 154 genes in the maternal samples. This is in keeping with the decreasing dispersion seen both in species composition and ARG profiles. 50 ARGs were shared among all infant samples, with the

only gene found in the infant core resistome and not in the maternal core being *leuO* (ARO:3003843), although this gene was found in 43 of 44 maternal samples.

Overall, there was a strong correlation between the relative abundances of both Proteobacteria and *E.coli* and mapping rates to CARD (R^2=0.59 and 0.61, respectively, p<0.001 for both tests, Pearson's product-moment test). Looking just at infants in the 3- and 6-month age groups these correlations were also apparent (p<0.001 for all correlations, Table S1), while at 12 months, correlations were non-significant, and associations were not discernable at all in the mothers.

23 mothers had received antibiotic treatment during the first six months of pregnancy, including at least one penicillin treatment. In order to investigate the effects of maternal use of antibiotics during pregnancy on the infant gut microbiome, children of mothers that had used antibiotics were compared with children whose mothers had not used antibiotics. We did not observe significant effects on the species composition in the microbiota in any infant age bracket (p>0.35 for all three comparisons, PERMANOVA), nor did we see significant effects on the composition of the resistome (p>0.5 for all three comparisons, PERMANOVA). We also did not observe significant differences in CARD mapping rates between the two infant groups at 3 or 6 months, and only a marginally lower mapping rate at 12 months in infants whose mothers had not used antibiotics (p=0.056, Wilcoxon rank sum test; Fig. S5). Further, we did not find any difference in general beta-lactam resistance carriage in any age bracket, nor did we find significant differences in the carriage of 6 sub-classes of beta-lactam (carbapenem, cephalosporin, cephamycin, penam, penem and monobactam).

Infants that had received antibiotics during the first year was a considerably smaller and less homogenous group in terms of antibiotic exposure. Therefore, analysis was restricted to comparing general CARD mapping rates for investigating ARG responses to antibiotic exposures in infants. Here, we compared infants of 12 months age, that had been exposed to antibiotics at any earlier point (n=12)

with those that had not been exposed at all (n=34). We also checked for differences in CARD mapping rates between infants at 3 months age who had received antibiotics within the first 90 days of life (n=3) and 3-month-olds that had not received antibiotics during that period. We carried out similar testing for infants that had received antibiotics between 90 and 180 days (n=10), as well as 180-270 days (n=3). We did not find significant differences in overall CARD mapping rates (p=0.32), considering all 12 infants that had received antibiotics within the first year of life. We also did not find significant differences in mapping rates in 3-, 6- and 12-month-old infants that had used antibiotics within the 3 last months prior to sampling compared to those who had not (p>0.35 for all comparisons).

VFG content

Mapping of reads to the VFDB found 602 VFGs. VFG and ARG carriage were highly correlated with a Pearson's R^2=0.83 (Spearman's Rho=0.87) between the mapping rates to their respective databases. The high correlation was observed in all infant age brackets, as well as in the adults (R^2=0.79, 0.83, 0.64 and 0.88, for 3-, 6- and 12-month-old infants and mothers, respectively, p<0.001 for all correlations by Pearson's product-moment tests). As for ARGs, the VFG profile of a sample showed strong structure according to sampling age or maternal provenance (p<0.001, R^2=0.22, PERMANOVA; Fig 4a). Between-sample distances showed a strong decline over time, with much smaller distances between maternal samples relative to infants of any age (Fig. 4b), while VFG diversity showed a clear increasing trend (Fig. 4c). Similarly to ARGs, the mapping rates to VFDB declined over time and appeared to stabilize at maternal levels at 12 months (Fig. 4d). Mode of birth or gender did not significantly impact mapping rates to VFDB in any of the age brackets (Wilcoxon rank sum test, p>0.5 for all comparisons), and we did not see significant effects on VFG composition (p>0.3 for all comparisons, PERMANOVA).

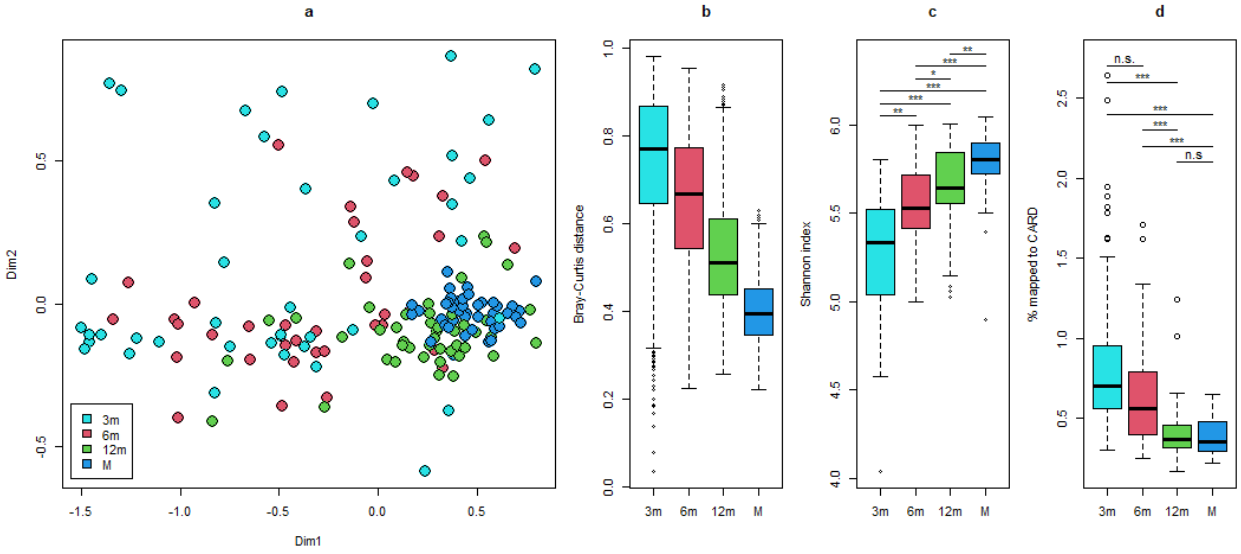


Fig. 4. a) NMDS plot based on Bray-Curtis distances of VFG abundance profiles. b) Between-samples Bray-Curtis distances of VFG profiles. Each comparison is highly significant ($p<0.001$, Wilcoxon rank sum test with Benjamini-Hochberg p-value correction). c) VFG diversity as measured by Shannon entropy. d) Percentage of sequence reads mapped to VFDB. The difference between 3 and 6 months was marginally significant at $p=0.052$. Significance testing and indication is as in figure 2.

