



HHS Public Access

Author manuscript

Sci Transl Med. Author manuscript; available in PMC 2019 June 28.

Published in final edited form as:

Sci Transl Med. 2015 April 29; 7(285): 285ra62. doi:10.1126/scitranslmed.aaa5680.

Urinary metabolic signatures of human adiposity

Paul Elliott^{1,*†}, Joram M. Posma^{1,2,†}, Queenie Chan¹, Isabel Garcia-Perez², Anisha Wijeyesekera², Magda Bictash², Timothy M. D. Ebbels², Hirotsugu Ueshima³, Liancheng Zhao⁴, Linda van Horn⁵, Martha Daviglus^{5,6}, Jeremiah Stamler⁵, Elaine Holmes², and Jeremy K. Nicholson^{2,*}

¹Department of Epidemiology and Biostatistics, Medical Research Council-Public Health England (MRC-PHE) Centre for Environment and Health, School of Public Health, Imperial College London, London W2 1PG, UK.

²Biomolecular Medicine, Division of Computational and Systems Medicine, MRC-National Institute for Health Research (MRC-NIHR) National Phenome Centre, MRC-PHE Centre for Environment and Health, Department of Surgery and Cancer, Imperial College London, London SW7 2AZ, UK

³Department of Health Science, Shiga University of Medical Science, Otsu 520-2192, Japan.

⁴Department of Epidemiology, Fu Wai Hospital and Cardiovascular Institute, Chinese Academy of Medical Sciences, Beijing, Beijing 100037, People's Republic of China.

⁵Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA.

⁶College of Medicine at Chicago, University of Illinois at Chicago, Chicago, IL 60612, USA.

Abstract

Obesity is a major public health problem worldwide. We used 24-hour urinary metabolic profiling by proton (¹H) nuclear magnetic resonance (NMR) spectroscopy and ion exchange chromatography to characterize the metabolic signatures of adiposity in the U.S. ($n = 1880$) and UK ($n = 444$) cohorts of the INTERMAP (International Study of Macro- and Micronutrients and Blood Pressure) epidemiologic study. Metabolic profiling of urine samples collected over two 24-hour time periods 3 weeks apart showed reproducible patterns of metabolite excretion associated with adiposity. Exploratory analysis of the urinary metabolome using ¹H NMR spectroscopy of the U.S. samples identified 29 molecular species, clustered in interconnecting metabolic pathways, that were significantly associated ($P = 1.5 \times 10^{-5}$ to 2.0×10^{-36}) with body mass index (BMI); 25 of these species were also found in the UK validation cohort. We found multiple associations

*Corresponding author. p.elliott@imperial.ac.uk (P.E.); j.nicholson@imperial.ac.uk (J.K.N.).

†These authors contributed equally to this work.

Author contributions: P.E., J.M.P., E.H., and J.K.N. wrote the manuscript; J.M.P. performed data analysis; Q.C. processed and provided dietary and population data; J.M.P., I.G.-P., A.W., and M.B. identified metabolites; I.G.-P. performed the experiments; T.M.D.E. provided statistical support; P.E. and J.K.N. designed the study; P.E., J.S., H.U., L.Z., L.v.H., and M.D. lead the INTERMAP study; J.K.N., P.E., and E.H. lead the INTERMAP metabolomic study.

Competing interests: J.K.N. and E.H. are both nonexecutive directors of Metabometrix Ltd., one of several fee-for-service companies that provide metabolic profiling services. The company does not work directly in the area of obesity research, historically mainly providing toxicological screening capabilities. I.G.P. and T.M.D.E. have been paid consultants for Metabometrix Ltd. The other authors declare that they have no competing interests.

between urinary metabolites and BMI including urinary glycoproteins and *N*-acetyl neuraminic acid (related to renal function), trimethylamine, dimethylamine, 4-cresyl sulfate, phenylacetylglutamine and 2-hydroxyisobutyrate (gut microbial co-metabolites), succinate and citrate (tricarboxylic acid cycle intermediates), ketoleucine and the ketoleucine/leucine ratio (linked to skeletal muscle mitochondria and branched-chain amino acid metabolism), ethanolamine (skeletal muscle turnover), and 3-methylhistidine (skeletal muscle turnover and meat intake). We mapped the multiple BMI-metabolite relationships as part of an integrated systems network that describes the connectivities between the complex pathway and compartmental signatures of human adiposity.

INTRODUCTION

The prevalence of obesity and being overweight is rising worldwide and now affects a large proportion of the adult population in the United States, UK, and many other countries (1, 2). Adiposity is associated with increased mortality from heart disease, stroke, diabetes, and cancer (3), contributing to an estimated 3.4 million deaths per year worldwide (4). However, the mechanisms by which adiposity may contribute to increased risk of these diseases are not fully understood. Better knowledge about the effects of adiposity on human metabolism and the complex metabolic pathway connections underlying them is needed to provide new insights into perturbations that link adiposity to human disease risks. This in turn may help efforts to control the obesity epidemic and its sequelae through better targeting of preventive efforts to those most at risk.

Exploratory metabolic phenotyping offers a powerful means of capturing systems-level information that reflects both genetic (host, gut bacterial metagenome) and environmental influences, hence helping to elucidate the metabolic disturbances and pathways associated with chronic disease (5, 6). For example, we have demonstrated use of a metabolic phenotyping approach for discovering urinary molecular markers associated with interindividual blood pressure variation and cardiovascular disease risk (7). Here, we apply this approach to elucidate the urinary metabolic signatures associated with body mass index (BMI) and adiposity in the U.S. and UK cohorts of the International Study of Macro- and Micronutrients and Blood Pressure (INTERMAP) epidemiologic study.

RESULTS

We have used a unique collection of urinary aliquots sampled from pooled timed 24-hour urine collections obtained over two separate 24-hour periods from men and women in the population-based INTERMAP epidemiologic study (7–9) in the United States and UK. We used these urine samples to delineate a metabolic signature of adiposity that was reproducible in individuals over time and in different populations (Fig.1). We applied a combination of untargeted proton (^1H) nuclear magnetic resonance (NMR) spectroscopy to aliquots from the two 24-hour urinary collection periods and also performed targeted analyses of urinary amino acids and related compounds using ion exchange chromatography (IEC). We analyzed urinary amino acids because of the reported importance of amino acids, especially branched-chain amino acids (BCAAs), in adiposity (10–16). We identified metabolites significantly and reproducibly associated with BMI after correction for multiple

testing and located these in an interconnecting systems framework. The few previous metabolic phenotyping studies of human adiposity have analyzed serum or plasma metabolite profiles (11–16) or urine from small numbers of insulin-resistant, morbidly obese, and lean individuals (17, 18). Here, we take a population-wide approach using BMI as a continuous measure across the full range of BMI from lean to obese individuals. Use of 24-hour urine samples as reported here has the advantage over single blood or spot urine samples collected at one point in time, because it captures the end products of metabolism integrated over a 24-hour time period and therefore is not influenced by sampling time and diurnal variation in metabolite concentrations (19).

We have included data on 1880 individuals (962 men and 918 women) ages 40 to 59 years from eight American population samples in INTERMAP with a replication data set of 444 individuals (245 men and 199 women) ages 40 to 59 years in INTERMAP UK (8) (Fig. 1). We obtained height and weight measurements for the calculation of BMI, extensive covariate information (for example, demographic variables and medical and life-style factors), as well as data on dietary variables and energy intake from four interviewer-administered multipass 24-hour dietary recalls per person (8, 9) (table S1).

We present our results as partial correlations between BMI and urinary metabolites along with *P* values to indicate the strength of associations and *Q* values to indicate the false discovery rate (see below). This approach made no assumptions as to whether the observed cross-sectional associations were causally related or were the consequence of adiposity. Rather, our results captured the metabolic signatures of the multitude of dietary, environmental, and other life-style characteristics associated with the obesity epidemic. We used the eight U.S. population samples as the discovery set and calculated the partial correlation of individual BMI against (i) each of 7100 ¹H NMR detected spectral variables and (ii) the targeted set of urinary amino acids and related compounds measured by IEC. In multiple regression analyses, we adjusted initially for age, gender, and population sample (defined as model 1, see footnotes of tables and Materials and Methods) because of the known associations of these variables with both BMI and metabolite concentrations (7,15, 20, 21). For these discovery analyses, we used a *Q* value (22) threshold of 1% based on the Storey-Tibshirani false discovery rate (23) (ST-FDR) to identify spectral variables significantly associated with BMI after correction for multiple testing (also see Materials and Methods). For a spectral variable to be retained, both the variable itself and the two adjacent spectral variables needed to pass the *Q* value threshold of 1% (to avoid spurious findings based on a single data point) in addition to the correlation with adiposity of all three variables being in the same direction (that is, the correlation is either less than 0 or greater than 0 for all three variables). We then further adjusted for multiple potential confounders including history of heart disease or stroke, physical activity, medication and dietary supplement use, special diet, smoking, education (as an indicator of socioeconomic status), and total energy intake (defined as model 2, see footnotes of tables and Materials and Methods). Finally, we additionally adjusted for urinary creatinine (model 3) as a marker of meat intake (24, 25) and muscle turnover (26), because BMI correlates with muscle mass as well as fat mass (27). BMI-metabolite associations that decreased or became apparent after correction could reflect a link between (skeletal) muscle mass and metabolites.

We used a range of statistical spectroscopic tools (28, 29) and experimental analytical techniques (30) to identify structurally the metabolites significantly associated with BMI (table S2). We assessed the reproducibility of our findings using the NMR spectral data from the second urine collection obtained from each individual in the U.S. population samples and then validated our findings using independent data from the INTERMAP UK cohort. For these confirmatory analyses, a threshold for the ST-FDR of 5% was used to correct for multiple testing. Furthermore, for both the U.S. and UK data sets, we determined the intraclass correlations (ICCs) (table S3) between the 24-hour metabolite excretion measured in the two urine collections to assess the reliability of urinary excretion patterns over the, on average, 3-week period between the two sampling occasions.

We found significant metabolic associations with BMI involving an extensive interconnected set of biochemical pathways and physiological processes, as well as evidence for involvement of symbiotic gut microbial-human co-metabolism. The average ^1H NMR spectrum from the discovery analysis of BMI, adjusted for age, gender, and population sample (model 1, see footnotes of Table 1), is shown in Fig. 2A. The significantly associated peaks, representing multiple metabolites, are labeled in the figure with the $-\log_{10} Q$ values and direction of association shown in the accompanying Manhattan plot (Fig. 2B). We structurally identified 29 metabolites significantly associated with BMI, with P values (first 24-hour urine collection) ranging from 1.46×10^{-5} to 2.04×10^{-36} (model 1, Table 1). After adjustment for multiple potential confounders (model 2, Table 1), associations remained significant for all identified metabolites. Table 2 shows results of the targeted analyses of 22 amino acids and related compounds by IEC. The majority of these metabolites were found to be significantly associated with BMI after Bonferroni correction for multiple testing ($P = 4.55 \times 10^{-4}$, models 1 and 2); exceptions are taurine, serine, asparagine, glycine, methionine, and arginine.

Gut microbial co-metabolic signatures

We found associations of nine gut microbial co-metabolites with BMI reflecting five different host-gut transformation microbial pathways. First, we find associations with BMI of metabolites related to the gut microbial metabolism of choline. We found previously unreported associations with BMI (direct) of trimethylamine (TMA) ($P = 2.04 \times 10^{-9}$, Table 1 and Fig. 2, A and B) and dimethylamine ($P = 1.74 \times 10^{-8}$). Also, we found another product of choline degradation and glycinebetaine oxidation, dimethylglycine ($P = 1.43 \times 10^{-7}$) (31), to be directly associated with BMI. These metabolites are closely related to formate ($P = 4.50 \times 10^{-12}$), the major hub metabolite in one-carbon metabolism and a major product of gut microbial origin. The second pathway class relates to distal colonic microbial protein putrefaction and the metabolism of tyrosine and related compounds. We found a relationship (inverse) with BMI of the gut microbial-host co-metabolite 4-cresyl sulfate ($P = 8.09 \times 10^{-11}$), a hepatic phase 2 α conjugation (detoxification) product of 4-cresol, which is a distal colonic bacterial degradation product of tyrosine. We also found direct associations of BMI with urinary tyrosine ($P = 2.02 \times 10^{-45}$, Table 2), confirming previous findings in blood (11–14, 16), and with the ^1H NMR signal representing tyrosine and the tyrosine metabolite 4-hydroxymandelic acid ($P = 1.35 \times 10^{-19}$, Table 1). The association of BMI with 4-cresyl sulfate in particular is largely unaffected by the multiple covariate adjustments,

including creatinine, indicating that this relationship is independent of confounders and muscle turnover (Table 1) and thus can be considered to be directly linked to gut microbial activity differences. In addition, we show a previously unreported association (inverse) of BMI with urinary 3-hydroxymandelate ($P = 2.81 \times 10^{-8}$, Table 1), a metabolite of the α -adrenergic agonist *p*-synephrine (found in bitter orange extract) (32), and closely related to tyrosine metabolism. The third host-gut transformation microbial pathway, involving the distal colonic microbial metabolism of phenylalanine, is closely related to that of tyrosine. We found a relationship (inverse) of BMI with the gut microbial-host co-metabolite phenylacetylglutamine ($P = 1.39 \times 10^{-8}$), a bacterial degradation product of phenylalanine. A direct association with BMI of phenylalanine ($P = 8.35 \times 10^{-19}$, Table 2), based on the IEC amino acid analysis data, confirmed previous findings linking obesity with the blood concentration of phenylalanine (11–13, 16). The fourth pathway concerns the gut microbial production of benzoic acid from dietary polyphenolics, flavonoids, and related compounds in the proximal colon; benzoic acid is phase 2 conjugated with glycine in the mitochondria (mainly in the liver) to form hippurate, an abundant urine excretory product (33). We found a strong inverse association of urinary hippurate across the range of BMI ($P = 1.52 \times 10^{-14}$, Table 1); a weak inverse association with adiposity has been reported previously only in morbidly obese humans (17) as well as in dogs and mice (34, 35). Last, we found associations of BMI with the symbiotic gut microbiota-associated metabolite 2-hydroxyisobutyrate ($P = 7.24 \times 10^{-12}$), which is related to *N*-butyrate produced by the bacterium *Faecalibacterium prausnitzii* (36).

BCAA, skeletal muscle, and energy metabolism

We observed direct associations of BMI with urinary excretion of the BCAAs leucine, valine, and isoleucine, consistent with previous reports of the importance of the BCAAs in adiposity based on measurements in plasma (10–16). Here, we report associations (inverse) with BMI ($P = 7.19 \times 10^{-22}$) for ketoleucine, the first step metabolic product of leucine metabolism and ketogenesis in skeletal muscle mitochondria (37), and the ketoleucine/leucine ratio ($P = 8.85 \times 10^{-26}$). These findings reflect the importance of the enzymatic conversion of leucine to ketoleucine as a rate-limiting step in BCAA metabolism and a critical energy process in skeletal muscle (37, 38). We found a strong direct association of urinary 24-hour creatinine excretion with BMI ($P = 8.88 \times 10^{-14}$, Table 1); adjustment for creatinine (model 3, Table 2) greatly attenuated the association of BMI with 3-methylhistidine, the most strongly associated ($P = 8.28 \times 10^{-80}$, Table 2) of the variables measured by IEC. Like creatinine, 3-methylhistidine is a marker of muscle turnover (39) and meat intake (24, 25, 40). The significant direct association between BMI and ethanolamine ($P = 2.00 \times 10^{-29}$, Table 2) is similarly attenuated by creatinine adjustment, as are the direct associations with BMI of carnosine, 1-methylhistidine (both markers of meat intake) (40), threonine, alanine, valine, isoleucine, and histidine (Table 2). Some metabolites remain relatively unaffected by creatinine adjustment, including glycoproteins, tyrosine, ketoleucine, succinate, and 4-cresyl sulfate, indicating that these are largely independent of muscle turnover. Others such as taurine, serine, and asparagine only became significant after creatinine adjustment, indicating their interdependence with creatinine in association with BMI (Table 2). We also show associations (inverse) of BMI with the tricarboxylic acid

(TCA) cycle intermediates succinate ($P = 3.22 \times 10^{-18}$) and citrate ($P = 8.61 \times 10^{-15}$) (Table 1) in urine.

