



Gut microbiome and function are altered for individuals living in high fluoride concentration areas in Pakistan

Sara Bibi^{a,c}, Caroline Kerbirou^b, Uzma^c, Shona Mckirdy^b, Anastasiia Kostrytsia^c, Hifza Rasheed^d, Syed Ali Mustjab Akber Shah Eqani^a, Konstantinos Gerasimidis^b, Syed Muhammad Nurulain^{e,*}, Umer Zeeshan Ijaz^{c,f,g,**,1}

^a Department of Biosciences, COMSATS University Islamabad, 45550, Pakistan

^b School of Medicine, Dentistry & Nursing, Glasgow Royal Infirmary, Glasgow G31 2ER, UK

^c Water & Environment Research Group, University of Glasgow, Mazumdar-Shaw Advanced Research Centre, Glasgow G11 6EW, UK

^d National Water Quality Laboratory, Pakistan Council of Research in Water Resources (PCRWR), Islamabad, Pakistan

^e Department of Biosciences, Grand Asian University Sialkot, Pakistan

^f Department of Molecular and Clinical Cancer Medicine, University of Liverpool, Liverpool L69 7BE, UK

^g National University of Ireland, University Road, Galway H91 TK33, Ireland

ARTICLE INFO

Edited by Dr Muhammad Zia-ur-Rehman

Keywords:

Endemic fluorosis
Gut microbiota
MetaCyc pathways
Short chain fatty acids

ABSTRACT

Background: Endemic fluorosis refers to the condition when individuals are exposed to excessive amounts of fluoride ion due to living in a region characterized by elevated levels of fluorine in the drinking water, food, and/or air. In Pakistan, a substantial proportion of the population is thereby affected, posing a public health concern. **Objectives:** Assessing how the gut microbiota and its metabolic profiles are impacted by chronic exposure to fluoride in drinking water (that caused Dental Fluorosis) as well as to perceive how this microbiota is connected to adverse health outcomes prevailing with fluoride exposure.

Methods: Drinking water (n=27) and biological samples (n=100) of blood, urine and feces were collected from 70 high fluoride exposed (with Dental Fluorosis) and 30 healthy control (without Dental Fluorosis) subjects. Water and urinary fluoride concentrations were determined. Serum/plasma biochemical testing was performed. Fecal DNA extraction, 16S rRNA analysis of microbial taxa, their predicted metabolic function and fecal short chain fatty acids (SCFAs) quantification were carried out.

Results: The study revealed that microbiota taxonomic shifts and their metabolic characterization had been linked to certain host clinical parameters under the chronic fluoride exposure. Some sets of genera showed strong specificity to water and urine fluoride concentrations, Relative Fat Mass index and SCFAs. The SCFAs response in fluoride-exposed samples was observed to be correlated with bacterial taxa that could contribute to adverse health effects.

Conclusions: Microbial dysbiosis as a result of endemic fluorosis exhibits a structure that is associated with risk of metabolic deregulation and is implicated in various diseases. Our results may form the development of novel interventions and may have utility in diagnosis and monitoring.

1. Introduction

Fluoride occurrence in the biosphere is inevitable and groundwater fluoride exposure is an environmental hazard that poses risks to human health. People are severely affected around the globe where the problem is endemic and are unable to substitute the high fluoride containing

groundwater source. Naturally occurring fluoride (F⁻) minerals in soil and aquifer sediments among other hydrogeochemical factors, cause fluoride to accumulate in groundwater. The groundwater fluoride hazard map's machine learning tools have identified hotspots in numerous regions of Africa and Asia, as well as parts of eastern Brazil, western North America and central Australia. The majority of the 180 million

* Corresponding author.

** Correspondence to: Mazumdar-Shaw Advanced Research Centre (ARC), University of Glasgow, 11 Chapel Lane, Western Site, Glasgow, G11 6EW, UK.

E-mail addresses: syednurulain@gaus.edu.pk, syed.nurulain@comsats.edu.pk (S.M. Nurulain), Umer.Ijaz@glasgow.ac.uk (U.Z. Ijaz).

¹ <http://userweb.eng.gla.ac.uk/umer.ijaz>

<https://doi.org/10.1016/j.ecoenv.2024.116959>

Received 9 March 2024; Received in revised form 16 August 2024; Accepted 26 August 2024

Available online 3 September 2024

0147-6513/© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

people who are impacted globally live in Asia (51–59% of total population) and Africa (37–46% of total population) (Podgorski and Berg, 2022). Geographical location, geological structure and climatic factors of regions contribute towards distribution of fluoride in contaminated water sources. The countries severely impacted by fluoride contamination of groundwaters are Kenya, Libya, Ghana, Sudan, Burundi, Tanzania, Ethiopia, Algeria, China, India, Sri Lanka, Iran, Pakistan, Argentina, Germany, Chile, Jordan and Türkiye (Shaji et al., 2024). Pakistan, being located on the Tethyan Mobile belt (belt 3 of the 5 global fluoride belts), is facing the geogenic fluoride levels in groundwater and maximum fluoride concentration equal to 30 mg/L has been reported (Chowdhury et al., 2019; Rafique et al., 2009). As such fluoride is one water contaminant that the WHO believes (fluoride concentration in drinking water is allowed up to 1 mg/L for temperate climate regions) it has potential to impact the human health in an unfavorable way (Ahmad et al., 2022). Chronic fluoride exposure produces the recognized adverse effect of dental fluorosis (Buzalaf, 2018) while much excessive F⁻ consumption over an extended period of time can damage both skeletal and soft tissues by disrupting cellular metabolism and ionic transport (Barbier et al., 2010). Water fluoride levels raised above 1 mg/L are known to cause a linear decline in IQ. However, dose-response meta-analysis has indicated that even at low exposure levels, fluoride had a negative impact on children's IQ (Veneri et al., 2023). After a certain threshold of exposure (around 2.5 mg/L F⁻), drinking water with high fluoride content appears to have a non-linear effect on thyroid function, increase TSH release in children and raise the risk of several thyroid diseases (Jamandii et al., 2024). According to a recent study, 6% of the Pakistani population (13 million people) is susceptible to groundwater fluorosis (Ling et al., 2022). Chronic fluoride toxicity subjects the human body under physiological stress and it is anticipated that it would also affect the human gut microbiome. According to previous studies, the ingested fluoride has altered the intestinal symptomatology and microbial communities of the several tested animal models (Dionizio et al., 2021; Li et al., 2021; Melo et al., 2017; Miao et al., 2020). High fluoride concentration was demonstrated to be detrimental by reducing some of the beneficial bacteria while enhancing some of the harmful bacteria in an *in vitro* fecal fermentation model (Chen et al., 2021). It has been discovered that the existence of acidogenic oral microbiota, that promotes the development of caries, is a substantial factor in dental fluorosis (Wang et al., 2021). Another study uncovered the significant impact of fecal microbiota depleted carbohydrate metabolism pathways on the risk of dental fluorosis in children (Zhou et al., 2023).

Bacteria predominantly make up the current knowledge of gut microbiota. Among humans, more than 1000 species and more than 500 distinct genera have been identified, demonstrating high diversity at lower taxonomic levels (Falony et al., 2016; Huttenhower et al., 2012; Qin et al., 2010). The specific gut environmental factors (such as acidity, substrate concentration, molecular oxygen etc.) influence the microbial communities differently at distinct gastro-intestinal (GI) compartments (Costello et al., 2009; Ding and Schloss, 2014). Diet, lifestyle and the host's DNA impact the gut microbiota since birth. Gut microbiota interacts by regulating the host metabolic pathways, forming a network of metabolic, signalling and immune-inflammatory axes that physiologically link the gut with other organs, such as the brain and liver. A greater understanding of these axes is essential to define therapeutic options to manage the gut microbiota in an effort to avoid disease and improve health (Nicholson et al., 2012). Focusing on the functional capabilities of the microbiome could be more informative rather than just its taxonomic makeup. The functional capacity of the microbiome is more consistent across different people than the specific microbial taxa they host (Visconti et al., 2019). Moreover, the metabolic activities of the microbiome have a more extensive and stronger influence on the metabolites present in both blood and feces compared to the influence of specific microbial species. This highlights the importance of microbial metabolic pathways in shaping the metabolic landscape of the human body.

Under unwanted conditions, the microbiota is susceptible to variations in its composition and in its interactions with the host. Consequently, the altered microbiota lends itself to the emergence of metabolic disorders and influences the process of a disease development (Berg et al., 2020). Such dysbiosis refers to the loss of good bacteria and increase of potentially bad bacteria when compared to eubiosis (Aldars-García et al., 2021). SCFAs, being the final byproduct of microbial fermentation, perform immunomodulatory and metabolic regulatory roles within the human host. Variations in the concentration of these SCFAs produced in the gut can have adverse consequences on the human physiology (O'Reilly et al., 2023). Finding associations between the gut microbes and the host health factors is considered a robust strategy to comprehend health and illness, and can help with the design of a clinical study (Manor et al., 2020). The gut flora acts as the mediator by producing the host stress response and related outcomes. Characterizing the human gut microbiota in relation to high fluoride exposure is in its inception stage. There is currently insufficient knowledge on the theory that fluoride exposure may disrupt the homeostasis between the gut microbiota and the human host. Moreover, the patterns of SCFAs profiles are defined here for the healthy controls and the fluorosis subjects. Together with the gut microbial metabolic pathways, these findings have implications in the aetiology of fluorosis.

2. Materials and methods

2.1. Recruitment of study subjects

This case-control study followed certain criteria. The inclusion criteria were that the exposed participants lived in the fluorosis endemic regions (with atleast past 6 months exposure to groundwater fluoride) and had dental fluorosis. The exclusion criteria were to avoid people with chronic pathologies like diabetes, cardiovascular disease, liver dysfunction, neurological and gastrointestinal disorder, pregnancy or cancer etc. Fluoride endemic areas were identified according to the reports: Draban (District D.I.Khan); Lalmato, Thor Sheikhan, Zaggi Village and Sher Bridge (District Mullagori); Sailab Colony, Wanda Baharawala, Wanda Bopanwala, Ali Khel Wanda and Wanda Gul Khan Khel (District Mianwali) (Bibi et al., 2023). The residents of these areas used groundwater for drinking purpose (having fluoride concentration >1 mg/L) and were physically examined for dental and skeletal fluorosis manifestations. Skeletal fluorosis was not observed, however, some subjects complained of its early symptoms. Several sectors of Islamabad (non-endemic regions) were visited for control subjects' recruitment. Control subjects drank bottled water (containing no detectable fluoride) and did not exhibit any symptoms of fluorosis seen in endemic areas. The chronically groundwater fluoride exposed (n=70) and healthy controls (n=30) voluntarily participated in the study. The participants were aged between 17 and 60 years. Anthropometrics, blood pressure and blood glucose were recorded. All participants followed an omnivorous diet. The history of GI illnesses, other chronic pathologies and use of any medication in the three months prior to sampling was not reported.

2.2. Samples collection and processing

Drinking water samples, venous blood, urine and fecal samples were taken from all the participants of the study. Water samples were obtained from the different water sources like handpump, water bore, tubewell and bottle. Venous blood (5 mL) was drawn once from all participants, separated using different tubes with and without anticoagulants for the plasma and serum fractions. The aliquoted samples were kept at -80 °C for subsequent biochemical profiling. Urine collected was filtered through 0.22 µm pore size syringe filters and kept frozen till analysis. Intact fecal samples were collected in clean specimen containers, homogenized, weighed and separated for DNA extraction (150 mg) and SCFAs quantification (1 g). The aliquot for DNA isolation was placed at -80 °C till further processing. The 1 mL of NaOH (1 M)

was used to homogenize the 1 g of the fecal sample (for SCFA), freeze-dried for 36 hrs. and kept at room temperature.

2.3. Determination of host biochemical parameters

Hemoglobin was determined as cyanmethemoglobin. Plasma was used to test butyrylcholinesterase (BChE) activity whereas erythrocyte-acetylcholinesterase (E-AChE) activity was determined using whole blood dilutions (Worek et al., 1999). Modified procedures were followed for serum malondialdehyde (MDA), reduced glutathione (GSH), superoxide dismutase (SOD) and catalase (CAT) quantification respectively (Boutin et al., 1989; Ellman, 1959; Misra and Fridovich, 1977; Wilbur et al., 1949). ELISA techniques were used to assess serum interleukin-1 β (IL-1 β), interleukin-6 (IL-6), tumor necrosis factor-alpha (TNF- α) and C-reactive protein (CRP) levels. Serum total cholesterol (TC), triglycerides (TG), high-density lipoprotein (HDL) and low-density lipoprotein (LDL) were detected by their specific colorimetric assay kits. The atherogenic-index of plasma was calculated ($AIP = \log(TG/HDL)$). Serum amylase, lipase and insulin were measured with their specific commercial kits and HOMA-IR was calculated ($HOMA-IR = \text{FastingInsulin} * \text{FastingGlucose}/22.5$). Enzymatic activities for serum aspartate transaminase (AST), alanine transaminase (ALT), lactate dehydrogenase (LDH), alkaline phosphatase (ALP) and gamma-glutamyltransferase (γ -GT) activities were confirmed by their reagent kits. Serum creatinine and urea were quantified using kinetic method and colorimetric method kits respectively. Blood urea nitrogen (BUN) and estimated glomerular filtration rate (eGFR) calculated by Chronic Kidney Disease Epidemiology for Pakistan (eGFR CKD EPI PK) were computed (Ahmed et al., 2017). Fluoride ion concentration was quantified based on SPADNS colorimetric method in water and urine samples. The details of the instruments and reagent kits used are provided in the [supplementary information](#).

