

RESEARCH

Open Access



A meta-analysis study of the robustness and universality of gut microbiome-metabolome associations

Efrat Muller¹, Yadid M. Algavi² and Elhanan Borenstein^{1,2,3*}

Abstract

Background: Microbiome-metabolome studies of the human gut have been gaining popularity in recent years, mostly due to accumulating evidence of the interplay between gut microbes, metabolites, and host health. Statistical and machine learning-based methods have been widely applied to analyze such paired microbiome-metabolome data, in the hope of identifying metabolites that are governed by the composition of the microbiome. Such metabolites can be likely modulated by microbiome-based interventions, offering a route for promoting gut metabolic health. Yet, to date, it remains unclear whether findings of microbially associated metabolites in any single study carry over to other studies or cohorts, and how robust and universal are microbiome-metabolites links.

Results: In this study, we addressed this challenge by performing a comprehensive meta-analysis to identify human gut metabolites that can be predicted based on the composition of the gut microbiome across multiple studies. We term such metabolites “robustly well-predicted”. To this end, we processed data from 1733 samples from 10 independent human gut microbiome-metabolome studies, focusing initially on healthy subjects, and implemented a machine learning pipeline to predict metabolite levels in each dataset based on the composition of the microbiome. Comparing the predictability of each metabolite across datasets, we found 97 robustly well-predicted metabolites. These include metabolites involved in important microbial pathways such as bile acid transformations and polyamines metabolism. Importantly, however, other metabolites exhibited large variation in predictability across datasets, suggesting a cohort- or study-specific relationship between the microbiome and the metabolite. Comparing taxonomic contributors to different models, we found that some robustly well-predicted metabolites were predicted by markedly different sets of taxa across datasets, suggesting that some microbially associated metabolites may be governed by different members of the microbiome in different cohorts. We finally examined whether models trained on a control group of a given study successfully predicted the metabolite’s level in the disease group of the same study, identifying several metabolites where the model was not transferable, indicating a shift in microbial metabolism in disease-associated dysbiosis.

Conclusions: Combined, our findings provide a better understanding of the link between the microbiome and metabolites and allow researchers to put identified microbially associated metabolites within the context of other studies.

Keywords: Gut microbiome, Metagenomics, Metabolomics, Multi-omic, Meta-analysis, Machine learning

* Correspondence: elbo@tauex.tau.ac.il

¹The Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel

²Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

The microbial communities that reside in the human gut wield a multitude of activities with pervasive effect on human health and disease. Among these activities, perhaps the most important and clinically relevant one is the microbiota's metabolic activity. Indeed, metabolites of microbial origin or metabolites that undergo microbial transformation have been implicated in various host processes, including immune system development and activity, host metabolism, and even brain function [1–4]. Several recent studies have also established a causal link between microbially produced metabolites and various medical conditions (or benefits). One well-known example is that of trimethylamine *N*-oxide (TMAO), a product of microbial metabolism of nutrients found in eggs and red meat, which accelerates atherosclerotic cardiovascular diseases [5, 6]. Other examples include imidazole propionate, a metabolite over-produced by Type 2 Diabetes-associated bacteria, which was found to impair glucose tolerance [7], *Akkermansia muciniphila*-associated nicotinamide, which was shown to improve motor symptoms in a mouse model of ALS [8], and microbially produced short-chain fatty acids (SCFAs), which have a protective role against intestinal inflammation [9]. Other microbial metabolites protect from influenza [10] and even improve mental quality of life [11]. Moreover, microbiome-associated metabolites such as the organic acid taurine or some SCFAs have been studied as potential novel therapeutic agents, intended to “correct” negative effects of microbial dysbiosis [12].

This appreciation for the metabolic role of the gut microbiome in maintaining host health or promoting disease calls for a better understanding of which gut metabolites are governed by the microbiota, and specifically by which members of the microbiota. To address this challenge, *in vitro* culturing of common gut strains [13, 14], *in vivo* experiments comparing germ-free mice to humanized mice [15], as well as bioinformatic analyses of bacterial genomes [16], have all been applied in an attempt to map the metabolic potential of the human gut microbiota. However, additional factors shaping the human gut metabolome, including host genetics, diet, medications, as well as other exogenous factors, render an extremely complex system in which the exact role of the microbiota remains challenging to untangle [3, 17].

One promising approach for systematically evaluating microbe-metabolite links in the human gut relies on integrative data analysis of paired microbiome and metabolome profiles [18–32]. Such studies, referred to here as ‘microbiome-metabolome’ studies, typically apply high-throughput sequencing and metabolomics technologies to a set of fecal samples from a cohort of interest, and then utilize various statistical and computational methods to identify potential links between the obtained

microbiome and metabolome profiles. Such analyses commonly involve correlation- and linear regression-based methods, aiming to estimate relations between specific metabolites and specific microbes or between specific metabolites and the entire microbial community. More recently, classic machine learning and deep learning models have also been applied to predict metabolite levels based on microbiome data and to highlight the main taxa associated with each metabolite [33–35]. MellonPan, for example, uses an elastic net model to predict metabolite levels based on functional profiles of microbiome samples, and was shown to well-predict 107 out of 466 identified metabolites, including sphingolipids, fatty acids, and B-group vitamins [33].

Such microbiome-metabolite associations reported in individual studies of the human gut, however, raise several questions. First, it remains unclear what do such associations mean biologically, and whether or not they reflect underlying mechanisms. In fact, our lab has recently shown that microbe-metabolite correlations have extremely high false positive rates in predicting mechanistic links in simulated data [36]. Second, in the context of case-control studies, it is often unclear how the association is related to the phenotype/disease that is under study and how to interpret findings of associations that exist in one study group but not the other vs those that exist in both groups. Third, it is unclear how generalizable are such associations, given that they were identified in a specific cohort, often using a specific computational method and a specific metagenomics/metabolomics processing protocol. Put differently, it is often not known how replicable are reported associations in other human gut datasets of varying geographies, ages, sample processing protocols, and profiling platforms [37, 38]. Moreover, even when analyzing large cohorts, the transferability of microbiome-based metabolite predictions to new cohorts is not guaranteed, as recently demonstrated on serum metabolites [35].

In this study, we focus primarily on the third question above and attempt to characterize the landscape of gut metabolites that are consistently well-predicted by the gut microbiome, as captured in fecal microbiome-metabolome datasets. We hypothesize that some of the reported microbiome-associated metabolites are biologically meaningful, representing metabolites that are universally governed by the microbiome across different cohorts and biological backgrounds, and are robust to the variation in the technical settings used in each study. We accordingly aim to identify a set of gut *microbiome*-metabolite associations, consistent across multiple cohorts and settings, and explore the specific genera contributing to the predictability of each such metabolite. To this end, we obtained multiple paired fecal microbiome-metabolome datasets and focused on the healthy individuals in each dataset. For each dataset, we then

trained and evaluated a machine learning model per metabolite, aiming to predict the metabolite level in each sample based on microbiota composition. We examined which metabolites were consistently well-predicted across multiple datasets, using rigorous statistical methods inspired by traditional meta-analysis techniques to define “robustly well-predicted” metabolites. Lastly, we analyze which genera contribute most to the predictability of each robustly well-predicted metabolite, and whether these too are consistent across datasets and in healthy vs. disease cohorts. We believe that this cross-study perspective on suspected microbial-governed metabolites in the human gut is crucial for prioritizing further research hypotheses and for enhancing interpretations of future microbiome-metabolome studies by providing relevant context.