Other metabolite markers of adiposity

A composite *N*-acetyl signal from mixed urinary glycoproteins from the ^1H NMR-detected metabolite panel has a strong and significant association with BMI ($P = 2.04 \times 10^{-36}$) (Fig. 2, A and B, and Table 1). The association of BMI with urinary glycoproteins is thought to reflect the known higher glomerular filtration rate associated with obesity and may also provide a useful noninvasive indicator of the glomerular hyperfiltration process that is known to precede diabetes (41). Additionally, we found an association of BMI with *N*-acetyl neuraminic acid ($P = 2.02 \times 10^{-16}$), a hydrolytic breakdown product of urinary glycoproteins with sialic acid subunits; the assignment was confirmed by *N*-acetyl neuraminidase incubation experiments (42) (table S2). This finding may also reflect the fact that α -1 acid glycoprotein and other low molecular weight glycoproteins are part of the acute phase reactive proteins associated with inflammatory conditions including obesity (43). Inverse association of BMI with urinary excretion of *N*-methyl nicotinate (NMNA) ($r = -0.16$, $P = 9.04 \times 10^{-13}$) is consistent with previous observations in morbidly obese individuals (17). NMNA is a niacin-related (vitamin B3) metabolite reported to be a biomarker of coffee drinking (44). We found a strong partial correlation between urinary NMNA excretion and dietary caffeine intake ($r = 0.54$, $P = 1.20 \times 10^{-140}$, adjusted for model 1), although when caffeine intake was added to the model 1 covariates, it did not materially alter the BMI-NMNA association ($r = -0.16$, $P = 5.01 \times 10^{-12}$), suggesting that the BMI-NMNA association was independent of coffee consumption. The direct association of BMI with urinary pseudouridine excretion ($P = 2.10 \times 10^{-12}$) reflected increased whole-body nucleic acid turnover with higher BMI, consistent with enhanced adipose tissue turnover associated with obesity (45).

We found a direct association of urinary glucose with BMI ($r = 0.14$, $P = 3.27 \times 10^{-9}$), which is materially unchanged when dietary glucose intake, total sugar, or total carbohydrate intake are added to the regression model, indicating possible latent insulin resistance and undiagnosed diabetes related to adiposity (11, 13, 14, 16). We also found associations between urinary markers of dietary intake and BMI. Thus, we found an inverse association with BMI of proline betaine ($P = 7.29 \times 10^{-9}$), a biomarker of citrus fruit consumption (46, 47) associated with a healthy eating pattern (48), and a direct association of BMI with urinary excretion of *O*-acetyl carnitine ($r = 0.10$, $P = 1.46 \times 10^{-5}$), a marker of red meat intake (24, 49), which in turn has been associated with long-term weight gain (50). As expected, the association of *O*-acetyl carnitine with BMI was reduced when adjusted for total dietary animal protein intake, although it remained significant ($r = 0.07$, $P = 2.93 \times 10^{-3}$).

We found several differences in metabolic associations with BMI between men and women based on significant gender-interaction term (tables S5 and S6, see Materials and Methods). For example, there was a greater association of BMI with cystine (dicysteine) in women than men in both the U.S. and UK samples; the important role of cystine (and cysteine) in insulin signaling and glucose and lipid metabolism (51) might be manifested as different patterns of

fat deposition between men and women (52). We also found a number of ^1H NMR signals from as yet unidentified metabolites that were associated with BMI (tables S4 and S7).

Replication and validation of BMI-metabolite associations in independent samples

We show that the significant associations with BMI of all 29 ^1H NMR-identified metabolites from analysis of the first set of urine samples could be replicated in the second set of urine samples from the U.S. cohort, obtained on average 3 weeks later (table S8). We then analyzed data from INTERMAP UK to validate our findings from the U.S. samples in an independent data set. The same methods and protocol (8) were used for the U.S. and UK INTERMAP sample data collection, and we have reported that urinary metabolite excretion patterns measured by ^1H NMR spectroscopy were similar (7). We show here that the metabolic signatures of adiposity are consistent across the U.S. and UK population samples (Table 3 and table S9). Specifically, the UK population samples replicated 25 of the 29 structurally identified significant BMI-metabolite associations at a 5% ST-FDR; the exceptions were citrate, NMNA, proline betaine, and 3-hydroxymandelate, possibly reflecting reduced statistical power for these analyses because of the smaller sample size in the UK cohort.

Reliability of urinary metabolite excretion patterns

We previously demonstrated that the analytical (technical) reproducibility of the ^1H NMR urinary metabolite data in INTERMAP was >98% by comparing data from identical samples that were split in the field and analyzed blindly using different identification numbers (53). To check the stability of the metabolic network (below), we calculated the reliability of the urinary metabolite excretion data by comparing results from ^1H NMR analysis of the first and second urine collections for each individual. Specifically, we calculated the ICCs (table S3) of the BMI-associated urinary metabolites detected by ^1H NMR spectroscopy. The median ICC coefficient across all 7100 ^1H NMR spectral variables was 0.37 for the U.S. cohort [interquartile range (IQR), 0.14 to 0.51] and 0.30 for the UK cohort (IQR, 0.11 to 0.49) (fig. S1). For the 29 identified metabolites significantly associated with BMI from ^1H NMR data, the ICC coefficients ranged from 0.34 (pseudouridine) to 0.72 (NMNA) in the U.S. cohort and up to 0.78 (tyrosine + 4-hydroxymandelate) in the UK cohort (table S3) (median ICC coefficient of identified metabolites is 0.53). These ICC coefficients are similar to those described previously for a range of metabolites in serum (54, 55) and urine (56) based on mass spectrometry (MS) and gas chromatography. They indicate a modest degree of attenuation of observed (simple) correlations of metabolites with, for example, BMI in the U.S. samples of between 20 and 70% for ^1H NMR measurements from a single 24-hour urine collection per person.

Metabolic network map of adiposity

Metabolic pathways, as generally depicted, describe the sequential chemical reactions and conversions of substrates and products involved in multiple anabolic and catabolic processes. However, these representations do not easily capture the differential compartmental activities of these pathways or indeed the symbiotic co-metabolic interactions between, say, the gut microbiome and the host or their complex stochastic interdependencies (57–59). We present an integrated metabolic reaction network of human adiposity based on the results from the

U.S. population samples, using the publicly available MetaboNetworks program (60), to show the shortest metabolic paths linking the individual BMI-associated metabolites. The software matches the metabolite biomarker excretion patterns to a preexisting model of all possible scalar distances between metabolites in the Kyoto Encyclopedia of Genes and Genomes (KEGG) with the possibility to include gut microbial species in the model (60). The metabolic systems map includes the significant metabolites from NMR discovery and IEC after adjustment for confounders in model 2 and those that became significant only after the adjustment for creatinine (model 3). The resultant network effectively summarizes the human systemic urinary metabolic signatures of adiposity (Fig. 3). The network map shows (i) the metabolites associated with BMI and the direction of association and (ii) major metabolic functions classified via a colored overlay showing the class or compartment associations of the metabolites (Fig. 3 and tables S10 and S11). We show here that BCAA energy metabolism is linked to various energy-related pathways including both the TCA cycle and a series of TCA anaplerotic replenishment reactions, that is, chemical reactions that form intermediates that are then used in the TCA cycle for energy production. Also, the metabolic map indicates links between muscle and BCAA energy metabolism with a large proportion of the network being related to lipid metabolism. The amino acid excretion data shown here suggest a much broader involvement in multiple pathways disturbed in adiposity than for the BCAAs alone. Notably, the gut microbial-related markers of BMI map onto several pathways embedded in the host multi-compartmental network. For instance, 4-cresyl sulfate, hippurate, phenyl-acetylglutamine, and 3- and 4-hydroxymandelate map to the aromatic amino acid pathway cluster, and dimethylamine and TMA map to choline metabolism, as does formate. Also, hippurate (benzoylglycine) production is mediated via coenzyme A adduct formation, which is mitochondrial (as is the case for all coenzyme A-mediated amino acid conjugations), and therefore is linked indirectly to mitochondrial energy metabolism that powers the reaction (61).

Joint covariate and metabolite models

We investigated which of the urinary metabolites appeared to contribute most to the metabolic signature of adiposity in two ways. First, we constructed a saturated regression model that includes model 1 and 2 covariates and all of the significantly associated (with BMI) metabolites together, and examined each metabolite in turn for its “residual” association with BMI (table S12). Then, we used the elastic net (EN) approach (62) to fit a more parsimonious joint model of the covariate data and the metabolic variables (table S13) to give a sense of the important variables correlated with BMI. The optimal EN parameters (λ and α) were found using 10-fold cross-validation to avoid overfitting the model. Some metabolites found by ^1H NMR spectroscopy have signals in heavily overlapping spectral regions containing resonances from compounds such as ethanolamine, 3-methylhistidine, and lysine, which were also measured independently by IEC (Table 2). When there was a choice between a metabolite measured with NMR and IEC, the most significant was retained.

In the fully saturated model with model 2 covariates, based on a conservative Bonferroni threshold of $P = 2.50 \times 10^{-4}$ for the U.S. samples (table S12A), urinary glycoproteins, cystine, tyrosine, ethanolamine, and 3-methylhistidine remained significantly and directly

associated with BMI ($P= 4.44 \times 10^{-6}$ to $P= 1.99 \times 10^{-35}$), whereas glutamine, 4-cresyl sulfate, O-acetyl carnitine, the ketoleucine/leucine ratio, taurine, and asparagine were inversely associated ($P= 1.60 \times 10^{-4}$ to $P= 6.17 \times 10^{-9}$). This confirmed the importance in relation to BMI of muscle metabolism, urinary amino acids, gut microbial co-metabolism, and dietary meat-related markers; urinary glycoproteins, tyrosine, and 3-methylhistidine predominated in the smaller UK sample data (table S12B).

A more comprehensive set of discriminatory variables emerges with use of the EN to select significant BMI-associated metabolites in a joint modeling approach (table S13). In addition to the metabolites significantly associated with BMI in the fully saturated model for the U.S. samples (table S12A), *N*-acetyl neuraminic acid was directly associated with BMI, whereas succinate, citrate, hippurate, NMNA, proline betaine, 3-hydroxymandelate, and serine were inversely associated, further supporting the importance of gut microbial co-metabolites, urinary amino acids, and diet-related variables, as well as TCA cycle intermediates in determining the metabolic signature of adiposity (tables S13 and S14). Finally, as further validation of significant BMI-metabolite associations, we applied the U.S. EN model (excluding terms for population sample) to the UK sample data and found that the U.S. model provides a good fit to the UK data (table S14 and fig. S2) (F test: $F_{52,391} = 4.04$, $P= 9.12 \times 10^{-16}$).

DISCUSSION

We have found multiple strong associations between adiposity and urinary metabolites. These include metabolites associated with renal function, gut microbial co-metabolites, TCA intermediates, skeletal muscle turnover and mitochondrial metabolism, BCAA metabolism, and dietary intake. We have linked the urinary metabolites together in an integrated biochemical network that delineates the urinary metabolic signature of human adiposity, reproducible both across populations and among individuals over time. The network map shows dependencies and connectivities between metabolic pathways associated with adiposity and provides a systems and compartmental overview of these metabolic disturbances. Many of the metabolites remain independently associated with BMI when accounting for all variables simultaneously, and point to the importance of metabolism, environment, diet, and life-style in the ongoing obesity epidemic.

Notably, in the Results section, we observed the products of five distinct gut microbial-host co-metabolic pathways linked to BMI. The first concerns the action of the gut microbiota on dietary choline to produce TMA via choline TMA-lyases (63), with subsequent conversion of TMA to TMA-N-oxide via the FMO3 enzyme (64). This pathway has been linked to the development of atherosclerosis and cardiovascular disease (65, 66). Increased microbial co-metabolism of choline has also been associated with steatosis, fatty liver, and insulin resistance in animals fed high-fat diets (67). In humans, an increase in energy intake induces rapid changes in the gut microbiota and concomitant changes in energy absorption and energy loss in the stool (68); this may contribute to the associations observed for BMI with total energy intake as well as with energy from animal protein and total and saturated fat (69).

The second and third major pathways of gut microbial activity concern the distal colonic putrefaction of proteins by the microbiota, which leads to the production of 4-cresol from tyrosine and phenylacetylglutamine from phenylalanine; 4-Cresol is O-sulfated in the liver and excreted in the urine. 4-cresyl sulfate and phenylacetylglutamine are two of the three major human gut microbial urinary co-metabolites (the other is hippurate) (21). The production of 4-cresol from 4-hydroxyphenylacetic acid is a purely microbial enzymatic conversion performed by clostridial species (Firmicutes) (70) and reflects variation in microbial genomic activity and composition. Our findings of an association between 4-cresyl sulfate and adiposity is consistent with previous findings regarding differences in the Firmicutes/Bacteroidetes ratio in obese versus lean individuals (71), although there remains debate regarding the mechanistic connection between this microbial ratio and adiposity (72).

Another major gut microbial-host co-metabolic pathway connected to adiposity concerns the production and excretion of hippurate. The proximal colonic metabolism of polyphenolic compounds (including catechins and chlorogenic acid) results in production of benzoic acid, which is absorbed and then phase 2 glycine-conjugated to form hippurate in the host hepatic mitochondria (33). Hippurate excretion is inversely correlated with BMI (Table 1) and is reproducibly sensitive to dietary manipulation in individuals (73). A close connection of hippurate modulation with TCA cycle intermediates has been observed toxicologically where mitochondrial activity has been compromised. These metabolic connectivities may arise because of the common compartmental location (61) as the glycine conjugation of benzoic acid is coenzyme A-mediated and occurs in the hepatic mitochondria (33). This co-compartmentalization of the metabolic processes rather than simple enzymatic activity explains why the close relationship between hippurate and TCA cycle intermediate excretion is not directly apparent from Fig. 3, which maps shortest path connectivities between metabolites based on the KEGG pathways. The final gut microbial pathway associated with BMI concerns the urinary excretion of 2-hydroxyisobutyrate, which has been reported previously only among morbidly obese individuals (17) including those undergoing bariatric surgery (18).

The close connections between metabolites in multiple human body compartments and the products of gut microbial activities have not previously been demonstrated in an embedded symbiotic metabolic network (Fig. 3). Gut microbial disorders and microbe-host metabolic activities and signaling abnormalities have been implicated in the etiopathogenesis of several common chronic diseases (74) as well as premorbid conditions including obesity (71, 75) and insulin resistance (76). Specifically, there is increasing evidence for a key role of the microbiota in the modulation and control of immune responses and of inflammatory pathways (74, 77, 78). The possible connection between microbiome composition and obesity in humans (71) may also reflect the fact that calorific availability can be affected by microbial activity, for example, short-chain fatty acid production from dietary fiber by colonic microorganisms (78). Transplant of fecal material from third-trimester pregnant women or obese individuals into germ-free mice results in an obese phenotype (79,80). We show here independent evidence of the multiple functional consequences of micro-biome activities in human adiposity.

The TCA cycle intermediates citrate and succinate are key energy metabolites in the Krebs cycle and are inversely correlated with BMI (Table 1). Clearly, the greater loss of these metabolites to the urine in individuals with lower BMI represents a small calorific loss to the body. However, the main driver for excretion of TCA cycle intermediates likely relates to renal physiological function, such as renal tubular acid-base balance. In renal tubular acidosis, for example, there is reduced excretion of TCA cycle intermediates because of their increased utilization in proximal tubular mitochondria, a well-known effect of drugs that cause renal tubular acidosis such as acetazolamide and certain heavy metal toxins (81).

Given the high proportion of skeletal muscle mass in the body (~30% in women and 38% in men) (82), and the high level of resting energy consumption and even higher level of exercise energy consumption (90% of whole-body O₂ uptake) in skeletal muscle (83), it is unsurprising that markers of muscle turnover and mass correlate with BMI. Our finding that the ketoleucine/leucine ratio correlates strongly with BMI is a new observation related to the utilization of BCAAs in muscle metabolism. The BCAAs have been extensively studied with respect to obesity; typically, blood concentrations of BCAAs are markedly elevated in obese individuals (10–16). Data in humans show increased utilization of leucine and conversion to ketoleucine in muscle with increased exercise; the hyperaminoacidemia associated with lack of exercise may contribute to insulin resistance in obesity (10,11). The leucine aminotransferase enzyme that controls the key leucine to ketoleucine step in the muscle mitochondria is induced by exercise to respond to increased energy demand by enhancing BCAA metabolism in skeletal muscle (37); the inverse association of the ketoleucine/leucine ratio with BMI observed here may reflect the important role of exercise in controlling adiposity. Most enzymes connected with leucine metabolism and other metabolites detected here have not been found in large-scale genome-wide association studies (GWAS) of obesity (84). This may reflect the relative lack of sensitivity of GWAS to detect effects in inducible enzyme systems under a high degree of environmental control, in this case, physical activity.

