Early life gut microbial and metabolic shifts reflected in stool consistency- a gut transit time proxy

Anna-Katariina Aatsinki[1,2], Matilda Kråkström[3], Katja Salonen[3], Heidi Isokääntä[1,2,4], Abhijit Paul[3], Thomaz Bastiaanssen[5], Leo Lahti[6], Minna Lukkarinen[1,2,7], Eveliina Munukka[2,8], Tiina Paunio[9], Katri Kantojärvi[9], Pieter C. Dorrestein [10], Hasse Karlsson[1,2,11], Linnea Karlsson[1,2,12,13], Matej Oresic[3,14], Alex M Dickens[3,15], Santosh Lamichhane [3,4]

[1] Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland

[2] FinnBrain Birth Cohort Study, Turku Brain and Mind Center, Department of Clinical Medicine, University of Turku, Turku, Finland.

[3] Turku Bioscience Centre, University of Turku and Åbo Akademi University, 20520 Turku, Finland

[4] Research Center for Infections and Immunity, Institute of Biomedicine, University of Turku,Turku, Finland.

[5] Amsterdam University Medical Center, Amsterdam, Netherlands

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

[7]Department of Pediatrics and Adolescent Medicine, Turku University Hospital and University of Turku, Turku, Finland

[8]Turku Clinical Microbiome Bank, Clinical Microbiology, Turku University Hospital, The Wellbeing services county of Southwest Finland, Turku, Finland

[9] Department of Psychiatry and SleepWell Research Program, Faculty of Medicine, University of Helsinki, Helsinki University Central Hospital, Helsinki, Finland.

[10]Department of Public Health, University of Turku, Turku, Finland

[10] Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, USA

[11]Department of Psychiatry, University of Turku and Turku University Hospital, Turku, Finland

[12] Department of Public Health, University of Turku, Turku, Finland

[13]Department of Child Psychiatry, Turku University Hospital Turku, Finland

[14]School of Medical Sciences, Örebro University, 702 81 Örebro, Sweden

[15]Department of Chemistry, University of Turku, 20520 Turku, Finland

Correspondence: alex.dickens@utu.fi, ankaaa@utu.fi, santosh.lamichhane@utu.fi

ABSTRACT

Transit time, fluid intake, diet, and overall gut health all influence stool consistency. However, the relationships between the gut microbiota, metabolome (microbial metabolites), and stool consistency in infants and young children remain poorly understood. In this study, we analyzed the microbiota and metabolome of 618 stool samples from children aged 2.5 months (n = 360), 6 months (n = 229), 14 months (n = 274), and 30 months (n = 169) from the FinnBrain Birth Cohort Study, correlating these data with stool water content and parent-reported stool consistency. We found that breastfeeding had the strongest association with both stool consistency and fecal water content. We observed that the concentrations of newly discovered microbial bile acid amidates were associated with constipation and negatively correlated with stool water content. Partial versus exclusive breastfeeding was further linked to higher levels of specific bile acid amidates through mediation by constipation. Bile salt hydrolase, an enzyme involved in deconjugating primary bile acids and generating novel microbial bile acids, was also predicted to be negatively associated with stool water content. In addition, short-chain fatty acids (SCFAs), particularly acetic acids, were positively associated with stool water content, whereas branched SCFAs showed a negative association. These findings suggest longer transit time appeared to allow more extensive bacterial metabolism of microbial metabolites, including the transformation of bile acid amidates. Here, we also observed that the relative abundances of specific genera such as *Clostridium*, *Hungatella*, and *Lactobacillus* were associated with firmer stool consistency. Overall, our results highlight the importance of documenting stool consistency in fecal metabolomics and microbiome research, and provide new insights into how breastfeeding, microbial metabolism, and gut transit time shape early-life gut development.

Introduction

The concept of assessing stool consistency and its link with gastrointestinal pathophysiology dates back over a century [1]. These days, stool consistency is generally assessed using the Bristol Stool Scale. Recent reports suggest that stool consistency is strongly associated with gut microbiota richness, composition, and microbial metabolism [2-5] . Additionally, pathological conditions related to alterations in stool consistency or defecation frequency, such as chronic constipation, are associated with gut microbiota composition in children [6].

Emerging literature, including our own work, suggests that the gut microbiota in early life plays a pivotal role in host health and the development of various diseases across the lifespan [7, 8] . There is a growing awareness that the impact of microbes on various physiological processes including initiation and pathogenesis of various diseases may be mediated by microbially regulated metabolites [9-12]. Current microbiota research related to human health and disease has shifted its focus from characterizing community composition towards understanding the intricate interplay between the gut microbiota and microbiota-dependent metabolites. Efforts have been made to enhance the identification of microbiota composition [2, 13-16]. Accounting for these factors will enable us to more robustly identify disease markers in human microbiota studies [17]. Particularly high intra-individual variance in gut microbiota community composition is explained by proxies for gut transit time [18] . However, measuring the gut transit time is often inconvenient due to logistic challenges and its invasive nature [16, 19]. Stool consistency measurement has been identified as a potentially useful proxy for gut transit time within specific cohorts [20]; although other explanations are possible - in general hard stools suggest slow transit while loose stools are linked with fast transit through the colon. Nevertheless, stool consistency and water content measurement are often neglected in human microbiota studies.

To the best of our knowledge, there are no studies in children investigating stool metabolites in relation to stool consistency. Several adult human studies from both healthy and pathological conditions have identified associations with gut microbiota composition, stool consistency, or transit time measurement [17]. Although

understanding in adult populations is increasing, there is a dearth of studies examining the impact of the gut microbiota and metabolome on both stool consistency and total water content of feces or *vice versa* in children [21]. Here, we aim to characterize associations between stool consistency, total water content of stool, stool microbiota composition, and metabolite levels in infants and small children. We used stool samples and complementary markers for gut transit time: Stool water content and parental reports of crude stool consistency from 2.5-, 6-, 14-, and 30-month-old children.

Results

Longitudinal measurements

We measured the gut microbiota, metabolome, stool consistency, and water content from 618 stool samples (Figure 1). We obtained measurements at four time points of 2.5 months (n=360), 6 months (n=229), 14 months (n=274), and 30 months (n=169) of age. The distribution of participants providing samples was as follows: those who provided only one sample (n=326), two samples (n=190), three samples (n=82), and all four samples (n=20). We performed targeted metabolomics analysis of six short-chain fatty acids (SCFAs; acetic, propionic, iso-butyric, butyric, isovaleric, and valeric acids), and bile acids (BAs; including both taurine and glycine conjugates (n=33) and those microbial conjugated to 22 different amino acids (n=110)). We also detected metabolites panel including a wide range of chemical classes, such as amino acids, carboxylic acids (mainly free fatty acids and other organic acids), hydroxy acids, phenolic compounds, alcohols, and sugar derivatives in an untargeted metabolomics assay. Parents reported on crude stool consistency. Stool water content was measured as the percentage of water present in the stool sample. Table 1 shows clinical characteristics of the study subjects and Table S1 shows the sample sizes for available data types including measured metabolites, stool consistency, and water content analyses.

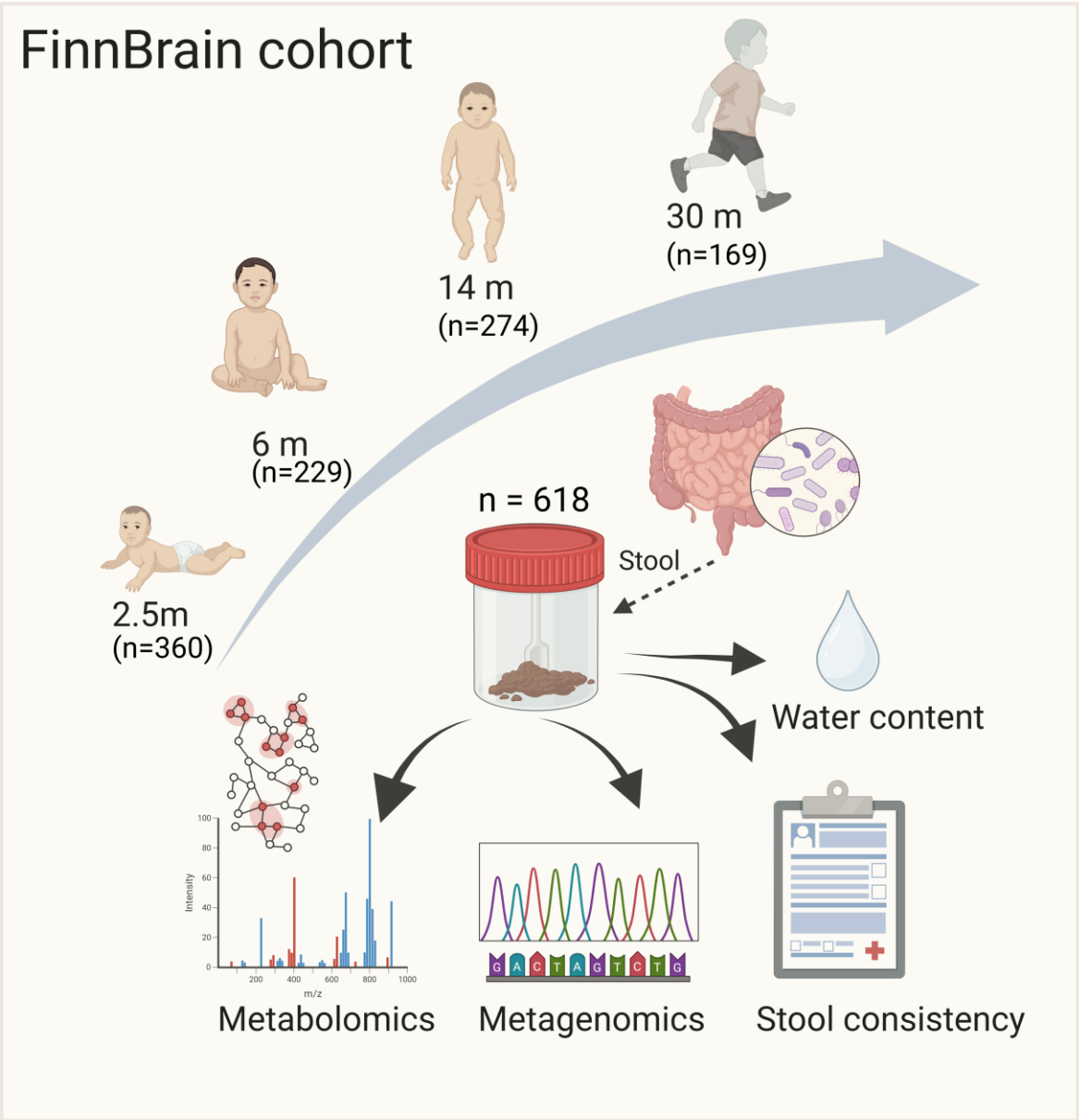

Figure 1. An overview of the study setting. We obtained longitudinal gut microbiota (16S rRNA), metabolome (targeted (bile acid, SCFAs) and untargeted (polar metabolites)), stool consistency, and water content measurements at four time points of 2.5 months (n=360), 6 months (n=229), 14 months (n=274), and 30 months (n=169) of age. Short-chain fatty acids (SCFAs). The figure is made in Biorender.

Table 1. Characteristics of the study subjects who had either stool consistency or water content were reported.

