



Investigating the Role of Diet and Exercise in Gut Microbe-Host Com metabolism

 N. Penney,^a W. Barton,^{b,c,d}  J. M. Posma,^{e,f} A. Darzi,^a  G. Frost,^g  P. D. Cotter,^{b,c}  E. Holmes,^g F. Shanahan,^{b,d} O. O'Sullivan,^{b,c}  I. Garcia-Perez^g

^aDivision of Surgery, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, United Kingdom

^bAPC Microbiome Ireland, University College Cork, National University of Ireland, Cork, Ireland

^cTeagasc Food Research Centre, Moorepark, Co. Cork, Ireland

^dDepartment of Medicine, University College Cork, National University of Ireland, Cork, Ireland

^eSection of Bioinformatics, Division of Systems Medicine, Department of Metabolism, Digestion and Reproduction, Faculty of Medicine, Imperial College London, London, United Kingdom

^fHealth Data Research UK, London, United Kingdom

^gSection for Nutrition Research, Division of Digestive Diseases, Department of Metabolism, Digestion and Reproduction, Faculty of Medicine, Imperial College London, London, United Kingdom

N. Penney, W. Barton, O. O'Sullivan, and I. Garcia-Perez contributed equally. Author order was decided as N. Penney led the manuscript write-up.

ABSTRACT We investigated the individual and combined effects of diet and physical exercise on metabolism and the gut microbiome to establish how these lifestyle factors influence host-microbiome com metabolism. Urinary and fecal samples were collected from athletes and less active controls. Individuals were further classified according to an objective dietary assessment score of adherence to healthy dietary habits according to WHO guidelines, calculated from their proton nuclear magnetic resonance (¹H-NMR) urinary profiles. Subsequent models were generated comparing extremes of dietary habits, exercise, and the combined effect of both. Differences in metabolic phenotypes and gut microbiome profiles between the two groups were assessed. Each of the models pertaining to diet healthiness, physical exercise, or a combination of both displayed a metabolic and functional microbial signature, with a significant proportion of the metabolites identified as discriminating between the various pairwise comparisons resulting from gut microbe-host com metabolism. Microbial diversity was associated with a combination of high adherence to healthy dietary habits and exercise and was correlated with a distinct array of microbially derived metabolites, including markers of proteolytic activity. Improved control of dietary confounders, through the use of an objective dietary assessment score, has uncovered further insights into the complex, multifactorial relationship between diet, exercise, the gut microbiome, and metabolism. Furthermore, the observation of higher proteolytic activity associated with higher microbial diversity indicates that increased microbial diversity may confer deleterious as well as beneficial effects on the host.

IMPORTANCE Improved control of dietary confounders, through the use of an objective dietary assessment score, has uncovered further insights into the complex, multifactorial relationship between diet, exercise, the gut microbiome, and metabolism. Each of the models pertaining to diet healthiness, physical exercise, or a combination of both, displayed a distinct metabolic and functional microbial signature. A significant proportion of the metabolites identified as discriminating between the various pairwise comparisons result from gut microbe-host com metabolism, and the identified interactions have expanded current knowledge in this area. Furthermore, although increased microbial diversity has previously been linked with health, our observation of higher microbial diversity being associated with increased proteolytic

Citation Penney N, Barton W, Posma JM, Darzi A, Frost G, Cotter PD, Holmes E, Shanahan F, O'Sullivan O, Garcia-Perez I. 2020. Investigating the role of diet and exercise in gut microbe-host com metabolism. *mSystems* 5:e00677-20. <https://doi.org/10.1128/mSystems.00677-20>.

Editor Janet K. Jansson, Pacific Northwest National Laboratory

Copyright © 2020 Penney et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to O. O'Sullivan, orla.osullivan@teagasc.ie, or I. Garcia-Perez, i.garcia-perez@imperial.ac.uk.

Received 18 July 2020

Accepted 4 November 2020

Published 1 December 2020

activity indicates that it may confer deleterious as well as beneficial effects on the host.

KEYWORDS diet, exercise, metabolism, microbiome

Recent studies have shown that increased physical activity and aerobic fitness may mediate some health benefits through modulation of the gut microbiome (1–4). Human studies have shown differences in the gut microbiome, including increased diversity, in both habitual exercisers (5, 6) and professional athletes (7–10) compared to more sedentary controls. Longitudinal studies have also found exercise-dependent compositional and functional changes in the gut microbiome that result in altered levels of a number of microbially derived bioactive metabolites such as short-chain fatty acids (SCFA) as well as improved glucose homeostasis (7, 11, 12).

However, understanding the relationship between the gut microbiome, diet, and exercise remains elusive in part because of the dietary adaptations that accompany habitual physical activity. Difficulties understanding this relationship are compounded by inaccuracies in assessing dietary habits, with 30 to 88% misreporting using traditional self-reported assessment tools (13–15). In an earlier report, taxonomic and functional diversity were positively correlated with both dietary protein intake and physical exercise in a population of athletes (7, 8). We hypothesized that more accurate dietary assessment would help separate the contributions of diet and exercise to the modulation of the gut microbiota and enhance our understanding of the contribution of diet and exercise to the modulation of microbe-host cometabolism. To achieve this, we applied metabolic profiling combined with a mathematical modeling strategy to provide objective evidence of adherence to World Health Organization (WHO) healthy eating guidelines (increased fruits, vegetables, whole grains, and dietary fiber and decreased fats, sugars, and salt) (16, 17). Our approach characterizes volunteers more accurately according to physical activity and dietary status. By minimizing dietary variation in volunteer subsets the study has uncovered insights into the distinct influences of physical activity and dietary status on host metabolism, the gut microbiome, and subsequent gut microbiome-host cometabolism. In addition, we have explored the functional and metabolic implications of higher microbial diversity, previously reported in athletes (7, 8), on the host.

RESULTS

To assess the separate contributions of diet and exercise in modulating the gut microbiota, we investigated urine and fecal samples from male athletes ($n = 40$) in the national Irish Rugby Football team that were collected while attending an intensive training camp and healthy controls ($n = 46$) matched for age and gender, as previously described (8). We used a previously validated, metabolic profiling dietary assessment tool (17) to objectively assess and score participants' dietary habits (Fig. 1). Higher calculated scores indicate more complete adherence to WHO dietary guidelines of increased fruits, vegetables, whole grains, and dietary fiber and decreased fat, sugar, and salt consumption. Briefly, the dietary profiling model was built using proton nuclear magnetic resonance ($^1\text{H-NMR}$) global urinary metabolic profiles derived from healthy participants that attended an inpatient randomized controlled crossover trial and were assigned to diets with differing levels of adherence to WHO healthy eating guidelines (16, 17) (see Materials and Methods).

As expected, controls had diverse adherence to a healthy diet, while professional athletes were predominantly assigned scores that met or exceeded the healthy eating guidelines with healthy eating scores of $\geq 100\%$, reflecting the known association between habitual exercise and good dietary habits (18). The scores for healthy eating behavior for the less active control participants ranged from 12.8% to 112.2%, with a relatively even distribution between the two extremes, while those of the athletes ranged from 15.7% to 112.5%, with only 5 of the 40 professional athlete participants

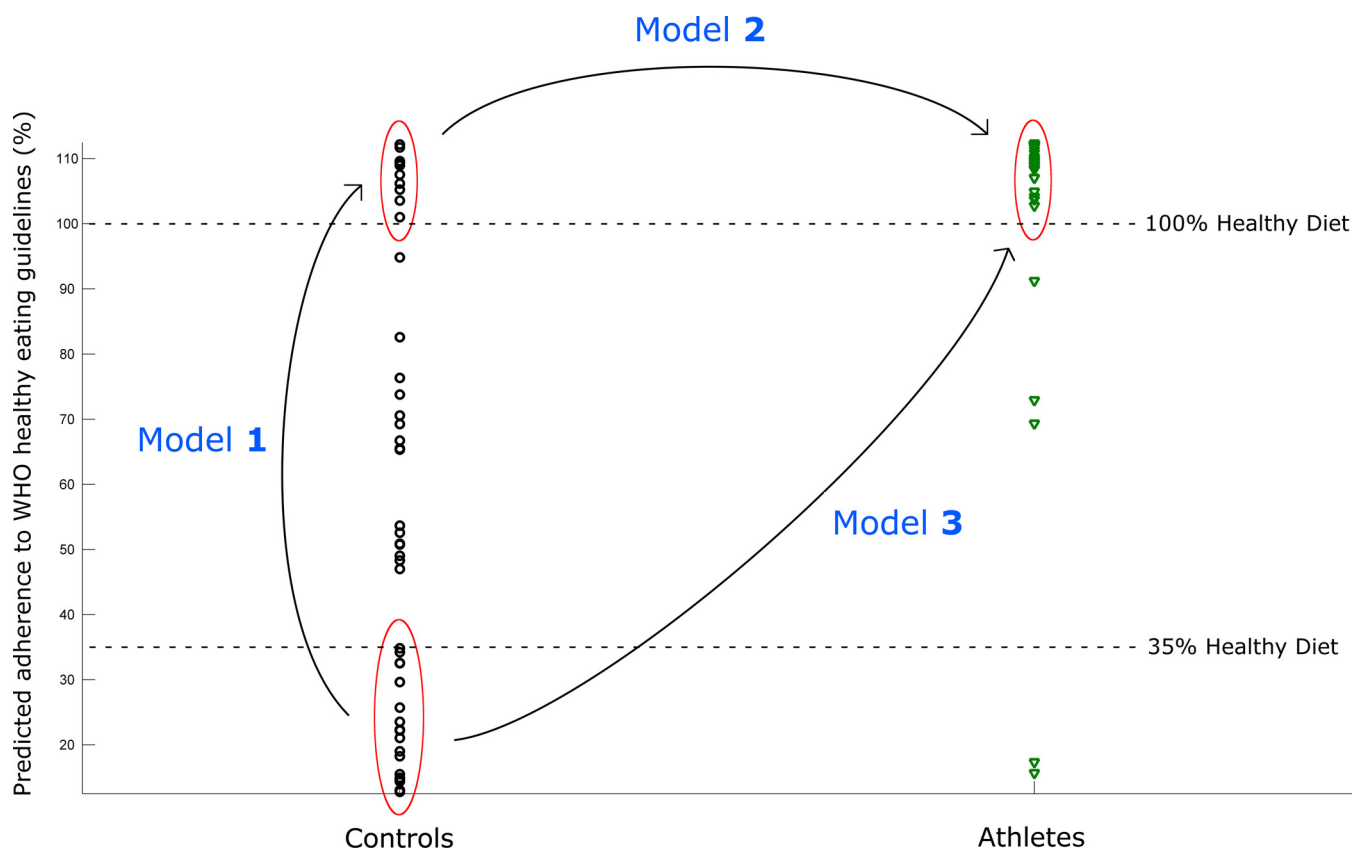


FIG 1 Clustering of individuals according to exercise status and adherence to healthy eating guidelines. Predicted adherence to WHO healthy eating guidelines calculated from $^1\text{H-NMR}$ urinary profile of each individual using a validated metabolic profiling tool (see Materials and Methods for dietary assessment methodology). Individuals were subsequently clustered to form comparative models. Model 1 (healthy diet effect): comparing controls with a $<35\%$ adherence versus controls with a $\geq 100\%$ adherence to WHO healthy eating guidelines. Model 2 (effect of exercise): controls with a $\geq 100\%$ adherence versus professional athletes with a $\geq 100\%$ adherence to WHO healthy eating guidelines. Model 3 (combined diet and exercise effect): controls with a $<35\%$ adherence versus athletes with a $\geq 100\%$ adherence to WHO healthy eating guidelines.

