Article

Gut physiology and environment explain variations in human gut microbiome composition and metabolism

Check for updates

Nicola Procházková [1], Martin F. Laursen [2], Giorgia La Barbera [1], Eirini Tsekitsidi[3], Malte S. Jørgensen [1], Morten A. Rasmussen[4,5], Jeroen Raes [6,7], Tine R. Licht [2], Lars O. Dragsted [1] & Henrik M. Roager [1] ✉

The human gut microbiome is highly personal. However, the contribution of gut physiology and environment to variations in the gut microbiome remains understudied. Here we performed an observational trial using multi-omics to profile microbiome composition and metabolism in 61 healthy adults for 9 consecutive days. We assessed day-to-day changes in gut environmental factors and measured whole-gut and segmental intestinal transit time and pH using a wireless motility capsule in a subset of 50 individuals. We observed substantial daily fluctuations, with intra-individual variations in gut microbiome and metabolism associated with changes in stool moisture and faecal pH, and inter-individual variations accounted for by whole-gut and segmental transit times and pH. Metabolites derived from microbial carbohydrate fermentation correlated negatively with the gut passage time and pH, while proteolytic metabolites and breath methane showed a positive correlation. Finally, we identified associations between segmental transit time/pH and coffee-, diet-, host- and microbial-derived metabolites. Our work suggests that gut physiology and environment are key to understanding the individuality of the human gut microbial composition and metabolism.

Diet influences the gut microbial composition and metabolism[1,2]. However, even with identical dietary intake, the gut microbiome varies[3,4], suggesting that other factors in the gut contribute to these variations. Gut transit time accounts for substantial variation in the microbiome composition of healthy populations[5–8], with longer transit time associated with increased microbial protein degradation and methane production[9]. While short-chain fatty acids (SCFAs), the main microbial products of saccharolysis, are typically considered beneficial[10], microbial proteolysis results in metabolites associated with poor health outcomes, including hydrogen sulfide, ammonia, branched-chain fatty acids (BCFAs), p-cresol, indole and phenylacetate[11,12].

Changes in pH along the gut are also linked to gut microbial composition and metabolism[13]. The presence of SCFAs and other organic acids lowers colonic pH[13], inhibiting bacteria sensitive to acidic environments[14]. Yet, little is known about how the gut environment, determined by physiological factors such as transit time and luminal pH, associates with diet–host–microbiota metabolism. Understanding these factors could be crucial for future personalized dietary microbiome-based strategies.

[1]Department of Nutrition, Exercise and Sports, University of Copenhagen, Frederiksberg, Denmark. [2]National Food Institute, Technical University of Denmark, Kgs Lyngby, Denmark. [3]Laboratory of Hygiene, Social & Preventive Medicine and Medical Statistics, School of Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece. [4]Department of Food Science, University of Copenhagen, Frederiksberg, Denmark. [5]Copenhagen Studies on Asthma in Childhood (COPSAC), Herlev-Gentofte Hospital, University of Copenhagen, Gentofte, Denmark. [6]Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven, Belgium. [7]Center for Microbiology, VIB, Leuven, Belgium. ✉e-mail: hero@nexs.ku.dk

Transit time can be assessed using wireless motility capsules (SmartPills), which measure intraluminal pH, temperature and pressure along the gastrointestinal tract that allow for the determination of segmental transit times. SmartPills are one of the standard clinical methods for directly assessing transit time, along with radio-opaque markers and scintigraphy[15], and have been used in previous microbiome studies[16,17].

In this Article, we conducted a 9 day observational study with 61 healthy volunteers and assessed whole-gut and segmental transit time and pH by SmartPills. We also collected data on bowel habits, dietary intake, and breath hydrogen and methane levels and profiled urine and faecal samples using multi-omics techniques. This allowed us to follow and link inter-individual and day-to-day changes in the gut environment, gut microbiota and microbiota-derived metabolites.

Results

Study design and participant characteristics

We enrolled 61 healthy participants (aged 39 ± 13.5 years, with body mass index (BMI) of 23.6 ± 2.8 kg m⁻²; Table 1 and Extended Data Fig. 1) and asked them to maintain their habitual lifestyle and diet for 9 consecutive days (Fig. 1a). The study included two visits (day 2 and day 9) where fasting blood glucose, insulin and C-peptide, as well as breath hydrogen and methane, were measured (Table 1). On the first visit, participants were given a breakfast that accounted for 25% of their daily energy needs (Supplementary Table 1) to provide a standardized meal before a subset of the volunteers ($n = 50$) ingested a wireless motility capsule (SmartPill) to measure whole-gut and segmental transit time and pH[18]. While previous investigations used granola bars (SmartBar) before the monitoring[19,20], we used a complex meal similar to a recent study[21] to investigate diet–microbiota interactions. Postprandial breath and urine samples were obtained as indicated in Fig. 1a. The participants recorded daily 24 h dietary records (days 1–8) using the myfood24 nutrition platform (https://www.myfood24.org); noted daily bowel habits including defecation time, stool consistency assessed by the Bristol Stool Form Scale (BSS)[22] and stool frequency (number of bowel movements per day); and collected daily urine (the first morning sample) and faecal samples (the first bowel movement). The study population had normal bowel habits (Table 1) with a median BSS of type 4 and 1 bowel movement per day. Transit time was also estimated by a self-administered sweet-corn transit time test[23] on days 3 and 5 (corn TT). We measured faecal water content (indication of stool moisture, a proxy marker of transit time[23]), pH and microbial load in all collected faecal samples ($n = 484$). All collected urine samples (daily spot and postprandial samples, $n = 1,154$) and a subset of faecal samples ($n = 170$) were profiled by untargeted liquid chromatography–mass spectrometry (LC-MS)-metabolomics to obtain urine and faecal metabolomes. Finally, we obtained the gut microbiome composition via 16S ribosomal RNA (rRNA) gene sequencing of a subset of faecal samples ($n = 362$) and assessed both relative microbiome profiles (RMPs) and quantitative microbiome profiles (QMPs) after adjusting for microbial load[24].

Gut environment stability varies for each individual

Daily sampling allowed us to evaluate the fluctuations in gut environmental factors, faecal and urine metabolomes, gut microbiomes and diets within and between healthy adults over time (Supplementary Fig. 1). First, we observed varying degrees of day-to-day fluctuations within individuals for faecal pH (coefficient of intra-individual variation (CV_{Intra} 0.3–8.1%), BSS (0–57.8%), stool frequency (0–73.1%), stool moisture (2.2–24%) and microbial load (7.6–72.7%) (Fig. 1b and Supplementary Table 2), suggesting that some individuals have more stable gut environments than others. Most of the gut environmental factors varied within individuals over the 9 days, whereas faecal pH remained relatively stable (Fig. 1b and Supplementary Table 3a). Participant ID

Table 1 | Participants' characteristics (N=61)

	Mean±s.d./Median (25th–75th percentiles)	Range
Sex, male/female	18/43	-
Age (years)	39±13.5	20–66
BMI (kg m⁻²)	23.6±2.8	17.6–29.5
Fasting glucose (mmol l⁻¹)ᵃ	5.1 (4.9–5.4)	4.4–6.9
Fasting insulin (mmol l⁻¹)ᵃ	32.5 (24.2–51.7)	14.7–132.0
Fasting C-pep (pmol l⁻¹)ᵃ	407 (321–520)	186–771
Dietary intakeᵃ		
Total energy intake (kcal d⁻¹)	2256±605	1,276–5,091
Carbohydrate (g d⁻¹)	231.4±76.9	51–521
Carbohydrate (E%)	41.1±8.7	14.2–69.4
Protein (g d⁻¹)	86.1±29.5	32–209
Protein (E%)	15.5±4.4	5.5–35.6
Fat (g d⁻¹)	99.0±34.4	34–256
Fat (E%)	39.4±8.2	16.5–62.3
Fibre (g d⁻¹)	24.0±10.3	3–62
Fibre intake (g per 1,000 kcal d⁻¹)	10.8±3.9	2–23
Gut environmental factorsᵃ		
Stool consistency, BSS	4 (3–5)	1–7
Stool frequency (n per day)	1 (1–2)	0–5
Stool moisture (%)	73 (69–77)	28–93
Faecal pH	6.8 (6.3–7.0)	5.4–7.3
Faecal SCFAs (μmol per g of faeces)ᵃ		
Acetate	16.11 (8.29–26.06)	0.85–76.07
Propionate	2.49 (1.56–3.93)	0.01–33.80
Butyrate	1.43 (0.78–2.09)	0.04–5.96
Valerate	1.20 (0.89–1.69)	0.33–4.10
Caproate	0.54 (0.13–1.07)	0.01–6.88
Faecal BCFAs (μmol per g of faeces)ᵃ		
2-Methylbutyrate	0.54 (0.39–0.71)	0.07–3.24
Isovalerate	0.41 (0.28–0.55)	0.08–2.23
Isobutyrate	0.29 (0.22–0.37)	0.07–1.39
Breathᵃ		
Fasting hydrogen (p.p.m.)	6.5 (4.0–12.0)	0.5–51
Fasting methane (p.p.m.)	1.0 (0–18)	0–67.5

E%, energy per cent; p.p.m., parts per million. ᵃMean of all records/measurements.

explained a significant proportion of variation for day-to-day fluctuations in all of the gut environmental factors (Supplementary Table 3b), indicating that the stability of the gut environment is to some extent personal.

Next, we performed a permutational multivariate analysis of variance (PERMANOVA) on the QMPs, urine and faecal metabolomes and found that the individual explained more than 50% of the inter-individual variations in all three cases (Fig. 1c). The sampling day explained on average 6.7% of the urine metabolome variation but did not explain day-to-day variations in the gut microbiome and faecal metabolome. (Fig. 1c). By inspecting the β-diversities of individual microbiome and metabolome profiles, we observed that some individuals showed less variation over the study period than others (Extended Data Fig. 2).
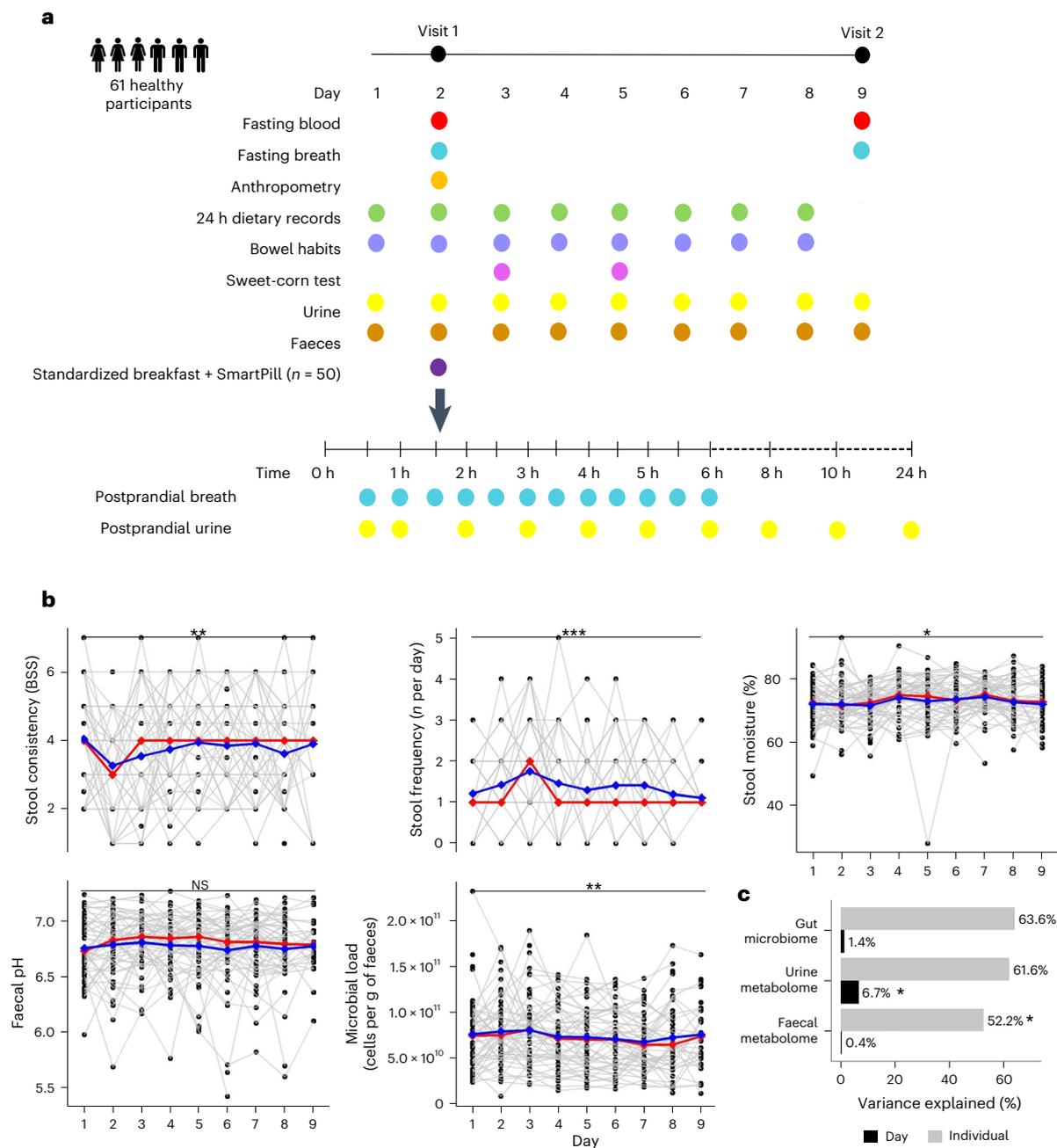
Fig. 1 | PRIMA study design and variations in gut environmental factors, gut microbiome and metabolomes. a, PRIMA study design. The study included two site visits, at which fasting blood and breath samples were taken. At visit 1, anthropometric measurements were attained, and all participants were given a standardized breakfast; a subset of 50 volunteers ingested SmartPills immediately after. Postprandial breath hydrogen and methane were measured every 30 min for 6 h, and postprandial urine was collected at 0.5 h and every hour until 24 h as indicated. On days 3 and 5, participants performed a sweet-corn test to measure WGTT. In addition, daily 24 h dietary records (days 1–8), records of bowel habits (stool consistency, stool frequency and time of defecation) and daily urine and faecal samples were obtained. Solid line indicates sample collection on site and dashed line sample collection at home. **b**, Inter- and intra-individual variations in the gut environmental factors over the 9 consecutive days. The red and blue lines represent median and mean values, respectively. Grey lines represent intra-individual fluctuations over time. Asterisks indicate the statistical significance of mixed-effect models accounting for repeated measures (two-sided ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$; NS, not significant; see Supplementary Table 3 for details; no adjustment for multiple testing was applied). **c**, Percentage of variation explained by individual and study day in the gut microbiome and urine and faecal metabolomes based on PERMANOVA tests (two-sided *$P < 0.05$).