We confirmed the strong association of 3-methylhistidine with BMI, reflecting its association with muscle mass and dietary meat intake (24, 25, 39, 40). We also report a strong association of BMI with ethanolamine. However, unlike 3-methylhistidine, we find here that ethanolamine is largely uncorrelated with dietary intake. We thus propose that ethanolamine may be an intrinsic marker of skeletal muscle turnover, reflecting its important structural role in the skeletal muscle sarcolemma (85).

We identified a wide range of urinary markers associated with adiposity. These findings may have clinical utility. For example, it may be possible to identify nonobese individuals with an obesogenic urinary profile, who are at high risk of developing obesity and related metabolic disorders. Such individuals may benefit most from a personalized approach to obesity prevention including tailored individual life-style advice/interventions (for example, improved diet, increased physical activity, and pre- or probiotics) as an early preventive strategy to reduce or reverse the adverse metabolic patterns associated with obesity. In this way, the future disease burden associated with the obesity epidemic may be reduced. Such possibilities should be assessed in future studies.

Our study has a number of strengths. We included eight representative U.S. population samples from the INTERMAP study, and validated our findings using data from the INTERMAP UK population samples, with commonality of data collection methods, urinary specimen storage, and analysis. This design enabled us both to replicate our findings robustly in independent population samples and increase generalizability across geographically distant countries at the forefront of the obesity epidemic (1,2). Standardized data collection in INTERMAP included collection of two 24-hour urine samples, providing integrated urinary metabolite excretion concentrations over a 24-hour period, and four multipass 24-hour dietary recalls per person (8, 9). The repeated urinary measures allowed us to replicate our findings in the second set of urine collections and to demonstrate reproducibility of urinary metabolite excretion over about 3-week period (table S3). We included both untargeted and targeted approaches to the analysis of urinary metabolites, thus maximizing discovery of metabolites in relation to BMI while including known associations with the BCAAs and possible relationships with other amino acids and related compounds. By using a network mapping approach, we were able to show pathway connectivities between the urinary metabolites significantly associated with BMI, reflecting different systems and sources of metabolites contributing to the metabolic signatures of adiposity.

Our study also has a number of limitations. Most importantly, it was largely cross-sectional involving a follow-up period of just 3 weeks, and therefore, cause and effect associations of urinary metabolites with BMI cannot be inferred directly. Specifically, we used mainly an untargeted metabolic profiling approach to identify combinations of correlations between urinary metabolites (reflecting whole system activity including that of the gut microbiota) and adiposity in two Western populations. We have identified multiple markers of adiposity and have extended the range of pathways and biochemical processes that can be linked to adiposity. In the absence of a strong causal model and directly supportive experimental data, we can point to the metabolic consequences of adiposity and their interconnections but rely on published network and reaction information (KEGG) to identify the intermediate steps and pathways linking the metabolic signatures of adiposity. Further downstream work, beyond the scope of this study, is needed to identify the specific alterations leading to the metabolic disturbances documented here. We have identified a series of interrelated pathways, linking both human metabolism and gut microbial cometabolism, which provide a systems-level overview of the metabolic disturbances and consequences of adiposity. This approach is also scalable and translatable to other types of epidemiological study where tabulated data can be expressed in an interconnected metabolic network format.

BMI is a well-established measure of adiposity, highly correlated with the percentage of body fat (27, 86, 87) and strongly associated with cardiometabolic outcomes such as diabetes and cardiovascular disease (3). However, BMI is a relatively crude measure of general adiposity; inclusion of other measures, for example, waist and hip circumference (unavailable in INTERMAP) as a measure of central adiposity, may add more specificity to the phenotype (88) and allow refinement of metabolic phenotypes between different adiposity measures. In addition to percentage of body fat, BMI correlates with total muscle mass (27), so it was important to adjust in our analyses for urinary creatinine (model 3), as a marker of muscle turnover (26). Also, BMI does not take into account gender differences in fat distribution. We therefore adjusted the BMI-metabolite associations for gender (model 1)

and looked separately at gender interactions in the urinary excretion of metabolites in relation to BMI (tables S5 and S6).

Our untargeted analyses are based on ^1H NMR spectroscopy, which has both advantages and disadvantages compared with MS-based assays (19). Whereas NMR spectroscopy gives superior spectral reproducibility to chromatography linked to MS, which is important for large-scale metabolic profiling studies as undertaken here, it is usually less sensitive than MS for a given analytical problem (5, 6, 19, 89). However, many NMR-detectable metabolites are in major metabolic pathways and represent the key currencies of metabolic exchange around the body and thus are highly informative of relative pathway activities (6). Here, the spectra have been used as semiquantitative or relative quantification readouts where the patterns and combinations of metabolites are of interest rather than their absolute values. Finally, structural identification of metabolites can be challenging particularly where peaks overlap or low-intensity signals are detected close to baseline; we made considerable efforts through both statistical approaches and experimentation to identify the structure of adiposity-associated metabolites so these could be placed into the network model.

In summary, the relationships shown here between BMI and the 24-hour urinary metabolite profile, summarized in the network map, characterize a reproducible metabolic signature associated with the modern obesity epidemic. These relationships point to a variety of poorly understood avenues of metabolism that link to BMI, including gut microbial- and skeletal muscle-related energy pathways. Our findings reveal multiple connections between many metabolic compartments and pathways, and provide possible starting points for new approaches to prevention and treatment, for example, functional microbiome modulation (90) and stimulation of skeletal muscle mitochondrial metabolism. Tackling the obesity epidemic has now become an urgent priority; otherwise, gains in life expectancy enjoyed over recent decades may be halted or reversed, and healthcare provision to deal with the sequelae of obesity-related diseases may be overtaken by the demand.

MATERIALS AND METHODS

Study design

The INTERMAP study is investigating dietary and other factors associated with raised blood pressure (8), the major modifiable risk factor underlying cardiovascular disease (91). The INTERMAP study surveyed 4680 men and women ages 40 to 59 years from 17 population samples in four countries (Japan, People's Republic of China, UK, and United States). Participants were randomly recruited from general and occupational population samples in 1996 to 1999. Each participant made four clinic visits: the first two on consecutive days and the second two on average 3 weeks later, also on consecutive days. The data obtained include eight blood pressure measurements (two per visit), four 24-hour dietary recalls obtained by a trained interviewer using the multipass recall method, measurements of height and weight, and extensive questionnaire information. Two-timed 24-hour urine collections were obtained on the second and fourth visits. The present report relates to data from 1880 of the 2195 U.S. INTERMAP participants used in the discovery and replication phase and 444 of the 501 UK INTERMAP participants used in the validation phase (Fig. 1). Participant exclusion criteria are outlined below and in Fig. 1. This study is reported in accordance with

the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (92).

Data collection and pretreatment (NMR)

The urine specimens were prepared for, and analyzed with, high-resolution NMR spectroscopy using a Bruker Avance III spectrometer, operating at 600.29 MHz for ^1H , equipped with a 5-mm TCI Z-gradient CryoProbe using a standard one-dimensional (1D) pulse sequence with water suppression (7, 93). Free induction decays were Fourier-transformed, referenced to an internal standard (trimethylsilyl propionate), and baseline- and phase-corrected by use of an in-house software. The spectral regions containing water and urea (δ 6.4 to 4.5), the internal standard (δ 0.2 to -0.2), and regions δ 0.5 to 0.2, δ -0.2 to -4.5 , and δ 15.5 to 9.5 were removed before normalization via the median fold change method (94). The remaining variables were “binned” to 7100 variables with bin widths of 0.001 ppm. Metabolic outliers were excluded from the data set by use of Hotelling’s T^2 statistic on the scores of the principal components analysis using Pareto-scaled data (dividing each variable by the square root of the SD). Metabolic outliers were defined as participants whose scores, for either urine collection, mapped outside the Hotelling’s T^2 ellipse (confidence interval = 0.95) in a cross-validated seven-component model; these participants often had high ethanol or nonsteroidal antiinflammatory drug (NSAID) excretion (7). Participants with a doctor diagnosis of diabetes mellitus were also excluded (Fig. 1).

Data collection and pretreatment (IEC)

Urine specimens (first urine collection) were analyzed with IEC to quantify the concentration of amino acids and related compounds using the Biochrom 20+ and 30+ Amino Acid Analyser with Midas autosampler. The amino acids and related compounds were separated using mobile phases with different pH and by varying temperature. Ninhydrin was used to derivatize the compounds after separation followed by measurement of the absorbance at 570 or 440 nm (depending on the compound) of the colored complexes. EZChrom Elite was used to integrate the peaks automatically and calculate the concentration using an internal standard (2,4-diaminobutyrate) as a reference. The robustness of the method was checked using quality control samples, and compounds were excluded where the variation in concentration was bigger than ± 2.5 SD. Zeros in the data matrix were replaced by the lowest nonzero value in each column before log normalization.

Metabolite identification

Statistical total correlations spectroscopy (STOCSY) (28) and subset optimization by reference matching (STORM) (29) were used to identify structural correlations in the spectral data. A Bruker reference library was used to assign the chemical shifts to metabolites. Confirmation of metabolite identification was achieved using a combination of strategies. Specimens containing the largest amount of each unidentified metabolite were selected using an in-house MATLAB script and/or STORM. A catalog of standard 1D ^1H NMR pulse sequence with water peak presaturation and 2D NMR methods such as 2D J -resolved, 2D ^1H - ^{13}C heteronuclear single quantum coherence (HSQC), ^1H - ^1H correlation spectroscopy (COSY), ^1H - ^1H total correlation spectroscopy (TOCSY), and diffusion-edited

NMR experiments were applied to the selected specimens to confirm the signals and their chemical shifts. Pretreatment of urine with solid-phase extraction (SPE) methods was carried out for some metabolites to simplify urine specimens and isolate specific signals according to their polarity and dissociation constant. All the SPE fractions were analyzed by 1D NMR spectroscopy. Metabolites were confirmed by in situ spiking experiments using authentic chemical standards or enzymatic methods. The strategy used for each identified metabolite is shown in table S2. Table S3 shows the (3-week) ICC (Eq. 1) from comparison of excretion levels in the first and second urine collections for each metabolite. Here, σ_B^2 indicates the between-sample variance and σ_W^2 indicates the within-sample variance.

$$\text{ICC} = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2} \quad (1)$$

The ICCs were calculated in R using the “irr” package (95). The degree of bias for standard correlation was calculated as in Eq. 2, where a BMI attenuation factor of 0.98 (96) was used.

$$\frac{1}{\sqrt{\text{ICC}_{\text{metabolite}} \times \text{ICC}_{\text{BMI}}}} - 1 \quad (2)$$

Statistical analysis

Partial correlation models.—Partial correlation models were run for the NMR data (7100 variables) and IEC data (22 variables) of the first urine collection with adjustment for progressively increasing numbers of confounding variables. Model 1 adjusts for demographic variables: age, gender, and sample; model 2 adds medical and life-style factors: history of heart disease or stroke, moderate to heavy physical activity (hours/day), medication for hypertension, prescribed lipidlowering drugs, NSAID use, dietary supplement use, special diet, smoking status, years of education, and total energy intake per day (kcal/day); and model 3 adds 24-hour urinary creatinine (mmol/day). ^1H NMR data of the second urine collection are used as replication data set.

The partial (Pearson) correlation is calculated in two steps. First, the data are adjusted for covariates (Eq. 3). Here, X_j indicates the j^{th} variable of X , C indicates a matrix of a constant (column of ones) and the covariates (confounding variables) for a model (M) combined, and i indicates the i^{th} sample.

$$\begin{aligned} R &= Z - C(C^{-1}Z) \text{ where } R = [\tilde{x}\tilde{Y}], \\ Z &= [X; Y], \text{ and } C_i = [1M_i] \end{aligned} \quad (3)$$

Second, the partial correlation is calculated (Eq. 4) using the two columns of R , \tilde{x} , and \tilde{Y} representing the adjusted variable X_j and outcome variable, respectively. Here, r_j is the partial correlation for variable j and n is the number of samples.

$$r_j = \tilde{x}^T \tilde{Y} \sqrt{\left(\sum_{i=1}^n \tilde{x}_i^2 \right) \times \left(\sum_{i=1}^n \tilde{Y}_i^2 \right)} \quad (4)$$

In addition to the adjustment for gender, we investigated the gender-metabolite interaction term, where gender was coded as 0 (men) and 1 (women).

Multiple testing.—The ST-FDR (23) is used to calculate the Q values to correct for multiple testing for the 7100 ^1H NMR variables. The proportion of truly null P values (π_0) was determined using bootstrap resampling (97) as described in (23). ST-FDR takes a list of P values (P) sorted from lowest (P_1) to highest (P_p), where p is the total number of variables (Eq. 5). Next, the algorithm loops back from Q_{p-1} to Q_1 to investigate whether the list of Q values is still sorted from smallest to largest (Eq. 6).

$$Q_i = \frac{\pi_0 \times P_i \times p}{i} \quad (5)$$

$$Q_i = \begin{cases} Q_{i+1} & \text{if } Q_i > Q_{i+1} \\ Q_i & \text{otherwise} \end{cases} \quad (6)$$

For the ^1H NMR discovery analyses (U.S. samples, first urine collection), a threshold of $Q = 0.01$ was used to indicate a significant association with BMI. For the replication analyses (U.S. samples, second urine collection), the threshold was 5%; the same threshold (5%) was used for the validation data sets (UK ^1H NMR data). To avoid spurious associations, spectral variables were designated as significant only if they met the ST-FDR threshold, above, of 1 or 5%, and the two adjacent spectral signals also attained the same threshold in addition to all three variables having the same sign of the correlation.

For the targeted IEC data (22 amino acids and related compounds), a Bonferroni threshold was used to determine significance. For the discovery analyses (U.S.), a threshold of $P = 4.55 \times 10^{-4}$ ($P = 0.05/2$) was used. For the validation analyses (UK), a threshold of $P = 2.50 \times 10^{-3}$ ($P = 0.05/20$) was used (methionine and arginine were excluded). Second urine data were not available for the IEC analyses.

The gender-interaction analyses were adjusted for multiple testing using a Bonferroni threshold of $P = 3.33 \times 10^{-4}$ ($P = 0.01/30$), $P = 1.67 \times 10^{-3}$ ($P = 0.05/30$), and $P = 4.55 \times 10^{-4}$ ($P = 0.05/2$) for ^1H NMR discovery, ^1H NMR replication, and IEC discovery,

respectively. For the validation analyses, the thresholds were $P = 1.67 \times 10^{-3}$ ($P = 0.05/30$) for ^1H NMR and $P = 2.50 \times 10^{-3}$ ($P = 0.05/20$) for IEC.

Last, partial correlation models were run where the association of each metabolite (^1H NMR and IEC combined) was adjusted for the association of the remaining metabolites (in addition to the covariates in models 1 and 2). In case of overlap between ^1H NMR and IEC, the most significant variable of the two was included. For the discovery analyses, a Bonferroni threshold of $P = 2.56 \times 10^{-4}$ ($P \times 0.01/39$) was used, and for the validation analyses, $P = 1.28 \times 10^{-3}$ ($P \times 0.05/39$) was used.