Results

A unified human fecal microbiome-metabolome multi-study dataset collection

We collected and processed data from 10 different human gut microbiome-metabolome studies, totaling 779 samples from 629 individuals in ‘healthy’ (‘control’) groups and 954 samples from 729 individuals in 7 ‘disease’ (‘case’) groups (Table 1, Fig. 1A, Additional file 1: Table S1). Notably, these studies spanned various ages, geographies, health conditions, metagenomics/metabolomics platforms, and 16S rRNA gene hypervariable regions, all of which are expected to introduce heterogeneity between datasets, as demonstrated in

previous microbiome and metabolome meta-analysis studies in various fields [44–46]. Importantly, in case-control studies, we treated healthy and disease subgroups separately (considering only study groups with ≥ 40 samples) to avoid the confounding impact of the disease state on both the microbiome and metabolome compositions. Moreover, to first focus on robustly well-predicted metabolites in general population-like cohorts, in our analysis below we initially considered *only* the healthy datasets (Fig. 2A).

Microbiome data was processed to obtain genus-level profiles, providing more comparable taxonomic profiles across 16S rRNA gene sequencing and whole genome shotgun sequencing (WGSS) datasets at the expense of sensitivity and resolution (Fig. 1B), as done in several recent microbiome-related meta-analysis studies [44, 47]. Specifically, when possible, 16S rRNA gene sequencing raw data were re-processed using QIIME2 [48] to obtain genus-level relative abundances and MetaPhlan2 tables were collapsed to genus-level profiles (see full details in Additional file 1: Table S2). Metabolite identifications were unified by converting the identifications given in each dataset to the Human Metabolome Database (HMDB) metabolite identifiers [49]. Further processing details, statistics, and study-specific required adjustments are detailed in the “Methods” section and in Additional file 1: Tables S2–S4.

Clearly, the set of detected features in both microbiome and metabolome profiles (i.e., genera and

Table 1 Studies included in the main analysis (additional details provided in Additional file 1: Table S1)

Dataset ID [abbreviation] ¹	Reference	# samples – control ²	# samples – cases (disease/condition) ²	Microbiome data description	Metabolome data description	# metabolites ³
KIM_ADENOMAS [KI]	[27]	102	138 (Colorectal adenomas and CRC)	16S, V3-V5	Untargeted, MS	410
YACHIDA_CRC [YA]	[39]	127	220 (CRC)	WGSS	Targeted, MS	387
FRANZOSA_IBD [FR]	[22]	56	157 (IBD)	WGSS	Untargeted, MS	294
HE_INFANTS [HE]	[40]	68	–	16S, V4	Targeted, NMR	113
iHMP_IBD [iH]	[19]	72 [26]	212 [79] (IBD)	WGSS	Untargeted, MS	504
JACOBS_IBD_RELATIVES [JA]	[21]	54	– ⁵	16S, V4	Untargeted, MS	45
MARS_IBS [MA]	[41]	70 [24]	143 [51] (IBS)	WGSS	Targeted, NMR, and MS	40
POYET_BIO_ML [PO]	[26]	141 [83]	–	16S ⁴ , V4	Untargeted, MS	313
SINHA_CRC [SI]	[42]	89	42 (CRC)	16S, V3-V4	Untargeted, MS	351
ERAWIJANTARI_GC [ER]	[43]	– ⁵	42 (History of gastrectomy for GC)	WGSS	Untargeted, MS	342

¹The “Dataset ID” is formatted as follows: <First author/cohort name>_<Short cohort description>. The 2 letter abbreviations are used for plotting purposes

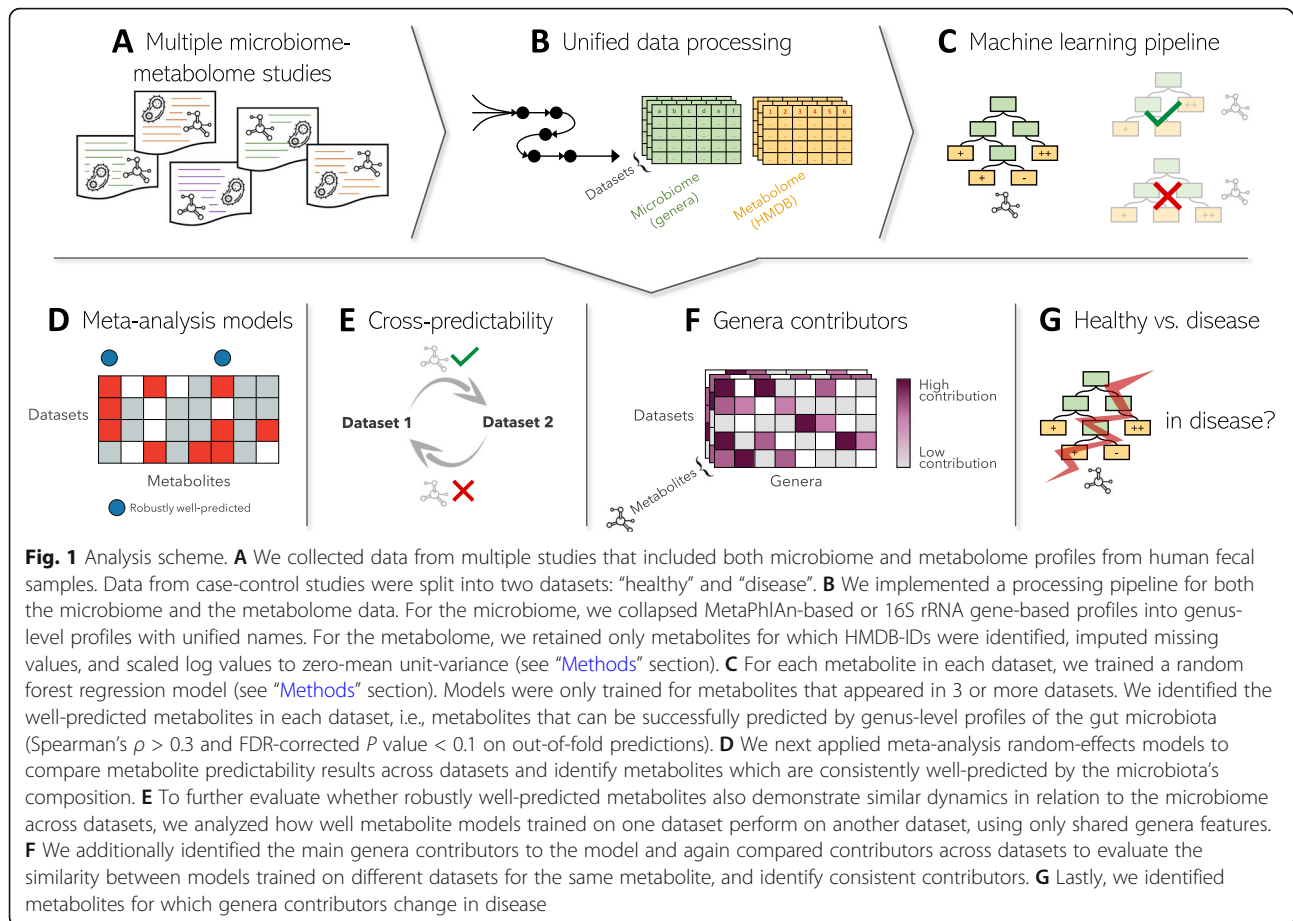
²Sample numbers refer only to samples included in the analysis herein. In cases of temporal datasets with multiple samples per individual, the number of individuals is noted in square brackets