2.4. Quantification of Fecal SCFAs concentrations

Freeze dried fecal sample (100 mg) was weighed in 15 mL polypropylene centrifugation tubes. Then the sample was mixed with 300 μ l distilled water, 100 μ l of internal standard solution (73.25 mM; 2-Ethylbutyric acid) and 100 μ l of orthophosphoric acid (OPA) until complete homogenization. Diethyl ether (1.5 mL) was poured into the sample and mixed with it using an IKA (VXR Vibrax, Germany) orbital shaker at 1200 rpm for 1 minute. The diethyl ether containing SCFAs was recovered three times and collected in a sterile tube. This pooled extract was put in a clean glass vial, crimped tightly to avoid evaporation and loaded onto the GC autosampler. One microliter of each sample was injected into the GC-FID (Agilent 7820 A, USA), using an Agilent DB-Wax Ultra Inert (15 m, 0.53 mm, 1.00 μ m) capillary column. Nitrogen was the carrier gas. The GC was programmed with the following parameters: hold at 80 $^{\circ}$ C for one minute, increase by 15 $^{\circ}$ C per minute until maximum temperature of 210 $^{\circ}$ C, held for 1 minute. To determine the concentration of each fatty acid, the area under the curve (AUC) of each fatty acid was computed and integrated against the calibration curve in relation to the AUC of the internal standard. Results are presented as percentages of the total SCFA as well as per mass of wet and dry fecal material (μ mol/g).

2.5. Fecal DNA isolation

DNA extraction from 150 mg of well homogenized fecal sample was carried out with DNeasy PowerSoil Pro kit (QIAGEN, Hilden, Germany). After extraction, DNA purity and concentration were measured with Invitrogen Qubit 2.0 Fluorometer (Life Technologies Corporation, 5791 Van Allen Way, Carlsbad, California).

2.6. 16S rRNA gene amplicon library and sequencing

The DNA samples were processed using the Glasgow Polyomics standard 16S rRNA protocol which encompasses a two-step amplification process to produce libraries encompassing the V3 and V4 regions of the bacterial 16S ribosomal subunit. The first amplification used degenerate primers to well conserved areas, the second included the addition of barcodes and sequencing adapters. Sequencing of the libraries were performed using V3 chemistry with 2 \times 300 bp on the MiSeq.

2.7. Bioinformatics

A total of 10,261,693 paired end reads were obtained from 100 samples. We then constructed the Operational Taxonomic Units (OTUs) at 99% similarity using the methods described previously (Ijaz et al., 2018), with minor alterations: a) taxonomic assignment using SILVA SSU Ref NR database release v.138 (Quast et al., 2013) was performed; and b) phylogenetic tree and other post-processing were made within the QIIME2 framework (Bolyen et al., 2019). To obtain predictive metabolic potential of the microbial communities, PICRUSt2 (Douglas et al., 2020) was utilised within the QIIME environment to give the corresponding KEGG enzymes (10,543 enzymes for 100 samples) and MetaCyc pathways (488 enzymes for 100 samples) table. As per authors' recommendation (Douglas, 2022), we had used the standard parameters `-p-hsp-method pic -p-max-ntsi 2` in the pipeline: `qiime picrust2 full-pipeline`. Once the abundance table and taxonomy were obtained, QIIME2 was used to consolidate both tables in a single BIOM file (summary statistics of reads/sample as follows: [Min: 37,728; 1st Quartile: 59,263; Median: 65,128; Mean: 65,480; 3rd Quartile: 70,480; Max: 107,212]). The BIOM file along with the phylogenetic tree, and the associated meta data were then subjected to statistical analyses using the hypotheses considered in this study.

2.8. Statistical analyses

For the biochemical features, the non-parametric Kruskal-Wallis test was used. Before applying the test, we auto scaled the data, and after adjusting the *p*-values (using Benjamini Hochberg method for multiple comparisons) returning from Kruskal-Wallis test, we fitted a random forest classifier on the selected features using fluoride exposure groups.

As a pre-processing step, contaminants such as *Mitochondria* and *Chloroplasts*, along with unassigned OTUs at all levels were removed, following the same strategy as given in <https://docs.qiime2.org/2022.8/tutorials/filtering/>. This gave a final table of *n*=100 samples \times *P*=17,971 OTUs on which we applied different multivariate statistical algorithms.

The R's vegan package (Dixon, 2003) was used for computing community-level metrics including alpha (Shannon entropy, rarefied richness, and Simpson index) and beta diversity analyses. To facilitate the pair-wise analysis of variance (ANOVA), R's `aoV()` function was used to calculate *p*-values, which were later incorporated in the alpha diversity figures.

The beta diversity distances of OTU abundance table were used in the Principal Coordinate Analysis (PCoA) to compare samples. Three type of distances were used: (i) *Bray-Curtis distance* to facilitate analysis of compositional changes; (ii) *Unweighted UniFrac distance* estimated using R's Phyloseq package (McMurdie and Holmes, 2013) to facilitate phylogenetic changes between samples; and (iii) *Hierarchical Meta-Storms (HMS)*, a beta diversity distance specifically targeting KEGG Orthologs (KOs) and computes functional changes by considering multi-level pathway hierarchy these KOs are part of (Zhang et al., 2021a). Using `adonis()` function from Vegan package, PERMANOVA analysis was performed to establish percentage variability in microbiome structure with different sources of variation. In addition to the above-mentioned beta diversity distances, weighted UniFrac distance

using the Phyloseq package was also incorporated in the PERMANOVA analysis.

To report on the fraction of the microbiome that prevails in most of the samples (the core microbiome), a minimum of 85 % prevalence in all samples was used as a cut-off. For this purpose, R's microbiome package (Lahti et al., 2017) was utilised.

To find genera that were at 2 log2 fold different ($p < 0.05$) between multiple cohorts in this study, DESeqDataSetFromMatrix() function from the DESeq2 package (Love et al., 2014) was used. The reported p -values were adjusted for multiple comparisons internally by the procedure.

To find a minimal subset of genera that had changed with respect to continuous predictors considered in this study, CODA LASSO regression using the coda_glmnet() function from R's coda4microbiome package (Calle et al., 2023) was used. The fitted regression was of the form $y_i = \beta_0 + \beta_1 \log(x_{1i}) + \dots + \beta_j \log(x_{ji}) + \epsilon_i$ (for i -th sample and j -th feature, with x_{ji} being the abundance of feature), and where the outcome y_i was a continuous outcome variable. The model used two constraints: a) all β -coefficients summed up to zero, that gave two sets of features: those that were positively associated with continuous outcome, and those that were otherwise; and b) the optimization function incorporated a LASSO shrinkage that forced some of the beta coefficients to go to zero, particularly those that were statistically insignificant in relation to the outcome. We had employed CODA LASSO on the top 100 most abundant genera (OTUs collated at genus level using taxonomy recovered from the SILVA SSU Ref NR database release v.138, and then with the MetaCyc pathways table recovered from the PICRUSt2 software.

To explore if certain genera existed within a narrow range of continuous covariates, we had utilised R's Specificity Package (Darcy et al., 2022). The Rao's Quadratic Entropy (RQE) as $RQE = \sum_{i=1}^{s-1} \sum_{j=i+1}^s D_{ij} p_i p_j$ was calculated where genus abundance $p_i p_j$ was the multiplication of the abundances of a specific genus in samples i and j , respectively, and was weighted by the difference in the continuous covariate value D_{ij} . A null modelling procedure using 999 random permutations was used for the RQE values. Deviation of the original RQE from the average of RQEs of these random permutations gave a "Spec" number, ranging from -1 to $+1$, with 0 as the null hypothesis that the genus weights were randomly ordered with regard to sample identity, with perfect specificity when Spec approached -1 and perfect cosmopolitanism when spec approached $+1$. The null modelling procedure provided additional p -values for significance. For visualization purposes, the plotting was restricted to the lowest 25th quartile (i.e., those genera that were specific). Furthermore, we had used the algorithm twice, once for biochemical parameters (including water, urine and blood parameters), and once for SCFAs.

We then applied DIABLO algorithm (Rohart et al., 2017) to integrate $M = 3$ datasets denoted by $X^{(1)} (N \times P_1)$, $X^{(2)} (N \times P_2)$, $X^{(3)} (N \times P_3)$, where $X^{(1)}$ represented the TSS+CLR (Total Sum Scaling + Centralised Log Ratio) normalised OTUs abundance table collated at genus level, $X^{(2)}$ represented the autoscaled SCFA table (whether Dry, Wet or % SCFAs), and $X^{(3)}$ denoted the dummified (i.e., binary representation of labelling such as Control and High Exposure). The DIABLO algorithm then factorized the datasets into scores and loading vectors in such a way that the covariance between the scores of these datasets was maximized, i.e., for $q = 1, 2, \dots, Q$, DIABLO solved for each component $h = 1, \dots, H$:

$$\begin{aligned} \operatorname{argmax}_{a_h^{(1)}, \dots, a_h^{(Q)}} \quad & \sum_{q,j=1, q \neq j}^Q c_{qj} \operatorname{cov}(X_h^{(q)} a_h^{(q)}, X_h^{(j)} a_h^{(j)}) \quad \text{s.t.} \quad \|a_h^{(q)}\|_2 \\ & = 1 \text{ and } \|a_h^{(q)}\|_1 \leq \lambda^{(q)} \end{aligned}$$

Where $a_h^{(q)}$ for component h was the loading vector associated with the matrix $X_h^{(q)}$ of the data set $X^{(q)}$, and $C = \{c_{qj}\}_{q,j}$ was the design matrix.

With additional l_1 penalty $\|a_h^{(q)}\|_1 \leq \lambda^{(q)}$ (where $\lambda^{(q)}$ being the penali-

zation parameter), we were ensuring some of the loading components (acting as weighting for each microbial genus) to go to zero filtering out genera that were insignificant. The other constraint $\|a_h^{(q)}\|_2 = 1$ ensured the loading vector to have unit magnitude. C was a $Q \times Q$ design matrix where we had used a full weighted design with $c_{qj} = 0.1$ between data matrices ($X^{(1)}$ and $X^{(2)}$) and $c_{qj} = 1$ where the outcome was involved ($X^{(1)}$ and $X^{(3)}$; $X^{(2)}$ and $X^{(3)}$). These were based on the recommendations by the mixOmics package to suggest a reasonable trade-off between maximizing correlation across datasets and maximizing the discrimination across different groups coded in the outcome $X^{(3)}$.

To predict the number of principle components, associated loading vectors, and the number of discriminants features in DIABLO algorithm, block.splsda() and tune.block.splsda() functions were used. The optimisation process involved a) we fine-tuned the model using leave-one-out cross-validation by splitting the data into training and the testing sets, and then identified the number of loading components that maximised the class separation using any of the distance measures; and b) we found the non-zero coefficients for each of the loading components. For dry SCFAs, we had used ncomp=2 components and dist="max.dist", for wet SCFAs, we used ncomp=2 components and dist="centroids.dist", and for %SCFAs, we had used ncomp=2 and dist="centroids.dist" in tune.block.splsda() function, respectively.

3. Results

3.1. Characteristics of the study subjects

The target population was grouped by drinking-water fluoride exposure status which was confirmed by the presence of dental fluorosis ("severe" according to Dean's modified index based on full-mouth scoring) as clinical outcome. Table S1 shows the different properties of sampling areas. Fluoride exposed cases were recruited from fluorosis endemic areas whereas controls were chosen from non-endemic areas. Table 1 displays the baseline characteristics for the two groups i.e. high fluoride exposure and healthy control groups. Additionally, coupling Kruskal Wallis with the importance measures from the classifier, we were able to rank which parameters had an altered response in the high fluoride exposed group (Figs. S1, S2 and S3). The control subjects had higher BMI, RFM index, HDL, triglycerides, catalase activity, eGFR CKD EPI PK and fecal valeric acid (C5) percentage (1.785 %), iso-valeric acid (IC5) percentage (1.557 %), caproic acid (C6) percentage (0.366 %) and heptanoic acid (C7) percentage (0.046 %) values. The high fluoride exposed group had elevated blood pressure variability, oxidative stress rendered by the MDA levels, E-AChE and BChE activities, BUN to creatinine ratio, ALP, γ GT and lipase activities. The pro-inflammatory cytokines (TNF- α , IL-1 β and IL-6) and CRP were significantly deregulated for the high fluoride exposed individuals. The calculated HOMA-IR and AI-P values depicted the high fluoride exposed individuals to mostly have significant insulin resistance and risk of dyslipidemia. The fluoride exposed individuals had increased urine fluoride concentrations than controls. Acetic acid, butyric acid and iso-caproic acid in dry fecal matter were quantified significantly between the two groups with high fluoride exposure group revealing high levels of these SCFAs as compared to the control group.

3.2. Microbial diversity measures segregate control and high fluoride exposed samples

For the OTUs, alpha diversity analyses (Fig. 1A) indicated higher mean richness in controls than fluoride exposed samples. Though the Shannon and Simpson indices were slightly higher for the fluoride exposed samples, these were not significant. For KEGG orthologs (KOs) (Fig. 1B), alpha diversity was predominantly ($p < 0.05$) elevated in the control gut samples than high fluoride exposed samples. For the recovered MetaCyc Pathways (Fig. 1C), the richness metric was significantly

Table 1

Summary statistics of the subjects analyzed. A total of 100 subjects were assigned to two groups: High Fluoride Exposure and Healthy Controls. Summary statistics are given, including the median and interquartile range (IQR) of continuous variables such as age and height.

