

The fecal metabolome as a functional readout of the gut microbiome

Jonas Zierer^{1,2,8}, Matthew A. Jackson¹, Gabi Kastenmüller², Massimo Mangino^{1,3}, Tao Long^{1,3}, Amalio Telenti⁴, Robert P. Mohn⁵, Kerrin S. Small¹, Jordana T. Bell¹, Claire J. Steves¹, Ana M. Valdes^{1,6,7}, Tim D. Spector^{1,10*} and Cristina Menni^{1,10*}

The human gut microbiome plays a key role in human health¹, but 16S characterization lacks quantitative functional annotation². The fecal metabolome provides a functional readout of microbial activity and can be used as an intermediate phenotype mediating host-microbiome interactions³. In this comprehensive description of the fecal metabolome, examining 1,116 metabolites from 786 individuals from a population-based twin study (TwinsUK), the fecal metabolome was found to be only modestly influenced by host genetics (heritability (H^2) = 17.9%). One replicated locus at the *NAT2* gene was associated with fecal metabolic traits. The fecal metabolome largely reflects gut microbial composition, explaining on average 67.7% ($\pm 18.8\%$) of its variance. It is strongly associated with visceral-fat mass, thereby illustrating potential mechanisms underlying the well-established microbial influence on abdominal obesity. Fecal metabolic profiling thus is a novel tool to explore links among microbiome composition, host phenotypes, and heritable complex traits.

There is growing evidence that the gut microbiome contributes to maintaining homeostasis of host metabolism⁴. Disruption of this intricate system is associated with diseases such as obesity^{4,5} and insulin resistance⁶. Metabolomics and the gut microbiome are strongly related, and microbes produce many of the body's chemicals, hormones, and vitamins⁷. The gut microbiome has been reported to affect circulating levels of several metabolites such as branched-chain amino acids, thus potentially causing insulin resistance⁸. However, despite the advances in next-generation-sequencing platforms, which allow for profiling of complex microbial communities through 16S sequencing, annotation is sparse. Moreover, the microbiome provides information on possible microbial entities rather than their actual activity: it cannot indicate the transcriptional activity of the genes within each bacterial genome² or differentiate between alive and dead microbes⁹. Fecal metabolomics, however, reports specifically on the metabolic interplay among the host, diet and gut microbiota³, and it complements sequencing-based approaches by providing a functional readout of the microbiome. Here, we provide a comprehensive description of the fecal metabolome in a large population-based setting, by taking advantage of the twin model. We report fecal metabolite (i) associations with age, sex, and obesity; (ii) associations with host genetic

influences; and (iii) uni-multivariable associations with gut microbiome composition.

We analyzed fecal samples of 786 predominantly female twins in the TwinsUK cohort, 65.2 (± 7.6) years of age, with an average body mass index (BMI) of 26.1 (± 4.7) (Supplementary Table 1), and we replicated our genetic results in an independent sample of 230 individuals, 66.9 (± 8.6) years of age, with an average BMI of 27.2 (± 5.2). Untargeted metabolomics profiling of the participants' fecal samples was conducted through mass spectrometry performed by Metabolon, Inc. A total of 1,116 metabolites, including 866 with known chemical identity, were measured. Among the metabolites identified, 570 were common and detected in at least 80% of the samples, whereas 345 were detected in at least 20% but less than 80% of all samples (Fig. 1a). The latter 345 samples were analyzed as dichotomous traits (on the basis of presence/absence in a sample), and metabolites measured in less than 20% of the samples were discarded from further analysis. Of the 1,116 measured metabolites, 647 were not detected in blood samples of the same individuals profiled on the same platform (Fig. 1b). This result suggests that the fecal metabolome provides information complementary to that from blood metabolomics. We did not find significant associations between 915 fecal metabolites and age after correcting for multiple testing. However, a multivariate partial-least-squares discriminant analysis incorporating all the common 570 metabolites was able to distinguish the oldest decile (>75 years) from the youngest decile (<56 years) of the study population (area under the curve = 0.71, $P = 6.8 \times 10^{-6}$; Fig. 2b), and one metabolite, phytanate, was significantly different between the oldest and youngest deciles ($P = 5.0 \times 10^{-3}$; Fig. 2a). These results suggest nonlinear associations between the fecal metabolome and age, in line with previous reports on the effects of age on the gut microbiome^{9,10}.

BMI was associated with eight metabolites at a false discovery rate (FDR, Benjamini-Hochberg corrected) of 5%: five fecal lipids, including arachidonate (β [95% confidence interval (CI)] = 0.13 [0.07–0.19], $P = 1.1 \times 10^{-5}$), the hemoglobin metabolite bilirubin (β = 0.13 [0.06–0.19], $P = 8.9 \times 10^{-5}$), and two unknown metabolites (Supplementary Table 2). We then looked for associations with visceral-fat mass, a measure of abdominal obesity, correcting for BMI, and found a total of 102 statistically significant associations (FDR $< 5\%$, 13 passing Bonferroni correction), which together explained 28.4% of the observed total variance in visceral fat ($P < 2.2 \times 10^{-16}$).

¹Department for Twin Research & Genetic Epidemiology, King's College London, London, UK. ²Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg, Germany. ³NIHR Biomedical Research Centre at Guy's and St Thomas' Foundation Trust, London, UK. ⁴Human Longevity, Inc, San Diego, CA, USA. ⁵Metabolon, Inc, Durham, NC, USA. ⁶NIHR Nottingham Biomedical Research Centre, Nottingham, UK. ⁷School of Medicine, University of Nottingham, Nottingham, UK. ⁸Present address: Institute for Computational Biomedicine, Weill Cornell Medical College, New York, NY, USA. ⁹Present address: Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, USA. ¹⁰These authors jointly directed this work: Tim D. Spector, Cristina Menni. *e-mail: tim.spector@kcl.ac.uk; cristina.menni@kcl.ac.uk

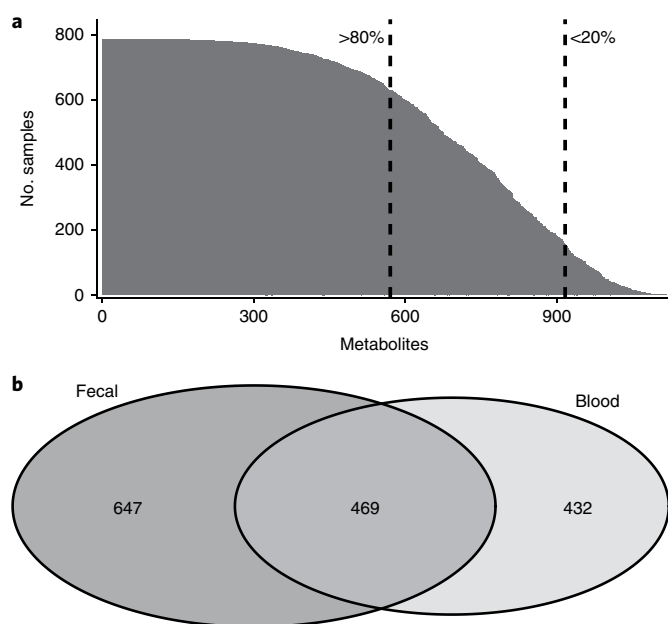


Fig. 1 | Number of measured fecal metabolites. **a**, 1,116 metabolites were detected in 786 fecal samples; 570 metabolites were detected in at least 80% of all samples, and 345 metabolites were detected in less than 80% but more than 20% of all samples. The former 570 were analyzed continuously, whereas the latter 345 were dichotomized as present/absent. 210 metabolites that were present in less than 20% of samples were excluded from further analysis. **b**, 469 metabolites were observed in both fecal and blood samples from the sampled individuals, whereas 647 metabolites were unique to feces. 499 of these 647 metabolites were observed in at least 20% of the fecal samples.