falling below the 100% boundary. Individuals were then stratified according to their score to construct comparative groups (Fig. 1). Participants with a scored adherence to WHO healthy eating guidelines of $\geq 100\%$ were considered healthy eaters. A cluster of controls was observed with a healthy eating adherence score of $<35\%$. There is evidence that those at the bottom third of a healthy eating scale have a 25% higher all-cause mortality than those in the top third, as well as a 40% higher mortality from cardiovascular disease (19). Therefore, 35% was chosen as the cutoff to define unhealthy eaters. We used these groupings to construct a variety of models to statistically measure the differences in host metabolism, the gut microbiome, and their subsequent cometabolism between various groups (Fig. 1). Model 1 (healthy diet effect) was constructed comparing controls with a $<35\%$ adherence ($n=19$) versus controls with a $\geq 100\%$ adherence to healthy eating ($n=11$); model 2 (exercise effect) compared controls with a $\geq 100\%$ adherence to healthy eating ($n=11$) versus athletes with a $\geq 100\%$ adherence to healthy eating ($n=35$); model 3 (combined healthy diet and exercise effect) compared controls with a $<35\%$ adherence to healthy eating ($n=19$) versus athletes with a $\geq 100\%$ adherence to healthy eating ($n=35$). Other lifestyle factors, including smoking levels and alcohol consumption, were also measured. There were no significant differences in these factors between groups.

Diet and exercise status were associated with distinct urinary and fecal metabolomes. Based on these comparative groups, six statistically robust Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) models (Fig. 2A to F) were obtained to assess the effects of diet, exercise, and the two in combination on participants' urinary and fecal metabolic profiles, measured by $^1\text{H-NMR}$. The urinary

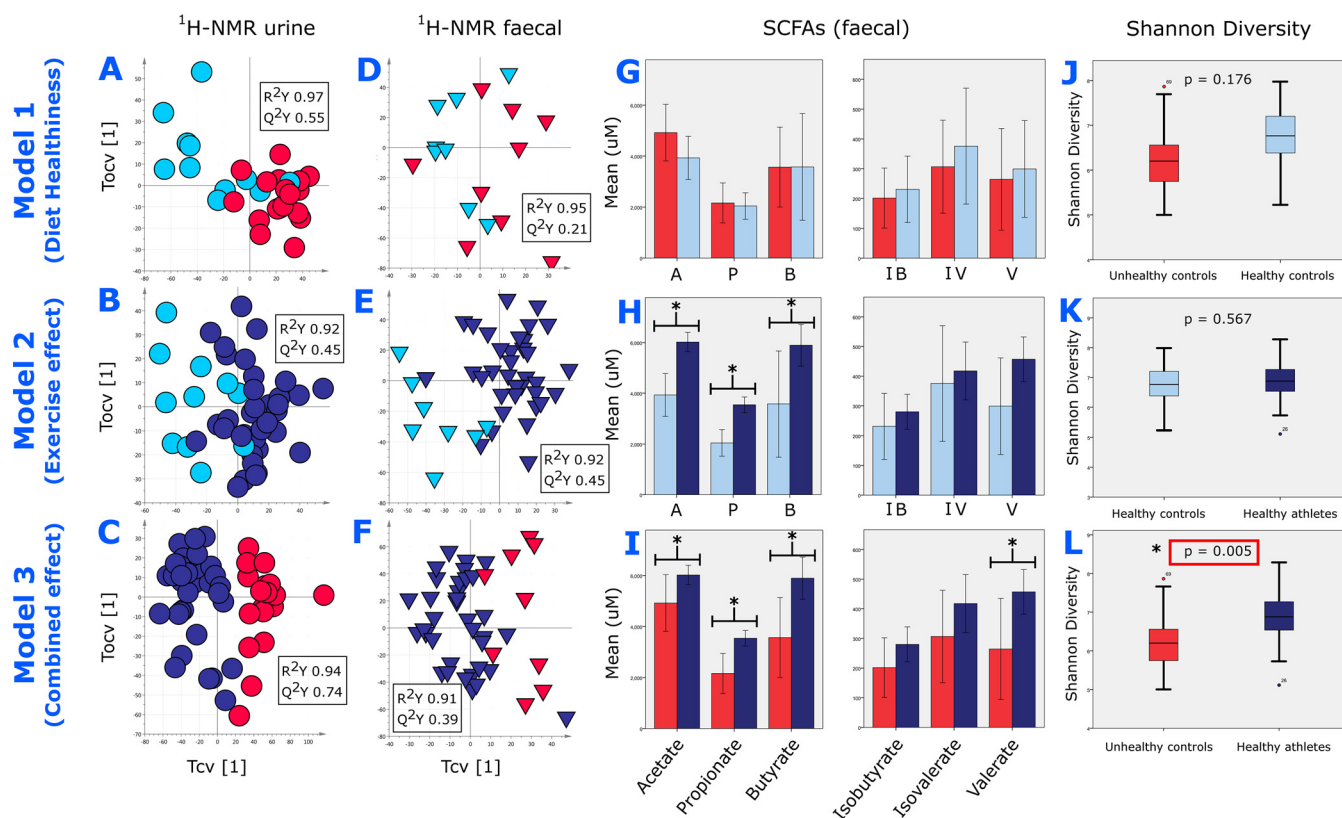


FIG 2 ¹H-NMR metabolic phenotyping, SCFA, and microbial diversity results. (A to F) Cross-validated OPLS-DA score plots, generated with one predictive (Tcv) and one orthogonal (Tocv) component, of ¹H-NMR urinary profiles comparing (A) controls with a <35% adherence (red) to controls with a ≥100% adherence (light blue) to WHO healthy eating guidelines (model 1, healthy diet effect), (B) controls with a ≥100% adherence to healthy eating (light blue) to professional athletes with a ≥100% adherence to healthy eating (dark blue) (model 2, exercise effect), (C) controls with a <35% adherence to healthy eating (red) versus professional athletes with a ≥100% adherence to healthy eating (dark blue) (model 3, combined effect), and ¹H-NMR fecal profiles comparing (D) model 1, healthy diet effect, (E) model 2, exercise effect, and (F) model 3, combined effect. (G to I) Bar charts of mean SCFA levels measured through quantitative GC-MS comparing (G) model 1, healthy diet effect, (H) model 2, exercise effect, and (I) model 3, combined effect. (J to L) Box plots showing mean Shannon diversity index levels of taxa described by 16S profiling comparing (J) model 1, healthy diet effect, (K) model 2, exercise effect, and (L) model 3, combined effect. The 95% confidence intervals shown. Significant results (pFDR, <0.05) are marked with *.

model characterizing the combined effect of diet healthiness and exercise was strongest (model 3; Fig. 2C, $R^2_Y = 0.94$, $Q^2_Y = 0.74$), but statistically robust urinary models were also obtained defining the impact of diet healthiness (model 1; Fig. 2A, $R^2_Y = 0.97$, $Q^2_Y = 0.55$) and exercise alone (model 2; Fig. 2B, $R^2_Y = 0.92$, $Q^2_Y = 0.45$). With regard to fecal profiles, the strongest model was that defining the effect of exercise (model 2; Fig. 2E, $R^2_Y = 0.92$, $Q^2_Y = 0.45$). The combination of diet healthiness and exercise produced a relatively robust model (model 3; Fig. 2F, $R^2_Y = 0.91$, $Q^2_Y = 0.39$), while investigation of diet alone showed the least impact (model 1; Fig. 2D, $R^2_Y = 0.95$, $Q^2_Y = 0.21$).

For each pairwise OPLS-DA model built with the global ¹H-NMR metabolic profiles, a number of metabolites discriminating between the two classes were identified from the model coefficients. These are listed in Table S1 (urine) and Table S2 (feces). As expected, in model 1 (healthy diet effect), urinary metabolites with well-known dietary associations such as proline-betaine (found in citrus fruit, particularly oranges [20]) and hippurate (associated with the consumption of fruits and vegetables) were significantly higher in the group with healthier eating scores, whereas markers of oxidative stress (2-hydroxybutyrate) and anaerobic metabolism (lactate) were found in lower concentrations. Fecal metabolic profiles in subjects with healthier diet scores were associated with lower concentrations of 2-aminobutyrate and 2-hydroxybutyrate and higher levels of markers of cruciferous vegetable intake (*S*-methyl-L-cysteine-sulfoxide derivatives). In the exercise effect model (model 2), the athlete group was characterized by a number of urinary metabolites derived from the gut microbiome, such as

phenylacetylglutamine (PAG) and 3-indoxyl sulfate (3IS), as well as the ketone body acetoacetate and markers of red meat intake and fatty-acid (FA) metabolism (carnitine, *O*-acetyl carnitine). In contrast, markers of oxidative stress (allantoin [21]) and tricarboxylic acid (TCA) cycle intermediates (succinate and citrate) were higher in the control group. Analysis of fecal metabolites revealed higher concentrations of the short-chain fatty acids (SCFAs) acetate, propionate, butyrate, and valerate and lower levels of the amino acids glycine, phenylalanine, and tyrosine in the athlete group compared to controls consuming a comparatively healthy diet. Model 3 (combined effect) consisted predominantly of metabolites derived from models 1 and 2. Healthy eating, exercise, and the combined effect of both (models 1 to 3) were all characterized by higher markers of choline metabolism, with higher urinary trimethylamine-*N*-oxide (TMAO). In the exercise effect model (model 2), this corresponded with higher fecal methylamine and trimethylamine and lower dimethylamine in athletes relative to controls.