Variables		High Fluoride Exposure		Healthy Controls		Variables		High Fluoride Exposure		Healthy Controls				
		(n=70)	(n=30)	(n=70)	(n=30)			(n=70)	(n=30)					
1	Sex	Male	34	13	16	TNF- α (ng/L)	Median	20.32	6.62	17	IL-6 (ng/L)	Median	12.05	6.24
		Female	36	17			IQR	12.05	6.24			IQR	12.05	6.24
2	Age (Years)	Median	27.5	28.5	17	IL-6 (ng/L)	Median	16.93	8.49	18	IL-1 β (ng/L)	Median	13.87	5.97
		IQR	17	10.75			IQR	13.87	5.97			IQR	11.16	8.23
3	Dental Fluorosis	Yes	70	0	19	Hemoglobin (μ mol/L)	Median	5.4	2.37	20	E-AChE Activity (mU/ μ mol Hb)	Median	0.44	0.16
		No	0	30			IQR	5.4	2.37			IQR	0.44	0.16
4	Height (m)	Median	1.68	1.67	21	BChE Activity (μ mol/L/min)	Median	28.38	41.69	22	BUN (mmol/L)	Median	9.12	12.6
		IQR	0.18	0.1			IQR	9.12	12.6			IQR	10.27	5.82
5	RFM Index	Median	25.55	31.38	23	eGFR _{CKD EPI PK} (mL/min/1.73 m ²)	Median	0.44	0.16	24	Alkaline Phosphatase (IU/L)	Median	0.05	0.03
		IQR	9.23	12.89			IQR	0.05	0.03			IQR	0.02	0.01
6	Systolic BP (mm Hg)	Median	131	118.5	25	γ -glutamyltransferase (IU/L)	Median	10.27	5.82	26	Lactate dehydrogenase (IU/L)	Median	10.27	5.82
		IQR	15.5	18			IQR	5.51	1.07			IQR	5.51	1.07
7	Diastolic BP (mm Hg)	Median	79.5	76	27	Urine Fluoride Concentration (mg/L)	Median	95.75	114.5	28	Water Fluoride Concentration (mg/L)	Median	327.5	42.5
		IQR	12.75	11			IQR	95.75	114.5			IQR	117.75	233
8	Blood Glucose (mg/dL)	Median	115.9	94.76	29	Water pH	Median	37.75	9.43	30	Water ORP (mV)	Median	327.5	42.5
		IQR	26.23	11.3			IQR	37.75	9.43			IQR	117.75	233
9	Serum Insulin (μ IU/mL)	Median	16.14	6.47	26	Lactate dehydrogenase (IU/L)	Median	28.3	41	27	Urine Fluoride Concentration (mg/L)	Median	117.75	233
		IQR	10.37	2.93			IQR	28.3	41			IQR	117.75	233
10	HOMA IR	Median	4.69	1.58	28	Water Fluoride Concentration (mg/L)	Median	32.83	10.25	29	Water pH	Median	117.75	233
		IQR	3.15	0.71			IQR	32.83	10.25			IQR	117.75	233
11	Total Cholesterol (mg/dL)	Median	203.13	194.8	29	Water pH	Median	261	241.5	30	Water ORP (mV)	Median	261	241.5
		IQR	80.1	12.06			IQR	261	241.5			IQR	261	241.5
12	HDL (mg/dL)	Median	43.48	61	27	Urine Fluoride Concentration (mg/L)	Median	1.84	0.06	28	Water Fluoride Concentration (mg/L)	Median	1.84	0.06
		IQR	11.35	15.48			IQR	1.84	0.06			IQR	0.55	0.06
13	Atherogenic Index Plasma	Median	0.39	0.29	29	Water pH	Median	2.3	0	30	Water ORP (mV)	Median	2.3	0
		IQR	0.21	0.05			IQR	2.3	0			IQR	0.7	0
14	Catalase (U/mg protein)	Median	0.41	0.55	29	Water pH	Median	8.43	8.52	30	Water ORP (mV)	Median	8.43	8.52
		IQR	0.21	0.17			IQR	8.43	8.52			IQR	0.41	0.96
15	MDA (nmol/mg protein)	Median	2.07	0.99	30	Water ORP (mV)	Median	72	62	30	Water ORP (mV)	Median	72	62
		IQR	1.14	0.37			IQR	72	62			IQR	17	28

($p < 0.05$) higher for the control group samples. These results elucidate that the high fluoride exposure is possibly linked to modulation of the gut microbiome to the occurrence of decremented bacterial KOs and metabolic pathways. By comparing the samples between the two groups (Fig. 1D, E and F), the OTUs from control samples were markedly differed from that in high fluoride exposed samples based on Bray-Curtis ($R^2 = 0.03192$, $p = 0.001$), UniFrac ($R^2 = 0.02704$, $p = 0.001$) and HMS ($R^2 = 0.02352$, $p = 0.029$) distances.

The Table S2 summarizes the overall difference between the two microbial communities' structure in relation to factors/covariates of interest. Among these variables, the SCFA values (except Butyric acid percentage and Caproic acid dried), BChE activity, dental fluorosis, study group, sex, lipase activity, fecal water percentage, residence, RFM index, urine fluoride concentration and water fluoride concentration notably ($p < 0.05$) varied the composition, phylogeny and function of gut microbial communities simultaneously.

3.3. Key representative genera of high fluoride exposed samples

While looking at the topmost abundant genera present across both cohorts, controls and high fluoride exposed samples (Fig. 2), the immediate observation was that the sample profiles looked quite similar with *Bifidobacterium*, *Lactobacillus* and *Enterococcus* being the top 3 most dominant genera. This seemed to suggest that the dysbiosis or changes were subtle and were then further explored using the core microbiome analyses (Fig. 3) where we looked at genera that persisted in at least 85 % of the samples. Amongst the core microbiome, the unique genera found in the control samples were *Escherichia-Shigella*, *Terrisporobacter* and *Ruminococcaceae;uncultured* and those that were unique to the high fluoride exposed samples were *Catenibacterium*, *Erysipelotrichaceae;uncultured*, *Prevotella* and *Eggerthellaceae;Slackia*. Finally, we wanted to see

which genera were differentially abundant between the control and the high fluoride exposed group. We assumed two approaches: a) where we had explicit labeling of the samples available as high fluoride exposed group (> 1 mg/L fluoride content), and used DESeq2 analysis to find genera that were at least 2 log₂ fold different, and b) not assuming discrete labeling of the samples, and incorporating the quantitative value of fluoride concentration in a regression model (CODA-LASSO) to find two subsets of genera (those that increased in abundance as fluoride concentration increased, and those that decreased). DESeq2 suggested *Sarcina*, *[Eubacterium]ruminantium_group*, *Peptococcus*, *[Eubacterium]_xylanophylum_group*, *Lachnospiraceae_NK4A136_group*, *Lachnospira* *ceae_UCG-003*, *Enterorhabdus* and *Solobacterium* were more abundant in high fluoride exposure group while *Pediococcus*, *Ruminococcaceae-CAG_352*, *Dialister*, *Bacteroides*, *Fusicatenibacter* and *Erysipelotrichaceae_UCG_003* were highly abundant in controls (Fig. 4). On the other hand, CODA-LASSO suggested that increase in fluoride concentration led to increase in abundance of *Collinsella*, *Lachnospiraceae_FCS020_group*, *[Eubacterium]_siraeum_group*, *Subdoligranulum*, *Corynebacterium*, *Slackia*, *Marvinbryantia*, *Lactobacillus*, *Holdemanella*, *Peptoniphilus*, *Lachnospiraceae_UCG_003*, *Negativicoccus*, *Lachnospiraceae_NK4A136_group*, *Actinomyces*, *Catenibacterium*, *Gordonibacter*, *Rothia*, *Peptococcus* and *Clostridium_sensu_stricto_1* whereas increasing fluoride concentration caused decrease in abundance of *Dorea*, *Dialister*, *Ruminococcaceae_UBA1819*, *Pediococcus*, *Escherichia-Shigella*, *Christensenellaceae_R7_group*, *Butyricoccus*, *Intestinibacter*, *Family-XIII_AD3011_group*, *Fusicatenibacter*, *Coriobacteriaceae_UCG_003*, *Butyrivibrio*, *Sellimonas*, *Akkermansia*, *Coprobacillus*, *Libanicoccus* and *Ruminococcaceae_CAG_352* (Fig. 5A). The common genera between both approaches in association of fluoride exposure were given as additional annotation (*) and were *Peptococcus*, *Lachnospiraceae_NK4A136_group* and *Lachnospiraceae_UCG_003*.

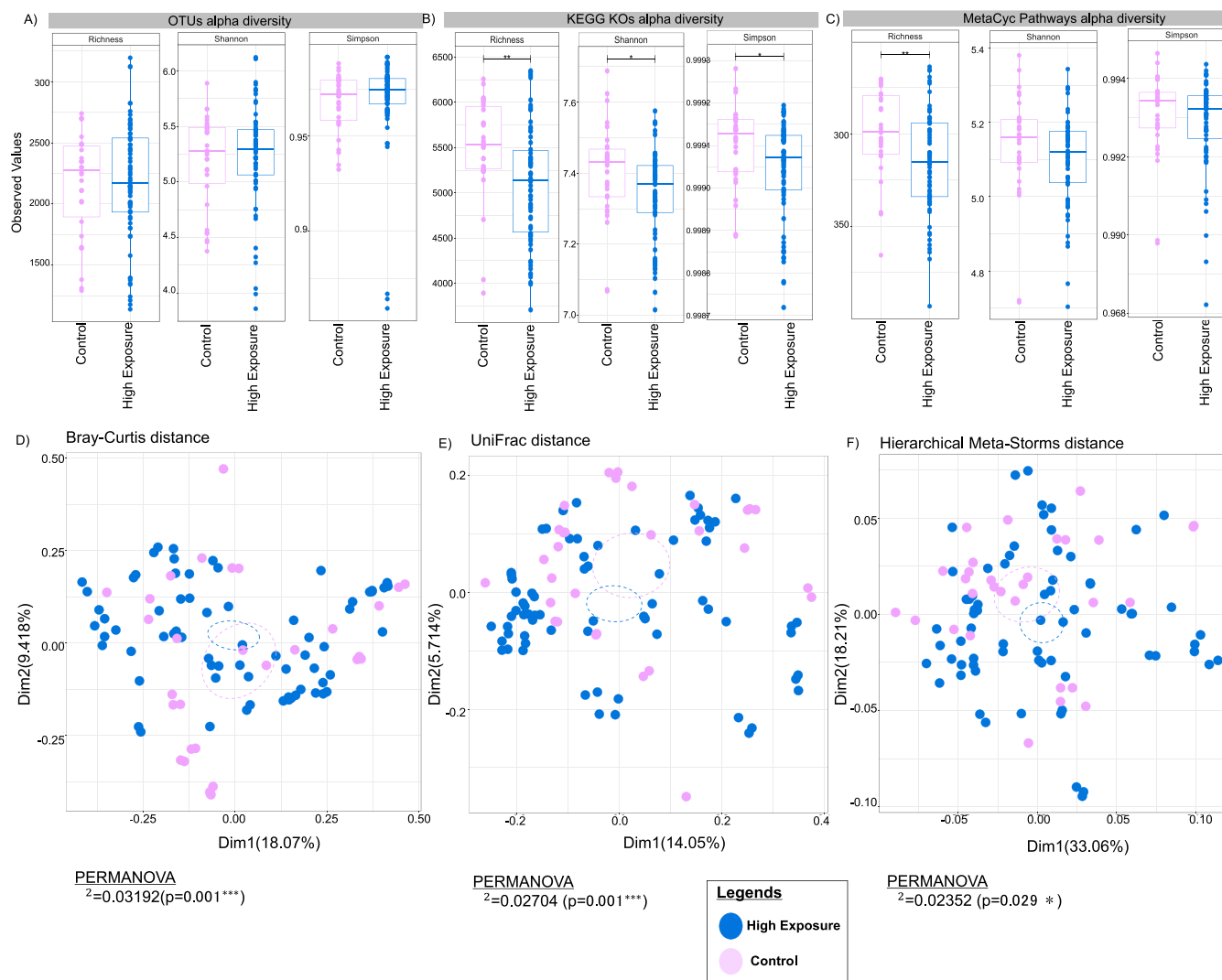


Fig. 1. Microbiota (Bacterial) Diversity Metrics for Control and High Fluoride Exposure Groups. Alpha diversity described at three levels (A) OTUs (B) KEGG Orthologs and (C) MetaCyc Pathways. The three indices used for alpha diversity are Richness (Observed taxa), Shannon index (richness and evenness) and Simpson index (taxa dominance). Lines connect two categories where the differences were significant (ANOVA) with $*p < 0.05$, $**p < 0.01$, or $***p < 0.001$. Distances between groups' samples were determined via Principal Coordinates Analysis using beta diversity estimates such as (D) Bray-Curtis distance i.e. between-samples compositional dissimilarity, (E) UniFrac distance i.e. the fraction of unshared branch lengths between all OTUs in two samples and (F) Hierarchical Meta-Storms distance i.e. the effect of inter-function relations on microbiome distances. PERMANOVA multivariate analysis was used to test for significant differences between groups.