(Supplementary Table 2). In contrast, we found only eight metabolites associated with BMI, an imprecise measure of adiposity that is based on overall mass without making a distinction between lean and fat mass¹¹. However, the gut microbiome is known to play a major role in fatty acid metabolism and adiposity, which may be better reflected by visceral-fat measures^{12,13}. Emerging evidence also suggests a role of the intestinal microbiota in visceral development through interaction with dietary components¹⁴. We have shown strong associations between visceral-fat mass and gut microbiome composition^{15,16}. The much larger number of fecal metabolite associations with visceral fat than with BMI is consistent with these findings and highlights the strong influence of gut metabolic processes on abdominal adiposity.

Visceral-fat-associated metabolites were significantly enriched in amino acids (43 metabolites, enrichment P value $< 2 \times 10^{-4}$) but also included 14 fatty acids, including arachidonate ($\beta = 5.07$ [2.55–7.59], $P = 8.2 \times 10^{-5}$), 8 nucleotides, 6 sugars, and 6 vitamins. The strong association between the fecal metabolome and central obesity confirms hypotheses on the involvement of microbial amino acid metabolism in obesity and suggests new mechanisms, such as microbial vitamin B metabolism. In previous work, we have found several microbe families associated with lower visceral-fat mass¹⁵ and reduced weight gain in germ-free mice receiving human fecal transplants compared with germ-free mice not receiving human fecal transplants¹⁷. By analyzing the fecal metabolome, we found the abundance of the same families to be strongly associated with lower abundance of amino acids (described below), thus suggesting that their association with visceral fat may be mediated by amino acid availability (Fig. 3). This association may be due to increased utilization or decreased production of amino acids by these bacteria, or may be the result of more complex host–microbe interactions.

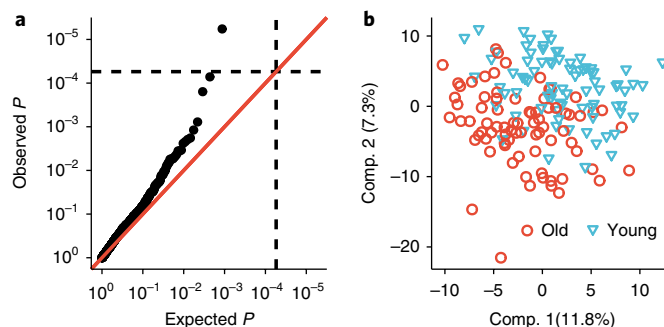


Fig. 2 | Association of the fecal metabolome with age. We compared the fecal metabolome between the oldest (> 75 years, $n = 79$) and youngest decile (< 56 years, $n = 80$) of the study population. **a**, First, we investigated the age effect for all metabolites individually by using logistic regression models and found one metabolite, phytanate, that was significantly different between the youngest and the oldest decile of the data. **b**, Then we fitted a multivariate partial-least-squares discriminant analysis model to distinguish the older (red) from the younger (blue) group. We estimated the area under the receiver operating characteristic curve to be 0.71 ($P = 6.8 \times 10^{-6}$) in a tenfold cross-validation. Comp, component. P values were determined by partial-least-squares discriminant analysis.

The gut microbiome is heritable^{17,18}, and we found a heritable variance component for 210 operational taxonomic units (OTUs), which explained 22.7% of the observed total variance, on average. To test whether host genetics also influences the fecal metabolome, we first estimated its heritability, taking advantage of the twin structure in our data (Fig. 4) by using structural equation modeling (148 monozygotic (MZ) pairs and 155 dizygotic (DZ) pairs). For 428 metabolites, the best-fitting model contained a heritable variance component (A), which explained 17.9% ($\pm 9.7\%$) of the metabolite variation, on average. Long-chain fatty acid-containing metabolites, such as 1-palmitoyl-2-arachidonoyl-GPC ($H^2 = 60.7\%$ [43.4–78.0]) and stearyl carnitine ($H^2 = 54.3\%$ [36.4–72.3]), were among the most heritable metabolites. For 279 metabolites, including the coffee metabolite 5-acetylamin-6-amino-3-methyluracil (common environment component (C) = 30.3% [20.0–40.6]), the best fitting model was the common environment–unique environment (CE) model, in which C explained 14.8% ($\pm 8.1\%$) of the variance, on average. For the remaining 208 metabolites, the best-fitting model was the unique environment (E) model, wherein the entire variation of the metabolite was due to individual differences such as the microbiome or individual diet (Supplementary Table 2). We found a significantly stronger environmental effect on lipids than other metabolites (enrichment P value $< 2 \times 10^{-4}$).

We subsequently conducted a genome-wide association study for the 428 metabolites with a heritable component (Supplementary Table 3) and identified three metabolites (the amino acid 3-phenylpropionate and two lipids, eicosapentaenoate and 3-hydroxyhexanoate) that were significantly associated with genetic loci after correction for multiple testing ($P < 1.2 \times 10^{-10} = 5 \times 10^{-8}/428$) (Table 1 and Fig. 5a). We also tested for genetic associations with metabolite ratios, which can be better proxies for chemical reactions than single metabolites¹⁹. After correcting for 31,226 tested ratios, we found that the ratio of 5-acetylamin-6-amino-3-methyluracil and 1,3-dimethylurate was associated with a locus on chromosome 8 (rs35246381, $P = 7.0 \times 10^{-21}$, P gain = 7.5×10^9) (Table 1 and Fig. 5b). We replicated the results of our genome-wide association study in an independent sample of 230 individuals. Of the four loci tested, only the metabolite ratio of 5-acetylamin-6-amino-3-methyluracil to 1,3-dimethylurate was significantly associated in the replication cohort ($P = 3.6 \times 10^{-10}$; meta-analysis $P = 3.3 \times 10^{-36}$). The two metabolites 5-acetylamin-6-amino-3-methyluracil and