Targeted gas chromatography-mass spectrometry (GC-MS) of SCFA in feces revealed significantly higher levels of acetate (Benjamini-Hochberg false-discovery rate [pFDR] < 0.001), propionate (pFDR < 0.001), and butyrate (pFDR = 0.006) in healthy eating athletes relative to healthy eating controls (model 2; Fig. 2H, exercise effect). In addition to propionate (pFDR = 0.004) and butyrate (pFDR = 0.019), valerate (pFDR = 0.043) was found in higher concentrations in healthy eating athletes relative to controls with unhealthy eating profiles, while acetate was higher but not significant after Benjamini-Hochberg (BH) multiple testing corrections (pFDR = 0.056) (model 3; Fig. 2I, combined effect). There were no significant differences in the diet healthiness model (model 1; Fig. 2G) and no differences in isobutyrate or isovalerate concentrations in any model.

Targeted analysis of 10 organic acids in urine by GC-MS demonstrated lower levels of lactate (pFDR = 0.002) and 2-hydroxybutyrate (pFDR = 0.014) in healthy eating controls versus controls with unhealthy eating habits (model 1, healthy diet effect). 2-Methylbutyrate (pFDR = 0.043) was lower in healthy eating athletes versus healthy eating controls (model 2, exercise effect). In contrast to the fecal SCFA model, urinary acetate (pFDR = 0.006), propionate (pFDR = 0.003), 2-methylbutyrate (pFDR = 0.001), isovalerate (pFDR = 0.001), and lactate (pFDR < 0.001) were all lower in athletes with $\geq 100\%$ adherence to healthy eating versus controls with <35% adherence to healthy eating (model 3, combined effect).

Microbial diversity. No difference in gut microbial diversity was found between comparative groups when investigating the effect of healthy eating (model 1; Fig. 2J) or exercise (Model 2; Fig. 2K). However, there was increased diversity of genera detected by 16S profiling in healthy eating athletes versus unhealthy eating controls, both in richness and evenness (model 3; Fig. 2L, combined effect) using Shannon (pFDR = 0.012), Simpson (pFDR = 0.020), whole-tree (phylogenetic diversity [PD]; pFDR = 0.001), Chao1 (pFDR = 0.025), and observed feature (pFDR = 0.012) indices. Similar comparisons for each of the models with microbial metabolic pathways, species, and genera identified with metagenomic sequencing were not significantly different despite having similar trends (data not shown).

Linear regression analysis of ¹H-NMR fecal water global profiles against Shannon, Simpson, and PD diversity indices, corrected for confounding factors (age, lean mass, and fat mass), revealed multiple metabolites associated with diversity (see Table S2). In contrast, no metabolites from urinary ¹H-NMR global profiles were found to significantly correlate with measures of gut diversity. Fecal metabolites associated with higher diversity included increased short/medium-chain FAs valerate and caproate; branched SCFAs isobutyrate, isovalerate, and 2-methylbutyrate; branched-chain amino acid (BCAA) degradation products 2-oxoisocaproate and 2-oxoisovalerate; as well as products of phenylalanine (phenylacetate and 3-phenylpropionate), choline (dimethylamine), proline (2-methylproline), and uracil (ureidopropionate) metabolism. Increased diversity was also associated with lower glucose, isoleucine, asparagine, and histidine levels. See Table S2 for corrected *P* values (pFDR).

Diet-gut microbiome-metabolome interactions. A total of 47 fecal and 27 urinary metabolites of interest were subsequently quantified. Selection of these metabolites was based on discriminatory capacity (model weighting) from models 1 to 3 and/or

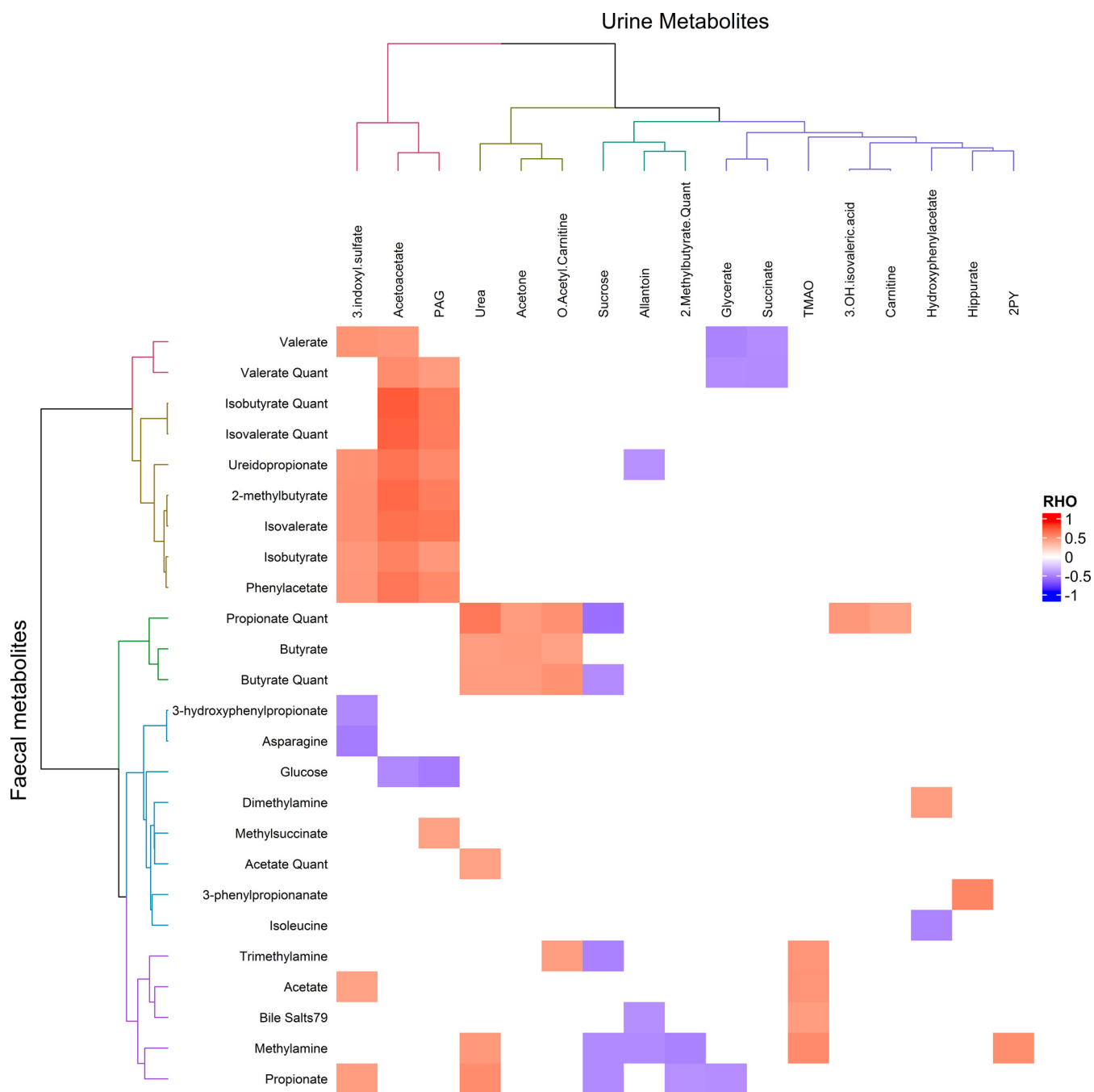


FIG 3 Faecal-urinary metabolic interactions. Significant Spearman correlations ($pFDR, <0.05$) between faecal and urinary metabolic data sets are shown, shaded according to the strength of the correlation coefficient (Rho). Correlations are clustered according to Euclidean distances.

metabolites associated with microbial diversity. A number of correlation matrices were generated using these quantified metabolites to explore gut metabolome-host metabolome interactions. First, faecal-urinary metabolome interactions were derived from Spearman correlations between the two data sets (see Fig. 3). Findings included the correlation of urinary 3IS with faecal SCFAs, urinary PAG with faecal phenylacetate, urinary hippurate with faecal 3-phenylpropionate, and urinary TMAO with faecal trimethylamine and methylamine. Second, Spearman correlations between metabolite levels and relative abundance of (i) microbial species (Fig. S1) and (ii) microbial metabolic pathways (Fig. 4) were explored, revealing numerous significant correlations ($pFDR < 0.01 = 257$ and 58 correlations, respectively). These included the correlation of