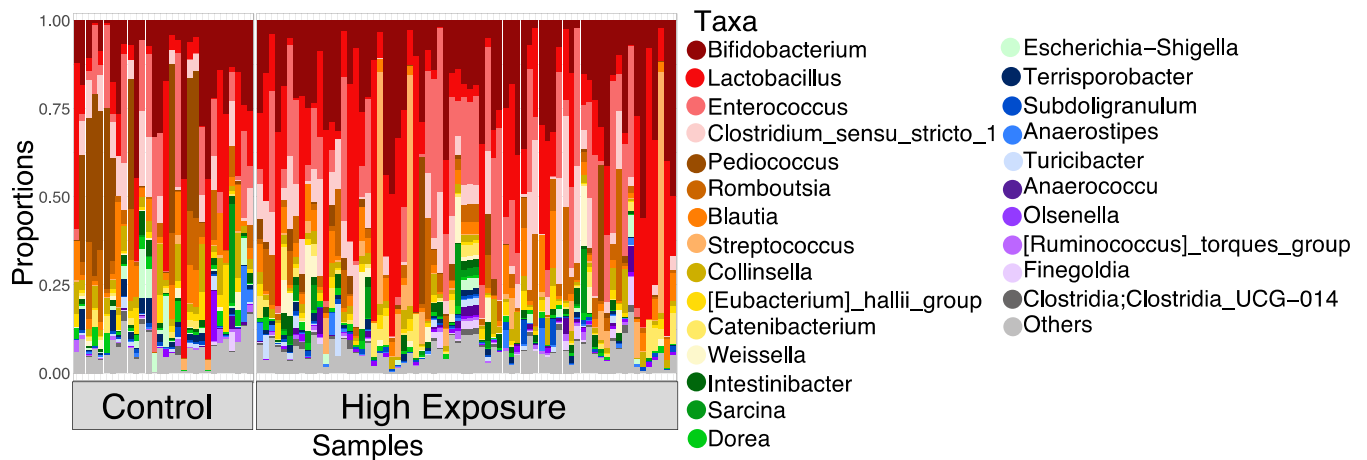


Fig. 2. Taxa bars showing the relative abundance of 25 most abundant bacterial genera of gut microbiome of the control (n=30) and high fluoride exposed (n=70) samples. The bars with different colours correspond to different genera. The taxa key is arranged according to the genera abundance levels. The top samples profiles looked similar between the groups but the relative abundance of the constituents varied.

3.4. Key changes in potential metabolic function of microbial communities

Similar to the previous approach, we have also used CODA-LASSO to associate MetaCyc pathways recovered using PICRUSt2 approach (Fig. 5B). Those that associated positively with increasing concentration of water fluoride were PWY-3001: superpathway of L-isoleucine biosynthesis I, HSERMETANA-PWY: L-methionine biosynthesis III, PWY-7242: D-fructuronate degradation, MET-SAM-PWY: superpathway of S-adenosyl-L-methionine biosynthesis, RIBOSYN2-PWY: flavin biosynthesis I (bacteria and plants), PWY-6906: chitin derivatives degradation, HOMOSER-METSYN-PWY: L-methionine biosynthesis I, PWY-6397: mycolyl-arabinogalactan-peptidoglycan complex biosynthesis and LACTOSECAT-PWY: lactose and galactose degradation. MetaCyc pathways that associated negatively with fluoride concentration were GLYCOCAT-PWY: glycogen degradation I, ARGORNPROST-PWY: L-arginine degradation (Stickland reaction), PWY-5104: L-isoleucine biosynthesis IV, GALLATE-DEGRADATION-II-PWY: gallate degradation I, PWY-6165: chorismate biosynthesis II (archaea), PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis, PWY-5973: cis-vaccenate biosynthesis, PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II and PWY-7446: sulfoquinovose degradation I.

3.5. Taxa and function of microbial communities associated with clinical parameters and other sources of variation

CODA-LASSO approach was then used with several other quantitative clinical outcomes, water oxidation reduction potential (ORP) and pH as shown in supplementary Figs. S4-S29 respectively. Their significance and implications published in scientific literature are presented in Table S3 for genera and in Table S4 for metabolic pathways. The top most positive genera and pathways signatures for these parameters were: for alkaline phosphatase [*Lachnospiraceae_FCS020_group*, PWY-5676: acetyl-CoA fermentation to butanoate]; for atherogenic index plasma [*Oscillospiraceae_UCG_005*, PWY-7616: methanol oxidation to carbon dioxide]; for blood glucose [*Staphylococcus*, PWY-6148: tetrahydromethanopterin biosynthesis]; for blood urea nitrogen [*Lachnospiraceae_FCS020_group*, COLANSYN-PWY: colanic acid building blocks biosynthesis]; for catalase [*Fournierella*, PWY-5181: toluene degradation III (aerobic) (via p-cresol)]; for diastolic blood pressure [*Lachnospiraceae_FCS020_group*, PWY-5743: 3-hydroxypropanoate cycle]; for estimated Glomerular Filtration Rate calculated by Chronic Kidney Disease Epidemiology for Pakistan [*Sellimonas*, PWY-6339:syringate degradation]; for erythrocytes

acetylcholinesterase [*Lachnospiraceae_FCS020_group*, PWY-5861: superpathway of demethylmenaquinol-8 biosynthesis]; for butyrylcholinesterase [*Oscillospiraceae_UCG_005*, PWY-6507: 4-deoxy-L-threo-hex-4-enopyranuronate degradation]; for high-density lipoprotein [*Erysipelotrichaceae_UCG_003*, PPGPPMET-PWY: ppGpp metabolism]; for height [*Catenibacterium*, HISTSYN-PWY:L-histidine biosynthesis]; for haemoglobin [*Lachnospiraceae_UCG_010*, CRNFORCAT-PWY: superpathway of demethylmenaquinol-8 biosynthesis]; for Homeostasis Model Assessment for Insulin Resistance [*Lachnospiraceae_ND3007_group*, PWY-6944: androstenedione degradation I (aerobic)]; for Interleukin-1beta [*Lachnospiraceae_ND3007_group*, PWY-5743:3-hydroxypropanoate cycle]; for Interleukin-6 [*Fournierella*, PWY-6944: androstenedione degradation I (aerobic)]; for lactate dehydrogenase [*Romboutsia*, PWY-7295:L-arabinose degradation IV]; for whole blood acetylcholinesterase [*Lachnospiraceae_FCS020_group*, GALACT-GLUCUROCAT-PWY: superpathway of hexuronide and hexuronate degradation]; for malondialdehyde [*[Eubacterium]_ruminantium_group*, P621-PWY: nylon-6 oligomer degradation]; for relative fat mass index [*[Eubacterium]_hallii_group*, METHYL-GALLATE-DEGRADATION-PWY:methylgallate degradation]; for serum insulin [*Lachnospiraceae_ND3007_group*, PWY-6944:androstenedione degradation I (aerobic)]; for TNF-alpha [*Gastranaerophilales*, PWY-181:photorespiration]; for total cholesterol [*Incertae_Sedis*, ORN-DEG-PWY: superpathway of ornithine degradation]; for urine fluoride [*Collinsella*, P161-PWY: acetylene degradation (anaerobic)]; for systolic blood pressure [*Lachnospiraceae_FCS020_group*, GLUCONEO-PWY: gluconeogenesis]; for water oxidation reduction potential [*Enterococcus*, PWY-5392:reductive TCA cycle II] and for Water pH [*Collinsella*, HEXITOLDEGSUPER-PWY:superpathway of hexitol degradation (bacteria)].

3.6. Species specific to a narrow range of observed clinical parameters

Whilst majority of the statistical tests explored associations or correlations (whether positive or negative) between microbial species and functions, and the observed clinical outcomes, we also applied a recent "Specificity" analysis. This as opposed to previous approaches explored if a particular genus existed only inside the narrow range of covariates considered. The method used a null modelling procedure to compute a "spec" number. The number offered a threshold to determine between genera that were *Cosmopolitan* (>0; existed in full range), or *Specific* (<0; existed in narrow range). The Fig. S30A shows violin plot of "spec" values to subject characteristics, biochemical parameters, water, urine and fecal properties with width of the violin representing the number of

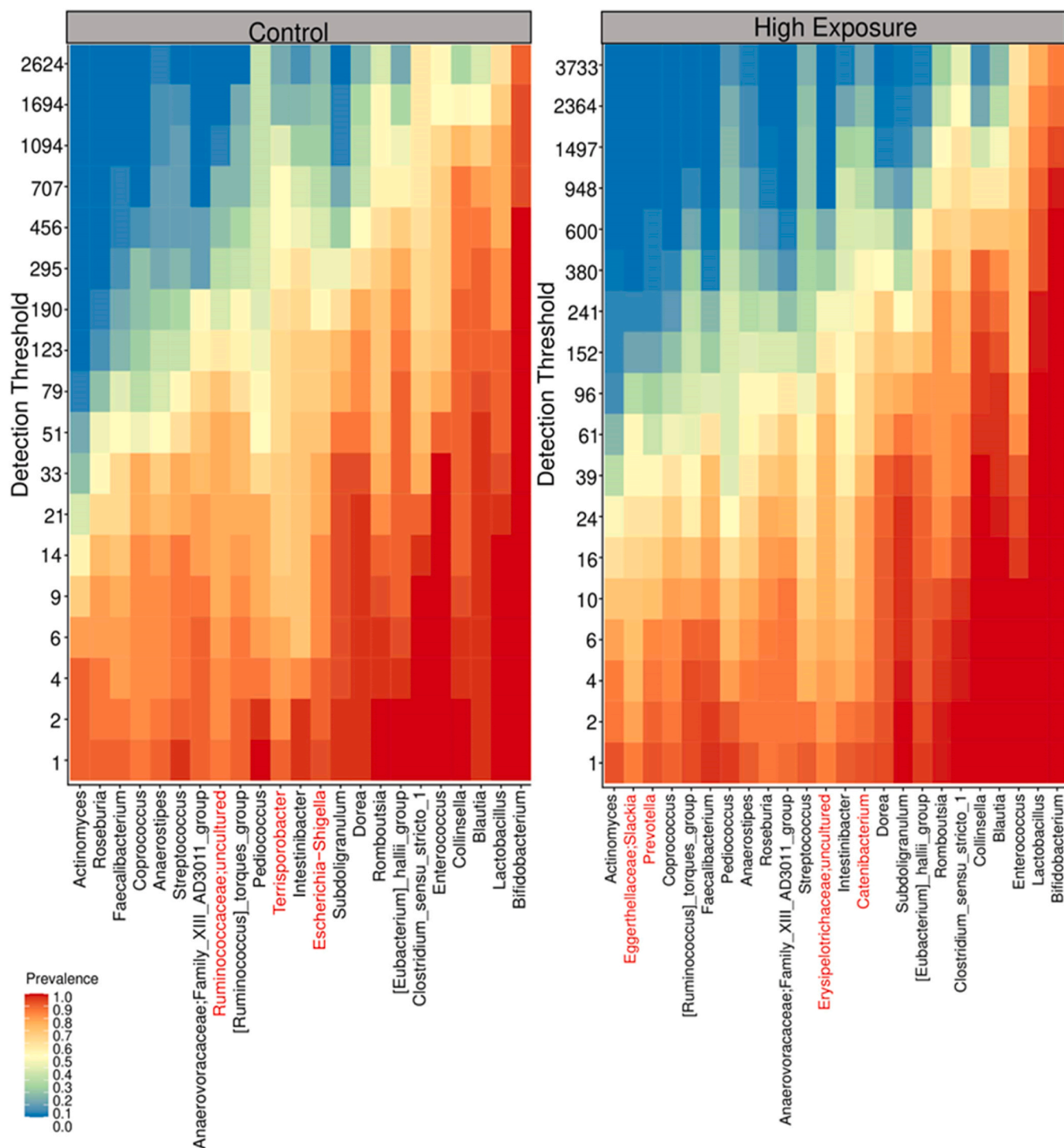


Fig. 3. The core microbiome heatmaps for the control and high fluoride exposed samples; made up of the presence of any genus with a minimum of 85 % prevalence in all samples. The y-axis represents the detection thresholds from lower to higher abundance values and color shading shows the prevalence of each taxon among samples for each abundance threshold. The prevalence decreases as the detection threshold increases. The highlighted (red) undermentioned genera were unique to each group.

genera with those particular value. Pairwise spec correlations are displayed in Fig. S30B, along with correlation coefficients (r) for each pairwise comparison (i.e., whether those genera that are specific to one clinical covariate, are also specific to other clinical covariate). The similar analysis is given for short chain fatty acid values in Fig. S31A and B. As such, the pairwise specificity correlation relationships between the any two variables in the dataset is shown to spot any patterns between the various variables and their complicity with microbiota.

Overall, this seems like a helpful tool for understanding the interplay of clinical parameters and how they may be related to the corresponding active microbiota performing roles synergically.