³Number of HMDB-annotated metabolites. See “Methods” section

⁴WGSS data was also available but not used in the analysis herein

⁵This study group was not used

CRC colorectal cancer, MS mass spectrometry, 16S 16S rRNA gene sequencing, WGSS whole genome shotgun sequencing, IBD inflammatory bowel disease, IBS irritable bowel syndrome, GC gastric cancer

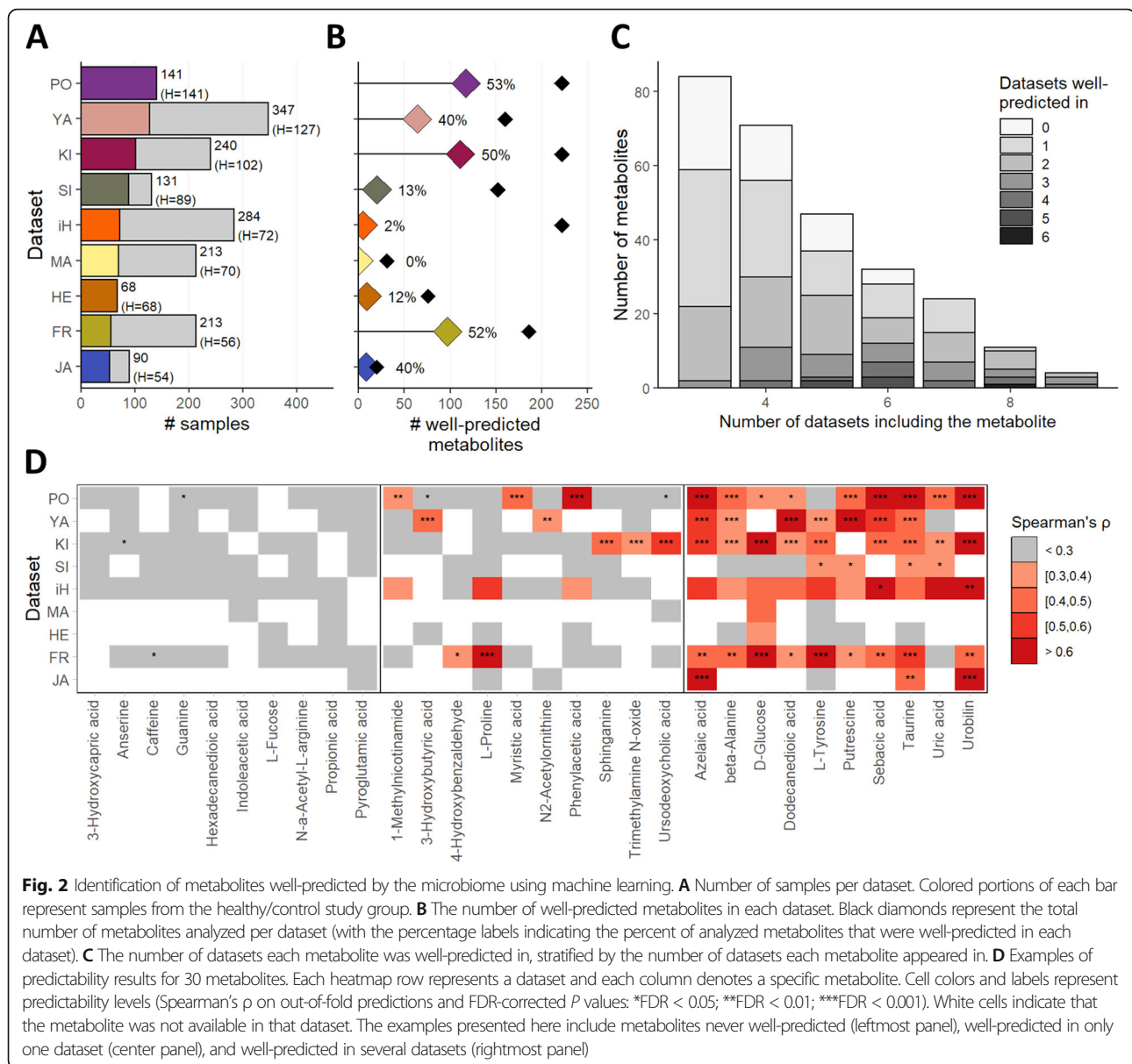


metabolites, respectively) are expected to differ between studies due to both technical [2, 50, 51] and biological factors [20, 52, 53]. This limited overlap across datasets hinders in and of itself our ability to carry microbiome-metabolite links over different studies, and is discussed further in the “Discussion” section below. Here, in order to detect consistently well-predicted metabolites, we first examined which gut metabolites are shared between datasets, limiting our analysis to those that appear in 3 or more datasets. Out of 951 unique, non-rare, HMDB compound IDs found across all datasets, 273 (29%) were shared among 3 or more datasets (Additional file 3: Figure S1A). We additionally examined which genera are shared between datasets, to facilitate specific analysis of contributors to well-predicted metabolites. Out of 85 unique, none-rare genera found across all datasets, 55 (65%) were shared between 3 or more datasets (Additional file 3: Figure S1B). Pairwise comparisons between studies in terms of overlap in microbial and metabolic features further demonstrate the large variability and often limited overlap (Additional file 3: Figure S1C-D).

Predictability of metabolite levels based on microbiome data

We implemented a machine learning pipeline to estimate how well genus-level profiles can predict metabolite levels in each dataset and for each metabolite (Fig. 1C). Specifically, for each HMDB-annotated metabolite that appeared in 3 or more datasets, we trained a random forest regression model to predict metabolite levels based on genera relative abundances. Alternative pipelines with either a different machine learning model or a different hyperparameter tuning process were examined as well (see Additional file 2: Supplementary Note 1). We evaluated each model’s performance using leave-one-out cross validation by calculating the Spearman’s correlation coefficient, ρ , between actual vs predicted left out metabolite levels. Spearman’s correlation P value was also recorded, and FDR-correction was applied to all metabolite-models in each dataset (see “Methods” section). Metabolites with a predictability of $\rho > 0.3$ and an FDR < 0.1 were referred to as ‘well-predicted’ metabolites.

Overall, 1255 metabolite predictor models (i.e., for a specific metabolite in a specific dataset) were trained,



covering 273 unique metabolites in 9 healthy datasets. Of these, 418 models were able to successfully predict the metabolite level (with $\rho > 0.3$ and FDR < 0.1), and accordingly defined as well-predicted. In each individual dataset, 0–53% of the analyzed metabolites were well-predicted (Fig. 2B). Of the 273 unique metabolites, 219, 125, and 49 metabolites were well-predicted in at least 1, 2, or 3 datasets, respectively, while 54 were never well-predicted (Fig. 2C, D). Full predictability results for each metabolite in each dataset are provided in Additional file 1: Table S5.