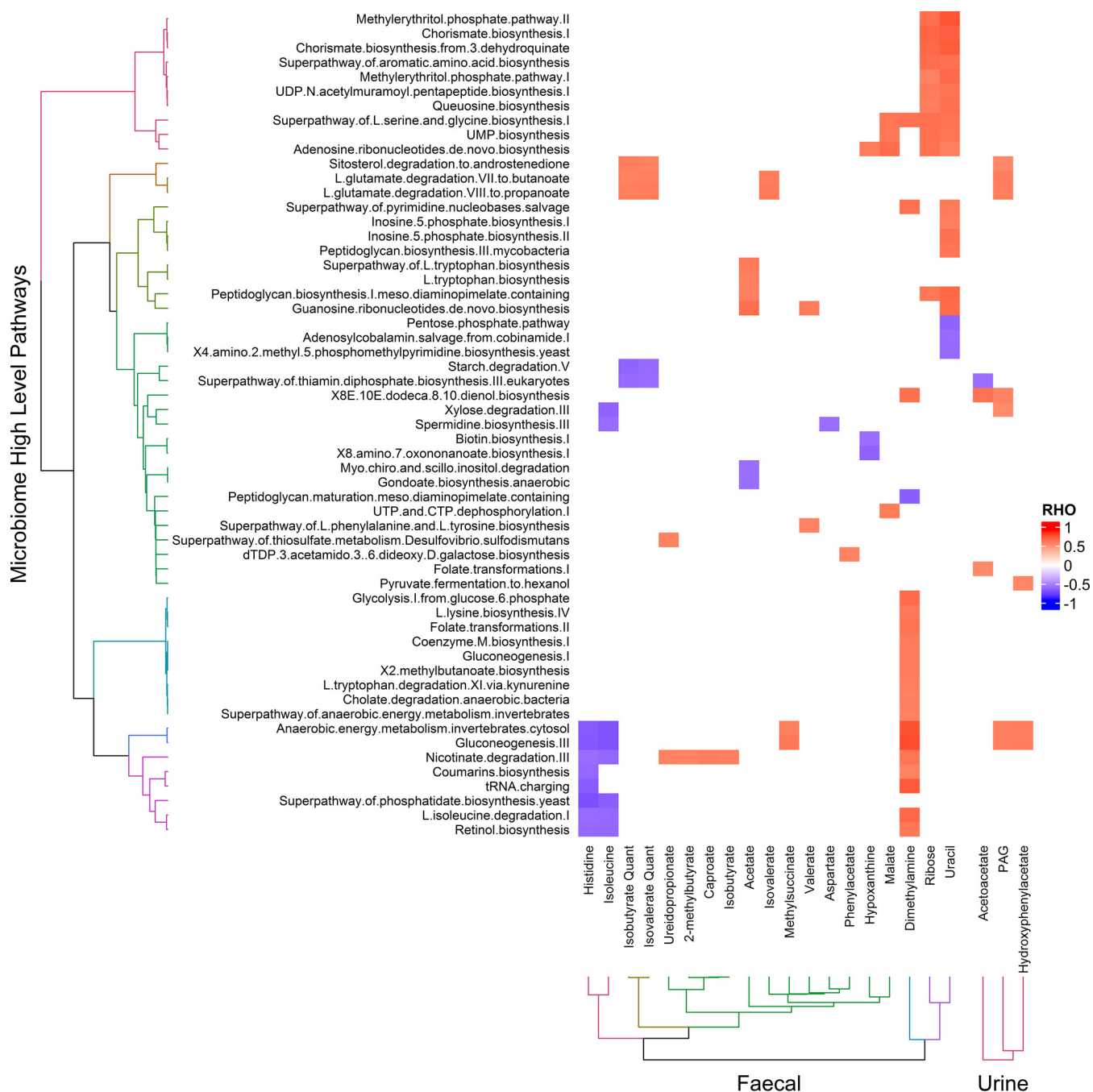


FIG 4 Microbial metabolic pathway-metabolite interactions. Significant Spearman correlations ($pFDR, <0.01$) between microbial metabolic pathways versus faecal and urinary metabolic data sets are shown, shaded according to the strength of the correlation coefficient (Rho). Correlations are clustered according to Euclidean distances.

urinary PAG levels with bacterial species from the *Actinobacteria* and *Proteobacteria* phyla as well as a number of bacterial metabolic pathways, including those related to gluconeogenesis, anaerobic energy metabolism, and glutamate degradation to SCFAs. We found a number of positive correlations between bacterial species and pathways with SCFAs, including the correlation of acetate with *Roseburia hominis*, *Streptococcus suis*, and *Halobacteriovorax marinus*, propionate with *Leuconostoc citreum*, and valerate with *Pyrococcus horikoshii*, *Simkania negevensis*, and *Streptococcus suis*. We identified a large number of species and pathways whose abundance correlated with faecal dimethylamine (DMA) excretion but which did not appear to influence TMAO levels. Whereas the

heterofermentative lactic acid bacterium *Leuconostoc citreum* did correlate with higher TMAO levels. Interestingly, we also identified a number of lactococcal phages that correlated with TMAO. Furthermore, lactococcal phage levels correlated with increased TMAO:DMA and TMA:DMA ratios. Further correlations between dietary, metabolite, microbial pathway, and microbial diversity data sets were also performed (Fig. S2 to S5). Fecal butyrate correlated strongly with fiber intake, whereas propionate correlated most strongly with protein intake. Diversity was correlated with a number of metabolites, including valerate, medium-chain fatty acids (MCFA), branched SCFA, and branched-chain keto acids (BCKA), as well as lower amino acid levels. Increased diversity also correlated with a number of microbial pathways, including higher amino acid, BCAA, and nicotinate degradation, cofactor and amino acid biosynthesis, and gluconeogenesis pathways, along with a reduced activity in starch degradation, methionine, and fatty acid biosynthesis pathways.

DISCUSSION

Models of diet healthiness, exercise, or both each displayed metabolic and functional microbial signatures, with a number of discriminatory metabolites. Microbial diversity was associated with a combination of increased diet healthiness and exercise and also correlated with distinct microbially derived metabolites. Our results suggest that dietary changes alone did not significantly impact bacterial diversity, although they did affect gut microbial metabolites in urine and feces.

Metabolic phenotype associated with diet and exercise. As expected, the effect of adherence to WHO healthy eating guidelines was associated with higher markers of fruit and vegetable intake (urinary proline-betaine, formate, and hippurate in addition to lower sugar levels) (17, 22). Healthy eating was also characterized by lower urinary levels of lactate and 2-hydroxybutyrate, markers of anaerobic respiration and oxidative stress, respectively (22).

High levels of physical activity were linked with higher concentrations of *O*-acetyl carnitine, a metabolite associated with the intake of animal protein. Although protein consumption was higher in the athlete group and included whey protein supplementation in several individuals, higher *O*-acetyl carnitine may also have been due to increased fatty acid mobilization and oxidation (23). In addition, there were lower concentrations of urinary succinate in athletes, likely due to increased systemic utilization through the TCA cycle.

Diet and exercise induced alterations in gut microbe-host cometabolism. Tryptophan metabolism. In control participants assigned high healthy eating scores (model 1), we found lower levels of *N*-methylnicotinamide (NMND). NMND is generated from tryptophan metabolism through the kynurenine pathway in humans. The resulting nicotinamide is methylated by nicotinamide *N*-methyltransferase (NNMT), to form NMND, which is then further metabolized to 2-methyl-2-pyridone-5-carboxamide (2PY) (24). Nicotinamide serves an important role in energy regulation. Indeed, NNMT knockout in mice—resulting in higher nicotinamide and reduced NMND and 2PY—is protective of diet-induced obesity (25). Thus, the lower levels of NMND seen here suggest increased energy expenditure in healthy eaters.

Conversely, in the exercise effect model we find higher levels of urinary 2PY associated with physical activity (model 2). In addition, there is increased gut microbial metabolism of tryptophan through the indole pathway, leading to increased 3IS. This is a similar metabolic signature to that seen in malnourished children (24). Increased 2PY can be explained through indoleamine 2,3-dioxygenase (IDO)-mediated activation of the kynurenine pathway due to increased inflammation (8). Increased 3IS is likely due to a shift toward proteolytic fermentation of increased colonic protein by gut bacteria due to higher protein intake in healthy eating athletes relative to healthy eating controls (median intake, 240 g/day versus 104 g/day; $P < 0.001$). We also found that urinary 3IS was positively correlated with fecal SCFAs (Fig. 3), suggesting it was linked to increased fermentation of other substrates, including dietary fiber and complex carbohydrates.

Phenylalanine metabolism. Higher physical activity (model 2) and the combined effect of healthy diet and exercise (model 3) resulted in higher levels of urinary

phenylacetylglutamine (PAG), produced through gut microbial conversion of phenylalanine to phenylacetate and subsequent conversion to PAG in the liver. In keeping, fecal phenylacetate was positively correlated with urinary PAG (Fig. 3), and we observed lower fecal phenylalanine and tyrosine concentrations in model 2. Consistent with these findings, PAG has previously been associated with a lean phenotype (26). Our study highlights a number of correlations between PAG levels and bacterial species and metabolic pathways. The ketone body acetoacetate was similarly associated with higher physical activity and the combined effect of healthy diet and exercise and can also be produced through ketogenic metabolism of phenylalanine.

While PAG was positively associated with higher physical activity (model 2), the more minor microbial phenylalanine metabolites 3-hydroxyphenylacetate (HPA) and 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA) were inversely correlated. Whereas healthy eating (model 1) was positively correlated with HPA levels in urine. The biological role of HPA and HPHPA is disputed. Some studies have associated HPA and HPHPA with neurological disorders (27). Others, however, have shown the contrary (28), and furthermore, HPA is thought to be a dietary marker of rutin intake, a flavonoid and antioxidant.

Benzoate metabolism. An alternative metabolic route for phenylalanine and other dietary aromatic compounds, such as catechin is via their metabolism to benzoate, which is further glycine conjugated predominantly in the liver to form hippurate (29). Here, we found higher urinary hippurate in individuals with a healthy diet score (model 1). Higher hippurate levels have been correlated with a lean phenotype and lower blood pressure as well as fruit intake (30, 31). In this study, urinary hippurate levels correlated strongly with fecal 3-phenylpropionate (Fig. 3), suggesting that the main source of hippurate production observed here was through microbial degradation of catechin from the diet.

Choline metabolism. We found higher levels of TMAO in both those that adhered to healthy eating guidelines (model 1) and those with high levels of physical activity (model 2) in addition to the combination of both. Circulating TMAO is thought to predict cardiovascular disease (CVD), possibly through altered cholesterol metabolism and oxidative stress (32, 33). However, the causative effect in humans is disputed and may be due to confounders such as kidney function and poor metabolic control (34). Furthermore, high concentrations of TMAO are present in the tissues of cold-water-dwelling fish, where it acts as an antifreeze agent, and is consequently also found in high levels in the urine of Japanese populations whose diet contains a high portion of fish and who do not have high risk for CVD (30). Moreover, TMAO was recently found to protect against impaired glucose tolerance and reduce endoplasmic reticulum stress (35).

