medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The Impact of Chlorinated Drinking Water Exposure on Gut Microbiota Development in Infants: a Randomised Controlled Trial.

Kimberley Parkin[1,2], Claus T. Christophersen[4,5], Valerie Verhasselt[3,6], Debra J. Palmer[2,6], Matthew N. Cooper[2,6], Susan L. Prescott[2,7,8,9,10], Desiree Silva[2,5,7,,10], David Martino[1,11,*]

1 Wal-yan Respiratory Research Centre, The Kids Research Institute Australia, Nedlands, Perth, WA 6009, Australia

2 Medical School, University of Western Australia, Nedlands 6009, WA, Australia

3 Larsson-Rosenquist Centre for Immunology and Breastfeeding, Medical School, University of Western Australia, Nedlands 6009, WA, Australia

4 WA Human Microbiome Collaboration Centre, Curtin University, Bentley, Perth, WA 6102, Australia

5 Nutrition & Health Innovation Research Institute, School of Medical and Health Sciences, Edith Cowen University, Joondalup, Perth, WA 6027, Australia

6 Telethon Kids Institute, Nedlands, Perth, WA 6009, Australia

7 Joondalup Health Campus, Joondalup, Perth, WA 6027, Australia

8 Nova Institute for Health, Baltimore, Maryland 21231, USA

9 Department of Family and Community Medicine, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA

10 The ORIGINS Project, The Kids Research Institute Australia, Nedlands, WA 6009, Australia

11 School of Molecular Science, University of Western Australia, Nedlands, WA 6009, Australia

*Correspondence: david.martino@telethonkids.org.au

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

ABSTRACT

Background: Early gut microbiota disruptions can increase the risk for dysbiosis and predispose to chronic diseases. While chlorinated drinking water is generally considered safe, the antimicrobial effects of chlorine-based disinfectants may negatively impact the developing infant microbiota, which is particularly vulnerable during this critical period. This study investigates the specific effects of chlorinated water on the gut microbiome in infants.

Methods: The waTer qUality and Microbiome Study (TUMS) is a double-blinded, randomized controlled trial (RCT). Six-month old infants (n=197) received either de-chlorinated drinking water (via benchtop filtration), or regular tap water for twelve months. Stool samples were collected at six-months (baseline) and at eighteen-months of age. Metagenomic sequencing was used for faecal microbiome analysis. Samples of participant tap water was also collected pre- and post-intervention.

Results: Participants were balanced for sex, delivery method, and breast-feeding status at time of recruitment between the control (n=98) and treatment (n=99) groups. We collected a total of 170 baseline stool samples (83 in the control group and 87 in the treatment group), and 130 follow-up stool samples (65 in the control group and 65 in the treatment group). Interindividual variation was higher at six-months compared to eighteen months. No significant differences in overall community structure (beta diversity (p=0.35), richness (p=0.98) or Shannon Index (p=0.45)) were found between treatment groups. However, 55 genera were differentially abundant at follow-up (adjusted p-value < 0.01), primarily within the Firmicutes phylum. Analysis of metabolic pathways revealed the chlorinated water group showed significantly higher abundance of antibiotic resistance genes mostly attributable to *Escherichia* and *Klebsiella* species.

Conclusion: Water chlorination induces predominantly minor changes in infant gut microbiome composition; but appears to increase the abundance of antibiotic-resistant genes. While water chlorination remains a vital public health tool for ensuring safe drinking water, our findings underscore the need for continued research into the potential for increased antibiotic resistance and suggest there may be value in exploring alternative disinfectant strategies.

Australian New Zealand Clinical Trial Registry: ACTRN12619000458134, registered 20/03/2019,
https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=377164&isReview=true

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Key words: Infant microbiota, water quality, water chlorination, disinfection by-products, whole genome shotgun metagenomics, antibiotic resistance

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

INTRODUCTION

The gut microbiota, a complex ecosystem of microorganisms residing in the gastrointestinal (GI) tract, is essential for healthy development and protection against non-communicable disease. In infancy, environmental factors significantly influence the establishment of this microbiota [1], with long-term consequences for gut health and disease risk. Dysbiosis, a disruption of the normal microbial balance, can lead to altered composition, diversity, and function, potentially allowing harmful species to colonize [2]. This early life dysbiosis has been repeatedly linked to an increased risk of chronic diseases in later childhood, including allergies, asthma, atopic dermatitis, diabetes, obesity, cardiovascular diseases, and neurological disorders [3-5]. Given the global rise in chronic non-communicable diseases, understanding the factors that contribute to gut dysbiosis is a critical research priority.

One understudied environmental factor with the potential to disrupt infant gut microbiota assembly is early-life exposure to chlorinated drinking water [6]. Water chlorination, a widely implemented and effective method for controlling microbial contamination, has been instrumental in safeguarding public health from waterborne diseases. Chlorine-based disinfectants effectively inactivate harmful pathogens such as *Legionella pneumophila*, *Salmonella typhi*, and *Vibrio cholerae*, known to cause severe illnesses [7]. To ensure adequate disinfection throughout the distribution system, current practices aim to maintain a controlled residual chlorine level, typically below 5 mg/L [8], as recommended by international standards. This level is considered sufficient for disinfection while minimizing adverse health effects from disinfection by-products, which are formed when chlorine reacts with organic matter. Notably, current infant feeding guidelines in many countries, including Australia, recommend boiling and cooling tap water before giving it to infants under twelve months of age to minimize potential risks [9].

Previous research in animal models has consistently shown that chlorine exposure alters gut microbiome composition, diversity, and function [10-14]. Additionally, observational human studies have revealed associations between household water sources and distinct gut microbiota signatures [6,15]. However, only one intervention study to date has directly examined the impact of chlorine on the gastrointestinal microbiota in children [16]. Conducted in a low-middle income urban setting in Bangladesh, this study delivered chlorine via a dosing device fitted to communal taps. Using metagenomic sequencing, the authors found small, age-specific effects on the relative abundance of several human enterobacteria, as well as an increased abundance of antibiotic resistance genes (ARGs) in children aged 15-30 months [16]. While these findings support the overall safety of chlorine residuals for the GI

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

microbiota, they also suggest heightened susceptibility to alterations in microbial composition and ARG enrichment in children, particularly during specific developmental windows. Whether these effects on the gut microbiome are also evident in high-income settings with more advanced water treatment infrastructure and demographically unique microbiota remains an important unanswered question.

This study directly investigates the impact of chlorinated drinking water consumption on the developing infant gut microbiome within a high-income setting. We conducted a randomized controlled trial (RCT) in which six-month-old infants were assigned to receive either chlorinated tap water or de-chlorinated water for twelve months. Our primary objective was to compare measures of microbiome community composition, diversity and functional potential between infants exposed to chlorinated or de-chlorinated water, to identify any chlorine-specific alterations. These findings shed light on the impact of early-life chlorine exposure on the developing microbiota and the potential public health implications.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

METHODS

Ethics approval and consent to participate

The study protocol was approved by the Ramsay Health Care (RHC) WA | SA Human Research Ethics Committee (EC00266; HREC reference number 1911). Mothers provided written informed consent on behalf of their children. All procedures were conducted in accordance with the Declaration of Helsinki.

Study design

The Water Quality and the Microbiome Study (TUMS) is a parallel (ratio 1:1), double-blinded randomized controlled trial (RCT), nested within the ORIGINS Project birth cohort [17]. The ORIGINS Project is a longitudinal observational study of 10,000 pregnant women and their families, recruited from the Joondalup and Wanneroo communities of Western Australia during pregnancy, and followed until their offspring are five years of age [17]. We recruited six-month-old infants participating in the ORIGINS Project and randomly assigned them to receive either de-chlorinated tap water (intervention group) or regular tap water (control group) for twelve months. De-chlorinated water was provided via a benchtop water filtration device, while the control group received an identical sham filter that did not alter the chemical composition of the water.

Stool samples were collected from participants at study entry (baseline, six months of age) and again after one year (follow-up, eighteen months of age). Metagenomic sequencing was performed on both samples to profile the gut microbiome. Tap water samples were collected at baseline (prior to filter installation) and post-intervention for detailed chemical analysis. Health assessments were conducted on approximately half of the cohort (opt-in) at twelve months of age as part of the standard ORIGINS Project protocol. The study design is presented in Figure 1.

Eligibility criteria

Inclusion criteria were: living within the Perth metropolitan area (Western Australia), born healthy, at least 37 weeks' gestation, and using tap water connected to the mains water supply to prepare infant drinks. Exclusion criteria were: participating in another diet-based research study, using bottled or filtered water to prepare infant drinks, using childcare facilities for three or more days per week, planning overseas holidays lasting longer than four weeks, having received antibiotics since birth, or being in their seventh month of life.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Intervention

All participants received a benchtop water filter attached to their kitchen faucet. The intervention group's filters contained a Hi-Tech Optimizer carbon filtration block rated to remove >97% of chloramines and >99% of chlorine, while preserving fluoride. The control group received identical filters with a sediment-only filter that did not alter the chemical composition of the water. Participants were encouraged to use the provided filter for all infant drinks, maintain normal drinking habits, and carry study-supplied filtered water when outside the home. Filter cartridges were rated for 5,000 liters or twelve months to last the study duration. Participants were advised to follow Australian Infant Feeding Guidelines, regardless of group assignment. Pure Water Systems (Burleigh Heads, Queensland, Australia) supplied the intervention devices (unbranded) but did not sponsor this study. Both participants and investigators were blinded during the intervention period.

Randomization and blinding

Following enrolment, participants were randomized to either treatment group 'A' or 'B' using minimization (biased coin; 0.7) stratified by sex, breastfeeding status at recruitment, and mode of delivery. Allocation of filters to treatment groups was done by an independent biostatistician, with both participants and investigators blinded throughout the intervention. The allocator labelled filters, removing identifying information, and revealed group assignments only after data collection was complete. Filters were installed by the principal investigator, and unblinding occurred at completion of the intervention.

Data collection

House visits: Home visits were conducted at study entry and post-intervention. Baseline data collection included a study entry questionnaire, stool sample collection, tap water sampling, and filter installation. Post-intervention, a water sample was collected from the filter before removal. Participants returned their stool sample kits via a pre-paid envelope.

Water quality testing: A sample of flushed (30 seconds) tap water was analyzed for heavy metals, major ions, nitrates, pH, phosphates, chloramines, and trihalomethanes at baseline before installing the intervention filter, and at follow-up. On-site testing for free/total chlorine and pH was also conducted using a Hach D300 pocket colorimeter. Sample collection bottles were filled to exclude air and kept out of direct sunlight. Water samples were delivered to accredited facility ChemCentre in Bentley, Perth.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Questionnaires: Participants completed ORIGINS Project questionnaires providing information on antenatal characteristics, water use, child health, breastfeeding, demographics, medical history, and diet. Additional questionnaires assessed filter usage at nine, twelve, and fifteen months. Data was entered into a secure REDCap database.

Health assessments: Infant eczema, recurrent wheeze, and allergen sensitization were assessed at twelve months in a subset of participants who opted for pediatrician visits. Medically diagnosed eczema was defined by typical skin lesions [18] clinically observed by a medical practitioner. Recurrent wheeze was defined as more than one separate episode of wheeze symptoms clinically observed by a medical practitioner. Infant allergen sensitisation is defined as a positive skin prick test (with mean weal diameter ≥ 3 mm above the control weal size) to at least one food and/or environmental allergen at one year of age. The panel of allergens tested included hen's egg, cow's milk, wheat, fish (tuna), peanut, cashew nut, ryegrass pollen, cat and house dust mite, with histamine and control solutions, using commercial extracts, and in accordance with standard clinical methods.