We further validated that these predictability results cannot be attributed to statistical noise or other artifacts in the data by comparing the fraction of well-predicted metabolites in each dataset to the fraction of well-

predicted metabolites obtained in a shuffled dataset (see “Methods” section). We found that in such shuffled datasets only 0.3% of the metabolites in each dataset were well-predicted on average, compared to 29.1% well-predicted metabolites in the real data. As an additional independent validation for our machine-learning pipeline, we confirmed that our set of well-predicted metabolites in one specific dataset exhibits marked overlap with the set of predictable metabolites found by another study that analyzed this same dataset (see Additional file 2: Supplementary Note 2 and Additional file 3: Figure S2A).

Robustness of metabolite predictability

While the results reported above already demonstrate intriguing variation in predictability and highlight several

metabolites that appear to be well-predicted across multiple datasets, we next applied a more rigorous statistical approach for synthesizing and combining the predictability results obtained for each dataset independently and for quantifying how *robustly* well-predicted is each metabolite. Specifically, we used random-effects models (REM)—a common meta-analysis statistical framework for integrating effect sizes from multiple studies with different sample sizes and study designs, and for calculating an estimate for the mean effect across studies and its associated P value [54] (see “**Methods**” section). We applied this framework to the predictability results obtained above for each metabolite, using the level of predictability (scored by Spearman’s ρ) as the measured effect size in each dataset. The resulting REM for each metabolite thus provides an estimation of the mean predictability of that metabolite across datasets, with an associated P value (that was FDR-corrected to account for multiple REMs). Finally, we defined metabolites with a REM predictability score > 0.3 and FDR < 0.1 as ‘*robustly well-predicted*’, resulting in a total of 97 robustly well-predicted metabolites (Fig. 3A, B and Additional file 1: Table S6). For comparison, only 87.7 ± 5.5 metabolites met the above threshold on average when applying our pipeline to datasets where metabolite labels were shuffled (see “**Methods**” section), suggesting that the number of robustly well-predicted metabolites in the original dataset is somewhat higher than expected by chance ($P = 0.053$).

Examining the resulting robustly well-predicted metabolites, we found that they span multiple metabolite classes including amino acids, carbohydrates, and bile acids, and are significantly enriched in the “fatty acids and conjugates” class (FDR-corrected P value: 0.026, Fisher’s exact test, Additional file 3: Figure S3A-B). Importantly, our set of robustly well-predicted metabolites included multiple metabolites that take part in clinically important pathways known to involve the gut microbiome. For example, this set highlights the microbiota’s known essential role in the transformation of primary bile acids to secondary bile acids and the metabolism of other bile components [55] (Fig. 3C). Additional pathways in which multiple metabolites were robustly well-predicted by the microbiome included trimethylamine N-oxide (TMAO) metabolism, tryptophan, and indole derivatives metabolism (Additional file 3: Figure S3C), polyamine biosynthesis (Fig. 3D), poly-unsaturated fatty acids (PUFAs), specifically omega-3 and omega-6 acids pathways (Additional file 3: Figure S3D-E), L-proline biosynthesis, sugars metabolism, and metabolites involved in the gut-brain axis, all of which are known to involve the gut microbiota. We additionally found that robustly well-predicted metabolites tend to have a larger proportion of their variance explained by the microbiome, as estimated by an independent study of the largest microbiome-metabolome dataset collected to date [20]

(see Additional file 2: Supplementary Note 2 and Additional file 3: Figure S2B).

Interestingly, we also found several robustly well-predicted metabolites that are only rarely considered in the context of the gut microbiota. One such example is medium-chain fatty acids (MCFAs), including specifically caproic acid (HMDB0000535), pentadecanoic acid (HMDB0000826), and heptanoic acid (HMDB0000666). MCFAs are highly abundant components in dairy foods and have an essential physiological role as efficient cell energy sources [56]. A recent study demonstrated a significant association between dairy foods intake and the microbiota’s composition, suggesting an intriguing interaction between gut microbes and these MCFAs [57]. Another example is dicarboxylic acids, including specifically malonic acid (HMDB0000691), undecanedioic acid (HMDB0000888), pimelic acid (HMDB0000857), suberic acid (HMDB0000893), sebacic acid (HMDB0000792), azelaic acid (HMDB0000784), and dodecanedioic acid (HMDB0000623). The clinical relevance of both metabolite classes, specifically in the context of type 2 diabetes [58, 59], and their relatively unexplored relation with the gut microbiome, suggest interesting directions for future research. Additional file 2: Supplementary Note 3 details additional intriguing information concerning the set of robustly well-predicted metabolites, their biological roles, and previously reported links to the microbiome.

Similarity and variation between metabolite models from different datasets

So far, we have identified metabolites that are consistently well-predicted by the composition of the human gut microbiota across multiple datasets. We next sought out to examine whether the models trained for predicting these metabolites across the different datasets are in fact similar to one another. To this end, we first estimated the contribution of each genus feature to each model using the permutation-based approach from Altman et al. (2010) to assign a P value for each feature in each model (“**Methods**” section and Additional file 1: Table S7). We refer to genera with $P < 0.1$ as significant contributors. Comparing models trained on different datasets for the same metabolite, we observed substantial variation in the numbers of shared significant contributors (Fig. 4 and Additional file 3: Figure S4), suggesting that even robustly well-predicted metabolites may not necessarily be predicted by the same microbiota features in different datasets.

To further examine the similarity between models, we also looked at cross-predictability, i.e., how well does a metabolite model trained on one dataset predict the levels of the metabolite in samples from another dataset (considering only shared genus features; see “**Methods**” section). Mirroring our findings above, for some