The main virulence mechanism in infants at 3- and 6- months were metabolic factors, followed by cell adherence, effector delivery systems and exotoxins (Fig S6a and b). At 12 months effector delivery systems were the main mechanism, followed by metabolic factors, cell adherence and immune modulation (Fig. S6c), while in maternal samples effector delivery systems were most prevalent, followed by metabolic factors and immune modulation (Fig S6d).

Similarly to ARGs, VFGs within the main mechanistic classes decreased in infants going from 3 to 6 to 12 months (Fig. S7), although in no case was the difference between abundances at 3- and 6 months statistically significant (Table S2). The maternal samples generally had the lowest abundances across classes. Two exceptions were the categories stress survival and motility, each represented by a single

297 gene. The *katA* gene encoding a catalase associated with *Neisseria meningitides* displayed increased
298 abundance (Table S2) in maternal samples relative to infants at 12 months. The *FlgN* gene, associated
299 with *Pseudomonas aeruginosa,* was significantly decreased in infants at 3 and 6 months compared to
300 infants at 12 months and maternal samples (Table S2).

301 96 out of a total of 602 detected VFGs were found in all 168 samples. In infants at 3 months of age 111
302 VFGs were shared by all infants in that group, while the corresponding number at 6- and 12 months
303 were 232 and 246. In the maternal samples 319 VFGs were ubiquitous.

304 As for ARGs, mapping rates to VFDB showed highly significant correlations with relative abundances of
305 Proteobacteria and *E. coli* at 3- and 6- months of age ($p<001$ for all correlations, Pearson's product-
306 moment test, Table S3), but unlike for ARGs this pattern persisted, although weakened, in samples taken
307 at 12 months ($p<0.001$ and $p=0.015$ for Proteobacteria and *E. coli*, respectively). In the mothers, there
308 was absolutely no association between VFDB mapping rates and Proteobacteria and *E. coli* abundances
309 (Table S3).

310 **Similarity between infants and mothers**

311 A decrease in the Bray-Curtis distance over time between the microbiomes of MI pairs was observed
312 with respect to the species composition, resistome and VFG profiles (Fig. 5). By 12 months, infants had
313 become significantly more similar to their mothers with respect to all three categories relative to at 3-
314 and 6- months. Somewhat surprisingly, at no age did the infants resemble their own mothers more than
315 the unrelated mothers, when comparing within group Bray-Curtis distances, with respect to any of the
316 three categories ($p>0.6$ for all comparisons, Wilcoxon rank sum test; Fig. S8). This result was maintained
317 when using Jaccard distances as an alternative metric ($p>0.5$ for all comparisons; Fig. S9a-c). When using
318 a database independent metagenome distance measure with Jaccard distances computed from k-mer

frequencies, we still did not observe statistically significant familiarity (p>0.1 for all comparisons; Fig. S9d).

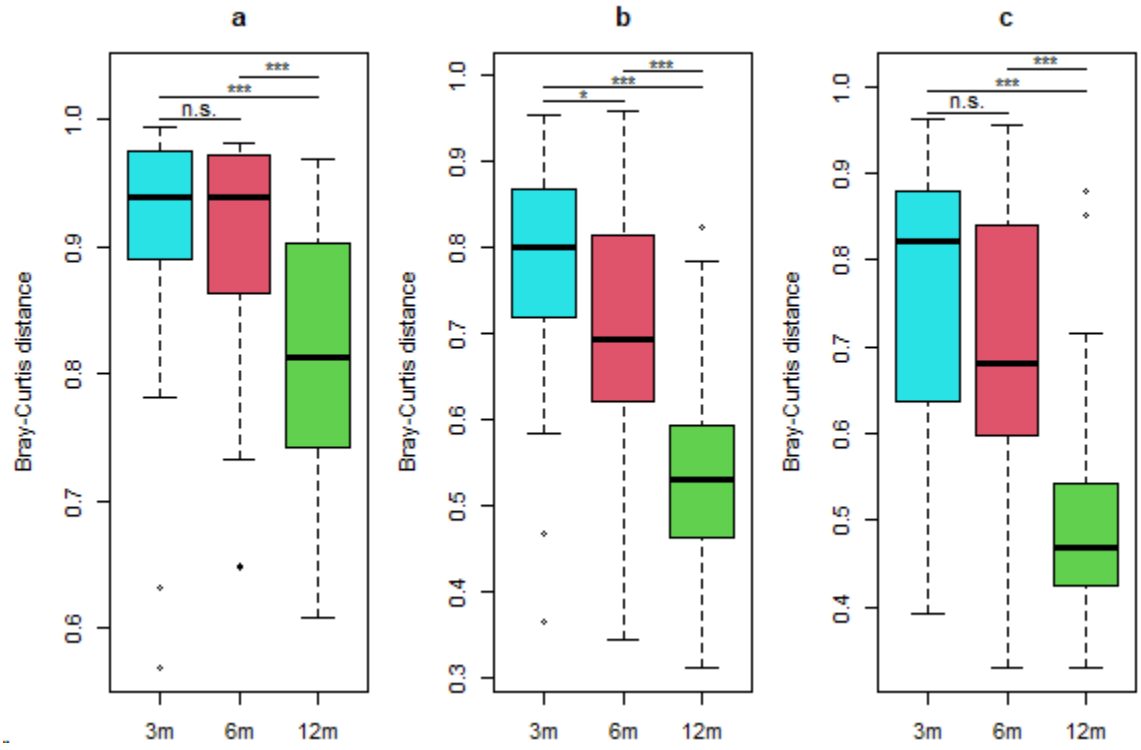


Fig. 5. Bray-Curtis distances between infants and their mothers based on a) species composition, b) ARG profiles c) VFG profiles. In the case of VFGs the difference in distance to mothers at 3 and 6 months was marginally significant at p=0.061. Significance testing and indication is as in figure 2.

Plasmid composition

Of the 4805 circular contigs assembled by metaplasmidSPAdes, 1937 were categorized as plasmid by geNomad. These plasmid contigs were clustered into 740 clusters and 82 systems using MobMess, with a mean of 31.28 clusters (sd 10.04, range 13-53) per MI pair. Most of the clusters were transient, while 4 clusters appeared in all 4 samples of a MI pair (Fig S10a). Interestingly, only a few plasmids were shared

between mother and infant at all ages (Fig S10b-d). However, a slight non-significant decline of between-sample distances between infants and their mothers, and a non-significant increase of shared plasmids, was seen at 12 months of age.