Stool moisture and pH explain daily gut microbiome fluctuations

To explore what drives the intra-individual fluctuation in the metabolomes and the microbiome, we performed distance-based redundancy analysis (db-RDA). We considered daily dietary macronutrients and fibres, as well as the gut environmental factors. None of the dietary components explained intra-individual fluctuations in the gut microbiome or metabolomes. By contrast, stool moisture, faecal pH, BSS and time of defecation markedly affected the gut microbiome (QMP, genus level; Fig. 2a), explaining 3.5%, 2.5%, 2% and 1.3% of the variations, respectively. Similar results were observed using the RMP data (Extended Data Fig. 3a) and in previous studies[6,25]. Notably, these

explanatory factors are proxies for gut transit time, suggesting that day-to-day variations in transit time are reflected in the gut microbiome variation.

Stool moisture and faecal pH further explained 3.1% and 3%, respectively, of the intra-individual variation in urine metabolomes, despite subtle day-to-day fluctuations (Fig. 2b). This suggests that even small changes in the colonic water content and pH may be associated with the host–microbiota metabolism. However, these observations could also be influenced by daily diet variations. Gut environmental factors did not contribute to the intra-individual fluctuations in the faecal metabolomes (Extended Data Fig. 3b). It should be noted that faecal metabolome data were derived only from three consecutive days, and stool moisture still tended ($P = 0.081$) to have an effect.

Transit times and pH vary between individuals

SmartPills were used to obtain whole-gut transit time (WGTT), gastric emptying time (GET), small-bowel transit time (SBT), colonic transit time (CTT) and intestinal transit time (ITT; SBT + CTT), as well as pH throughout the gastrointestinal tract (GIT) (Extended Data Fig. 3c). In 8 individuals, the capsule was retained in the stomach for over 8 h, a common event reported in other studies[19,26]. Therefore, GET and WGTT values from these participants were excluded from our analyses. In addition, we could not determine CTT and WGTT in one participant due to a signal loss.

The median values of transit time were as follows: GET, 4.8 h (range 3.1–6.2 h); WGTT, 23.3 h (12.4–72.3 h); CTT, 13.6 h (2.1–63.5 h); and SBT, 5.1 h (2.5–10.3 h), in agreement with previously reported data on healthy populations[27]. For comparison, the corn TT showed a median of 23.6 h (10.8–109.7 h) at day 3 and 19.7 h (12.0–84.5 h) at day 5. Furthermore, we found a strong correlation between the two corn TT measurements (Spearman correlation coefficient (SCC) = 0.8, $P < 0.001$) suggesting consistency within individuals. The median of the mean corn TT across the two days was 21.7 h (11.7–97.1 h) (Fig. 2c), similar to the WGTT obtained by the SmartPill. However, we did not observe any correlation between the WGTT and corn TT (Extended Data Fig. 4a), indicating that despite providing similar results on average, individually, these two methods showed different results.

When exploring the relationships between segmental transit times, corn TT, gut environmental factors and participant characteristics (Extended Data Fig. 4a), we found that the transit times recorded by both methods were negatively correlated to BSS, as also reported previously[23,28]. We also observed that women had significantly longer CTT compared with men, while there was no effect of menstruation status among the women (Extended Data Fig. 4b).

Large inter-individual variations in the gastrointestinal segmental pH were also observed (Fig. 2d) with the following median pH values in the upper GIT: the stomach (0.9, range 0.5–4.9), duodenum (6.1, 5.0–7.2) and small intestine (7.4, 6.4–8.2). pH in the proximal colon was slightly acidic (6.3, 5.3–7.0) followed by a gradual increase in the distal colon (6.9, 5.0–8.2) and sigmoid colon (7.2, 5.6–8.6). Interestingly, a small decrease in pH was observed from the sigmoid colon to the rectum (7.0, 5.7–8.6) and also in the faecal pH (6.9, 6.6–7.3), indicating that acidifying processes occur after entry into the rectum.

### CTT and pH contribute to gut microbiome variations

To quantify how participant characteristics, clinical variables, diet and gut environmental and physiological factors explain inter-individual variations in the gut microbiome and metabolomes, we performed a db-RDA using data derived from faecal and 24 h urine collections on day 2 from all participants ($n = 61$; Supplementary Table 4). Moreover, we performed the same analysis with whole-gut and segmental transit times and pH derived from the SmartPills on day 2 ($n = 50$). Stool moisture and distal colon pH were key factors associated with inter-individual variation in QMPs (Fig. 2e), accounting for 5.5% and 5% of the variation, respectively, on day 2 and also other days (Supplementary Table 4). Unlike previously reported data from larger cohorts[5], BSS did not explain a significant proportion of the variation in QMP in this population. WGTT, CTT, corn TT and faecal pH explained 9.1%, 6.2%, 4.9% and 5.4%, respectively, of the inter-individual variations in the 24 h urine metabolome, in comparison to age, which explained 4.5% of the variation (Fig. 2f). These contributions were consistent when testing against the urine metabolomes on different days (Supplementary Table 4). By contrast, segmental transit time did not contribute to the inter-individual variation in the faecal metabolomes, whereas pH in the distal colon and fibre intake showed the largest effects explaining 6.8% and 5.9% of the variations, respectively; however, this was not significant after adjusting for multiple testing (Supplementary Table 4).

We also tested the effect of menstruation during the study period for women (non-menstruating, $n = 30$; menstruating, $n = 13$), which showed effect sizes of 3.7% (gut microbiome) and 3.6% (urine metabolome), however without statistical significance. Considering the notable effect size of age on urine metabolome and a significant age difference between the two groups of women ($P = 0.01$), age might contribute to these observed effects. Our results emphasize that the personal gut environment contributes considerably to the inter-individual differences in the gut microbiota and urinary metabolic profiles.

Individual gut microbiota and metabolite profiles are dynamic

We next assessed intra-individual fluctuations in microbial-derived metabolites including breath hydrogen and methane, faecal SCFAs (acetate, propionate, butyrate, valerate and caproate) and BCFAs (isobutyrate, isovalerate and 2-methylbutyrate; Table 1), as well as 16 other microbial-derived metabolites detected in faeces and urine, including the proteolytic markers, p-cresol sulfate (PCS), phenylacetyl-glutamine (PAGln) and indoxyl sulfate. Substantial day-to-day fluctuations were observed (Fig. 3a and Supplementary Fig. 2). Breath methane and hydrogen had a median CV_{Intra} of 141% and 47%, respectively, with a moderate positive correlation between the two time points for both gases (hydrogen, SCC = 0.42, $P < 0.001$; methane, SCC = 0.66, $P < 0.001$). Faecal concentrations of the SCFAs and BCFAs fluctuated considerably from day to day (median CV_{Intra} ranging from 26% to 40%) with valerate varying the least and acetate the most. Similarly, the relative abundances of the proteolytic markers varied substantially from day to day with a median CV_{Intra} of 26%, 42% and 39% for PAGln, indoxyl sulfate and PCS, respectively. These findings suggest that microbial-derived metabolites in breath, faeces and urine fluctuate from day to day on a habitual diet.

Fig. 2 | Intra- and inter-individual variations in gut microbiome and urine metabolome explained by gut environment. a,b, Contributions of dietary and gut environmental factors on intra-individual variations in gut microbiome (QMP, all days) (**a**) and urine metabolome (all days, all features) (**b**). **c**, Boxplots showing segmental and WGTT measured by the SmartPill ($n = 50$) at day 2 and mean transit time of sweet corn ($n = 61$, day 3 and day 5) with each dot representing an individual. **d**, Boxplots showing pH throughout the gastrointestinal tract measured by the SmartPill ($n = 50$) and in faeces measured by pH meter at day 2 ($n = 61$) with each dot representing an individual. **e,f**, Contributions of clinical variables, dietary components, gut environmental and physiological factors and participant characteristics to inter-individual variations in the gut microbiome (QMP, sample closest to the capsule body exit) (**e**) and urine metabolome (24 h, day 2, all features) (**f**). Panels **a**, **b**, **e** and **f** were quantified by db-RDA with permutation tests using Bray–Curtis distances. Effect sizes are plotted. The asterisks indicate statistical significance after adjustments for multiple testing (*$q < 0.1$). See Extended Data Fig. 3a,b for RMPs and faecal metabolome. Boxplot centre in **c** and **d** represents median, and box represents interquartile range (IQR). Whiskers extend to most extreme data point <1.5 IQR. SB, small bowel; CH_4, breath methane; H_2, breath hydrogen.
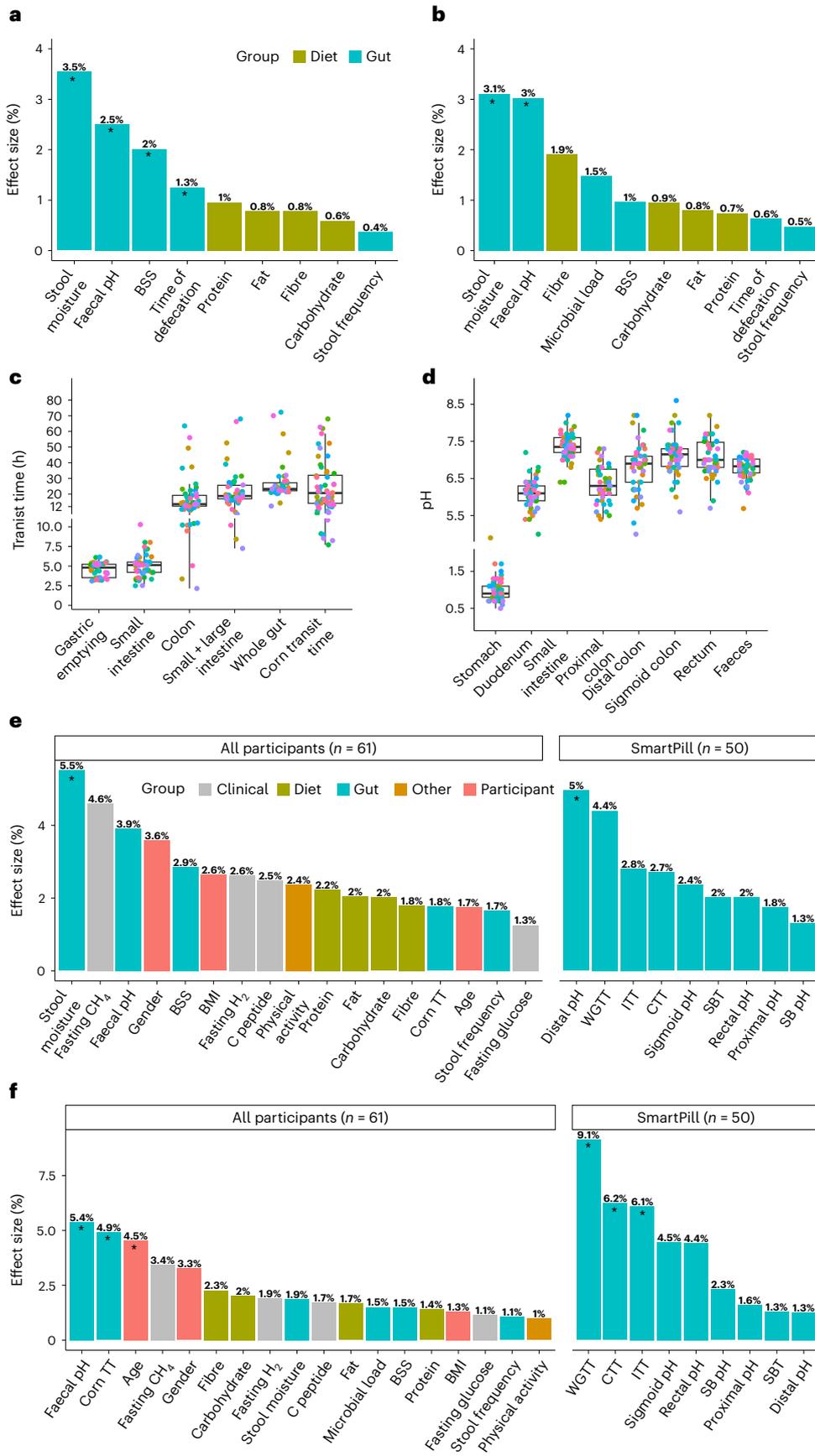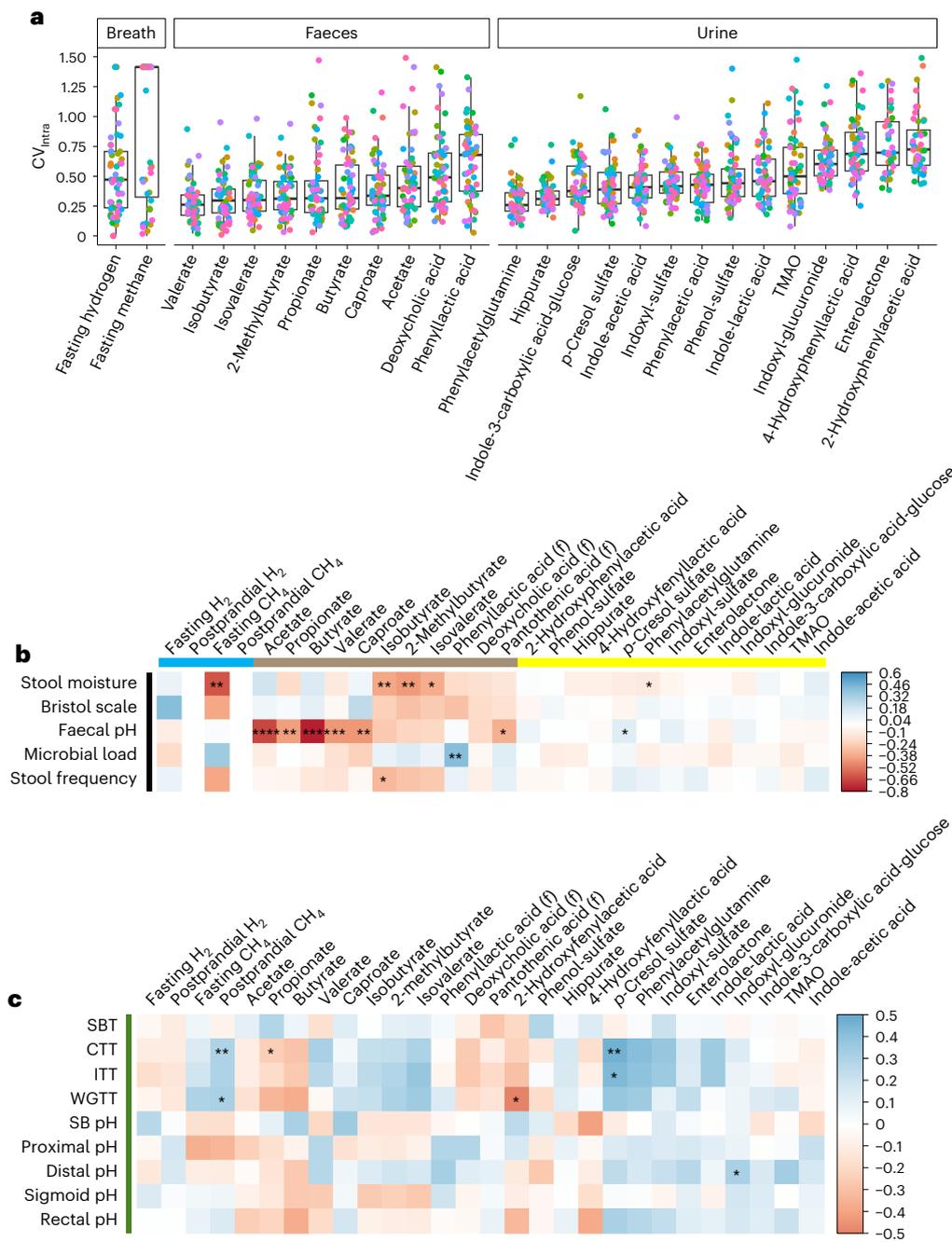
Fig. 3 | Fluctuations in microbial metabolites and their correlations to gut physiology and environment. a, Intra-individual fluctuations in microbial metabolites measured in breath, faeces and urine. Boxplots show coefficients of intra-individual variations. Each dot represents an individual ($n = 61$). Boxplot centre represents median, and box represents IQR. Whiskers extend to most extreme data point <1.5 IQR. **b,c**, Correlations between microbial metabolites and gut physiology and environmental factors as assessed by repeated measures correlation coefficient (**b**) or the Spearman correlation coefficient (**c**).