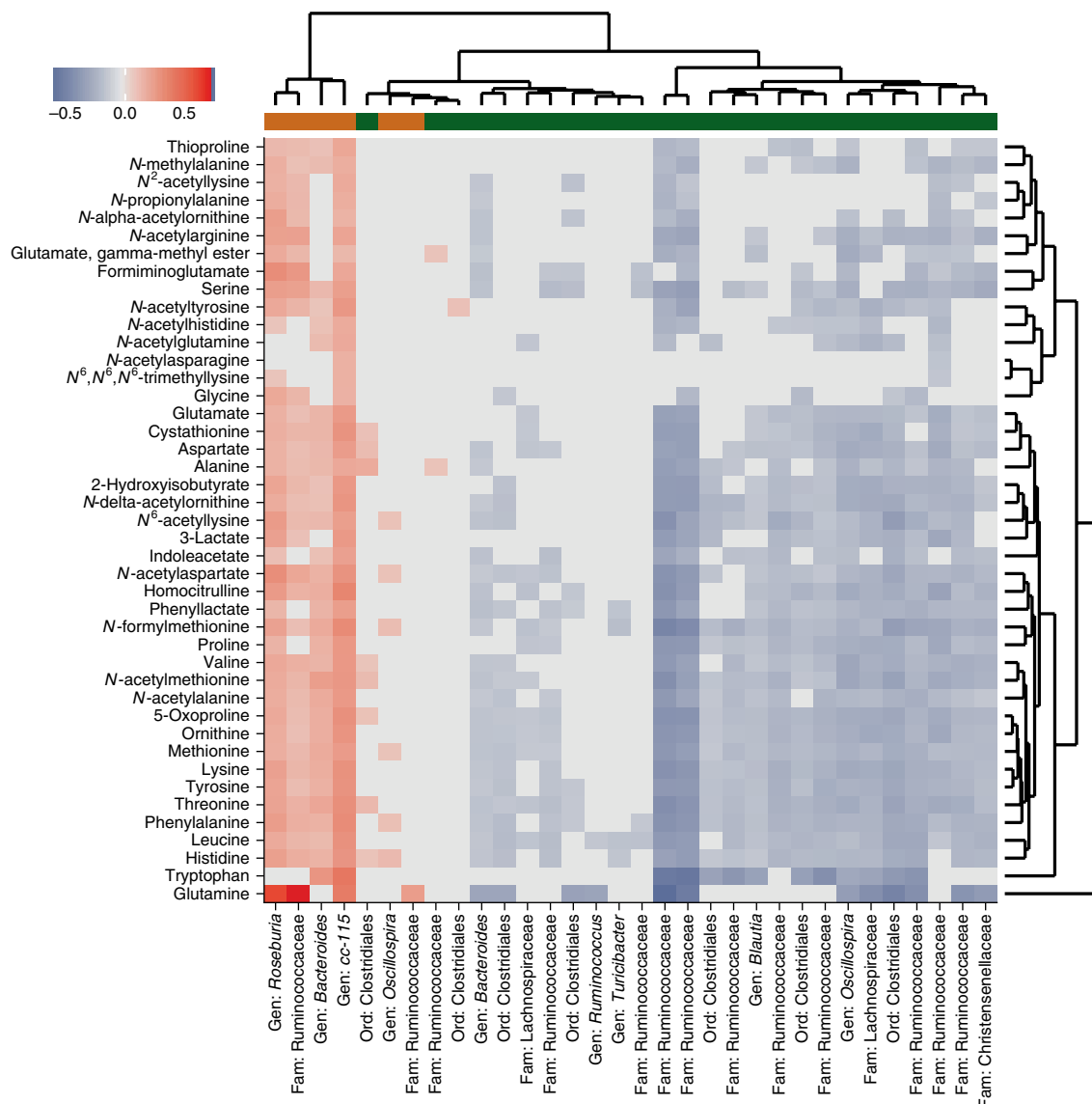


Fig. 3 | Associations between fecal metabolites and the gut microbiome correspond to microbial effects on visceral fat. Visceral-fat mass was significantly associated with 43 fecal amino acids (all positively) ($n=647$) and 32 OTUs ($n=540$) (6 positively in orange, 26 negatively in green). Red tiles indicate positive associations between these metabolites and OTUs ($\beta > 0$); blue tiles indicate negative associations ($\beta < 0$); gray tiles indicate nonsignificant associations ($\text{FDR} > 5\%$). Microbial associations with fecal metabolites correspond to their respective associations with visceral fat, thus indicating that the microbial metabolic profile is more closely related to the host phenotype than taxonomy. Gen, genus; fam, family; ord, order.

1,3-dimethylurate are products of caffeine metabolism²⁰. The associated locus at the *NAT2* gene encodes an *N*-acetyltransferase, which catalyzes the degradation of caffeine metabolites²¹ (Supplementary Fig. 3). Associations of this locus with other caffeine metabolites (1-methylxanthine, 4-acetamidobutanoate, and 1-methylurate) have been observed in blood²² and urine²³, and are likely to reflect the efficiency of caffeine degradation. We then explored whether there were any expression quantitative trait loci (eQTLs) or other functional variants in strong linkage disequilibrium (LD) with the top SNP. Although we found three eQTLs (rs11996129, rs1112005, and rs1799930) for *NAT2* (ref. ²⁴), they were in only weak LD ($r^2 < 0.16$) with rs35246381 (ref. ²⁵), and the associations between these SNPs and the metabolite ratio were weaker than that for the top SNP ($P = 3.6 \times 10^{-10}$ versus $P = 9.4 \times 10^{-7}$). *NAT2* expression is highest in the liver, then in the jejunal and colonic mucosa, duodenum, colon, and small intestine²⁶ (see URLs). This tissue expression is consistent with polymorphisms in the *NAT2* gene being associated with the

concentration of caffeine-derived metabolites in feces. We explored the relationship between caffeine and the fecal metabolites 5-acetylaminol-6-amino-3-methyluracil and 1,3-dimethylurate, and found that their ratio was positively correlated with both coffee intake and serum caffeine levels (Supplementary Fig. 3). These genetic association data illustrate how part of the complex metabolism of caffeine takes place in the intestine before the metabolites reach the liver, and that the links between host genetic makeup and xenobiotic concentrations can be captured in fecal metabolites. Beyond its function in caffeine metabolism, the *NAT2* enzyme is involved in metabolism of various xenobiotics and is therefore related to variance in drug response and toxicity²⁷. The composition of the gut microbiome has been shown to regulate xenobiotic enzymes; for instance, the expression of *NAT2* in the large intestine is 1.5 times higher in germ-free animals than control animals²⁸. These data together suggest that xenobiotic metabolism may be jointly regulated by host genetic variation and gut microbiome composition.

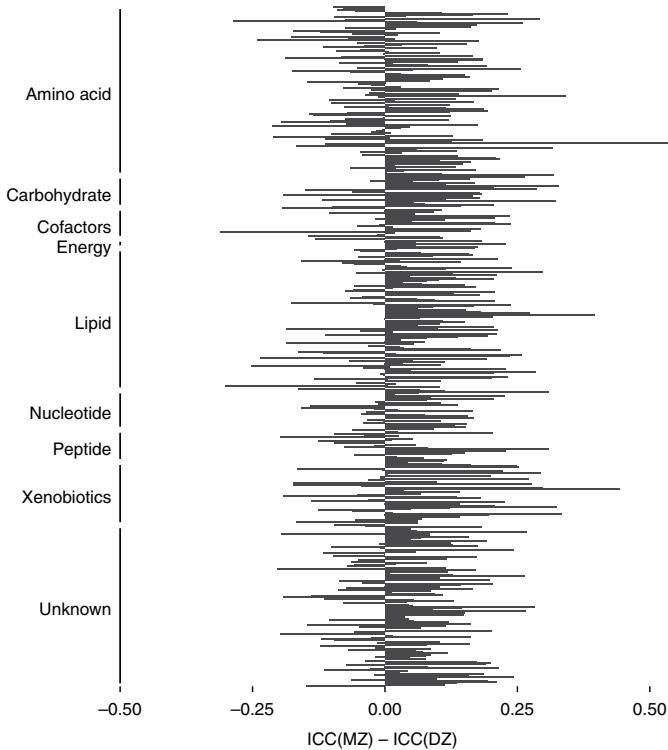


Fig. 4 | Intraclass correlation of fecal metabolites in monozygotic and dizygotic twins. The intraclass correlation coefficients (ICCs) were calculated from variance components of a one-way analysis of variance separately for MZ ($n=148$ pairs) and DZ ($n=155$ pairs) twins for each metabolite. Positive values of their respective differences indicate more similar metabolic profiles between MZ than DZ twins.

We then investigated the extent to which the fecal metabolome reflects metabolic processes of the gut microbiome. We regressed metabolite levels against microbial diversity (quantified with the Shannon index), and found that 575 metabolites across all pathways showed a significant association with microbial diversity at an FDR of 5%, 347 of which remained significant after Bonferroni correction. We estimated the proportion of variance in each metabolite

explained by microbiome composition by using the unweighted UniFrac beta-diversity metric, a measure of overall phylogenetic dissimilarity between individuals' microbiota⁹. We found that gut microbial composition explained a substantial proportion of the observed variance of 710 metabolites, 67.7% ($\pm 18.8\%$) of the observed variance on average, ranging from 22.1% for 1-linolenoyl-glycerol to 100% for several amino acids (Supplementary Table 2). The microbiome explained a significant proportion of the variance of the 8 BMI-related and 101 visceral-fat-related metabolites, among others. Xenobiotics showed the strongest associations with microbial composition (enrichment P value $< 1 \times 10^{-4}$), which explained the entire observed variance for some associations, including the B vitamins nicotinate and pantothenate.