As expected, urinary TMAO levels correlated with fecal trimethylamine and methylamine, although not with dimethylamine (DMA) (Fig. 3). Furthermore, TMAO correlated with markers of protein intake: urinary carnitine, acetyl-carnitine, and urea (Fig. S5), which is consistent with the higher reported intake of protein in the healthy eating athletes represented in the exercise and combined models. The microbial abundance data also demonstrated a dichotomy with respect to association between TMAO and DMA. We identified a large number of species and pathways whose abundance correlated with fecal DMA excretion but which did not appear to influence TMAO levels, whereas the heterofermentative lactic acid bacterium *Leuconostoc citreum* did correlate with higher TMAO levels. Intriguingly, the identification of a number of lactococcal phages that correlated with TMAO raises the possibility that changes to the gut environment, such as altered pH from changes to lactic acid bacteria, can alter levels of TMA/TMAO-producing and -metabolizing bacteria, leading to subsequent changes in host TMAO levels. Indeed, lactococcal phage levels also correlated with increased TMAO:DMA and TMA:DMA ratios, suggesting they may influence bacterial TMA dehydrogenase or TMAO aldolase activity.

SCFA and branched-chain amino acid metabolism. Fecal SCFAs were found in

higher concentrations in participants with increased physical activity (model 2) and participants with combined healthy eating and increased exercise (model 3) relative to controls. SCFAs are produced through the bacterial fermentation of dietary fiber and complex carbohydrates. The most abundant SCFAs are acetate (C2), propionate (C3), and butyrate (C4). Valerate and the branched SCFAs isobutyrate, isovalerate, and 2-methyl-butyrate are produced by gut bacteria in smaller quantities (36). Butyrate is the preferred source of energy for colonic epithelial cells (37), whereas propionate, along with remaining butyrate, is used predominantly by hepatocytes for gluconeogenesis (38). Acetate is mainly utilized by muscle cells to generate energy. Importantly, SCFAs have been associated with reduced appetite and weight loss through stimulating release of the anorectic gut hormones peptide YY (PYY) and glucagon-like peptide-1 (GLP-1) (36, 39). Interestingly, it has been reported that undigested proteins and amino acids in the colon may serve as an additional substrate for SCFA production (40). Indeed, in this cohort, we found that while fecal butyrate correlated strongly with fiber intake, propionate correlated most strongly with protein intake, which is increased in healthy eating athletes (Fig. S2). We found a number of positive correlations between bacterial species and SCFAs (Fig. S1). These included the correlation of acetate with *Roseburia hominis*, *Streptococcus suis*, and *Halobacteriovorax marinus*, propionate with *Leuconostoc citreum*, and valerate with *Pyrococcus horikoshii*, *Simkania negevensis*, and *Streptococcus suis*.

Medium-chain fatty acids (MCFA), branched SCFA, and branched-chain keto acids (BCKA) were correlated with higher microbial diversity, seen after modeling the combined effects of healthy diet and exercise. The branched SCFAs isobutyrate, isovalerate, and 2-methylbutyrate are produced through the fermentation of BCAAs and have also been shown to modulate energy metabolism (41). Here, isobutyrate and isovalerate were both correlated with bacterial amino acid (glutamate) degradation pathways to propionate and butyrate and had an inverse correlation to pathways encoding starch degradation (Fig. 4). It is likely that bacteria were using similar pathways to ferment BCAAs in place of glutamate. Further, in fecal samples with higher diversity, we detected higher levels of 2-oxoisocaproate and 2-oxoisovalerate, BCKAs produced through the catabolism of the BCAAs leucine and valine, respectively. This initial transamination reaction, catalyzed by branched-chain aminotransferase, produces glutamate (also higher in fecal samples with higher diversity) and the respective BCKA (42). Subsequent leucine metabolism via 2-oxoisocaproate is ketogenic, while valine metabolism via 2-oxoisovalerate is glucogenic (43). This constellation of raised BCKAs, glutamate, and branched-SCFAs suggests that the increased diversity observed here was associated with an adaptation toward increased degradation of protein and amino acids in the colon. Additionally, we observed lower glucose and amino acid levels (histidine, asparagine, and isoleucine) with higher levels of diversity. It is not clear whether these changes were a result of microbial adaptation to a high protein/low sugar environment in the distal colon or if they were driven by an increase in microbial utilization of amino acids. Intriguingly, the lower fecal amino acid levels, seen here with higher microbial diversity, indicate that enhanced protein degradation increased beyond any increased protein intake. Corresponding changes were also noted in microbial metabolic pathways related to increased diversity, including higher abundance of higher amino acid, BCAA, and nicotinate degradation pathways (Fig. S3).

In addition, SCFAs can also be utilized by bacteria for *de novo* amino acid biosynthesis (44). Here, increased fecal valerate correlated with increased exercise (models 2 and 3) and microbial diversity and was positively correlated with bacterial phenylalanine and tyrosine biosynthesis pathways (Fig. 4). Furthermore, increased fecal acetate also correlated with increased exercise (models 2 and 3) and was positively correlated with bacterial tryptophan biosynthesis pathways. Lastly, fecal acetate and valerate were positively associated with guanosine ribonucleotide biosynthesis pathways. These correlations suggest that SCFAs have an important role in enabling the biosynthesis of aromatic compounds.

Conclusion. We have shown that each exercise/diet group exhibited a distinct metabolic and functional microbial phenotype. The incorporation of an objective measure of dietary habits into the models enabled us to more accurately reduce the confounding effect of diet when investigating exercise and therefore better ascertain the individual biological sequelae of diet healthiness and exercise. Limiting the participants to males removed confounders resulting from gender. However, further studies are needed to confirm the reproducibility of these results in females. In addition, we were unable to isolate the effect of diet in participants with high physical activity levels due to a paucity of athletes that follow an unhealthy diet.

Although increased microbial diversity has previously been linked with health (45–48), increased diversity in the present cohort was associated with amplified proteolytic fermentation by gut bacteria. Since protein metabolism within the gut has been linked with the production of toxic compounds such as ammonia, amines, and sulfides (49, 50), increased gut bacterial diversity may include deleterious as well as beneficial effects on the host. This emphasizes the value of in-depth analysis at a metabolic pathway and taxonomic level in future microbiome studies.

MATERIALS AND METHODS

Study population. Professional male athletes from the national Irish Rugby Football team ($n=40$) and healthy controls ($n=46$) matched for age and gender were enrolled in this study as previously described (8). Exclusion criteria included receiving antibiotics within 2 months before screening and prior diagnosis with any cardiovascular, gastrointestinal, or immunological condition. Control subgroups were established with body mass index (BMI) ranges matching the body types of the athletes. Approval was granted by the Cork Clinical Research Ethics Committee. Urine and fecal samples were collected from both groups and stored at -80°C until analysis.

Acquisition of clinical exercise and dietary data. Urine and fecal samples were collected from athletes while they attended an intensive training camp and from healthy but less active controls. Activity levels were validated using the EPIC-Norfolk questionnaire (51) and creatine kinase levels. Dietary intake was calculated from food frequency questionnaires (FFQ) administered by a research nutritionist (8).

Gut microbiota analysis. As determined previously, 16S taxonomic profiles had more robust characterization of taxa at and above the level of genus (7). Due to the removal of samples falling outside healthy diet thresholds, species and strain-level designations were unreliably varied within the study groups. As a result, the 16S data were used in order to more conservatively represent alpha diversity above the species level, whereas metagenomic sequencing was used to establish species and pathway data.

DNA preparation. Extraction and purification of DNA from fresh fecal samples were accomplished with the QIAmp DNA stool minikit (Qiagen, UK). The manufacturer's protocol for the extraction kit was followed with the addition of a bead-beating step ($30\text{s} \times 3$) to better disrupt cell walls. The resulting DNA was initially stored at -20°C prior to 16S rRNA sequencing before being stored at -80°C until the samples were prepared for shotgun metagenomics sequencing.

16S rRNA sequencing. The complete procedures used for 16S rRNA gene amplicon sequencing were outlined previously (8). Briefly, 16S rRNA gene (v4) PCR amplicons were generated using a combination of universal 16S rRNA primers estimated to bind to 94.6% of all 16S genes. Sample-specific identifier tags and 454 adaptor sequences were combined with the primers. The AMPure magnetic bead purification system was used to clean both 16S amplicons and shotgun metagenomics libraries (Beckman Coulter; catalogue number 9A63880). The 16S amplicons were sequenced on a 454 genome sequencer FLX platform at the Teagasc sequencing facility using the manufacturer's protocols. The 16S sequencing data were originally processed accordingly as follows: quality trimming of raw 16S sequences was done with the Qiime (v1.2) (52) software suite, using the SILVA 16S rRNA database (v106) (53). Reads were discarded from further analysis if they fell below a minimum quality score of 25 or were shorter than 150 bp. BLAST with default parameters was used with the SILVA database to generate input for METaGenome ANalyzer (MEGAN v4.70.4), which assigned reads to taxonomies. Following read clustering to operational taxonomical units (OTUs) and chimera removal, Qiime was then used to generate measures of alpha diversity. Although the methods to process 16S sequencing data are continuously evolving, minimal changes have occurred with OTU assignment and the calculation of alpha diversity measures. Due to this, diversity indices as previously reported were used to describe alpha diversity above the species level (8).

Shotgun metagenomic sequencing. As described previously (7), metagenomic library preparation was performed with the Illumina Nextera XT DNA library preparation kit (Illumina, Inc., USA) in explicit accordance with the manufacturer's protocol (15031942, Illumina). Prior to library preparation, DNA samples were normalized to $0.2\text{ ng}/\mu\text{l}$ using the Qubit v2.0 fluorometric quantification system (Thermo Fisher Scientific). Library fragment size was assessed with the Agilent 2100 bioanalyzer system (Agilent Technologies; catalogue number G2939BA). Finalized libraries were combined in equimolar concentration (2 nM) before sequencing. Metagenomic libraries were sequenced on the Illumina HiSeq 2500 (chemistry v4.0) next-generation sequencing (NGS) platform by Eurofins Genetic

Services Ltd. (Ebersberg, Germany) using the high-output run mode for 2×125 -bp paired-end reads with the addition of a PhiX library (1%) to estimate sequence quality. Contaminant reads from humans were removed from raw FASTQ sequence files with NCBI Best Match Tagger (BMTagger) software. Remaining reads were quality checked with Picard and SAMTools software, removing duplicate and substandard-quality reads. The resulting high-quality sequence data were subjected to functional profiling by the Human Microbiome Project Unified Metabolic Analysis Network (HUMAN2 v.0.5.0) pipeline (54). Here, models of microbial metabolic pathways derived from the MetaCyc database were generated. Metagenomic taxonomic profiling was performed with the Kraken software package (v0.10.6) (55). Species-level shotgun metagenomics data were converted to relative abundance, and HUMAN2 pathway profiles were normalized to copies per million units prior to statistical analysis.