The asterisks indicate statistical significance after adjustment for multiple testing (****$q < 0.001$, ***$q < 0.01$, **$q < 0.05$, *$q < 0.1$). Blue, brown and yellow bars indicate breath, faecal and urine metabolites, respectively. The black bar in **b** indicates repeated measure correlations where daily values for each variable have been used (Extended Data Fig. 5), whereas the green bar in **c** indicates analysis based on data collected on day 2. Postprandial hydrogen and methane were only measured at one time point and therefore were not included in the repeated measure analysis. (f), faecal; TMAO, trimethylamine N-oxide.

Next, we used repeated measures (Fig. 3b, Extended Data Fig. 5 and Supplementary Table 5) and Spearman correlation analysis (Fig. 3c) to find links between specific microbial metabolites and the gut environment along with diet. Faecal SCFAs were negatively correlated to faecal pH with butyrate showing the strongest correlation ($r = −0.77$, $q < 0.001$) in line with previous human studies[29]. Moreover, higher faecal propionate (SCC = −0.25, $q < 0.1$) was linked to shorter CTT with a similar tendency observed for faecal butyrate (SCC = −0.29, $P < 0.05$,

$q = 0.3$). Faecal butyrate also tended to negatively correlate with rectal pH (SCC = −0.37, $P < 0.05$, $q = 0.2$) but not with pH in other segments of the colon, suggesting that butyrate production may contribute to the reduced pH observed in the rectum and faeces.

Proteolytic markers including urinary PAGln and faecal BCFAs were negatively correlated to stool moisture, and urinary PCS correlated positively with faecal pH (PAGln, $r = −0.12$, $q < 0.1$; isobutyrate, $r = −0.39$, $q < 0.05$; isovalerate, $r = −0.37$, $q < 0.1$; 2-methylbutyrate, $r = −0.43$,

$q < 0.05$; PCS, $r = 0.12$, $q < 0.1$). Similarly, higher urinary levels of PCS were associated with longer CTT and ITT (SCC = 0.48, SCC = 0.44, respectively, $q < 0.05$), with similar tendencies observed for PAGln and indole-lactic acid. Furthermore, breath methane was linked to lower stool moisture and longer CTT. It is worth noting that none of this was shown for SBT, indicating that CTT determines the abundance of these metabolites and supports the hypothesis that longer passage through the colon is linked to microbial proteolysis possibly due to the depletion of substrates for saccharolytic fermentation[9,30]. Urinary indoxyl-glucuronide was positively associated with pH in the distal colon (SCC = 0.33, $q < 0.1$), and a similar trend was found between urinary PCS and rectal pH (SCC = 0.32, $P < 0.05$). These metabolites did not correlate to pH in the small intestine and the proximal colon, indicating a higher contribution of microbial proteolysis to pH in the distal gut compared with the proximal gut. In summary, CTT and colonic pH, but not SBT and small-intestinal pH, are associated with levels of several microbial metabolites in breath, faeces and urine. In addition, we found several associations between microbial metabolites and dietary components (Supplementary Table 5) with notable inverse correlations between the intake of dietary fibres and faecal BCFAs, urinary PCS and urinary PAGln, respectively.

Faecal and urine metabolomes are linked to transit time and pH

To explore unknown metabolic features related to gut physiology, we used untargeted metabolomics to profile the urine and faecal metabolomes. We applied univariate and multivariate statistical models on all molecular features identified in urine and faeces. We first used sparse partial least squares (SPLS) models on the SmartPill-derived data and urine metabolomes from 24 h postprandial urine collected on day 2 and faecal metabolomes collected closest to the SmartPill egestion. We then performed linear regression models on the same data and further investigated features selected by both models (446 unique features; Supplementary Table 6).

Several metabolic features in urine and faeces were associated with whole-gut and segmental transit time and pH (Fig. 4a,b). To investigate these features in further detail, the corresponding samples were analysed by tandem mass spectrometry (MS^2) and by matching with authentic standards when available, resulting in the identification of 33 metabolites (Supplementary Tables 7 and 8).

Apart from urinary levels of PCS and PAGln, several other urinary metabolites derived from the breakdown of aromatic amino acids tryptophan and tyrosine by gut microbes were found to be linked with gut transit time and faecal pH. Specifically, 5-hydroxy-2-oxindole sulfate, 3-hydroxy-2-oxindole sulfate and 4-hydroxybenzoic acid sulfate were associated with longer WGTT/CTT, while 3-hydroxy-2-oxindole glucuronide correlated with higher faecal pH. By contrast, faecal tryptophan was negatively linked to faecal pH. In addition, higher faecal proline and urinary picolinoylglycine levels were linked with increased faecal and rectal pH, respectively.

Several dicarboxylic acids in faeces, pimelic, suberic and sebacic acids were positively associated with WGTT and CTT. By contrast, faecal glutaric acid and pipecolic acid were negatively correlated with WGTT/CTT and sigmoid, rectal and faecal pH. Pipecolic acid is highly abundant in plants; however, it can also be produced by the gut microbiota from lysine[31]. Furthermore, higher urinary levels of citric acid were positively associated with pH in the proximal colon.

Moreover, faecal levels of 2-oxindole-3-acetic acid, previously linked to the New Nordic Diet and Mediterranean diet[32,33], were negatively associated with WGTT, CTT and faecal pH. Similarly, faecal pantothenic and nicotinic acids were negatively associated with CTT and faecal pH, respectively. In addition, dihydroferulic acid glucuronide and argininic acid in urine were negatively associated with rectal pH, while p-hydroxyphenyllactic acid in faeces was negatively linked to faecal pH.

4-Hydroxyhippuric acid and several urinary markers of coffee intake, including 1-methyluric acid, 1-methylxanthine, 1,3-dimethyluric acid, 1,7-dimethyluric acid and 1,3,9-trimethyluric acid, were negatively associated with small-intestinal pH. 1-Methylxanthine and 1,3,9-trimethyluric acid in faeces were also negatively associated with WGTT or faecal pH, suggesting a link between coffee consumption and gut function. In addition, a positive correlation was observed between rectal pH and urinary 4-methylcatechol sulfate, a metabolite of quercetin found in plant-based foods[34]. Urinary taurine and faecal cholic acid were also positively associated with small-intestinal pH, supporting the role of bile acids in neutralizing the acidic chyme coming from the stomach[35].

Finally, urinary pseudouridine, a primary constituent of RNA, was found to be inversely associated with CTT and sigmoid colon pH, in line with our previous work[9]. Pseudouridine was also found in faeces and showed a similar inverse relationship with faecal pH, as did deoxy-xanthosine and xanthine. This suggests a link between increased cell turnover and lower colonic pH.

Altogether, by using untargeted LC-MS metabolomics, we identified several host-, microbial- and food-derived metabolites associated with WGTT, CTT and pH in the distal part of the colon emphasizing an interplay between diet, the gut environment, the host and the microbiota.

Microbial alpha diversity is linked to long passage

To explore potential links between the identified metabolites and the gut microbiota, Spearman correlation analysis was performed (Fig. 5). Strong positive correlations between microbial alpha diversity measures and microbial proteolysis, CTT and ITT were found. On the contrary, alpha diversity correlated negatively with stool moisture and microbial saccharolysis.

Products of microbial proteolysis and dicarboxylic acids were positively correlated with the absolute abundances of several bacterial genera including *Intestimonas*, *Flavonifractor*, *Eubacterium*, *Lachnospira*, *Clostridium*, *Oscillibacter*, *Alistipes*, *Dialister* and *Akkermansia*. The same genera negatively correlated with faecal levels of tryptophan, oxindole-3-acetic acid and various coffee-derived metabolites. Not surprisingly, these genera were also positively associated with longer ITT and CTT and higher faecal pH, and negatively associated with stool moisture and/or BSS. Conversely, SCFAs-producing genera including *Agathobacter*, *Faecalibacterium* and *Blautia*[36,37], along with lactate-producing *Streptococcus*, were all positively associated with faecal nicotinic acid, pantothenic acid and the coffee-derived metabolites. Notably, *Oscillibacter*, *Alistipes* and *Akkermansia* have repeatedly been found elevated in samples linked to longer transit time and/ or constipation[8,9,28], whereas butyrate-producing genera including *Faecalibacterium* and *Agathobacter* have been associated with shorter transit time[28,38]. In summary, these observations highlight the interdependency between gut bacteria, metabolites and gut physiology.

Discussion

Gut transit time and pH are important determinants of gut microbiota composition and metabolism[7]. Here we showed substantial variation in whole-gut and segmental transit time, along with luminal pH among healthy individuals. These variations explained differences in microbiome composition and host–microbiota co-metabolism. As pH and transit time influence microbial growth and enzyme activities[39], these factors could play a key role in shaping the gut microbial composition and metabolism along the GIT as well as microbiome responses to foods. A recent study confirmed that microbiome and metabolome compositions differ along the GIT[40]. Future studies with sampling along the GIT combined with measurements of regional pH and transit time are needed to ultimately disentangle this. Our study emphasizes that person-specific differences in the luminal pH may pose challenges for studies using pH-sensitive ingestible devices[40,41]. It is worth noting that
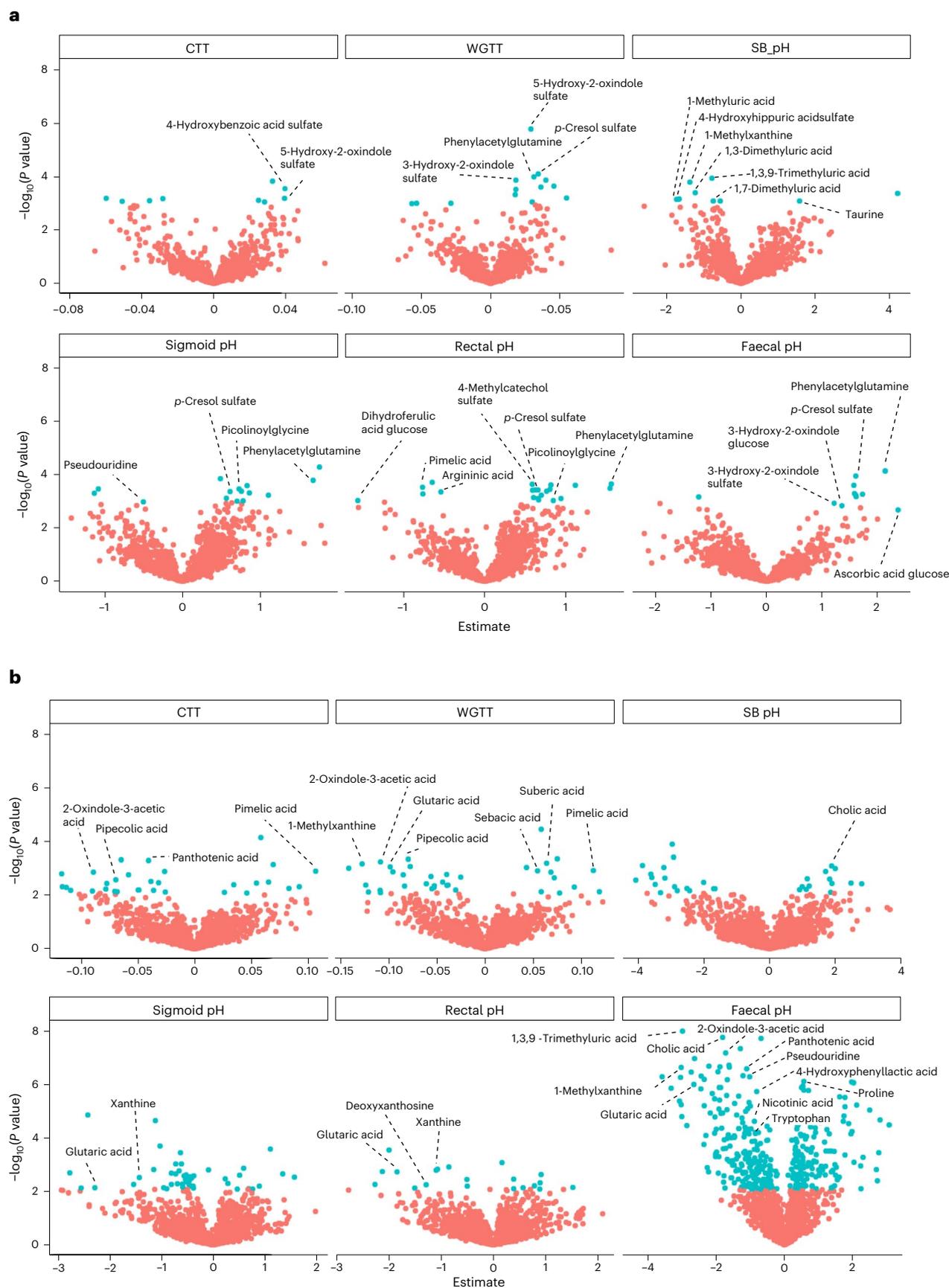
Fig. 4 | Metabolites identified via LC-MS untargeted metabolomics associated with segmental transit time and pH. a,b, Volcano plots derived from regression models where each dot represents a metabolic feature with blue representing statistically significant associations (FDR-adjusted P < 0.1) in urine (**a**) and faeces (**b**). The x axis shows the regression coefficient values (estimate) indicating either positive or negative associations and the y axis represents the $-\log_{10}$ transformation of the P value.