Annotation, classification, and taxonomy of the putative host of all plasmid clusters can be found in the supplement (Table S1-2). In short, 381 plasmids were found to be mobilizable (MOB), 38 plasmids had genes encoding for a complete conjugation machinery (CONJ), and 321 plasmids cannot be mobilized (MOBless). Putative Broad Host Range (pBHR) plasmids within the Bacteroidales order were the most abundant, with the genera *Bacteroides*, *Phocaeicola*, and *Parabacteroides* providing the highest plasmid abundances when regarding genera only (Table S3). Other putative plasmid hosts of high abundant plasmids were *Collinsella*, *Bifidobacterium*, and pBHR Enterobacterales (mostly from *Escherichia*, *Klebsiella*, and *Citrobacter*). Plasmid diversity was highest in *Collinsella*, but none of these plasmids were categorized as pBHR (Table S4). *Escherichia* was predicted as host for a similar number of plasmids as *Collinsella*, although many of these plasmids had several putative hosts and were grouped as pBHR Enterobacterales or pBHR other. In addition, pBHR Bacteroidales, *Bifidobacterium*, and *Clostridium* hosted a diverse set of plasmids.

A significant separation of plasmid composition was seen between the different timepoints ($p<0.001$, $R^2=0.031$, PERMANOVA; Fig. 6a). Between-sample distances declined with infant age, while plasmid diversity increased from 3- to 12- months, with the respective lowest and highest values in the maternal samples (Fig. 6b and c). In addition, plasmid abundance was significantly higher in the infant samples compared to the maternal samples (Fig. 6d).

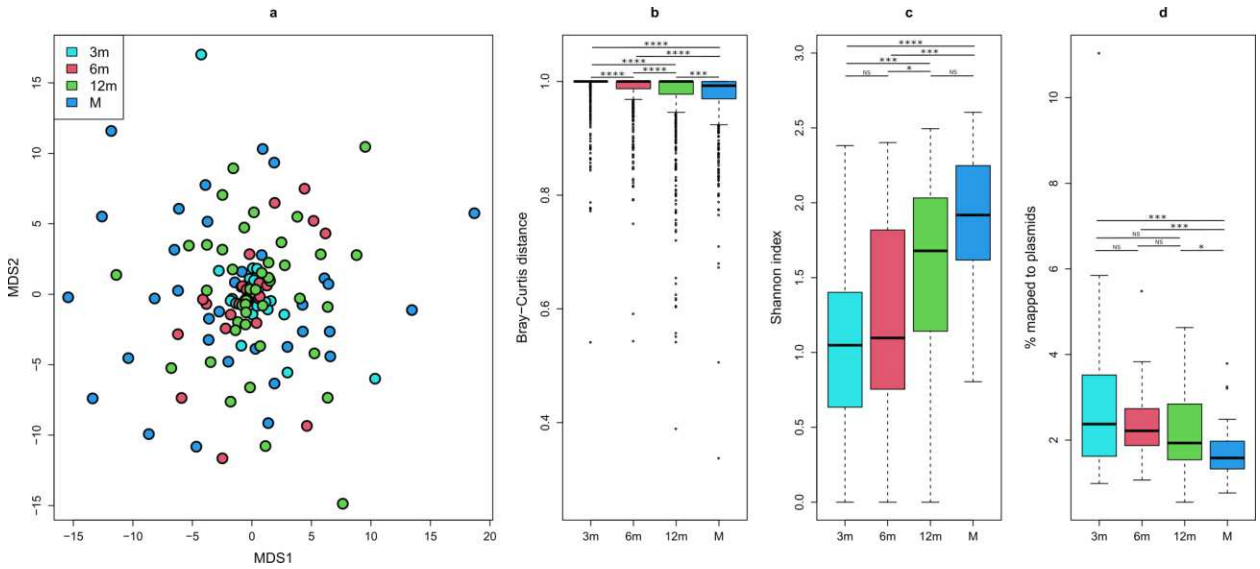


Fig. 6. a) NMDS plot based on Bray-Curtis distances of plasmid presence/absence profiles, colored by sampling timepoint. b) Between-samples Bray-Curtis distances of plasmid profiles. c) Plasmid diversity as measured by Shannon entropy. d) Percentage of reads mapped to own MI pair plasmid reference. Significance testing and indication is as in figure 2.

ARG plasmids

In total, 43 plasmids were found to carry at least one ARG, and ARG carrying plasmids were spread across 33 MI pairs. The most widespread ARG plasmids were clusters 667 and 680, carried by 14 and 12 MI pairs respectively. Interestingly, all found contigs of cluster 667 were 100% identical to each other. In addition, 100% identical plasmids were previously found in rat gut metagenomes from a Danish hospital sewer (Fig S11) [64, 65]. Another plasmid, similar to 667 but with an additional insertion sequence, was found in a *Parabacteroides faecis* isolate [66]. Plasmid system PS8/PS64 was very similar across most MI pairs, except for pair 26 which contained an additional insertion sequence (Fig S12). A similar plasmid was also assembled from the previously mentioned rat gut metagenome [64]. The plasmid could not be

366 linked to any putative host but was classified as MOB, suggesting conjugation potential with the help of
367 a second plasmid.

368 The most interesting case, plasmid cluster 689, carried a *Tet(W)* gene in 5 MI pairs, but is part of plasmid
369 systems PS58 and PS67 containing 4 clusters found in 18 MI pairs (Fig. 7). Three of the non-*Tet(W)*
370 carrying plasmids had a deletion in the same region where *Tet(W)* was inserted in the other plasmids,
371 which might be a recombination site of this plasmid system. When comparing the assembled contigs
372 within MI pairs, the infant of MI pair 11 at 6 months carried the plasmid with the *Tet(W)* gene (689),
373 while at 12 months the plasmid without the *Tet(W)* gene was assembled (649). A similar case was found
374 in MI pair 44, where the mother carried the *Tet(W)* gene, but the infant at 3 months old did not.
375 Comparing the full-length sequences of all contigs from plasmid system PS58/67 with each other
376 showed that the plasmids within the two MI pairs were more similar to other clusters of the system than
377 to each other (Fig S13a). This suggests that the plasmid was not shared within the two pairs but was
378 carried by a different bacterial host species or strain. In comparison, the two cases of assembly of
379 plasmid 680 within a MI pair were identical to each other, suggesting sharing of the plasmid between
380 mother and infant (Fig S13b).