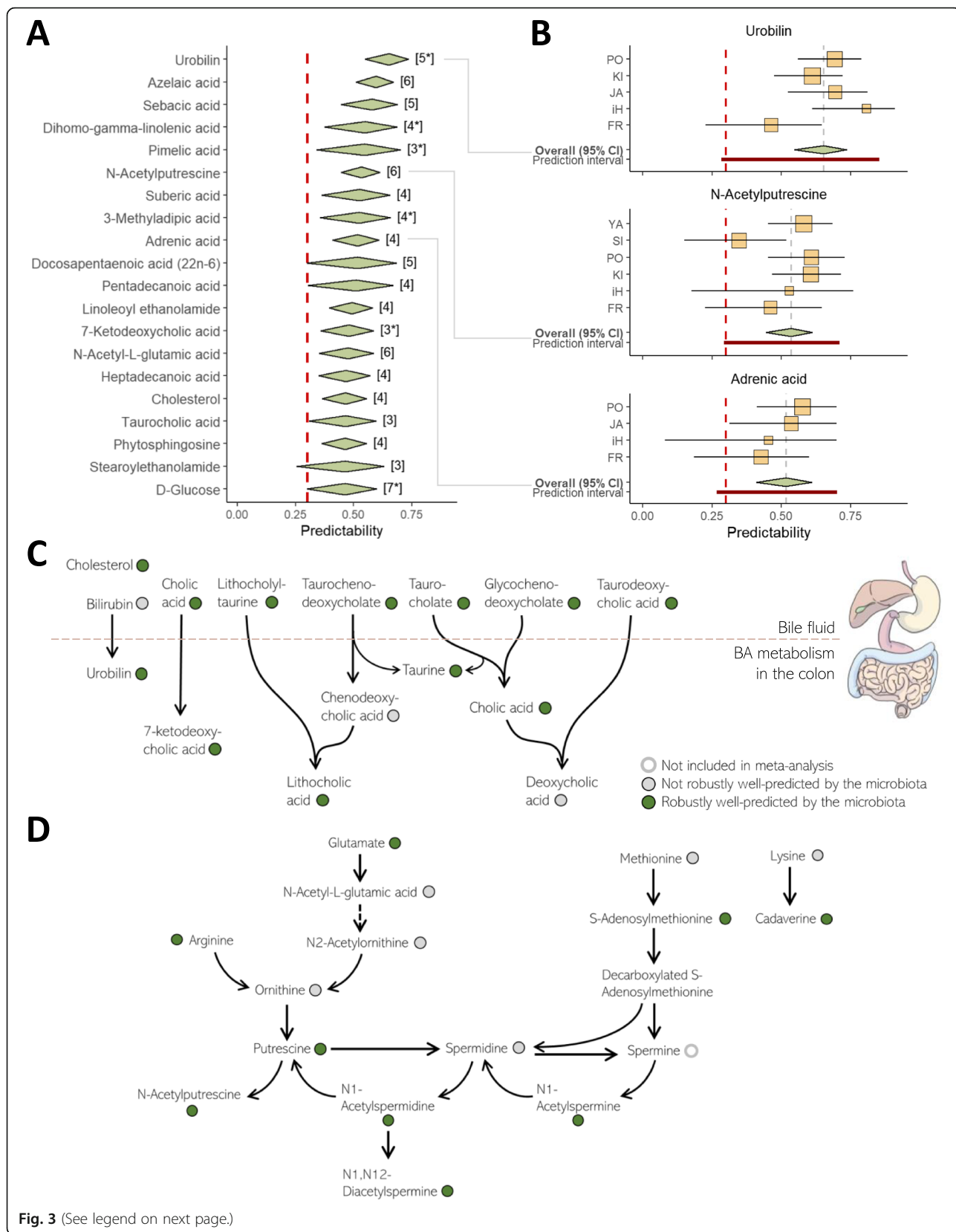


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Robustly well-predicted metabolites. **A** Top 20 robustly well-predicted metabolites. Diamonds' centers represent the random-effects model's estimated mean effect size (mean predictability) and diamonds' widths represent the mean's 95% confidence interval. The numbers in square brackets represent the number of datasets in which the metabolite was available. A star beside the number of datasets indicates that in one or more of the datasets, this metabolite was annotated with low confidence (see "Methods" section). The red dashed line represents a Spearman's correlation of 0.3, which we defined as the threshold for a successful predictive model. **B** Forest plots of 3 robustly well-predicted metabolites. Yellow boxes represent predictability estimates from each dataset, with box sizes proportional to the number of samples and horizontal lines denoting the 95% confidence interval for each estimate. The diamonds and red dashed line are as above. Prediction intervals (dark red interval) further indicate the predicted range of the effect in 95% of unobserved studies [54]. **C** An illustration of a part of the bile metabolism pathway (including primary bile acids, secondary bile acids, and other bile components). Colored circles indicate whether the metabolite was robustly well-predicted by the gut microbiome (green), not robustly well-predicted (gray), or not included in the analysis (white). **D** An illustration of parts of the polyamines metabolic pathways. Color coding is as in panel **C**.

metabolites such as beta-alanine, sebacic acid, taurine, and N-acetylputrescine, models were successfully transferred between most dataset pairs (Fig. 4B, C, Additional file 3: Figure S4A-B), whereas for other metabolites, such as L-tyrosine and xanthine, models were not transferable between *any* pair of studies (Fig. 4D, Additional file 3: Figure S4C).

Importantly, model comparisons, either by comparison of shared contributors or by cross-predictability analysis, are somewhat impeded by technical variation between studies, and specifically by differences in the microbiome profiling technology used (16S rRNA gene vs. WGSS). Indeed, dataset pairs that used the same metagenomic technology have a significantly higher number of shared features in comparison with dataset pairs that used different technologies (29.6 vs. 22.3 on average; Mann-Whitney P value < 0.0001), which in turn affects the number of shared significant contributors (1.4 vs. 0.6, P value < 0.0001). Yet, we did find examples where despite a large overlap in available features, models still markedly differed in the set of detected significant contributors, as was the case, for example, for L-tyrosine and cholic acid and when comparing the models obtained for the datasets YACHIDA_CRC_HEALTHY (YA) vs FRANZOSA_IBD_HEALTHY (FR) (Fig. 4D, Additional file 3: Figure S4D).

Notably, comparing the set of significant contributors across models and datasets can further provide intriguing and valuable insights. For example, we defined genera that contributed to over 50% of the metabolite models they participated in, "consistent contributors" to that metabolite ("Methods" section; Additional file 1: Table S8), and found that often such genus-metabolite links coincide with known metabolic capacities of certain gut bacteria. *Bacteroides*, for example, consistently contributed to models of several bile components, whereas *Alistipes* consistently contributed to various polyamine models, both in agreement with experimental findings [55, 60, 61] (see Additional file 2: Supplementary Note 3 for additional examples). We also note that even though genera more abundant in the gut were more likely to significantly contribute to more

metabolite models (Spearman's correlation 0.45, P value < 0.0001), some genera strongly contributed to the predictability of multiple gut metabolites even though their abundance in the gut is typically low. *Odoribacter*, for example, a member of the *Bacteroidetes* phylum whose mean relative abundance in our data was 0.4%, was a significant contributor in 94 of 223 models it participated in. This genus was also a consistent contributor to models of multiple bioactive metabolites such as L-phenylalanine, pimelic acid, linoleoyl ethanolamide, and L-tryptophan. Similarly, *Haemophilus* (mean relative abundance in our data of 0.4%), from the *Proteobacteria* phylum, was a significant contributor in 49 of 126 models. Finally, we note that the average number of significant contributors per metabolite ranged from 5.6 (L-Arginine) to 10.6 (Cholic acid and Dihomo-gamma-linolenic acid), suggesting an association with multiple different genera.

Metabolite predictability in health vs. disease

We eventually sought to determine whether various patterns identified above for the healthy datasets also hold in disease. To this end, we additionally considered the 7 disease datasets obtained from the same studies as described above (Table 1). These included 3 datasets of patients with colorectal adenomas or cancer, 2 of patients with inflammatory bowel disease (IBD), 1 of patients with a history of gastric cancer, and 1 of patients with irritable bowel syndrome (IBS), totaling 954 samples (Table 1, Additional file 1: Table S1).