Stool sample collection: Mothers collected stool samples at six and eighteen months using provided stool collection kits that utilized a Copan FLOQSwab in an active drying tube (FLOQSwab-ADT, Copan Diagnostics, Murrieta, CA, USA). Samples were shipped directly to Microba Life Sciences Limited (for storage and processing).

Laboratory processing and bioinformatics profiling

Metagenomic sequencing: DNA was extracted from stool samples and subjected to shotgun metagenomic sequencing and microbiome profiling using standard workflows developed by Microba Life Sciences (Brisbane, Queensland, Australia). A more detailed description of the methodology; DNA extraction, library preparation, and sequencing, has been previously published [1]. To summarise, DNA extraction was performed using a modified protocol, optimising the mechanical lysis step, and following the manufacturer's instructions for the QIAmp 96 PowerSoil QIAcube HT Kit (Qiagen 47021) on the QIAcube HT automated extraction system. Samples were required to reach a minimum of 0.2 ng/µL to pass quality control requirements. Libraries were prepared according to the manufacturer's instructions using the Illumina DNA Prep (M) Tagmentation Kit and indexed with IDT for Illumina Nextera DNA Unique Dual Indexes Set A-D (Illumina 20027213-16) with a modification to volume to accommodate processing in a 384-well plate format. Nextera XT libraries were pooled in equimolar amounts to create a sequencing pool and assessed using QuantIT (Thermo Fisher Q33120) and visualised by capillary gel electrophoresis using a QIAxcel

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

DNA High Resolution Kit (Qiagen 929002). Sequencing pools were loaded and sequenced on a NovaSeq6000 (Illumina) according to manufacturer's instructions and sequenced with 2 x 150bp paired-end chemistry in the Microba accredited laboratories.

Raw data processing: Data quality control and processing was performed at Microba Life Sciences Limited. Paired-end DNA sequencing data was demultiplexed and adaptor-trimmed using the Illumina BaseSpace Bcl2fastq2 (v2.20), accepting one mismatch in index sequences. Reads were quality trimmed and residual adaptors were removed using the Trimmomatic v0.39 [19] software package with the following parameters: -phred33 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 CROP:100000 HEADCROP:0 MINLEN:100. Human DNA was identified and removed by aligning reads to the human genome reference assembly 38 (GRCh38.p12, GCF_000001405) using bwa-mem v0.7.17 [20] with default parameters, except that the minimum seed length was set to 31 (-k31). Human genome alignments were filtered using SAMtools v1.7 [21], with flags -ubh-f1-F2304. Any read pairs in which at least one read mapped to the human genome with >95% identity and >90% of the read length were flagged as human DNA and removed. All samples were randomly subsampled to a standard depth of seven million read pairs to standardize the number of reads across all samples to ensure they could be statistically compared.

Taxonomic and metabolic Profiling: Species-level metagenomes were assembled using the Microba Community Profiler (MCP) v1.0 (www.microba.com) using the Microba Genome Database (MGDB) v1.0.3 as the genome reference database. Reads were assigned to genomes within the MGDB, and the relative cellular abundance of species clusters were estimated and reported. Quantification of gene and pathway abundance in the metagenomic samples were performed using the Microba Gene and Pathway Profiler (MGPP) v1.0 against the Microba Genes (MGENES) database v1.0.3. MGPP is a two-step process. In step one, all open reading frames (ORFs) from all genomes in the MGDB were clustered against UniRef90 [22] release 2019/04 using 90% identity over 80% of the read length with MMSeqs2 Release 10-6d92c [23]. To determine metabolic pathways encoded by metagenome profiles, we used the Enzyme Commission annotations of MetaCyc [24] pathways using the EnrichM tool (https://github.com/geronimp/enrichM), and identified MetaCyc pathways with a minimum completeness >= 80%. In step two, all DNA sequencing read pairs aligned with one or more bases to the gene sequence from any protein within an MGENES protein cluster were summed. The abundance of encoded pathways of species reported as detected by MCP was calculated by averaging the read counts of all genes for each enzyme in that pathway.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Statistical analyses

All analyses were performed using R version 4.3.1. All boxplots represent the first and third quartiles, with a median as a middle line and whiskers at the last value within a 1.5 × IQR distance from the upper or lower quartile, where IQR is the interquartile range. An alpha level 0.05 was used to define statistical significance for unadjusted and adjusted p-values as appropriate.

Sample size: Our sample size was based on a previously published randomized control crossover intervention, which investigated the effects of alkaline and neutral water pH on faecal microbiota in 29 Danish men. The authors noted that their post hoc analysis indicated a sample size of 180 would have 80% power to detect differences alpha diversity at a p-value of 0.05 [25].

Study outcomes: The primary outcome of this study was the difference in infant gut microbiome composition between the intervention and control groups, measured from stool samples collected at six and eighteen months of age. Secondary outcomes included the incidence of allergic sensitization, wheeze, and atopic eczema, as assessed by physician diagnosis at 12 months of age.

Community comparison: The overall composition of the gut microbiome was visualized using taxonomic bar plots of the average relative abundance of the top eight most abundant phyla. Data were stratified by age and intervention group.

Diversity analysis: Diversity measures were calculated using the phyloseq R package (version 1.38.0) [26]. The Shannon index and observed richness were used to estimate alpha diversity

and differences between groups was assessed using the Kruskal-Wallis test, followed by Dunn's post hoc test with Bonferroni correction for multiple comparisons. Beta diversity was calculated using the Bray-Curtis dissimilarity index, and non-metric multidimensional scaling (NMDS) was employed to visualize beta diversity. Permutational multivariate analysis of variance (PERMANOVA) was used to formally test for differences in beta diversity between groups [27].

Differential analysis: To identify differentially abundant amplicon sequence variants (ASVs) and metabolic pathways between the intervention and control groups at 18 months of age, we utilized DESeq2 (version 1.40.2) [28]. This approach models count data using a negative binomial distribution and applies a Wald test for significance. ASVs present in at least 5% of samples (n=9/170) and with a total count exceeding 1000 were considered for analysis, and

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

taxonomic classification was determined at the species level. P-values were adjusted for multiple comparisons using the Benjamini-Hochberg false discovery rate (FDR) correction. Similarly, differentially abundant MetaCyc pathways between groups were identified using DESeq2. Data visualization was performed using the R packages phyloseq (24), microViz (version 0.10.0) (27), ggplot2 (version 3.4.0) (28), and ggstatsplot (package version 0.12.2) (29).

SAGER guidelines. Subject sex was recorded in clinical notes as parent-reported male or female. Regression modelling was adjusted for sex such that findings are generalizable to both sexes.

Inclusion and Ethics. This research was conducted in partnership with local researcher consumers who were involved in the design, data collection, analysis and reporting of the study. The research is locally relevant, and the study protocol was acceptable to parents and local IRB.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

RESULTS

Study Design and Participant Flow

Recruitment occurred between November 2019 and March 2022, enrolling 197 participants randomized to the control (n=98) or treatment (n=99) groups. Baseline characteristics were balanced between groups, including sex, delivery method (assisted vaginal, unassisted vaginal, elective caesarean, or emergency caesarean), and breastfeeding status at recruitment (predominantly formula-fed, predominantly breastfed, or mixed-fed), as shown in Table 1. Further demographic details on the cohort were described previously (ref). During the study, 76 adverse events (AEs) were reported, defined as instances where a participant presented to hospital for any reason. These were mostly due to unrelated common childhood infections or injuries. Two participants experienced serious adverse events (SAEs) including anaphylaxis, and septic arthritis which were assessed and deemed unrelated to the intervention.

Of the 426 infants assessed for eligibility, 229 were excluded (Figure 2). Reasons for exclusion included: not meeting inclusion criteria (n=190), declining participation (n=36), residing outside the eligible area (n=1), and logistical constraints preventing the initial home visit (n=1). Of the 197 randomized participants, 10 did not receive the intervention (4 treatment, 6 control) due to filter incompatibility with their tap fittings. A total of 130 participants (66%) completed the study and provided stool samples at 18 months of age.

Among those with 12-month health assessment data available, no significant differences were observed in health outcomes between intervention and control groups, including antibiotic use (Table 1). Due to the logistical disruptions related to the COVID-19 pandemic, the rate of missing data in these health outcomes was high (approximately 50%).

Effectiveness of the Intervention on Water Quality

Participants primarily resided in the northern regions of Perth, Western Australia (Figure 3A). Pre- and post-intervention water samples were collected from 83 and 57 treatment group households, and 82 and 57 control group households, respectively. Pre-intervention total chlorine levels varied across homes (0.3-1.1 mg/L), but all were below the 5 mg/L recommended by the Australian Drinking Water Guidelines (Figure 3A). Disinfection by-products (bromoform, bromodichloromethane, chloroform, chlorodibromomethane) also varied at baseline but remained below recommended levels (Supplementary Figure 1).

The intervention significantly reduced total chlorine levels in the treatment group from a mean of 0.74 mg/L (SD = 0.2447) pre-intervention to 0.10 mg/L (SD = 0.000) post-intervention (paired t-test, $p = 2.2^{-16}$). In contrast, the control group showed no significant

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

change in mean total chlorine levels (0.71 mg/L pre-intervention vs. 0.63 mg/L post-intervention, paired t-test, p = 0.06). The intervention also significantly reduced bromoform, bromodichloromethane, and chlorodibromomethane levels, while the control group showed no significant changes in these disinfection by-products (Figures 3C, 3D, and 3E).

Participant compliance with the intervention, assessed through three monthly surveys, revealed no significant differences between the treatment and control groups. Survey questions encompassed the daily consumption of filtered water (cups per day) by the infant (mean in control group=3.1, mean in intervention=2.9, P=0.06, Welch t-test), usage of study-provided filters outside the home (never, sometimes, always, P=0.574, Chi-Squared test), and the age at which unboiled tap water was introduced (mean in control group=10.71, mean in intervention=10.60, P=0.781, Welch t-test). These findings suggest consistent adherence to the intervention protocol across both groups (Supplementary Figure 2).

Microbial community profiling

We analysed 170 stool samples collected at baseline (six months of age) and 130 samples collected at follow-up (18 months of age), totalling 300 unique faecal metagenomes. Batch effect analysis did not identify any significant differences in microbial community composition between sequencing batches (Supplementary Figure 3).

After quality filtering and rarefaction to 7,000,000 read pairs per sample, a total of 3,936,112,872 reads were obtained (mean per sample: 12,990,472 ± 301,687 SD). On average, 47,087 human DNA reads per sample (SD ± 655,372) were removed before analysis. Across all samples, we identified 1,305 bacterial ASVs classified at the species level, belonging to 420 genera, 98 families, 42 orders, 17 classes, and 14 phyla. Additionally, we identified two archaeal ASVs (*Methanobrevibacter smithii* and *Methanobrevibacter smithii_A*) and ten eukaryotic ASVs (various *Candida*, *Clavispora*, *Geotrichum*, *Kluyveromyces*, *Saccharomyces*, *Wickerhamiella*, and *Blastocystis* species).