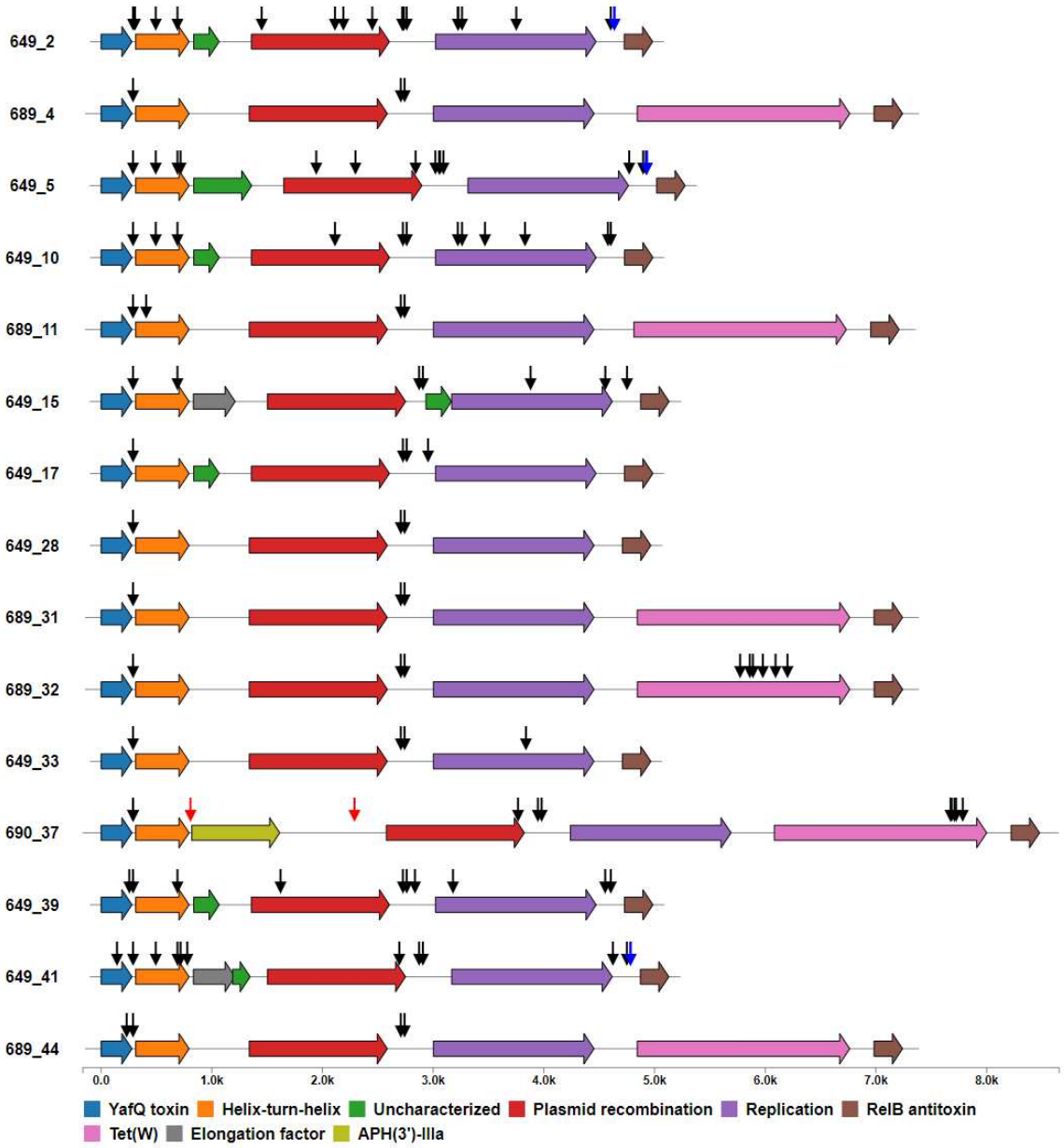


Fig. 7. Genetic map of plasmid system PS58/PS67. The plasmid was found in 15 MI pairs out of 50 total. Based on mapping data, cluster 689 was found nine times in both mother and infant, one time in only the infant, and five times only in the mother. Six out of 15 pairs carrying this plasmid carried the Tet(W) gene, the others are without ARGs. Cluster 690 of MI pair 37 had an insertion with an additional ARG. Clusters 689 and 690 of MI pairs 32 and 37 had nonsynonymous single nucleotide variants in the Tet(W) gene. The plasmid of MI pairs 2, 5, and 41 had small deletions on a similar position where the Tet(W)

was inserted in some of the other plasmids. Based on UniRef the host of this plasmid might be *Holdemania* or *Bacteroides* (or *Bifidobacterium* in 37). Where there is a black arrow in all sequences, there were two common nucleotides in the position compared to all other sequences. Black arrow is a single nucleotide variation, red arrow is an inserted sequence (compared to the other contigs), blue arrow is a deletion position.

In two of the plasmids from the previously mentioned plasmid system PS58/67, nonsynonymous single nucleotide variants were found within the *Tet(W)* gene. The other four *Tet(W)* genes were identical to CARDs *Tet(W)* amino acid sequence. Using Uniprot's online blast tool, we aligned all 122 protein sequences with at least 90% identity to CARD's Tet(W) sequence (Fig S14). Almost half of the found proteins were identical to *Tet(W)* and clustered together in the phylogenetic tree (Fig. S15). *Tet(W)* on cluster 689 in MI pair 32 was closest related to a predicted protein from *butyrate-producing bacterium SM4/1* and *TetM* of *Berryella intestinalis*. The cluster 690 *Tet(W)* of MI pair 37 was most closely related to a group of *Tet(W)* genes from three *Eubacterium* species and *Qingrenia yutianensis*.

Discussion

Here, demographic trends were in line with previous studies [5, 47], with species diversity increasing over the first 12 months of life, and taxonomic composition approaching an adult profile dominated by Firmicutes and Bacteroidetes [48, 49]. This development was accompanied by decreasing β-diversity with age, with much lower dispersion in the maternal samples compared to infants at 12 months. This phenomenon has been observed by others [49, 50], and indicates an early stage characterized by unstable dynamics and transient colonization. We did not find significant effects of gender on the taxonomic composition of the microbiota, which is in line with findings in a larger infant cohort [48].