Joint covariate and metabolite models.—An EN (62) model was used to determine the important contributors to the BMI model with covariates and metabolites combined into one model. The EN model allows for the inclusion of correlated groups of variables in the model while avoiding the inclusion of all variables in the model at the same time, which may lead to overfitting. ENs are a trade-off between the lasso (ℓ_1) (98) and ridge (ℓ_2) (99) penalty; the ℓ_2 penalty forces regression coefficients (β) toward 0, whereas the ℓ_1 forces certain coefficients to be exactly 0 (Eq. 7). Here, β_0 denotes the constant; n , the number of samples; p , the number of variables; Y , the outcome variable; X , the data matrix; λ , the shrinkage factor; and α , the weighing factor between the ℓ_1 and ℓ_2 penalties.

$$\beta_0, \beta \in \mathbb{R}^{p+1} \left[\frac{1}{2n} \sum_{i=1}^n (Y_i - \beta_0 - X_i^T \times \beta)^2 + \lambda \left(\frac{1-\alpha}{2} \|\beta\|_{\ell_2}^2 + \alpha \|\beta\|_{\ell_1} \right) \right] \quad (7)$$

A grid search was performed to find the optimal parameter settings for λ and α . For all combinations of nonnegative λ 's and α 's between 0 and 1, the mean squared error (MSE) of (10-fold) crossvalidation was calculated. To avoid overfitting, the optimal λ and α were chosen as the combination for which the MSE falls within 1 SD of the overall minimal MSE. The coordinate descent algorithm (100) was used to calculate the regularization path for the EN.

Two EN models were calculated; the first used the U.S. data with all covariates and a unique set of 39 metabolites across the ^1H NMR and IEC data sets to identify the important contributors to the BMI model. A second model was then calculated that applied the U.S. model coefficients to the UK data; this required that the variables representing the U.S. population samples were left out of the model; thus, the model applied to the UK data included all covariates (except the population sample variables) and the unique set of 39 metabolites (coefficients for some of which were set to zero).

Metabolic reaction network

Using MetaboNetworks (60), reactions that occur spontaneously or by means of an enzyme linked to a *Homo sapiens* or microbial gene were identified in KEGG. The gut microbiota that are included are those of the most abundant endosymbionts—the phyla Firmicutes, Bacteroidetes, Alphaproteobacteria, Betaproteobacteria, Deltaproteobacteria, Gammaproteobacteria, and Actinobacteria—because these make up 99% of the phylotypes

found in the human gut (101). For each of the reactions, the main reaction pairs were identified, and an adjacency matrix was calculated for all compounds based on the main reactant pairs. The shortest paths between all metabolites significantly associated with BMI were calculated from the adjacency matrix, and a network graph was drawn for all compounds needed to connect all identified metabolites with the shortest paths. In the network, the three dashed lines indicate the closest related metabolite in the network based on literature for metabolites that are not found in the KEGG database.

All calculations were performed in MATLAB (2013a, The MathWorks) and R (3.0.2., R Foundation for Statistical Computing).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank E. Maibaum for carrying out the ^1H NMR spectroscopy of the INTERMAP urine samples, I. Yap for data analyses, and the staff at local, national, and international centers for collecting the INTERMAP data and samples. A partial listing of colleagues can be found in (8).

Funding: The INTERMAP and INTERMAP metabonomic study are supported by the U.S. National Heart, Lung, and Blood Institute (grants R01-HL50490 and R01-HL84228). INTERMAP data collection was also supported by Chicago Health Research Foundation and national agencies in the People's Republic of China, Japan [Ministry of Education, Science, Sports, and Culture, Grant-in-Aid for Scientific Research (A), No. 090357003], and UK. P.E. acknowledges support from Medical Research Council-Public Health England (MRC-PHE) Centre for Environment and Health, National Institute for Health Research (NIHR) Biomedical Research Centre at Imperial College Healthcare National Health Service Trust and Imperial College London, and the NIHR Health Protection Research Unit on Health Impact of Environmental Hazards. P.E. is an NIHR senior investigator. J.M.P. was supported by an MRC-PHE Centre for Environment and Health PhD studentship. We also thank the MRC-NIHR National Phenome Centre for facilitating this and related work. The data reported in this manuscript are tabulated in the main paper and in the Supplementary Materials.

REFERENCES AND NOTES

1. Finucane MM, Stevens GA, Cowan MJ, Danaei G, Lin JK, Paciorek CJ, Singh GM, Gutierrez HHR, Lu Y, Bahalim AN, Farzadfar F, Riley LM, Ezzati M, National, regional, and global trends in body-mass index since 1980: Systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet* 377, 557–567 (2011). [PubMed: 21295846]
2. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, Mullany EC, Biryukov S, Abbafati C, Abera SF, Abraham JP, Abu-Rmeileh NME, Achoki T, AlBuhairan FS, Alemu ZA, Alfonso R, Ali MK, Ali R, Guzman NA, Ammar W, Anwar P, Banerjee A, Barquera S, Basu S, Bennett DA, Bhutta Z, Blore J, Cabral N, Nonato IC, Chang J-C, Chowdhury R, Courville KJ, Criqui MH, Cundiff DK, Dabhadkar KC, Dandona L, Davis A, Dayama A, Dharmaratne SD, Ding EL, Durrani AM, Esteghamati A, Farzadfar F, Fay DFJ, Feigin VL, Flaxman A, Forouzanfar MH, Goto A, Green MA, Gupta R, Hafezi-Nejad N, Hankey GJ, Harewood HC, Havmoeller R, Hay S, Hernandez L, Husseini A, Idrisov BT, Ikeda N, Islami F, Jahangir E, Jassal SK, Jee SH, Jeffreys M, Jonas JB, Kabagambe EK, Khalifa SEAH, Kengne AP, Khader YS, Khang Y-H, Kim D, Kimokoti RW, Kinge JM, Kokubo Y, Kosen S, Kwan G, Lai T, Leinsalu M, Li Y, Liang X, Liu S, Logroscino G, Lotufo PA, Lu Y, Ma J, Mainoo NK, Mensah GA, Merriman TR, Mokdad AH, Moschandreas J, Naghavi M, Naheed A, Nand D, Narayan KMV, Nelson EL, Neuhouser ML, Nisar MI, Ohkubo T, Oti SO, Pedroza A, Prabhakaran D, Roy N, Sampson U, Seo H, Sepanlou SG, Shibuya K, Shiri R, Shiue I, Singh GM, Singh JA, Skirbekk V, Stapelberg NJC, Sturua L, Sykes BL, Tobias M, Tran BX, Trasande L, Toyoshima H, van de Vijver S, Vasankari TJ, Veerman JL, Velasquez-Melendez G,

- Vlassov VV, Vollset SE, Vos T, Wang C, Wang X, Weiderpass E, Werdecker A, Wright JL, Yang YC, Yatsuya H, Yoon J, Yoon S-J, Zhao Y, Zhou M, Zhu S, Lopez AD, Murray CJL, Gakidou E, Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 384, 766–781 (2014). [PubMed: 24880830]
3. Whitlock G, Lewington S, Sherliker P, Clarke R, Emberson J, Halsey J, Qizilbash N, Collins R, Peto R, Body-mass index and cause-specific mortality in 900,000 adults: Collaborative analyses of 57 prospective studies. *Lancet* 373, 1083–1096 (2009). [PubMed: 19299006]
 4. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, Amann M, Anderson HR, Andrews KG, Aryee M, Atkinson C, Bacchus LJ, Bahalim AN, Balakrishnan K, Balmes J, Barker-Collo S, Baxter A, Bell ML, Blore JD, Blyth F, Bonner C, Borges G, Bourne R, Boussinesq M, Brauer M, Brooks P, Bruce NG, Brunekreef B, Bryan-Hancock C, Bucello C, Buchbinder R, Bull F, Burnett RT, Byers TE, Calabria B, Carapetis J, Carnahan E, Chafe Z, Charlson F, Chen H, Chen JS, Cheng AT-A, Child JC, Cohen A, Colson KE, Cowie BC, Darby S, Darling S, Davis A, Degenhardt L, Dentener F, Des Jarlais DCD, Devries K, Dherani M, Ding EL, Dorsey ER, Driscoll T, Edmond K, Ali SE, Engell RE, Erwin PJ, Fahimi S, Falder G, Farzadfar F, Ferrari A, Finucane MM, Flaxman S, Fowkes FGR, Freedman G, Freeman MK, Gakidou E, Ghosh S, Giovannucci E, Gmel G, Graham K, Grainger R, Grant B, Gunnell D, Gutierrez HR, Hall W, Hoek HW, Hogan A, Hosgood HD, Hoy D, Hu H, Hubbell BJ, Hutchings SJ, Ibeanusi SE, Jacklyn GL, Jasrasaria R, Jonas JB, Kan H, Kanis JA, Kassebaum N, Kawakami N, Khang Y-H, Khatibzadeh S, Khoo J-P, Kok C, Laden F, Lalloo R, Lan Q, Lathlean T, Leasher JL, Leigh J, Li Y, Lin JK, Lipshultz SE, London S, Lozano R, Lu Y, Mak J, Malekzadeh R, Malingler L, Marcenes W, March L, Marks R, Martin R, McGale P, McGrath J, Mehta S, Mensah TR, Merriman GA, Micha R, Michaud C, Mishra V, Hanafiah K.Mohd, Mokdad AA, Morawska L, Mozaffarian D, Murphy T, Naghavi M, Neal B, Nelson PK, Nolla JM, Norman R, Olives C, Omer SB, Orchard J, Osborne R, Ostro B, Page A, Pandey KD, Parry CDH, Passmore E, Patra J, Pearce N, Pelizzari PM, Petzold MR, Phillips MR, Pope D, Pope CA, Powles J, Rao M, Razavi H, Rehfuess EA, Rehm JT, Ritz B, Rivara FP, Roberts T, Robinson C, Rodriguez-Portales JA, Romieu I, Room R, Rosenfeld LC, Roy A, Rushton L, Salomon JA, Sampson U, Sanchez-Riera L, Sanman E, Sapkota A, Seedat S, Shi P, Shield K, Shivakoti R, Singh GM, Sleet DA, Smith E, Smith KR, Stapelberg NJC, Steenland K, Stöckl H, Stovner LJ, Straif K, Straney L, Thurston GD, Tran JH, Van Dingenen R, van Donkelaar A, Veerman JL, Vijayakumar L, Weintraub R, Weissman MM, White RA, Whiteford H, Wiersma ST, Wilkinson JD, Williams HC, Williams W, Wilson N, Woolf AD, Yip P, Zielinski JM, Lopez AD, Murray CJL, Ezzati M, AlMazroa MA, Memish ZA, A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 380, 2224–2260 (2012). [PubMed: 23245609]
 5. Holmes E, Wilson ID, Nicholson JK, Metabolic phenotyping in health and disease. *Cell* 134, 714–717 (2008). [PubMed: 18775301]
 6. Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC, Metabolic phenotyping in clinical and surgical environments. *Nature* 491, 384–392 (2012). [PubMed: 23151581]
 7. Holmes E, L Loo R, Stamler J, Bictash M, Yap IKS, Chan Q, Ebbels T, De Iorio M, Brown IJ, Veselkov KA, Daviglus ML, Kesteloot H, Ueshima H, Zhao LC, Nicholson JK, Elliott P, Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* 453, 396–400 (2008). [PubMed: 18425110]
 8. Stamler J, Elliott P, Dennis B, Dyer AR, Kesteloot H, Liu K, Ueshima H, Zhou BF, INTERMAP: Background, aims, design, methods, and descriptive statistics (nondietary). *J. Hum. Hypertens.* 17, 591–608 (2003). [PubMed: 13679950]
 9. Dennis B, Stamler J, Buzzard M, Conway R, Elliott P, Moag-Stahlberg A, Okayama A, Okuda N, Robertson C, Robinson F, Schakel S, Stevens M, Van Heel N, Zhao L, Zhou BF, Grp IR, INTERMAP: The dietary data—Process and quality control. *J. Hum. Hypertens.* 17, 609–622 (2003). [PubMed: 13679951]
 10. Felig P, Marliss E, Cahill GF Jr., Plasma amino acid levels and insulin secretion in obesity. *N. Engl. J. Med.* 281, 811–816 (1969). [PubMed: 5809519]
 11. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, Haqq AM, Shah SH, Arlotto M, Slentz CA, Rochon J, Gallup D, Ilkayeva O, Wenner BR, Yancy WS Jr., Eisenson H, Musante

- G, Surwit RS, Millington DS, Butler MD, Svetkey LP, A branched- chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* 9, 311–326 (2009). [PubMed: 19356713]
12. Cheng S, Rhee EP, Larson MG, Lewis GD, McCabe EL, Shen D, Palma MJ, Roberts LD, Dejam A, Souza AL, Deik AA, Magnusson M, Fox CS, O'Donnell CJ, Vasani RS, Melander O, Clish CB, Gerszten RE, Wang TJ, Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation* 125, 2222–2231 (2012). [PubMed: 22496159]
 13. Jourdan C, Petersen A-K, Gieger C, Döring A, Illig T, Wang-Sattler R, Meisinger C, Peters A, Adamski J, Prehn C, Suhre K, Altmaier E, Kastenmüller G, Romisch-Margl W, Theis FJ, Krumsiek J, Wichmann H-E, Linseisen J, Body fat free mass is associated with the serum metabolite profile in a population-based study. *PLOS One* 7, e40009 (2012).
 14. Moore SC, Matthews CE, Sampson JN, Stolzenberg-Solomon RZ, Zheng W, Cai Q, Tan YT, Chow W-H, Ji B-T, Liu DK, Xiao Q, Boca SM, Leitzmann MF, Yang G, Xiang YB, Sinha R, Shu XO, Cross AJ, Human metabolic correlates of body mass index. *Metabolomics* 10, 259–269 (2014). [PubMed: 25254000]
 15. Xie G, Ma X, Zhao A, Wang C, Zhang Y, Nieman D, Nicholson JK, Jia W, Bao Y, The metabolite profiles of the obese population are gender-dependent. *J. Proteome Res.* 13, 4062–4073 (2014). [PubMed: 25132568]
 16. Würtz P, Wang Q, Kangas AJ, Richmond RC, Skarp J, Tiainen M, Tynkkynen T, Soininen P, Havulinna AS, Kaakinen M, Viikari JS, Savolainen MJ, Kähönen M, Lehtimäki T, Männistö S, Blankenberg S, Zeller T, Laitinen J, Pouta A, Mäntyselkä P, Vanhala M, Elliott P, Pietiläinen KH, Ripatti S, Salomaa V, Raitakari OT, Jarvelin MR, Smith GD, Ala-Korpela M, Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. *PLOS Med.* 11, e1001765 (2014).
 17. Calvani R, Miccheli A, Capuani G, Tomassini Miccheli A, Puccetti C, Delfini M, Iaconelli A, Nanni G, Mingrone G, Gut microbiome-derived metabolites characterize a peculiar obese urinary metabolite. *Int. J. Obes.* 34, 1095–1098 (2010).
 18. Friedrich N, Budde K, Wolf T, Jungnickel A, Grotevendt A, Dreßler M, Völzke H, Blüher M, Nauck M, Lohmann T, Wallaschofski H, Short-term changes of the urine metabolome after bariatric surgery. *OMICS* 16, 612–620 (2012). [PubMed: 23095112]
 19. Bictash M, Ebbels TM, Chan Q, Loo RL, Yap IKS, Brown IJ, de Iorio M, Daviglus ML, Holmes E, Stamler J, Nicholson JK, Elliott P, Opening up the “Black Box”: Metabolic phenotyping and metabolome-wide association studies in epidemiology. *J. Clin. Epidemiol.* 63, 970–979 (2010). [PubMed: 20056386]
 20. Mittelstrass K, Ried JS, Yu Z, Krumsiek J, Gieger C, Prehn C, Roemisch-Margl W, Polonikov A, Peters A, Theis FJ, Meitinger T, Kronenberg F, Weidinger S, Wichmann HE, Suhre K, Wang-Sattler R, Adamski J, Illig T, Discovery of sexual dimorphisms in meta-bolic and genetic biomarkers. *PLOS Genet.* 7, e1002215 (2011).
 21. Wijeyesekera A, Clarke PA, Bictash M, Brown IJ, Fidock M, Ryckmans T, Yap IKS, Chan Q, Stamler J, Elliott P, Holmes E, Nicholson JK, Quantitative UPLC-MS/MS analysis of the gut microbial co-metabolites phenylacetylglutamine, 4-cresyl sulphate and hippurate in human urine: INTERMAP study. *Anal. Methods* 4, 65–72 (2012). [PubMed: 23946767]
 22. Storey JD, A direct approach to false discovery rates. *J. R. Stat. Soc. B* 64, 479–498 (2002).
 23. Storey JD, Tibshirani R, Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9440–9445 (2003). [PubMed: 12883005]
 24. Stella C, Beckwith-Hall B, Cloarec O, Holmes E, Lindon JC, Powell J, van der Ouderaa F, Bingham S, Cross AJ, Nicholson JK, Susceptibility of human metabolic phenotypes to dietary modulation. *J. Proteome Res.* 5, 2780–2788 (2006). [PubMed: 17022649]
 25. Cross AJ, Major JM, Sinha R, Urinary biomarkers of meat consumption. *Cancer Epidemiol. Biomarkers Prev.* 20, 1107–1111 (2011). [PubMed: 21527577]
 26. Heymsfield SB, Arteaga C, McManus C, Smith J, Moffitt S, Measurement of muscle mass in humans: Validity of the 24-hour urinary creatinine method. *Am. J. Clin. Nutr.* 37, 478–494 (1983). [PubMed: 6829490]