Effect of chlorinated water on microbiota taxonomic structure and diversity

The taxonomic characterization of participants' microbiota aligned with previous reports on young infants, with *Bifidobacterium infantis*, *Bifidobacterium longum*, and *Bifidobacterium breve* as the predominant species at six months of age [29,30]. Interindividual variation in microbiome composition decreased from baseline to follow-up (Supplementary Figure 4).

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

While at each timepoint both groups displayed similar phylum-level structures (Figure 4A), these structures overall changed significantly from baseline to follow-up.

Beta diversity analysis revealed significant variation associated with age (11.3% of variance explained, PERMANOVA, $p = 0.001$), but no significant clustering by treatment group (0.6% of variance explained, $p = 0.35$) in both longitudinal (Figure 4B) and cross-sectional analyses (Supplementary Figure 5). Alpha diversity was not significantly different between treatment groups at either time point (observed richness index $p = 0.98$, Shannon diversity index $p = 0.45$) (Figures 4C & 4D).

Water chlorination-induced shifts in taxa abundance

Differential abundance analysis using DESeq2 revealed 55 genera across 7 phyla were significantly different in abundance between the intervention (de-chlorinated water) and control (chlorinated water) groups in the 18-month (post-intervention) stool samples (adjusted p-value < 0.01) (Figure 5). These differentially abundant genera were primarily members of the Firmicutes phylum. However, due to high inter-individual variation, most of these taxa were present at low frequencies in both groups, although they were more frequently detected in the intervention group (Supplementary Figure 6). A comprehensive list of all differentially abundant ASVs, along with their statistical significance, is provided in Supplementary Table 1.

Higher antibiotic resistance gene abundance in children receiving chlorinated water

To assess the impact of water chlorination on gut microbiome function, we analysed the abundance of metabolic pathways annotated in the MetaCyc database [24] using DESeq2 [28]. In the 18-month (post-intervention) samples, we identified 901 metabolic pathways belonging to 51 major groups. Two of these groups were differentially abundant between the intervention and control groups , with chlorinated water exposure most significantly associated with higher abundance of the 'Antibiotic Resistance' major group, as well as the sundry group 'Other' (adjusted $p < 0.05$) (Figure 6A, Table 2). At the pathway level, 29 metabolic pathways were upregulated in the chlorinated water group (adjusted $p < 0.05$), including polymyxin resistance which has membership in the Antibiotic Resistance major group (Figure 6B, Table 2). To determine which species contributed to increased expression of antibiotic resistance genes, we examined the species abundance in individual microbiomes that were annotated to the MetaCyc "Antibiotic Resistance" functional pathway. This analysis revealed that antibiotic resistance genes were distributed across multiple species, including

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Escherichia, *Klebsiella*, *Enterobacter*, *Clostridia*, and *Citrobacter*, among others (Supplementary Figure 7).

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

DISCUSSION

In this double-blinded randomized controlled trial conducted in a high-income setting, infants exposed to chlorinated drinking water exhibited no differences in the overall richness and diversity of the microbiota, although subtle shifts in taxonomic abundance were observed. Several species known to be part of a healthy gut microbiome, including *Bacteroides*, *Clostridium*, and *Blautia* showed reductions in taxonomic abundance. While some beneficial bacteria like *Faecalibacterium prauznitzii* and *Bifidobacterium* increased in abundance, others *Streptococcus*, *Clostridia* and *Klebsiella pneumoniae* also increased, which could be opportunistic pathogens under certain circumstances. It is important to note that carriage of differentially abundant taxa was generally low across the cohort against a background of strikingly individualized microbiota compositions in this age group. On balance, our data suggest chlorination-induced shifts in the microbiota vary substantially from infant to infant, possibly due to a range of factors including genetics, diet and environmental exposures.

Consistent with the recent study by Nadimpalli [16], we found that exposure to chlorinated drinking water for a 12-month period was associated with a higher carriage of antibiotic resistance genes in stool samples. These ARGs, which encompass genes encoding efflux pumps and antibiotic-degrading enzymes, were detected across diverse bacterial species, including *Escherichia*, *Klebsiella*, *Enterobacter*, *Clostridia*, and *Citrobacter*. While not directly indicative of antibiotic-resistant species, this finding aligns with the Bangladeshi trial suggesting chlorinated water may selectively promote the growth of bacteria harbouring these genes. It's noteworthy that the age group in the previous study where chlorination effects were observed (15-30 months) overlaps with the age range of our study participants. This suggests there might be specific developmental windows during which the gut microbiome is particularly susceptible to the influence of chlorinated water. Further investigation is required to ascertain the potential for the development or spread of antibiotic resistance in this population.

Our study has limitations, including a relatively small sample size and the inability to draw definitive conclusions about infant health outcomes. Antibiotic use, recorded in approximately one-third of participants equally across both groups, could not be controlled for in a per-protocol analysis due to power limitations. While compliance and water consumption practises were monitored, the potential for unmeasured confounders is acknowledged. Individual exposure patterns likely varied due to the complexities of Perth's water distribution network. Additionally, our findings cannot be solely attributed to chlorine, as the study filters also removed nitrates, heavy metals, and disinfection by-products. The

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

potential for unmeasured confounders related to diet or socioeconomic status is also acknowledged.

CONCLUSION

Our study extends previous findings and provides valuable insights into the impact of chlorinated drinking water on the infant gut microbiota. These findings are particularly relevant to the current Australian Infant Feeding Guidelines, which recommend introducing unboiled tap water at twelve months. Considering the high plasticity of the infant gut microbiome until three years of age, our findings raise questions about the potential benefits of continuing to boil or filter tap water to remove chlorine residuals until later ages. Further research is needed to explore the potential for increased antibiotic resistance.

FUNDING STATEMENT

This work was supported by Telethon Kids Institute Blue Sky award 2018. Individual authors are supported by: DM work was supported by the Western Australian Future Health Research and Innovation Fund, which is an initiative of the WA State Government. VV work was supported by the funding endowed by the Family Larsson-Rosenquist Foundation. DJP was supported by the NHMRC Medical Research Futures Fund (MRFF) Career Development Fellowship ID1144544 (2018-2021) and the Telethon Kids Institute Ascend Fellowship (2022-2026).

INSTITUTIONAL REVIEW BOARD STATEMENT

The study was conducted in accordance with the Declaration of Helsinki and was completed in accordance with the ethical guidelines of the Ramsay Health Care (RHC) WA | SA HREC (EC00266) and under the approved HREC reference number 1911 (approved 14 June 2019).

INFORMED CONSENT STATEMENT

Informed consent was obtained from all subjects involved in the study.

DATA AVAILABILITY STATEMENT

Raw data supporting the conclusions of the article will be make available by the authors upon request.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

COMPETING INTERESTS STATEMENT

The authors declare that this study received funding from the Water Corporation WA. This funding was used to support the costs associated with home tap water analysis and metagenomic sequencing. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors declare no other potential conflicts of interest.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the invaluable contributions of the ORIGINS Project families for their ongoing support of The Water Quality and Microbiome Study (TUMS). We extend our sincere thanks to the following teams and individuals who have made The ORIGINS Project possible: The ORIGINS Project team; Joondalup Health Campus (JHC); members of ORIGINS Project Community Reference and Participant Reference Groups; Research Interest Groups and the ORIGINS Project Scientific Committee; Telethon Kids Institute; City of Wanneroo; City of Joondalup; and Fiona Stanley.

We are also grateful for the contributions of Richard Theobald (WA Department of Health) for advice on water quality testing and study design considerations; the Telethon Kids Biostatistics team (Ali Hollingsworth, Zac Dempsey) for their assistance with randomisation, blinding and compliance monitoring; Nikki Schultz for her analysis of the compliance data, the Water Corporation WA (Cameron Gordon, Steve Christie, Gabrielle O'Dwyer) for their interest, advice, and engagement with the study. We acknowledge the support of our research partners, Microba Life Sciences, ChemCentre and Pure Water Systems for their additional in-kind support of the project. We would also like to thank Jen Rozier from Telethon Kids Institute for the geospatial mapping analysis, and Cynthia Joll from Curtin University for her expertise and advice on water quality guidelines in Australia.

AUTHORSHIP CONTRIBUTIONS. Conception and design of the work (D.M, D.S, M.C, D.P, C.C, V.V, S.P). Data collection (K.P.) Drafting the article (K.P, D.M). Critical revisions (all authors). Final approvals (all authors).

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

REFERENCES

1. Parkin K, Palmer DJ, Verhasselt V, et al. Metagenomic Characterisation of the Gut Microbiome and Effect of Complementary Feeding on Bifidobacterium spp. in Australian Infants. *Microorganisms* 2024; **12**(1): 228.
2. Bidell MR, Hobbs ALV, Lodise TP. Gut microbiome health and dysbiosis: A clinical primer. *Pharmacotherapy* 2022; **42**(11): 849-57.
3. Sarkar A, Yoo JY, Valeria Ozorio Dutra S, Morgan KH, Groer M. The Association between Early-Life Gut Microbiota and Long-Term Health and Diseases. *J Clin Med* 2021; **10**(3).
4. Abrahamsson TR, Jakobsson HE, Andersson AF, Björkstén B, Engstrand L, Jenmalm MC. Low gut microbiota diversity in early infancy precedes asthma at school age. *Clin Exp Allergy* 2014; **44**(6): 842-50.
5. Łoś-Rychalska E, Gołębiewski M, Grzybowski T, Rogalla-Ładniak U, Krogulska A. The microbiome and its impact on food allergy and atopic dermatitis in children. *Advances in Dermatology and Allergology/Postępy Dermatologii i Alergologii* 2020; **37**(5): 641-50.
6. Bowyer RCE, Schillereff DN, Jackson MA, et al. Associations between UK tap water and gut microbiota composition suggest the gut microbiome as a potential mediator of health differences linked to water quality. *Sci Total Environ* 2020; **739**: 139697.
7. Magana-Arachchi DN, Wanigatunge RP. Chapter 2 - Ubiquitous waterborne pathogens. In: Vara Prasad MN, Grobelak A, eds. Waterborne Pathogens: Butterworth-Heinemann; 2020: 15-42.
8. Organization WH. Guidelines for drinking-water quality: World Health Organization; 2004.
9. Council. CNHaMR. Infant Feeding Guidelines. *National Health and Medical Research Council (2012)* 2012.
10. Zhang Y, Zhao F, Deng Y, Zhao Y, Ren H. Metagenomic and metabolomic analysis of the toxic effects of trichloroacetamide-induced gut microbiome and urine metabolome perturbations in mice. *Journal of proteome research* 2015; **14**(4): 1752-61.
11. Sasada T, Hinoi T, Saito Y, et al. Chlorinated water modulates the development of colorectal tumors with chromosomal instability and gut microbiota in Apc-deficient mice. *PLoS One* 2015; **10**(7): e0132435.
12. Dias MF, Reis MP, Acurcio LB, et al. Changes in mouse gut bacterial community in response to different types of drinking water. *Water Res* 2018; **132**: 79-89.
13. Zhou K, Liu W, Chen Z, et al. The effect of different drinking water in culture medium on feces microbiota diversity. *J Water Health* 2021; **19**(2): 267-77.
14. Zhou Z, Lin Z, Shuai X, Achi C, Chen H. Antibiotic resistance genes alterations in murine guts microbiome are associated with different types of drinking water. *J Hazard Mater* 2024; **465**: 133422.
15. Baumann-Dudenhoeffer AM, D'Souza AW, Tarr PI, Warner BB, Dantas G. Infant diet and maternal gestational weight gain predict early metabolic maturation of gut microbiomes. *Nat Med* 2018; **24**(12): 1822-9.
16. Nadimpalli ML, Lanza VF, Montealegre MC, et al. Drinking water chlorination has minor effects on the intestinal flora and resistomes of Bangladeshi children. *Nat Microbiol* 2022; **7**(5): 620-9.
17. Silva DT, Hagemann E, Davis JA, et al. Introducing the ORIGINS project: a community-based interventional birth cohort. *Rev Environ Health* 2020; **35**(3): 281-93.
18. Hanifin JM, Rajka G. Diagnostic features of atopic dermatitis. 1980.
19. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; **30**(15): 2114-20.
20. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv: Genomics* 2013.
21. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**(16): 2078-9.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

22. Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* 2015; **31**(6): 926-32.