To explore the associations between the fecal metabolome and gut microbes at a finer taxonomic resolution, we regressed each metabolite against the 581 most abundant OTUs, adjusting for potential confounding factors including Shannon diversity. We found 42,645 significant associations of 907 different metabolites with 579 different OTUs after adjusting for multiple testing (FDR $< 5\%$). We also calculated associations of fecal metabolites with collapsed taxonomical levels, ranging from the genus to the phylum level (Supplementary Table 4). We found that 264 metabolites were associated with microbes only at the OTU level, and the remainder were also associated with broader taxonomic groupings.

Finally, to investigate the connectivity of the fecal metabolome with microbes, we calculated a Gaussian graphical model combining 435 common metabolites with a known chemical identity with 241 OTUs with complete taxonomy assignment to at least the genus level. The resulting model consisted of 2,553 independent associations: 1,035 among metabolites, 946 among microbes, and 572 connecting metabolites and microbes (Supplementary Table 5 and Fig. 4). All but nine variables formed one connected component. We detected 19 clusters in the largest component, 9 of which contained both microbes and fecal metabolites, and 10 of which consisted of metabolites only. Xenobiotics had higher node degrees ($P < 3 \times 10^{-4}$) and were more densely connected with OTUs ($P < 2.4 \times 10^{-3}$). Our model demonstrated a high degree of interrelatedness between the gut microbiome and fecal metabolome, despite the very different technologies used.

In conclusion, although state-of-the-art metagenomic sequencing allows for quantitative and functional annotation of species and microbial pathways³⁰, 16S sequencing data have limitations, includ-

Table 1 | Genetic associations of fecal metabolites

Metabolite	H^2	MAF	SNP	Gene	Discovery			Replication		Meta-analysis	
					Effect	P	P gain	Effect	P	Effect	P
1,3-Dimethylurate/5-acetylamino-6-amino-3-methyluracil	40.2%	24.7%	rs35246381 (chr 8: 1841502)	NAT2	-0.17 ± 0.02	7.0×10^{-21}	7.5×10^9	-0.22 ± 0.03	3.5×10^{-10}	-0.18 ± 0.01	3.3×10^{-36}
3-Hydroxyhexanoate	20.7%	3.7%	rs62311177 (chr 4: 92962004)	GRID2	0.41 ± 0.06	2.9×10^{-12}		0.07 ± 0.09	0.429	0.32 ± 0.05	3.0×10^{-11}
Eicosapentaenoate (20:5 n3)	16.1%	1.4%	rs149572251 (chr 20: 34322936)	ITCH	1.45 ± 0.21	3.4×10^{-11}		-0.23 ± 0.38	0.544	1.06 ± 0.18	6.8×10^{-9}
3-Phenylpropionate (hydrocinnamate)	23.9%	1.6%	rs58539483 (chr 11: 28830501)	AC090833.1	-1.31 ± 0.19	3.6×10^{-11}		-0.39 ± 0.22	0.076	-0.92 ± 0.14	1.5×10^{-10}

Three metabolites and one metabolite ratio were significantly associated with genetic loci in the discovery cohort ($n=739$). We report their respective heritabilities (H^2), the associated variant along with its chromosomal position and the nearest gene, and the effect (\pm s.e.). P values in the discovery and replication cohorts were calculated with the two-sided score test implemented in GEMMA and combined through fixed-effects inverse-variance meta-analysis. The P gain describes the strength of the association of the metabolite ratio relative to the associations of each individual metabolite. MAF, minor allele frequency; chr, chromosome.

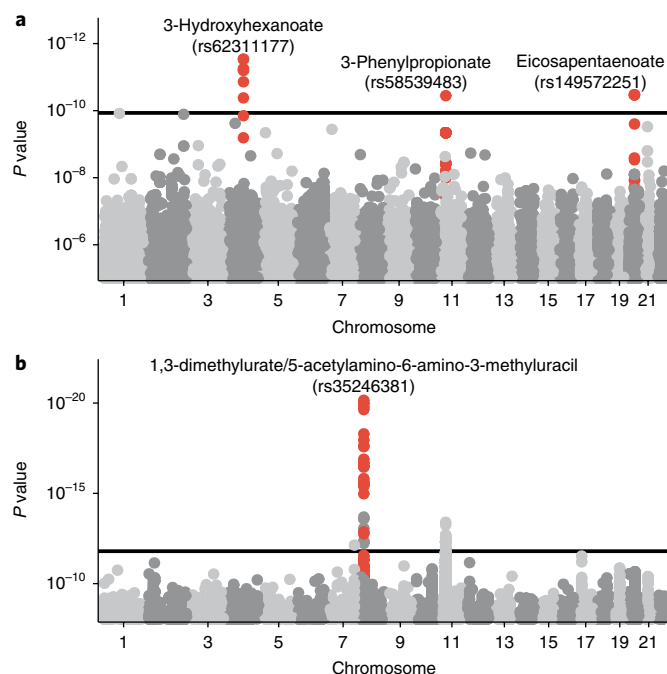


Fig. 5 | Host genetic influence on the fecal metabolome. Genome-wide association studies were conducted for 428 heritable fecal metabolites and 31,226 fecal metabolite ratios ($n = 739$). P values were calculated with two-sided score test implemented in GEMMA. **a**, Manhattan plot illustrating the genetic associations with fecal metabolites in the discovery sample. The horizontal line indicates the Bonferroni cutoff of 1.2×10^{-10} . Three loci (red) passed the Bonferroni threshold. **b**, Manhattan plot showing genetic associations with metabolite ratios in the discovery sample. The horizontal line indicates the Bonferroni cutoff of $P < 1.6 \times 10^{-12}$. Two loci passed the threshold; however, only the association with 1,3-dimethylurate/5-acetyluracil-6-amino-3-methyluracil ($P = 6.2 \times 10^{-21}$, red) passed filtering by P gain ($P \text{ gain} > 8.9 \times 10^5$) and thus was considerably stronger than the association of each individual metabolite. Box plots, quantile–quantile plots, and regional association plots for each of the four loci are shown in Supplementary Figs. 1 and 2, and Supplementary Data 1.

ing the lack of quantitative functional annotations. Fecal metabolomics provides a complementary functional readout of microbial metabolism as well as its interaction with host and environmental factors. We focused on the relationships among fecal metabolites and host and microbial genetics. Future studies should further investigate the influence of environmental factors, particularly nutrition, and should consider the influence of stool frequency and/or type on fecal metabolite measurements, which are known to be associated with fecal microbiome composition^{31,32}. The fecal metabolome can thus be used as intermediate phenotype that promotes microbial effects on the host and vice versa. Using associations with obesity, we demonstrated that fecal metabolomics is a useful tool to complement future genomic and microbiome studies.

URLs. NAT2 human tissue expression, http://bgee.org/?page=gene&gene_id=ENSG00000156006.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41588-018-0135-7>.