Timepoint (months)	2.5	6	14	30	Overall
Sample sizes	360	229	274	169	1032
Stool consistency					
constipation	23 (6.4%)	31 (13.5%)	9 (3.3%)	10 (5.9%)	73 (7.1%)
loose	32 (8.9%)	4 (1.7%)	14 (5.1%)	2 (1.2%)	52 (5.0%)
normal	189 (52.5%)	89 (38.9%)	105 (38.3%)	80 (47.3%)	463 (44.9%)
other or gas	58 (16.1%)	8 (3.5%)	4 (1.5%)	4 (2.4%)	74 (7.2%)
missing	58 (16.1%)	97 (42.4%)	142 (51.8%)	73 (43.2%)	370 (35.9%)
Water content					
mean (SD)	79.3 (5.70)	77.4 (6.43)	75.4 (5.50)	75.0 (5.86)	76.6 (6.06)
median [Min, Max]	79.3 [60.8, 98.6]	78.4 [47.2, 90.9]	76.2 [60.5, 87.3]	75.7 [53.6, 86.3]	77.2 [47.2, 98.6]
missing	236 (65.6%)	42 (18.3%)	30 (10.9%)	23 (13.6%)	331 (32.1%)
Delivery mode					
C-section	57 (15.8%)	37 (16.2%)	45 (16.4%)	23 (13.6%)	162 (15.7%)
vaginal	298 (82.8%)	191 (83.4%)	226 (82.5%)	145 (85.8%)	860 (83.3%)
missing	5 (1.4%)	1 (0.4%)	3 (1.1%)	1 (0.6%)	10 (1.0%)
Gestational age (weeks)					
mean (SD)	40.1 (1.36)	40.0 (1.38)	39.8 (1.52)	39.9 (1.36)	40.0 (1.41)
median [Min, Max]	40.1 [34.6, 42.3]	40.0 [34.1, 42.3]	40.1 [32.9, 42.3]	40.0 [35.6, 42.3]	40.1 [32.9, 42.3]

Antibiotic treatment prior 6 months					
no	300 (83.3%)	128 (55.9%)	111 (40.5%)	75 (44.4%)	614 (59.5%)
yes	2 (0.6%)	4 (1.7%)	22 (8.0%)	21 (12.4%)	49 (4.7%)
missing	58 (16.1%)	97 (42.4%)	141 (51.5%)	73 (43.2%)	369 (35.8%)
Current feeding					
full breastfeeding	220 (61.1%)	9 (3.9%)	0 (0%)	0 (0%)	229 (22.2%)
no breastfeeding	7 (1.9%)	1 (0.4%)	3 (1.1%)	1 (0.6%)	12 (1.2%)
partial breastfeeding	50 (13.9%)	145 (63.3%)	32 (11.7%)	1 (0.6%)	228 (22.1%)
ceased breastfeeding	10 (2.8%)	53 (23.1%)	217 (79.2%)	155 (91.7%)	435 (42.2%)
missing	73 (20.3%)	21 (9.2%)	22 (8.0%)	12 (7.1%)	128 (12.4%)
Maternal age (years)					
mean (SD)	30.8 (4.42)	30.7 (4.52)	31.1 (4.33)	31.0 (4.29)	30.9 (4.39)
median [Min, Max]	31.0 [19.0, 45.0]	30.0 [19.0, 42.0]	31.0 [19.0, 45.0]	30.0 [21.0, 45.0]	30.0 [19.0, 45.0]
Primiparity					
no	158 (43.9%)	105 (45.9%)	137 (50.0%)	74 (43.8%)	474 (45.9%)
yes	202 (56.1%)	124 (54.1%)	137 (50.0%)	95 (56.2%)	558 (54.1%)
Sex					
boy	189 (52.5%)	125 (54.6%)	150 (54.7%)	98 (58.0%)	562 (54.5%)
girl	171 (47.5%)	104 (45.4%)	124 (45.3%)	71 (42.0%)	470 (45.5%)

Demographic factors and stool characteristics

We investigated the association between demographic factors including age, breastfeeding status, delivery mode, recent antibiotics intake and premature birth (born < 37 gestational weeks) on stool consistency and water content. As expected, individuals with parent-rated constipation exhibited lower overall water content (Figure 2A, Kruskal-Wallis test, Dunn's test q=0.003), particularly at 6 and 14 months (Figure 2A, Kruskal-Wallis test, Dunn's test q = 0.032, q= 0.017, respectively). Normal stool consistency was associated with 3.6 percentage point higher water content compared to constipation, adjusting for sampling age and accounting for within-subject correlation using random intercepts (p=0.00025).

Our analysis revealed a declining trend in water content with increasing age (Figure 2B, Linear mixed model, $p = 5 \cdot 10^{-8}$). On average, when accounting for the intra-individual covariance, water content was 1.6 percent lower after a year. Moreover, subjects who were breastfed demonstrated higher water content compared to those who were not (Figure 2C, Wilcoxon signed-rank test $q = 8 \cdot 10^{-7}$). However, the magnitude of the association was small, and breastfeeding vs no breastfeeding increased the water content by 1 % (linear mixed model with breastfeeding and age as fixed effects, water content as dependent variable and subject identity as random effect, p=0.055).

Overall, breastfeeding was associated with stool consistency (χ^2 q = 0.00110), but delivery mode, antibiotic intake or prematurity at any specific time point or overall did not have an association. However, we observed a distinct pattern between stool consistency and age (Figure 2D). Specifically, we noted fewer occurrences of loose stool at 6 months compared to 2.5 months (Figure 2D, Dunn's test q = 0.028). Additionally, there were more reports of gas at 2.5 months compared to 6, 14, or 30 months (Figure 2D, Dunn's test $q = 2.3 \times 10^{-6}$, $q = 8.8 \times 10^{-11}$, $q = 1.2 \times 10^{-7}$, for 6, 14, 30 months, respectively). The prevalence of constipation was higher at 6 months compared to 2.5, 14, and 30 months (Figure 2D, Dunn's test $q = 3 \cdot 10^{-7}$, $q = 8 \cdot 10^{-6}$, q = 0.0026 for 2.5, 14, and 30 months, respectively. Furthermore, reports on normal stool

consistency were more common at 14 and 30 months compared to 2.5 months (Figure 2D, Dunn's test q = 0.0017, q = 0.0017 for 14 and 30 months, respectively).

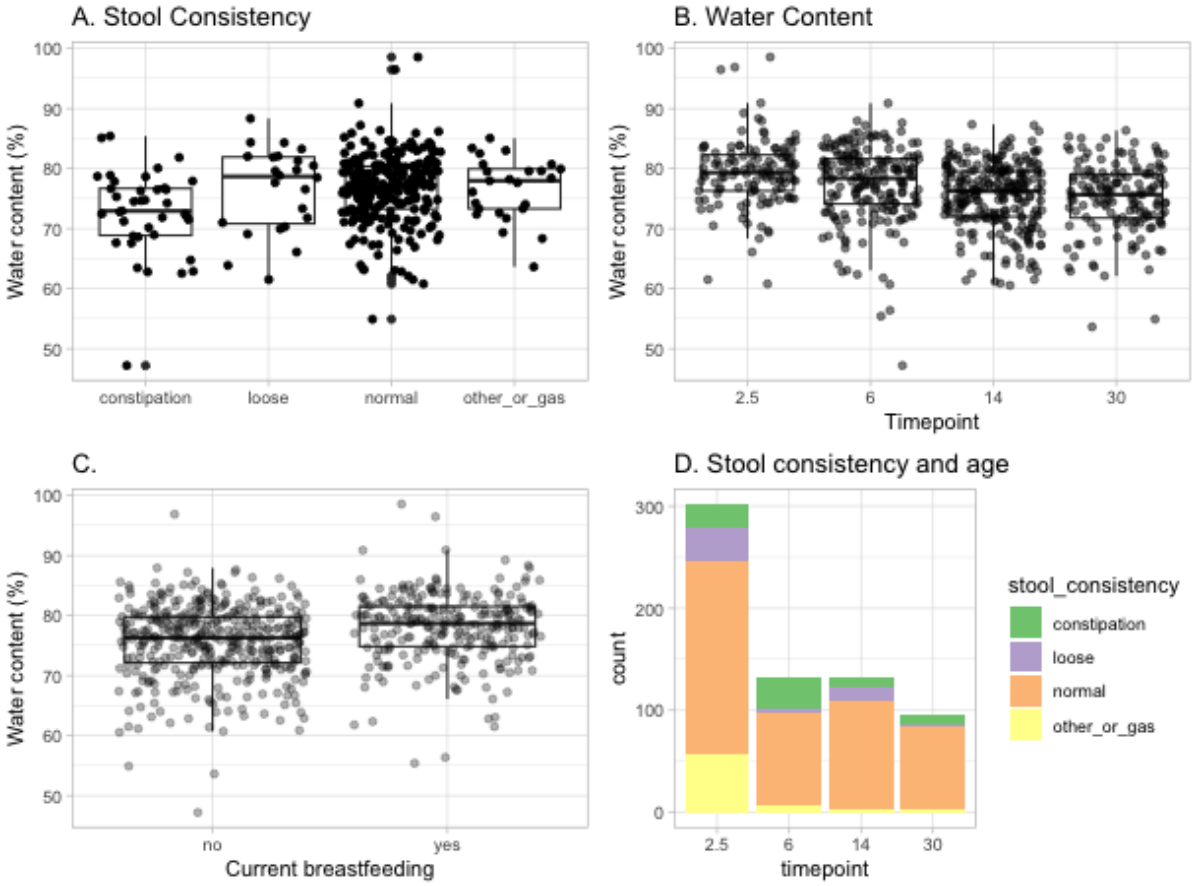

Figure 2. A. Constipated subjects had lower water content. B. Water content gradually decreased as the child aged, and **C.** consequently water content was associated with breastfeeding. D. There were more loose stools and reports of gas and other complaints at 2.5 months, whereas constipation was reported more at 6 months.

Taxonomic richness, diversity and composition associate with water content and stool consistency

We observed that both the Shannon index and observed richness at 6 months (Figure 3, Spearman ρ = -0.29, q = 0.003) and observed richness at 30 months (Figure 3, Spearman ρ= -0.2, q = 0.039) associated with stool water content. These associations were significant also after adjusting for breastfeeding (q = 0.0001 for Shannon index at 6 months, q = 0.0001 for observed richness at 6 months, q = 0.0166 for observed

richness at 30 months). Moreover, in a mixed model, both observed richness and Shannon index associated (Table S2) with water content. However, neither observed richness nor Shannon index were associated with stool consistency across time points or longitudinally (Table S2).

Figure 3: Taxonomic richness (observed richness) and diversity (Shannon index) and water content.

Community composition was associated with water content in distance-based redundancy analysis (Figure 4, Table S3, Bray-Curtis dissimilarity, db-RDA PERMANOVA $R^2 = 0.006$, $q = 0.012$) at both 6 (Table S3, $R^2 = 0.015$, $q = 0.012$) and 30 months (Table S3, $R^2 = 0.015$, $q = 0.016$). This association persisted after adjusting for breastfeeding in the entire dataset for 6 months. None of the subjects were breastfed for 30 months. The associations were similar with Aitchison distance, and additionally water content at 14 months was associated with water content while adjusting for breastfeeding. (Table S4, $R^2 = 0.012$, $q = 0.02$). At 6 months, variance in water content was positively associated with amplicon sequence variants (ASVs) annotated to *Clostridium, Escherichia,* and *Veillonella*, while it was inversely associated with *Bifidobacterium*. ASVs in *Bacteroides* and *Enterobacteriaceae* showed loading in both directions. Conversely, in 30-month-olds, the relative abundance of *Veillonella,*

Sutterella, and *Ruminococcus* were associated with lower water content, while *Bacteroides, Bifidobacterium,* and *Enterobacteriaceae* were associated with higher water content.