Metabolic profiling. Metabolic analyses of urine and fecal biofluids were conducted using established $^1\text{H-NMR}$ and GC-MS methods (56, 57), as previously described (7).

Global, untargeted $^1\text{H-NMR}$ urine and fecal metabolic profiling analyses were performed on a 600-MHz spectrometer (Bruker BioSpin, Germany) using established methods (56, 58). Briefly, urine and fecal samples were prepared with a pH 7.4 phosphate buffer for $^1\text{H-NMR}$ spectroscopy as described previously and analyzed at 300 K using the following standard one-dimensional pulse sequence with saturation of the water resonance: RD- g_{z1} - 90° - t_1 - 90° - t_m - g_{z2} - 90° -ACQ. The relaxation delay (RD) was set at 4 s, 90° represents the applied 90° radio frequency pulse, the interpulse delay (t_1) was set to an interval of 4 μs , the mixing time (t_m) was 10 ms, magnetic field gradients (g_{z1} and g_{z2}) were applied for 1 ms, and the acquisition period (AQA) was 2.7 s. Water suppression was achieved through continuous wave irradiation at the water resonance frequency. Each spectrum was acquired using 4 dummy scans, 32 scans, and 64 K time domain points.

Targeted analysis to quantify SCFAs was conducted using GC-MS as previously described (57), using an Agilent 7890B gas chromatography system, equipped with an automatic liquid sampler, coupled to an Agilent 7000C single quadrupole mass selective detector (Agilent Technologies, USA).

Urinary and fecal quality control samples were used for the nuclear magnetic resonance (NMR) and GC-MS analyses to ensure data quality.

Objective assessment of dietary intake. We applied a novel and validated mathematical tool capable of objectively assessing free-living individuals' dietary patterns based on their urine composition, without the need to collect dietary data. The tool is based on a Monte Carlo cross-validated PLS-DA model built from global urinary metabolic profiles derived from a highly controlled environment (inpatient randomized controlled crossover trial) to ensure that healthy participants were fully adherent to four dietary interventions that reflected four different levels of adherence (25, 50, 75, and 100%) to WHO healthy eating recommendations (increased fruits, vegetables, whole grains, and dietary fiber and decreased fats, sugars, and salt) (17). The application of this model has been successfully validated in free-living populations (17, 59). $^1\text{H-NMR}$ urinary metabolic profiles from athletes and controls were projected into the MCCV-PLS-DA model to calculate a predicted score for each participant that reflected their adherence to healthy eating. We considered a score that is predicted at 100% or higher to be reflective of healthy eating; e.g., the global profile of such a score has (on average) higher concentrations of biomarkers reflective of the 100% diet than trial participants following the 100% diet.

Relationship between self-reported dietary intake and independent urinary metabolite assessment. Based on the self-reported food frequency questionnaires (FFQ), there were no significant differences in the frequency of citrus fruit or fruit and vegetable consumption comparing controls with a $\leq 35\%$ adherence to healthy eating versus controls with a $\geq 100\%$ adherence to healthy eating (citrus consumption, $P = 0.96$; fruits and vegetables, $P = 0.67$) or controls with a $\leq 35\%$ adherence to healthy eating versus athletes with a $\geq 100\%$ adherence to healthy eating (citrus consumption, $P = 0.85$; fruits and vegetables, $P = 0.77$).

However, we and others have demonstrated the use of urinary biomarkers to independently assess the consumption of specific food groups, including proline-betaine for citrus fruits (20, 60) and hippurate for fruits and vegetables (31, 61). Investigation of urinary proline-betaine and hippurate found higher intake of citrus and fruits and vegetables in the healthy eating athletes and controls. This suggests a level of misreporting in the self-reported data that makes it difficult to interpret. Therefore, participants' dietary intakes were scored using the validated metabolic profiling dietary assessment tool that relies on urinary metabolic profiles and dietary biomarkers instead of self-reported FFQ (17).

Statistical analysis. $^1\text{H-NMR}$ spectra were phased and digitized over the range $\delta 0.5$ to 9.5 and imported into MATLAB to undergo median fold change normalization (62). Multivariate statistical analysis was performed using SIMCA v14.1 (Umetrics) (62). The three comparative groups previously defined were modeled using OPLS-DA: (i) controls with a $< 35\%$ adherence to healthy eating versus controls with a $\geq 100\%$ adherence to healthy eating (model 1), (ii) controls with a $\geq 100\%$ adherence to healthy eating versus athletes with a $\geq 100\%$ adherence to healthy eating (model 2), and (iii) controls with a $< 35\%$ adherence to healthy eating versus athletes with a $\geq 100\%$ adherence to healthy eating (model 3). The OPLS-DA models were established based on one predictive component and one orthogonal component. Unit variance scaling was applied to $^1\text{H-NMR}$ spectral data. The fit and predictability of the models obtained were determined by the R^2_Y and Q^2_Y values, respectively. Significant metabolites differentiating between groups were obtained from $^1\text{H-NMR}$ OPLS-DA models after investigating $^1\text{H-NMR}$ signals with correlation coefficient values higher than 0.4. Jack-knifed 95% confidence intervals of the coefficients were used to confirm the significance of the variables.

Univariate statistical analysis (two-sided Mann-Whitney U test) was used to identify discriminatory

metabolites in SCFA quantitative data sets. *P* values were adjusted for multiple testing using the Benjamini-Hochberg (BH) false-discovery rate method (pFDR).

Linear regression analysis of ¹H-NMR global spectra was performed against Shannon, Simpson, and whole-tree bacterial diversity indices, corrected for confounding factors (age, lean mass, and fat mass), using an in-house MATLAB (release 2014a) script. Metabolites with adjusted pFDR values of <0.01, were considered significant and were subsequently visualized in a Manhattan plot.

Finally, heat maps showing significant Spearman correlations between data sets were generated using the ComplexHeatmap script in R (63). Spearman correlations were first calculated in MATLAB. Correlations between metabolite-metabolite data sets with a pFDR <0.05 were included. Due to the large number of variables, a false-discovery rate of 1% was set for correlations, including microbiome taxa and pathway data sets. Hierarchical clustering of correlations was performed using Euclidean distances.

Metabolite identification. A combination of data-driven strategies such as such as SubseT Optimization by Reference Matching (STORM) (64) and Statistical TOtal Correlation Spectroscopy (STOCSY) (65) and analytical identification strategies were used to aid structural identification of significant discriminatory metabolites. Specifically, a catalogue of 1D ¹H-NMR sequences with water presaturation and 2D NMR experiments such as J-Resolved spectroscopy, ¹H-¹H Total Correlation Spectroscopy (TOCSY), ¹H-¹H COrrrelation Spectroscopy (COSY), ¹H-¹³C Hetero-nuclear Single Quantum Coherence (HSQC) and ¹H-¹³C Hetero-nuclear Multiple-Bond Correlation (HMBC) spectroscopy were performed. Finally, when possible, metabolites were confirmed by *in situ* spiking experiments using authentic chemical standards.

Metabolite quantification. Semiquantification of identified metabolites from ¹H-NMR profiles was performed using an in-house MATLAB script to annotate maximal intensity of median fold change normalized spectral peaks. The highest-intensity spectral peak was identified for each measured metabolite and precisely quantified for each individual. Manual annotation enabled precise marking of peaks, allowing for variations in chemical shift in individual samples. Where spectral peaks overlapped with other metabolites, the highest-intensity unambiguous spectral region was used.

Data availability. All presented data are tabulated and detailed in the main text and supporting information. Codes used to analyze these data are referenced in the main text. Metagenomic sequencing data are available at the European Nucleotide Archive repository (accession number [PRJEB15388](https://www.ebi.ac.uk/ena/record/PRJEB15388)). The 16S rRNA sequence reads are available from the Sequence Read Archive (accession number [PRJEB4609](https://www.ncbi.nlm.nih.gov/sra/PRJEB4609)).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 10.1 MB.

FIG S2, TIF file, 3.8 MB.

FIG S3, TIF file, 4.2 MB.

FIG S4, TIF file, 4.8 MB.

FIG S5, TIF file, 3.4 MB.

TABLE S1, DOCX file, 0.03 MB.

TABLE S2, DOCX file, 0.04 MB.

ACKNOWLEDGMENTS

We express gratitude to all participants for the donation of time and samples, in particular, staff and players at the Irish Rugby Football Union.

This research was funded by Science Foundation Ireland in the form of a center grant (APC Microbiome Institute grant number SFI/12/RC/2273). Research in the Cotter laboratory is funded by SFI through the PI award “Obesibiotics” (11/PI/1137). N.P. is funded by the Diabetes Research and Wellness Foundation through the Sutherland-Earl Clinical Research Fellowship 2015. O.O. and W.B. are funded by Science Foundation Ireland through a Starting Investigator Research Grant award (13/SIRG/2160). I.G.-P. is supported by an NIHR career development research fellowship (NIHR-CDF-2017-10-032). J.M.P. is supported by a Rutherford Fund Fellowship at Health Data Research (HDR) UK (MR/S004033/1). G.F. is an NIHR senior investigator. Infrastructure support was provided by the NIHR Imperial Biomedical Research Centre (BRC) in line with the gut health research theme based at Imperial College Healthcare National Health Service (NHS) Trust and Imperial College London.

The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

N.P. and I.G.-P. devised the study experimental design and approach. E.H., I.G.-P., and N.P. performed metabolomic processing and statistical analysis thereof. J.M.P. provided

statistical tools. G.F. provided support on dietary analysis. W.B. prepared DNA samples for metagenomic sequencing. O.O. and W.B. processed and analyzed the metagenomic data. N.P. wrote the manuscript, which was revised by all the authors. Results were discussed by all authors.

F.S. is a founder shareholder in Atlantia Food Clinical Trials, Tucana Health, and Alimentary Health. He is director of the APC Microbiome Institute, a research center funded in part by Science Foundation Ireland (APC/SFI/12/RC/2273) and which is/has recently been in receipt of research grants from Abbvie, Alimentary Health, Cremo, Danone, Janssen, Friesland Campina, General Mills, Kerry, MeadJohnson, Nutricia, 4D Pharma, Second Genome, and Sigmoid Pharma.