Fig. 5 | Associations between metabolites, bacterial genera and gut environmental factors. Hydrogen and methane were measured in breath, and SCFAs and BCFAs were measured in faeces; other metabolites as indicated in brackets. (b), breath; (f), faecal metabolite; (u), urine metabolite. The colour gradient shows the Spearman correlation coefficient, and the asterisks indicate statistical significance after adjustment for multiple testing (***$q < 0.001$, **$q < 0.01$, *$q < 0.05$).

repeated measurements in human studies are needed as we observed substantial daily fluctuations in microbial-derived metabolites.

We identified several metabolites associated with longer transit time that have been reported to be elevated in patient groups with constipation[42–44]. Among these, dicarboxylic acids including pimelic acid were associated with longer ITT and/or higher pH, and bacteria consistently associated with constipation[8,9,28]. Pimelic acid, possibly originating from microbial fatty acid metabolism[45,46], has been found at higher faecal levels in patients with chronic kidney disease[47] and colorectal cancer[48], often associated with constipation[49,50]. A recent study showed an increased abundance of dicarboxylic acids towards the distal gut[40], and the authors speculated that it could be due to the catabolism of host epithelial cells. Whether longer ITT might be associated with increased epithelial cell turnover and shedding needs further research.

A negative association between daily fibre intake and several proteolytic markers was also observed. As dietary fibres can regulate microbial tryptophan metabolism[51], availability of fibre in the colon may affect microbial protein fermentation associated with negative health outcomes[11,12]. Further research is needed to understand these mechanisms and to explore the relationship between gut physiology and microbiome under controlled diets, possibly involving dietitians or providing whole diets.

Despite its limited cohort size, our study shows significant associations between intestinal segmental transit time and pH with intra- and inter-individual differences in the gut microbiome composition and metabolism in a healthy population. Potential limitations to consider are the choice of breakfast made before the SmartPill measurement; given that past validation studies and their normative data rely on specific meal/nutrient combinations[52,53], any deviations from these could likely influence gut motility and transit time. Furthermore, the sweet-corn test is not a validated tool to assess WGTT despite being cost effective. Finally, the introduction of corn and the meals during the first visit constitute small dietary changes, but they could possibly have impacted gut physiology. While this study included a rather homogenous group of healthy volunteers, it provides valuable insights into longitudinal variations in gut microbial metabolism and pH over more than 1 week. Our results highlight the important role of transit time and pH for the gut microbiome composition and levels of microbial-derived metabolites, emphasizing the importance of considering gut physiology and environment in human microbiome studies. This may be key for understanding the healthy gut microbiome and for disentangling personal microbiome responses to foods and other lifestyle factors.

Methods

Study participants

A 9 day human observational trial (PRIMA, toward Personalized dietary Recommendations based on the Interaction between diet, Microbiome and Abiotic conditions in the gut) among healthy participants was conducted at the Department of Nutrition, Exercise and Sports at the University of Copenhagen in Denmark from April to December 2021. The research protocol was approved by the Municipal Ethical Committee of the Capital Region of Denmark (H-20074067), and all participants provided written informed consent to participate according to Case Report (CARE) guidelines and in compliance with the principles of the Declaration of Helsinki. The study was registered at ClinicalTrails.gov (ID NCT04804319).

Out of the initially anticipated 85 individuals, 63 healthy participants living in Denmark were enrolled, and 61 completed the study (43 women and 18 men; Extended Data Fig. 1). The two drop-outs were excluded due to illness and antibiotic administration. Among the 61 study participants, a subset of 50 volunteers (37 women and 13 men) underwent wireless motility capsule monitoring at visit 1 as anticipated. Volunteers were compensated with gift cards (500 or 800 DKK) but received no direct financial compensation. The criteria for inclusion in the study specified for participants who were healthy by self-report (did not suffer from inflammatory bowel syndrome, small-intestinal overgrowth, inflammatory bowel disease, chronic or infectious disease, diabetes or cancer), ages 18–75 years old with a BMI between 18.5 and 29.9 kg m^{-2}, with no intake of medication, except for mild antidepressants and contraceptive pills. Intake of antibiotics, diarrhoea inhibitors and laxatives 1 month before the trial was not allowed. Furthermore, pregnant or lactating women were not included in the trial.

The PRIMA study was an explorative study. The primary outcome was to investigate associations between faecal pH and gut microbial saccharolytic/proteolytic metabolism (assessed by targeted and untargeted metabolomics). The secondary outcomes included to explore relationships between the gut environmental factors (small-intestinal and colonic pH and transit time assessed by wireless motility capsules and various transit time proxy markers measured in faeces) and gut microbiome and metabolome assessed by 16S rRNA sequencing and metabolomics.

Experimental design and sample collection

Seven days before the study, the participants were asked not to consume any sweet corn as two self-administered sweet-corn tests to evaluate the WGTT were part of the study. Before both visits, the participants were asked to abstain from alcohol intake, smoking and strenuous exercise.

The participants were asked to maintain their habitual diet and register their food intake online via the Myfood24 tool (myfood24.org) with nutritional values based on the Danish food composition database FRIDA version 4.1 (frida.fooddata.dk) for eight consecutive days during the study. During the trial, the participants collected daily stool samples (first bowel movement whenever possible), stored the samples in their domestic freezers and transported them to the laboratory while being kept cold. Moreover, the participants self-reported daily their defecation patterns including time of defecations, stool consistency assessed by the BSS and stool frequency, intake of dietary supplements and medication (limited to pain killers for a few participants), and their gastrointestinal symptoms. The gastrointestinal symptoms were assessed based on a 10 cm visual analogue scale (0, no symptoms; 10, the most severe symptoms) in regard to stomach ache, bloating, constipation, diarrhoea and overall comfort. Women were asked to note down whether they had menstruation during the study period (yes/no). Furthermore, the participants collected seven daily spot morning urine samples (days 1, 2, 4, 5, 6, 7, 8; the first morning sample) and two 24 h urine samples (days 2–3 and days 8–9) during the study period. The collected urine samples were stored in the participants' domestic freezers, transported to the study site in a cooling bag and stored at −20 °C overnight. After thawing at 5 °C, aliquots of 1 ml were taken and stored at −80 °C until further use. In addition, the participants consumed 100 g of sweet corn before their evening meal on days 3 and 5 and recorded the time of the corn egestion[23].

At both visits (days 2 and 9), fasting blood and breath samples were collected. During the first visit, anthropometric measurements (height, body weight and BMI) were obtained. Furthermore, the first visit also included a standardized meal test for all participants (n = 61). The test meal consisted of rye bread (with butter and jam), a boiled egg, a portion of natural yoghurt along with nuts, walnuts, blueberries and a glass of water (100 ml) with 250 mg of dissolved paracetamol (Table S1), which was used as a marker of postprandial gastric emptying of liquids[54]. The meal portion size was calculated as 25% of the daily energy demand of each participant based on the Harris–Benedict equation[22]. Postprandial urine samples (at 30 min, 60 min, 120 min, 180 min, 240 min, 300 min and 360 min and between 6 and 8 h, between 8 and 10 h and between 10 and 24 h) and postprandial breath exhalations (at 30 min, 60 min, 90 min, 120 min, 150 min, 180 min, 210 min, 240 min, 270 min, 300 min, 330 min and 360 min) were collected. A subset of participants (n = 50) ingested a SmartPill capsule immediately after

the meal with a bit of additional water if needed. All participants drank 150 ml of water at 2 h and 4 h after the meal, respectively. At 6 h, all participants received a sandwich and 500 ml of water and left the study site.

SmartPill data collection and analysis

The SmartPill capsule is a single-use wireless gastrointestinal capsule (26.8 mm × 13 mm) that transmits data on luminal pH, temperature and pressure to a portable receiver, which was worn by the participants from ingestion to egestion and thereafter returned to the study personnel. The capsule measures a pH range of 1–9, with an accuracy of ±0.5 pH units, pressure at a range of 0–350 mmHg (±5 mmHg) and temperature ranging between 20 °C and 40 °C (±1 °C)[23]. Upon receiving the portable receiver, the study personnel downloaded the raw data from the receiver to the manufacturer's software via a docking station (Motility GI v 3.1). Intestinal segmental transit times were determined based on landmark changes in the pH values as follows: gastric emptying was defined as the time point with an abrupt increase of ≥3 pH units indicating passage from the stomach into the duodenum. The passage from the small intestine into the ileocaecal junction was defined as the first time point with a decrease of at least one pH unit. The body exit of the capsule was defined as the time point with a decrease in temperature and/or a loss of data. The time of capsule residence in each of the gastrointestinal segments corresponds to GET, small-intestinal transit time, CTT and combined WGTT. Regional pH and pressure profiles were also obtained, and the median values were determined. The segmental transit time and pH values in the colon were further segmented into proximal, distal and recto-sigmoid, respectively. The proximal colon pH and transit time were estimated as median values of the first 32.3% of the total CTT, the distal colon pH were median values of the next 32.6% and the recto-sigmoid pH were median values of the last 35.4%; this was based on previously reported data, which determined the percentages of total CTT according to the location of radio-opaque markers (visualized by X-rays) in the different segments of the colon[16]. In addition, the median pH value measured during the last 10 min before the capsule egestion was registered as rectal pH.

Dietary records

Detailed 24 h weighted food intakes were recorded for 8 consecutive days by the participants via the online Myfood24 tool (myfood24.org) with nutritional values based on the Danish food composition database FRIDA version 4.1 (frida.fooddata.dk). The collected data included information about the intake of macronutrients (carbohydrate, protein, fat) and dietary fibre (AOACFIB), in addition to information about more than 80 nutrients. Under-reporting was identified by calculating the reported caloric intake divided by the average daily energy demand for each person with a cut-off value of 0.8 (ref. [55]). Accordingly, approximately 25% of the daily dietary records were under-reported, and the data were removed in the subsequent analyses in this study (this essentially affected 10 participants who under-reported more than 4 out of the 8 days, while the other participants occasionally under-reported daily intakes). By contrast, no over-reporters (cut-off >2.5) were detected. The total dietary profiles (all macro- and micronutrients available in Myfood24) were used in the principal component analysis, whereas macronutrient and fibre intake were used in the redundancy analyses. The daily intake was used for intra-individual analysis, whereas mean intake across the 8 days was used for the inter-individual analysis.

Breath exhalations measurements

Fasting and postprandial levels of hydrogen and methane were measured in all breath samples by the M.E.C. Lactotest 202 Xtend device (M.E.C. R&D sprl).

Biochemical analysis of blood

Blood samples were immediately put on ice upon collection until they were centrifuged for precipitation of blood cells and stored at −80 °C.

Glucose was measured in plasma samples by using Pentra ABX 400 (HORIBA ABX) with a detection limit of 0.11 mmol l^{-1}. Serum insulin and C-peptide levels were measured by using Immulite 2000 XPi (Siemens Healthcare Diagnostics) with the detection limit of 14.4 pmol l^{-1} and 27 pmol l^{-1}, respectively. Before the analyses, both instruments' performances were validated using external and internal insulin, c-peptide and glucose controls. Three participants arrived for the second visit in a postprandial state; the blood was collected and analysed accordingly, but the glucose, insulin and c-peptide values were not included in the data analysis.

Faecal measurements

Upon receipt, faecal samples were stored at −20 °C overnight, thawed and homogenized in sterile water with a sample-to-water ratio of 1:1 (w/v) (faecal slurry). Subsequently, pH was measured in the faecal slurry using a digital pH meter (Mettler Toledo). The homogenized samples were subsequently aliquoted to cryotubes and stored at −80 °C until further analyses. Stool moisture was determined by evaporating the water of one aliquot (approximately 1 ml) using a vacuum concentrator (Speed-Vac, Christ RVC 2-25) and by calculating the faecal weight difference before and after evaporation.

Faecal SCFAs and BCFAs were quantified by LC-MS in samples collected between day 2 and day 5 (n = 170) as previously described[32]. In brief, the samples were thawed, mixed with ethanol and purified by filtration (0.2 μm filter). Subsequently, the samples were derivatized with 3-nitrophenylhydrazine, and labelled internal SCFA standards were added. Dilution series of external SCFA standards were spiked with internal SCFA standards, and all derivatized samples were analysed on ultra-performance liquid chromatography (UPLC)-quadrupole time-of-flight mass spectrometry (QTOF-MS) (Synapt G2, Waters) in negative ionization mode (cone voltage 3.0 kV) with an ACQUITY BEH C18 guard column (2.1 × 5 mm, 1.7 μm, Waters) coupled to an ACQUITY BEH C18 column (2.1 × 100 mm, 1.7 μm, Waters) and with collision energy of 6.0 eV. The faecal concentrations of SCFAs and BCFAs were determined using vendor software (Quanlynx, Waters).

Bacterial load in faeces was determined using approximately 500 μl of frozen faecal slurry (238–816 mg) and diluting it 400,000 times in physiological saline (8.5 g l^{-1} NaCl; VWR International). Next, 1 ml of the microbial cell suspension obtained was stained with 1 μl SYBR Green I (1:100 dilution in dimethylsulfoxide; shaded during 20 min incubation at 37 °C; 10,000 concentrate, Thermo Fisher Scientific). The flow cytometry analysis of the bacterial cells present in the suspension was performed using a Cytoflex flow cytometer (CytoFLEX 3; Beckman) as previously described (Supplementary Fig. 3)[24]. The final microbial load was calculated per gram of faeces.