Community composition was associated with overall stool consistency (Figure4A, Table S3, PERMANOVA $R^2 = 0.003$, $q = 0.014$), even after adjusting for breastfeeding (Table S3, PERMANOVA $R^2 = 0.041$, $q = 0.038$). The associations remained similar while using Aitchison distance (Table S4). While in age-wise analysis, the association of stool consistency with community composition remained insignificant, there was differential dispersion in stool consistency groups (Figure 4G. PERMDISP2 $p = 0.003$), with constipated subjects showing a smaller distance to the centroid. Therefore, the statistical significance found in the db-RDA test may be due to heterogeneity in the group dispersions.

Next, we determined the link between individual genus relative abundances and stool consistency and water content, conducting both longitudinal analyses with Linear models for differential abundance analysis (LinDA) and age-wise analyses with ANOVA-Like Differential Expression (ALDEx2). Longitudinal analysis revealed a higher abundance of *Terrisporobacter* and *Varibaculum* in constipation compared to normal stool when adjusted for breastfeeding (Table S5). Multiple genera such as *Clostridium*, *Hungatella*, and *Lactobacillus* were indirectly associated with water content, with most associations persisting after adjusting for breastfeeding (Table S5, Figure S1). However, associations regarding stool consistency or water content were not observed in age-wise analysis (Table S4, S6).

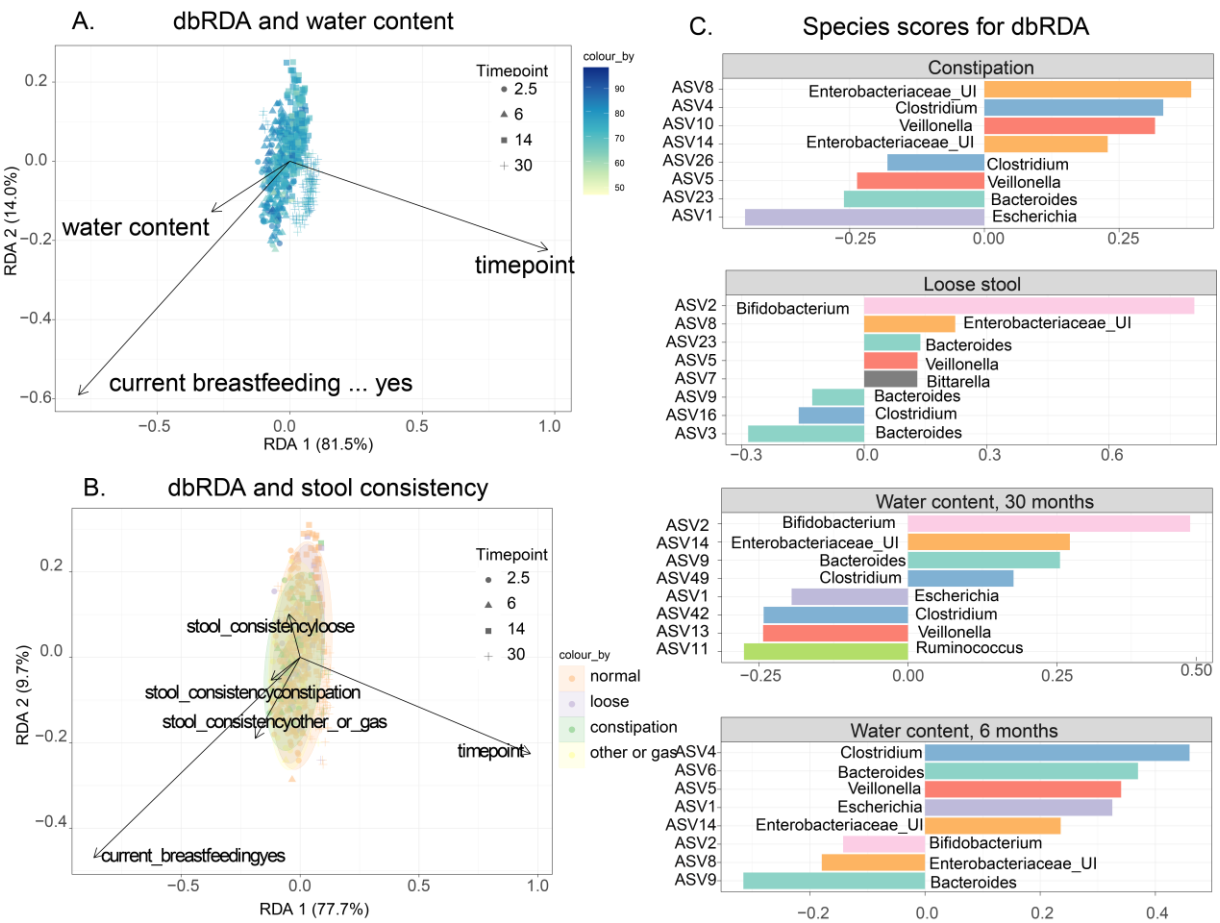

Figure 4: Stool consistency and community composition. db-RDA on A. water content and B. stool consistency. C. shows the ASVs that have the largest absolute coefficients in the RDA models. The bar colors represent the genus-level annotation of the ASV. Here, enterobacteriaceae UI is Enterobacteriaceae unidentified genus.

Stool metabolite concentrations and stool consistency and water content

Next, we sought to investigate whether stool metabolites were associated with stool consistency and water content (Figure 5A-B). We found that constipated subjects had lower levels of tauro- and glycoconjugated bile acids compared to subjects with normal or loose stools at both 2.5 and 14 months (Figure 5, Table S8, S9, Kruskal-Wallis test, q= 0.020 for tauroconjugated BA at 2.5 months, q = 0.029 for glycoconjugated BA at 14 months, q = 0.037 for tauroconjugated BA at 14 months < 0.05). Acetic acid concentration was also lower in constipated subjects compared to normal subjects at 6 months (Figure 5, Figure S4, Table S8 Kruskal-Wallis test q = 0.019). Similarly, the aggregate sum of all SCFA concentrations seemed to be lower in

constipated individuals at 6 months (Figure 5C-D, Kruskal-Wallis test q =0.085, Dunn's test q = 0.042). Additionally, several polar metabolites, including 3-3-hydroxyphenylpropionic acid and fructose were lower in constipated subjects (Table S8, Figure S5). We also found that recently discovered bile acid amidates—microbially derived CA-Asn, CA-Thr, and CA-Arg had lower concentration in constipated children, whereas CDCA-Val, UDCA/HDCA-Lys, UDCA-Ile/Leu had higher concentration (Table S 9-10, Figure S7, S12).

We also investigated whether the stool metabolite content was associated with the water content (Figure 5A, Table S9-10, Figures S6-12). Consistent with microbiota findings, multiple metabolite associations remained in age-wise analysis (n = 237 associations, Table S9). Most of these associations remained unchanged even after adjusting for breastfeeding (n = 203 associations, Table S10). Tauro and glycol conjugated bile acids were positively associated with water content across all time points (Figure S6). Primary bile acids were also positively correlated with water content at 14 and 30 months, although with low coefficients (Spearman ρ 0.17 and 0.14, respectively, Table S9). On the other hand, we found that most of the novel conjugated bile acid amidates were negatively associated with water content at 2.5, 6, 14, and 30 months (Figure S11, S12, Table S9) with exception for CDCA-Asn, CA-Asn, CDCA-Trp, UDCA-Met, CA-Thr which showed consistently positive association at 6 and 30 months (Figure S11, S12, Table S9). We also found that acetic acid was positively associated with water content at 6, 14, and 30 months (Figure S9). Butyric acid was positively associated at 6 months (Figure S9). Branched SCFAs, particularly isobutyric and isovaleric acids, were negatively associated with water content at 6, 14, and 30 months (Figure S9). The sum of all SCFA concentrations was directly correlated with water content at 6, 14, and 30 months (Figure 5C, Table S10). Furthermore, fucose, fructose, glycerol, and pinitol were positively associated with water content, consistent with the findings from constipated subjects (Figure 5B, Figure S10).

In the longitudinal analyses with mixed models, taurine and glycine conjugated bile acids inversely associated with constipation when compared to normal stool consistency after adjusted for age and further with breastfeeding (Figure 5A, Table

S12). Moreover, we observed in linear mixed models that tauroconjugated bile acids, primary bile acids, acetic acid, fructose, pinitol and several other polar metabolites were positively associated with water content, whereas isobutyric acid and isovaleric acid were negatively associated, both adjusted and unadjusted for breastfeeding (Figure 5B, Table S13). Additionally, water content is associated directly, and constipation associated indirectly with the sum of all SCFA (Figure 5C-D, linear mixed-model, $p = 0.9 \times 10^{-17}$ for water content, $p = 0.047$ for constipation).

Microbially conjugated bile acid amidates had mostly negative associations with water content (DCA/HDCA/UDCA-Val, CA/CDCA/UDCA/HDCA-Ala, UDCA-Trp, DCA/HDCA-Ile/Leu, Figure 5B, Table S12-S13, Figure S11). On the other hand, methionine and alanine conjugated bile acid had positive associations with water content (Figure 5B, Table S12-S13, Figure S11).