23. Steinegger M, Söding J. Clustering huge protein sequence sets in linear time. *Nat Commun* 2018; **9**(1): 2542.

24. Caspi R, Billington R, Keseler IM, et al. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Res* 2020; **48**(D1): D445-d53.

25. Hansen TH, Thomassen MT, Madsen ML, et al. The effect of drinking water pH on the human gut microbiota and glucose regulation: results of a randomized controlled cross-over intervention. *Sci Rep* 2018; **8**(1): 16626.

26. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013; **8**(4): e61217.

27. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 2001; **26**(1): 32-46.

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

29. Sakata S, Tonooka T, Ishizeki S, et al. Culture-independent analysis of fecal microbiota in infants, with special reference to Bifidobacterium species. *FEMS Microbiol Lett* 2005; **243**(2): 417-23.

30. Tannock GW, Lawley B, Munro K, et al. Comparison of the compositions of the stool microbiotas of infants fed goat milk formula, cow milk-based formula, or breast milk. *Appl Environ Microbiol* 2013; **79**(9): 3040-8.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Tables

Table 1. Characteristics of 197 children participating in TUMS

Characteristics	**Treatment, *n* (%)**	**Control, *n* (%)**
Sex		
Male	49/99 (49.5)	46/98 (46.9)
Female	50/99 (50.5)	52/98 (53.1)
Milk-based diet at six months of age		
Predominantly breastfed	55/99 (55.5)	50/98 (51.0)
Predominantly formula-fed	32/99 (32.3)	36/98 (36.7)
Mixed fed	12/99 (12.1)	12/98 (12.2)
Delivery method		
Vaginal	55/99 (55.5)	57/98 (58.2)
Emergency caesarean	20/99 (20.2)	18/98 (18.4)
Elective caesarean	24/99 (24.2)	23/98 (23.5)
Health outcomes at twelve months of age		
Antibiotic use	33/156 (21.2)	30/156 (19.2))
Allergen sensitisation	4/76 (5.3)	4/76 (5.3)
Wheeze	3/15 (20)	2/15 (13.3)
Eczema	5/16 (31.3)	6/16 (37.5)

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 2. Significantly differentially abundant MetaCyc groups ($p<0.05$) and pathways (adjusted $p<0.05$) between the intervention and control groups at follow-up (rounded to four decimal places)

MetaCyc major group	**Log2FoldChange**	**p-value**	**Adjusted p-value**
Antibiotic Resistance	1.5876	0.0006	0.0300
Other	1.7332	0.0017	0.0407
Methylglyoxal Detoxification	1.7602	0.0058	0.0587
TCA cycle	1.7502	0.0060	0.0587
Respiration	0.7191	0.0052	0.0587
Entner-Duodoroff Pathways	1.4993	0.0119	0.0975
Fatty Acid and Lipid Degradation	0.47441	0.0226	0.1583
Metabolic Regulator Biosynthesis	-0.2116	0.0430	0.2448
O-Antigen Biosynthesis	1.2175	0.0450	0.2448
MetaCyc pathway			
PWY-6618~guanine and guanosine salvage III	0.5285	0.0001	0.0480
PWY-6536~4-aminobutanoate degradation III	0.5265	0.0002	0.0480
PWY-1881~formate oxidation to CO2	0.3693	0.0003	0.0480
KDO-LIPASYN-PWY~(Kdo)2-lipid A biosynthesis I	0.5522	0.0005	0.0480
PWY66-162~ethanol degradation IV	0.3786	0.0005	0.0480
PWY0-1338~polymyxin resistance	0.5610	0.0006	0.0480
PWY-6537~4-aminobutanoate degradation II	0.5072	0.0007	0.0480
PWY0-1471~uracil degradation III	0.5748	0.0007	0.0480
GLUCARGALACTSUPER-PWY~superpathway of D-glucarate and D-galactarate degradation	0.4901	0.0007	0.0480
PWY-6896~thiamine salvage I	0.5748	0.0008	0.0480
PWY-46~putrescine biosynthesis III	0.2671	0.0009	0.0480
CHOLINE-BETAINE-ANA-PWY~choline degradation I	0.5296	0.0009	0.0480
BETSYN-PWY~glycine betaine biosynthesis I (Gram-negative bacteria)	0.5296	0.0009	0.0480
PWY66-394~aspirin triggered resolvin E biosynthesis	0.4882	0.0010	0.0480
P283-PWY~hydrogen oxidation I (aerobic)	0.2789	0.0010	0.0480
SO4ASSIM-PWY~assimilatory sulfate reduction I	0.6237	0.0011	0.0480
PWY-6050~dimethyl sulfoxide degradation	0.2765	0.0011	0.0480
TCA-GLYOX-BYPASS~superpathway of glyoxylate bypass and TCA	0.5565	0.0012	0.0480
FERMENTATION-PWY~mixed acid fermentation	0.5438	0.0013	0.0480
PWY-6535~4-aminobutanoate degradation I	0.5876	0.0013	0.0480

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

DETOX1-PWY-1~reactive oxygen species degradation	0.6156	0.0013	0.0480
PWY-7417~phospholipid remodeling (phosphatidate, yeast)	0.5341	0.0014	0.0480
PWY0-1241~ADP-L-glycero-Î²-D-manno-heptose biosynthesis	0.3240	0.0015	0.0480
PWY0-981~taurine degradation IV	0.5531	0.0015	0.0480
PUTDEG-PWY~putrescine degradation I	0.5320	0.0016	0.0480
PWY-5137~fatty acid Î²-oxidation III (unsaturated, odd number)	0.5200	0.0017	0.0480
TCA~TCA cycle I (prokaryotic)	0.5852	0.0018	0.0480
PWY-5755~4-hydroxybenzoate biosynthesis II (bacteria)	0.5693	0.0018	0.0480
PWY-7856~heme a biosynthesis	0.5484	0.0018	0.0480

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure Legends

Figure 1. Graphical overview of trial design. A timeline of data collection events is shown as a function of infant age in months.

Figure 2. CONSORT flow diagram of participant recruitment, enrolment, allocation, follow-up, and analysis. This diagram illustrates the flow of participants through the different stages of the TUMS randomized controlled trial. It shows the number of infants assessed for eligibility, reasons for exclusion, randomization into the intervention and control groups, loss to follow-up, and the final number of participants included in the analysis.

Figure 3. Impact of the intervention on water quality in TUMS participant homes. A geospatial heatmap map displaying mean total chlorine levels (mg/mL), grouped by suburb. Total n=197 (**A**). Boxplot comparison between mean levels of total chlorine (**B**) and disinfection by-products: bromoform (**C**), bromodichloromethane (**D**), and chlorodibromomethane (**E**) between the treatment and control groups. The box indicates the interquartile range (25th to 75th percentile), the horizontal line within the box represents the median, and the whiskers extend to 1.5 times the interquartile range. Outliers are shown as individual points. Total n=140.

Figure 4. No effect of chlorinated drinking water on the taxonomic structure and diversity of the infant gut microbiome. (**A**) Average relative abundance of bacterial phyla across all samples, stratified by age (6 and 18 months) and treatment group (chlorinated or de-chlorinated water). (**B**) Non-metric multidimensional scaling (NMDS) plot visualizing beta diversity (inter-sample dissimilarity) based on Bray-Curtis distances. Each point represents a sample, colored by age and shaped by treatment group. (**C**) Boxplots comparing observed richness (number of unique genera) between treatment groups at 6 and 18 months of age. (**D**) Boxplots comparing Shannon diversity indices (a measure of both richness and evenness) between treatment groups at 6 and 18 months of age. The boxes indicate the interquartile range (25th to 75th percentile), the horizontal line within the box represents the median, and the whiskers extend to 1.5 times the interquartile range. Outliers are shown as individual points.

Figure 5. Differential abundance of bacterial genera in infant gut microbiota between treatment groups. The dot plot displays the log2 fold change in abundance for each bacterial

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

genus identified as significantly different (adjusted p-value < 0.01) between the de-chlorinated water (intervention) and chlorinated water (control) groups at 18 months of age. Each dot represents a genus, coloured by phylum. A positive log2 fold change indicates higher abundance in the intervention group, while a negative value indicates higher abundance in the control group. Data are plotted at the genus level for ease of visualisation.

Figure 6. Impact of water chlorination on infant gut microbiome metabolic pathways. (**A**) Volcano plot displaying the differential abundance of MetaCyc metabolic groups between the de-chlorinated water (intervention) and chlorinated water (control) groups. Each point represents a metabolic group, with the x-axis showing log2 fold change and the y-axis showing -log10 adjusted p-value. Red points indicate groups with significantly higher abundance in the control group (adjusted p < 0.05). (**B**) Volcano plot displaying the differential abundance of individual MetaCyc metabolic pathways between the intervention and control groups, following the same format as panel A.

Supplementary Figure 1. Geospatial heatmap of disinfection by-products (DBPs). Visualisation represents DBPs expressed as ug/mL averaged over participant homes and represented at the suburb level.

Supplementary Figure 2. Participant compliance with the intervention assessed through three-monthly questionnaires. (A) Frequency of parent-reported consumption of intervention water (n=??). (B) Counts of the frequency that parents report taking intervention water with them when outside the home (n=?). (C) Histogram showing distribution of mean age of introduction of unboiled tap water into the infant diet for intervention and control groups (n=?).

Supplementary Figure 3. Lack of batch effect on gut microbiota composition. Stool samples were sequenced in two batches. The first batch included all 170 baseline samples and 35 follow-up samples. The remaining follow-up samples, along with three follow-up samples re-sequenced from the first batch, were included in the second batch. Non-metric multidimensional scaling (NMDS) ordination demonstrates no significant difference in gut microbial community composition between sequencing batches. Each point represents an individual sample. Red symbols denote three randomly selected samples sequenced in batch 1, while blue symbols represent the same three samples re-sequenced in batch 2. The

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

proximity of red and blue symbols for each sample indicates minimal variation introduced by sequencing batch.

Supplementary Figure 4. Inter-individual variation in gut microbiota composition with age. Taxonomic bar plot illustrating the top twelve most abundant genera at baseline (top) and at follow-up (bottom), ranked by Bray-Curtis dissimilarity within each time point. Each bar represents an individual participant, coloured according to the relative abundance of each genus (x-axis).