Received: 30 November 2016; Accepted: 5 April 2018;
Published online: 28 May 2018

References

- O'Hara, A. M. & Shanahan, F. The gut flora as a forgotten organ. *EMBO Rep.* **7**, 688–693 (2006).
- Frias-Lopez, J. et al. Microbial community gene expression in ocean surface waters. *Proc. Natl. Acad. Sci. USA* **105**, 3805–3810 (2008).
- Marcobal, A. et al. A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. *ISME J.* **7**, 1933–1943 (2013).
- Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–1023 (2006).
- Turnbaugh, P. J. et al. A core gut microbiome in obese and lean twins. *Nature* **457**, 480–484 (2009).
- Pedersen, H. K. et al. Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* **535**, 376–381 (2016).
- Clarke, G. et al. Minireview: Gut microbiota: the neglected endocrine organ. *Mol. Endocrinol.* **28**, 1221–1238 (2014).
- Cangelosi, G. A. & Meschke, J. S. Dead or alive: molecular assessment of microbial viability. *Appl. Environ. Microbiol.* **80**, 5884–5891 (2014).
- O'Toole, P. W. & Claesson, M. J. Gut microbiota: changes throughout the lifespan from infancy to elderly. *Int. Dairy J.* **20**, 281–291 (2010).
- Yatsunenko, T. et al. Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227 (2012).
- Romero-Corral, A. et al. Accuracy of body mass index in diagnosing obesity in the adult general population. *Int. J. Obes. (Lond)* **32**, 959–966 (2008).
- Arora, T. & Bäckhed, F. The gut microbiota and metabolic disease: current understanding and future perspectives. *J. Intern. Med.* **280**, 339–349 (2016).
- Parséus, A. et al. Microbiota-induced obesity requires farnesoid X receptor. *Gut* **66**, 429–437 (2017).
- Shoae, S. et al. Quantifying diet-induced metabolic changes of the human gut microbiome. *Cell Metab.* **22**, 320–331 (2015).
- Beaumont, M. et al. Heritable components of the human fecal microbiome are associated with visceral fat. *Genome Biol.* **17**, 189 (2016).
- Pallister, T. et al. Untangling the relationship between diet and visceral fat mass through blood metabolomics and gut microbiome profiling. *Int. J. Obes. (Lond)* **41**, 1106–1113 (2017).
- Goodrich, J. K. et al. Human genetics shape the gut microbiome. *Cell* **159**, 789–799 (2014).
- Goodrich, J. K. et al. Genetic determinants of the gut microbiome in UK Twins. *Cell Host Microbe* **19**, 731–743 (2016).
- Petersen, A.-K. et al. On the hypothesis-free testing of metabolite ratios in genome-wide and metabolome-wide association studies. *BMC Bioinformatics* **13**, 120 (2012).
- Weimann, A., Sabroe, M. & Poulsen, H. E. Measurement of caffeine and five of the major metabolites in urine by high-performance liquid chromatography/tandem mass spectrometry. *J. Mass Spectrom.* **40**, 307–316 (2005).
- Nyéli, A., Buclin, T., Biollaz, J. & Decosterd, L. A. NAT2 and CYP1A2 phenotyping with caffeine: head-to-head comparison of AFMU vs. AAMU in the urine metabolite ratios. *Br. J. Clin. Pharmacol.* **55**, 62–67 (2003).
- Shin, S.-Y. et al. An atlas of genetic influences on human blood metabolites. *Nat. Genet.* **46**, 543–550 (2014).
- Raffler, J. et al. Genome-wide association study with targeted and non-targeted NMR metabolomics identifies 15 novel loci of urinary human metabolic individuality. *PLoS Genet.* **11**, e1005487 (2015).
- GTEX Consortium. et al. Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).
- Zerbino, D. R. et al. Ensembl 2018. *Nucleic Acids Res.* **46**, D754–D761 (2018).
- Bastian, F. et al. *Data Integration in the Life Sciences* (Springer, Berlin and Heidelberg, 124–131, 2008).
- McDonagh, E. M. et al. PharmGKB summary: very important pharmacogene information for N-acetyltransferase 2. *Pharmacogenet. Genomics* **24**, 409–425 (2014).
- Meinl, W., Sczesny, S., Brigelius-Flohé, R., Blaut, M. & Glatt, H. Impact of gut microbiota on intestinal and hepatic levels of phase 2 xenobiotic-metabolizing enzymes in the rat. *Drug Metab. Dispos.* **37**, 1179–1186 (2009).
- Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**, 8228–8235 (2005).
- Keegan, K. P., Glass, E. M. & Meyer, F. MG-RAST, a metagenomics service for analysis of microbial community structure and function. *Methods Mol. Biol.* **1399**, 207–233 (2016).
- 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).
- Tigheelaar, E. F. et al. Gut microbiota composition associated with stool consistency. *Gut* **65**, 540–542 (2016).

Acknowledgements

The study was funded by the Wellcome Trust, European Community's Seventh Framework Programme (FP7/2007–2013). The study also received support from the

National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and the Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London, by the Chronic Disease Research Foundation and by the Denise Coates Foundation. HLI, Inc., collaborated with King's College London to produce [the metabolomics data from Metabolon](#) Inc. C.M. was funded by the MRC AIM HY (MR/M016560/1) project grant. We thank J. Goodrich and R. Ley for support in sequencing the fecal samples.

Author contributions

Conceived and designed the experiments: A.T., T.D.S., and C.M. Performed the experiments: R.P.M. Analyzed the data: J.Z., M.A.J., T.L., and C.M. Contributed reagents/materials/analysis tools: M.M., G.K., T.L., A.T., K.S.S., C.J.S., J.T.B., and A.M.V. Wrote the manuscript: J.Z., M.A.J., R.P.M., A.M.V., T.D.S., and C.M.. All authors revised the manuscript.

Competing interests

R.P.M. is an employee of Metabolon, Inc. T.L. and A.T. were employees of HLI, Inc. at the time this work was conducted. TDS is a co-founder of MapMyGut Ltd. All other authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41588-018-0135-7>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to T.D.S. or C.M.

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

Methods

Study population. The study participants were 786 twins from the TwinsUK cohort. TwinsUK (a national twin registry) has recruited subjects since 1992 through media campaigns and is representative of the population of the UK in terms of lifestyle³³. The study population is predominantly female (93.4% females) and has an average age of 65.2 (± 7.6) years and an average BMI of 26.1 (± 4.7). Ethical approval was granted by the St Thomas' Hospital ethics committee; all participants provided informed written consent.

Results of the genome-wide association study were replicated in an independent set of 230 individuals (98.3% female) from the TwinsUK study, 66.9 (± 8.6) years of age, with an average BMI of 27.2 (± 5.2) (Supplementary Table 1).

Data collection. Sample collection, DNA extraction, and sequencing of the samples in this study has been described previously^{17,18}. Briefly, the fecal samples were collected, refrigerated, and kept in ice packs until they were frozen at -80°C (mostly within 24 h after collection) before further processing. A number of participants (15%) sent their samples by post.

Metabolomics profiling. Metabolite concentrations were measured from fecal samples by Metabolon, Inc., Durham, North Carolina, USA, by using an **untargeted LC/MS** platform as previously described^{22,34}.

Sample preparation for global metabolomics. Samples were stored at -80°C until processing. Sample preparation was carried out as described previously³⁴ at Metabolon, Inc. Lyophilized fecal samples were extracted at a constant per-mass basis. Briefly, recovery standards were added before the first step in the extraction process for quality-control purposes. To remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills Genogrinder 2000), then centrifuged. The resulting extract was divided into five fractions: (i) acidic positive-ion conditions chromatographically optimized for more hydrophilic compounds; (ii) acidic positive-ion conditions chromatographically optimized for more hydrophobic compounds; (iii) basic negative-ion-optimized conditions with a separate dedicated C18 column; (iv) negative ionization after elution from a HILIC column; (v) reserved for backup.