That same study, however, showed significant differences between the developmental trajectories of the microbiomes of infants delivered vaginally or by caesarian section, as have other studies [50, 51]. We failed to observe such an effect in any age bracket, possibly due to the small number of caesarian-delivered infants in the cohort.

ARGs followed the same trends seen for taxonomic composition and diversity, with decreasing β-diversity and increasing α-diversity in the infants, and the lowest and highest levels, respectively, observed in the mothers. The total ARG carriage, on the other hand, decreased to a level comparable to the mothers by 12 months of age. The heightened carriage in infants relative to adults has been reported elsewhere [2, 52], as has the general decrease in carriage as the infants age [4, 53].

In a previous study of a smaller cohort of healthy, born-to-term infants, we found a significant negative, albeit non-linear, trend in ARG diversity over the first year of life [4]. A previous study of MI pairs did not find significant differences in ARG diversity in infants at 1- and 6 months, nor between infants and mothers [2]. Yet another recent study found significantly higher ARG richness in adults compared to infants [54], in line with the results presented here. We do not have a good answer for these discrepancies, other than possibly cohort differences related to size and sampling scheme. The results obtained here do seem quite robust, but further studies will be needed in order to confirm typical time trends in ARG diversity during infancy and the relationship between the infant and adult resistome in this regard.

Some rare ARG classes conferring resistance to glycopeptides and sulfonamides did not show decreasing abundances over time, or in infants relative to mothers. This is similar to Trosvik et al. [4], where these classes showed positive time trends in a cohort of 12 infants sampled on a monthly basis over the first year of life. Here, glycopeptide resistance, represented by a total of 22 vancomycin resistance genes, was found in all 168 samples. Total carriage showed a positive trend from 3 to 12 months, but without

significant differences in abundance between time points. Higher temporal resolution may have provided more statistical power to detect a time trend.

As we have reported previously in a smaller cohort of infants, beta-lactam resistance genes were the most prevalent class of ARGs [4], although this was no longer the case when comparing with subclasses of these genes. Similarly to Pärnanen et al. [2], tetracycline resistance was the most common ARG class observed in the mothers. Also, efflux pumps constituted the most common mechanism of resistance in infants. These genes often encode multi-drug transporter complexes associated with gram negative bacteria such as *E.coli* and other Proteobacteria. As noted by several authors [3, 4, 53] antibiotic resistance in the infant gut is strongly linked with Proteobacteria, and particularly *E. coli*. Here, we also found a strong general correlation between CARD mapping rates and *E. coli* relative abundance. Interestingly, this association was strong in infants at 3- and 6- months of age, but much weakened at 12 months. In the mothers there was no discernable correlation between ARG abundances and abundances of these taxa. This does not necessarily mean that ARGs are less prevalent in *E. coli* strains associated with more mature infants and adults, but rather that *E. coli* is less abundant, and that other taxa play a more important role in the resistome.

Although the core resistome is not a well-defined concept, but it's reasonable to assume that many ARGs confer resistance only incidentally and may have additional cellular functions that are not specifically related to resistance to a particular antibiotic [55]. For example, the gene encoding 2'-N-acetyltransferase in *Providencia stuartii* confers intrinsic resistance to aminoglycoside because the enzyme's natural substrate, peptidoglycan, resembles that antibiotic [56]. Other examples are multi-drug transporters such as the TolC membrane channel protein that functions in conjunction with a number of efflux complexes in *E.coli*, like AcrAB-TolC, and which plays a role in the transport of numerous substrates, including many antibiotics, out of the cell [57]. Here, we found that 49 ARGs were shared between all infant and adult samples, 32 of which form part of efflux systems. In fact, only a

458 single gene was found in the infant core resistome that was not also found in all of the maternal
459 samples, namely *leuO*, the product of which is an activator of the mdpNOP efflux system that confers
460 resistance to nucleoside antibiotics and is widespread in gamma Proteobacteria. The product of this
461 gene is also involved in regulation of leucine biosynthesis, and possibly other cellular functions [58]. It
462 should be noted that this gene was found in 43 out of 44 maternal samples. The finding that the infant
463 core resistome is basically a subset of a substantially larger adult core resistome suggests that some of
464 these genes may confer a resistant phenotype incidentally. A previous study of a smaller cohort of
465 infants sampled monthly for the first year of life identified a core resistome of 47 ARGs [4], 27 of which
466 were shared with the 50 ARG infant core resistome found here, indicating a reasonable degree of
467 agreement between the studies.

468 We were not able to find significant effects of antibiotic use during the two first trimesters of pregnancy
469 on the infant microbiome taxonomic and ARG composition, nor on ARG abundance. This is perhaps not
470 so surprising given the separation in time between usage and the points of fecal sampling from the
471 infants. Antibiotic use during any trimester of pregnancy has been found to alter the vaginal microbiome
472 by promoting increased colonization of *Staphylococcus spp. [59]*. It is, however, not clear how sustained
473 these changes are or how they may affect the infant microbiome upon birth. A systematic review from
474 2021 identified a few studies of the effects of antibiotic use during pregnancy, identifying, among other
475 things, a decrease in *Bifidobacterium* colonization relative to infants whose mothers had not received
476 antibiotics [6]. This meta-study also included interpartum antibiotic use, which was the case for the
477 majority of the studies included, and potential effects of exposure earlier during pregnancy on the infant
478 microbiome is clearly a data deficient area of study. The fact that we were not able to detect any effects
479 of antibiotic exposure in infants is perhaps less expected. ARGs have been found to be enriched in the
480 infant GI microbiome upon treatment with antibiotics [60], although a recent study found no effect on
481 ARG abundance and minimal impact on the infant microbiome as a whole upon exposure to

predominantly amoxicillin [61]. Yassour et al. [51] found an increase in chromosomally encoded ARGs during antibiotic treatment which subsided rapidly upon cessation, although this was not the case for episomally encoded ARGs.

Here, we did not observe any effects of antibiotic use during any stage of the 12 first months of life on ARG carriage in infants at 1 year of age. This may also be explained by the prolonged interval between the time of exposure and sampling. Furthermore, we did not observe differences in ARG abundance in infant samples from individuals that had been treated within the last three months before the time of sampling. Again, this fits with a scenario of swift microbiome recovery upon cessation of treatment. It should also be noted that the number of individuals in the various treatment groups was small, which is likely to make detection of significant differences unlikely even in the case of underlying effects.