Three genera namely *Family_XIII_UCG_001*, *Lachnospiraceae_UCG_003* and *Fournierella* revealed strong specificity to water fluoride content (Fig. 6). Genera like *Lachnospiraceae_UCG_001*, *[Eubacterium]_xylanophilum_group* and RF39 were specific to propionic acid percent (Fig. S32). Butyric acid percent-specific genera were

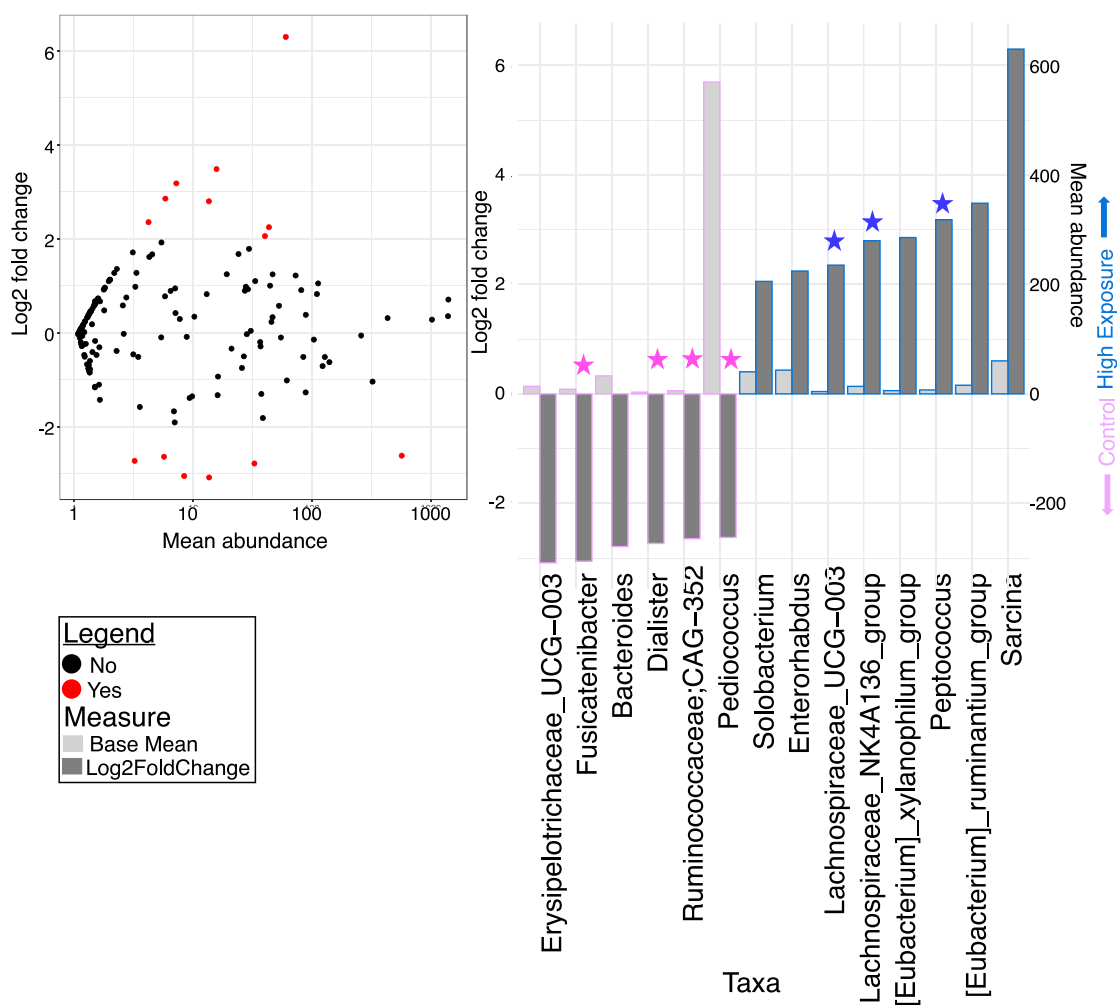


Fig. 4. Differential abundance of bacterial genera between control and high fluoride exposed groups. The plots show the log2 fold changes and base means for genus level taxa that were statistically significant at the 0.05 level as implemented in DESeq2. Positive values indicate the log2fold change for taxa over represented in high fluoride exposure. Negative values indicate the log2fold change for taxa under-represented in high fluoride exposure. The bar plots show Log2 fold change in abundance between groups (dark grey bar) and the mean abundance across all the samples (light grey bar). The starred taxa are common between Figs. 4 and 5.

Muribaculaceae, *Sellimonas*, *[Clostridium]_methylpentosum_group* and *Alloscardovia* (Fig. S33). The genera showing specificity to valeric acid dry were *Oscillibacter*, *Deftuviitaleaceae_UCG_011*, *Fournierella*, *Lachnospiraceae_NK4B4_group*, *Alloprevotella*, *Finogoldia*, *Incertae_Sedis*, *Gastranaerophilales* (Fig. S34) while genera associated with valeric acid percent were *Deftuviitaleaceae_UCG_011*, *Enterobacter*, *Fournierella*, *Aerococcus*, *CAG_873*, *Finogoldia*, *Peptoniphilus* and *Sarcina* (Fig. S35). Octanoic acid wet-specific genera were *Oscillospiraceae_UCG_005*, *[Eubacterium]_eligens_group*, *Clostridia_vadinBB60_group* and *Gastranaerophilales* (Fig. S36). *Methanobrevibacter*, *Megamonas* and *Varibaculum* showed specificity to isobutyric acid dry (Fig. S37) while *Phascolarctobacterium*, *Megamonas* and *Varibaculum* showed specificity to isobutyric acid percent (Fig. S38). Iso-caproic acid percent-specific genera (Fig. S39) were *Family_XIII_UCG_001*, *Deftuviitaleaceae_UCG_011* and *Clostridia_vadinBB60_group* while iso-caproic acid wet-specific genera were *[Clostridium]_innocuum_group* and *Erysipelotrichaceae_UCG_003* (Fig. S40). The genera corresponding specificity to iso-caproic acid dry were *[Clostridium]_innocuum_group*, *Eubacterium*, *[Eubacterium]_nodatum_group*, *Erysipelotrichaceae_UCG_003*, *Deftuviitaleaceae_UCG_011* and *Sellimonas* (Fig. S41). Five genera showed specificity to urine fluoride content [*Ruminococcus]_gnavus_group*, *Comamonas*, *Eubacterium*, *Klebsiella* and *CAG_873* (Fig. S42). *Lachnospiraceae_NK4B4_group* and *Oscillospiraceae_UCG_003* showed specificity to RFM-index (Fig. S43).

3.7. Identification of coherent patterns between Genera/SCFAs that are predictive of high fluoride exposure

The algorithm found two components and showed the reduced order representation of samples in the Fig. S44A, B and C. Circos plots for the analysis (Fig. 7A, B and C) represented all the features in the DIABLO model, and the strong correlations between variables of different data types. Loading plots (Fig. S45-S50) displayed discriminating signatures for genera and SCFAs by the control and high exposed samples. Propionic acid-dry (C3-Dry) was found in a strong negative correlation with *Bacteroides*, that were highly abundant in control samples. The positively associated genera with C3-Dry in fluoride exposed outcome were *Peptoniphilus*, *Sarcina* and *Prevotella* etc. Iso-caproic acid-wet (IC6-Wet) showed potent negative correlations with *Escherichia-Shigella*, *Bacteroides*, *Fuscatenibacter* and *Erysipelotrichaceae_UCG-003* whose abundance in control samples was markedly high. The positively correlated genera with IC6-Wet in fluoride exposed outcome were *Lachnospiraceae_NK4A136_group*, *Lachnospiraceae_ND3007_group*, *[Eubacterium]_xylanophilum_group*, *Enterorhabdus* etc. Butyric acid percentage (C4 %) exhibited negative correlations with *Bacteroides* and *Erysipelotrichaceae_UCG-003* which were abundant in control samples. Genera that showed positive correlations with C4 % in fluoride exposed outcome were *Lachnospiraceae_NK4A136_group*, *Lachnospiraceae_ND3007_group*, *[Eubacterium]_xylanophilum_group*, *Enterorhabdus*,

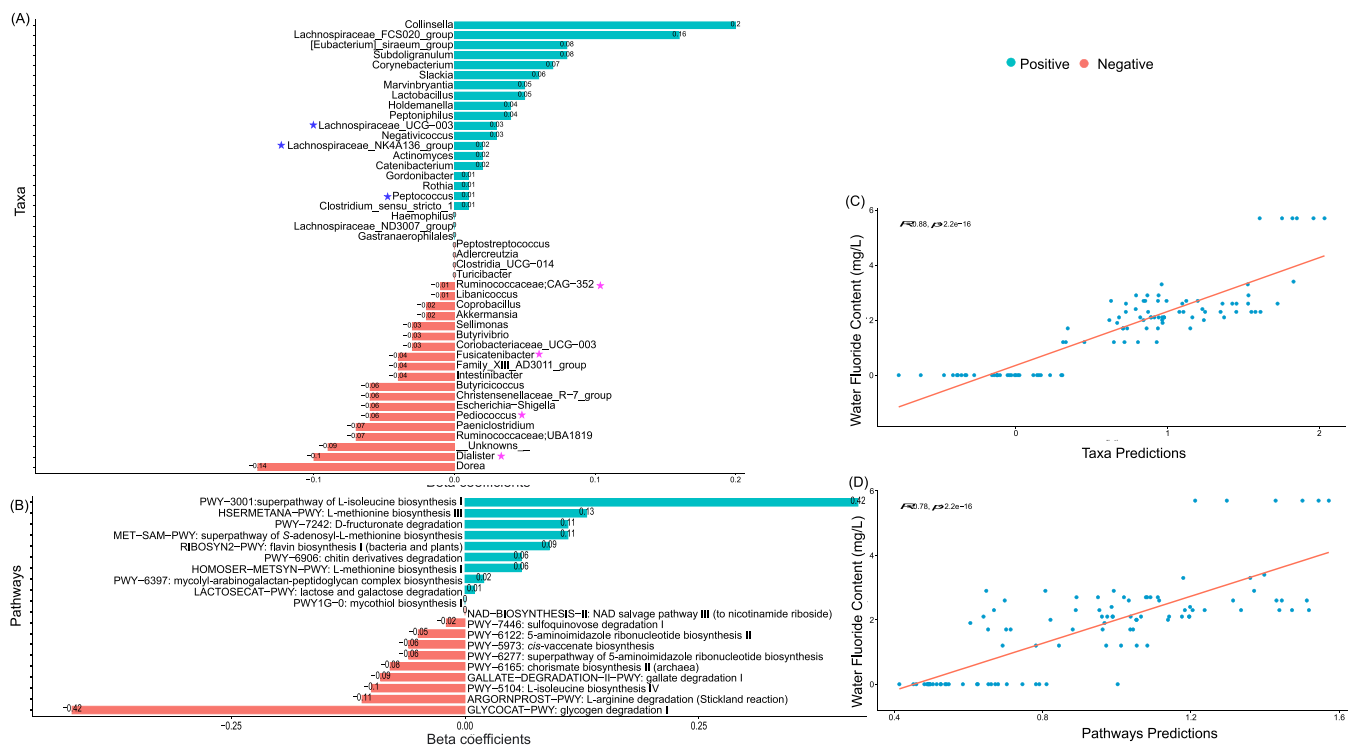


Fig. 5. CODA-LASSO plots for Microbial (A) Taxa signature and (B) Pathways signature identification. The bar graphical representations constitute the signature and their coefficients. The scatter plots (C and D) show the prediction for classification accuracy i.e. relationship/correlation between the water fluoride concentration and microbial taxa/pathways. CODA-LASSO identified the signature with maximum discrimination accuracy between the control and high fluoride exposure groups. The signature is defined by the relative abundances of two groups of taxa/pathways where one group is composed of taxa/pathways with positive coefficients (green) in the regression model and the other group composed of taxa/pathways with negative coefficients (red).

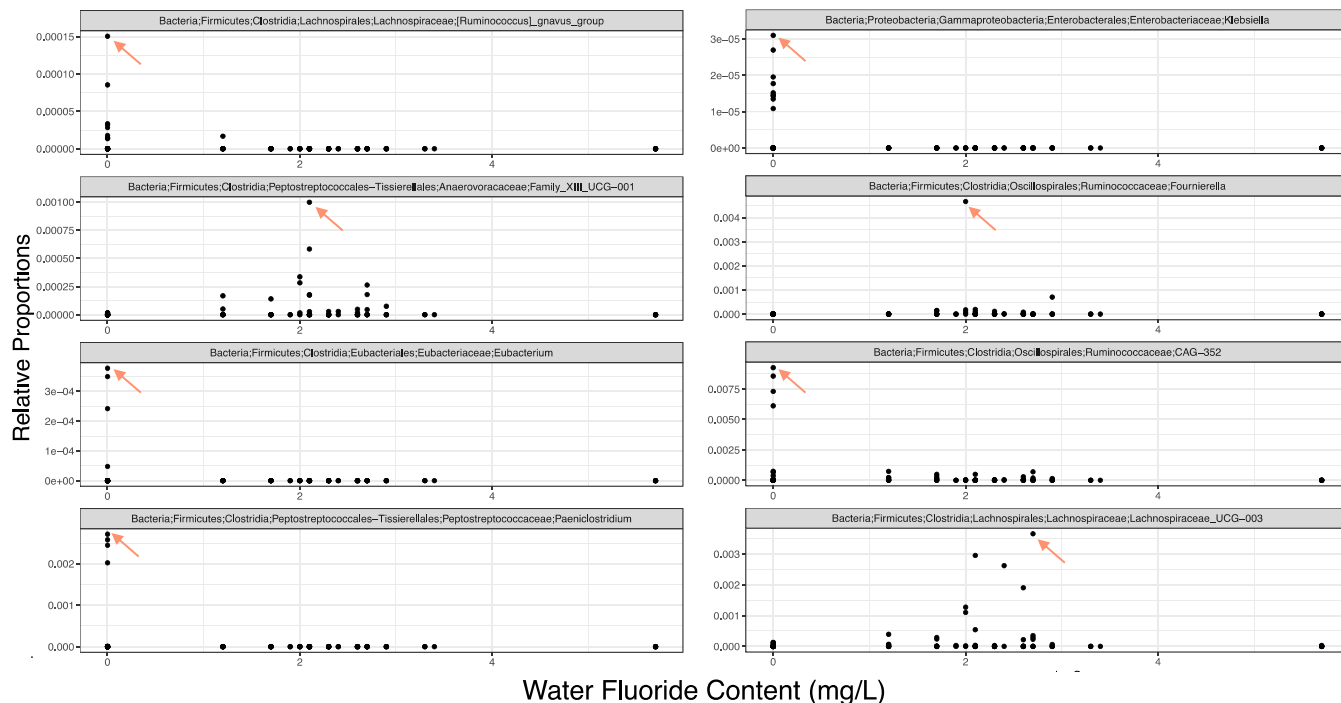


Fig. 6. The genera identified as “Specific” in Fig. S30A (darker) for Water Fluoride Content i.e. the genera in lowest 25th quartile of *Spec* values are shown with their abundances in the whole range of Water Fluoride Content values. This water fluoride concentration-specific analysis ensured the variations observed at genus level in microbiome composition were specific between control and high fluoride exposure groups. In eight out of total genera, there were significant differences in the microbial composition of control and fluoride exposed samples.