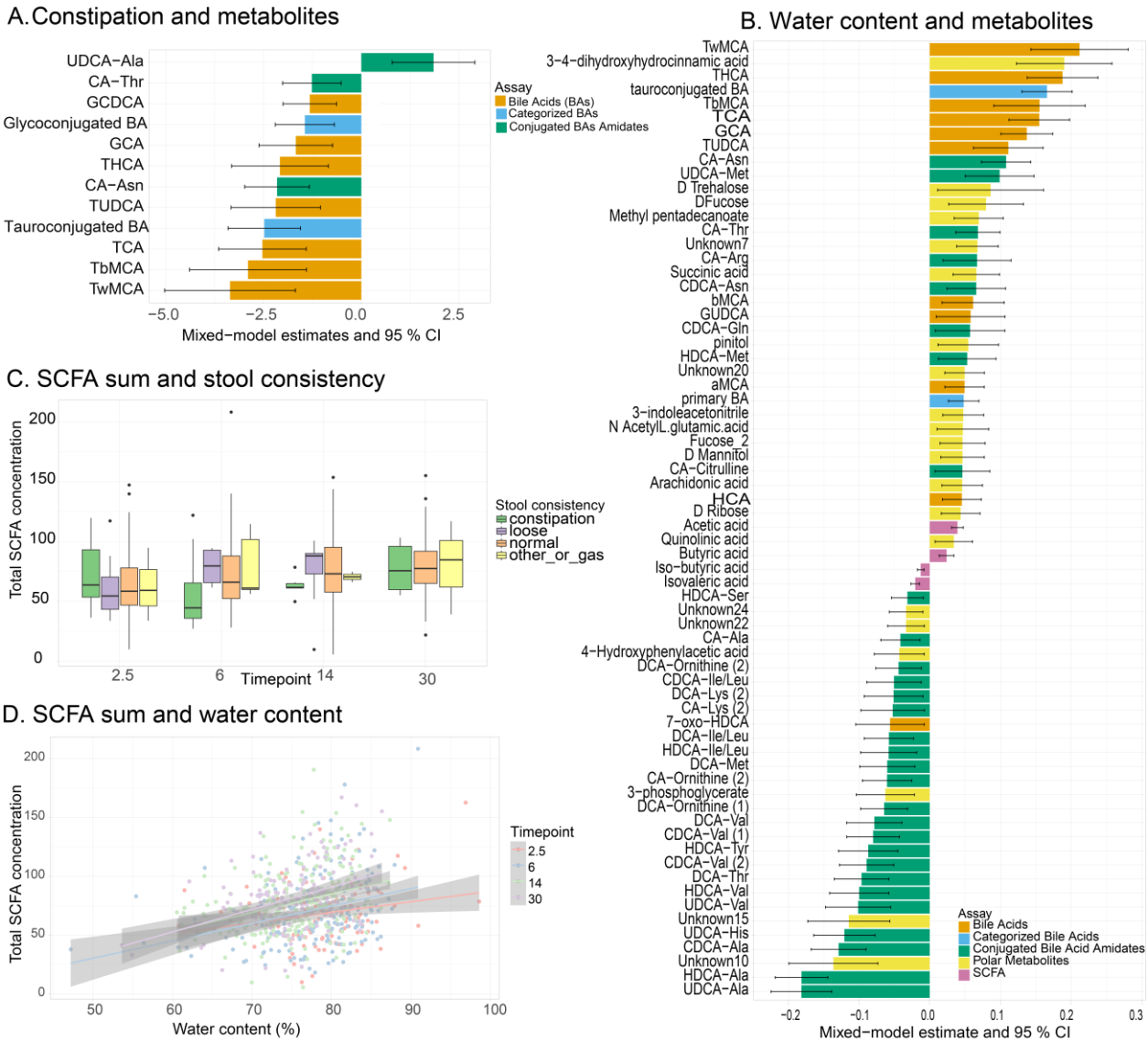

Figure 5. Metabolite concentrations associated with water content and constipation in mixed models with metabolite concentration as the dependent variable, stool consistency or water content as the fixed effect, any current breastfeeding as fixed effect, and child identity as random effect. A. Constipation compared with normal stool consistency associated with taurine- and glycine conjugated bile acids. B. Water content associated positive with taurine- and glycine conjugated bile acids, bile acid amidates, polar metabolites such as fucose, acetic and butyric acids. Water content associated negatively with isobutyric and isovaleric acids, and several bile acid amidates. C. The constipated individuals had lower total SCFA at 6 months. The only

study subject with loose stools at 30 months was removed from the figure. D. Water content associated positively with overall SCFA concentrations.

Breastfeeding, stool consistency and fecal metabolites

In clinical practice, breastfeeding is considered to be beneficial for alleviating constipation in infants [22, 23]. We have previously reported that breastfeeding is associated with metabolite levels (), and our data also shows that breastfeeding was associated with higher water content and stool consistency. Hence, we wanted to explore if stool water content or stool consistency can mediate the associations between breastfeeding and stool metabolites. However, stool water content did not seem to mediate the association between breastfeeding on metabolites levels, including the conjugated bile acid amidates, based on average causal mediated effect (Table S15). At 2.5 months, partial vs full breastfeeding was partially mediated via constipation to higher levels of UDCA-Ornithine (1) and UDCA/HDCA-Ornithine and lower levels of 4-Hydroxyphenylacetic acid (Figure S, Table S16). However, the UDCA-Ornithine (1) and UDCA/HDCA-Ornithine were detected only in few individuals at 2.5 months.

Further, the interaction between breastfeeding and water content was significantly associated with concentrations of CA-His, CA-Met, CA-Ornithine, CA-Ser in addition to ribitol, L-arginine and TUDCA (Table S17, Figure S14). This indicates that breastfeeding and higher stool water content associated with lower concentrations of these CA-conjugated amidates, whereas non-breastfed infants were associated with showed the opposite pattern or no linear trend (Figure S14). Moreover, non-breastfed subjects with higher water content had increased concentrations of TUDCA, when breastfed subjects didn't seem to differ in TUDCA concentration regardless the water content (Figure S14).

Predicted enzymatic capacity for bile acid and SCFA metabolism

Next, we also examined whether predicted bacterial enzymes involved in SCFA and BA metabolism were associated with stool consistency and water content using PICRUSt2

(Phylogenetic Investigation of Communities by Reconstruction of Unobserved States). We found mainly genes involved in acetic acid metabolism, and to a lesser extent with propionic acid, butyric acid and bile acids in our data (Figure S15).

Longitudinal analyses with mixed models revealed that the abundance of SCFA-metabolizing enzymes was positively associated with constipation compared to lose stool consistency, after adjusting for age and breastfeeding. Specifically, enzymes involved in acetic and butyric acid metabolism were more abundant in individuals with constipation than in those with loose stools at both 2.5 and 6 months. Furthermore, in a linear mixed models adjusted for age and breastfeeding as fixed effects and child identity as random effects, the abundance of acetic, butyric, and propionic acid-metabolizing bacterial enzymes was negatively associated with stool water content. Similarly, for bile acid metabolism, we found that the abundance of chologlycine hydrolase, a bile salt hydrolase primarily involved in the deconjugation of primary bile acids, was negatively associated with stool water content (Figure S16).

Polygenic score for stool frequency and stool microbiota and metabolites

To further understand the interplay between stool consistency and transit time, we also examined other known potential surrogates for gut transit time. A genome-wide association study has shown that there are multiple loci that associate with a questionnaire-based assessment of stool frequency, a proxy for gut motility (PMID: 34957435), that is negatively related to gut transit time (PMID: 33722860). In the report, the genetic information was used to calculate a polygenetic score (PGS) for stool frequency, and the high PGS scores were related to risk for irritable bowel syndrome and particularly the diarrhea-dominant type (PMID: 34957435). In our sample, the mean of PGS was lower among those with normal stool consistency, but this was not statistically significant. Moreover, PGS was not associated the water content.

PGS was not associated with taxonomic alpha diversity or community composition. Neither did PGS associate with any taxa cross-sectionally or longitudinally. UDCA/HDCA-Val associated positively with the PGS in a mixed model adjusted for age and breastfeeding as fixed effects and child identity as random effects (q = 0.0329

for both, Figure S17, Table S18). At nominal significance level, CA-Ornithine, UDCA-His, UDCA/HDCA-Glu, and UDCA-Trp among other bile acid amidates negatively associated with PGS (Table S18). These were partially replicated in cross-sectional linear regressions adjusted for breastfeeding (CA-Ornithine, CDCA-Lys, UDCA-His, UDCA-Ala), but other nominally significant findings, such as positive association with HDCA-Ala at 6 months, 7-oxo-converted BA at 6 months, 3,3-hydroxyphenylpropionic acids at 14 months, UDCA-Gln at 14 months, and D-fucose at 14 months were observed (Table S18).

Discussion

We show in infants and small children that both stool microbiota and metabolome link with the stool consistency and the total water content. This corroborates earlier findings indicating that stool consistency and gut transit time correlate with microbiota and metabolism in adults [2, 14, 24]. However, data especially from infants and with longitudinal study design are scarce. Our findings indicate more extensive microbial metabolism in low-moisture stool and in constipated children. Moreover, our findings also underscore the potential of using the total water content as an objective measure in microbiota studies, particularly during early infancy when stool consistency vary considerably. We also found agreement between the total stool water content and parental crude reports of constipation.

Fluctuations in stool consistency correlate with microbiota taxonomic richness, and community composition[2, 17, 24, 25]. Vandeputte et al. (2016) found that firmer stools in Bristol stool form scale were associated with a higher Bacteroidetes:Firmicutes ratio and elevated levels of *Akkermansia* and *Methanobrevibacter* whereas loose stools were associated with lower taxonomic richness in adults [2]. In line, our study in infants and small children reveals that stool consistency associates with the bacterial community composition, particularly through variations in *Bacteroides, Bifidobacterium, Ruminococcus,* and *Clostridium* abundances. We also found an ASV in *Bacteroides* genus associated with the water content-related community composition positively at 30 months and negatively at 6 months. However, a previous study reported a different

direction in the association between stool consistency and microbiota composition during infancy and childhood [21]. In the longitudinal analysis, genera including *Clostridium*, *Prevotella*, and *Eggarthella* associated negatively with water content. Many of these genera are known as dietary fiber fermenters, specifically *Clostridium* and *Prevotella*, and slower transit in in vitro simulator increases the growth of fiber fermenters such as *Prevotella* [PMID: 37926855]. Apart from fiber fermentation, both *Clostridium* and *Eggarthella* contain species implicated in the bile-acid metabolism [26, 27]. Intriguingly, Kwiatkowska *et al.* 2024 reported that functional constipation in three-year-olds was associated with depletion of rare genera, such as butyrate-producer *Subdoligranulum* [28]. We observed that constipated individuals had more homogenous community composition compared to other stool consistency groups, and the lack of rare taxa could partly explain our observation, however, we did not directly assess prevalence or abundance of rare taxa. Overall, these findings align with previous research on the relationship between stool consistency and microbiota, although differences in associated taxonomic groups may result from gut microbiota alterations across different life stages [29].

Regarding the metabolic features, we show that the associations between metabolites and low stool water content and parent-report of constipation are qualitatively similar: children with constipation and lower water content also showed decreased concentrations of host-conjugated bile acids (glycine and taurine-bound). Gut microbes regulate human bile acid metabolism [30]. Primary bile acids are synthesized and conjugated in the liver, stored in the gallbladder, secreted in the intestine after meal consumption, and later modified (de-conjugated or again re-conjugated) by microbes in the gut before reabsorption (95%) or excretion (5%) [31]. Thus, the decreased level of host-conjugated BAs in egested constipated stool suggests that longer transit time allows for more extensive bacterial metabolism of bile acids in the gut[24]. Moreover, bacteria can then conjugate bile acids with other amino acids and produce novel conjugated bile acid amidates. In line, many microbially conjugated bile acid amidates were negatively associated with water content as well as the polygenic score for defecation frequency. Aberrant bowel movement frequencies have been

linked to constipation and alterations in microbially derived metabolites in the circulation [32]. This again indicates that longer transit time allows for more microbially driven re-conjugation of bile acids. This might have health implications, since biliary excretion of xenobiotics is an important elimination route, and deconjugation might delay clearance of substances such as estrogens [33].

Besides bile acids, constipation and lower water content were associated with less SCFA, specifically acetic acid, while branched short-chain fatty acids (BCFAs) concentrations were inversely associated with stool water content at 14 and 30 months. In fact, SCFAs, primarily acetate, propionate, and butyrate, are end products of fermentation of complex carbohydrates such as dietary fiber [34, 35], while increased BCFAs reflect ongoing putrefaction or/and proteolytic activities in the colon. We hypothesize that prolonged colonic transit time depleted available carbohydrates for fermentation which consequently induced proteolytic activity, regarded as an alternative pathway for carbon and nitrogen, and an energy source for the microbes in the colon. Interestingly, another study showed that by reducing gut transit time with drugs, the concentrations of total SCFAs could be increased [36]. Our finding is consistent with similar works in adults, where lower water content or longer gut transit time showed to shift in colonic metabolism towards proteolytic activity [14, 24, 36-38]. Roager *et al*. found that longer transit time was associated with higher levels of protein-derived molecules in the urine such as p-cresol-sulfate and phenylacetylglutamine, indicating a shift from saccharolytic to proteolytic activity in the colon [14]. Also, Zhang *et al*. found that the Bristol stool form scale scores were found to negatively associate with the levels of fecal amino acids, suggesting increased protein catabolism in irritable bowel syndrome patients [39]. In adults, metabolites from protein catabolism have been implicated in atherosclerotic cardiovascular diseases and colon cancer [40-42]. Hence our observation that higher BCFA concentrations were associated with constipation might have further health-implications, although this has not yet been studied in infants or small children.