Supplementary Figure 5. Developmental shifts in gut microbiota composition. Non-metric multidimensional scaling (NMDS) ordination illustrating pairwise genomic distances between samples at (A) six months and (B) eighteen months of age. Each point represents an individual sample. The distinct clustering patterns at different ages suggest significant shifts in gut microbial community structure during early childhood development.

Supplementary Figure 6. Differentially Abundant Gut Microbiota Associated with Intervention. Box plots display the log-transformed abundance of the top four Amplicon Sequence Variants (ASVs) exhibiting significant differences between the intervention and control groups: (A) *Clostridium MIC8573*, (B) *Bacteroides faecis*, (C) *Bacteroides fragilis_A*, and (D) *TF01-11 sp000436755*.

Supplementary Figure 7. Antibiotic Resistance Genes in Gut Microbiota Across Treatment Groups. Heatmap depicting the abundance of bacterial species harbouring genes associated with antibiotic resistance (as defined by the MetaCyc database) within the gut microbiome of each infant. Rows represent individual bacterial species, while columns represent treatment groups. Cell color intensity reflects the relative abundance of each species within the respective group, highlighting potential differences in antibiotic resistance profiles among infants receiving different treatments.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

682 **SUPPLEMENTARY TABLES**

683 **Supplementary Table 1:** Differentially abundant species between the intervention (de-
684 chlorinated water) and control (chlorinated water) groups at follow-up.

Species	log2FoldChange	p-value	Adjusted p-value	Effect of chlorinated water
Clostridium MIC8573	-30.00	1.34E-29	4.46E-27	Reduces abundance
Bacteroides faecis	-30.00	4.63E-28	7.69E-26	Reduces abundance
Bacteroides fragilis_A	-30.00	3.87E-25	1.43E-23	Reduces abundance
TF01-11 sp000436755	-30.00	3.87E-25	1.43E-23	Reduces abundance
Bacteroides clarus	-30.00	3.87E-25	1.43E-23	Reduces abundance
CAG-83 sp000435555	-30.00	3.87E-25	1.43E-23	Reduces abundance
Butyricimonas synergistica_A	-30.00	3.87E-25	1.43E-23	Reduces abundance
Eubacterium_R sp003526845	-29.57	1.81E-24	5.45E-23	Reduces abundance
CAG-74 MIC7649	-29.15	8.10E-24	1.94E-22	Reduces abundance
CAG-45 sp000438375	-28.90	1.91E-23	3.87E-22	Reduces abundance
Parabacteroides goldsteinii	-28.63	4.80E-23	8.43E-22	Reduces abundance
CAG-1427 sp000435675	-27.93	5.31E-22	8.39E-21	Reduces abundance
Dialister MIC7247	-27.85	6.92E-22	1.04E-20	Reduces abundance
Oscillibacter MIC7169	-27.75	9.55E-22	1.38E-20	Reduces abundance
Monoglobus pectinilyticus	-27.73	1.02E-21	1.41E-20	Reduces abundance
Tyzzerella MIC7714	-27.59	1.63E-21	2.16E-20	Reduces abundance
Anaerotignum sp001304995	-27.53	1.99E-21	2.48E-20	Reduces abundance
Collinsella sp003487125	-27.53	2.02E-21	2.48E-20	Reduces abundance
Bacteroides nordii	-27.48	2.37E-21	2.80E-20	Reduces abundance
Lachnoclostridium _A edouardi	-27.38	3.33E-21	3.81E-20	Reduces abundance
Streptococcus parasanguinis_B	-27.17	6.60E-21	7.00E-20	Reduces abundance
Hungatella hathewayi	-27.16	6.75E-21	7.00E-20	Reduces abundance

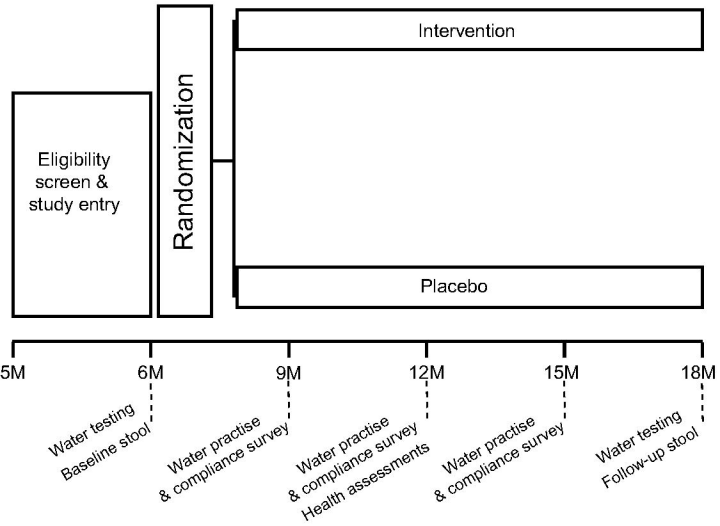
medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Tyzzerella MIC8139	-26.77	2.39E-20	2.27E-19	Reduces abundance
Acutalibacteraceae MIC8041	-26.64	3.66E-20	3.28E-19	Reduces abundance
Blautia hominis	-26.62	3.90E-20	3.41E-19	Reduces abundance
Pauljensenia sp001838165	-25.89	3.98E-19	3.00E-18	Reduces abundance
GCA-900066495 MIC8689	-25.85	4.48E-19	3.23E-18	Reduces abundance
Haemophilus_D sp001679485	-25.04	5.43E-18	3.60E-17	Reduces abundance
Veillonella_A sp000431435	-24.73	1.35E-17	8.63E-17	Reduces abundance
Turicibacter sanguinis	23.84	1.85E-16	1.12E-15	Increases abundance
Corynebacterium argentoratense	23.89	1.63E-16	1.00E-15	Increases abundance
Streptococcus sp001587175	24.56	2.29E-17	1.43E-16	Increases abundance
QAMM01 sp003150405	24.76	1.27E-17	8.29E-17	Increases abundance
Dorea sp900066555	25.63	8.79E-19	5.95E-18	Increases abundance
Clostridium_M lavalense	25.75	6.16E-19	4.26E-18	Increases abundance
Absiella dolichum	25.77	5.79E-19	4.09E-18	Increases abundance
Dorea phocaeense	25.87	4.16E-19	3.07E-18	Increases abundance
CAG-306 sp000980375	25.92	3.62E-19	2.80E-18	Increases abundance
Clostridium_M MIC9401	26.01	2.68E-19	2.12E-18	Increases abundance
Klebsiella pneumoniae	26.08	2.18E-19	1.77E-18	Increases abundance
Holdemania sp900120005	26.31	1.07E-19	8.84E-19	Increases abundance
UBA11774 sp003507655	26.49	5.88E-20	5.01E-19	Increases abundance
Massiliomicrobiota sp002160815	26.70	2.98E-20	2.75E-19	Increases abundance
Streptococcus MIC7545	26.90	1.58E-20	1.54E-19	Increases abundance
Eubacterium callanderi	27.15	7.00E-21	7.04E-20	Increases abundance
Holdemanella	27.22	5.59E-21	6.19E-20	Increases abundance

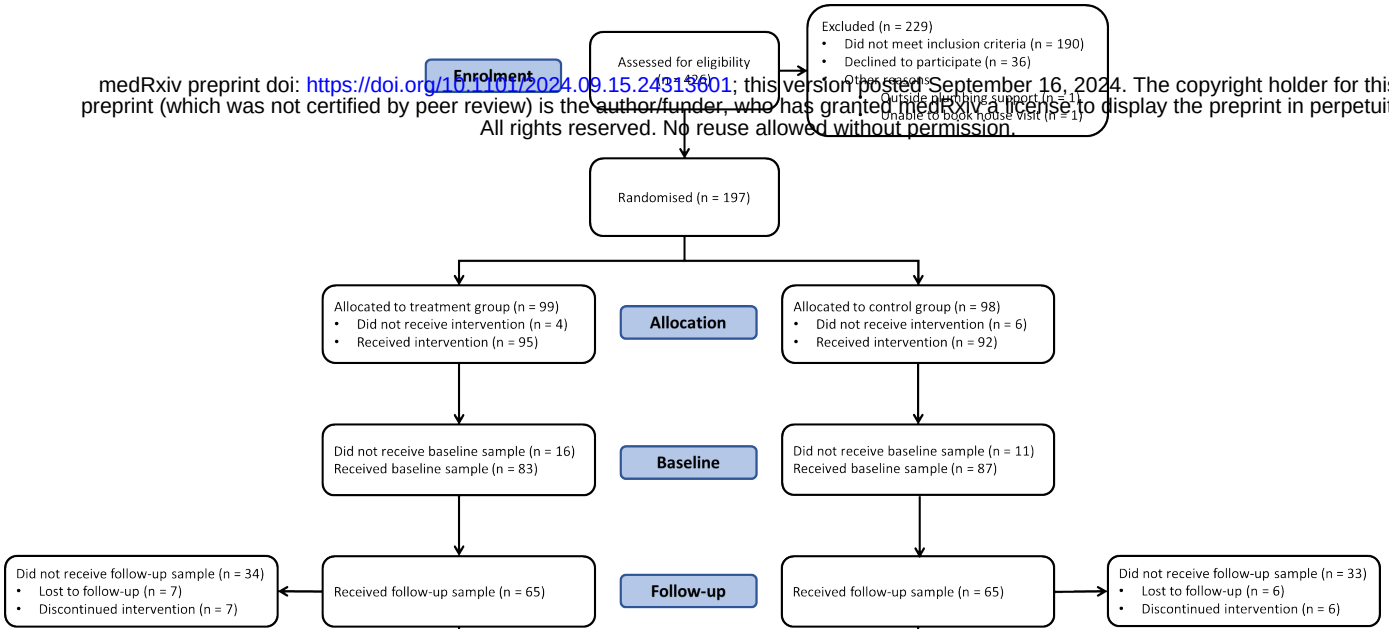
medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

sp002299315				
Eubacterium_G sp000432355	28.35	1.26E-22	2.09E-21	Increases abundance
Catenibacterium sp000437715	28.63	4.82E-23	8.43E-22	Increases abundance
Faecalibacterium prausnitzii_E	28.89	1.98E-23	3.87E-22	Increases abundance
CAG-245 sp000435175	28.91	1.87E-23	3.87E-22	Increases abundance
UBA1691 MIC9213	29.14	8.20E-24	1.94E-22	Increases abundance
Eubacterium_G sp000435815	29.27	5.24E-24	1.45E-22	Increases abundance
GCA-900066995 sp900291955	29.86	6.44E-25	2.14E-23	Increases abundance
CAG-303 sp000437755	30.00	3.87E-25	1.43E-23	Increases abundance
Bifidobacterium sp002742445	30.00	3.87E-25	1.43E-23	Increases abundance