Three types of controls were analyzed in concert with the experimental samples: a pooled sample generated from a small portion of each experimental sample of interest served as a technical replicate throughout the platform run; extracted water samples served as process blanks; and a cocktail of standards spiked into every analyzed sample allowed for instrument performance monitoring. Instrument variability was determined by calculation of the median relative s.d. (RSD) for the standards that were added to each sample before injection into the mass spectrometers (median RSDs were determined to be 5%; $n = 31$ standards). Overall process variability was determined by calculating the median RSD for all endogenous metabolites (i.e., noninstrument standards) present in 90% or more of the pooled technical-replicate samples (median RSD = 12%, $n = 832$ metabolites). Experimental samples and controls were randomized across the platform run.

Mass spectrometry analysis. Extracts were subjected to **UPLC-MS/MS**³⁵. The chromatography was standardized, and no further changes were made after the method was validated. As part of Metabolon's general practice, all columns were purchased from a single manufacturer's lot at the outset of experiments. All solvents were similarly purchased in bulk from a single manufacturer's lot in sufficient quantity to complete all related experiments. For each sample, vacuum-dried samples were dissolved in injection solvent containing eight or more injection standards at fixed concentrations, depending on the platform. The internal standards were used to ensure both injection and chromatographic consistency. Instruments were tuned and calibrated for mass resolution and mass accuracy daily.

All methods used a Waters Acquity UPLC and a Thermo Scientific Q-Exactive high-resolution/accurate mass spectrometer interfaced with a **heated electrospray ionization (HESI-II)** source and an Orbitrap mass analyzer operated at 35,000 mass resolution. The sample extract was dried, then reconstituted in solvents compatible with each of the four methods. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed by using acidic positive-ion conditions, which were chromatographically optimized for relatively hydrophilic compounds. In this method, the extract was gradient eluted from a C18 column (Waters UPLC BEH C18, 2.1×100 mm, $1.7 \mu\text{m}$) with water and methanol containing 0.05% perfluoropentanoic acid and 0.1% formic acid. Another aliquot was also analyzed by using acidic positive-ion conditions; however, it was chromatographically optimized for relatively hydrophobic compounds. In this method, the extract was gradient eluted from the same aforementioned C18 column with methanol, acetonitrile, water, 0.05% perfluoropentanoic acid, and 0.01% formic acid, and was operated at an overall higher organic content. Another aliquot was analyzed by using basic negative-ion-optimized conditions and a separate dedicated C18 column. The basic extracts were gradient eluted from the column with methanol and water, as well as 6.5 mM ammonium bicarbonate at pH 8. The fourth aliquot

was analyzed via negative ionization after elution from a HILIC column (Waters UPLC BEH Amide 2.1×150 mm, $1.7 \mu\text{m}$) with a gradient consisting of water and acetonitrile with 10 mM ammonium formate, pH 10.8. The MS analysis alternated between MS and data-dependent MSn scans using dynamic exclusion. The scan range varied slightly between methods but covered 80–1,000 m/z .

Compound identification, quantification, and data curation. Metabolites were identified by automated comparison of the ion features in the experimental samples to a reference library of chemical standard entries that included **retention time, molecular weight (m/z), preferred adducts, and in-source fragments** as well as associated MS spectra, and were curated by visual inspection for quality control in software developed at Metabolon^{35,36}. Identification of known chemical entities was based on comparison to **metabolomic library** entries of purified standards. Commercially available purified standard compounds have been acquired and registered into LIMS for distribution to the various **UPLC-MS/MS platforms** for determination of their detectable characteristics. Additional mass-spectral entries have been created for structurally unnamed biochemicals, which have been identified on the basis of their recurrent nature (both chromatographic and mass spectral). These compounds have the potential to be identified by future acquisition of a matching purified standard or by **classical structural analysis**. Peaks were quantified through **area-under-the-curve analysis**. Raw area counts for each metabolite in each sample were normalized to correct for variation resulting from instrument interday tuning differences by the median value for each run day, and the medians were therefore set to 1.0 for each run. This procedure preserved variation among samples but allowed metabolites of widely different raw peak areas to be compared on a similar graphical scale.

A total of 1,116 different metabolites were measured in the 786 fecal samples, of which 210 metabolites were observed in less than 20% of the samples and thus were excluded from further analysis because of lack of power. 345 metabolites were observed in more than 20% but less than 80% of the samples and were thus analyzed qualitatively as dichotomous traits (observed in a sample versus not observed). The remaining 570 metabolites, which were observed in at least 80% of all samples, were scaled by **run-day medians, log-transformed and scaled to uniform mean 0 and s.d. 1 and analyzed quantitatively** (Fig. 1). Metabolite ratios were calculated from **the run-day median-normalized metabolite levels** and subsequently log-transformed and scaled to a mean of 0 and s.d. of 1.

We analyzed effects of sample storage time (i) in the refrigerator before samples were frozen and (ii) in the freezer before further analysis. To this end, we regressed metabolite concentrations against storage times. After correcting for multiple testing, we found significant storage effects on seven metabolites (FDR < 0.05 ; Supplementary Fig. 5). We thus corrected all further analyses for both storage time in the refrigerator and freezer, to avoid spurious results. Despite correcting all models for the storage time, we cannot ultimately eliminate a potential confounding effect due to storage time, and future studies should investigate the influence of storage time on fecal metabolites.

Microbial sequencing. 16S rRNA was extracted from fecal samples, PCR amplified, barcoded per sample and sequenced with the **Illumina MiSeq platform**, as previously described¹⁸. DNA was isolated from the samples with a PowerSoil kit. The V4 region of bacterial 16S rRNA gene sequences was PCR amplified with the 515F and 806R primers³⁷. Reads were barcoded per sample and combined for multiplexed sequencing with an Illumina MiSeq instrument to generate 250-bp paired-end reads. Paired-end reads were merged with a minimum overlap of 200 nt by using **fastq join within QIIME**, which was also used to demultiplex the sequence data.

Preprocessing of sequences and their clustering to operational taxonomic units followed **the Sumacust de novo approach**, as previously described for a subset of samples from ref. ³⁸. In brief, **16S sequence reads** for all samples were filtered to remove chimeric sequences produced during PCR, by using **USEARCH for chimera identification**³⁹. The remaining reads were then collapsed to OTUs by using the **Sumacust algorithm for de novo clustering in QIIME 1.9.0**; this method selected group reads at the 97% level as de novo OTUs⁴⁰, and Sumacust has been found to be one of the best-performing greedy clustering algorithms in previous comparisons³⁸. Taxonomy was assigned to OTUs on the basis of alignment of representative sequences to the Greengenes 13_8 database with a 97% similarity threshold in UCLUST.

De novo clustering across all samples within the TwinsUK cohort produced ~300,000 OTUs after singleton removal; however, most of those were of low abundance and found in very few samples (table density 0.002). We subdivided the OTU table by discarding samples with fewer than 10,000 reads and OTUs that were not found in at least 25% of these samples. This procedure resulted in a table of 581 OTUs (table density 0.547) (Supplementary Table 7). OTU counts were converted to relative abundance values (over all reads in each sample), a pseudocount of 10^{-6} was added to account for zero counts, and the abundance values were log-transformed. The transformed abundance values were then used as the response in models with sequencing run, sequencing depth, individual who extracted the DNA, individual who loaded the DNA, and sample-collection method as technical covariates. The residuals of these models were then used in downstream analysis of OTU abundance values. The same normalization and control for technical effects

was also carried out on taxonomic abundance values collapsed at each taxonomic level. Collapsed taxonomies included counts from all OTUs. The Shannon alpha diversity was also calculated from the complete OTU table. Each sample was rarefied to a depth of 10,000 reads 50 times. Diversity metrics were calculated for each sample in each table, and the mean across all tables was taken as the final measure. Beta diversity was calculated from all OTUs, and singletons were excluded with the unweighted UniFrac algorithm³⁹.