Microbiome profiling

DNA was extracted in random order from the faecal slurries (n = 484) using DNeasy PowerLyzer PowerSoil kit (Qiagen, 12855-100), and the V3 region of the 16S rRNA gene was PCR amplified using 0.2 μl Phusion High-Fidelity DNA polymerase (ThermoFisher Scientific, F-553L), 4 μl high-fidelity buffer, 0.4 μl dNTP (10 mM of each base), 1 μM forward primer (primer binding upstream); 5′-A-adapter-TCAG-barcode-CCTACGGGAGGCAGCAG-3′), 1 μM reverse primer (primer binding reverse); 5′-trP1-adapter-ATTACCGCGGCTGCTGG-3′) and 0.05–5 ng faecal DNA in 20 μl total reaction volume. Both primers (TAG Copenhagen A/S) were linked to sequencing adaptors, and the forward primer additionally contained a unique 10 bp barcode (Ion Xpress Barcode Adapters) for each sample. The PCR program consisted of an initial denaturation for 30 s at 98 °C, followed by 24 cycles of 98 °C for 15 s and 72 °C for 30 s and a final extension at 72 °C for 5 min. The PCR products were purified by the HighPrep PCR clean-up system (AC-60500 Magbio) according to the manufacturer's protocol. The resulting DNA concentrations were determined by Qubit HS assay and libraries constructed with

mixing equimolar amounts of each PCR product. Partial 16S rRNA gene sequencing was performed on an Ion S5 System (ThermoFisher Scientific) using OneTouch 2 Ion5: 520/530 kit-OT2 400 bp and an Ion 520 Chip. The raw data were pre-processed into an amplicon sequence variant (ASV) table using our in-house pipeline[56] based on the DADA2 algorithm and settings recommended for IonTorrent reads[57], with taxonomy assigned to the ASVs using the Ribosomal Database Project (RDP, v18). The resulting ASV table, taxonomy and ASV sequences were merged into a phyloseq object for further analysis. For quantitative microbiome profiling analyses, the relative abundances derived from the pre-processed 16S rRNA sequencing analysis were adjusted for the bacterial loads as previously published[58]. In brief, samples with <10,000 reads were removed (n = 362) and downsized to even sampling depth, defined as the ratio between sample size (16S rRNA gene copy number corrected sequencing depth) and bacterial load. 16S rRNA gene copy numbers were retrieved from the rRNA operon copy number database rrnDB73 (ref. [59]). The copy-number-corrected sequencing depth of each sample was rarefied to the level necessary to equate the minimum observed sampling depth in the cohort while assuring a minimum number of 10,000 reads in each sample and optimizing the chosen sampling depth to exclude as few samples as possible. In case of no copy number correction, an average copy number of 3.88 was used[6].

Metabolic profiling

Preparation of urine and faecal samples. Untargeted urine and faecal metabolomics were performed as previously published[32]. All urine samples were thawed on ice, centrifuged at 10,000 g at 4 °C for 2 min and transferred to a new tube to remove solid particles. The urine samples were kept cold on ice during preparation. Samples were randomized and pipetted into 15 plates (96-well). All urine samples from the same individual were placed on the same 96-well plate. Subsequently, they were diluted to 1:5 with an internal standard mixture (L-adenine-8-^{13}C (Cambridge Isotope Lab), L-phenyl-d5-alanine-2,3,3-d3 (Cambridge Isotope Lab), caffeic acid ^{13}C$_3$ (Toronto Research Chemicals), caffeine ^{13}C$_3$ (Toronto Research Chemicals), L-tyrosine ^{13}C$_9$ (Sigma Aldrich), *para*-aminobenzoic acid (Sigma Aldrich), L-tryptophan-(indole-d$_5$) (Sigma Aldrich), hippuric acid-[^{13}C$_6$] (IsoSciences), cortisone-d8 (Sigma Aldrich) and glycocholic acid-[^{2}H$_4$] (IsoSciences)). Quality control (QC) samples were obtained by mixing 20 µl of each urine sample in each plate (plate pools) and by mixing 20 µl of each plate pool to create the global pool. The QC samples, blank assays (0.1% formic acid) and mixtures of known standards (including 33 microbial-derived compounds) were included in each plate. The plates were sealed and stored at 4 °C until analysis (24 h maximum, otherwise stored at −80 °C). If the plate was frozen and thawed again before analysis, the plate was gently mixed by vortex stirring for 30 min immediately before analysis.

Faecal homogenates collected between day 2 and day 5 (n = 170) were thawed at room temperature for 30 min and vortexed. Approximately 50 ±5 mg (≈50 µl) of the homogenates were mixed with 96% ethanol and internal standard mixture (L-adenine-8-^{13}C (Cambridge Isotope Lab), L-phenyl-d5-alanine-2,3,3-d3 (Cambridge Isotope Lab), caffeic acid ^{13}C$_3$ (Toronto Research Chemicals), caffeine ^{13}C$_3$ (Toronto Research Chemicals), L-tyrosine ^{13}C$_9$ (Sigma Aldrich), lysophosphatidylcholine (17:1d$_7$) (Avanti Polar Lipids), L-tryptophan-(indole-d$_5$) (Sigma Aldrich), hippuric acid-[^{13}C$_6$] (IsoSciences), cortisone-d8 (Sigma Aldrich) and glycocholic acid-[^{2}H$_4$] (IsoSciences)) resulting in a 1:60 dilution. The diluted samples were vortexed for 30 s and subsequently mixed at 60 °C for 2 min in a Thermo mixer at 1,400 r.p.m., before being centrifuged at 20,000 g (Eppendorf centrifuge 5417R) at 4 °C for 2 min. The supernatants were filtered through a 0.2 µm filter, and 200 µl of each faecal suspension was transferred to a 96-well plate, evaporated using a cooled vacuum centrifuge and re-dissolved in 200 µl 0.1% formic acid before the UPLC-MS. All faecal samples from the same individual were placed on the same 96-well plate, and QC samples were prepared in the same way as for the urine samples. In addition, each 96-well plate contained blank assays (96% ethanol) and mixtures of known standards (including 33 microbial-derived compounds).

UPLC-electrospray ionization-QTOF-MS analysis

Both urine and faecal samples were profiled by UPLC coupled with a QTOF mass spectrometer equipped with electrospray ionization (Synapt G2, Waters) in both positive and negative ionization modes[32]. Blank samples (0.1% formic acid), assay blanks, standard mixtures and QC samples were injected regularly to evaluate LC-MS system stability, possible contamination and/or loss of metabolites. The injected samples (5 µl) were separated on a reversed-phase column (ACQUITY HSS T3 C18 column, 2.1 × 100 mm, 1.8 µm) coupled with a pre-column (ACQUITY VanGuard HSS T3 C18 column, 2.1 × 5 mm, 1.8 µm). The mobile phases consisted of 0.1% formic acid in water (solvent A) and 0.1% formic acid in 70:30 acetonitrile/methanol (solvent B). The duration of the analytical run was 7 min with the following flow rate: start condition (0.5 ml min^{-1}), 1 min (0.5 ml min^{-1}), 2 min (0.6 ml min^{-1}), 3 min (0.7 ml min^{-1}), 4 min (0.8 ml min^{-1}), 4.5 min (1.0 ml min^{-1}), 6.4 min (1.1 ml min^{-1}), 6.6 min (1.0 ml min^{-1}), 6.8 min (0.5 ml min^{-1}), 7.0 min (0.5 ml min^{-1}), and the following gradient: start condition (5% B), 1 min (8% B), 2 min (15% B), 3 min (40% B), 4 min (70% B), 4.5 min (100% B), 6.6 min (5% B) and 7 min (5% B). Mass spectrometry data were acquired in full scan mode with a scan range of 50–1,000 mass/charge (m/z). Data-dependent acquisition was performed on the top three most abundant ions on QC samples (only urine) to provide MS^2 data. Electrospray settings were the following: the cone voltage was 2.5 kV and 3.2 kV; the collision energy was 6.0 and 4.0 eV; and the temperature of the ion source and desolvation nitrogen gas temperature were 120 °C and 400 °C for positive and negative ionization mode, respectively.

Metabolomics data processing

The raw data obtained by UPLC-MS were converted to mzML format by publicly available msConvert (ProteoWizard Toolkit)[60]. The converted data were pre-processed using the open-source R package XCMS (v3.18) using the centWave algorithm (requiring three consecutive scans with an intensity of over ten counts)[61]. The pre-processing steps included noise filtering, peak picking, retention time alignment and feature grouping across samples, and filling of missing features, which were done separately for the urine and faecal samples (and for positive and negative mode), respectively. The detailed pre-processing parameter settings can be found in Supplementary Table 9. Noise filtering settings included that features should be detected in a minimum of 10% of all samples. Features with a retention time below 0.5 min or above 6.8 min were excluded. Data tables were generated comprising mass-to-charge ratio (m/z), retention time (rt) and intensity (peak area) for each variable in every sample. Each detected peak is represented by a feature defined by a rt and a m/z. The obtained data were corrected for within- and between-batch intensity drift using the locally estimated scatterplot smoothing correction method[62]. The processed data were normalized by the probabilistic quotient normalization[63] method to correct for variations in urine and faecal concentrations within and between batches. Upon analyses of 15 plates with urine samples, QC samples clustered closely together in the principal component analysis score plots, confirming a stable UPLC system during the course of analysis with the exception of two plates in the negative mode and one plate in the positive mode, which had to be removed from further statistical analyses (Supplementary Fig. 4).

Moreover, features with high variability after normalization across the pooled QC samples were filtered out (coefficient of variation, CV% >50%). Finally, the CAMERA package[64] (v1.52) was used to group features together based on retention time (tolerance = 0.1 s) and to annotate possible adducts and isotopes. Upon pre-processing, 641 and 651 molecular features were detected in the urine in positive and negative modes, respectively, whereas 453 and 445 molecular features were detected in faeces in positive and negative modes,

respectively. MzMine 3 (ref. 65) and MassLynx (Waters) were used for data visualization.

Metabolite identification and structure elucidation

MS^2 analyses were performed by an ultra-high performance LC system coupled to a Vion IMS QTOF mass spectrometer (Waters) for obtaining spectra with higher mass accuracy. The samples were separated on a reversed-phase column (ACQUITY HSS T3 C18 column, 2.1×100 mm, 1.8 µm) coupled with a pre-column (ACQUITY VanGuard HSS T3 C18 column, 2.1×5 mm, 1.8 µm) at a temperature of 50 °C. The mobile phases consisted of 0.1% formic acid in water (solvent A), methanol (solvent B), 0.1% formic acid in 70:30 acetonitrile/methanol (solvent C) and isopropanol (solvent D). The duration of the analytical run was 10 min with the following flow rate: start condition (0.4 ml min⁻¹), 0.75 min (0.4 ml min⁻¹), 6 min (0.5 ml min⁻¹), 6.5 min (0.5 ml min⁻¹), 8 min (0.6 ml min⁻¹), 8.1 min (0.4 ml min⁻¹), 9 min (0.4 mlmin⁻¹), 10 min (0.4 ml min⁻¹), and the following gradient: start condition (100% A), 0.75 min (100% A), 6 min (100% B), 6.5 min (70% C, 30% D), 8 min (70% C, 30% D), 8.1 min (70% C, 30% D), 9 min (100% A) and 10 min (100% A). Full scan acquisition was performed on selected urine samples with a scan range of 50–1,500 m/z. Data-dependent acquisition was performed on a selected list of precursors at three different collision dissociation energies, 10, 30 and 50 eV.

Mass spectra were manually interpreted, and metabolites were identified by matching the precursor ion and fragmentation patterns with databases such as Human Metabolome Database (https://hmdb.ca/), Metlin (https://metlin.scripps.edu/), mzCloud (https://www.mzcloud.org/), combinatorial database of bile acid conjugates[66] (http://melolab.org/smilib/) and an in-house database (https://gitlab.com/stanstrup_R_packages/mscurate and https://gitlab.com/stanstrup_R_packages/xcms-annotator). In addition, we used several software annotations including GNPS (v30)[67] (https://gnps.ucsd.edu/), microbeMASST (v2024.08.26)[68] (https://masst.gnps2.org/microbe-masst/) and SIRIUS (v6.0.5)[69] (https://bio.informatik.uni-jena.de/software/sirius/), without obtaining additional plausible matches. Furthermore, authentic standards were run together with the samples with the highest intensity on the same batch and platform. If needed, the authentic standards were sulfated or glucuronidated with either biomimetic synthesis[70] or chemical synthesis[32]. The identification level of metabolites that were identified was classified according to Sumner et al.[71] as level I (confirmed by matching to a standard with two orthogonal measures (rt, m/z), level II (matching MS^2 fragmentation to a spectral library), level III (compound classification) or level IV (unknown)[25]. See Supplementary Tables 7 and 8 for further details. 3-Hydroxy-2-oxindole, 5-hydroxyoxindole, 2-picolinic acid, 4-methylcatechol, xanthine, 2-oxindole-3-acetic acid, pantothenic acid, nicotinic acid, tryptophan, sebacic acid, pipecolic acid, glutaric acid, citric acid, psedouridine, taurine, 1,3-dimethyluric acid, suberic acid and 1,3,7-trimethyluric acid were purchased from Sigma-Aldrich. 4-Hydroxyhippuric acid, 1-methylxanthine and 1-methyluric acid were purchased from Toronto Research Chemicals.

Statistical analysis

Statistical analyses were conducted in R (v 4.2). The area under the curves for hydrogen and methane concentrations during the postprandial period was calculated using the trapezoid rule in GraphPad Prism (v 9.2.0). The normality of data was assessed with the Gaussian distribution and Shapiro–Wilk test procedure.

Mixed-effects linear regression models were used to examine the day-to-day fluctuations and inter-individual variation in gut environmental factors using data from all 9 days. The models were generated using the *lme4* R package (v 1.1-31) as lmer (gut environmental factor - factor(day) + (1 | Participant ID)); moreover, ranova function from the *lmerTest* package (v 3.1-3) was used to perform the random effects-likelihood ratio tests to infer whether

Participant ID significantly contributes to explaining the variation in the gut environmental factors. P value of < 0.05 was considered statistically significant. Coefficients of intra-individual variation were calculated as $CV_{Intra} = (s.d._{intra}/Mean_{intra}) \times 100$ where mean and s.d. were based on all measurements from a single individual over the 9 days.

Gut microbiome beta-diversity analysis using Bray–Curtis distances as well as metabolome and diet beta-diversity analyses using Euclidian distances were performed with the *phyloseq* package (v 1.42.0) and PERMANOVA tests by adonis2 function from the *vegan* package (v 2.6) with 999 permutations and strata = Participant ID when testing the day-to-day fluctuations.

Single time point correlations were calculated using standard Spearman's rank correlation, as implemented in the *Hmisc* R package (v 4.7), and heat maps were generated by the *corrplot* package (v0.92). Repeated measure correlations were performed using the *rmcorr* (v 0.5)[72].

db-RDA was performed to quantify the effect sizes of gut environmental factors and other variables on the intra-individual and inter-individual variation in the gut microbiome (both relative and quantitative profiles at genus level) and faecal and urine metabolomes (untargeted data, all features). The analyses were performed with Bray–Curtis dissimilarity using the *capscale* function as implemented in the *vegan* package (v 2.6). With regards to intra-individual analyses, data available from all samples (day 1 to day 9) and *strata* = Participant ID were used. For the inter-individual analyses, data collected on day 2 (visit 1) were used separately for all participants ($n = 61$) and for the SmartPill subgroup ($n = 50$). The statistical significance was determined by permutation test with 9,999 random permutations (*anova.cca* function), and P values were adjusted for multiple testing by false discovery rate (Benjamin–Hochberg)[73]. An adjusted P value (q value) below 0.1 was considered significant.