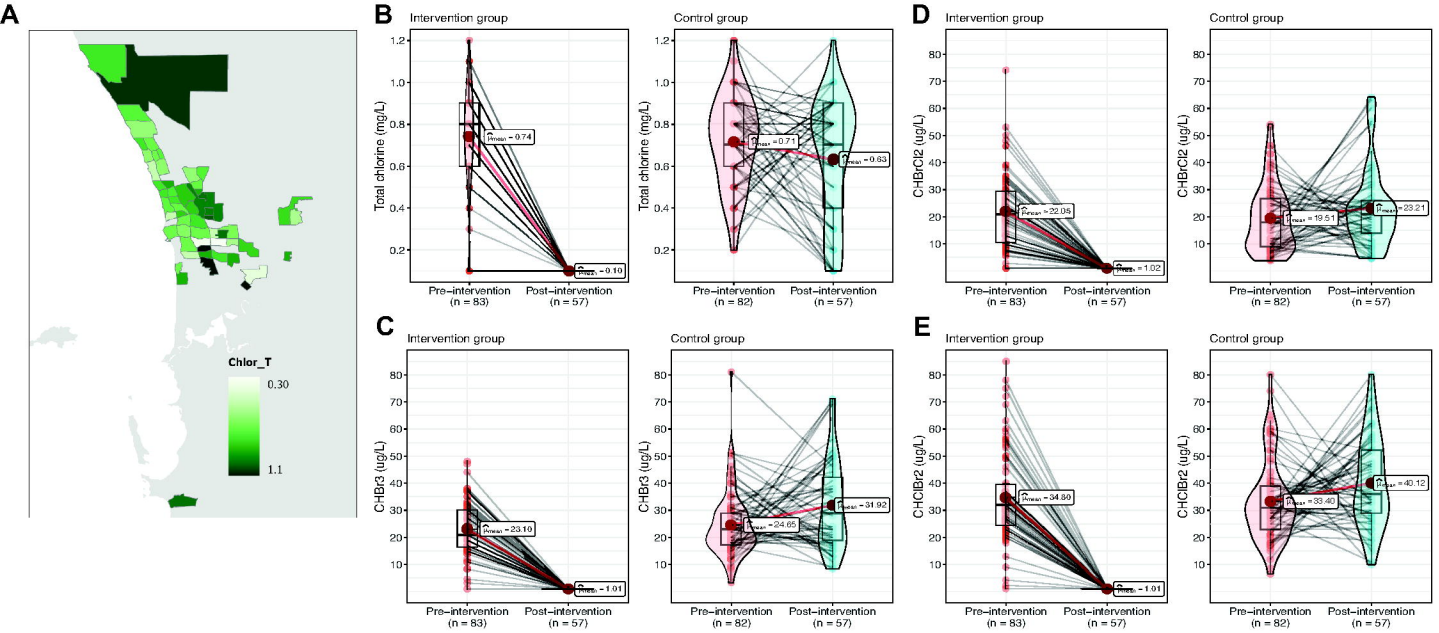
medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



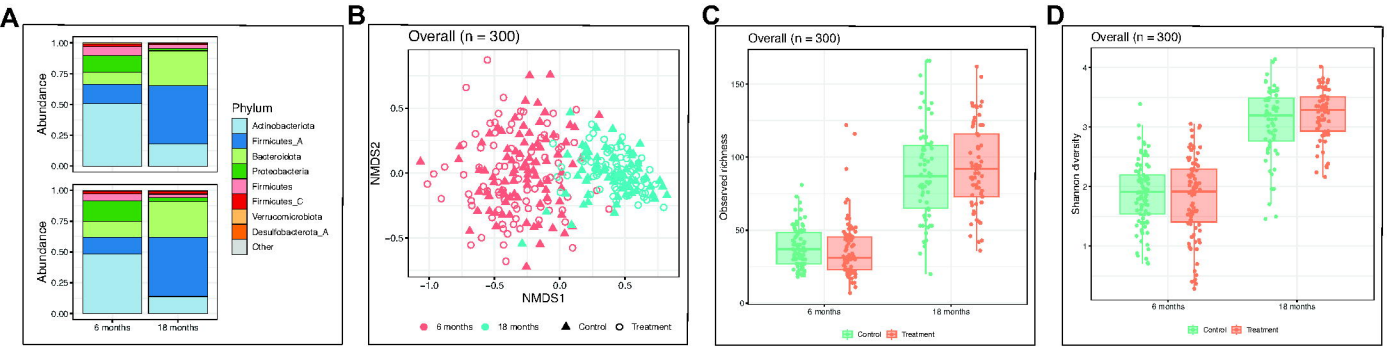
medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



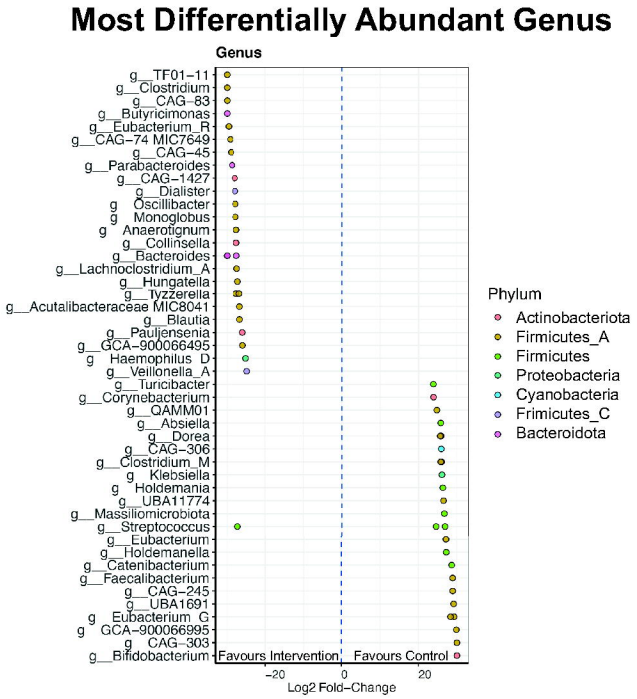
medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

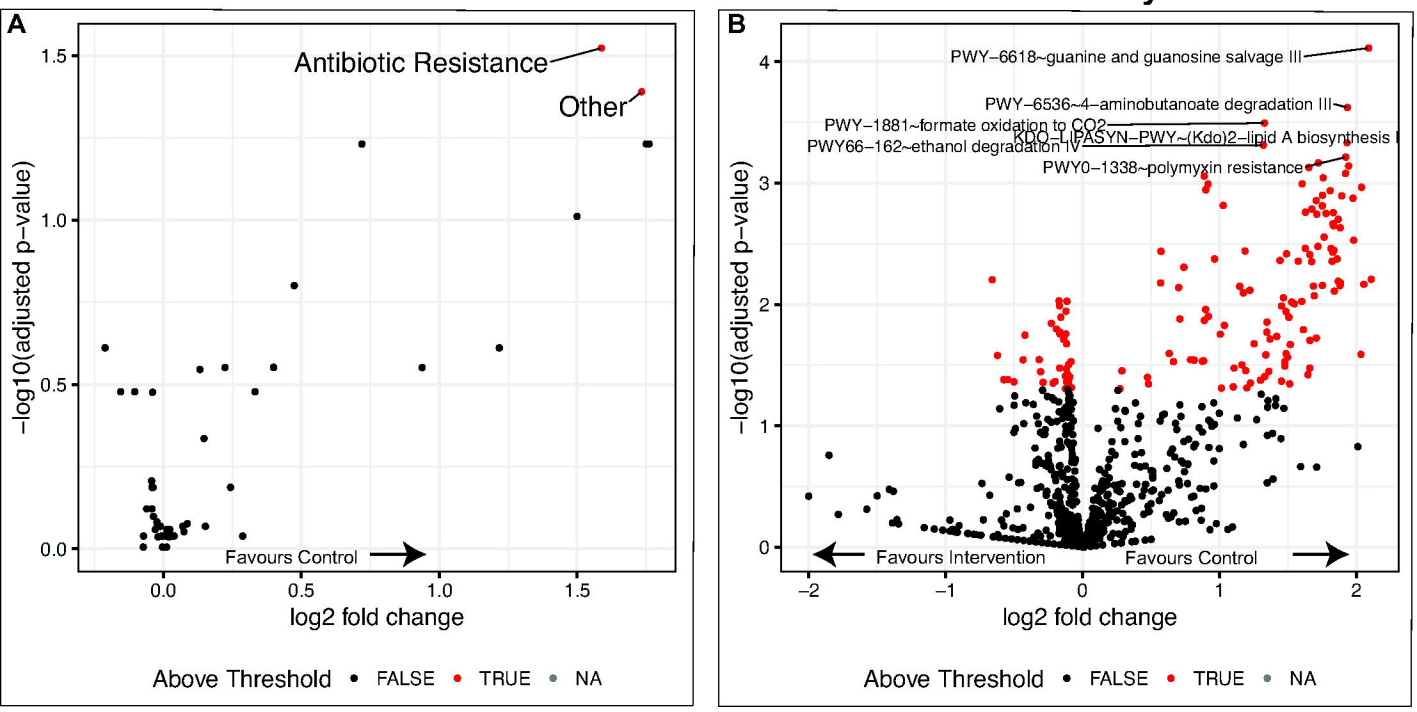


medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



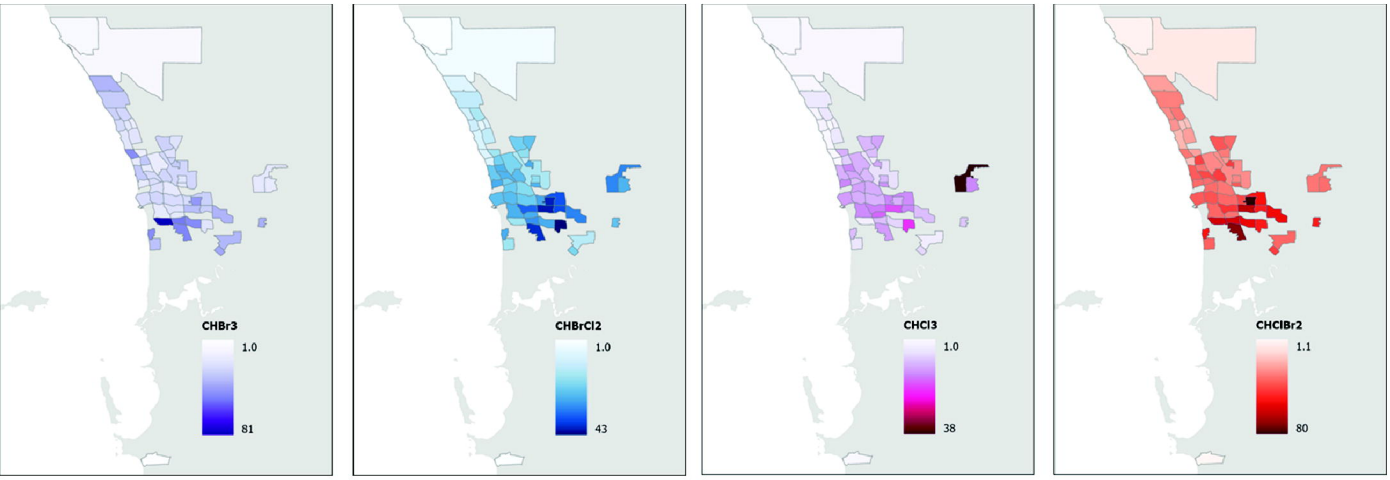
medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Volcano Plot of Treatment Effect on Metabolic Pathways