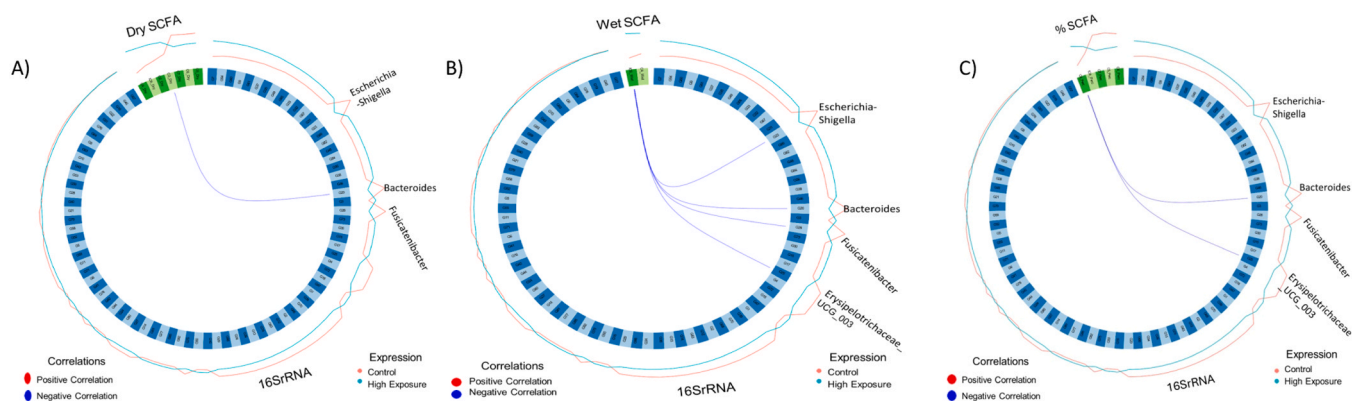


Fig. 7. Significant correlations ($|r| > 0.6$) whether positive (red) or negative (blue) between the two datasets i.e. (A) Genera (16SrRNA) and SCFAs (Dry SCFA); (B) Genera (16SrRNA) and SCFAs (Wet SCFA); (C) Genera (16SrRNA) and SCFAs (% SCFA) returned from the DIABLO algorithm as the Circos plots. Note that these correlations are above a value of 0.6 (cutoff = 0.6). All the interpretations made above are only relevant for features with very strong correlations. Lines along the outside of the circle represent the mean “expression” levels. Greater expression levels are in accordance with the line being farther away from the circle. Expression of Genera/SCFAs in control and high exposed groups are ranked in terms of significantly differing expression.

Olsenella, [*Eubacterium*]*_ruminantium_group etc.*

4. Discussion

The fluoride concentrations in the sampled groundwater of the endemic regions have exceeded the WHO guideline. EPA Pakistan follows the same standard for fluoride in drinking water (i.e. ≤ 1 mg/L for health-related priority). In these regions, groundwater chemistry is predominantly regulated by geochemical processes such as the interaction between rocks and water and the cation exchange. Groundwater with high alkalinity, poor calcium content and sodium bicarbonate water type facilitates the leaching processes. The arid climate of the regions enhance the groundwater residence times resulting increased fluoride contamination of groundwater (Ali et al., 2019; Noor et al., 2022). Groundwater fluoride exposure is known to have long-term effects, including fluorosis of the teeth and skeleton. The effects of fluorosis on the host’s serum metabolites, metabolic pathways for fatty acid oxidation and energy production, protein and purine breakdown and ω -6 fatty acid linoleate signatures have recently been assessed (Usman et al., 2022). It has been anticipated that chronic exposure of fluoride through drinking water causes susceptibility of gut microbiome towards unexpected changes that might have health repercussions. Ours is one of the first studies reporting the link between the biochemical parameters for fluorosis subjects and their associated gut microbiome structure profiles. It appears that the exposed group’s increased oxidative stress is a result of their excessive fluoride intake. As a consequence, the activity of cholinesterases (AChE and BChE) may have increased, which might have also diminished the cholinergic anti-inflammatory pathways, resulting in low-grade inflammation (Bibi et al., 2023). In addition, more deregulated metabolic markers were found in the fluoride-exposed group compared to the control group, including significant insulin resistance, atherogenic index of plasma, an elevated BUN to creatinine ratio, a high alkaline phosphatase level and a lower eGFR. Regarding the urinary fluoride concentrations of the control group, it is to be realized that food and beverages contribute to exposure though high fluoride drinking water is typically responsible for daily fluoride intake. Similarly, a study has linked unhealthy cardiometabolic outcomes in school-aged children to dietary fluoride exposures during early and mid-childhood. Stronger associations were observed in girls prior to year 8 than boys posing sex-prone risk (India Aldana et al., 2024). Another study found that increased proportion of patients in the high blood pressure group had elevated levels of water or plasma fluoride compared to the normal blood pressure group (Hung et al., 2023). Regarding microbial abundance, statistically significant changes were not observed in OTUs richness of the control and fluoride-exposed fecal samples but

there were significant shifts in α -diversity for enzymes and functional pathways. It is assumed that high fluoride exposure has somewhat reduced the functional properties of the gut microbial communities of the exposed group. While fluoride as an environmental factor has the ability to bring about functional changes in microbes (Carda-Diéguez et al., 2022; Dionizio et al., 2021; Marquis et al., 2003), we can assume that with high fluoride exposure, the gut microbiome may exist as an adapted ecosystem with less coordination for exchange of nutrients and other signals. Comparing the β -diversity metrics revealed considerably different microbial profiles across the two groups and further illustrated (PERMANOVA) the influential factors shaping the composition and metabolism of microbial communities with water fluoride concentration, sex and residence being the more relevant along the SCFAs values.

The taxonomic output for the top 25 bacterial genera exposed some of the ubiquitous inhabitants found in healthy people’s guts, with *Bifidobacterium*, *Lactobacillus* and *Enterococcus* being the most prevalent (Derrien et al., 2022; Krawczyk et al., 2021; Rastogi and Singh, 2022). *Bifidobacterium* species are some of the first microorganisms to invade the human gut system and are considered to improve the health of their host (O’Callaghan and van Sinderen, 2016). Humans and *Lactobacilli* have a mutualistic connection in which *Lactobacillus* species provide assistance in the digestion of specific dietary components and protection against infections (Dempsey and Corr, 2022). The *Enterococci* are adaptable species that can endure under challenging conditions, making them ideally suited to the medical setting environments (García-Solache and Rice, 2019). The core microbiome looked for regularly observed taxa in our studied host groups. The persistent genera contributed towards stability of gut microbial community. *Ruminococcaceae-uncultured* (associated with lowering HOMA-IR), *Terrisporobacter* (commensal) and *Escherichia-Shigella* (pathobiont) existed in control samples only (Baltazar-Díaz et al., 2022; Galié et al., 2021; Radwan et al., 2020) while *Catenibacterium* and *Prevotella* (polysaccharide-degrading genera) and *Eggerthellaceae-Slackia* (roles in lipid and xenobiotic metabolism) were found in high fluoride exposed samples solely (Cho et al., 2016; Garcia-Mantrana et al., 2018). Importantly, the genera that were differentially enhanced in control samples were found to be relevant to probiotics, mental health, immune factors meditation, hunger and healthy ageing (Gu et al., 2021; Martínez et al., 2013; Rios-Covian et al., 2017; Singh et al., 2019; Todorov et al., 2023; Yu et al., 2022). On the other hand, differentially abundant genera in the high fluoride exposed samples exhibited propensity to proliferation of opportunistic pathogens, delayed gastric emptying, inflammatory bowel diseases, and certain SCFA producers mainly associated with poor health (Amaruddin et al., 2020; Barrak et al., 2020; Lam-Himlin et al., 2011; López-Montoya et al., 2022; Vacca et al., 2020; Zhang et al., 2021b). The genera that

were most positively associated with fluoride exposure markers (such as water fluoride and urine fluoride concentrations) also correlated positively with those host biochemical features that could turn out to be risk factors when elevated (alkaline phosphatase, E-AChE, BChE, BUN, blood glucose, blood pressure, IL-6, TNF- α , MDA) in relation to high fluoride exposure. These genera were listed in the Table S3 and majority of these genera showed negative association with host haemoglobin, RFM index and eGFR CKD EPI PK noticeably. On the contrary, the genera that were negatively associated with fluoride exposure markers showed positive association with either host haemoglobin (*Dialister*, *Escherichia shigella*), RFM index (*Dialister*, *Fusicatenibacter* and *Ruminococcaceae-CAG-352*) or eGFR CKD EPI PK (*Escherichia Shigella*). The most common microbial metabolic pathways prevailing under fluoride exposure were described in the Table S4. These pathways are mainly concerned with the higher sugar degradation; biosynthesis of cysteine, phospholipids and polyamines; SCFAs production and degradation of xenobiotic compounds; mucosal microbiome and systemic inflammation; ulcerative colitis and non-alcoholic fatty liver disease. *Collinsella*, being the top most positively associated genus with high water fluoride concentration, has been implicated in literature with positive association with atherosclerosis, altering gut permeability, production of proinflammatory cytokine IL-17A, rheumatoid arthritis and nonalcoholic steatohepatitis (Astbury et al., 2020). PWY-3001: superpathway of L-isoleucine biosynthesis I, being the strongly associated pathway with high water fluoride concentration, has been found in positive correlation with LDL-cholesterol and platelet count while in negative correlation with age, hemoglobin A1c, insulin, prothrombin time and direct bilirubin (Oh et al., 2020).

We have further identified some bacterial genera that exhibited strong specificity towards the covariates in this study. Fluoride levels in the water and urine, the RFM index and several SCFAs (propionic acid, butyric acid, valeric acid, octanoic acid, iso-butyric acid, iso-caproic acid) have been shown to alter the microbial populations in the gut. In another study, *Family XIII UCG_001* and *Fournierella* had shown positive correlation with HbA1c in children with T1DM (Tamahane et al., 2023). Numerous studies link *Ruminococcus gnavus* to Crohn's disease and CAD (Henke et al., 2019; Toya et al., 2020). *Comamonas spp.*, commonly considered as having low virulence, have led to health issues in many healthy people (Ryan et al., 2022). In the gut, *Eubacterium spp.* transform cholesterol and bile acids (Mukherjee et al., 2020). Certain *Klebsiella species* can act as opportunistic human infections (Ristuccia and Cunha, 1984). *Lachnospiraceae_NK4B4_group* are common butyrate producers (Bui et al., 2023) while *Oscillospiraceae_UCG_003* was negatively correlated with the triglycerides (Portela et al., 2023) and was one of the dominant genera prevailing in relation to vegetarian diet (Šik Novak et al., 2023). The altered production of SCFAs in fluoride-exposed samples indicates a potential disruption in the normal metabolic activities of gut bacteria. The correlation with specific bacterial genera suggests that certain types of bacteria may be more sensitive or responsive to fluoride exposure. The perturbations in bacterial genera associated with changes in SCFA production may have implications for the general health of the host organism.

Regarding the elevated levels of fecal SCFAs, there is an evidence of association of increased SCFAs with gut dysbiosis and permeability, hypertension and cardiometabolic risk factors (de la Cuesta-Zuluaga et al., 2018). Previous studies on human illnesses have shown that the gut bacteria produced significantly large levels of SCFAs with concentrations reaching as high as 80 \pm 11 mmol/kg in the descending colon and 13 \pm 6 mmol/kg in the terminal ileum (Cong et al., 2022; Cummings et al., 1987). Furthermore, existing literature also corroborates that individuals with higher butyrate and lower propionate and acetate production, are the people that may have low abundance and diversity of SCFA-producing bacteria (Murphy et al., 2010; You et al., 2022). At the same time, it is important to consider that exposure to butyrate has epigenetic effects that are strongly correlated with glucose metabolism (Donohoe et al., 2012). Butyrate may block histone deacetylases by attaching to the Zn²⁺ in the catalytic site (Riggs et al., 1977) while the

multiplication of stem/progenitor cells in the intestinal crypt was likewise demonstrated to be inhibited by abundantly produced butyrate (Kaiko et al., 2016). Additionally, branched-chain fatty acids (BCFA), being markers of colonic protein fermentation, have also been linked to the control of lipid and glucose metabolism (Heimann et al., 2016). Among the SCFA-specific bacterial genera, many are reported in the literature as SCFA-producing genera like *Muribaculaceae* (Tian et al., 2021), *Clostridium spp.* (Guo et al., 2020), *Clostridia vadinBB60_group* (Wassie et al., 2022), *Phascolarctobacterium* (Wu et al., 2017), *Lachnospiraceae* (Parada Venegas et al., 2019), *Oscillibacter* and *Erysipelotrichaceae spp.* (Martin-Gallausiaux et al., 2021).