The predicted metabolic capacity for SCFA and BA showed that genes related to SCFA, mainly acetic acid, metabolism were negatively associated with water content and increased in constipated individuals. Although 16s sequencing and limited bile acid

metabolism coverage in the database, we could show that the microbiome relates with proxies of transit time not only on compositional level, but also on predicted functional properties. Interestingly, bile salt hydrolase was negatively associated with water content. Bile salt hydrolase possesses dual functions in bile acid metabolism; it has been shown to both deconjugate as well re-conjugate these novel microbial bile acids amidate [43]. Hence it is plausible that faster transit time relates to lower abundance of bile salt hydrolase.

As expected, [22, 23], breastfeeding was related to higher water content, although with a small magnitude. Breastfeeding has been considered to promote healthy gut transit time and alleviate constipation, that is common in infants receiving solid foods, and hence it was interesting to test how breastfeeding modulates the link between stool consistency and gut microbiota and metabolites. However, most of our associations remained similar regardless of breastfeeding. In addition, in this general population based sample, stool water content did not mediate the associations between breastfeeding and fecal metabolites. In our data, infants receiving formula and with constipation had higher concentrations of 4-hydroxyphenylacetic acid. We did observe that, in 2.5-month-olds, constipation partially mediated the link between full vs. partial breastfeeding and 4-hydroxyphenylacetic acid, a phenolic compound in olive oil and oats as well as a microbial metabolite [44] . This may reflect that, in a young infant, formula feeding relates to increased microbial fermentation of aromatic amino acids in the context of slow gastrointestinal transit time. In addition, the interaction analysis suggested that the combination of breastfeeding and higher water content related with lower concentrations of cholic acid conjugated amidates, whereas no such pattern was visible for non-breastfed infants. This potentially indicates that breastfeeding together with slower transit time may relate to increased bacterial re-conjugation of cholic acid. In all, this reflects that early nutrition in conjunction with gut transit time, relates to bacterial metabolism.

We expect that our findings may reflect the influence of longer colonic transit time on intestinal microbial communities potentially via selectively favoring certain microbes and their activity and levels of microbe-regulated metabolites[25]. However, other mechanisms may underline the links between microbiota and gut motility. Although

we observed concurrent associations and we did not study predictive capacity of gut microbiome on later markers of transit time, it should be noted that early life microbiota may affect the development of enteric nervous system and consequently (later) gut motility and secretions [45-48]. Although we studied generally healthy sample, our results highlight a need to study whether similar microbiota and microbially regulated metabolite signatures could be seen in pathological context, such as in disorders of gut-brain interaction in older children [28, 49-51].

Our findings of parent-reported stool consistency, water content, child age and breastfeeding are in line with earlier literature and clinics. First, the information that parent-reported constipation associated with objective measure of lower water content may be clinically relevant to support future research to avoid invasive examination techniques of colonic transit time [52, 53]. Second, the association between breastfeeding and stool consistency and water content is in line with general clinical observation that breastfed infants seldom suffer from constipation [54], and prolonged breastfeeding may protect from functional constipation[22]. Moreover, parental reports of gas were prevalent at 2.5-months-of-age when over half of the infants were breastfed and they typically swallow air during breastfeeding which leads to gas problems [54] . Third, constipation becomes more prevalent at 6 months that aligns well with the introduction of solid foods and reduction of breastfeeding. Our observation that water content gradually decreases by age likely reflects the normative prolonging of gut transit time by age [55, 56]. Interestingly, polygenic score for defecation frequency was not related to the parent-reported stool consistency not the water content. Our study suggests that genetic surrogate for defecation frequency is linked to microbially driven bile acid metabolism, aligning with recent observations of other microbially derived compounds such as 3-indoxyl sulfate and aberrant bowel movement frequencies[32]. The metabolite signatures were not comparable between high water content and polygenic score and warrants for future studies. To the best of our knowledge, the polygenic score for defecation frequency has not been applied in a child cohort before nor has it been studied in relation to gut microbiome.

Our study has limitations. We used crude parental report of child stool consistency, which is not as refined as the Bristol stool scale, and this may have implications on

comparability or generalizability of our results but still, it was associated with the objective measure of total stool water content. Stool water content is different from *e.g.* gastrointestinal transit time assessment with radio-opaque markers, but we studied vulnerable study subjects of healthy infants and young children, where utilizing invasive measuring techniques has ethical constrains. Utilizing stool samples does not allow proper inference of absorption of the microbiota-produced metabolites in the whole gut. Serum or urinary metabolome analysis may have given greater insight into the systemic alterations related to the saccharolytic *vs.* proteolytic fermentation and enterohepatic circulation. Moreover, one can speculate that extremely low water content in the faeces may alter the DNA extraction success, which may subsequently lead to changes in the microbiota composition. However, we are not aware of methodological studies addressing this.

In summary, our findings in infants and small children support the concept that stool consistency and total water content are associated with gut microbiota composition and metabolic activity. We found microbiota community composition to be associated with stool consistency, and these varied partly by age. Our result suggests metabolic shifts from saccharolytic to proteolytic activity and decreased concentrations of host-conjugated bile acids and increased concentrations of microbially conjugated bile acids related to constipation and lower water content, providing insights into gut health and function. Stool consistency and water content related to gut microbiota composition and metabolome mostly independently of breastfeeding, but combination of both breastfeeding and low water content related to higher concentrations of cholic acids microbial conjugates. This indicates an interaction between early nutrition and gut physiology is important for microbial metabolism.

Methods

Cohort and Questionnaire Data

The study subjects are children from the FinnBrain Cohort Study [57]that is a general population birth cohort study located in the southwestern Finland. The Ethical Committee of Southwestern Finland approved the study. Parents provided informed consent on behalf of their children. STORMS guideline was used for reporting the methods and materials[58].

The sample used in the current study were a subset of children that participated in a study visit at 2.5, 6, 14 and 30 months postpartum. The fecal samples were collected from May 2013 to May 2018. The samples were collected in plastic tubes, and parents were instructed to store the sample in a refrigerator, and bring the sample to the laboratory within 24 h. The sample collection time was reported.

Information on breastfeeding was collected with questionnaires to the mother at 3, 6, 12, and 24 months postpartum and during study visits 2.5, 6, 14, and 30 months. Information on delivery mode and gestational age was collected from the hospital records (The Wellbeing Services County of Southwest Finland, VARHA). Breastfeeding was categorized as any current breastfeeding (yes vs. no). At the sample collection, the parent reported the child stool consistency as: normal, loose, constipated, gas or other.

DNA Extraction, Sequencing and Data Processing

The samples were divided into cryotubes and freezed in -80C within 2 days after arriving at the laboratory. Samples were kept at +4C before freezing. Only samples that were freezed within 48 h of sample collection were sequenced. DNA was extracted using a semi-automatic extraction instrument Genoxtract with DNA stool kit (HAIN life science, Germany). The DNA extraction was performed according to the manufacturer's protocol except that 1 ml of lysis buffer was added and the samples were homogenized with glass beads 1000 rpm / 3 min. The samples were centrifuged at high speed (> 13000 rpm) for 5 min. The lysate (800µL) was then transferred to tubes and the extraction proceeded according to the manufacturer’s protocol. DNA yields were measured with Qubit fluorometer using Qubit dsDNA High Sensitivity

Assay kit (Thermo Fisher Scientific, USA). Bacterial community composition was determined by sequencing the V4 region of 16S rRNA gene using Illumina MiSeq platform (Illumina, USA). The sequence library was constructed with an in-house developed protocol as described in [59, 60].Positive control (DNA 7-mock standard) and negative control (PCR grade water) were included in library preparation and sequencing runs.

DADA2-pipeline (version 1.14) was used to preprocess the 16S rRNA gene sequencing data to infer exact amplicon sequence variants (ASVs) (PMID: 27214047). The reads were truncated to length 225 and reads with more than two expected errors were discarded (maxEE = 2). SILVA taxonomy database (version 138) and RDP Naive Bayesian Classifier algorithm ref were used for the taxonomic assignments of the ASVs [61].

Stool Water Content

To quantify the water content, stool samples underwent freeze-drying. Prior drying, samples were weighted and briefly centrifuged in the bottom of the tubes. Subsequently, they were transferred to a custom freeze-drier equipped with a vacuum pump. The samples were left to dry overnight and then reweighed to determine their dry weight.

Metabolite Analyses

The BAs, SCFA and untargeted metabolome were previously described and adapted for this publication[59].

Targeted BAs measurement

Conjugated BA amidates were extracted by adding 80 µL fecal homogenate to 400 µL crash solvent (methanol containing 62,5 ppb each of the internal standards TCA-d4, GUDCA-d4, GCA-d4, CA-d4, UDCA-d4, GCDCA-d4, CDCA-d4, DCA-d4, GLCA-d4 and 625 ppb of LCA-d4) and filtering them using a Supelco protein precipitation filter plate after vortexing. The samples were dried under a gentle nitrogen flow and resuspended using 40 µL resuspension solution (Methanol:water (40:60) as in injection standard). Quality control (QC) samples were prepared by combining an

aliquot of every sample into a tube, vortexing it and preparing QC samples in the same way as the other samples. Blank samples and conjugated BA standard mixture (made initially by combinatorial chemistry) were prepared by pipetting 400 µL crash solvent into a 96-well plate, then drying and resuspending them the same way as the other samples.

The LC separation was performed on an Acquity Premier system consisting of a binary pump, an autosampler set to 10 °C and a column oven set to 45 °C. An Acquity premier HSS T3 1.8um (2.1x100 mm) column with a precolumn with the same material was used. Eluent A was 2 mM ammonium acetate in water and eluent B was 2 mM ammonium acetate in methanol.

The gradient started with 33.5 % B for half a minute and increased to 72 % B over three minutes. Next, the gradient increased to 100 % B over 7.5 minutes and was kept at 100 % B for 5 minutes. The flow rate was 0.4 mL/min and the injection volume was 5 µL. The mass spectrometer used for this method was a Sciex 7600 ZenoTof mass spectrometer operating in information-dependent mode with positive polarity. The ion source gas 1 was 60 psi and gas 2 was 80 psi. The curtain gas was 35 psi, the CAD gas was 7 and the temperature was 450 °C. The spray voltage was 5500 V.