The Cork Clinical Research Ethics Committee granted ethics approval.

REFERENCES

- Mika A, Van Treuren W, Gonzalez A, Herrera JJ, Knight R, Fleshner M. 2015. Exercise is more effective at altering gut microbial composition and producing stable changes in lean mass in juvenile versus adult male F344 rats. *PLoS One* 10:e0125889. <https://doi.org/10.1371/journal.pone.0125889>.
- Choi JJ, Eum SY, Rampersaud E, Daunert S, Abreu MT, Toborek M. 2013. Exercise attenuates PCB-induced changes in the mouse gut microbiome. *Environ Health Perspect* 121:725–730. <https://doi.org/10.1289/ehp.1306534>.
- Queipo-Ortuno MI, Seoane LM, Murri M, Pardo M, Gomez-Zumaquero JM, Cardona F, Casanueva F, Tinahones FJ. 2013. Gut microbiota composition in male rat models under different nutritional status and physical activity and its association with serum leptin and ghrelin levels. *PLoS One* 8:e65465. <https://doi.org/10.1371/journal.pone.0065465>.
- Lamoureux EV, Grandy SA, Langille MGI. 2017. Moderate exercise has limited but distinguishable effects on the mouse microbiome. *mSystems* 2:e00006-17. <https://doi.org/10.1128/mSystems.00006-17>.
- Estaki M, Pither J, Baumeister P, Little JP, Gill SK, Ghosh S, Ahmadi-Vand Z, Marsden KR, Gibson DL. 2016. Cardiorespiratory fitness as a predictor of intestinal microbial diversity and distinct metagenomic functions. *Microbiome* 4:42. <https://doi.org/10.1186/s40168-016-0189-7>.
- Bressa C, Bailen-Andrino M, Perez-Santiago J, Gonzalez-Soltero R, Perez M, Montalvo-Lominchar MG, Mate-Munoz JL, Dominguez R, Moreno D, Larrosa M. 2017. Differences in gut microbiota profile between women with active lifestyle and sedentary women. *PLoS One* 12:e0171352. <https://doi.org/10.1371/journal.pone.0171352>.
- Barton W, Penney NC, Cronin O, Garcia-Perez I, Molloy MG, Holmes E, Shanahan F, Cotter PD, O'Sullivan O. 2018. The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level. *Gut* 67:625–633. <https://doi.org/10.1136/gutjnl-2016-313627>.
- Clarke SF, Murphy EF, O'Sullivan O, Lucey AJ, Humphreys M, Hogan A, Hayes P, O'Reilly M, Jeffery IB, Wood-Martin R, Kerins DM, Quigley E, Ross RP, O'Toole PW, Molloy MG, Falvey E, Shanahan F, Cotter PD. 2014. Exercise and associated dietary extremes impact on gut microbial diversity. *Gut* 63:1913–1920. <https://doi.org/10.1136/gutjnl-2013-306541>.
- Petersen LM, Bautista EJ, Nguyen H, Hanson BM, Chen L, Lek SH, Sodergren E, Weinstock GM. 2017. Community characteristics of the gut microbiomes of competitive cyclists. *Microbiome* 5:98. <https://doi.org/10.1186/s40168-017-0320-4>.
- Scheiman J, Luber JM, Chavkin TA, MacDonald T, Tung A, Pham LD, Wibowo MC, Wurth RC, Punthambaker S, Tierney BT, Yang Z, Hattab MW, Avila-Pacheco J, Clish CB, Lessard S, Church GM, Kostic AD. 2019. Metabolomics analysis of elite athletes identifies a performance-enhancing microbe that functions via lactate metabolism. *Nat Med* 25:1104–1109. <https://doi.org/10.1038/s41591-019-0485-4>.
- Allen JM, Mailing LJ, Niemi GM, Moore R, Cook MD, White BA, Holscher HD, Woods JA. 2018. Exercise alters gut microbiota composition and function in lean and obese humans. *Med Sci Sports Exerc* 50:747–757. <https://doi.org/10.1249/MSS.0000000000001495>.
- Liu Y, Wang Y, Ni Y, Cheung CKY, Lam KSL, Wang Y, Xia Z, Ye D, Guo J, Tse MA, Panagiotou G, Xu A. 2020. Gut microbiome fermentation determines the efficacy of exercise for diabetes prevention. *Cell Metab* 31:77–91.e5. <https://doi.org/10.1016/j.cmet.2019.11.001>.
- Poslusna K, Ruprich J, de Vries JH, Jakubikova M, van't Veer P. 2009. Misreporting of energy and micronutrient intake estimated by food records and 24 hour recalls, control and adjustment methods in practice. *Br J Nutr* 101(Suppl 2):S73–S85. <https://doi.org/10.1017/S0007114509990602>.
- Rennie KL, Coward A, Jebb SA. 2007. Estimating under-reporting of energy intake in dietary surveys using an individualised method. *Br J Nutr* 97:1169–1176. <https://doi.org/10.1017/S0007114507433086>.
- Lafay L, Mennen L, Basdevant A, Charles MA, Borys JM, Eschwege E, Romon M. 2000. Does energy intake underreporting involve all kinds of food or only specific food items? Results from the Fleurbaix Laventie Ville Sante (FLVS) study. *Int J Obes Relat Metab Disord* 24:1500–1506. <https://doi.org/10.1038/sj.ijo.0801392>.
- World Health Organization T. 2004. Global strategy on diet, physical activity and health. WHO, Geneva, Switzerland.
- Garcia-Perez I, Posma JM, Gibson R, Chambers ES, Hansen TH, Vestergaard H, Hansen T, Beckmann M, Pedersen O, Elliott P, Stamler J, Nicholson JK, Draper J, Mathers JC, Holmes E, Frost G. 2017. Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. *Lancet Diabetes Endocrinol* 5:184–195. [https://doi.org/10.1016/S2213-8587\(16\)30419-3](https://doi.org/10.1016/S2213-8587(16)30419-3).
- Cobb-Clark DA, Kassenboehmer SC, Schurer S. 2014. Healthy habits: the connection between diet, exercise, and locus of control. *J Econ Behav Organ* 98:1–28. <https://doi.org/10.1016/j.jebo.2013.10.011>.
- Akbaraly TN, Ferrie JE, Berr C, Brunner EJ, Head J, Marmot MG, Singh-Manoux A, Ritchie K, Shipley MJ, Kivimaki M. 2011. Alternative Healthy Eating Index and mortality over 18 y of follow-up: results from the Whitehall II cohort. *Am J Clin Nutr* 94:247–253. <https://doi.org/10.3945/ajcn.111.013128>.
- Heinzmann SS, Brown IJ, Chan Q, Bictash M, Dumas ME, Kochhar S, Stamler J, Holmes E, Elliott P, Nicholson JK. 2010. Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. *Am J Clin Nutr* 92:436–443. <https://doi.org/10.3945/ajcn.2010.29672>.
- Il'yasova D, Spasojevic I, Wang F, Tolun AA, Base K, Young SP, Marcom PK, Marks J, Mixon G, DiGiulio R, Millington DS. 2010. Urinary biomarkers of oxidative status in a clinical model of oxidative assault. *Cancer Epidemiol Biomarkers Prev* 19:1506–1510. <https://doi.org/10.1158/1055-9965.EPI-10-0211>.
- Pechlivanis A, Kostidis S, Sarasilanidis P, Petridou A, Tsalis G, Mougios V, Gika HG, Mikros E, Theodoridis GA. 2010. (1)H NMR-based metabolomic investigation of the effect of two different exercise sessions on the metabolic fingerprint of human urine. *J Proteome Res* 9:6405–6416. <https://doi.org/10.1021/pr100684t>.
- Nieman DC, Gillitt ND, Sha W, Meaney MP, John C, Pappan KL, Kinchen JM. 2015. Metabolomics-based analysis of banana and pear ingestion on exercise performance and recovery. *J Proteome Res* 14:5367–5377. <https://doi.org/10.1021/acs.jproteome.5b00909>.
- Mayneris-Perxachs J, Lima AA, Guerrant RL, Leite AM, Moura AF, Lima NL, Soares AM, Havt A, Moore SR, Pinkerton R, Swann JR. 2016. Urinary N-methylnicotinamide and beta-aminoisobutyric acid predict catch-up growth in undernourished Brazilian children. *Sci Rep* 6:19780. <https://doi.org/10.1038/srep19780>.
- Kraus D, Yang Q, Kong D, Banks AS, Zhang L, Rodgers JT, Pirinen E, Puliniikunnil TC, Gong F, Wang YC, Cen Y, Sauve AA, Asara JM, Peroni OD, Monia BP, Bhanot S, Alhonen L, Puigserver P, Kahn BB. 2014. Nicotinamide N-methyltransferase knockdown protects against diet-induced obesity. *Nature* 508:258–262. <https://doi.org/10.1038/nature13198>.