We first examined whether robustly well-predicted metabolites were also consistently well-predicted in the disease datasets. We accordingly used an expansion of REMs (termed subgroup meta-analysis) that considers studies from two or more different cohort types (here: healthy and disease) and estimates both the mean effect within each subgroup of studies, as well as the between-subgroup differences (see "Methods" section). Out of 97 robustly well-predicted metabolites identified above, 81 were also present in at least 3 disease datasets, and of these 57 (70%) were also robustly well-predicted in disease datasets using

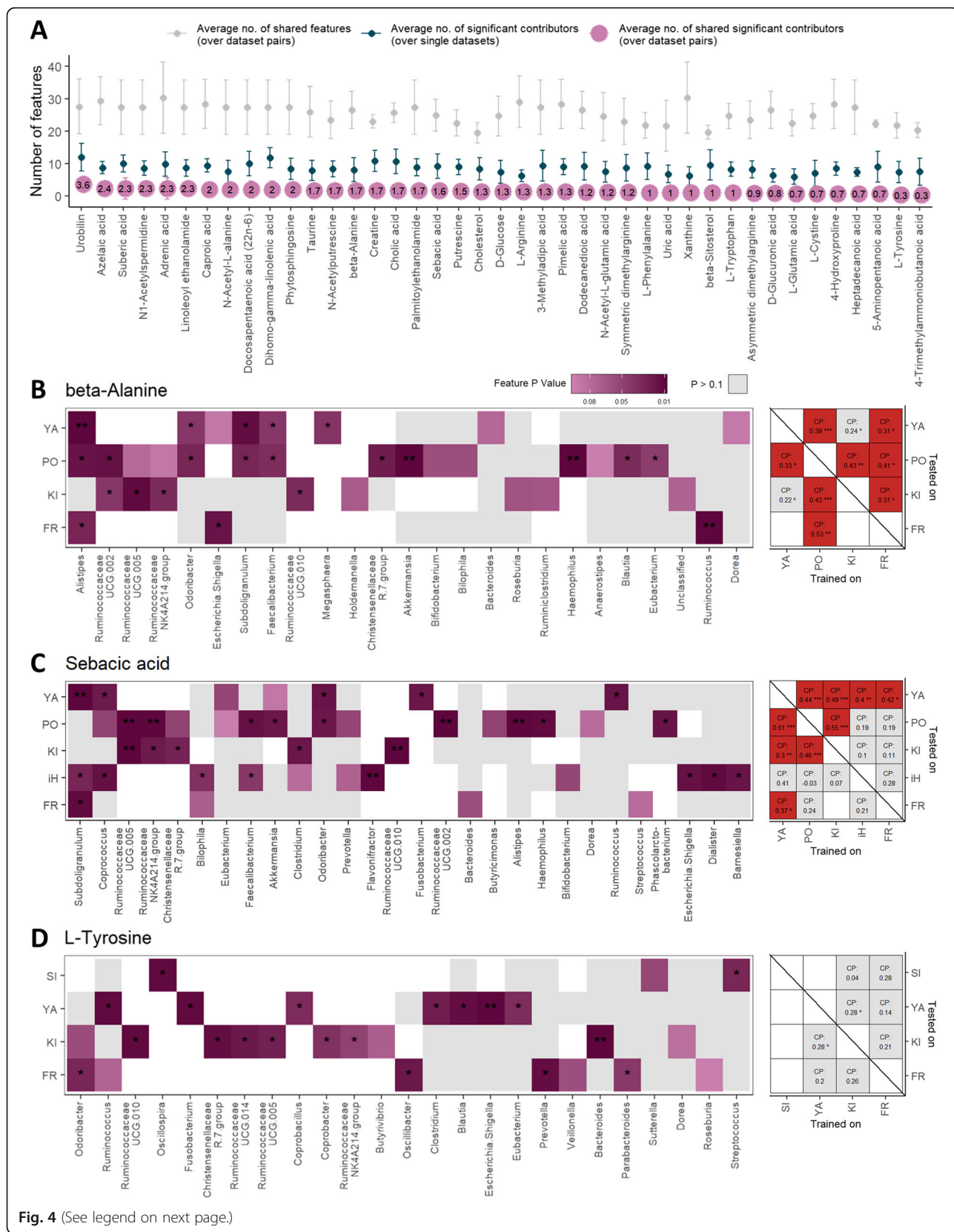


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Comparisons of metabolite models between datasets. **A** An overview of pairwise model comparisons for each robustly well-predicted metabolite. For each metabolite, we compared the number of shared features (i.e., genera; grey points) and the number of shared significant contributors (purple points) for every pair of models trained on two different datasets. For each metabolite we additionally note the average number of significant contributors over all models trained for that metabolite (turquoise points). Error bars represent standard deviations. **B** A detailed comparison of models for predicting beta-alanine levels. Left panel: contributors comparison. Each row represents a dataset in which the metabolite was well-predicted by the microbiome and each column represents a genus feature (only features significantly contributing to at least one model are included). Purple-colored cells denote the significance of the specific feature in the specific model ($P < 0.1$). White cells indicate that the feature was not available in the specific dataset. Right panel: Cross-predictability analysis. Matrix columns indicate the dataset used for training and matrix rows indicate the dataset used for testing. Numbers in cells indicate Spearman's correlation between predicted and actual metabolite levels in the test dataset. Red cells denote cases where the model was well transferred from the column-dataset to the row-dataset. **C, D** include similar plots as **B**, for sebacinic acid and L-tyrosine metabolites, respectively. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; CP cross predictability

the same thresholds as before (Additional file 1: Table S9, Additional file 3: Figure S5). Importantly, even though the remaining 24 metabolites did not meet our criteria, none exhibited a significantly different effect than the one observed in the healthy datasets (as estimated by the Q statistic), suggesting that statistical power was insufficient in these cases.

Finally, we examined whether metabolites that are well-predicted in both the healthy and disease datasets of a specific study, are well-predicted by similar models. As before, we compared models both by comparing their significant contributors and by testing how well the healthy-trained model transfers to the pertaining disease dataset. Notably, here we analyzed each case-control study independently, allowing us to eliminate potential heterogeneity between datasets, and focus on model-differences associated with the disease. Overall, we performed healthy vs. disease comparisons for 128 metabolites over 5 studies in which both healthy and disease datasets were available (covering 83 distinct metabolites). We observed that many metabolite models transferred well from healthy to disease in different studies, suggesting that microbiome-metabolite dynamics in these cases are not substantially altered in the disease state. Docosapentaenoic acid (22n-6), 12,13-DHOME, and phytosphingosine, for example, transferred well from healthy to disease in both an IBD cohort and a colorectal adenomas cohort. N-acetylputrescine, pimelic acid, taurine, and tryptamine similarly transferred well in two different colorectal adenomas/cancer cohorts, and sebacinic acid transferred well in all 3 datasets analyzed (Additional file 3: Figure S6). In some of these cases, we additionally observed a substantial overlap in significant contributors between the healthy and disease models (Additional file 3: Figure S7). In contrast, several metabolite models did not transfer well even though the metabolite was well-predicted in both the healthy dataset and linked disease dataset independently, potentially indicating a shift in microbiome metabolism in disease (Additional file 3: Figure S6; Additional file 3: Figure S7). The full catalog of significant contributors to both the healthy and the disease models can be found in Additional file 1: Table S7.