Although the results look promising, our study has some limitations. Our cases and controls were heterogeneous with respect to age and body weight. Whilst the study investigates the connection between fluoride exposure, the biochemical markers and the gut microbiota, it is difficult to fully evaluate the gut-organ axis under fluoride exposure. Identification of fecal metabolites was missing that could serve as possible diagnostic biomarkers to reveal the relationships between fecal metabolites and fluoride toxicity. Furthermore, due to the limitations in the reference database for 16S rRNA gene amplicon sequencing, OTUs could only be resolved up to the genus level due to high similarity of the gene from the closely related species.

5. Conclusion

In conclusion, our study was a preliminary step in prospecting the host-microbe interaction in the development of fluorosis. The microbiota dynamics investigated in relation to the control and high fluoride exposure conditions, had significant discriminatory shifts in the gut bacterial community structure. The healthy gut microbiota predicted to maintain vital body processes like digestion, immune system regulation, metabolism and pathogen resistance. In high fluoride exposure, the commensal and opportunistic bacteria could modify the severity of the health issues. Our study identified microbial and metabolic pathways' signatures that were positively associated with the host's variables. The changes in the gut microbiota can, in turn, influence metabolic activities within the body. Such metabolic disturbances may contribute to the development of various diseases. These valuable insights of the relationship between microbiota and host health can shed light on the underlying mechanisms of fluoride toxicity. This suggests a potential mechanism by which fluoride may exert its harmful effects beyond dental fluorosis. Such studies may help design strategies for mitigating the health risks associated with chronic fluoride exposure and developing targeted interventions to restore gut microbial balance and metabolic homeostasis. It would be useful to explore these findings for a larger cohort, and to look at the impact of fluoride on other human associated microbiomes. Also, the exposed subjects in this study were drinking water that caused dental fluorosis and symptomatic complains about excessive fluoride ingestion but skeletal fluorosis was largely unseen. A future study looking at severity of fluorosis beyond what we have considered will be potentially useful.

Abbreviations

SCFAs: Short chain fatty acids
DF: Dental fluorosis
BMI: Body mass index
RFM: Relative fat mass
HDL: High density lipoprotein
eGFR: Estimated glomerular filtration rate
CKD EPI PK: Chronic Kidney Disease Epidemiology for Pakistan
MDA: Malondialdehyde
E-AChE: Erythrocytes acetylcholinesterase
BChE: Butyrylcholinesterase
BUN: Blood urea nitrogen
ALP: Alkaline phosphatase

γ GT: Gamma glutamyltransferase
 HOMA-IR: Homeostasis Model Assessment for Insulin Resistance
 AI-P: Atherogenic Index Plasma
 ORP: Oxidation reduction potential
 PWY: Pathway
 WHO: World Health Organization
 EPA: Environmental Protection Agency

Ethics approval and consent to participate

This study was approved by the Ethics Review Board (ERB) at COMSATS University Islamabad (ERB No. CUI/Bio/ERB/2021/51). The purpose of the study was explained to each subject and written informed consent was obtained.

Funding

SB acknowledges support from International Research Support Initiative Program from Higher Education Commission, Pakistan Project No. 1–8/HEC/HRD/2021/11509. UZI is funded by EPSRC (EP/V030515/1).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The raw sequence files supporting the results of this article are available in the European Nucleotide Archive under the project accession number PRJEB60561

Acknowledgments

We acknowledge all the participants in the study. For open access, the authors have applied for a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.ecoenv.2024.116959](https://doi.org/10.1016/j.ecoenv.2024.116959).

References

- Ahmad, S., Singh, R., Arfin, T., Neeti, K., 2022. Fluoride contamination, consequences and removal techniques in water: a review. *Environ. Sci.: Adv.* 1, 620–661. <https://doi.org/10.1039/D1VA00039J>.
- Ahmed, S., Jafri, L., Khan, A.H., 2017. Evaluation of “CKD-EPI Pakistan” Equation for estimated Glomerular Filtration Rate (eGFR): A Comparison of eGFR Prediction Equations in Pakistani Population. *J. Coll. Physicians Surg. Pak.* 27, 414–418.
- Aldars-García, L., Chaparro, M., Gisbert, J.P., 2021. Systematic Review: The Gut Microbiome and Its Potential Clinical Application in Inflammatory Bowel Disease. *Microorganisms* 9, 977. <https://doi.org/10.3390/microorganisms9050977>.
- Ali, W., Aslam, M.W., Junaid, M., Ali, K., Guo, Y., Rasool, A., Zhang, H., 2019. Elucidating various geochemical mechanisms drive fluoride contamination in unconfined aquifers along the major rivers in Sindh and Punjab, Pakistan. *Environ. Pollut.* 249, 535–549. <https://doi.org/10.1016/j.envpol.2019.03.043>.
- Amaruddin, A.I., Hamid, F., Koopman, J.P.R., Muhammad, M., Brienen, E.A.T., van Lieshout, L., Geelen, A.R., Wahyuni, S., Kuyper, E.J., Sartono, E., Yazdanbakhsh, M., Zwiittink, R.D., 2020. The Bacterial Gut Microbiota of Schoolchildren from High and Low Socioeconomic Status: A Study in an Urban Area of Makassar, Indonesia. *Microorganisms* 8, 961. <https://doi.org/10.3390/microorganisms8060961>.
- Astbury, S., Atallah, E., Vijay, A., Aithal, G.P., Grove, J.I., Valdes, A.M., 2020. Lower gut microbiome diversity and higher abundance of proinflammatory genus *Collinsella* are associated with biopsy-proven nonalcoholic steatohepatitis. *Gut Microbes* 11, 569–580. <https://doi.org/10.1080/19490976.2019.1681861>.

- Baltazar-Díaz, T.A., González-Hernández, L.A., Aldana-Ledesma, J.M., Peña-Rodríguez, M., Vega-Magaña, A.N., Zepeda-Morales, A.S.M., López-Roa, R.I., del Toro-Arreola, S., Martínez-López, E., Salazar-Montes, A.M., Bueno-Topete, M.R., 2022. Escherichia/Shigella, SCFAs, and Metabolic Pathways—The Triad That Orchestrates Intestinal Dysbiosis in Patients with Decompensated Alcoholic Cirrhosis from Western Mexico. *Microorganisms* 10, 1231. <https://doi.org/10.3390/microorganisms10061231>.
- Barbier, O., Arreola-Mendoza, L., Del Razo, L.M., 2010. Molecular mechanisms of fluoride toxicity. *Chem.-Biol. Interact.* 188, 319–333. <https://doi.org/10.1016/j.cbi.2010.07.011>.
- Barrak, I., Stájer, A., Gajdacs, M., Urbán, E., 2020. Small, but smelly: the importance of *Solobacterium moorei* in halitosis and other human infections. *Heliyon* 6. <https://doi.org/10.1016/j.heliyon.2020.e05371>.
- Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C.C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G.H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J.A., Maguin, E., Mauchline, T., McClure, R., Mitter, B., Ryan, M., Sarand, I., Smidt, H., Schelkle, B., Roume, H., Kiran, G.S., Selvin, J., Souza, R.S.C. de, van Overbeek, L., Singh, B.K., Wagner, M., Walsh, A., Sessitsch, A., Schloter, M., 2020. Microbiome definition re-visited: old concepts and new challenges. *Microbiome* 8, 103. <https://doi.org/10.1186/s40168-020-00875-0>.
- Bibi, S., Habib, R., Shafiq, S., Abbas, S.S., Khan, S., Eqani, S.A.S., Nepovimova, E., Khan, M.S., Kuca, K., Nurulain, S.M., 2023. Influence of the chronic groundwater fluoride consumption on cholinergic enzymes, AChE and BCHE gene SNPs and pro-inflammatory cytokines: a study with Pakistani population groups. *Sci. Total Environ.* 880, 163359. <https://doi.org/10.1016/j.scitotenv.2023.163359>.
- Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K.B., Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciolk, T., Kreps, J., Langille, M.G.L., Lee, J., Ley, R., Liu, Y.-X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A.V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K.C., Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y., Zhu, Q., Knight, R., Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. <https://doi.org/10.1038/s41587-019-0209-9>.
- Boutin, J.A., Kass, G.E.N., Moldéus, P., 1989. Drug-induced hydrogen peroxide production in isolated rat hepatocytes. *Toxicology* 54, 129–137. [https://doi.org/10.1016/0300-483X\(89\)90039-5](https://doi.org/10.1016/0300-483X(89)90039-5).
- Bui, T.V.A., Hwangbo, H., Lai, Y., Hong, S.B., Choi, Y.-J., Park, H.-J., Ban, K., 2023. The Gut-Heart Axis: Updated Review for the Roles of Microbiome in Cardiovascular Health. *Korean Circ. J.* 53, 499–518. <https://doi.org/10.4070/kcj.2023.0048>.
- Buzalaf, M. a R., 2018. Review of Fluoride Intake and Appropriateness of Current Guidelines. *Adv. Dent. Res* 29, 157–166. <https://doi.org/10.1177/0022034517750850>.
- Calle, M.L., Pujolassos, M., Susin, A., 2023. coda4microbiome: compositional data analysis for microbiome cross-sectional and longitudinal studies. *BMC Bioinforma.* 24, 82. <https://doi.org/10.1186/s12859-023-05205-3>.
- Carda-Diéguez, M., Moazzez, R., Mira, A., 2022. Functional changes in the oral microbiome after use of fluoride and arginine containing dentifrices: a metagenomic and metatranscriptomic study. *Microbiome* 10, 159. <https://doi.org/10.1186/s40168-022-01338-4>.
- Chen, G., Hu, P., Xu, Z., Peng, C., Wang, Y., Wan, X., Cai, H., 2021. The beneficial or detrimental fluoride to gut microbiota depends on its dosages. *Ecotoxicol. Environ. Saf.* 209, 111732. <https://doi.org/10.1016/j.ecoenv.2020.111732>.
- Cho, G.-S., Ritzmann, F., Eckstein, M., Huch, M., Briviba, K., Behnsilian, D., Neve, H., Franz, C.M.A.P., 2016. Quantification of *Slackia* and *Eggerthella* spp. in Human Feces and Adhesion of Representative Strains to Caco-2 Cells. *Front. Microbiol.* 7, 1499. <https://doi.org/10.3389/fmicb.2016.01499>.
- Chowdhury, A., Adak, M.K., Mukherjee, A., Dhak, P., Khatun, J., Dhak, D., 2019. A critical review on geochemical and geological aspects of fluoride belts, fluorosis and natural materials and other sources for alternatives to fluoride exposure. *J. Hydrol.* 574, 333–359. <https://doi.org/10.1016/j.jhydrol.2019.04.033>.
- Cong, J., Zhou, P., Zhang, R., 2022. Intestinal Microbiota-Derived Short Chain Fatty Acids in Host Health and Disease. *Nutrients* 14, 1977. <https://doi.org/10.3390/nu14091977>.
- Costello, E.K., Lauber, C.L., Hamady, M., Fierer, N., Gordon, J.I., Knight, R., 2009. Bacterial Community Variation in Human Body Habitats Across Space and Time. *Science* 326, 1694–1697. <https://doi.org/10.1126/science.1177486>.
- Cummings, J.H., Pomare, E.W., Branch, W.J., Naylor, C.P., Macfarlane, G.T., 1987. Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* 28, 1221–1227. <https://doi.org/10.1136/gut.28.10.1221>.
- Darcy, J.L., Amend, A.S., Swift, S.O.I., Sommers, P.S., Lozupone, C.A., 2022. specificity: an R package for analysis of feature specificity to environmental and higher dimensional variables, applied to microbiome species data. *Environ. Micro* 17, 34. <https://doi.org/10.1186/s40793-022-00426-0>.