The BAs were measured in fecal samples as described previously [59, 62]. Only samples frozen within 24 h of sample collection were included in the metabolome analyses. The order of the samples was randomized before sample preparation. Two aliquots (50 mg) of each fecal sample were weighed. An aliquot was freeze-dried prior to extraction to determine the dry weight. The second aliquot was homogenized by adding homogenizer beads and 20 µL of water for each mg of dry weight in the fecal sample, followed by samples freezing to at least -70 °C and homogenizing them for five minutes using a bead beater. The BAs analysed were Litocholic acid (LCA), 12-oxo-litocholic acid(12-oxo-LCA), Chenodeoxycholic acid (CDCA), Deoxycholic acid (DCA), Hyodeoxycholic acid (HDCA), Ursodeoxycholic acid (UDCA), Dihydroxycholestanoic acid (DHCA), 7-oxo-deoxycholic acid (7-oxo-DCA), 7-oxo-hyocholic acid (7-oxo-HCA), Hyocholic acid(HCA), β-Muricholic acid (b-MCA), Cholic acid (CA), Ω/α-Muricholic acid (w/a-MCA), Glycolitocholic acid (GLCA), Glycochenodeoxycholic acid (GCDCA), Glycodeoxycholic acid (GDCA), Glycohyodeoxycholic acid (GHDCA),

Glycoursodeoxycholic acid (GUDCA), Glycodehydrocholic acid (GDHCA), Glycocholic acid (GCA), Glycohyocholic acid (GHCA), Taurolitocholic acid (TLCA), Taurochenodeoxycholic acid (TCDCA), Taurodeoxycholic acid (TDCA), Taurohyodeoxycholic acid (THDCA), Tauroursodeoxycholic acid (TUDCA), Taurodehydrocholic acid (TDHCA), Tauro-α-muricholic acid (TaMCA), Tauro-β-muricholic acid (TbMCA), Taurocholic acid (TCA), Trihydroxycholestanoic acid (THCA) and Tauro-Ω-muricholic acid (TwMCA). BAs were extracted by adding 40 µL fecal homogenate to 400 µL crash solvent (methanol containing 62,5 ppb each of the internal standards LCA-d4, TCA-d4, GUDCA-d4, GCA-d4, CA-d4, UDCA-d4, GCDCA-d4, CDCA-d4, DCA-d4 and GLCA-d4) and filtering them using a Supelco protein precipitation filter plate. The samples were dried under a gentle flow of nitrogen and resuspended using 20 µL resuspenstion solution (Methanol:water (40:60) with 5 ppb Perfluoro-n-[13C9]nonanoic acid as in injection standard). Quality control (QC) samples were prepared by combining an aliquot of every sample into a tube, vortexing it and preparing QC samples in the same way as the other samples. Blank samples were prepared by pipetting 400 µL crash solvent into a 96-well plate, then drying and resuspending them the same way as the other samples. Calibration curves were prepared by pipetting 40 µL of standard dilution into vials, adding 400 µL crash solution and drying and resuspending them in the same way as the other samples. The concentrations of the standard dilutions were between 0.0025 and 600 ppb. The LC separation was performed on a Sciex Exion AD 30 (AB Sciex Inc., Framingham, MA) LC system consisting of a binary pump, an autosampler set to 15 °C and a column oven set to 35 °C. A waters Aquity UPLC HSS T3 (1.8µm, 2.1×100mm) column with a precolumn with the same material was used. Eluent A was 0.1 % formic acid in water and eluent B was 0.1 % formic acid in methanol. The gradient started from 15 % B and increased to 30 % B over 1 minute. The gradient further increased to 70 % B over 15 minutes. The gradient was further increased to 100 % over 2 minutes. The gradient was held at 100 % B for 4 minutes then decreased to 15 % B over 0.1 minutes and re-equilibrated for 7.5 minutes. The flow rate was 0.5 mL/min and the injection volume was 5 µL. The mass spectrometer used for this method was a Sciex 5500 QTrap mass spectrometer operating in scheduled multiple reaction monitoring mode in negative mode. The ion source gas1 and 2 were both 40 psi. The curtain gas was 25 psi, the CAD

gas was 12 and the temperature was 650 °C. The spray voltage was 4500 V. Data processing was performed on Sciex MultiQuant.

Quantification of SCFA. We adapted and modified the targeted SCFA analysis from previous work [63]. Fecal samples were homogenized by adding water (10 µL per mg of dry weight as determined for the BA analysis) to wet feces, the samples were homogenized using a bead beater. Analysis of SCFA was performed on fecal homogenate (50 µL) crashed with 500 µL methanol containing internal standard (propionic acid-d6 and hexanoic acid-d3 at 10 ppm). Samples were vortexed for 1 min, followed by filtration using 96-Well protein precipitation filter plate (Sigma-Aldrich, 55263-U). Retention index (RI, 8 ppm C10-C30 alkanes and 4 ppm 4,4-Dibromooctafluorobiphenyl in hexane) was added to the samples. Gas chromatography (GC) separation was performed on an agilent 5890B GC system equipped with a Phenomenex Zebron ZB-WAXplus (30 m × 250 µm × 0.25 µm) column a short blank pre-column (2 m) of the same dimensions was also added. A sample volume of 1 µL was injected into a split/splitless inlet at 285°C using split mode at 2:1 split ratio using a PAL LSI 85 sampler. Septum purge flow and split flow were set to 13 mL/min and 3.2 mL/min, respectively. Helium was used as carrier gas, at a constant flow rate of 1.6 mL/min. The GC oven program was as follows: initial temperature 50°C, equilibration time 1 min, heat up to 150°C at the rate of 10°C/min, then heat at the rate of 40°C/min until 230°C and hold for 2 min. Mass spectrometry was performed on an Agilent 5977A MSD. Mass spectra were recorded in Selected Ion Monitoring (SIM) mode. The detector was switched off during the 1 min of solvent delay time. The transfer line, ion source and quadrupole temperatures were set to 230, 230 and 150°C, respectively. Dilution series of SCFA standards of acetic, propionic, butyric, valeric, hexanoic acid, isobutyric, and iso-valeric acid were prepared in concentrations of 0.1, 0.5, 1, 2, 5, 10, 20, 40, and 100 ppm for the construction of standard curves for quantification. SCFAs were quantified using Agilent Masshunter.

Polar metabolites were extracted in methanol. The method was adapted from the method used by [64]. Fecal homogenate (60 µL) were diluted with 600 µL methanol crash solvent containing internal standards (heptadecanoic acid (5 ppm) valine-d8 (1

ppm) and glutamic acid-d5 (1 ppm)). After precipitation the samples were filtered using Supelco protein precipitation filter plates. One aliquot (50 µL) was transferred to a shallow 96-well plate to create a QC sample. The rest of the sample volume was dried under a gentle stream of nitrogen and stored in -80 °C until analysis. After thawing the samples were again dried to remove any traces of water.
Derivatization was carried out on a Gerstel MPS MultiPurpoe Sampler using the following protocol: 25 µL methoxamine (20 mg/mL) was added to the sample followed by incubation on a shaker heated to 45 °C for 60 minutes. N-Methyl-N-(trimethylsilyl) trifluoroacetamide (25 µL) was added followed by incubation (60 min). After that, 25 µL retention index was added, the sample was allowed to mix for one min followed by injection. The automatic derivatization was carried out using the Gerstel maestro 1 software (version 1.4). Gas chromatographic (GC) separation was carried out on an Agilent 7890B GC system equipped with an Agilent DB-5MS (20 m x 0,18 mm (0,18 µm)) column. A sample volume of 1 µl was injected into a split/splitless inlet at 250°C using splitless mode. The system was guarded by a retention gap column of deactivated silica (internal dimensions 1.7 m, 0.18 mm, PreColumn FS, Ultimate Plus Deact; Agilent Technologies, CA, USA). Helium was used as carrier gas at a flow rate of 1.2 ml/min for 16 min followed by 2 mL/min for 5.75 min. The temperature programme started at 50°C (5 min), then a gradient of 20°C/min up to 270°C was applied and then finally a gradient of 40°/min to 300°C, where it was held stable for 7 min. The mass spectrometry was carried out on a LECO Pegasus BT system (LECO). The acquisition delay was 420 sec. The acquisition rate was 16 spectra/sec. The mass range was 50 – 500 m/z and the extraction frequency was 30 kHz. The ion source was held at 250 °C and the transferline heater temperature was 230 °C. ChromaTOF software (version 5.51) was used for data acquisition. The samples were run in 9 batches, each consisting of 100 samples and a calibration curve. In order to monitor the run a blank, a QC and a standard sample with a known concentration run between every 10 samples. Between every batch the septum and liner on the GC were replaced, the precolumn was cut if necessary and the instrument was tuned. The retention index was determined with ChromaTOF using the reference method function. For every batch a reference file was created. The reference file contained the spectras and approximate retention times of the alkanes from C10 to C30 as determined manually). A reference method was

implemented for every sample in order to determine the exact retention time of the alkanes. Text files with the names and retention times of the alkanes were then exported and converted to the correct format for MSDIAL using an in-house R script. The samples were exported from ChromaTOF using the netCDF format. After this they were converted to abf files using the abfConverter software (Reifycs). Untargeted data processing was carried out using MSDIAL (version 4.7). The minimum peak height was set to an amplitude of 1000, the sigma window value was 0.7 and the EI spectra cut off was 10. The identification was carried out using retention index with the help of the GCMS DB-Public-kovatsRI-VS3 library provided on the MSDIAL webpage. A separate RI file was used for each sample. The RI tolerance was 20 and the m/z tolerance was 0.5 Da. the EI similarly cut off was 70 %. The identification score cut off was 70 % and retention information was used for scoring. Alignment was carried out using the RI with an RI tolerance of 10. The EI similarity tolerance was 60 %. The RI factor was 0.7 and the EI similarity factor was 0.5. The results were exported as peak areas and further processed with excel. In excel the results were normalized using heptadecanoic acid as internal standard and the features with a coefficient of variance of less than 30 % in QC samples were selected. Further filtering was carried out to remove alkanes and duplicate features. The IDs of the features which passed the CV check were further checked using the Golm Metabolome Database.

Predicted functional potential of microbiome We used PICRUSt2 for predicting community functions[65]. We used reference sequences and count abundances of detected ASVs with abundance more than 0.1/100 and prevalence more than 2/100. We used default settings in PICRUSt2 version 2.5.0. Then, for downstream analyses, we focused on enzymes predicted by Enzyme Commission (EC) numbers involved in metabolizing short-chain fatty acids or bile acids, as these metabolites are robustly associated with water content or stool consistency. Relevant ECs were identified based on substrate and product information available on KEGG [66].

Polygenic score for defecation frequency

DNA was extracted from cord blood samples and DNA was sequenced at the National Institute for Health and Welfare and genotyped with Illumina Infinium PsychArray

and Illumina Infinium Global Screening Array at the Estonian Genome Centre as previously described [67]). Genotyped data was pre-phased with Eagle 2.3.5 [68]and imputed with Beagle 4.1 [69]using the Finnish population-specific SISu v2 imputation reference panel. Polygenic score for defecation frequency was calculated based on[70]. PGS was computed using the PRS-CS method [71]with Bayesian regression to estimate posterior variant effect sizes, employing the European 1000 Genomes Project samples as a reference panel. The final PGS for each participant was calculated as a weighted sum of about one million risk allele counts and standardized to a mean of 0 and a standard deviation of 1 using R 3.6.0.

Statistical analyses

All of the analyses were conducted, and figures created in R (version 4.3.2) using packages mia (version 1.10.2), miaViz (version 1.10.2), scater (version 1.30.1) (PMID: 2808876), vegan (version 2.6.2004), nlme (version 3.1.164) and tidyverse (version 2.0.0) [72-75]. The markers of stool consistency were modeled with: 1. water content as continuous variable, 2. stool consistency as a 4-class variable, 3. PGS as continuous variable. All analyses were conducted with all the timepoints as well as separately for each age group.