Methods

Data acquisition

In this study, we acquired data from several published studies of the human gut microbiome and metabolome. We focused on studies that included at least 40 individuals in each study group (or total, in non-case-control studies), for which both the microbiome and the metabolome were profiled from fecal samples. In case-control studies, cases and controls were treated as separate datasets to avoid the confounding impact of the disease state on the composition of both the microbiome and metabolome. Metabolomic datasets with only metabolite m/z info and without identification by name, KEGG IDs [62], or HMDB IDs [49] were discarded. For microbiome data, we considered both 16S rRNA gene sequencing and WGSS datasets. For 16S rRNA gene studies, we obtained raw fastq files and for WGSS datasets we obtained MetaPhlan2 tables (see Additional file 1: Table S2 for additional information). For longitudinal datasets, where multiple samples were collected from each participant over time, we randomly chose up to 3 samples per subject, to avoid having a few subjects potentially dominating the dataset. Datasets were either downloaded from public repositories (e.g. NCBI Sequence Read Archive, Qiita [63]) or shared by the corresponding authors. As to June 2020, we identified over 70 published microbiome-metabolome studies of the human gut (reflected by fecal samples specifically), approximately 40 of which contained a sufficient number of samples as noted above. Of these, we were able to obtain the necessary data and metadata from only 10 datasets (either from online repositories or through the authors). Table 1 and Additional file 1: Table S1 describe the included datasets in detail and reference their corresponding publications. Since the difference between infants' and adults' microbiomes may introduce a particularly prominent source of heterogeneity, we also conducted a supplementary analysis, excluding the one infant dataset included in our study (the HE dataset). Overall, this exclusion did not markedly impact our findings and its results are reported in Additional file 2: Supplementary Note 4.

Pre-processing of metabolomics datasets

We obtained processed metabolomics datasets from each study. As metabolomics technology, processing method, and output differed substantially between studies (see Additional file 1: Table S1 and S4), we pre-processed the metabolome datasets to allow comparison of metabolite-related findings across datasets. Importantly, we focused on unifying metabolite identifiers, and did not attempt to unify the measured values themselves. Specifically, we mapped metabolite identifications in each dataset to HMDB IDs. In many datasets, metabolites in the obtained data were identified by their names only, and we mapped metabolites to HMDB identifiers using MetaboAnalyst's compound ID conversion utility [64]. Additional manual curation was performed in order to map metabolite names that were not identified by MetaboAnalyst (due to, for example, typos or special characters in the name, or in cases where MetaboAnalyst found more than one match). In datasets where metabolite masses were available, manual curation of such metabolites was also validated by mass.

To track cases where metabolite mappings to HMDB identifiers could not be done with high-confidence, we flagged metabolite annotations as “representative” in any of the following cases: Metabolites were marked as “representative” or “low-confidence annotation” in the original data; Annotation was ambiguous (e.g., due to mass spectrometry inability to differentiate two or more metabolites with the same mass and retention time); When the mapping was not automatic by MetaboAnalyst due to some uncertainty (e.g., when the metabolite name was listed with a different spelling); Or when there was some conflict between MetaboAnalyst mappings and original mappings. These “representative” annotations are marked in the supplementary tables and in Fig. 3A to indicate lower confidence.

Finally, metabolites present in 10 or fewer samples in each cohort were discarded, missing or zero values were replaced with 90% of the minimal metabolite concentration observed across all other samples, and metabolite values were log-transformed and scaled to zero-mean and unit-variance, as commonly accepted for such data [65]. Additional file 1: Table S4 details which data was available in each dataset and how the mapping was conducted.

Pre-processing of 16S datasets

16S rRNA gene sequencing raw data was processed using QIIME2 version 2019-1 [48] as follows. When raw data was multiplexed, we demultiplexed the data using QIIME2's demux plugin. We applied DADA2 [66] in order to denoise the data and extract amplicon sequence variants (ASVs). Whenever a sufficient high-quality overlap of forward and reverse reads was available,

DADA2 was also used for merging paired end reads. We trimmed reads in each dataset to the first position with a median quality score under 30. In cases where this resulted in reads shorter than 100 base pairs, reads were trimmed according to the original publication.

To assign ASVs to taxonomy, we trained a Naive Bayes classifier per dataset using QIIME2's feature-classifier plugin [67]. Classifiers were trained on reads extracted from the SILVA 99-OTU database [68], according to the specific 16S rRNA gene hypervariable region used in each dataset. We then collapsed the ASV table to genus-level counts, joining all reads without a genus-level annotation to an “Unclassified” category. This step was performed in order to make genera “entities” as comparable as possible between 16S-based and WGSS-based taxonomic profiles, given the substantial methodological gaps. In each dataset, samples with less than 1000 reads were removed. Read counts were normalized to sum to 1 in each sample, resulting in a table of relative abundances. Lastly, we removed rare genera, defined by either less than 25% non-zero values, or less than 10 non-zero values, or a mean relative abundance over samples of less than 0.1%. Further study-specific parameters and details about the 16S rRNA gene data processing of each dataset are detailed in Additional file 1: Table S2.

Pre-processing of WGSS datasets

For WGSS datasets, we obtained taxonomic relative abundance tables precomputed by MetaPhlan2 [69], with one exception where another detailed taxonomic profile was available (see Additional file 1: Table S2). In order to unify the WGSS and 16S rRNA-based taxonomic profiles, non-bacteria entities were discarded, all data was collapsed to genus-level relative abundances, and genera names were translated when necessary to the genera names used in the 16S rRNA gene processed data. Rare genera were removed from each dataset using the same criteria described above. All downstream analyses were performed on this unified genus-level relative abundance data. Further dataset-specific details about the processing of each WGSS data and unification of genus-level taxonomic annotations are detailed in Additional file 1: Tables S2 and S3.

Implementation of a machine learning pipeline for predicting metabolite levels based on genus abundances

We implemented a machine learning pipeline to estimate how well genus-level profiles can predict metabolite levels in each dataset and for each metabolite. We specifically tried 4 pipeline settings: Random forest (RF) regression with default hyperparameters, RF regression with hyperparameters tuning, elastic net (ENet) with default hyperparameters (effectively making it a lasso

regression), and ENet with hyperparameters tuning. We intentionally focused on relatively simple and commonly used machine learning models (as opposed to top-performing and much more complex models such as XGBoost [70]) to avoid the risk of increased overfitting. A comparison of pipeline results and further details are presented in Additional file 2: Supplementary Note 1. We eventually focused on results from the first pipeline, as it yielded the overall highest number of well-predicted metabolites.

RF regressors were trained for each HMDB-annotated metabolite in each dataset, as long as the metabolite appeared in 3 or more datasets overall. The performance of each such model was estimated using leave-one-out cross validation (LOOCV) to maximize the amount of data for training. In datasets where multiple samples existed per subject, we used a leave-one-subject-out approach and estimated performance using only a single sample from each out-of-fold subject. We calculated the RMSE, R^2 , and Spearman's correlation coefficient, ρ , between actual metabolite levels and predicted ones for out-of-fold samples. Spearman's correlation P value was also recorded, and FDR was applied to all metabolite-models in each dataset separately. Due to the randomness introduced by the RF models, we ran the entire pipeline 5 times per metabolite, verified that results are sufficiently stable (see Additional file 2: Supplementary Note 1), and reported performance metrics that were averaged over these 5 independent runs. We defined metabolites with a Spearman's $\rho > 0.3$ and an FDR < 0.1 as "well-predicted" metabolites, similar to previous related works [33]. Lastly, we compared the number of well-predicted metabolites in each dataset to the number of well-predicted metabolites in a shuffled dataset, where metabolite values were shuffled across samples.