- Melo, C.G. de S., Perles, J.V.C.M., Zanoni, J.N., Souza, S.R.G. de, Santos, E.X., Leite, A. de L., Heubel, A.D., e Souza, C.O., Souza, J.G. de, Buzalaf, M.A.R., 2017. Enteric innervation combined with proteomics for the evaluation of the effects of chronic fluoride exposure on the duodenum of rats. *Sci. Rep.* 7, 1070. <https://doi.org/10.1038/s41598-017-01090-y>.
- Miao, L., Zhu, M., Li, H., Xu, Q., Dong, X., Zou, X., 2020. Dietary High Sodium Fluoride Impairs Digestion and Absorption Ability, Mucosal Immunity, and Alters Cecum Microbial Community of Laying Hens. *Animals* 10, 179. <https://doi.org/10.3390/ani10020179>.
- Misra, H.P., Fridovich, I., 1977. Superoxide dismutase: "Positive" spectrophotometric assays. *Anal. Biochem.* 79, 553–560. [https://doi.org/10.1016/0003-2697\(77\)90429-8](https://doi.org/10.1016/0003-2697(77)90429-8).
- Mukherjee, A., Lordan, C., Ross, R.P., Cotter, P.D., 2020. Gut microbes from the phylogenetically diverse genus *Eubacterium* and their various contributions to gut health. *Gut Microbes* 12, 1802866. <https://doi.org/10.1080/19490976.2020.1802866>.
- Murphy, E.F., Cotter, P.D., Healy, S., Marques, T.M., O'Sullivan, O., Fohuy, F., Clarke, S. F., O'Toole, P.W., Quigley, E.M., Stanton, C., Ross, P.R., O'Doherty, R.M., Shanahan, F., 2010. Composition and energy harvesting capacity of the gut microbiota: relationship to diet, obesity and time in mouse models. *Gut* 59, 1635–1642. <https://doi.org/10.1136/gut.2010.215665>.
- Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., Pettersson, S., 2012. Host-gut microbiota metabolic interactions. *Science* 336, 1262–1267. <https://doi.org/10.1126/science.1223813>.
- Noor, S., Rashid, A., Javed, A., Khattak, J.A., Farooqi, A., 2022. Hydrogeological properties, sources provenance, and health risk exposure of fluoride in the groundwater of Batkhela, Pakistan. *Environ. Technol. Innov.* 25, 102239. <https://doi.org/10.1016/j.eti.2021.102239>.
- O'Callaghan, A., van Sinderen, D., 2016. Bifidobacteria and Their Role as Members of the Human Gut Microbiota. *Front Microbiol* 7, 925. <https://doi.org/10.3389/fmicb.2016.00925>.
- Oh, T.G., Kim, S.M., Caussy, C., Fu, T., Guo, J., Bassirian, S., Singh, S., Madamba, E.V., Bettencourt, R., Richards, L., Yu, R.T., Atkins, A.R., Huan, T., Brenner, D.A., Sirlin, C. B., Downes, M., Evans, R.M., Loomba, R., 2020. A Universal Gut-Microbiome-Derived Signature Predicts Cirrhosis. *Cell Metab.* 32, 901. <https://doi.org/10.1016/j.cmet.2020.10.015>.
- O'Reilly, C., Mills, S., Rea, M.C., Lavelle, A., Ghosh, S., Hill, C., Ross, R.P., 2023. Interplay between inflammatory bowel disease therapeutics and the gut microbiome reveals opportunities for novel treatment approaches. *Micro Res Rep.* 2, 35. <https://doi.org/10.20517/mrr.2023.41>.
- Parada Venegas, D., De la Fuente, M.K., Landskron, G., González, M.J., Quera, R., Dijkstra, G., Harmsen, H.J.M., Faber, K.N., Hermoso, M.A., 2019. Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. *Front Immunol.* 10, 277. <https://doi.org/10.3389/fimmu.2019.00277>.
- Podgorski, J., Berg, M., 2022. Global analysis and prediction of fluoride in groundwater. *Nat. Commun.* 13, 4232. <https://doi.org/10.1038/s41467-022-31940-x>.
- Portela, N.D., Galván, C., Sanmarco, L.M., Bergero, G., Aoki, M.P., Cano, R.C., Pesoa, S. A., 2023. Omega-3-Supplemented Fat Diet Drives Immune Metabolic Response in Visceral Adipose Tissue by Modulating Gut Microbiota in a Mouse Model of Obesity. *Nutrients* 15, 1404. <https://doi.org/10.3390/nu15061404>.
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D.R., Li, J., Xu, J., Li, Shaochuan, Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.-M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H.B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, Shengting, Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, Songgang, Qin, N., Yang, H., Wang, Jian, Brunak, S., Doré, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Bork, P., Ehrlich, S.D., Wang, 2010. A human gut microbial gene catalogue established by metagenomic sequencing (Jun). *Nature* 464, 59–65. <https://doi.org/10.1038/nature08821>.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41, D590–596. <https://doi.org/10.1093/nar/gks1219>.
- Radwan, S., Gilfillan, D., Eklund, B., Radwan, H.M., Menofy, N.G.E., Lee, J., Kapuscinski, M., Abdo, Z., 2020. A comparative study of the gut microbiome in Egyptian patients with Type I and Type II diabetes. *PLOS ONE* 15, e0238764. <https://doi.org/10.1371/journal.pone.0238764>.
- Rafique, T., Naseem, S., Usmani, T.H., Bashir, E., Khan, F.A., Bhangar, M.I., 2009. Geochemical factors controlling the occurrence of high fluoride groundwater in the Nagar Parkar area, Sindh, Pakistan. *J. Hazard Mater.* 171, 424–430. <https://doi.org/10.1016/j.jhazmat.2009.06.018>.
- Rastogi, S., Singh, A., 2022. Gut microbiome and human health: Exploring how the probiotic genus *Lactobacillus* modulate immune responses. *Front. Pharmacol.* 13.
- Riggs, M.G., Whittaker, R.G., Neumann, J.R., Ingram, V.M., 1977. n-Butyrate causes histone modification in HeLa and Friend erythroleukaemia cells. *Nature* 268, 462–464. <https://doi.org/10.1038/268462a0>.
- Rios-Covian, D., Salazar, N., Gueimonde, M., de los Reyes-Gavilan, C.G., 2017. Shaping the Metabolism of Intestinal Bacteroides Population through Diet to Improve Human Health. *Front. Microbiol.* 8.
- Ristuccia, P.A., Cunha, B.A., 1984. *Klebsiella*. *Infect. Control* 5, 343–348.
- Rohart, F., Gautier, B., Singh, A., Cao, K.-A.L., 2017. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLOS Comput. Biol.* 13, e1005752. <https://doi.org/10.1371/journal.pcbi.1005752>.
- Ryan, M.P., Sevjahova, L., Gorman, R., White, S., 2022. The Emergence of the Genus *Comamonas* as Important Opportunistic Pathogens. *Pathogens* 11, 1032. <https://doi.org/10.3390/pathogens11091032>.
- Shaji, E., Sarath, K.V., Santosh, M., Krishnaprasad, P.K., Arya, B.K., Babu, M.S., 2024. Fluoride contamination in groundwater: A global review of the status, processes, challenges, and remedial measures. *Geosci. Front.* 15, 101734. <https://doi.org/10.1016/j.gsf.2023.101734>.
- Šik Novak, K., Bogataj Jontez, N., Petelin, A., Hladnik, M., Baruca Arbeiter, A., Bandelj, D., Pražnikar, J., Kenig, S., Mohorko, N., Jenko Pražnikar, Z., 2023. Could Gut Microbiota Composition Be a Useful Indicator of a Long-Term Dietary Pattern? *Nutrients* 15, 2196. <https://doi.org/10.3390/nu15092196>.
- Singh, H., Torralba, M.G., Moncera, K.J., DiLello, L., Petrini, J., Nelson, K.E., Pieper, R., 2019. Gastro-intestinal and oral microbiome signatures associated with healthy aging. *Geroscience* 41, 907–921. <https://doi.org/10.1007/s11357-019-00098-8>.
- Tamahane, V., Bhanushali, S., Shah, N., Gupta, A., Khadilkar, V., Gondhalekar, K., Khadilkar, A., Shouche, Y., 2023. A comparative study of the gut microbiome in Indian children with type 1 diabetes and healthy controls. *J. Diabetes.* <https://doi.org/10.1111/1753-0407.13438>.
- Tian, B., Zhao, J., Zhang, M., Chen, Z., Ma, Q., Liu, H., Nie, C., Zhang, Z., An, W., Li, J., 2021. Lycium ruthenicum Anthocyanins Attenuate High-Fat Diet-Induced Colonic Barrier Dysfunction and Inflammation in Mice by Modulating the Gut Microbiota. *Mol. Nutr. Food Res* 65, e2000745. <https://doi.org/10.1002/mnfr.202000745>.
- Todorov, S.D., Dioso, C.M., Liang, M.-T., Nero, L.A., Khosravi-Darani, K., Ivanova, I.V., 2023. Beneficial features of *pediococcus* from starter cultures and inhibitory activities to probiotic benefits. *World J. Microbiol. Biotechnol.* 39, 4. <https://doi.org/10.1007/s11274-022-03419-w>.
- Toya, T., Corban, M.T., Marrietta, E., Horwath, I.E., Lerman, L.O., Murray, J.A., Lerman, A., 2020. Coronary artery disease is associated with an altered gut microbiome composition. *PLoS ONE* 15, e0227147. <https://doi.org/10.1371/journal.pone.0227147>.
- Usman, M., Ali, A., Jabbar Siddiqui, A., Iftikhar, F., Kumari, S., Sibte-e-Hassan, S., Shad, R., Rafique, T., Kashif Raza, S., El-Seedi, H.R., Uddin, J., Musharraf, S.G., 2022. Evaluation of the chronic intoxication of fluoride on human serum metabolome using untargeted metabolomics. *Arab. J. Chem.* 15, 103928. <https://doi.org/10.1016/j.arabj.2022.103928>.
- Vacca, M., Celano, G., Calabrese, F.M., Portincasa, P., Gobetti, M., Angelis, M.D., 2020. The Controversial Role of Human Gut Lachnospiraceae. *Microorganisms* 8. <https://doi.org/10.3390/microorganisms8040573>.
- Veneri, F., Vinceti, M., Generali, L., Giannone, M.E., Mazzoleni, E., Birnbaum, L.S., Consolo, U., Filippini, T., 2023. Fluoride exposure and cognitive neurodevelopment: Systematic review and dose-response meta-analysis. *Environ. Res* 221, 115239. <https://doi.org/10.1016/j.envres.2023.115239>.
- Visconti, A., Le Roy, C.I., Rosa, F., Rossi, N., Martin, T.C., Mohney, R.P., Li, W., de Rinaldis, E., Bell, J.T., Venter, J.C., Nelson, K.E., Specter, T.D., Falchi, M., 2019. Interplay between the human gut microbiome and host metabolism. *Nat. Commun.* 10, 4505. <https://doi.org/10.1038/s41467-019-12476-z>.
- Wang, Q., Chen, X., Hu, H., Wei, X., Wang, X., Peng, Z., Ma, R., Zhao, Q., Zhao, J., Liu, J., Deng, F., 2021. Structural changes in the oral microbiome of the adolescent patients with moderate or severe dental fluorosis. *Sci. Rep.* 11, 2897. <https://doi.org/10.1038/s41598-021-82709-z>.
- Wassie, T., Cheng, B., Zhou, T., Gao, L., Lu, Z., Xie, C., Wu, X., 2022. Microbiome-metabolome analysis reveals alterations in the composition and metabolism of caecal microbiota and metabolites with dietary *Enteromorpha polysaccharide* and *Yeast glycoprotein* in chickens. *Front. Immunol.* 13.
- Wilbur, K.M., Bernheim, F., Shapiro, O.W., 1949. The thiobarbituric acid reagent as a test for the oxidation of unsaturated fatty acids by various agents. *Arch. Biochem.* 24, 305–313.
- Worek, F., Mast, U., Kiderlen, D., Diepold, C., Eyer, P., 1999. Improved determination of acetylcholinesterase activity in human whole blood. *Clin. Chim. Acta* 288, 73–90. [https://doi.org/10.1016/S0009-8981\(99\)00144-8](https://doi.org/10.1016/S0009-8981(99)00144-8).
- Wu, F., Guo, X., Zhang, J., Zhang, M., Ou, Z., Peng, Y., 2017. *Phascolarctobacterium faecium* abundant colonization in human gastrointestinal tract. *Exp. Ther. Med* 14, 3122–3126. <https://doi.org/10.3892/etm.2017.4878>.
- You, H., Tan, Y., Yu, D., Qiu, S., Bai, Y., He, J., Cao, H., Che, Q., Guo, J., Su, Z., 2022. The Therapeutic Effect of SCFA-Mediated Regulation of the Intestinal Environment on Obesity. *Front Nutr.* 9, 886902. <https://doi.org/10.3389/fnut.2022.886902>.
- Yu, D., Du, J., Pu, X., Zheng, L., Chen, Shuaishuai, Wang, N., Li, J., Chen, Shiyong, Pan, S., Shen, B., 2022. The Gut Microbiome and Metabolites Are Altered and Interrelated in Patients With Rheumatoid Arthritis. *Front Cell Infect. Microbiol* 11, 763507. <https://doi.org/10.3389/fcimb.2021.763507>.
- Zhang, Y., et al., 2021. The diversity of gut microbiota in type 2 diabetes with or without cognitive impairment. *Aging Clin. Exp. Res.* 33. <https://doi.org/10.1007/s40520-020-01553-9>.
- Zhang, Y., Jing, G., Chen, Y., Li, J., Su, X., 2021. Hierarchical Meta-Storms enables comprehensive and rapid comparison of microbiome functional profiles on a large scale using hierarchical dissimilarity metrics and parallel computing. *Bioinform Adv.* 1, vbab003. <https://doi.org/10.1093/bioadv/vbab003>.
- Zhou, G., Li, Q., Hou, X., Wu, H., Fu, X., Wang, G., Ma, J., Cheng, X., Yang, Y., Chen, R., Li, Z., Yu, F., Zhu, J., Ba, Y., 2023. Integrated 16S rDNA sequencing and metabolomics to explore the intestinal changes in children and rats with dental fluorosis. *Ecotoxicol. Environ. Saf.* 251, 114518. <https://doi.org/10.1016/j.ecoenv.2023.114518>.