For the untargeted metabolomics data, the area of each m/z feature was log-transformed, and missing values were imputed and replaced by values reflecting half of the minimum intensity of the given m/z feature. Linear regression models and SPLS models were performed to examine the relationship between the m/z features and the variables of interest (that is, segmental transit time and pH). The modelling was performed using the SmartPill-derived data and the 24 h postprandial urine metabolome collected at day 2 as well as the faecal metabolome closest to the time of the SmartPill egestion. The linear mixed models were performed with the *lme4* R package (v1.1-31). The multivariate SPLS models were performed with the *caret* R package (v 6.0-92). P values were corrected for multiple testing by the Benjamin–Hochberg false discovery rate (q value). Features with $q < 0.1$ were considered to be statistically significant, and only features selected by both the linear regression and SPLS were further submitted for identification including the MS^2.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All sequencing data have been submitted to the National Center for Biotechnology Information Sequence Read Archive. BioProject ID, PRJNA1044006. MS^2 data of global urine and faecal pool samples are deposited at MassIVE MSV000095466. Individual-level personally identifiable MS^2 data from the participants cannot be made freely available to protect the privacy of the participants, in accordance with the Danish Data Protection Act and European Regulation 2016/679 of the European Parliament and of the council (GDPR) that prohibits open distribution even in pseudoanonymized form. Metabolomics data and data tables can be shared upon request. For data inquiries, please contact the principal investigator, H.M.R., via email. Access will

be evaluated and granted upon signing a data processing agreement between the governing legal entities. Source data are provided with this paper.

Code availability

No custom code was generated for this work.

References

1. David, L. A. et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).
2. Wu, G. D. et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108 (2011).
3. Johnson, A. J. et al. Daily sampling reveals personalized diet–microbiome associations in humans. *Cell Host Microbe* **25**, 789–802.e5 (2019).
4. Guthrie, L. et al. Impact of a 7-day homogeneous diet on interpersonal variation in human gut microbiomes and metabolomes. *Cell Host Microbe* **30**, 863–874 (2021).
5. Falony, G. et al. Population-level analysis of gut microbiome variation. *Science* **352**, 560–564 (2016).
6. Vandeputte, D. et al. Temporal variability in quantitative human gut microbiome profiles and implications for clinical research. *Nat. Commun.* **12**, 6740 (2021).
7. Procházková, N. et al. Advancing human gut microbiota research by considering gut transit time. *Gut* https://doi.org/10.1136/gutjnl-2022-328166 (2022).
8. Vandeputte, D. et al. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut* **65**, 57–62 (2016).
9. Roager, H. M. et al. Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat. Microbiol.* **1**, 16093 (2016).
10. Gill, P. A., van Zelm, M. C., Muir, J. G. & Gibson, P. R. Review article: short chain fatty acids as potential therapeutic agents in human gastrointestinal and inflammatory disorders. *Aliment. Pharmacol. Ther.* **48**, 15–34 (2018).
11. Edamatsu, T., Fujieda, A., Ezawa, A. & Itoh, Y. Classification of five uremic solutes according to their effects on renal tubular cells. *Int. J. Nephrol.* **2014**, 512178 (2014).
12. O'Keefe, S. J. D. Diet, microorganisms and their metabolites, and colon cancer. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 691–706 (2016).
13. Macfarlane, G. T., Gibson, G. R. & Cummings, J. H. Comparison of fermentation reactions in different regions of the human colon. *J. Appl. Bacteriol.* **72**, 57–64 (1992).
14. Alakomi, H. L. et al. Lactic acid permeabilizes gram-negative bacteria by disrupting the outer membrane. *Appl. Environ. Microbiol.* **66**, 2001–2005 (2000).
15. Kim, E. R. & Rhee, P. L. How to interpret a functional or motility test - colon transit study. *J. Neurogastroenterol. Motil.* **18**, 94–99 (2012).
16. Ringel-Kulka, T. et al. Altered colonic bacterial fermentation as a potential pathophysiological factor in irritable bowel syndrome. *Am. J. Gastroenterol.* **110**, 1339–1346 (2015).
17. Steenackers, N. et al. Specific contributions of segmental transit times to gut microbiota composition. *Gut* **71**, 1443–1444 (2022).
18. Diaz Tartera, H. O. et al. Validation of SmartPill® wireless motility capsule for gastrointestinal transit time: intra-subject variability, software accuracy and comparison with video capsule endoscopy. *Neurogastroenterol. Motil.* **29**, 1–9 (2017).
19. Wang, Y. T. et al. Regional gastrointestinal transit and pH studied in 215 healthy volunteers using the wireless motility capsule: influence of age, gender, study country and testing protocol. *Aliment. Pharmacol. Ther.* **42**, 761–772 (2015).
20. Safarpour, D. et al. Gastrointestinal motility and response to Levodopa in Parkinson's disease: a proof-of-concept study. *Mov. Disord.* https://doi.org/10.1002/mds.29176 (2022).
21. Jensen, M. M. et al. Human gastrointestinal transit and hormonal response to different meal types: a randomized crossover study. *J. Nutr.* **152**, 1358–1369 (2022).
22. Lewis, S. J. & Heaton, K. W. Stool form scale as a useful guide to intestinal transit time. *Scand. J. Gastroenterol.* **32**, 920–924 (1997).
23. Nestel, N. et al. The gut microbiome and abiotic factors as potential determinants of postprandial glucose responses: a single-arm meal study. *Front. Nutr.* **7**, 1–9 (2021).
24. Vieira-Silva, S. et al. Quantitative microbiome profiling disentangles inflammation- and bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses. *Nat. Microbiol.* **4**, 1826–1831 (2019).
25. Reitmeier, S. et al. Arrhythmic gut microbiome signatures predict risk of type 2 diabetes. *Cell Host Microbe* **28**, 258–272.e6 (2020).
26. Kuo, B. et al. Comparison of gastric emptying of a nondigestible capsule to a radio-labelled meal in healthy and gastroparetic subjects. *Aliment. Pharmacol. Ther.* **27**, 186–196 (2008).
27. Nandhra, G. K. et al. Normative values for region-specific colonic and gastrointestinal transit times in 111 healthy volunteers using the 3D-Transit electromagnet tracking system: influence of age, gender, and body mass index. *Neurogastroenterol. Motil.* **32**, e13734 (2020).
28. Asnicar, F. et al. Blue poo: impact of gut transit time on the gut microbiome using a novel marker. *Gut* **70**, 1–10 (2021).
29. LaBouyer, M. et al. Higher total faecal short-chain fatty acid concentrations correlate with increasing proportions of butyrate and decreasing proportions of branched-chain fatty acids across multiple human studies. *Gut Microbiome* **3**, 1–14 (2022).
30. Falony, G., Vieira-Silva, S. & Raes, J. Richness and ecosystem development across faecal snapshots of the gut microbiota. *Nat. Microbiol.* **3**, 526–528 (2018).
31. Fujita, T., Hada, T. & Higashino, K. Origin of D- and L-pipecolic acid in human physiological fluids: a study of the catabolic mechanism to pipecolic acid using the lysine loading test. *Clin. Chim. Acta* **287**, 145–156 (1999).
32. Meslier, V. et al. Mediterranean diet intervention in overweight and obese subjects lowers plasma cholesterol and causes changes in the gut microbiome and metabolome independently of energy intake. *Gut* https://doi.org/10.1136/gutjnl-2019-320438 (2020).
33. Andersen, M. B. S. et al. Untargeted metabolomics as a screening tool for estimating compliance to a dietary pattern. *J. Proteome Res.* **13**, 1405–1418 (2014).
34. Manach, C., Scalbert, A., Morand, C., Rémésy, C. & Jiménez, L. Polyphenols: food sources and bioavailability. *Am. J. Clin. Nutr.* **79**, 727–747 (2004).
35. Boyer, J. L. Bile formation and secretion. *Compr. Physiol.* **3**, 1035–1078, https://doi.org/10.1002/cphy.c120027 (2013).
36. Bernalier, A., Willems, A., Leclerc, M., Rochet, V. & Collins, M. D. *Ruminococcus hydrogenotrophicus* sp. nov., a new H_2/CO_2-utilizing acetogenic bacterium isolated from human feces. *Arch. Microbiol.* **166**, 176–183 (1996).
37. Louis, P. & Flint, H. J. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol. Lett.* **294**, 1–8 (2009).
38. Tigchelaar, E. F. et al. Gut microbiota composition associated with stool consistency. *Gut* **65**, 540–542 (2016).
39. Raba, G., Adamberg, S. & Adamberg, K. Acidic pH enhances butyrate production from pectin by faecal microbiota. *FEMS Microbiol. Lett.* **368**, 1–8 (2021).
40. Folz, J. et al. Human metabolome variation along the upper intestinal tract. *Nat. Metab.* https://doi.org/10.1038/s42255-023-00777-z (2023).

41. Shalon, D. et al. Profiling the human intestinal environment under physiological conditions. *Nature* https://doi.org/10.1038/s41586-023-05989-7 (2023).

42. Parthasarathy, G. et al. Relationship between microbiota of the colonic mucosa vs feces and symptoms, colonic transit, and methane production in female patients with chronic constipation. *Gastroenterology* **150**, 367–379.e1 (2016).

43. Cirstea, M. S. et al. Microbiota composition and metabolism are associated with gut function in Parkinson's Disease. *Mov. Disord.* **35**, 1208–1217 (2020).

44. Gabriele, S. et al. Slow intestinal transit contributes to elevate urinary *p*-cresol level in Italian autistic children. *Autism Res.* **9**, 752–759 (2016).

45. Manandhar, M. & Cronan, J. E. Pimelic acid, the first precursor of the *Bacillus subtilis* biotin synthesis pathway, exists as the free acid and is assembled by fatty acid synthesis. *Mol. Microbiol.* **104**, 595–607 (2017).

46. Cronan, J. E. & Lin, S. Synthesis of the α,ω-dicarboxylic acid precursor of biotin by the canonical fatty acid biosynthetic pathway. *Curr. Opin. Chem. Biol.* **15**, 407–413 (2011).

47. Dong, Y. et al. Development and validation of diagnostic models for immunoglobulin A nephropathy based on gut microbes. *Front. Cell Infect. Microbiol.* **12**, 1–15 (2022).

48. Coker, O. O. et al. Altered gut metabolites and microbiota interactions are implicated in colorectal carcinogenesis and can be non-invasive diagnostic biomarkers. *Microbiome* **10**, 1–12 (2022).

49. Li, J. et al. Interplay between diet and gut microbiome, and circulating concentrations of trimethylamine *N*-oxide: findings from a longitudinal cohort of US men. *Gut* **71**, 724–733 (2022).

50. Murtagh, F. E. M., Addington-Hall, J. & Higginson, I. J. The prevalence of symptoms in end-stage renal disease: a systematic review. *Adv. Chronic Kidney Dis.* **14**, 82–99 (2007).

51. Sinha, A. K. et al. Dietary fibre directs microbial tryptophan metabolism via metabolic interactions in the gut microbiota. *Nat. Microbiol.* **9**, 1964–1978 (2024).

52. Tran, K., Brun, R. & Kuo, B. Evaluation of regional and whole gut motility using the wireless motility capsule: relevance in clinical practice. *Ther. Adv. Gastroenterol.* **5**, 249–260 (2012).

53. S, M., HP, P. & FK, F. Wireless capsule motility: comparison of the SmartPill GI monitoring system with scintigraphy for measuring whole gut transit. *Dig. Dis. Sci.* **54**, 2167–2174 (2009).

54. Bartholomé, R. et al. Paracetamol as a post prandial marker for gastric emptying, a food–drug interaction on absorption. *PLoS ONE* **10**, 1–9 (2015).

55. Ambrus, A. et al. Example of a protocol for identification of misreporting (under- and over-reporting of energy intake) based on the PILOT-PANEU project. *EFSA/EU J.* https://www.efsa.europa.eu/sites/default/files/efsa_rep/blobserver_assets/3944A-8-2-1.pdf (2013).

56. Mortensen, M. S. Optimized DADA2 pipeline for 16S rRNA gene sequencing using IonTorrent. *DTU Data* https://doi.org/10.11583/DTU.22657339.v1 (2023).

57. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

58. Vandeputte, D. et al. Quantitative microbiome profiling links gut community variation to microbial load. *Nature* **551**, 507–511 (2017).

59. Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. K. & Schmidt, T. M. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res.* **43**, D593–D598 (2015).

60. Chambers, M. C. et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).

61. Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R. & Siuzdak, G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **78**, 779–787 (2006).

62. Chambers, J. M. & Hastie, T. *Statistical Models in S* (Chapman & Hall/CRC, 1992).

63. Dieterle, F., Ross, A., Schlotterbeck, G. & Senn, H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal. Chem.* **78**, 4281–4290 (2006).

64. Kuhl, C. et al. CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal. Chem.* https://doi.org/10.1021/ac202450g (2011).

65. Schmid, R. et al. Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nat. Biotechnol.* **41**, 447–449 (2023).

66. Hoffmann, M. A. et al. High-confidence structural annotation of metabolites absent from spectral libraries. *Nat. Biotechnol.* **40**, 411–421 (2022).

67. Wang, M. et al. Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat. Biotechnol.* **34**, 828–837 (2016).

68. Zuffa, S. et al. microbeMASST: a taxonomically informed mass spectrometry search tool for microbial metabolomics data. *Nat. Microbiol.* **9**, 336–345 (2024).

69. Dührkop, K. et al. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods* **16**, 299–302 (2019).

70. Cuparencu, C. S. et al. Identification of urinary biomarkers after consumption of sea buckthorn and strawberry, by untargeted LC–MS metabolomics: a meal study in adult men. *Metabolomics* **12**, 1–20 (2016).

71. Sumner, L. W. et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **3**, 211–221 (2007).

72. Bakdash, J. Z. & Marusich, L. R. Repeated measures correlation. *Front. Psychol.* **8**, 456 (2017).

73. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).

Acknowledgements

Author contributions

N.P., T.R.L., L.O.D. and H.M.R. conceived and designed the human study as part of the PRIMA collaboration headed by T.R.L. N.P. conducted the study under the supervision of L.O.D. and H.M.R. Urine metabolomics was performed by N.P. Metabolite annotations were done by N.P., G.L.B. and L.O.D. Metabolite synthesis and fine identification were done by G.L.B. Faecal SCFAs were analysed by E.T.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41564-024-01856-x.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-024-01856-x.

Correspondence and requests for materials should be addressed to Henrik M. Roager.