Visceral-fat measurements. Measurements of whole-body composition were performed with DXA fan-beam technology (Hologic QDR; Hologic, Inc.) as previously described⁴¹. Briefly, subjects with their clothes removed and wearing gowns were positioned in a standardized manner, in a supine position. The DXA machine was calibrated on a daily basis by using a spine phantom and on a weekly basis by using a step phantom, as suggested by the manufacturer. The scans were analyzed in QDR System Software, Version 12.6.

Regions of interest were defined manually by the same operator following the SOP, which was derived from the manufacturer's guidelines. The lower horizontal margins were placed above the pelvis, just above the iliac crest, and the upper horizontal margins were placed at the half of the distance between the acromions and the iliac crest. The vertical margins were adjusted just at the external borders of the body so that all the soft tissue was included.

This DXA-based measurement of visceral fat has been validated against visceral fat measured by CT scans and shown to be reliable and reproducible⁴².

Statistical analysis. To assess the influence of age and sex on metabolite measurements, we regressed all metabolites against age and sex, correcting for family structure as a random intercept, in the R package lme4 (ref. ⁴³). Moreover, we calculated linear and logistic regression models to assess the relationship of the fecal metabolome with obesity, measured as BMI and visceral-fat mass (measured by double X-ray absorptiometry), and adjusted for age, sex, storage time, and family as the random intercept. Visceral-fat measurements were available for 647 individuals. The following regression models were used:

1. Associations of metabolites with age and sex
Metabolite ~ age + sex + storage time (refrigerator) + storage time (freezer) + (1 | family)
2. Associations with BMI
BMI ~ metabolite + age + sex + storage time (refrigerator) + storage time (freezer) + (1 | family)
3. Associations with visceral-fat mass
visceral fat ~ metabolite + age + sex + BMI + height + height² + storage time (refrigerator) + storage time (freezer) + (1 | family)
4. Associations with microbial alpha diversity
Metabolite ~ microbiome diversity + age + sex + BMI + storage time (fridge) + storage time (freezer) + (1 | family)
5. Associations with OTUs and taxa
Metabolite ~ microbe (OTU/taxa) + microbiome diversity + age + sex + BMI + storage time (refrigerator) + storage time (freezer) + (1 | family)

Partial-least-squares discriminant analysis. We used a partial-least-squares discriminant analysis (PLS-DA) to investigate global differences between the metabolic profiles of the youngest and oldest deciles of our study population. To this end, all missing values were imputed in the mice package⁴⁴, and metabolite levels were adjusted for storage times and family structure with linear mixed models. The residuals of these models were then used to train a PLS-DA model, as implemented in the mixOmics package⁴⁵. The predictive performance was assessed with a tenfold cross validation.

Heritability analysis. We used structural equation modeling to estimate the genetic (A), common environment (C), and unique environment (E) components of the total variance for each metabolite⁴⁶. To this end, we used the R package mets (version 1.1.1) to fit maximum-likelihood models, adjusting for age, sex, and storage time. For each metabolite, we fitted four models, estimating (i) A, C, and E components; (ii) A and E components; (iii) C and E components; and (iv) the E component only. The best model was selected by minimizing the Akaike information criterion. In the case of dichotomous metabolite abundance, a liability-threshold model was fitted by using the bptwin function in the mets package. Additionally, ICCs were calculated from variance components of a one-way analysis of variance for MZ and DZ twins individually, in the ICC package⁴⁷.

Genome-wide association study. Genetic variation was measured through whole-genome sequencing, as previously described⁴⁸. In brief, samples were sequenced on an Illumina HiSeqX sequencer with 150-base paired reads. Reads were then mapped to the hg38 genome in ISIS Analysis Software (v. 2.5.26.13; Illumina), and missing genotypes were filled in with reference homozygous calls⁴⁹. Genomes with a ratio of heterozygous to homozygous variants higher than 2.5 were excluded, thus leaving 739 individuals for further analysis. A cohort-based high-confidence region of the genome was constructed by concatenating positions with a 'pass' call rate greater than 90%, by using data from 3 sets of 100 randomly selected genomes.

Variants outside the high-confidence region and duplicated variants were removed. We moreover excluded 273,355 variants with Hardy-Weinberg $P < 10^{-6}$, calculated from 420 unrelated individuals, thus leaving 8,208,502 biallelic SNPs and 1,408,051 indels with MAF greater than 1% for further analysis.

We fitted linear mixed models to test for associations of heritable fecal metabolites with genetic variants, correcting for age, sex, and storage time, in GEMMA⁵⁰ by incorporating data from 739 individuals with fecal metabolomics and sequencing data. The twin structure of our data was taken into account by adjusting for the family relatedness by using the sample kinship matrix. The score test implemented in GEMMA was used to assess the significance of the associations. We considered metabolite associations with a P value lower than 1.2×10^{-10} significant, a threshold corresponding to a genome-wide-significance cutoff of 5.0×10^{-8} , corrected for 428 tested metabolites. Additionally, we tested for genetic associations with all pairwise metabolite ratios of fecal metabolites with known chemical identity and a heritable variance component. We used the P -gain statistic to assess the independence of the single metabolites⁵¹. The P gain is defined as the minimal P value of the associations of either of the single metabolites divided by the P value of the metabolite ratio. A high P -gain statistic indicates that the ratio carries additional information beyond that of individual metabolites. We considered metabolite ratios with $P < 1.6 \times 10^{-12}$ ($5 \times 10^{-8}/31,226$ metabolite ratios) and P gain $> 3.1 \times 10^5$ ($10 \times 31,226$ metabolite ratios) significant.

Four genome-wide-significant associations were replicated in 230 individuals from the TwinsUK study, with adjustment for the same confounding factors. The discovery and replication results were combined through fixed-effects inverse-variance meta-analysis.

Associations of the fecal metabolome with the gut microbiome. To assess the associations of the fecal metabolome with the gut microbiome, we first regressed metabolite concentrations against the Shannon alpha diversity, adjusting for age, sex, BMI, storage time, and family structure, by using 644 individuals with both fecal metabolomics and 16S sequencing data available.

We then estimated the proportions of variance of each metabolite explained by the microbiome by regressing the fecal metabolite concentration against the microbial beta diversity. This technique is commonly used to estimate heritability from genetic kinship matrices^{51,52}. To this end, we calculated a restricted maximum-likelihood model, regressing the metabolite level against the microbial similarity, adjusting for age, sex, BMI, storage time in the refrigerator and freezer, and technical covariates (sequencing run, sequencing depth, individual who extracted the DNA, individual who loaded the DNA, and sample-collection method), in the R package regress. The proportion of variance explained by microbial similarity (M^2) and its standard error were calculated from the variance components in the R package gap⁵³, and P values were calculated from the ratio of M^2 to standard error.

Next, we sought to identify microbes and taxonomical units that were associated with metabolite levels. To this end, we regressed 581 inverse-normalized OTUs against all 915 metabolites, adjusting for age, sex, BMI, sample storage times, family structure, and alpha diversity. Benjamini-Hochberg correction was applied to account for multiple testing. We further calculated associations at different taxonomical units, from the genus to the phylum level.

Finally, to assess multivariate dependencies between the fecal metabolome and the microbiome, we inferred a graphical model combining 423 metabolites with known chemical identity that were observed in at least 80% of the samples with 241 OTUs that were assigned complete taxonomy at least to the genus level. Sparse graphical models were inferred in the GeneNet package⁵⁴, and edges with FDR < 0.05 were included in the model. We used the Fruchterman-Reingold algorithm⁵⁵ to determine an unbiased graph layout and identified network modules by optimizing the modularity score, as implemented in the igraph package⁵⁶.