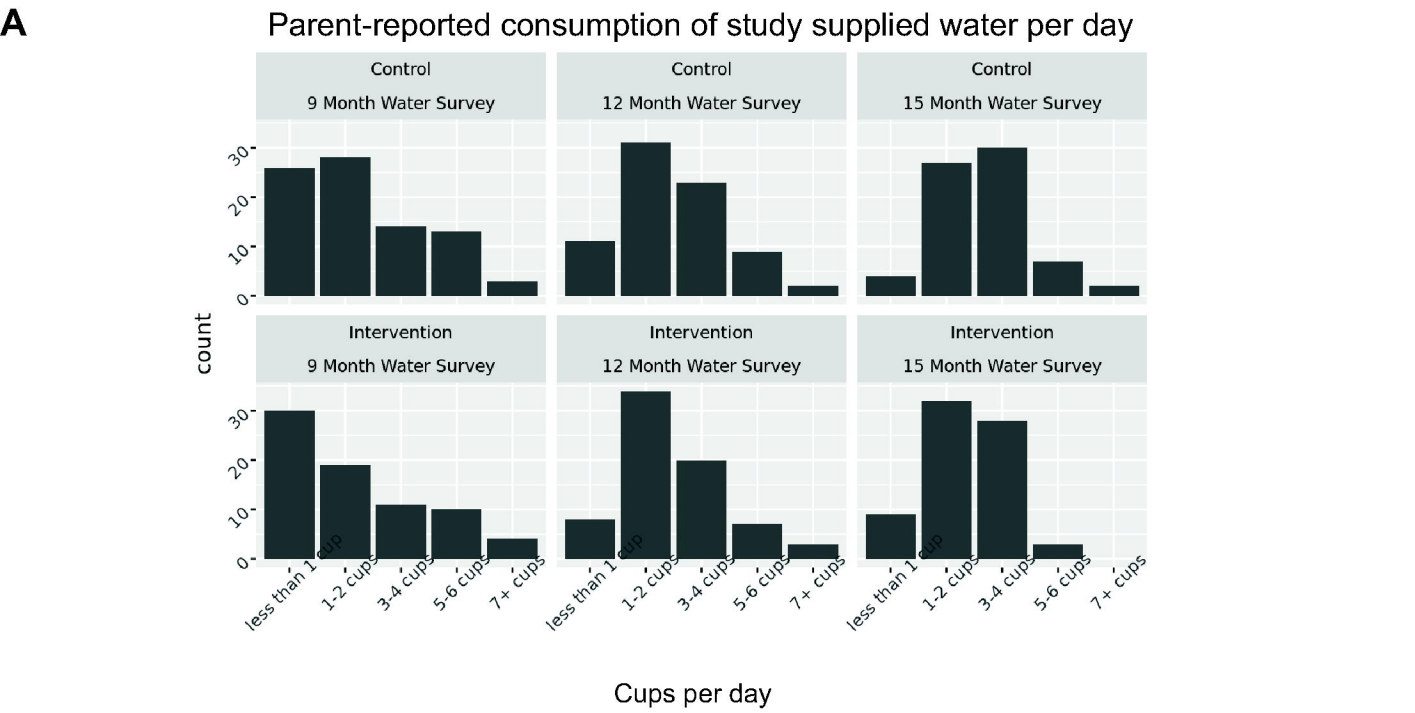
medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

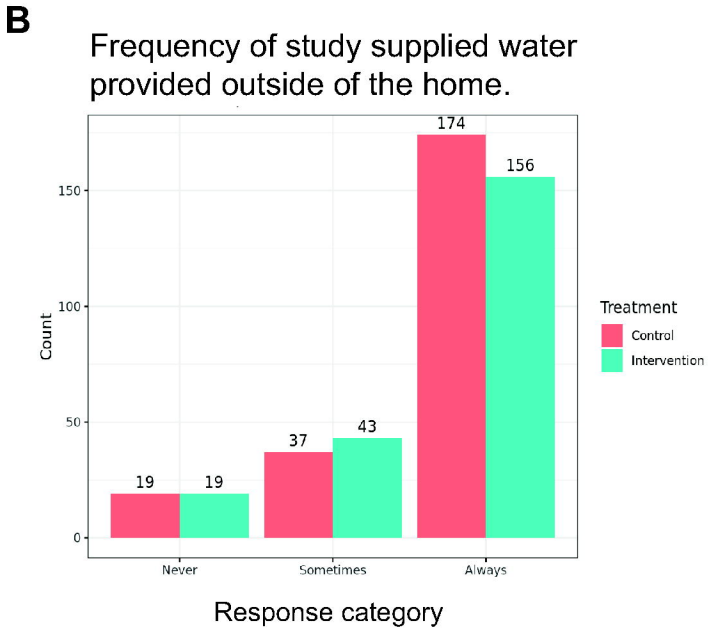
Disinfection by-products in Perth suburbs at study intake.

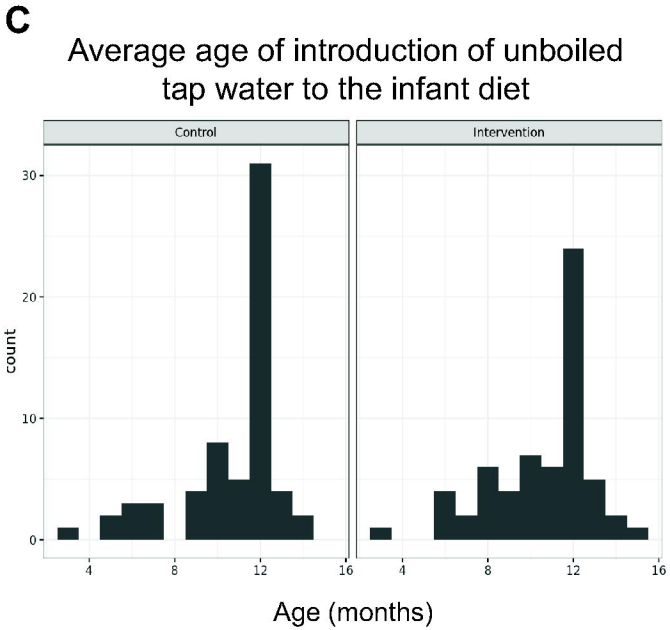


Bromoform **Bromodichloromethane** **Chloroform** **Chlorodibromomethane**

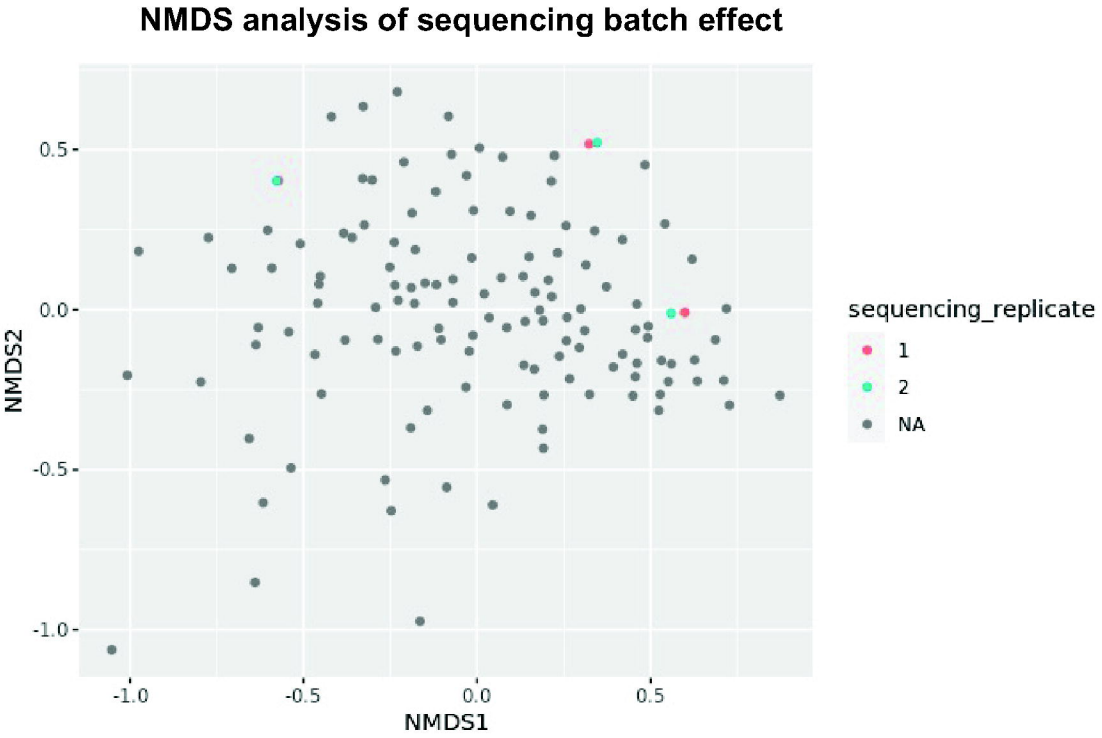
medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