27. Micozzi MS, Albanes D, Jones DY, Chumlea WC, Correlations of body mass indices with weight, stature, and body composition in men and women in NHANES I and II. *Am. J. Clin. Nutr.* 44, 725–731 (1986). [PubMed: 3788825]
28. Cloarec O, Dumas M-E, Craig A, Barton RH, Trygg J, Hudson J, Blancher C, Gauguier D, Lindon JC, Holmes E, Nicholson J, Statistical total correlation spectroscopy: An exploratory approach for latent biomarker identification from metabolic ¹H NMR data sets. *Anal. Chem.* 77, 1282–1289 (2005). [PubMed: 15732908]
29. Posma JM, Garcia-Perez I, De Iorio M, Lindon JC, Elliott P, Holmes E, Ebbels TMD, Nicholson JK, 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 (2012). [PubMed: 23151027]
30. Beckonert O, Keun HC, Ebbels TMD, Bundy J, Holmes E, Lindon JC, Nicholson JK, Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protoc.* 2, 2692–2703 (2007). [PubMed: 18007604]
31. Lever M, Slow S, The clinical significance of betaine, an osmolyte with a key role in methyl group metabolism. *Clin. Biochem.* 43, 732–744 (2010). [PubMed: 20346934]
32. Stohs SJ, Preuss HG, Shara M, A review of the receptor-binding properties of p-synephrine as related to its pharmacological effects. *Oxid. Med. Cell. Longev.* 2011, 482973 (2011).
33. Lees HJ, Swann JR, Wilson ID, Nicholson JK, Holmes E, Hippurate: The natural history of a mammalian-microbial cometabolite. *J. Proteome Res.* 12, 1527–1546 (2013). [PubMed: 23342949]
34. Schnackenberg LK, Sun J, Espandiari P, Holland RD, Hanig J, Beger RD, Metabo-nomics evaluations of age-related changes in the urinary compositions of male Sprague Dawley rats and effects of data normalization methods on statistical and quantitative analysis. *BMC Bioinformatics* 8, S3 (2007).
35. Wang Y, Lawler D, Larson B, Ramadan Z, Kochhar S, Holmes E, Nicholson JK, Meta- bomic investigations of aging and caloric restriction in a life-long dog study. *J. Proteome Res.* 6, 1846–1854 (2007). [PubMed: 17411081]
36. Li M, Wang B, Zhang M, Rantalainen M, Wang S, Zhou H, Zhang Y, Shen J, Pang X, Zhang M, Wei H, Chen Y, Lu H, Zuo J, Su M, Qiu Y, Jia W, Xiao C, Smith LM, Yang S, Holmes E, Tang H, Zhao G, Nicholson JK, Li L, Zhao L, Symbiotic gut microbes modulate human metabolic phenotypes. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2117–2122 (2008). [PubMed: 18252821]
37. Wagenmakers AJM, Brookes JH, Coakley JH, Reilly T, Edwards RHT, Exercise-induced activation of the branched-chain 2-oxo acid dehydrogenase in human muscle. *Eur. J. Appl. Physiol. Occup. Physiol.* 59, 159–167 (1989).
38. Brosnan JT, Brosnan ME, Branched-chain amino acids: Enzyme and substrate regulation. *J. Nutr.* 136, 207S–211S (2006). [PubMed: 16365084]
39. Rathmacher JA, Nissen SL, Development and application of a compartmental model of 3-methylhistidine metabolism in humans and domestic animals. *Adv. Exp. Med. Biol.* 445, 303–324 (1998). [PubMed: 9781398]
40. Altorf-van der Kuil W, Brink EJ, Boetje M, Siebelink E, Bijlsma S, Engberink MF, van 't Veer P, Tomé D, Bakker SJ, van Baak MA, Geleijnse JM, Identification of biomarkers for intake of protein from meat, dairy products and grains: A controlled dietary intervention study. *Br. J. Nutr.* 110, 810–822 (2013). [PubMed: 23452466]
41. Helal I, Fick-Brosnahan GM, Reed-Gitomer B, Schrier RW, Glomerular hyperfiltration: Definitions, mechanisms and clinical implications. *Nat. Rev. Nephrol.* 8, 293–300 (2012). [PubMed: 22349487]
42. Bell JD, Brown JC, Nicholson JK, Sadler PJ, Assignment of resonances for 'acute- phase' glycoproteins in high resolution proton NMR spectra of human blood plasma. *FEBS Lett* 215,311–315 (1987). [PubMed: 2438159]
43. Gabay C, Kushner I, Acute-phase proteins and other systemic responses to inflammation. *N. Engl. J. Med.* 340, 448–454 (1999). [PubMed: 9971870]

44. Lang R, Wahl A, Stark T, Hofmann T, Urinary N-methylpyridinium and trigonelline as candidate dietary biomarkers of coffee consumption. *Mol. Nutr. Food Res.* 55, 1613–1623 (2011). [PubMed: 21618426]
45. Rigamonti A, Brennand K, Lau F, Cowan CA, Rapid cellular turnover in adipose tissue. *PLOS One* 6, e17637 (2011).
46. Heinzmann SS, Brown IJ, Chan Q, Bictash M, Dumas M-E, Kochhar S, Stamler J, Holmes E, Elliott P, Nicholson JK, Metabolic profiling strategy for discovery of nutritional biomarkers: Proline betaine as a marker of citrus consumption. *Am. J. Clin. Nutr.* 92, 436–443 (2010). [PubMed: 20573794]
47. Lloyd AJ, Beckmann M, Favé G, Mathers JC, Draper J, Proline betaine and its biotransformation products in fasting urine samples are potential biomarkers of habitual citrus fruit consumption. *Br. J. Nutr.* 106, 812–824 (2011). [PubMed: 21736852]
48. Vázquez-Fresno R, Llorach R, Urpi-Sarda M, Lupianez-Barbero A, Estruch R, Corella D, Fitó M, Arós F, Ruiz-Canela M, Salas-Salvadó J, Andres-Lacueva C, Metabolomic pattern analysis after Mediterranean diet intervention in a nondiabetic population: A 1- and 3-year follow-up in the PREDIMED study. *J. Proteome Res.* 14, 531–540 (2014). [PubMed: 25353684]
49. O’Sullivan A, Gibney MJ, Brennan L, Dietary intake patterns are reflected in metabolomic profiles: Potential role in dietary assessment studies. *Am. J. Clin. Nutr.* 93, 314–321 (2011). [PubMed: 21177801]
50. Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB, Changes in diet and lifestyle and long-term weight gain in women and men. *N. Engl. J. Med.* 364, 2392–2404 (2011). [PubMed: 21696306]
51. Elshorbagy AK, Kozich V, Smith AD, Refsum H, Cysteine and obesity: Consistency of the evidence across epidemiologic, animal and cellular studies. *Curr. Opin. Clin. Nutr. Metab. Care* 15, 49–57 (2012). [PubMed: 22108094]
52. Durnin JVGA, Womersley J, Body fat assessed from total body density and its estimation from skinfold thickness: Measurements on 481 men and women aged from 16 to 72 years. *Br. J. Nutr.* 32, 77–97 (1974). [PubMed: 4843734]
53. Dumas M-E, Maibaum EC, Teague C, Ueshima H, Zhou BF, Lindon JC, Nicholson JK, Stamler J, Elliott P, Chan Q, Holmes E, Assessment of analytical reproducibility of 1H NMR spectroscopy based metabolomics for large-scale epidemiological research: The INTERMAP study. *Anal. Chem.* 78, 2199–2208 (2006). [PubMed: 16579598]
54. Floegel A, Drogan D, Wang-Sattler R, Prehn C, Illig T, Adamski J, Joost H-G, Boeing H, Pischon T, Reliability of serum metabolite concentrations over a 4-month period using a targeted metabolomic approach. *PLOS One* 6, e21103 (2011).
55. Guertin KA, Moore SC, Sampson JN, Huang WY, Xiao Q, Stolzenberg-Solomon RZ, Sinha R, Cross AJ, Metabolomics in nutritional epidemiology: Identifying metabolites associated with diet and quantifying their potential to uncover diet-disease relations in populations. *Am. J. Clin. Nutr.* 100, 208–217 (2014). [PubMed: 24740205]
56. Xiao Q, Moore SC, Boca SM, Matthews CE, Rothman N, Stolzenberg-Solomon RZ, Sinha R, Cross AJ, Sampson JN, Sources of variability in metabolite measurements from urinary samples. *PLOS One* 9, e95749 (2014).
57. Nicholson JK, Wilson ID, Understanding ‘global’ systems biology: Metabolomics and the continuum of metabolism. *Nat. Rev. Drug Discov.* 2, 668–676 (2003). [PubMed: 12904817]
58. Nicholson JK, Holmes E, Wilson ID, Gut microorganisms, mammalian metabolism and personalized health care. *Nat. Rev. Microbiol.* 3, 431–438 (2005). [PubMed: 15821725]
59. O’Keefe SJD, Li JV, L Lahti J, Ou, Carbonero F, Mohammed K, Poma JM, Kinross J, Wahl E, Ruder D, Vippera K, Naidoo V, Mtshali L, Tims S, Puylaert PGB, DeLany J, Krasinskas A, Benefiel AC, Kaseb HO, Newton K, Nicholson JK, De Vos WM, Gaskins HR, Zoetendal EG, Fat, fiber and cancer risk in African Americans and rural Africans. *Nat. Commun.*, in press.
60. Poma JM, Robinette SL, Holmes E, Nicholson JK, MetaboNetworks, an interactive Matlab-based toolbox for creating, customizing and exploring sub-networks from KEGG. *Bioinformatics* 30, 893–895 (2014). [PubMed: 24177720]

61. Bang J-W, Crockford DJ, Holmes E, Pazos F, Sternberg MJE, Muggleton SH, Nicholson JK, Integrative top-down system metabolic modeling in experimental disease states via data-driven Bayesian methods. *J. Proteome Res.* 7,497–503 (2008). [PubMed: 18179164]
62. Zou H, Hastie T, Regularization and variable selection via the elastic net. *J. R. Stat. Soc. B* 67, 301–320 (2005).
63. Craciun S, Balskus EP, Microbial conversion of choline to trimethylamine requires a glyceryl radical enzyme. *Proc. Natl. Acad. Sci. U.S.A.* 109, 21307–21312 (2012). [PubMed: 23151509]
64. Lang DH, Yeung CK, Peter RM, Ibarra C, Gasser R, Itagaki K, Philpot RM, Rettie AE, Isoform specificity of trimethylamine N-oxygenation by human flavin-containing monooxygenase (FMO) and P450 enzymes: Selective catalysis by FMO3. *Biochem. Pharmacol.* 56, 1005–1012 (1998). [PubMed: 9776311]
65. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, DuGar B, Feldstein AE, Britt EB, Fu X, Chung Y-M, Wu Y, Schauer P, Smith JD, Allayee H, Tang WH, DiDonato JA, Lusis AJ, Hazen SL, Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472, 57–63 (2011). [PubMed: 21475195]
66. Hartiala J, Bennett BJ, Tang WHW, Wang Z, Stewart AFR, Roberts R, McPherson R; CARDIoGRAM Consortium, Comparative genome-wide association studies in mice and humans for trimethylamine N-oxide, a proatherogenic metabolite of choline and L-carnitine. *Arterioscler. Thromb. Vasc Biol.* 34, 1307–1313 (2014). [PubMed: 24675659]
67. Dumas M-E, Barton RH, Toye A, Cloarec O, Blancher C, Rothwell A, Fearnside J, Tatoud R, Blanc V, Lindon JC, Mitchell SC, Holmes E, McCarthy MI, Scott J, Gauguier D, Nicholson JK, Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12511–12516 (2006). [PubMed: 16895997]
68. Jumpertz R, Le DS, Turnbaugh PJ, Trinidad C, Bogardus C, Gordon JI, Krakoff J, Energy-balance studies reveal associations between gut microbes, caloric load, and nutrient absorption in humans. *Am. J. Clin. Nutr.* 94, 58–65 (2011). [PubMed: 21543530]
69. Shay CM, Van Horn L, Stamler J, Dyer AR, Brown IJ, Chan Q, Miura K, Zhao L, Okuda N, Daviglus ML, Elliott P, Food and nutrient intakes and their associations with lower BMI in middle-aged US adults: The international study of macro-/micronutrients and blood pressure (INTERMAP). *Am. J. Clin. Nutr.* 96, 483–491 (2012). [PubMed: 22854407]
70. Selmer T, Andrei PI, p-Hydroxyphenylacetate decarboxylase from *Clostridium difficile*. A novel glyceryl radical enzyme catalysing the formation of p-cresol. *Eur. J. Biochem.* 268, 1363–1372 (2001). [PubMed: 11231288]
71. Ley RE, Turnbaugh PJ, Klein S, Gordon JI, Microbial ecology: Human gut microbes associated with obesity. *Nature* 444, 1022–1023 (2006). [PubMed: 17183309]
72. Duncan SH, Lobley GE, Holtrop G, Ince J, Johnstone AM, Louis P, Flint HJ, Human colonic microbiota associated with diet, obesity and weight loss. *Int. J. Obes.* 32, 1720–1724 (2008).
73. Heinzmann SS, Merrifield CA, Rezzi S, Kochhar S, Lindon JC, Holmes E, Nicholson JK, Stability and robustness of human metabolic phenotypes in response to sequential food challenges. *J. Proteome Res.* 11, 643–655 (2012). [PubMed: 21999107]
74. Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, Pettersson S, Host-gut microbiota metabolic interactions. *Science* 336, 1262–1267 (2012). [PubMed: 22674330]
75. Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI, Obesity alters gut microbial ecology. *Proc. Natl. Acad. Sci. U.S.A.* 102, 11070–11075 (2005). [PubMed: 16033867]
76. Tilg H, Obesity, metabolic syndrome, and microbiota: Multiple interactions. *J. Clin. Gastroenterol.* 44, S16–S18 (2010). [PubMed: 20535027]
77. Hooper LV, Littman DR, Macpherson AJ, Interactions between the microbiota and the immune system. *Science* 336, 1268–1273 (2012). [PubMed: 22674334]
78. Shapiro H, Thaiss CA, Levy M, Elinav E, The cross talk between microbiota and the immune system: Metabolites take center stage. *Curr. Opin. Immunol.* 30, 54–62 (2014). [PubMed: 25064714]
79. Koren O, Goodrich JK, Cullender TC, Spor A, Laitinen K, Backhed HK, Gonzalez A, Werner JJ, Angenent LT, Knight R, Backhed F, Isolauri E, Salminen S, Ley RE, Host remodeling of the gut

microbiome and metabolic changes during pregnancy. *Cell* 150, 470–480 (2012). [PubMed: 22863002]

80. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, Griffin NW, Lombard V, Henrissat B, Bain JR, Muehlbauer MJ, Ilkayeva O, Semenkovich CF, Funai K, Hayashi DK, Lyle BJ, Martini MC, Ursell LK, Clemente JC, Van Treuren W, Walters WA, Knight R, Newgard CB, Heath AC, Gordon JI, Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* 341, 1241214 (2013).
81. Nicholson JK, Higham DP, Timbrell JA, Sadler PJ, Quantitative high resolution ¹H NMR urinalysis studies on the biochemical effects of cadmium in the rat. *Mol. Pharmacol.* 36, 398–404 (1989). [PubMed: 2779524]
82. Janssen I, Heymsfield SB, Wang ZM, Ross R, Skeletal muscle mass and distribution in 468 men and women aged 18–88 yr. *J. Appl. Physiol.* 89, 81–88 (2000). [PubMed: 10904038]
83. Zurlo F, Larson K, Bogardus C, Ravussin E, Skeletal muscle metabolism is a major determinant of resting energy expenditure. *J. Clin. Invest.* 86, 1423–1427 (1990). [PubMed: 2243122]
84. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Lango Allen H, Lindgren CM, Luan J, Magi R, Randall JC, Vedantam S, Winkler TW, Qi L, Workalemahu T, Heid IM, Steinthorsdottir V, Stringham HM, Weedon MN, Wheeler E, Wood AR, Ferreira T, Weyant RJ, Segrè AV, Estrada K, Liang L, Nemesh J, Park J-H, Gustafsson S, Kilpeläinen TO, Yang J, Bouatia-Naji N, Esko T, Feitosa MF, Kutalik Z, Mangino M, Raychaudhuri S, Scherag A, Smith AV, Welch R, Zhao JH, Aben KK, Absher DM, Amin N, Dixon AL, Fisher E, Glazer NL, Goddard ME, Heard-Costa NL, Hoesel V, Hottenga JJ, Johansson Å, Johnson T, Ketkar S, Lamina C, Li S, Moffatt MF, Myers RH, Narisu N, Perry JRB, Peters MJ, Preuss M, Ripatti S, Rivadeneira F, Sandholt C, Scott LJ, Timpson NJ, Tyrer JP, van Wingerden S, Watanabe RM, White CC, Wiklund F, Barlassina C, Chasman DI, Cooper MN, Jansson J-O, Lawrence RW, Pellikka N, Prokopenko I, Shi J, Thiering E, Alavere H, Alibrandi MTS, Almgren P, Arnold AM, Aspelund T, Atwood LD, Baikou B, Balmforth AJ, Bennett AJ, Ben-Shlomo Y, Bergman RN, Bergmann S, Biebermann H, Blakemore AIF, Boes T, Bonnycastle LL, Bornstein SR, Brown MJ, Buchanan TA, Busonero F, Campbell H, Cappuccio FP, Cavalcanti-Proenca C, Chen Y-D, Chen C-M, Chines PS, Clarke R, Coin L, Connell J, Day INM, den Heijer M, Duan J, Ebrahim S, Elliott P, Elosua R, Eiriksdottir G, Erdos MR, Eriksson JG, Facheris MF, Felix SB, Fischer-Posovszky P, Folsom AR, Friedrich N, Freimer NB, Fu M, Gaget S, Gejman PV, Geus EJC, Gieger C, Gjesing AP, Goel A, Goyette P, Grallert H, Gräßler J, Greenawalt DM, Groves CJ, Gudnason V, Guiducci C, Hartikainen AL, Hassanal i N, Hall AS, Havulinna AS, Hayward C, Heath AC, Hengstenberg C, Hicks AA, Hinney A, Hofman A, Homuth G, Hui J, Igl W, Iribarren C, Isomaa B, Jacobs KB, Jarick I, Jewell E, John U, Jørgensen T, Jousilahti P, Jula A, Kaakinen M, Kajantie E, Kaplan LM, Kathiresan S, Kettunen J, Kinnunen L, Knowles JW, Kolcic I, König IR, Koskinen S, Kovacs P, Kuusisto J, Kraft P, Kvaløy K, Laitinen J, Lantieri O, Lanzani C, Launer LJ, Lecoeur C, Lehtimäki T, Lettre G, Liu J, Lokki M-L, Lorentzon M, Luben RN, Ludwig B, Manunta P, Marek D, Marre M, Martin NG, McArdle WL, McCarthy A, McKnight B, Meitinger T, Melander O, Meyre D, Midthjell K, Montgomery GW, Morken MA, Morris AP, Mulic R, Ngwa JS, Nelis M, Neville MJ, Nyholt DR, O'Donnell CJ, O'Rahilly S, Ong KK, Oostra B, Paré G, Parker AN, Perola M, Pichler I, Pietiläinen KH, Platou CGP, Polasek O, Pouta A, Rafelt S, Raitakari O, Rayner NW, Ridderstråle M, Rief W, Ruukonen A, Robertson NR, Rzehak P, Salomaa V, Sanders AR, Sandhu MS, Sanna S, Saramies J, Savolainen MJ, Scherag S, Schipf S, Schreiber S, Schunkert H, Silander K, Sinisalo J, Siscovick DS, Smit JH, Soranzo N, Sovio U, Stephens J, Surakka I, Swift AJ, Tammesoo M-L, Tardif J-C, Teder-Laving M, Teslovich TM, Thompson JR, Thomson B, Tönjes A, Tuomi T, van Meurs JBJ, van Ommen G-J, Vatin V, Viikari J, Visvikis-Siest S, Vitart V, Vogel CIG, Voight BF, Waite LL, Wallaschofski H, Walters GB, Widen E, Wiegand S, Wild SH, Willemsen G, Witte DR, Witterman JC, Xu J, Zhang Q, Zgaga L, Ziegler A, Zitting P, Beilby JP, Farooqi IS, Hebebrand J, Huikuri HV, James AL, Kähönen M, Levinson DF, Macciardi F, Nieminen MS, Ohlsson C, Palmer LJ, Ridker PM, Stumvoll M, Beckmann JS, Boeing H, Boerwinkle E, Boomsma DI, Caulfield MJ, Chanock SJ, Collins FS, Cupples LA, Smith GD, Erdmann J, Froguel P, Grönberg H, Gyllenstein U, Hall P, Hansen T, Harris TB, Hattersley AT, Hayes RB, Heinrich J, Hu FB, Hveem K, Illig T, Jarvelin M-R, Kaprio J, Karpe F, Khaw K-T, Kiemeny LA, Krude H, Laakso M, Lawlor DA, Metspalu A, Munroe PB, Ouwehand WH, Pedersen O, Penning BW, Peters A, Pramstaller PP, Quertermous T, Reinehr T, Rissanen A, Rudan I, Samani NJ, Schwarz PEH, Shuldiner AR,

- Spector TD, Tuomilehto J, Uda M, Uitterlinden A, Valle TT, Wabitsch M, Waeber G, Wareham NJ, Watkins H, Wilson JF, Wright AF, Zillikens MC, Chatterjee N, McCarroll SA, Purcell S, Schadt EE, Visscher PM, Assimes TL, Borecki IB, Deloukas P, Fox CS, Groop LC, Haritunians T, Hunter DJ, Kaplan RC, Mohlke KL, O'Connell JR, Peltonen L, Schlessinger D, Strachan DP, van Duijn CM, Wichmann H-E, Frayling TM, Thorsteinsdottir U, Abecasis GR, Barroso I, Boehnke M, Stefansson K, North KE, McCarthy MI, Hirschhorn JN, Ingelsson E, Loos RJ, MAGIC, Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* 42,937–948 (2010). [PubMed: 20935630]
85. Braverman NE, Moser AB, Functions of plasmalogen lipids in health and disease. *Biochim. Biophys. Acta* 1822, 1442–1452 (2012). [PubMed: 22627108]
 86. Barreira TV, Harrington DM, Staiano AE, Heymsfield SB, Katzmarzyk PT, Body adiposity index, body mass index, and body fat in white and black adults. *JAMA* 306, 828–830 (2011). [PubMed: 21862743]
 87. Lichtash CT, Cui J, Guo X, Chen Y-DI, Hsueh WA, Rotter JI, Goodarzi MO, Body adiposity index versus body mass index and other anthropometric traits as correlates of cardiometabolic risk factors. *PLOS One* 8, e65954 (2013).
 88. Pischon T, Boeing H, Hoffmann K, Bergmann M, Schulze MB, Overvad K, van der Schouw YT, Spencer A, Moons KG, Tjønneland A, Halkjaer J, Jensen MK, Stegger J, Clavel-Chapelon F, Boutron-Ruault M-C, Chajes V, Linseisen J, Kaaks R, Trichopoulou A, Trichopoulos D, Bamia C, Sieri S, Palli D, Tumino R, Vineis P, Panico S, Peeters PH, May AM, Bueno-de-Mesquita HB, van Duijnhoven FJ, Hallmans G, Weinehall L, Manjer J, Hedblad B, Lund E, Agudo A, Arriola L, Barricarte A, Navarro C, Martinez C, Quirós JR, Key T, Phil D, Bingham S, Khaw KT, Chir B, Boffetta P, Jenab M, Ferrari P, Riboli E, General and abdominal adiposity and risk of death in Europe. *N. Engl. J. Med.* 359, 2105–2120 (2008). [PubMed: 19005195]
 89. Bouatra S, Aziat F, Mandal R, Guo AC, Wilson MR, Knox C, Bjorndahl TC, Krishnamurthy R, Saleem E, Liu P, Dame ZT, Poelzer J, Huynh J, Yallou FS, Psychogios N, Dong E, Bogumil R, Roehring B, Wishart DS, The human urine metabolome. *PLOS One* 8, e73076 (2013).
 90. Flier JS, Mekalanos JJ, Gut check: Testing a role for the intestinal microbiome in human obesity. *Sci. Transl. Med.* 1, 6ps7 (2009).
 91. Danaei G, Finucane MM, Lin JK, Singh GM, Paciorek CJ, Cowan MJ, Farzadfar F, Stevens GA, Lim SS, Riley LM, Ezzati M, National, regional, and global trends in systolic blood pressure since 1980: Systematic analysis of health examination surveys and epidemiological studies with 786 country-years and 5.4 million participants. *Lancet* 377, 568–577 (2011). [PubMed: 21295844]
 92. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative, The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Ann. Intern. Med.* 147, 573–577 (2007). [PubMed: 17938396]
 93. Holmes E, Loo RL, Cloarec O, Coen M, Tang HR, Maibaum E, Bruce S, Chan Q, Elliott P, Stamler J, Wilson ID, Lindon JC, Nicholson JK, Detection of urinary drug metabolite (Xenometabolome) signatures in molecular epidemiology studies via statistical total correlation (NMR) spectroscopy. *Anal. Chem.* 79, 2629–2640 (2007). [PubMed: 17323917]
 94. Dieterle F, Ross A, Schlotterbeck G, Senn H, Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabolomics. *Anal. Chem.* 78, 4281–4290 (2006). [PubMed: 16808434]
 95. Gamer M, Lemon J, Fellows I, Singh P, Coefficients of interrater reliability and agreement for quantitative, ordinal and nominal data; <http://CRAN.R-project.org/package=irr>.
 96. Dyer AR, Elliott P, Shipley M, Urinary electrolyte excretion in 24 hours and blood pressure in the INTERSALT study: II. Estimates of electrolyte-blood pressure associations corrected for regression dilution bias. The INTERSALT Cooperative Research Group. *Am. J. Epidemiol.* 139, 940–951 (1994). [PubMed: 8166144]
 97. Efron B, Bootstrap Methods: Another look at the jackknife. *Ann. Statist.* 7, 1–26 (1979).
 98. Tibshirani R, Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B.* 58, 267–288 (1996).
 99. Hoerl AE, Kennard RW, Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* 12, 55–67 (1970).

100. Friedman J, Hastie T, Tibshirani R, Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22 (2010). [PubMed: 20808728]
101. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA, Diversity of the human intestinal microbial flora. *Science* 308, 1635–1638 (2005). [PubMed: 15831718]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

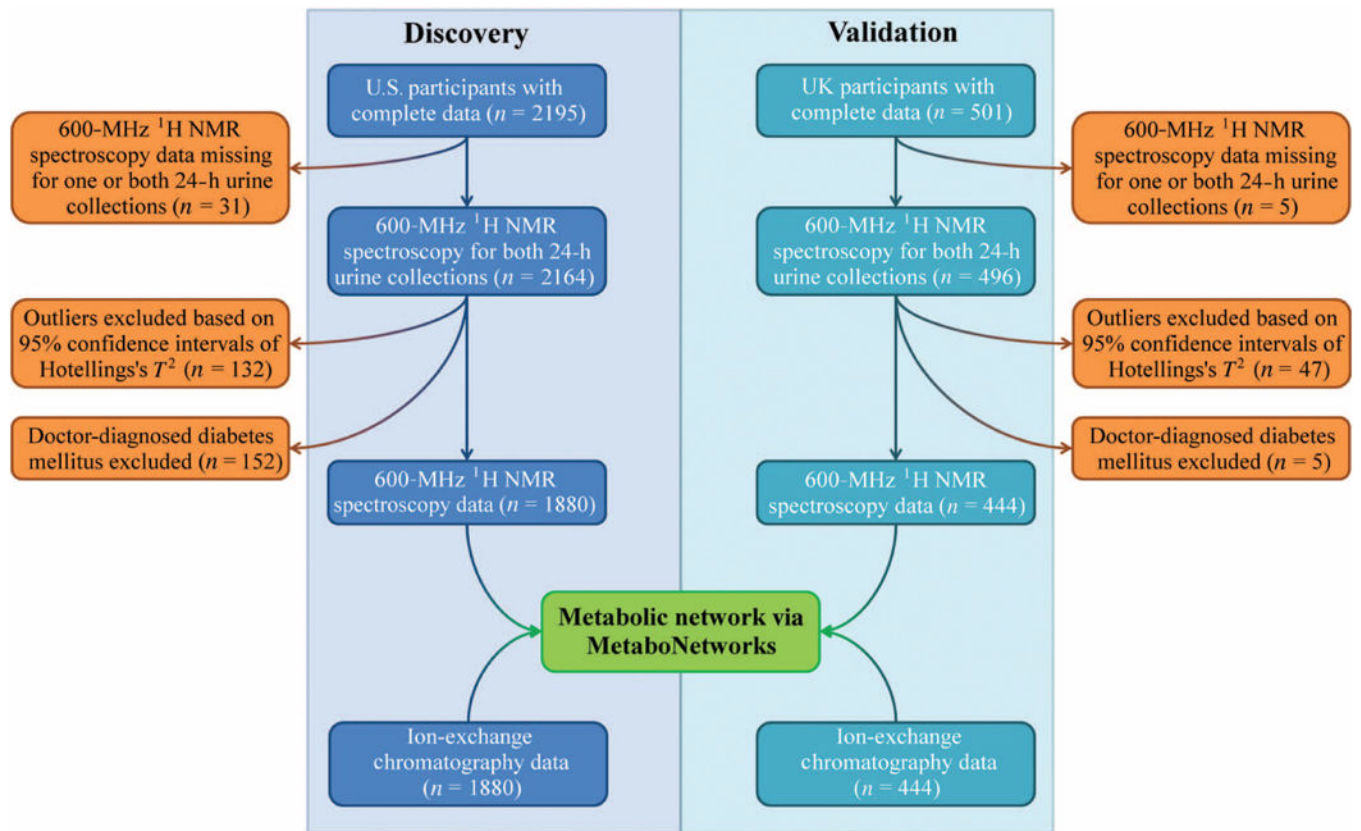


Fig. 1. Schema of study design.

Individuals were recruited into the study from general and occupational population samples in the United States and UK. Each individual made four clinic visits, the first two on consecutive days and the second two on consecutive days 3 weeks later. At each visit, blood pressure was measured twice, and a complete 24-hour dietary recall was obtained by trained interviewers. Data on height, weight, and extensive questionnaire information were also obtained. A timed 24-hour urine collection was commenced at the first (and third) visit and completed at the second (and fourth) visit the following day. Each 24-hour urine collection was mixed together, and urine aliquots were obtained from the pooled urine sample. Urinary aliquots from individuals with complete data were measured with ^1H NMR spectroscopy.