Peer review information *Nature Microbiology* thanks Pieter C. Dorrestein and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.
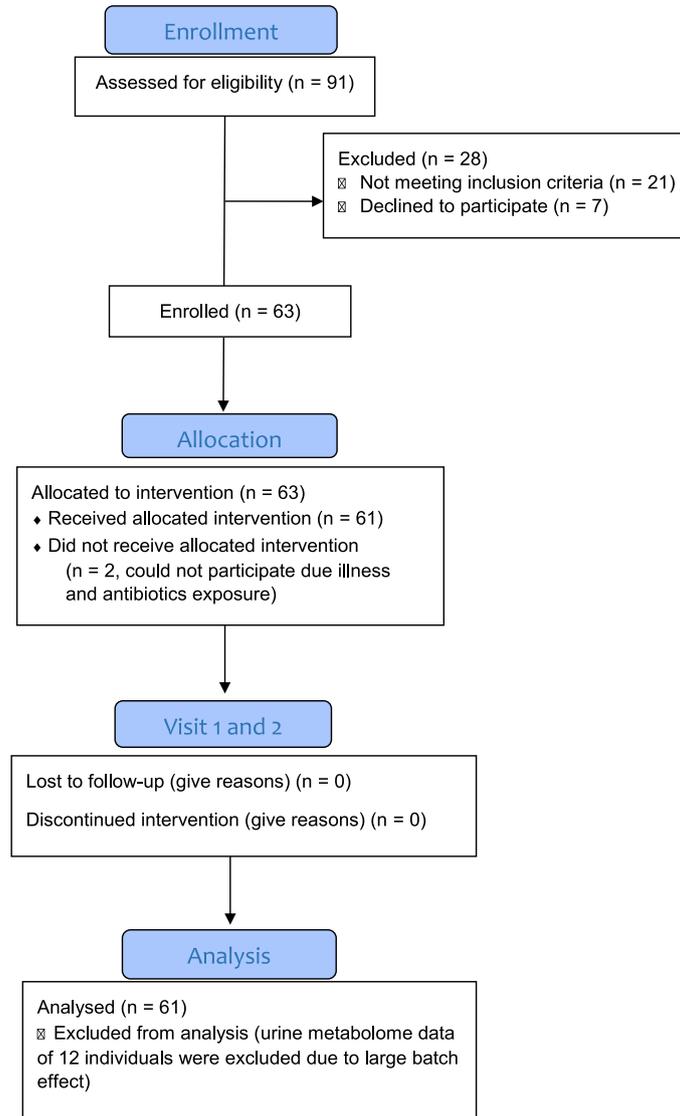
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
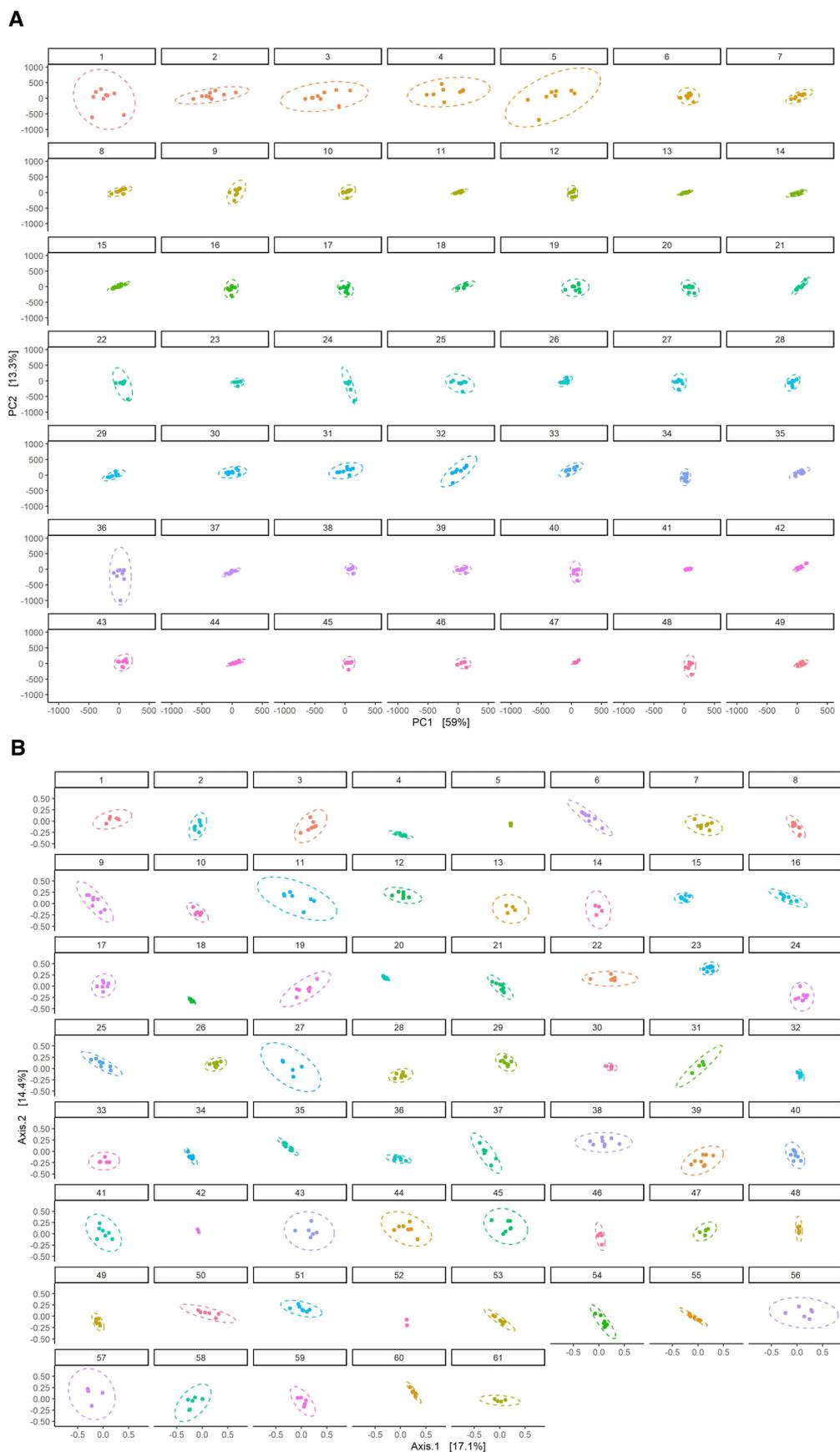
CONSORT Flow Diagram

Enrollment

Assessed for eligibility (n = 91)

Excluded (n = 28)
- ☒ Not meeting inclusion criteria (n = 21)
- ☒ Declined to participate (n = 7)

Enrolled (n = 63)

Allocation

Allocated to intervention (n = 63)
- ♦ Received allocated intervention (n = 61)
- ♦ Did not receive allocated intervention
  (n = 2, could not participate due illness
  and antibiotics exposure)

Visit 1 and 2

Lost to follow-up (give reasons) (n = 0)

Discontinued intervention (give reasons) (n = 0)

Analysis

Analysed (n = 61)
- ☒ Excluded from analysis (urine metabolome data
  of 12 individuals were excluded due to large batch
 effect)

Extended Data Fig. 1 | CONSORT flow diagram of the PRIMA study. This diagram illustrates the flow of participants through each stage of the PRIMA study, including enrollment, allocation, visit 1 and visit 2, and analysis.

A

B

Extended Data Fig. 2 | Individual microbiome and metabolome profiles. (a) Individual Bray Curtis β-diversity ordination of quantitative faecal microbiome profiles over 9 study days (n = 61) **(b)** Euclidian β-diversity ordination of urinary metabolomes profiles over 9 study days (n = 49, 12 individuals were removed due to LC-MS batch effect).
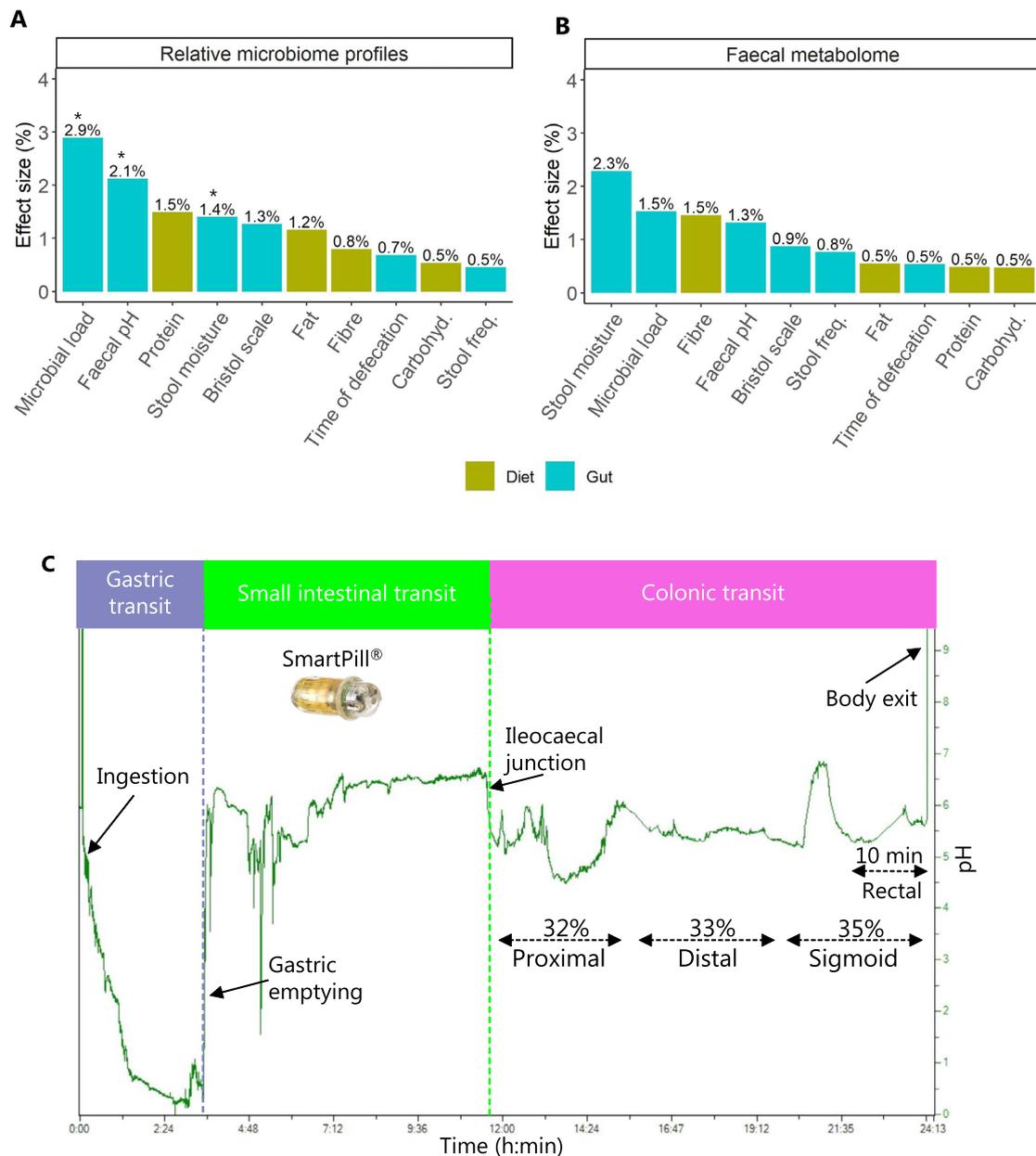
Extended Data Fig. 3 | Intra-individual variations in faecal microbiome and metabolome, and an example of a SmartPill profile. Contributions of dietary and gut factors on intra-individual variations in (a) relative microbiome profiles and (b) faecal metabolome. The analysis was performed with distance-based redundancy analysis (db-RDA) with permutation test on daily relative microbiome data (n = 61, 9 days) and untargeted faecal metabolome data (n = 61, 3 days). The asterisks indicate statistical significance after adjustment for multiple testing (*q-value < 0.05). **(c) An example of a pH profile measured by** **the SmartPill**. Segmental transit times were determined based on pH changes upon gastric emptying, ileocaecal junction and body exit as indicated. The proximal-, distal-, and sigmoid-colon pH were determined as median values in each of the segments of the colon based on an approximation of the transit time based on previous data showing that the first 32% followed by 33 % and 35 % of CTT corresponds to the proximal, distal, and sigmoid colon, respectively. In addition, the median pH of 10 min before the capsule egestion was registered as rectal pH.
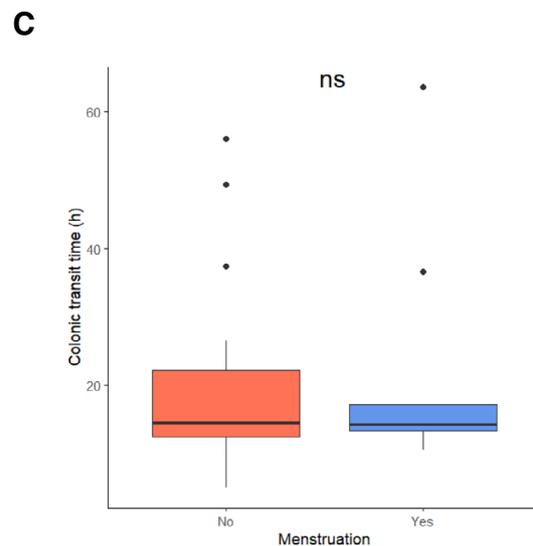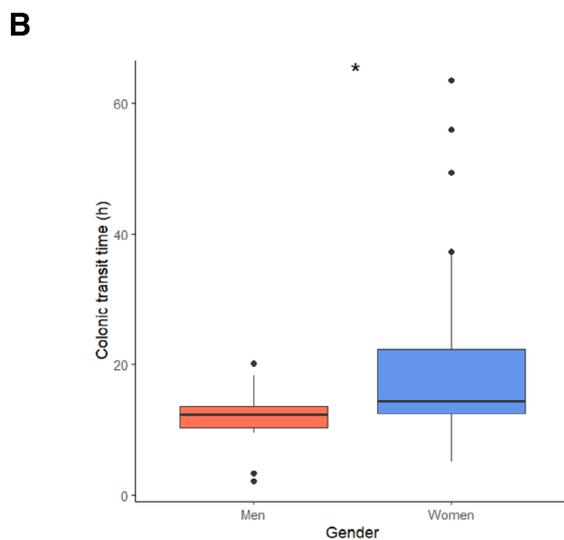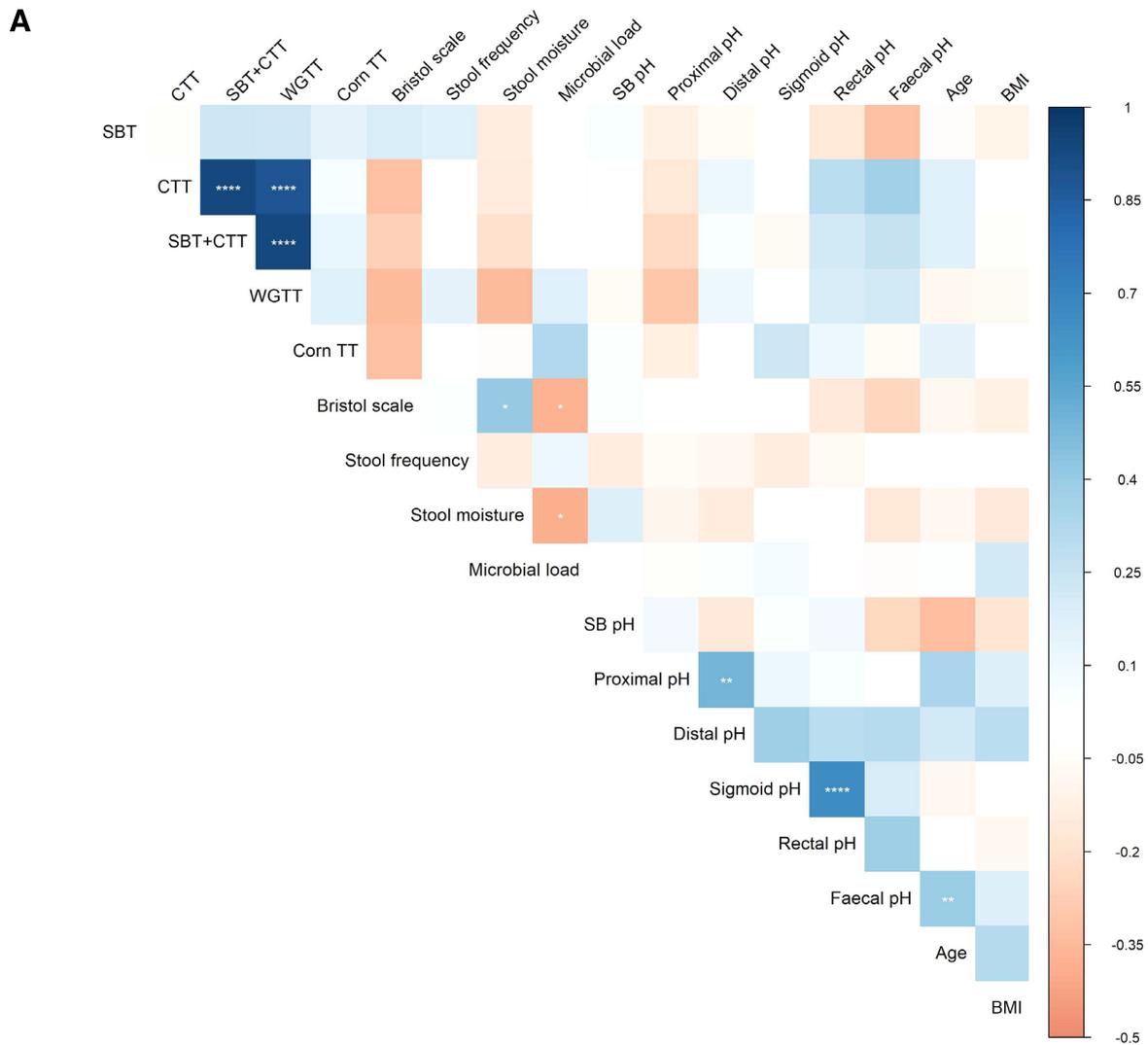
Extended Data Fig. 4 | Correlations between various factors assessed in the trial, and sex differences in colonic transit time. (a) Spearman correlation analysis between segmental transit times assessed by the SmartPill, corn transit time, various proxy markers of transit time, gut factors, and subject characteristics. The colour gradient shows the Spearman correlation coefficient and the asterisks indicate statistical significance (****q < 0.001,***q < 0.01,