To describe the longitudinal pattern of reported stool consistency, binary variables of a. loose vs no loose, b. constipation vs no constipation, c. gas and other complaints vs no were created. Subsequently, differences between timepoints were tested with Kruskal-Wallis and Dunn's posthoc test. To inform the interconnectedness of the stool consistency and water content as well as background factors, Spearman correlation, Wilcoxon signed rank test as well as Kruskal-Wallis test were used.

PGS association with water content was tested with Pearson correlation per time point as well as with linear regression with water content as the dependent variable and PGS and current breastfeeding as the independent variables per time point. PGS association with stool consistency was tested with Kruskal-Wallis test.

Taxonomic diversity, i.e. alpha diversity, was estimated with Shannon diversity and observed richness (number of observed species) from non-rarefied, non-transformed ASV-level data. The associations were calculated with Spearman correlation for water content and PGS and Wilcoxon Signed rank test and Kruskal-Wallis test for the stool consistency variables. The longitudinal analyses were conducted with mixed models with alpha diversity as the dependent variable, breastfeeding and water content, PGS or stool consistency as the independent variables and subject as the random effect.

Community composition i.e. beta diversity, was tested with distance-based redundancy analysis (db-RDA) with non-rarefied relative abundance ASV-level data and Bray-Curtis dissimilarity.

Differential abundance testing was conducted with ALDEx2 glm-module for each timepoint [76]. ALDEx2 is known to provide low false-positive rate[77]. However, we also wanted to utilize longitudinal models and ALDEx2 does not allow to random effects in the formula. Thus, the longitudinal data was tested with LinDA with following formula: abundance ~ water content / PGS + current breastfeeding + age, 1 | subject [78] .

Metabolome data was log2-transformed with half of the smallest value as a pseudocount. Associations with water content were tested with Spearman correlation. Associations with stool consistency were tested with Kruskal-Wallis test. The longitudinal data was tested with linear mixed model with metabolite concentration as the dependent variable, stool consistency, PGS or water content as fixed effects, and subject as the random effect.

Mediation analysis was conducted with the mediation (v. 4.5.1) [doi: 10.18637/jss.v059.i05] package in R. Metabolite concentration were the outcome variable, water content or stool consistency the mediator, and breastfeeding the treatment. At 2.5 months, exclusive vs partial breastfeeding was used as the treatment variable since it had more balanced sample sizes. Regarding water content, the analyses were conducted with linear regression models for each timepoint. Regarding stool consistency, binary variable constipation vs no constipation was used with glm-

function and binomial link function. Non-parametric bootstrap was used for estimating confidence intervals. Otherwise, default settings were used in the mediate-function. Interaction between breastfeeding and water content or stool consistency were tested with linear mixed models by including an interaction term as a predictor in addition to age as fixed effects and subject identity as random effect.

All EC abundances from PICRUSt2 were transformed to relative abundances and further log-transformed. The subset of ECs involved in SCFA or BA metabolism were selected for downstream testing. The associations were calculated with Spearman correlation for water content and Kruskal-Wallis test for the stool consistency variables. Dunn's post-hoc test was calculated following significant Kruskal-Wallis test. Linear regression models with EC abundance as dependent variable, water content and breastfeeding as independent variables were built per time point. The longitudinal analyses were conducted with mixed models with EC abundance as the dependent variable, breastfeeding and water content or stool consistency as the independent variables and subject identity as the random effect. P-values were corrected for multiple testing with Benjamini-Hochberg, and we consider $q<0.05$ as statistically significant.

Ethics declarations

Ethical issues have been considered and there is a research permit for the project. FinnBrain has a permit from The Wellbeing Services of the County of South-West Finland ethical committee (ETMK: 57/180/2011), which has approved Cohort profile and research protocol. FinnBrain parents have signed a consent form about their children's participation in research and given permission to use their samples for scientific purposes for the FinnBrain Birth Cohort Study Researchers and the collaborators. Samples went through the laboratory process anonymously with research code to protect participants' privacy.

Competing Interests

P.C.D. is an advisor and holds equity in Cybele, BileOmix, Sirenas and a scientific co-founder, advisor, holds equity and/or received income from Ometa, Enveda, and Arome with prior approval by UC San Diego. P.C.D. also consulted for DSM animal health in 2023. The other authors declare no competing interests.

Acknowledgments

Finnbrain Birth cohort Study (HK) has been funded by Research Council of Finland (grant numbers 253270, 134950), Jane and Aatos Erkko Foundation, as well as Signe and Ane Gyllenberg Foundation. LK was funded by the Research Council of Finland (grant number 308176 and 325292), Yrjö Jahnsson Foundation (6847, 6976), Signe and Ane Gyllenberg Foundation, Finnish State Grants for Clinical Research (P3654), Jalmari and Rauha Ahokas Foundation, and Waterloo Foundation (2110-3601). AKA was supported by Yrjö Jahnsson Foundation, Psychiatry Research Foundation, Emil Aaltonen Foundation, Brain Foundation, Instrumentarium Science Foundation, Signe and Ane Gyllenberg Foundation, Duodecim Finnish Medical Society, Juho Vainio Foundation and the Research Council of Finland (grant number 347640). S.L was supported by the Research Council of Finland (decision number 363417). LL was supported by Research Council of Finland (grant number 330887). AMD has been

funded by the Waterloo foundation and Research Council of Finland (347924). Further support was received by the "Inflammation in human early life: targeting impacts on life-course health" (INITIALISE) consortium funded by the Horizon Europe Program of the European Union under Grant Agreement 101094099 (to MO), and InFLAMES Flagship Programme of the Research Council of Finland (decision number: 337530). AKA and HI were supported by the Signe and Ane Gyllenberg Foundation (grant no. 6273). HI received funding from the Finnish Cultural Foundation (grant no. 00230482), the Finnish Foundation for Cardiovascular Research, the Diabetes Research Foundation, and the Doctoral Program in Clinical Research at the University of Turku.

We thank the FinnBrain Birth Cohort Study participants and personnel, Turku Metabolomics Center for the assistance and resources in the analyses of metabolites. We also thank Paulo Wender P Gomes for his excellent suggestions with the graphics. We like to acknowledge the Maternal and Pediatric Precision In Therapeutics (MPRINT) program, funded by NICHD P50 award (P50HD106463). We would like to acknowledge that the current version of this manuscript is available as a preprint on bioRxiv [79].

Author Contributions

Conception or design of the work: AKA, SL, AMD, TB. Acquisition, data analysis, or interpretation of data: AP, TP, KK, PCD, LK, HK, HI, MK, LL, ML, MO, AMD, AKA, SL, EM. SL and AKA wrote the manuscript. All authors critically reviewed and approved the final manuscript. All authors have approved the submitted revision.

Data and code availability

Due to Finnish national legislation and study participant consents, the individual-level metadata cannot be made available online, but data can potentially be shared with Research Agreement as part of research collaboration. Requests for collaboration can be sent to the Board of the FinnBrain Birth Cohort Study; please contact Linnea Karlsson (linnea.karlsson@utu.fi). The R scripts for data analyses can be found in Zenodo (DOI: 10.5281/zenodo.11222840).

References

1. Lyon L. 'All disease begins in the gut': was Hippocrates right? Brain : a journal of neurology 2018; 141:e20.
2. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Gut 2016; 65:57-62.
3. Tigchelaar EF, Bonder MJ, Jankipersadsing SA, Fu J, Wijmenga C, Zhernakova A. Gut microbiota composition associated with stool consistency. Gut 2016; 65:540-2.
4. Hadizadeh F, Walter S, Belheouane M, Bonfiglio F, Heinsen FA, Andreasson A, et al. Stool frequency is associated with gut microbiota composition. Gut 2017; 66:559-60.
5. Deutsch L, Stres B. The Importance of Objective Stool Classification in Fecal 1H-NMR Metabolomics: Exponential Increase in Stool Crosslinking Is Mirrored in Systemic Inflammation and Associated to Fecal Acetate and Methionine. Metabolites 2021; 11.
6. Yap CX, Henders AK, Alvares GA, Wood DLA, Krause L, Tyson GW, et al. Autism-related dietary preferences mediate autism-gut microbiome associations. Cell 2021; 184:5916-31.e17.
7. Fan Y, Pedersen O. Gut microbiota in human metabolic health and disease. Nature reviews Microbiology 2021; 19:55-71.
8. Lamichhane S, Sen P, Dickens AM, Alves MA, Härkönen T, Honkanen J, et al. Dysregulation of secondary bile acid metabolism precedes islet autoimmunity and type 1 diabetes. Cell reports Medicine 2022; 3:100762.
9. Fromentin S, Forslund SK, Chechi K, Aron-Wisnewsky J, Chakaroun R, Nielsen T, et al. Microbiome and metabolome features of the cardiometabolic disease spectrum. Nature medicine 2022; 28:303-14.
10. Talmor-Barkan Y, Bar N, Shaul AA, Shahaf N, Godneva A, Bussi Y, et al. Metabolomic and microbiome profiling reveals personalized risk factors for coronary artery disease. Nature medicine 2022; 28:295-302.
11. Lamichhane S, Sen P, Dickens AM, Orešič M, Bertram HC. Gut metabolome meets microbiome: A methodological perspective to understand the relationship between host and microbe. Methods 2018; 149:3-12.
12. Zierer J, Jackson MA, Kastenmüller G, Mangino M, Long T, Telenti A, et al. The fecal metabolome as a functional readout of the gut microbiome. Nature genetics 2018; 50:790-5.
13. Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, et al. Best practices for analysing microbiomes. Nature reviews Microbiology 2018; 16:410-22.
14. Roager HM, Hansen LB, Bahl MI, Frandsen HL, Carvalho V, Gøbel RJ, et al. Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. Nature microbiology 2016; 1:16093.
15. Hasan N, Yang H. Factors affecting the composition of the gut microbiota, and its modulation. PeerJ 2019; 7:e7502.
16. Asnicar F, Leeming ER, Dimidi E, Mazidi M, Franks PW, Al Khatib H, et al. Blue poo: impact of gut transit time on the gut microbiome using a novel marker. Gut 2021; 70:1665-74.
17. Procházková N, Falony G, Dragsted LO, Licht TR, Raes J, Roager HM. Advancing human gut microbiota research by considering gut transit time. Gut 2023; 72:180-91.
18. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. Science (New York, NY) 2016; 352:560-4.
19. Schindlbeck NE, Klauser AG, Müller-Lissner SA. [Measurement of colon transit time]. Zeitschrift fur Gastroenterologie 1990; 28:399-404.
20. Degen LP, Phillips SF. How well does stool form reflect colonic transit? Gut 1996; 39:109-13.
21. Jokela R, Ponsero AJ, Dikareva E, Wei X, Kolho KL, Korpela K, et al. Sources of gut microbiota variation in a large longitudinal Finnish infant cohort. EBioMedicine 2023; 94:104695.