26. Elliott P, Pasma JM, Chan Q, Garcia-Perez I, Wijeyesekera A, Bictash M, Ebbels TM, Ueshima H, Zhao L, van Horn L, Daviglius M, Stamler J, Holmes E, Nicholson JK. 2015. Urinary metabolic signatures of human adiposity. *Sci Transl Med* 7:285ra62. <https://doi.org/10.1126/scitranslmed.aaa5680>.
27. Xiong X, Liu D, Wang Y, Zeng T, Peng Y. 2016. Urinary 3-(3-hydroxyphenyl)-3-hydroxypropionic acid, 3-hydroxyphenylacetic acid, and 3-hydroxyhippuric acid are elevated in children with autism spectrum disorders. *Biomed Res Int* 2016:1–8. <https://doi.org/10.1155/2016/9485412>.
28. Emond P, Mavel S, Aidoud N, Nadal-Desbarats L, Montigny F, Bonnet-Brilhault F, Barthelemy C, Merten M, Sarda P, Laumonier F, Vourc'h P, Blasco H, Andres CR. 2013. GC-MS-based urine metabolic profiling of autism spectrum disorders. *Anal Bioanal Chem* 405:5291–5300. <https://doi.org/10.1007/s00216-013-6934-x>.
29. Lees HJ, Swann JR, Wilson ID, Nicholson JK, Holmes E. 2013. Hippurate: the natural history of a mammalian-microbial cometabolite. *J Proteome Res* 12:1527–1546. <https://doi.org/10.1021/pr300900b>.
30. Holmes E, Loo RL, Stamler J, Bictash M, Yap IK, Chan Q, Ebbels T, De Iorio M, Brown IJ, Veselkov KA, Daviglius ML, Kesteloot U, Ueshima H, Zhao L, Nicholson JK, Elliott P. 2008. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* 453:396–400. <https://doi.org/10.1038/nature06882>.
31. Heinzmann SS, Merrifield CA, Rezzi S, Kochhar S, Lindon JC, Holmes E, Nicholson JK. 2012. Stability and robustness of human metabolic phenotypes in response to sequential food challenges. *J Proteome Res* 11:643–655. <https://doi.org/10.1021/pr2005764>.
32. Tang WH, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, Wu Y, Hazen SL. 2013. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N Engl J Med* 368:1575–1584. <https://doi.org/10.1056/NEJMoa1109400>.
33. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, Feldstein AE, Britt EB, Fu X, Chung YM, Wu Y, Schauer P, Smith JD, Allayee H, Tang WH, DiDonato JA, Lusis AJ, Hazen SL. 2011. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472:57–63. <https://doi.org/10.1038/nature09922>.
34. Mueller DM, Allenspach M, Othman A, Saely CH, Muendlein A, Vonbank A, Drexel H, von Eckardstein A. 2015. Plasma levels of trimethylamine-N-oxide are confounded by impaired kidney function and poor metabolic control. *Atherosclerosis* 243:638–644. <https://doi.org/10.1016/j.atherosclerosis.2015.10.091>.
35. Dumas ME, Rothwell AR, Hoyles L, Aranas T, Chilloux J, Calderari S, Noll EM, Pean N, Boulange CL, Blancher C, Barton RH, Gu Q, Fearnside JF, Deshayes C, Hue C, Scott J, Nicholson JK, Gauguier D. 2017. Microbial-host co-metabolites are prodromal markers predicting phenotypic heterogeneity in behavior, obesity, and impaired glucose tolerance. *Cell Rep* 20:136–148. <https://doi.org/10.1016/j.celrep.2017.06.039>.
36. Lin HV, Frassetto A, Kowalik EJ, Jr, Nawrocki AR, Lu MM, Kosinski JR, Hubert JA, Szeto D, Yao X, Forrest G, Marsh DJ. 2012. Butyrate and propionate protect against diet-induced obesity and regulate gut hormones via free fatty acid receptor 3-independent mechanisms. *PLoS One* 7:e35240. <https://doi.org/10.1371/journal.pone.0035240>.
37. Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, Bunker MK, Bultman SJ. 2011. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* 13:517–526. <https://doi.org/10.1016/j.cmet.2011.02.018>.
38. Demigne C, Yacoub C, Remy S, Fafournoux P. 1986. Propionate and butyrate metabolism in rat or sheep hepatocytes. *Biochim Biophys Acta* 875:535–542. [https://doi.org/10.1016/0005-2760\(86\)90074-3](https://doi.org/10.1016/0005-2760(86)90074-3).
39. Chambers ES, Viardot A, Psichas A, Morrison DJ, Murphy KG, Zaccarelli SE, MacDougall K, Preston T, Tedford C, Finlayson GS, Blundell JE, Bell JD, Thomas EL, Mt-Isa S, Ashby D, Gibson GR, Kolida S, Dhillo WS, Bloom SR, Morley W, Clegg S, Frost G. 2015. Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults. *Gut* 64:1744–1754. <https://doi.org/10.1136/gutjnl-2014-307913>.
40. Bui TP, Ritari J, Boeren S, de Waard P, Plugge CM, de Vos WM. 2015. Production of butyrate from lysine and the Amadori product fructoselysine by a human gut commensal. *Nat Commun* 6:10062. <https://doi.org/10.1038/ncomms10062>.
41. Heimann E, Nyman M, Palbrink AK, Lindkvist-Petersson K, Degerman E. 2016. Branched short-chain fatty acids modulate glucose and lipid metabolism in primary adipocytes. *Adipocyte* 5:359–368. <https://doi.org/10.1080/21623945.2016.1252011>.
42. Holec M. 2018. Branched-chain amino acids in health and disease: metabolism, alterations in blood plasma, and as supplements. *Nutr Metab (Lond)* 15:33. <https://doi.org/10.1186/s12986-018-0271-1>.
43. Brosnan JT, Brosnan ME. 2006. Branched-chain amino acids: enzyme and substrate regulation. *J Nutr* 136:2075–2115. <https://doi.org/10.1093/jn/136.1.2075>.
44. Neis EP, Dejong CH, Rensen SS. 2015. The role of microbial amino acid metabolism in host metabolism. *Nutrients* 7:2930–2946. <https://doi.org/10.3390/nu7042930>.
45. Sonnenburg JL, Backhed F. 2016. Diet-microbiota interactions as moderators of human metabolism. *Nature* 535:56–64. <https://doi.org/10.1038/nature18846>.
46. Cotillard A, Kennedy SP, Kong LC, Prifti E, Pons N, Le Chatelier E, Almeida M, Quinquis B, Levenez F, Galleron N, Gougis S, Rizkalla S, Batto JM, Renault P, Dore J, Zucker JD, Clement K, Ehrlich SD, ANRM Consortium. 2013. Dietary intervention impact on gut microbial gene richness. *Nature* 500:585–588. <https://doi.org/10.1038/nature12480>.
47. Menni C, Jackson MA, Pallister T, Steves CJ, Spector TD, Valdes AM. 2017. Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain. *Int J Obes* 41:1099–1105. <https://doi.org/10.1038/ijo.2017.66>.
48. Zhernakova A, Kurilshikov A, Bonder MJ, Tigchelaar EF, Schirmer M, Vatanen T, Mujagic Z, Vila AV, Falony G, Vieira-Silva S, Wang J, Imhann F, Brandsma E, Jankipersadsing SA, Joossens M, Cenit MC, Deelen P, Swertz MA, Weersma RK, Feskens EJ, Netea MG, Gevers D, Jonkers D, Franke L, Aulchenko YS, Huttenhower C, Raes J, Hofker MH, Xavier RJ, Wijmenga C, Fu J, LifeLines Cohort Study. 2016. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* 352:565–569. <https://doi.org/10.1126/science.123369>.
49. Macfarlane GT, Cummings JH, Allison C. 1986. Protein degradation by human intestinal bacteria. *J Gen Microbiol* 132:1647–1656. <https://doi.org/10.1099/00221287-132-6-1647>.
50. Hughes R, Magee EA, Bingham S. 2000. Protein degradation in the large intestine: relevance to colorectal cancer. *Curr Issues Intest Microbiol* 1:51–58.
51. Wareham NJ, Jakes RW, Rennie KL, Mitchell J, Hennings S, Day NE. 2002. Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire. *Int J Epidemiol* 31:168–174. <https://doi.org/10.1093/ije/31.1.168>.
52. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336. <https://doi.org/10.1038/nmeth.f.303>.
53. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO. 2017. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35:7188–7196. <https://doi.org/10.1093/nar/gkm864>.
54. Franzosa EA, Mincer L, Rahnnavard G, Thompson LR, Schirmer M, Weingart G, Lipson KS, Knight R, Caporaso JG, Segata N, Huttenhower C. 2018. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* 15:962–968. <https://doi.org/10.1038/s41592-018-0176-y>.
55. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 15:R46. <https://doi.org/10.1186/gb-2014-15-3-r46>.
56. Dona AC, Jimenez B, Schafer H, Humpfer E, Spraul M, Lewis MR, Pearce JT, Holmes E, Lindon JC, Nicholson JK. 2014. Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Anal Chem* 86:9887–9894. <https://doi.org/10.1021/ac5025039>.
57. Garcia-Villalba R, Gimenez-Bastida JA, Garcia-Conesa MT, Tomas-Barberan FA, Carlos Espin J, Larrosa M. 2012. Alternative method for gas chromatography-mass spectrometry analysis of short-chain fatty acids in faecal samples. *J Sep Sci* 35:1906–1913. <https://doi.org/10.1002/jssc.201101121>.
58. Gratton J, Petchcharaburanin J, Mullish BH, Williams HR, Thursz M, Nicholson JK, Holmes E, Marchesi JR, Li JV. 2016. Optimized sample handling strategy for metabolic profiling of human feces. *Anal Chem* 88:4661–4668. <https://doi.org/10.1021/acs.analchem.5b04159>.
59. Garcia-Perez I, Pasma JM, Chambers ES, Mathers JC, Draper J, Beckmann M, Nicholson JK, Holmes E, Frost G. 2020. Dietary metabolite modelling predicts individual responses to dietary interventions. *Nat Food* 1:355–364. <https://doi.org/10.1038/s43016-020-0092-z>.
60. Lloyd AJ, Beckmann M, Fave G, Mathers JC, Draper J. 2011. Proline betaine and its biotransformation products in fasting urine samples are potential

- biomarkers of habitual citrus fruit consumption. *Br J Nutr* 106:812–824. <https://doi.org/10.1017/S0007114511001164>.
61. Krupp D, Doberstein N, Shi L, Remer T. 2012. Hippuric acid in 24-hour urine collections is a potential biomarker for fruit and vegetable consumption in healthy children and adolescents. *J Nutr* 142:1314–1320. <https://doi.org/10.3945/jn.112.159319>.
62. Veselkov KA, Vingara LK, Masson P, Robinette SL, Want E, Li JV, Barton RH, Boursier-Neyret C, Walther B, Ebbels TM, Pelczer I, Holmes E, Lindon JC, Nicholson JK. 2011. Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery. *Anal Chem* 83:5864–5872. <https://doi.org/10.1021/ac201065j>.
63. Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32:2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>.
64. Posma JM, Garcia-Perez I, De Iorio M, Lindon JC, Elliott P, Holmes E, Ebbels TM, Nicholson JK. 2012. Subset optimization by reference matching (STORM): an optimized statistical approach for recovery of metabolic biomarker structural information from 1H NMR spectra of biofluids. *Anal Chem* 84:10694–10701. <https://doi.org/10.1021/ac302360v>.
65. Cloarec O, Dumas ME, Craig A, Barton RH, Trygg J, Hudson J, Blancher C, Gauguier D, Lindon JC, Holmes E, Nicholson J. 2005. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Anal Chem* 77:1282–1289. <https://doi.org/10.1021/ac048630x>.