Pathway enrichment. We used pathway annotation as provided by Metabolon for pathway enrichment with the page algorithm. Enrichment P values were estimated with permutation tests with 10,000 random permutations, as implemented in the R package piano⁵⁷.

Graphical-model inference. To assess multivariate dependencies between the fecal metabolome and the microbiome, we inferred a graphical model combining 435 metabolites with known chemical identity that were observed in at least 80% of the samples with 241 OTUs that were annotated at least down to the genus level. To obtain a full data matrix, we first imputed missing metabolite levels in the program mice⁵⁸. A sparse graphical model was inferred with the GeneNet algorithm⁵⁴, by selecting edges with FDR < 0.05 . We used the Fruchterman-Reingold algorithm⁵⁵ to determine an unbiased graph layout and identified network modules by optimizing the modularity score, as implemented in the igraph package⁵⁶. For visualization purposes, we collapsed all metabolites of the same pathway into one node, which was connected to all neighbors of all its members. Microbes were collapsed by family. For both the full and the collapsed networks, we calculated several centrality measures of nodes, including node degrees, as the number of neighbors; clustering coefficients, as the proportion of neighbors that were connected; and betweenness centralities, as the proportion of shortest paths incorporating a node.

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

Data availability. 16S sequencing data used for this study have been deposited in the European Nucleotide Archive under accession code [ERP015317](https://www.ebi.ac.uk/ena/record/ERP015317). All other TwinsUK data are available upon reasonable request from the department website (<http://www.twinsuk.ac.uk/data-access/accessmanagement/>).

References

33. Moayyeri, A., Hammond, C. J., Valdes, A. M. & Spector, T. D. Cohort profile: TwinsUK and healthy ageing twin study. *Int. J. Epidemiol.* **42**, 76–85 (2013).
34. Evans, A. et al. High resolution mass spectrometry improves data quantity and quality as compared to unit mass resolution mass spectrometry in high-throughput profiling metabolomics. *J. Postgenomics Drug Biomark. Dev.* **4**, S24–S36 (2014).
35. Evans, A. M., DeHaven, C. D., Barrett, T., Mitchell, M. & Milgram, E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal. Chem.* **81**, 6656–6667 (2009).
36. Dehaven, C. D., Evans, A. M., Dai, H. & Lawton, K. A. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *J. Cheminform.* **2**, 9 (2010).
37. Caporaso, J. G. et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. USA* **108**, 4516–4522 (2011).
38. Jackson, M. A., Bell, J. T., Spector, T. D. & Steves, C. J. A heritability-based comparison of methods used to cluster 16S rRNA gene sequences into operational taxonomic units. *PeerJ* **4**, e2341 (2016).
39. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200 (2011).
40. Westcott, S. L. & Schloss, P. D. De novo clustering methods outperform reference-based methods for assigning 16S rRNA gene sequences to operational taxonomic units. *PeerJ* **3**, e1487 (2015).
41. Menni, C. et al. Metabolomic profiling to dissect the role of visceral fat in cardiometabolic health. *Obesity (Silver Spring)* **24**, 1380–1388 (2016).
42. Kaul, S. et al. Dual-energy X-ray absorptiometry for quantification of visceral fat. *Obesity (Silver Spring)* **20**, 1313–1318 (2012).
43. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 51 (2015).
44. van Buuren, S. & Groothuis-Oudshoorn, K. mice: multivariate imputation by chained equations in R. *J. Stat. Softw.* **45**, 1–67 (2011).
45. Dejean, S. et al. *mixOmics: Omics Data Integration Project*, <http://mixomics.org/> (2013).
46. Neale, M. & Cardon, L. *Methodology for Genetic Studies of Twins and Families* (Springer Netherlands, Houten, the Netherlands, 1994).
47. Wolak, M. E., Fairbairn, D. J. & Paulsen, Y. R. Guidelines for estimating repeatability. *Methods Ecol. Evol.* **3**, 129–137 (2012).
48. Long, T. et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nat. Genet.* **49**, 568–578 (2017).
49. Telenti, A. et al. Deep sequencing of 10,000 human genomes. *Proc. Natl. Acad. Sci. USA* **113**, 11901–11906 (2016).
50. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012).
51. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
52. Speed, D., Hemani, G., Johnson, M. R. & Balding, D. J. Improved heritability estimation from genome-wide SNPs. *Am. J. Hum. Genet.* **91**, 1011–1021 (2012).
53. Zhao, J. H. gap: genetic analysis package. *J. Stat. Softw.* **23**, 1–18 (2007).
54. Schaefer, J., Opgen-Rhein, R. & Strimmer, K. *GeneNet: Modeling and Inferring Gene Networks*, <https://CRAN.R-project.org/package=GeneNet> (2014).
55. Fruchterman, T. M. J. & Reingold, E. M. Graph drawing by force-directed placement. *Software-Practice Exp.* **21**, 1129–1164 (1991).
56. Csardi, G. & Nepusz, T. *The igraph Software Package for Complex Network Research* (InterJournal Complex Systems 1695, 2006).
57. Väre, L., Nielsen, J. & Nookaew, I. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Res.* **41**, 4378–4391 (2013).

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

► Experimental design

1. Sample size

Describe how sample size was determined.

This is the first study to examine fecal metabolites, therefore there is no prior knowledge on the effect sizes we are to expect. However, having included 786 individuals with fecal metabolites, anthropometric, microbial and GWAS data our study has 80% power to detect a difference of 0.175 SD for the anthropometric measures adjusting for multiple testing (FDR<0.05) and 80% power to detect a difference of 0.40 for GWAS metabolites ratio adjusting for 31226 tested ratios. The largest effect size found by us in the GWAS is -0.17 for a metabolite ratio with mean=0 SD=0.30 (i.e. an effect of 0.56 SD) which is larger than the effect needed to achieve 99% power for a $p < 2 \times 10^{-12}$ needed for the GWAS of metabolite ratios (adjusting for 31226 tested ratios)

2. Data exclusions

Describe any data exclusions.

No data were excluded from the analysis

3. Replication

Describe whether the experimental findings were reliably reproduced.

We replicated our GWAS-metabolomics results in an independent sample of 230 individuals

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

This is not a clinical trial and so randomization is not relevant

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

This is not a clinical trial and so blindness is not relevant

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or the Methods section if additional space is needed).

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- ☐ ☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.
- ☐ ☒ A statement indicating how many times each experiment was replicated
- ☐ ☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- ☐ ☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- ☐ ☒ The test results (e.g. p values) given as exact values whenever possible and with confidence intervals noted
- ☐ ☒ A summary of the descriptive statistics, including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- ☐ ☒ Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

R version 3.3.3
R package mets (version 1.1.1) to estimate heritability
GEMMA for the GWAS

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* [guidance for providing algorithms and software for publication](#) may be useful for any submission.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

N/A

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

N/A

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

N/A

b. Describe the method of cell line authentication used.

N/A

c. Report whether the cell lines were tested for mycoplasma contamination.

N/A

d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

N/A

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

N/A

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

We included 786 individuals, predominantly females (93.4%), with an average age of 65.2 (± 7.6) and an average of BMI of 26.1 (± 4.7) from the TwinsUK cohort. Results of the genome-wide association study were replicated in an independent set of 230 individuals (98.3% female) from the TwinsUK study, aged 66.9 (± 8.6) and an average BMI of 27.2 (± 5.2). Metabolite concentrations were measured from fecal samples by Metabolon Inc., Durham, USA, using an untargeted LC/MS platform. 16S rRNA microbial sequencing was extracted from fecal samples using Illumina MiSeq.

Genotypes were determined using whole-genome sequencing Ethical approval by St Thomas' Hospital ethics committee; all participants provided informed written consent.