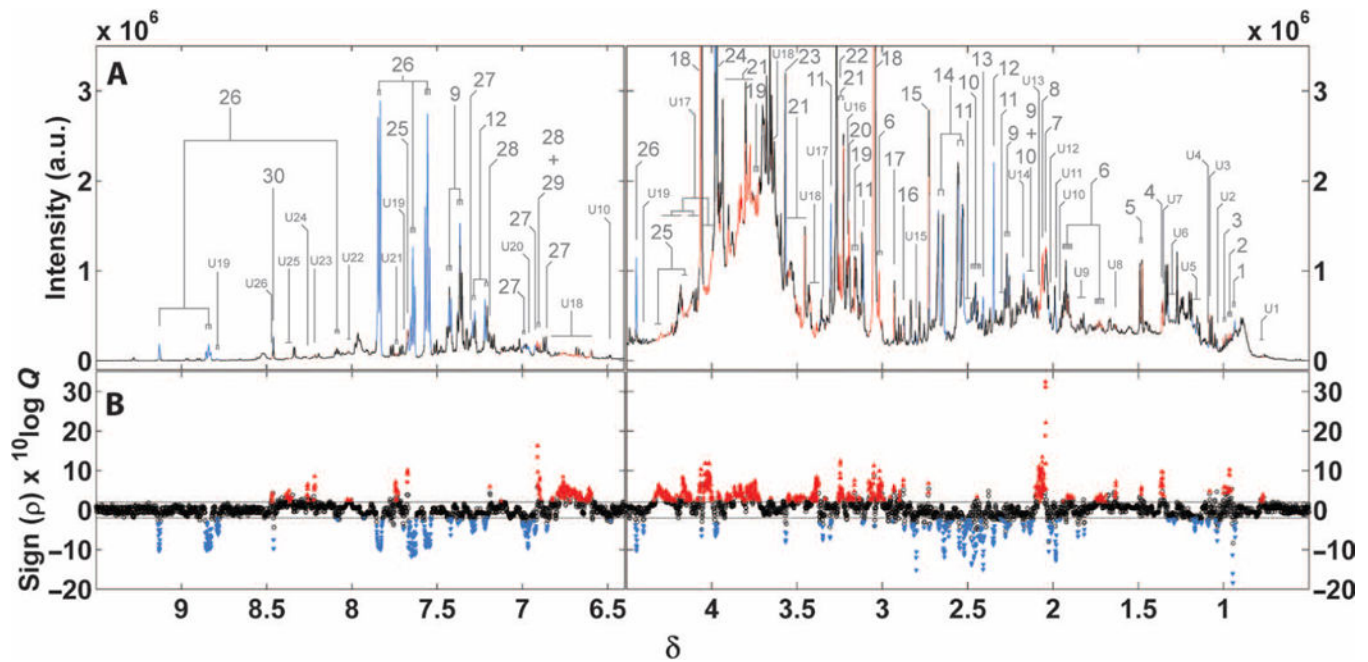


Fig. 2. Associations of BMI with urinary metabolites in the U.S. INTERMAP cohort ($n = 1880$). Urine was collected for 24 hours from individuals on two separate occasions 3 weeks apart. Each individual completed four 24-hour dietary recalls, and eight blood pressure measurements and questionnaire information were obtained. The partial correlation between each ^1H NMR variable and BMI was adjusted for age, gender, and population sample (model 1). (A) Average 600 MHz ^1H NMR spectrum of the first urine collection. (B) Manhattan plot showing $-\log_{10}(Q) \times \text{sign}$ of partial correlation for each of the 7100 spectral variables. Significance of ^1H NMR was determined based on a Q value threshold of 1%; in addition to this, both adjacent variables must also pass $Q < 0.01$ and must have the same sign. Statistically significant peaks are colored red if directly associated with BMI and blue if inversely associated. Metabolites significantly associated with BMI (numbered according to spectral position) were as follows: 1: ketoleucine, 2: leucine, 3: valine, 4: 2-hydroxyisobutyrate, 5: alanine, 6: lysine, 7: N -acetyl signals from urinary glycoproteins, 8: N -acetyl neuraminic acid, 9: phenylacetylglutamine, 10: glutamine, 11: proline betaine, 12: 4-cresyl sulfate, 13: succinate, 14: citrate, 15: dimethylamine, 16: TMA, 17: dimethylglycine, 18: creatinine, 19: ethanolamine, 20: O -acetyl carnitine, 21: glucose, 22: 3-methylhistidine, 23: glycine, 24: hippurate, 25: pseudouridine, 26: NMNA, 27: 3-hydroxymandelate, 28: tyrosine, 29: 4-hydroxymandelate, 30: formate, U1 to U26 unidentified metabolites. These data are tabulated in Table 1 and table S4. Some signals overlap; however, they were unequivocally identified using various statistical and experimental methods (table S2).

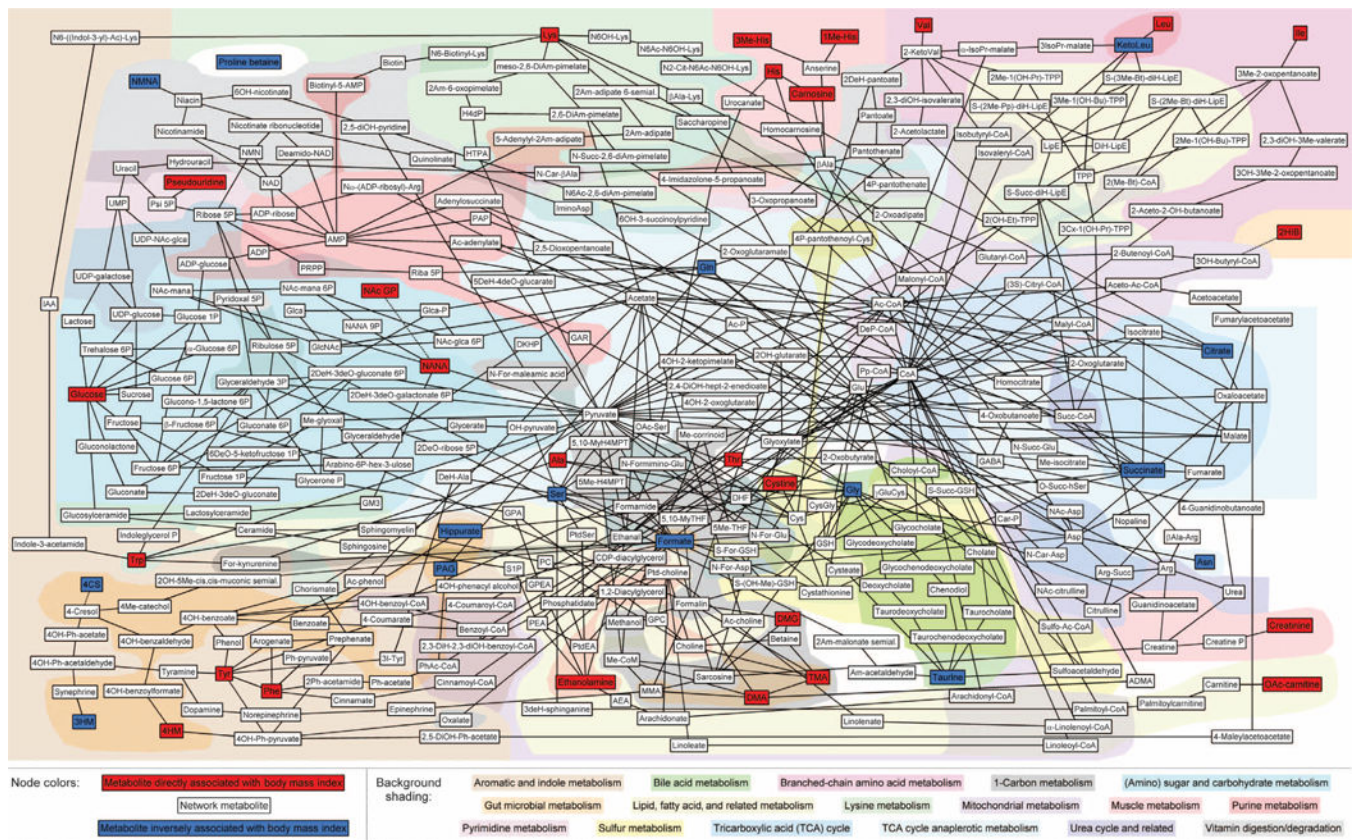


Fig. 3. Multicompartmental metabolic reaction network illustrating human adiposity-related urinary metabolic signatures in the U.S. population ($n = 1880$). Metabolites are included that passed the ST-FDR threshold of $Q = 0.01$ for both models 1 and 2 (^1H NMR metabolites, see Table 1) and the Bonferroni threshold of $P = 4.55 \times 10^{-4}$ for both models 1 and 2 (metabolites detected by IEC, see Table 2). The background shading illustrates different types of metabolism based on closest affinity classification. Table S10 lists the full names and abbreviations for the metabolites. Table S11 lists the closest affinity classifications for metabolites shown in this figure. Dotted lines indicate the closest related metabolite in the network for metabolites that are not listed in the KEGG database, based on available literature.

Structurally identified ¹H NMR-derived urinary metabolites associated significantly with BMI in 1880 U.S.

Table 1.

INTERMAP participants using first urine collection specimens. Excluding metabolic outliers based on Hotelling's T^2 test ($n = 132$) and participants with doctor-diagnosed diabetes mellitus ($n = 152$). Partial correlation (r) and corresponding P values are listed for each metabolite. Statistical significance is based on the ST-FDR, $Q < 0.01$ for each spectral variable and the two adjacent spectral variables.

Urinary variable	Model 1*			Model 2*†			Model 3*‡§		
	r	P	Q	r	P	Q	r	P	Q
Urinary glycoproteins	0.29	2.04×10^{-36}	4.15×10^{-33}	0.27	9.63×10^{-33}	3.22×10^{-29}	0.27	3.09×10^{-33}	1.23×10^{-29}
Ketoleucine	-0.22	7.19×10^{-22}	3.56×10^{-19}	-0.20	1.45×10^{-18}	8.07×10^{-16}	-0.23	6.84×10^{-23}	6.81×10^{-20}
Tyrosine + 4-hydroxymandelate [§]	0.21	1.35×10^{-19}	5.72×10^{-17}	0.20	3.64×10^{-18}	1.74×10^{-15}	0.17	3.60×10^{-14}	8.97×10^{-12}
Succinate	-0.20	3.22×10^{-18}	9.56×10^{-16}	-0.15	1.19×10^{-10}	8.44×10^{-9}	-0.14	2.42×10^{-9}	1.80×10^{-7}
Glutamine	-0.19	6.05×10^{-17}	1.50×10^{-14}	-0.15	1.38×10^{-10}	9.07×10^{-9}	-0.13	1.70×10^{-8}	8.78×10^{-7}
N-acetylneuraminic acid	0.19	2.02×10^{-16}	4.07×10^{-14}	0.19	5.25×10^{-17}	2.20×10^{-14}	0.22	3.75×10^{-22}	2.99×10^{-19}
3-Methylhistidine	0.18	2.87×10^{-15}	5.01×10^{-13}	0.17	2.79×10^{-13}	3.90×10^{-11}	0.12	5.68×10^{-7}	1.18×10^{-5}
Citrate	-0.18	8.61×10^{-15}	1.16×10^{-12}	-0.15	1.54×10^{-10}	9.86×10^{-9}	-0.14	3.45×10^{-9}	2.41×10^{-7}
Hippurate	-0.18	1.52×10^{-14}	1.77×10^{-12}	-0.14	3.39×10^{-10}	1.86×10^{-8}	-0.12	1.24×10^{-7}	4.02×10^{-6}
Creatinine	0.17	8.88×10^{-14}	6.28×10^{-12}	0.17	1.78×10^{-13}	2.84×10^{-11}			
NMNA	-0.16	9.04×10^{-13}	4.55×10^{-11}	-0.14	1.15×10^{-9}	4.31×10^{-8}	-0.12	8.60×10^{-8}	3.14×10^{-6}
Leucine	0.16	1.16×10^{-12}	5.45×10^{-11}	0.16	1.03×10^{-11}	1.01×10^{-9}	0.10	6.63×10^{-6}	7.03×10^{-5}
Pseudouridine	0.16	2.10×10^{-12}	8.53×10^{-11}	0.14	4.09×10^{-10}	2.11×10^{-8}	0.14	3.21×10^{-9}	2.29×10^{-7}
Lysine	0.16	3.67×10^{-12}	1.38×10^{-10}	0.14	5.23×10^{-10}	2.48×10^{-8}	0.13	4.39×10^{-8}	1.90×10^{-6}
Formate	-0.16	4.50×10^{-12}	1.65×10^{-10}	-0.13	3.17×10^{-8}	5.74×10^{-7}	-0.09	4.84×10^{-5}	3.28×10^{-4}
2-Hydroxyisobutyrate	0.16	7.24×10^{-12}	2.29×10^{-10}	0.11	4.89×10^{-6}	3.25×10^{-5}	0.09	7.60×10^{-5}	4.59×10^{-4}
4-Cresyl sulfate	-0.15	8.09×10^{-11}	1.57×10^{-9}	-0.15	1.32×10^{-10}	9.01×10^{-9}	-0.15	3.22×10^{-10}	3.13×10^{-8}
Glycine	-0.14	6.29×10^{-10}	8.82×10^{-9}	-0.12	2.91×10^{-7}	3.22×10^{-6}	-0.11	2.12×10^{-6}	3.09×10^{-5}
TMA	0.14	2.04×10^{-9}	2.49×10^{-8}	0.13	6.41×10^{-9}	1.66×10^{-7}	0.11	1.52×10^{-6}	2.46×10^{-5}
Ethanolamine	0.14	2.72×10^{-9}	3.22×10^{-8}	0.13	3.95×10^{-8}	6.65×10^{-7}	0.13	3.17×10^{-8}	1.45×10^{-6}
Glucose	0.14	3.27×10^{-9}	3.84×10^{-8}	0.13	3.21×10^{-8}	5.76×10^{-7}	0.12	6.85×10^{-8}	2.66×10^{-6}
Proline betaine	-0.13	7.29×10^{-9}	7.96×10^{-8}	-0.10	1.04×10^{-5}	6.10×10^{-5}	-0.09	1.88×10^{-4}	9.46×10^{-4}

Urinary variable	Model 1*			Model 2*†			Model 3*†‡		
	r	P	Q	r	P	Q	r	P	Q
Phenylacetylglutamine	-0.13	1.39×10^{-8}	1.37×10^{-7}	-0.12	1.79×10^{-7}	2.21×10^{-6}	-0.11	1.15×10^{-6}	1.97×10^{-5}
Dimethylamine	0.13	1.74×10^{-8}	1.69×10^{-7}	0.12	5.05×10^{-7}	5.09×10^{-6}	0.09	6.43×10^{-5}	4.07×10^{-4}
3-Hydroxymandelate	-0.13	2.81×10^{-8}	2.57×10^{-7}	-0.12	2.72×10^{-7}	3.05×10^{-6}	-0.10	1.98×10^{-5}	1.68×10^{-4}
Dimethylglycine	0.12	1.43×10^{-7}	1.04×10^{-6}	0.11	1.18×10^{-6}	1.01×10^{-5}	0.07	2.98×10^{-3}	8.49×10^{-3}
Valine	0.12	2.39×10^{-7}	1.62×10^{-6}	0.10	8.15×10^{-6}	4.97×10^{-5}	0.08	1.20×10^{-3}	4.02×10^{-3}
Alanine	0.12	3.32×10^{-7}	2.18×10^{-6}	0.11	2.29×10^{-6}	1.75×10^{-5}	0.09	4.72×10^{-5}	3.21×10^{-4}
O-acetyl carnitine	0.10	1.46×10^{-5}	5.37×10^{-5}	0.10	1.38×10^{-5}	7.62×10^{-5}	0.03	2.41×10^{-1}	2.48×10^{-1}
Ketoleucine/leucine ratio	-0.24	8.85×10^{-26}	5.84×10^{-23}	-0.23	1.02×10^{-23}	7.61×10^{-21}	-0.21	1.28×10^{-20}	6.80×10^{-18}

* Model 1 is adjusted for age, gender, and sample.

† Model 2 is adjusted for all factors in model 1 plus cardiovascular disease history (heart disease/stroke), physical activity, medication for hypertension, prescribed lipid-lowering drugs, NSAID use, dietary supplement use, special diet, smoking, education, and total energy intake per day (kcal/day).

‡ Model 3 is adjusted for all factors in models 1 and 2 plus 24-hour urinary creatinine.

§ The signal at δ 6.91 of tyrosine also includes a signal from its metabolite 4-hydroxymandelate.

Table 2.
Association with BMI of a set of metabolites measured by targeted IEC (log₁₀ values) in 1880 U.S. INTERMAP participants.

Excluding metabolic outliers based on Hotelling's T^2 test ($n = 132$) and participants with doctor-diagnosed diabetes mellitus ($n = 152$). Partial correlation (r) and corresponding P values are listed for each metabolite. Statistical significance based on a Bonferroni threshold of $P = 4.55 \times 10^{-4}$ ($P = 0.0/2$), n.s., not significant.