22. Motoki N, Inaba Y, Toubou H, Hasegawa K, Shibazaki T, Tsukahara T, et al. Impact of breastfeeding during infancy on functional constipation at 3 years of age: the Japan Environment and Children's Study. International breastfeeding journal 2023; 18:57.
23. Tunc VT, Camurdan AD, Ilhan MN, Sahin F, Beyazova U. Factors associated with defecation patterns in 0-24-month-old children. European journal of pediatrics 2008; 167:1357-62.
24. Müller M, Hermes GDA, Canfora EE, Smidt H, Masclee AAM, Zoetendal EG, et al. Distal colonic transit is linked to gut microbiota diversity and microbial fermentation in humans with slow colonic transit. American journal of physiology Gastrointestinal and liver physiology 2020; 318:G361-g9.
25. Minnebo Y, Delbaere K, Goethals V, Raes J, Van de Wiele T, De Paepe K. Gut microbiota response to in vitro transit time variation is mediated by microbial growth rates, nutrient use efficiency and adaptation to in vivo transit time. Microbiome 2023; 11:240.
26. Harris SC, Devendran S, Méndez-García C, Mythen SM, Wright CL, Fields CJ, et al. Bile acid oxidation by Eggerthella lenta strains C592 and DSM 2243(T). Gut microbes 2018; 9:523-39.
27. Guzior DV, Okros M, Shivel M, Armwald B, Bridges C, Fu Y, et al. Bile salt hydrolase acyltransferase activity expands bile acid diversity. Nature 2024; 626:852-8.
28. Kwiatkowska M, Gołębiewski M, Sikora M, Rycharska E, Krogulska A. The oral cavity and intestinal microbiome in children with functional constipation. Sci Rep 2024; 14:8283.
29. Martino C, Dilmore AH, Burcham ZM, Metcalf JL, Jeste D, Knight R. Microbiota succession throughout life from the cradle to the grave. Nature reviews Microbiology 2022; 20:707-20.
30. Mohanty I, Allaband C, Mannochio-Russo H, El Abiead Y, Hagey LR, Knight R, et al. The changing metabolic landscape of bile acids - keys to metabolism and immune regulation. Nature reviews Gastroenterology & hepatology 2024.
31. Guzior DV, Quinn RA. Review: microbial transformations of human bile acids. Microbiome 2021; 9:140.
32. Johnson-Martínez JP, Diener C, Levine AE, Wilmanski T, Suskind DL, Ralevski A, et al. Aberrant bowel movement frequencies coincide with increased microbe-derived blood metabolites associated with reduced organ function. Cell reports Medicine 2024; 5:101646.
33. Hu S, Ding Q, Zhang W, Kang M, Ma J, Zhao L. Gut microbial beta-glucuronidase: a vital regulator in female estrogen metabolism. Gut microbes 2023; 15:2236749.
34. Lamichhane S, Yde CC, Jensen HM, Morovic W, Hibberd AA, Ouwehand AC, et al. Metabolic Fate of (13)C-Labeled Polydextrose and Impact on the Gut Microbiome: A Triple-Phase Study in a Colon Simulator. J Proteome Res 2018; 17:1041-53.
35. Kråkström M, Dickens AM, Alves MA, Forssten SD, Ouwehand AC, Hyötyläinen T, et al. Dynamics of the Lipidome in a Colon Simulator. Metabolites 2023; 13.
36. El Oufir L, Flourié B, Bruley des Varannes S, Barry JL, Cloarec D, Bornet F, et al. Relations between transit time, fermentation products, and hydrogen consuming flora in healthy humans. Gut 1996; 38:870-7.
37. Lewis SJ, Heaton KW. Increasing butyrate concentration in the distal colon by accelerating intestinal transit. Gut 1997; 41:245-51.
38. Gabriele S, Sacco R, Altieri L, Neri C, Urbani A, Bravaccio C, et al. Slow intestinal transit contributes to elevate urinary p-cresol level in Italian autistic children. Autism research : official journal of the International Society for Autism Research 2016; 9:752-9.
39. Zhang WX, Zhang Y, Qin G, Li KM, Wei W, Li SY, et al. Altered profiles of fecal metabolites correlate with visceral hypersensitivity and may contribute to symptom severity of diarrhea-predominant irritable bowel syndrome. World journal of gastroenterology 2019; 25:6416-29.
40. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature 2011; 472:57-63.
41. Nemet I, Saha PP, Gupta N, Zhu W, Romano KA, Skye SM, et al. A Cardiovascular Disease-Linked Gut Microbial Metabolite Acts via Adrenergic Receptors. Cell 2020; 180:862-77.e22.

42. Al Hinai EA, Kullamethee P, Rowland IR, Swann J, Walton GE, Commane DM. Modelling the role of microbial p-cresol in colorectal genotoxicity. Gut microbes 2019; 10:398-411.
43. Rimal B, Collins SL, Tanes CE, Rocha ER, Granda MA, Solanki S, et al. Bile salt hydrolase catalyses formation of amine-conjugated bile acids. Nature 2024; 626:859-63.
44. Wang P, Wang R, Zhao W, Zhao Y, Wang D, Zhao S, et al. Gut microbiota-derived 4-hydroxyphenylacetic acid from resveratrol supplementation prevents obesity through SIRT1 signaling activation. Gut microbes 2025; 17:2446391.
45. Hung LY, Parathan P, Boonma P, Wu Q, Wang Y, Haag A, et al. Antibiotic exposure postweaning disrupts the neurochemistry and function of enteric neurons mediating colonic motor activity. American journal of physiology Gastrointestinal and liver physiology 2020; 318:G1042-g53.
46. Griffiths JA, Yoo BB, Thuy-Boun P, Cantu VJ, Weldon KC, Challis C, et al. Peripheral neuronal activation shapes the microbiome and alters gut physiology. Cell reports 2024; 43:113953.
47. Caputi V, Marsilio I, Filpa V, Cerantola S, Orso G, Bistoletti M, et al. Antibiotic-induced dysbiosis of the microbiota impairs gut neuromuscular function in juvenile mice. British journal of pharmacology 2017; 174:3623-39.
48. De Vadder F, Grasset E, Mannerås Holm L, Karsenty G, Macpherson AJ, Olofsson LE, et al. Gut microbiota regulates maturation of the adult enteric nervous system via enteric serotonin networks. Proceedings of the National Academy of Sciences of the United States of America 2018; 115:6458-63.
49. de Moraes JG, Motta ME, Beltrão MF, Salviano TL, da Silva GA. Fecal Microbiota and Diet of Children with Chronic Constipation. International journal of pediatrics 2016; 2016:6787269.
50. Hofman D, Kudla U, Miqdady M, Nguyen TVH, Morán-Ramos S, Vandenplas Y. Faecal Microbiota in Infants and Young Children with Functional Gastrointestinal Disorders: A Systematic Review. Nutrients 2022; 14.
51. Yuan Y, Lu Y, Zhang Z, Cheng W, Yan K, Zheng Y, et al. Characteristics of the Cajal interstitial cells and intestinal microbiota in children with refractory constipation. Microbial pathogenesis 2023; 184:106373.
52. Sharif H, Devadason D, Abrehart N, Stevenson R, Marciani L. Imaging Measurement of Whole Gut Transit Time in Paediatric and Adult Functional Gastrointestinal Disorders: A Systematic Review and Narrative Synthesis. Diagnostics (Basel, Switzerland) 2019; 9.
53. Popescu M, Mutalib M. Bowel transit studies in children: evidence base, role and practicalities. Frontline gastroenterology 2022; 13:152-9.
54. Salvatore S, Abkari A, Cai W, Catto-Smith A, Cruchet S, Gottrand F, et al. Review shows that parental reassurance and nutritional advice help to optimise the management of functional gastrointestinal disorders in infants. Acta paediatrica (Oslo, Norway : 1992) 2018; 107:1512-20.
55. Bowles A, Keane J, Ernest T, Clapham D, Tuleu C. Specific aspects of gastro-intestinal transit in children for drug delivery design. International journal of pharmaceutics 2010; 395:37-43.
56. Tran DL, Sintusek P. Functional constipation in children: What physicians should know. World journal of gastroenterology 2023; 29:1261-88.
57. Karlsson L, Tolvanen M, Scheinin NM, Uusitupa HM, Korja R, Ekholm E, et al. Cohort Profile: The FinnBrain Birth Cohort Study (FinnBrain). International journal of epidemiology 2018; 47:15-6j.
58. Mirzayi C, Renson A, Zohra F, Elsafoury S, Geistlinger L, Kasselman LJ, et al. Reporting guidelines for human microbiome research: the STORMS checklist. Nature medicine 2021; 27:1885-92.
59. Aatsinki A-K, Lamichhane S, Isokääntä H, Sen P, Kråkström M, Alves MA, et al. Dynamics of Gut Metabolome and Microbiome Maturation during Early Life. medRxiv 2023:2023.05.29.23290441.
60. Rintala A, Riikonen I, Toivonen A, Pietilä S, Munukka E, Pursiheimo JP, et al. Early fecal microbiota composition in children who later develop celiac disease and associated autoimmunity. Scandinavian journal of gastroenterology 2018; 53:403-9.

61. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic acids research 2013; 41:D590-6.
62. Kråkström M, Dickens AM, Alves MA, Forssten SD, Ouwehand AC, Hyötyläinen T, et al. Dynamics of the Lipidome in a Colon Simulator. Metabolites 2023; 13:355.
63. Huovinen V, Aatsinki AK, Kataja EL, Munukka E, Keskitalo A, Lamichhane S, et al. Infant gut microbiota and negative and fear reactivity. Dev Psychopathol 2023:1-16.
64. Lamichhane S, Kemppainen E, Trošt K, Siljander H, Hyöty H, Ilonen J, et al. Circulating metabolites in progression to islet autoimmunity and type 1 diabetes. Diabetologia 2019; 62:2287-97.
65. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, et al. PICRUSt2 for prediction of metagenome functions. Nature biotechnology 2020; 38:685-8.
66. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic acids research 2017; 45:D353-d61.
67. Shimizu N, Kawase C, Nakazono H, Hemmi H, Shimatake H, Aoki T. A novel RNA splicing mutation in Japanese patients with Wilson disease. Biochemical and biophysical research communications 1995; 217:16-20.
68. Loh PR, Danecek P, Palamara PF, Fuchsberger C, Y AR, H KF, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nature genetics 2016; 48:1443-8.
69. Browning BL, Browning SR. Genotype Imputation with Millions of Reference Samples. American journal of human genetics 2016; 98:116-26.
70. Bonfiglio F, Liu X, Smillie C, Pandit A, Kurilshikov A, Bacigalupe R, et al. GWAS of stool frequency provides insights into gastrointestinal motility and irritable bowel syndrome. Cell genomics 2021; 1:None.
71. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nature communications 2019; 10:1776.
72. Ernst F SS, Borman T, Lahti L. mia: Microbiome analysis. R package version 1.12.0, https://github.com/microbiome/mia., 2024.
73. McCarthy DJ, Campbell KR, Lun AT, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics (Oxford, England) 2017; 33:1179-86.
74. Dixon P. VEGAN, a package of R functions for community ecology. Journal of Vegetation Science 2003; 14:927-30.
75. Pinheiro J BD, R Core Team (2023). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-164, https://CRAN.R-project.org/package=nlme.
76. Fernandes AD, Macklaim JM, Linn TG, Reid G, Gloor GB. ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. PloS one 2013; 8:e67019.
77. Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al. Microbiome differential abundance methods produce different results across 38 datasets. Nature communications 2022; 13:342.
78. Li G, Yang L, Chen J, Zhang X. Robust Differential Abundance Analysis of Microbiome Sequencing Data. Genes 2023; 14.
79. Aatsinki A-K, Isokääntä H, Kråkström M, Lahti L, Lukkarinen M, Munukka E, et al. Fecal microbiota and metabolite composition associates with stool consistency in young children. bioRxiv 2024:2024.06.05.597641.