** q < 0.05,*q < 0.1). SBT; small bowel transit time, CTT; colonic transit time, WGTT; whole gut transit time). Differences in colonic transit time according to (**b**) gender and (**c**) menstruation (only women, n = 43). Wilcoxon test; two-sided *p < 0.05, boxplot center represents median and box interquartile range (IQR). Whiskers extend to most extreme data point <1.5 IQR.
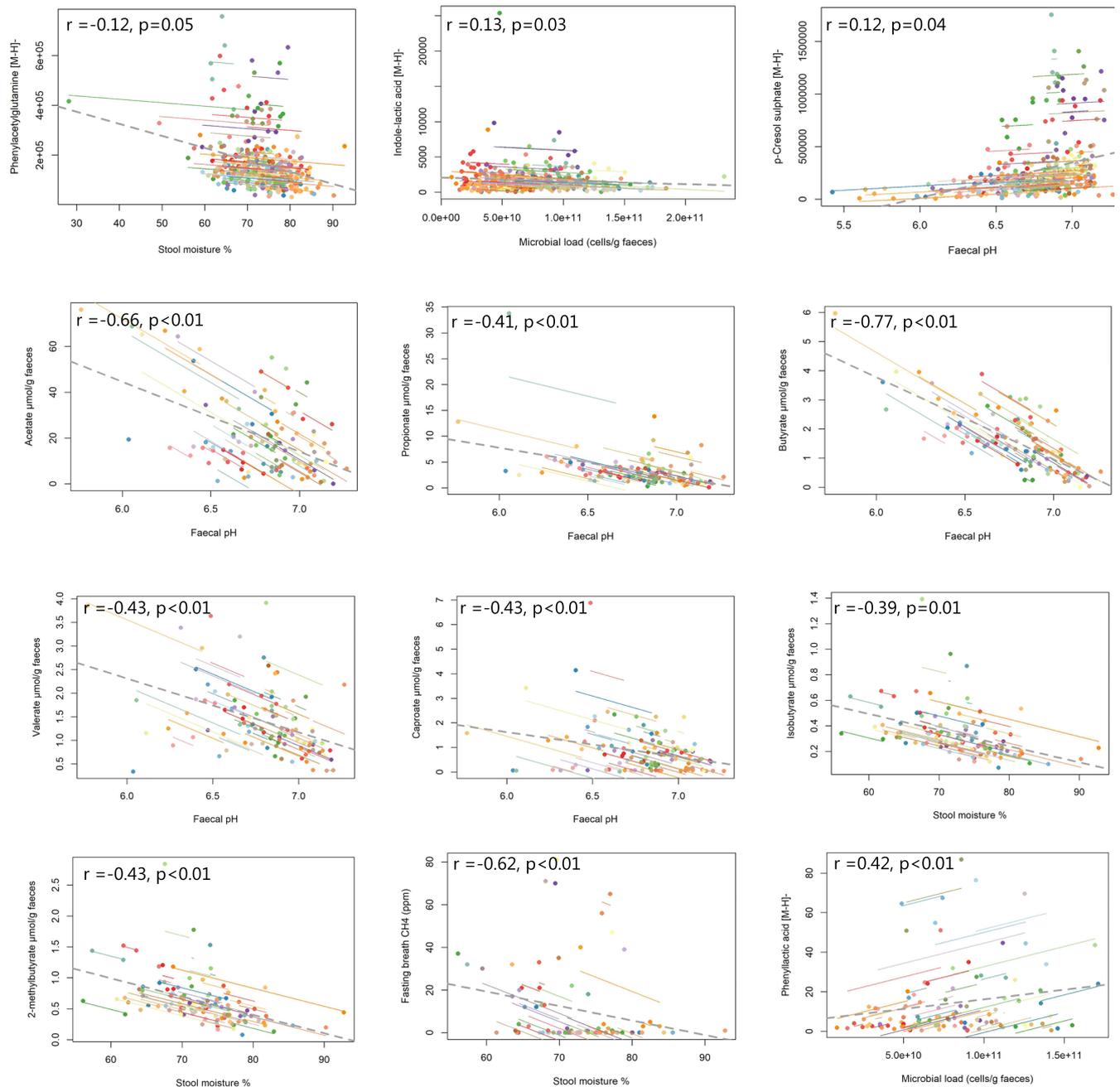
Extended Data Fig. 5 | Repeated measures correlation between gut factors and microbial metabolites. The colour lines show the individual correlation between each pair of tested variables for each of the study days using data from all 61 participants. The grey dashed lines, the correlation coefficients (r) and the p values indicate the overall trends.

nature portfolio

Corresponding author(s): Henrik Munch Roager

Last updated by author(s): Oct 8, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Motility GI software (v 3.1) was used to download and analyse data from the SmartPill receiver Dietary records were collected using the online tool MyFood24 (Research version 1).
Data analysis	The gut microbiota data were pre-processed into an ASV table using an in-house pipeline (https://doi.org/10.11583/DTU.22657339.v1) based on the DADA2 algorithm and the taxonomic classification was assigned using the RDP database v18 (https://mothur.org/wiki/rdp_reference_files/). The raw metabolomics data were converted to mzML files using msConvert from Proteowizard (v 3.0.24117); MzMine 3 and MassLynx (v4.2,Waters) were used for data visualization. Quantification of faecal short-chain fatty acids was performed by QuantLynx (v4.2, Waters). Statistical analysis and pre-processing of untargeted metabolomics data was performed using R(v4.0.2), RStudio(v1.3), and GraphPad Prism (v9.2.0). The following R packages were used: caret(v6.0.92), vegan(v2.6), lme4(1.1.31), phyloseq(v1.42.0), tidyverse(v1.3.0), XCMS(v3.18), CAMERA(v1.52), lmerTest(v3.1-3), corrplot(v0.92), rmcorr(v0.5), Hmisc(v4.7). Software feature annotations included GNPS(https://gnps.ucsd.edu/, v30), microbeMASST(https://masst.gnps2.org/microbemasst/, v2024.08.26), and SIRIUS (https://bio.informatik.uni-jena.de/software/sirius/, v6.0.5). Cytoflex flow cytometer (CytoFLEX 3; Beckman) was used for flow cytometry.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

All sequencing data have been submitted to the NCBI Sequence Read Archive (SRA). BioProject ID: PRJNA1027590.
MS/MS data of global urine and faecal pool samples are deposited at MassIVE MSV000095466. Individual-level personally identifiable MS/MS data from the subjects cannot be made freely available, to protect the privacy of the participants, in accordance with the Danish Data Protection Act and European Regulation 2016/679 of the European Parliament and of the Council (GDPR) that prohibit open distribution even in pseudoanonymised form. Metabolomics data and data tables can be shared upon request. For data inquiries, please contact the principal investigator, HMR (hero@nexs.ku.dk) via email. Access will be evaluated and granted upon signing a Data Processing Agreement between the governing legal entities.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	43 women and 18 men participated in the PRIMA study. The biological sex was assigned during the screening visit.
Reporting on race, ethnicity, or other socially relevant groupings	Race and ethnicity were not considered when designing the PRIMA study or consenting participants for the study.
Population characteristics	Sixty-three healthy participants living in Denmark were enrolled and 61 completed the study (43 women and 18 men). Participants were healthy by self-report (did not suffer from inflammatory bowel syndrome, small intestinal overgrowth, inflammatory bowel disease, chronic or infections disease, diabetes or cancer), aged 18-75 years with a BMI between 18.5 and 30.0 kg/m2 with no intake of medication with the exception of mild antidepressants and contraceptive pills. Intake of antibiotics, diarrhoea inhibitors and laxatives one month prior to the trial was not allowed. Furthermore, pregnant or lactating women were not included in the trial.
Recruitment	Participants were recruited via social media, internet (www.forsøgsperson.dk, www.nexs.ku.dk) and flyers distributed at the Department of Nutrition, Exercise and Sports at the University of Copenhagen. One potential source of bias in this study is self-selection, as participants were volunteers likely to be more interested in science and health, potentially skewing the sample towards a higher socioeconomic class. This could impact the results by over-representing health-conscious behaviors, such as diet and lifestyle, which may not reflect the broader population. As a result, the findings may have limited generalizability to individuals from diverse socioeconomic backgrounds.
Ethics oversight	The PRIMA study was approved by the Municipal Ethical Committee of the Capital Region of Denmark (H-20074067) and all participants provided written informed consent to participation.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to pre-determine sample size. Based on a previous study investigating the effect of colonic transit time on the human gut microbiome and urine metabolome, similar sample size was sufficient to obtain statistical power to identify associations between intestinal transit time, the faecal microbiome and metabolome (doi: 10.1038/nmicrobiol.2016.93).
Data exclusions	Microbiome profiling: 123 samples were removed due to low read number (< 10 000 reads). Urine metabolome: samples from 12 individuals were removed due to batch effect on the LC-MS SmartPill data: Gastric emptying times exceeding 8h and the corresponding whole gut transit times were excluded from the analyses.
Replication	Our study was an observational study and no replication was performed.
Randomization	No randomisation was performed as the study was observational.

| Blinding | No blinding was performed as the study was observational. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	https://www.clinicaltrials.gov/study/NCT04804319
Study protocol	The study protocol can be found in the Supplementary Data.
Data collection	The recruitment and data collection were carried out between April 2021 and December 2021.
Outcomes	The PRIMA study was an explorative study. The primary outcome was to investigate associations between faecal pH and gut microbial saccharolytic/proteolytic metabolism (assessed by targeted and untargeted metabolomics). The secondary outcomes included to explore relationships between the gut environmental factors (small intestinal and colonic pH and transit time assessed by wireless motility capsules, and various transit time proxy markers measured in faeces) and gut microbiome and metabolome assessed by 16s rRNA sequencing and metabolomics.

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Bacterial load in faeces was determined using approximately 500 μL of frozen faecal slurry (238 − 816 mg) and diluting it 400,000 times in physiological solution (8.5 g/L NaCl; VWR International). Next, 1 ml of the microbial cell suspension obtained was stained with 1 μL SYBR Green I (1:100 dilution in dimethylsulfoxide; shaded 20 min incubation at 37 °C; 10,000 concentrate, Thermo Fisher Scientific).
Instrument	The flow cytometry analysis of the bacterial cells present in the suspension was performed using a Cytoflex flow cytometer (CytoFLEX S; Beckman)
Software	CytExpert Software (Beckman)
Cell population abundance	n/a; In this study, total microbial cells present within fecal samples were measured. No subpopulations were assessed.
Gating strategy	Fluorescence events were measured using fluorescence channels FITC 525/40 nm combining with Side Scatter Channel SSC-A. In addition, backward gating at FSC-A combining SSC-A dot plot, FICT histogram plot and FITC combining PerCP dot plot was applied to double check the distribution of detected cells. Instrument and gating settings were identical for all samples (fixed staining–gating strategy).

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.