Urinary variable	Model 1*			Model 2*†			Model 3*†‡		
	r	P	Significance	r	P	Significance	r	P	Significance
Taurine	0.02	3.68×10^{-1}	n.s.	0.01	6.09×10^{-1}	n.s.	-0.11	1.13×10^{-6}	n.s.
Theonine	0.15	4.44×10^{-11}	n.s.	0.15	2.34×10^{-10}	n.s.	-0.02	3.82×10^{-1}	n.s.
Serine	0.05	3.06×10^{-2}	n.s.	0.06	1.58×10^{-2}	n.s.	-0.12	4.41×10^{-7}	n.s.
Asparagine	0.01	7.69×10^{-1}	n.s.	0.00	9.97×10^{-1}	n.s.	-0.11	7.93×10^{-7}	n.s.
Glutamine	0.10	1.29×10^{-5}	n.s.	0.12	3.34×10^{-7}	n.s.	-0.05	4.41×10^{-2}	n.s.
Glycine	-0.02	3.98×10^{-1}	n.s.	0.01	6.00×10^{-1}	n.s.	-0.10	6.08×10^{-6}	n.s.
Alanine	0.20	4.85×10^{-18}	n.s.	0.18	3.47×10^{-15}	n.s.	0.02	3.65×10^{-1}	n.s.
Valine	0.19	6.87×10^{-16}	n.s.	0.17	6.51×10^{-14}	n.s.	0.07	3.87×10^{-3}	n.s.
Cystine	0.33	4.85×10^{-49}	n.s.	0.30	4.59×10^{-41}	n.s.	0.18	6.80×10^{-15}	n.s.
Methionine	0.06	1.06×10^{-2}	n.s.	0.05	4.32×10^{-2}	n.s.	-0.03	1.57×10^{-1}	n.s.
Isoleucine	0.12	4.23×10^{-7}	n.s.	0.10	2.11×10^{-5}	n.s.	0.05	2.73×10^{-2}	n.s.
Leucine	0.08	3.66×10^{-4}	n.s.	0.06	7.53×10^{-3}	n.s.	0.02	5.08×10^{-1}	n.s.
Tyrosine	0.32	2.02×10^{-45}	n.s.	0.30	8.28×10^{-41}	n.s.	0.18	1.58×10^{-14}	n.s.
Phenylalanine	0.20	8.35×10^{-19}	n.s.	0.19	1.42×10^{-16}	n.s.	0.08	9.99×10^{-4}	n.s.
Ethanolamine	0.26	2.00×10^{-29}	n.s.	0.27	2.19×10^{-31}	n.s.	0.14	7.39×10^{-10}	n.s.
Lysine	0.27	4.64×10^{-33}	n.s.	0.26	7.81×10^{-30}	n.s.	0.12	1.06×10^{-7}	n.s.
1-Methylhistidine	0.11	9.47×10^{-7}	n.s.	0.09	7.20×10^{-5}	n.s.	-0.05	4.48×10^{-2}	n.s.
Histidine	0.14	3.02×10^{-9}	n.s.	0.15	6.23×10^{-11}	n.s.	0.01	6.43×10^{-1}	n.s.
Tryptophan	0.15	4.49×10^{-11}	n.s.	0.15	1.79×10^{-10}	n.s.	0.09	4.35×10^{-5}	n.s.
3-Methylhistidine	0.42	8.28×10^{-80}	n.s.	0.38	1.87×10^{-65}	n.s.	0.12	2.14×10^{-7}	n.s.
Camosine	0.18	7.16×10^{-15}	n.s.	0.17	2.28×10^{-13}	n.s.	0.02	2.90×10^{-1}	n.s.
Arginine	0.08	8.71×10^{-4}	n.s.	0.07	2.77×10^{-3}	n.s.	0.03	1.34×10^{-1}	n.s.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

* Model 1 is adjusted for age, gender, and sample.

† Model 2 is adjusted for all factors in model 1 plus cardiovascular disease history (heart disease/stroke), physical activity, medication for hypertension, prescribed lipid-lowering drugs, NSAID use, dietary supplement use, special diet, smoking, education, and total energy intake per day (kcal/day).

‡ Model 3 is adjusted for all factors in models 1 and 2 plus 24-hour urinary creatinine.

Table 3.

Association with BMI of (A) metabolites identified by ¹H NMR and (B) metabolites measured by targeted IEC (log₁₀ values) in 444 UK INTERMAP participants for first urine collection specimens.

Excluding metabolic outliers based on Hotelling's *T*² test (*n* = 47) and participants with doctor-diagnosed diabetes mellitus (*n* = 5). Partial correlation (*r*) and corresponding *P* values are listed for each metabolite. Statistical significance is based on the ST-FDR, *Q* = 0.05 for ¹H NMR metabolites identified and analyzed in the U.S. data. Statistical significance for IEC metabolites is based on a Bonferroni threshold of *P* = 2.50 × 10⁻³ (*P* = 0.05/20).

Urinary variable	Model 1*			Model 2**†			Model 3**‡		
	<i>r</i>	<i>P</i>	<i>Q</i>	<i>r</i>	<i>P</i>	<i>Q</i>	<i>r</i>	<i>P</i>	<i>Q</i>
A. ¹ H NMR spectroscopy									
Urinary glycoproteins	0.33	1.02 × 10 ⁻¹²	4.66 × 10 ⁻⁹	0.33	1.22 × 10 ⁻¹²	6.17 × 10 ⁻⁹	0.36	1.50 × 10 ⁻¹⁴	7.25 × 10 ⁻¹¹
Ketoleucine	-0.18	1.49 × 10 ⁻⁴	2.50 × 10 ⁻³	-0.14	2.66 × 10 ⁻³	2.25 × 10 ⁻²	-0.16	1.00 × 10 ⁻³	1.08 × 10 ⁻²
Tyrosine + 4-hydroxymandelate [§]	0.25	1.17 × 10 ⁻⁷	5.19 × 10 ⁻⁵	0.25	8.71 × 10 ⁻⁸	5.53 × 10 ⁻⁵	0.23	9.69 × 10 ⁻⁷	5.20 × 10 ⁻⁴
Succinate	-0.14	2.52 × 10 ⁻³	1.76 × 10 ⁻²	-0.15	2.19 × 10 ⁻³	1.98 × 10 ⁻²	-0.13	5.62 × 10 ⁻³	3.46 × 10 ⁻²
Glutamine	-0.14	2.62 × 10 ⁻³	1.80 × 10 ⁻²	-0.13	5.25 × 10 ⁻³	3.52 × 10 ⁻²	-0.12	1.49 × 10 ⁻²	6.70 × 10 ⁻²
N-acetylneuraminic acid	0.18	1.74 × 10 ⁻⁴	2.81 × 10 ⁻³	0.20	4.56 × 10 ⁻⁵	1.58 × 10 ⁻³	0.21	1.67 × 10 ⁻⁵	1.11 × 10 ⁻³
3-Methylhistidine	0.13	7.71 × 10 ⁻³	3.62 × 10 ⁻²	0.13	5.23 × 10 ⁻³	3.52 × 10 ⁻²	0.08	8.41 × 10 ⁻²	1.96 × 10 ⁻¹
Citrate	0.10	4.45 × 10 ⁻²	1.22 × 10 ⁻¹	-0.11	2.40 × 10 ⁻²	9.93 × 10 ⁻²	0.11	2.26 × 10 ⁻²	8.62 × 10 ⁻²
Hippurate	-0.24	2.75 × 10 ⁻⁷	5.19 × 10 ⁻⁵	-0.23	1.78 × 10 ⁻⁶	4.43 × 10 ⁻⁴	-0.23	1.80 × 10 ⁻⁶	6.70 × 10 ⁻⁴
Creatinine	0.17	3.96 × 10 ⁻⁴	5.11 × 10 ⁻³	0.18	1.68 × 10 ⁻⁴	3.73 × 10 ⁻³			
NMNA	-0.08	7.76 × 10 ⁻²	1.75 × 10 ⁻¹	-0.05	2.84 × 10 ⁻¹	4.29 × 10 ⁻¹	-0.04	3.69 × 10 ⁻¹	4.50 × 10 ⁻¹
Leucine	0.17	3.73 × 10 ⁻⁴	4.88 × 10 ⁻³	0.16	6.13 × 10 ⁻⁴	8.57 × 10 ⁻³	0.11	2.23 × 10 ⁻²	8.58 × 10 ⁻²
Pseudouridine	0.24	4.87 × 10 ⁻⁷	6.38 × 10 ⁻⁵	0.25	1.86 × 10 ⁻⁷	1.05 × 10 ⁻⁴	0.24	8.09 × 10 ⁻⁷	4.89 × 10 ⁻⁴
Lysine	0.32	4.30 × 10 ⁻¹²	9.84 × 10 ⁻⁹	0.29	7.11 × 10 ⁻¹⁰	9.03 × 10 ⁻⁷	0.24	5.78 × 10 ⁻⁷	4.29 × 10 ⁻⁴
Formate	-0.15	1.72 × 10 ⁻³	1.35 × 10 ⁻²	-0.14	4.66 × 10 ⁻³	3.29 × 10 ⁻²	-0.13	7.33 × 10 ⁻³	4.11 × 10 ⁻²
2-Hydroxyisobutyrate	0.19	4.82 × 10 ⁻⁵	1.11 × 10 ⁻³	0.16	1.20 × 10 ⁻³	1.32 × 10 ⁻²	0.12	9.74 × 10 ⁻³	5.01 × 10 ⁻²
4-Cresyl sulfate	-0.17	2.22 × 10 ⁻⁴	3.36 × 10 ⁻³	-0.18	2.09 × 10 ⁻⁴	4.30 × 10 ⁻³	-0.16	1.18 × 10 ⁻³	1.19 × 10 ⁻²
Glycine	-0.12	1.04 × 10 ⁻²	4.37 × 10 ⁻²	-0.14	4.65 × 10 ⁻³	3.29 × 10 ⁻²	-0.12	1.56 × 10 ⁻²	6.87 × 10 ⁻²
TMA	0.21	1.00 × 10 ⁻⁵	3.59 × 10 ⁻⁴	0.21	8.16 × 10 ⁻⁶	6.50 × 10 ⁻⁴	0.22	2.56 × 10 ⁻⁶	7.29 × 10 ⁻⁴
Ethanolamine	0.13	5.23 × 10 ⁻³	2.79 × 10 ⁻²	0.13	5.25 × 10 ⁻³	3.52 × 10 ⁻²	0.16	7.04 × 10 ⁻⁴	8.38 × 10 ⁻³

Urinary variable	Model 1*			Model 2**†			Model 3**††		
	r	P	Q	r	P	Q	r	P	Q
Glucose	0.19	4.97×10^{-5}	1.13×10^{-3}	0.17	3.71×10^{-4}	6.02×10^{-3}	0.21	1.47×10^{-5}	1.06×10^{-3}
Proline betaine	-0.12	1.31×10^{-2}	5.14×10^{-2}	-0.11	1.73×10^{-2}	7.81×10^{-2}	-0.11	1.85×10^{-2}	7.65×10^{-2}
Phenylacetylglutamine	-0.15	1.30×10^{-3}	1.13×10^{-2}	-0.17	5.02×10^{-4}	7.46×10^{-3}	-0.15	2.14×10^{-3}	1.76×10^{-2}
Dimethylamine	0.14	3.72×10^{-3}	2.27×10^{-2}	0.16	6.03×10^{-4}	8.48×10^{-3}	0.14	3.53×10^{-3}	2.51×10^{-2}
3-Hydroxymandelate	-0.09	4.95×10^{-2}	1.31×10^{-1}	-0.10	4.21×10^{-2}	1.41×10^{-1}	-0.07	1.31×10^{-1}	2.57×10^{-1}
Dimethylglycine	0.16	6.11×10^{-4}	7.00×10^{-3}	0.17	4.54×10^{-4}	6.94×10^{-3}	0.13	8.85×10^{-3}	4.66×10^{-2}
Valine	0.14	4.16×10^{-3}	2.47×10^{-2}	0.12	1.59×10^{-2}	7.35×10^{-2}	0.10	3.06×10^{-2}	1.03×10^{-1}
Alanine	0.21	1.10×10^{-5}	3.73×10^{-4}	0.22	5.07×10^{-6}	5.31×10^{-4}	0.19	5.81×10^{-5}	1.81×10^{-3}
O-acetyl carnitine	0.15	1.31×10^{-3}	1.13×10^{-2}	0.15	1.67×10^{-3}	1.66×10^{-2}	0.10	4.68×10^{-2}	1.35×10^{-1}
Ketoleucine/leucine ratio	-0.24	1.90×10^{-7}	5.19×10^{-5}	-0.24	7.11×10^{-7}	2.89×10^{-4}	-0.20	3.18×10^{-5}	1.41×10^{-3}
B. IEC									
Taurine	-0.03	5.88×10^{-1}	n.s.	-0.04	4.61×10^{-1}	n.s.	-0.12	1.38×10^{-2}	n.s.
Threonine	0.06	1.92×10^{-1}	n.s.	0.06	1.91×10^{-1}	n.s.	-0.06	2.28×10^{-1}	n.s.
Serine	0.01	9.16×10^{-1}	n.s.	0.01	7.60×10^{-1}	n.s.	-0.10	4.06×10^{-2}	n.s.
Asparagine	0.00	9.73×10^{-1}	n.s.	0.00	9.88×10^{-1}	n.s.	-0.09	7.05×10^{-2}	n.s.
Glutamine	0.09	5.75×10^{-2}	n.s.	0.10	4.08×10^{-2}	n.s.	-0.01	8.25×10^{-1}	n.s.
Glycine	-0.02	7.16×10^{-1}	n.s.	-0.02	7.28×10^{-1}	n.s.	-0.10	3.46×10^{-2}	n.s.
Alanine	0.16	7.26×10^{-4}	n.s.	0.16	7.36×10^{-4}	n.s.	0.06	1.83×10^{-1}	n.s.
Valine	0.23	7.14×10^{-7}	n.s.	0.23	1.29×10^{-6}	n.s.	0.08	1.16×10^{-1}	n.s.
Cysteine	0.33	1.40×10^{-12}	n.s.	0.31	4.59×10^{-11}	n.s.	0.20	4.52×10^{-5}	n.s.
Isoleucine	0.06	1.73×10^{-1}	n.s.	0.06	2.16×10^{-1}	n.s.	-0.03	5.18×10^{-1}	n.s.
Leucine	0.26	3.04×10^{-8}	n.s.	0.26	6.30×10^{-8}	n.s.	0.10	4.28×10^{-2}	n.s.
Tyrosine	0.32	9.21×10^{-12}	n.s.	0.32	1.56×10^{-11}	n.s.	0.22	4.99×10^{-6}	n.s.
Phenylalanine	0.25	1.57×10^{-7}	n.s.	0.25	8.54×10^{-8}	n.s.	0.12	1.40×10^{-2}	n.s.
Ethanolamine	0.31	3.44×10^{-11}	n.s.	0.31	8.26×10^{-11}	n.s.	0.16	8.32×10^{-4}	n.s.
Lysine	0.25	1.28×10^{-7}	n.s.	0.24	5.56×10^{-7}	n.s.	0.13	6.45×10^{-3}	n.s.

Urinary variable	Model 1*			Model 2 [†]			Model 3 [‡]		
	r	P	Q	r	P	Q	r	P	Q
1-Methylhistidine	0.12	8.90×10^{-3}	n.s.	0.08	7.99×10^{-2}	n.s.	-0.03	4.86×10^{-1}	n.s.
Histidine	0.11	1.85×10^{-2}	n.s.	0.12	1.39×10^{-2}	n.s.	0.01	8.26×10^{-1}	n.s.
Tryptophan	0.21	9.20×10^{-6}		0.20	1.80×10^{-5}		0.12	1.03×10^{-2}	n.s.
3-Methylhistidine	0.31	2.98×10^{-11}		0.29	5.05×10^{-10}		0.08	9.31×10^{-2}	n.s.
Carnosine	0.16	9.36×10^{-4}		0.17	2.90×10^{-4}		0.01	8.14×10^{-1}	n.s.

*Model 1 is adjusted for age, gender, and sample.

[†]Model 2 is adjusted for all factors in model 1 plus cardiovascular disease history (heart disease/stroke), physical activity, medication for hypertension, prescribed lipid-lowering drugs, NSAID use, dietary supplement use, special diet, smoking, education, and total energy intake per day (kcal/day).

[‡]Model 3 is adjusted for all factors in models 1 and 2 plus 24-hour urinary creatinine.

[§]The signal at δ 6.91 of tyrosine also includes a signal from its metabolite 4-hydroxymandelate.