We found that the presence of VFGs was closely linked with ARG abundances, and we observed the same patterns in terms of temporal trends of dispersion, diversity and abundance. In a previous study we also found this to be very much the case [4], and similarly to that study we found both gene categories to be strongly linked with Proteobacteria, and *E.coli* in particular. A study of a number of microbial ecosystems found that VFGs and ARGs tend to correlate, and that they tend to increase in abundance with microbiome diversity [9]. A study of a Danish infant cohort also found ARGs and VFGs to co-localize and predominantly associate with Proteobacteria [10]. Interestingly, we found that the correlation between VFG (and ARG) abundances subsides towards the first year of age, and it was no longer observable in the maternal samples, indicating that a more diverse group of taxa defines the complement of VFGs and ARGs in adults.

Another somewhat surprising finding in the present study was that infants did not resemble their own mothers more than they did the mothers of other infants. While the infants gradually developed toward a general maternal microbiome profile in terms of the composition of species, ARGs or VFGs, they did

not show significant familiar resemblance in any age group (Fig S6). This is in contrast to findings by Pärnanen et al. [2], who found significantly higher sharing of ARGs and species between MI pairs relative to unrelated infants and mothers. In that study the authors compared 16 MI pairs with both individuals sampled one month postpartum, and distances among samples were compared with respect to species and ARG composition. Here, all maternal samples were collected before the third trimester, while the earliest available infant samples were collected at an age of three months. Thus, the separation in time between MI pairs may help explain the discrepancy in results between the two studies. In the study by Pärnanen et al. the authors also used a database-independent approach based on k-mer frequencies for comparing the distance between infants and related and unrelated adult females. With that same approach, we were still unable to detect significant differences in the distance between related and unrelated infant/mother pairs.

With regard to the plasmidome, very few of the plasmids were shared between mother and infant in all three age categories, although the shared plasmid count was non-significantly higher at 12 months compared to 3- and 6- months. This finding is similar to a preprint from Fernández-Pato et al. [16], which found that plasmids were mostly shared between mothers and infants after 6 months. Another recent paper on the infant plasmidome found that ~60% of the infants' replicase groups were shared with the mothers, while only ~10% was shared with unrelated adults. Interestingly, Fernández-Pato et al. showed that there was no plasmid strain sharing at early time points (<6 months) when the infant was exclusively breastfed, suggesting co-colonization of a bacterial host carrying the same plasmid rather than transfer from mother to infant during or after birth. He et al. [62] even speculated about plasmid sharing from infants to mothers, based on a higher diversity of plasmids in infants compared to mothers, while mothers had a higher diversity of plasmid compared to other adults. However, we found plasmid diversity significantly higher in mothers compared to infants, with the lowest diversity at 3 months and

increasing diversity at 6- and 12 months. These results are in-line with Fernández-Pato et al. [16] who found the diversity of plasmid taxonomic units to be higher in mothers than in infants.

Similarly to Fernández-Pato et al., we found the majority of plasmids to be non-mobilizable, with less than 10% carrying a full conjugation machinery. The most abundant plasmids in our study were predicted to be carried by *Bacteroides* and *Bifidobacterium* species. These and most other genera in our top 10 most abundant plasmid carriers were similar to He et al. [62]. The most conspicuous differences were plasmids associated with *Clostridium* and *Collinsella*, which we found to be both relatively abundant and diverse in our study, while this was not the case in the study by He et al. A possible explanation for this could be the database bias of the PLSDB and the selected genomes used by He et al., which contained relatively few isolates and plasmids from *Clostridium* and *Collinsella*. We have tried to circumvent this problem by adding putative host assignment based on taxonomic classification in the UniProt90 protein annotation.

Lastly, we assembled a few plasmids that carried one or more ARG. The surprisingly low number of ARG-carrying plasmids is likely due to the difficulty of assembling plasmids containing ARGs using only short-read sequencing [63]. Instead, we focused on the three common plasmids in our cohort that carries ARGs. Interestingly, one of these three (carrying an ErmF gene) was described before with a 100% identical sequence [64, 65] and a version isolated from a *Parabacteroides faecis* including a short insertion sequence [66]. A very similar version of our second common ARG plasmid was also previously described by the Jørgensen et al. [65]. This plasmid was identical in two MI pairs suggesting transfer from one to the other. However, our last common ARG plasmid was also found in one MI pair, as well as twice in the same infant, but these were clearly different strains of the same plasmid. Together, the study of He et al. [62], Fernández-Pato et al. [16] and our results suggest that looking into plasmid strains is necessary for inferring plasmid sharing between MI pairs when using short-read shotgun metagenomics sequencing data.

Data availability

Metagenome sequence data has been deposited at the National Center for Biotechnology Information Sequence Read Archive (accession number: PRJNA1337172) available at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1337172.

References

1. Tamburini S, Shen N, Wu HC, Clemente JC. The microbiome in early life: implications for health outcomes. Nat Med. 2016;22(7):713–22; doi: 10.1038/nm.4142.
2. Parnanen K, Karkman A, Hultman J, Lyra C, Bengtsson-Palme J, Larsson DGJ, et al. Maternal gut and breast milk microbiota affect infant gut antibiotic resistome and mobile genetic elements. Nat Commun. 2018;9(1):3891; doi: 10.1038/s41467-018-06393-w.
3. Li X, Stokholm J, Brejnrod A, Vestergaard GA, Russel J, Trivedi U, et al. The infant gut resistome associates with E. coli, environmental exposures, gut microbiome maturity, and asthma-associated bacterial composition. Cell Host Microbe. 2021;29(6):975–87 e4; doi: 10.1016/j.chom.2021.03.017.
4. Trosvik P, Noordzij HT, de Muinck EJ. Antibiotic resistance gene dynamics in the commensal infant gut microbiome over the first year of life. Sci Rep. 2024;14(1):18701; doi: 10.1038/s41598-024-69275-w.
5. De Muinck EJ, Trosvik P. Individuality and convergence of the infant gut microbiota during the first year of life. Nature Communications. **2018**.
6. Grech A, Collins CE, Holmes A, Lal R, Duncanson K, Taylor R, et al. Maternal exposures and the infant gut microbiome: a systematic review with meta-analysis. Gut Microbes. 2021;13(1):1–30; doi: 10.1080/19490976.2021.1897210.
7. Qu W, Liu L, Miao L. Exposure to antibiotics during pregnancy alters offspring outcomes. Expert Opin Drug Metab Toxicol. 2021;17(10):1165–74; doi: 10.1080/17425255.2021.1974000.
8. Barlow M, Hall BG. Phylogenetic analysis shows that the OXA beta-lactamase genes have been on plasmids for millions of years. J Mol Evol. 2002;55(3):314–21; doi: 10.1007/s00239-002-2328-y.
9. Escudeiro P, Pothier J, Dionisio F, Nogueira T. Antibiotic Resistance Gene Diversity and Virulence Gene Diversity Are Correlated in Human Gut and Environmental Microbiomes. mSphere. 2019;4(3); doi: 10.1128/mSphere.00135-19.
10. Li X, Brejnrod A, Trivedi U, Russel J, Thorsen J, Shah SA, et al. Co-localization of antibiotic resistance genes is widespread in the infant gut microbiome and associates with an immature gut microbial composition. Microbiome. 2024;12(1):87; doi: 10.1186/s40168-024-01800-5.
11. Bengtsson-Palme J, Kristiansson E, Larsson DGJ. Environmental factors influencing the development and spread of antibiotic resistance. FEMS Microbiol Rev. 2018;42(1); doi: 10.1093/femsre/fux053.
12. Acman M, van Dorp L, Santini JM, Balloux F. Large-scale network analysis captures biological features of bacterial plasmids. Nat Commun. 2020;11(1):2452; doi: 10.1038/s41467-020-16282-w.
13. Yang LL, Mai GQ, Hu Z, Zhou HK, Dai L, Deng ZQ, et al. Global transmission of broad-host-range plasmids derived from the human gut microbiome. Nucleic Acids Research. 2023;51(15):8005–19; doi: 10.1093/nar/gkad498.