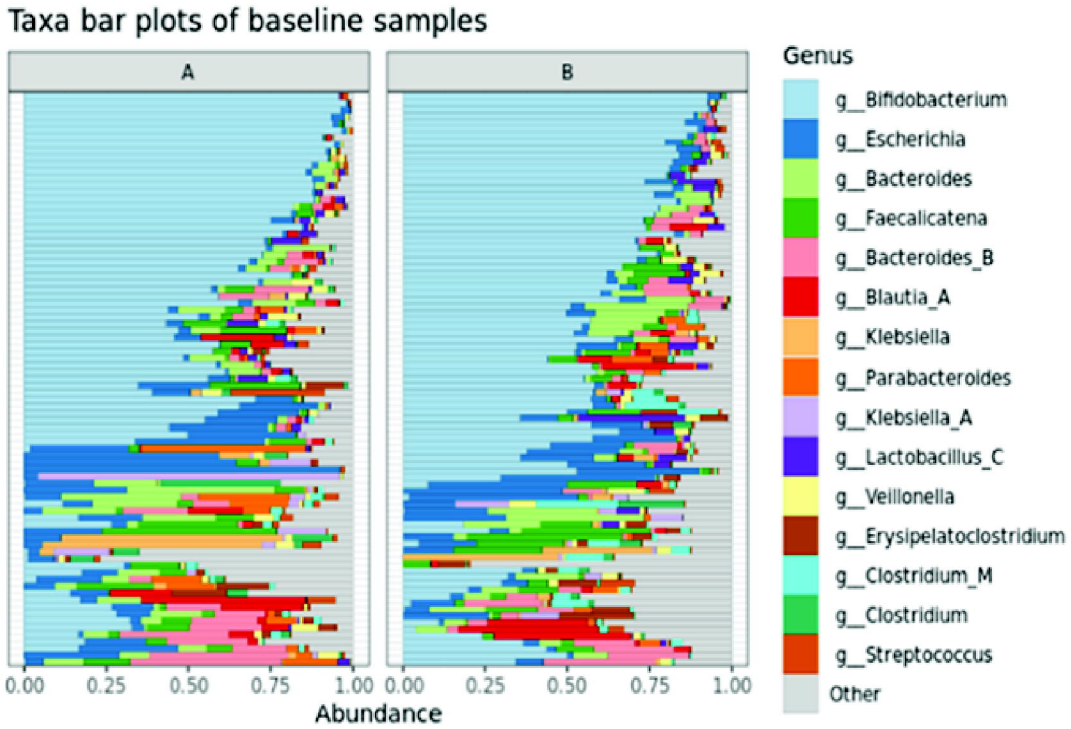


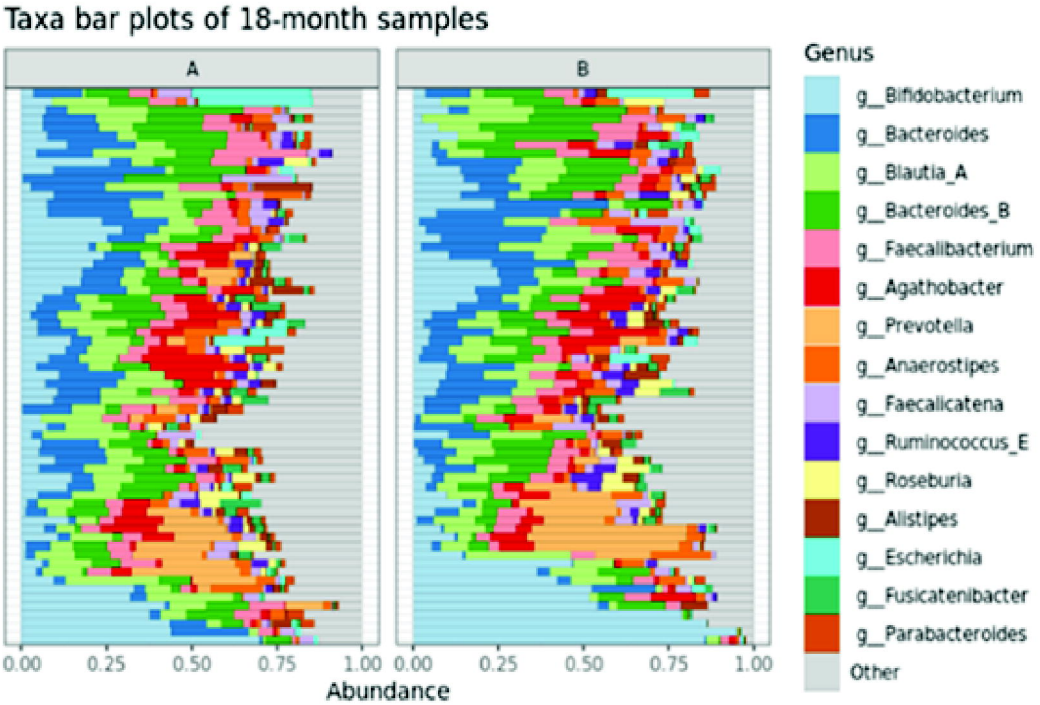


medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



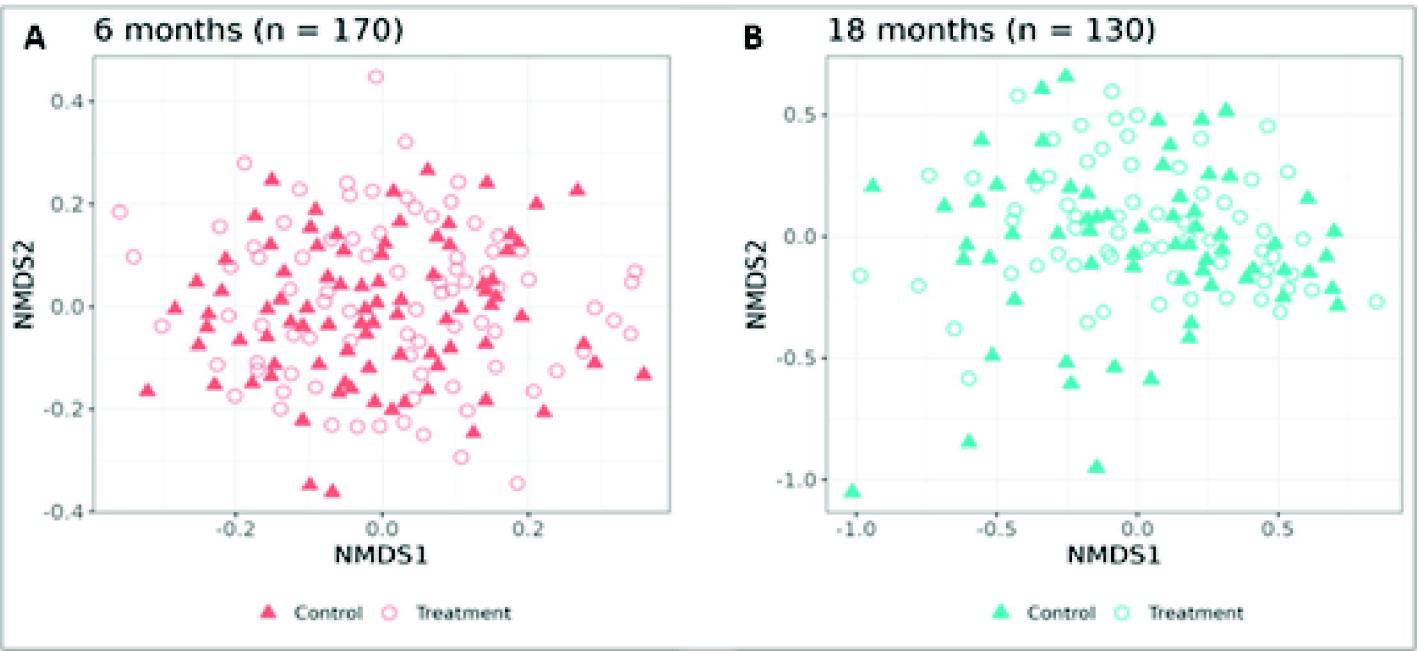
medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



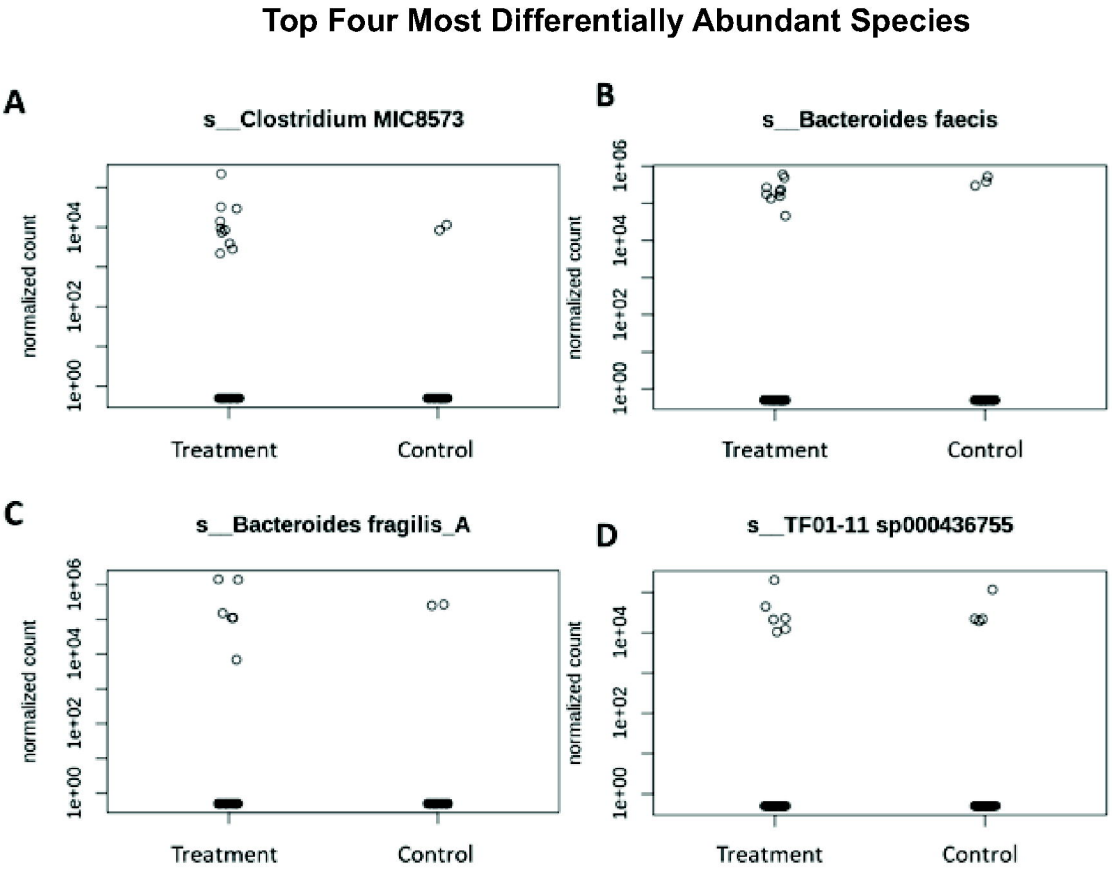


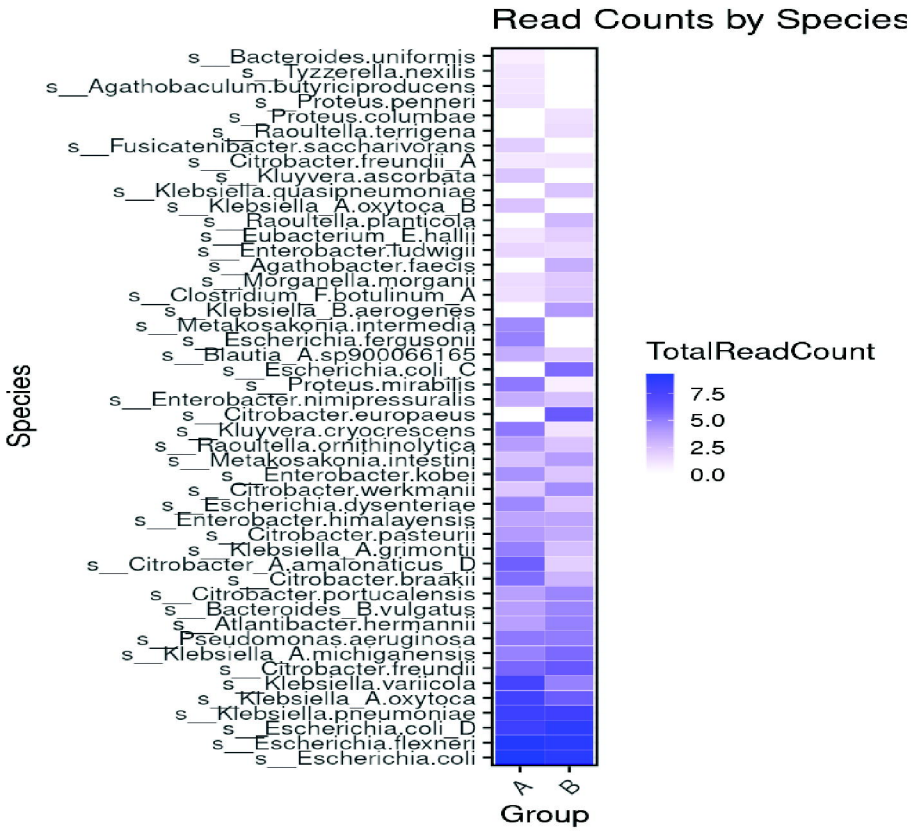
medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

NMDS analysis of treatment main effect



medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



medRxiv preprint doi: https://doi.org/10.1101/2024.09.15.24313601; this version posted September 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Log2 Sequencing reads for species annotated to the Antibiotic Resistance Metacyc pathway