14. Stockdale SR, Hill C. Incorporating plasmid biology and metagenomics into a holistic model of the human gut microbiome. Current Opinion in Microbiology. 2023;73; doi: ARTN 102307 10.1016/j.mib.2023.102307.

15. Rodriguez-Beltran J, DelaFuente J, Leon-Sampedro R, MacLean RC, San Millan A. Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. Nat Rev Microbiol. 2021;19(6):347–59; doi: 10.1038/s41579-020-00497-1.

16. Fernández-Pato A, Sinha T, Garmaeva S, Gulyaeva A, Kuzub N, Roux S, et al. Dynamics and Determinants of the Gut Mobilome in Early Life. bioRxiv. 2024.

17. Lodrup Carlsen KC, Rehbinder EM, Skjerven HO, Carlsen MH, Fatnes TA, Fugelli P, et al. Preventing Atopic Dermatitis and ALLergies in Children-the PreventADALL study. Allergy. 2018;73(10):2063–70; doi: 10.1111/all.13468.

18. Cobb L, de Muinck E, Kollias S, Skage M, Gilfillan GD, Sydenham MAK, et al. High-throughput sequencing of insect specimens with sub-optimal DNA preservation using a practical, plate-based Illumina-compatible Tn5 transposase library preparation method. PLoS One. 2024;19(3):e0300865; doi: 10.1371/journal.pone.0300865.

19. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20; doi: 10.1093/bioinformatics/btu170.

20. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9; doi: 10.1038/nmeth.1923.

21. Blanco-Miguez 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;41(11):1633–44; doi: 10.1038/s41587-023-01688-w.

22. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 2020;48(D1):D517–D25; doi: 10.1093/nar/gkz935.

23. Liu B, Zheng D, Zhou S, Chen L, Yang J. VFDB 2022: a general classification scheme for bacterial virulence factors. Nucleic Acids Res. 2022;50(D1):D912–D7; doi: 10.1093/nar/gkab1107.

24. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60; doi: 10.1093/bioinformatics/btp324.

25. Rognes T, Flouri T, Nichols B, Quince C, Mahe F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016;4:e2584; doi: 10.7717/peerj.2584.

26. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10(2); doi: 10.1093/gigascience/giab008.

27. Irber L, Pierce-Ward NT, Abuelanin M, Alexander H, Anant A, Barve K, et al. sourmash v4: A multitool to quickly search, compare, and analyze genomic and metagenomic data sets. Journal of Open Source Software. 2024;6830.

28. R Core Team (2025). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

29. Antipov D, Raiko M, Lapidus A, Pevzner PA. Metaviral SPAdes: assembly of viruses from metagenomic data. Bioinformatics. 2020;36(14):4126–9; doi: 10.1093/bioinformatics/btaa490.

30. Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, et al. Identification of mobile genetic elements with geNomad. Nat Biotechnol. 2024;42(8):1303–12; doi: 10.1038/s41587-023-01953-y.

31. Yu MK, Fogarty EC, Eren AM. Diverse plasmid systems and their ecology across human gut metagenomes revealed by PlasX and MobMess. Nat Microbiol. 2024;9(3):830–47; doi: 10.1038/s41564-024-01610-3.

32. Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, Alpi E, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Research. 2023;51(D1):D523–D31; doi: 10.1093/nar/gkac1052.
33. Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods. 2021;18(4):366–8; doi: 10.1038/s41592-021-01101-x.
34. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421; doi: 10.1186/1471-2105-10-421.
35. Ares-Arroyo M, Coluzzi C, Rocha EPC. Origins of transfer establish networks of functional dependencies for plasmid transfer by conjugation. Nucleic Acids Res. 2023;51(7):3001–16; doi: 10.1093/nar/gkac1079.
36. Luo H, Gao F. DoriC 10.0: an updated database of replication origins in prokaryotic genomes including chromosomes and plasmids. Nucleic Acids Res. 2019;47(D1):D74–D7; doi: 10.1093/nar/gky1014.
37. Néron B, Denise R, Coluzzi C, Touchon M, Rocha EPC, Abby SS. MacSyFinder v2: Improved modelling and search engine to identify molecular systems in genomes. Peer Community Journal. 2023;3.
38. Noordzij HT, Wortel MT, Heintz-Buschart A, Petrikonyte P, de Muinck EJ, Trosvik P. Assembly-based analysis of the infant gut microbiome reveals novel ubiquitous plasmids. Plasmid. 2025;134:102761; doi: 10.1016/j.plasmid.2025.102761.
39. Schmartz GP, Hartung A, Hirsch P, Kern F, Fehlmann T, Muller R, et al. PLSDB: advancing a comprehensive database of bacterial plasmids. Nucleic Acids Res. 2022;50(D1):D273–D8; doi: 10.1093/nar/gkab1111.
40. Ondov BD, Starrett GJ, Sappington A, Kostic A, Koren S, Buck CB, et al. Mash Screen: high-throughput sequence containment estimation for genome discovery. Genome Biol. 2019;20(1):232; doi: 10.1186/s13059-019-1841-x.
41. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014;15(3):R46; doi: 10.1186/gb-2014-15-3-r46.
42. 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;50(D1):D785–D94; doi: 10.1093/nar/gkab776.
43. Barrantes-Jimenez K, Lejzerowicz F, Tran T, Calderon-Osorno M, Rivera-Montero L, Rodriguez-Sanchez C, et al. Anthropogenic imprint on riverine plasmidome diversity and proliferation of antibiotic resistance genes following pollution and urbanization. Water Res. 2025;281:123553; doi: 10.1016/j.watres.2025.123553.
44. Lassmann T. Kalign 3: multiple sequence alignment of large data sets. Bioinformatics. 2019;36(6):1928–9; doi: 10.1093/bioinformatics/btz795.
45. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One. 2010;5(3):e9490; doi: 10.1371/journal.pone.0009490.
46. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 2019;35(3):526–8; doi: 10.1093/bioinformatics/bty633.
47. Bokulich NA, Chung J, Battaglia T, Henderson N, Jay M, Li H, et al. Antibiotics, birth mode, and diet shape microbiome maturation during early life. Sci Transl Med. 2016;8(343):343ra82; doi: 10.1126/scitranslmed.aad7121.
48. Stewart CJ, Ajami NJ, O'Brien JL, Hutchinson DS, Smith DP, Wong MC, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. Nature. 2018;562(7728):583–8; doi: 10.1038/s41586-018-0617-x.

49. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222–7; doi: 10.1038/nature11053.
50. Backhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, et al. Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. Cell Host Microbe. 2015;17(5):690–703; doi: 10.1016/j.chom.2015.04.004.
51. Yassour M, Vatanen T, Siljander H, Hamalainen AM, Harkonen T, Ryhanen SJ, et al. Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. Sci Transl Med. 2016;8(343):343ra81; doi: 10.1126/scitranslmed.aad0917.
52. Yassour M, Jason E, Hogstrom LJ, Arthur TD, Tripathi S, Siljander H, et al. Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life. Cell Host Microbe. 2018;24(1):146–54 e4; doi: 10.1016/j.chom.2018.06.007.
53. Lebeaux RM, Coker MO, Dade EF, Palys TJ, Morrison HG, Ross BD, et al. The infant gut resistome is associated with E. coli and early-life exposures. BMC Microbiol. 2021;21(1):201; doi: 10.1186/s12866-021-02129-x.
54. 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;14(1):8526; doi: 10.1038/s41467-023-44289-6.
55. van Schaik W. The human gut resistome. Philos Trans R Soc Lond B Biol Sci. 2015;370(1670):20140087; doi: 10.1098/rstb.2014.0087.
56. Macinga DR, Rather PN. The chromosomal 2'-N-acetyltransferase of Providencia stuartii: physiological functions and genetic regulation. Front Biosci. 1999;4:D132–40; doi: 10.2741/macinga.
57. Wright M, Kaur M, Thompson LK, Cox G. A historical perspective on the multifunctional outer membrane channel protein TolC in Escherichia coli. NPJ Antimicrob Resist. 2025;3(1):6; doi: 10.1038/s44259-025-00078-3.
58. Shimada T, Bridier A, Briandet R, Ishihama A. Novel roles of LeuO in transcription regulation of E. coli genome: antagonistic interplay with the universal silencer H-NS. Mol Microbiol. 2011;82(2):378–97; doi: 10.1111/j.1365-2958.2011.07818.x.
59. Stokholm J, Schjorring S, Eskildsen CE, Pedersen L, Bischoff AL, Folsgaard N, et al. Antibiotic use during pregnancy alters the commensal vaginal microbiota. Clin Microbiol Infect. 2014;20(7):629–35; doi: 10.1111/1469-0691.12411.
60. Gibson MK, Crofts TS, Dantas G. Antibiotics and the developing infant gut microbiota and resistome. Curr Opin Microbiol. 2015;27:51–6; doi: 10.1016/j.mib.2015.07.007.
61. Benitez AJ, Tanes C, Friedman ES, Zackular JP, Ford E, Gerber JS, et al. Antibiotic exposure is associated with minimal gut microbiome perturbations in healthy term infants. Microbiome. 2025;13(1):21; doi: 10.1186/s40168-024-01999-3.
62. He W, Russel J, Klincke F, Nesme J, Sorensen SJ. Insights into the ecology of the infant gut plasmidome. Nat Commun. 2024;15(1):6924; doi: 10.1038/s41467-024-51398-3.
63. Abramova A, Karkman A, Bengtsson-Palme J. Metagenomic assemblies tend to break around antibiotic resistance genes. BMC Genomics. 2024;25(1):959; doi: 10.1186/s12864-024-10876-0.
64. Jorgensen TS, Xu Z, Hansen MA, Sorensen SJ, Hansen LH. Hundreds of circular novel plasmids and DNA elements identified in a rat cecum metamobilome. PLoS One. 2014;9(2):e87924; doi: 10.1371/journal.pone.0087924.
65. Jørgensen TS, Hansen MA, Xu Z, A. TM, Sørensen SJ, Hansen LH. Plasmids, Viruses, And Other Circular Elements In Rat Gut. bioRxiv. 2017.

66. Tisza MJ, Smith DDN, Clark AE, Youn JH, Program NCS, Khil PP, et al. Roving methyltransferases generate a mosaic epigenetic landscape and influence evolution in group. Nature Communications. 2023;14(1); doi: 10.1038/s41467-023-39892-6.

Acknowledgements

We would like to acknowledge the excellent scientists at the Norwegian Sequencing Centre.

Funding

This work was funded by the Research Council of Norway grant 300948.

Contributions

HTN, EJDM and PT wrote the first draft of the manuscript. MN, KR, RV, HS and KLC contributed with sample collection and organization of metadata. EJDM processed samples for sequencing. HTN and PT analyzed data.

Corresponding author

Correspondence to Pål Trosvik.

Ethics declarations

Ethics approval and consent to participate

This study was carried out in accordance with relevant guidelines and regulations. All methods and protocols were approved by the Regional Ethics Committee of South-East Norway, REK application 2014/518. Informed consent was obtained from the legal guardians of all the infants in the study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementarymaterialv2.docx