The above pipeline was implemented in R using the "tidymodels" package suite [71]. Specifically, we used "ranger" for RF [72] algorithm and "glmnet" for ENet [73].

Defining robustly well-predicted metabolites

Meta-analysis is a well-established statistical framework for synthesizing results from multiple studies addressing the same topic [54]. Here, in order to synthesize results (correlation coefficients of predicted vs. actual metabolite levels) from multiple datasets, we adopted a classic meta-analysis statistical method, namely, random-effects models (REMs). REMs consider the effect sizes from each study, weighed proportionally to the study's sample size, and assume that in addition to within-study variance, there is true heterogeneity between trials (i.e., between-study variance), likely resulting from differences in study settings and cohort characteristics. True effects are further assumed to follow a normal distribution, with an average effect μ and variation τ^2 . In our case, we chose REMs as we assumed that multiple factors may be affecting the

predictive power of microbiota composition on metabolite levels, ranging from technical factors such as sample collection, storage, sequencing, or metabolomics instruments and processing, to cohort characteristics such as ages, diets, geographic locations, health status, and others.

Specifically, for each of the metabolites that appeared in 3 or more datasets, we used a REM to estimate the distribution of "true effects", i.e., the predictability level of that metabolite based on genus-level profiles of the microbiota. We used the DerSimonian-Laird (DL) estimator for between-study variance, as implemented in the "meta" R package [74]. From each REM we then recorded the estimated average effect μ , its confidence interval, and its associated P value. The estimated average effect of a specific REM can be roughly interpreted as the average predictability of this metabolite over a new set of microbiome-metabolome studies. Smaller P values indicate higher confidence that the mean effect is indeed robust across a wide range of studies and cohorts. FDR correction was applied to all REM P values. Following our definition of "well-predicted" for a specific metabolite in a specific study above, we defined metabolites with a REM average effect > 0.3 and FDR < 0.1 as "robustly well-predicted" by the microbiome. Notably, different estimators for between-study variance (i.e., alternatives to the DL estimator) result in slightly different lists of robustly well-predicted metabolites. Additional file 1: Table S6 includes the full DL-based REM results, as well as information about which metabolites are robustly well-predicted when using alternative estimators.

Finally, to assess the obtained overall number of robustly well-predicted metabolites, we shuffled metabolite labels within each dataset 1000 times. For each shuffled dataset, we then ran the REM pipeline described above and re-computed the number of robustly well-predicted metabolites.

All REM computations were conducted using the "meta" [74] and "metafor" [75] R packages.

Comparison of metabolite models in healthy datasets

For each robustly well-predicted metabolite that was also well-predicted independently in 3 or more datasets, totaling 41 metabolites, we used two analysis approaches to explore how similar are metabolite models trained on different datasets. For all of the below analyses, we only considered datasets in which the metabolite was well-predicted according to the definitions above.

In the first approach, we calculated permutation-based feature importance scores for each genus feature in each model using R's ranger package [72]. We additionally applied the method introduced by Altman et al. (2010) to compute a P value for each feature in each model [76], and considered features with a P value below 0.1 as

significant contributors. We then compared pairs of models by examining the number of shared significant contributors.

In the second analysis for assessing similarity between datasets, we tested how well a model trained on one dataset predicts metabolite levels in another dataset and vice versa, following previous microbiome-related meta-analysis studies [25, 77, 78]. We refer to this analysis as “cross-predictability analysis”. To overcome the differences between each pair of studies, we kept only genus features shared between both datasets, and additionally down-sampled the larger dataset to meet the sample size of the smaller dataset. After taking only the shared features, we re-normalized genus profiles to relative abundances. We then re-trained a RF regression model for the examined metabolite using the first dataset and evaluate its performance using LOOCV. If, due to down-sampling, the metabolite was no longer considered “well-predicted” by the new model, we discarded the specific comparison. Otherwise, we evaluated how well does the model perform on the second dataset and whether the performance meets the “well-predicted metabolite” definition used previously (i.e., the model transfers well). Due to the randomness presented by the down-sampling procedure and the RF itself, the final performance measures we report are averaged over 10 independent runs of the pipeline.

Analysis of consistent genus contributors to metabolite models

We identified genera that consistently contributed to the models of a specific metabolite by calculating the number of models in which a genus was a significant contributor (as previously defined). If a genus significantly contributed to over half of the metabolite models it appeared in, we referred to the genus as a consistent contributor to that metabolite.

Subgroup meta-analysis of healthy vs. disease datasets

Subgroup analysis is the process of comparing an effect (here, predictability of a metabolite) between two or more cohort “variants”. It is often used when some effect is suspected to differ between population subgroups (e.g., children vs. adults). Importantly, while large-scale studies often analyze these differences inherently, subgroup meta-analysis allows comparison of effect sizes between subgroups in smaller-scale studies and even when these subgroups appear in separate studies [54]. Here, we analyzed all metabolites with at least 3 healthy datasets and 3 disease datasets. We specifically used a REM for computing the average effect within subgroups and a fixed-effects model for estimating the between-subgroup differences. For each subgroup meta-analysis, we then recorded the mean predictability (and associated

P value) within healthy datasets (as before), mean predictability (and associated P value) within disease datasets, and the Q statistic and P value for differences between subgroups. Sub-group meta-analysis was conducted using the “meta” [74] R package, based on code from the “dmetar” [79] package.

Comparison of metabolite models in case-control datasets

As described before for comparison of models across healthy datasets, we used two analysis approaches to explore how similar are metabolite models trained on healthy vs. disease datasets from the same study. First, we compared significant contributors as previously described. Second, we performed a cross-predictability analysis to examine how well the healthy-model predicts metabolite levels in the pertaining disease dataset. To account for different study group sizes, the larger group was down-sampled to match the size of the smaller group. If, due to down-sampling, the metabolite was no longer well-predicted in the healthy group, we discarded that specific comparison. We limited our analysis to 5 case-control datasets that also had at least 1 well-predicted metabolite, and ran the entire analysis pipeline 10 times, averaging the results over these runs.

Discussion

Understanding how the gut microbiome shapes the gut metabolome is without a doubt crucial for any investigation of microbiota-related mechanisms affecting the host’s health and of microbiome-based therapy. Yet, the interaction between gut microbes and metabolites remains largely uncharacterized for various technical and methodological reasons [3, 13]. Microbiome-metabolome studies of the human gut aim to characterize these interactions using a data-driven approach, but the generalizability of reported associations is unclear given the substantial differences between studies and cohorts.

Here we perform a first large-scale meta-analysis of paired fecal microbiome-metabolome datasets. We specifically evaluated the robustness of human gut microbiome-metabolite associations in 9 datasets of healthy individuals totaling 779 samples. We implemented a bioinformatic pipeline for processing these paired microbiome-metabolome datasets, used machine learning to predict metabolite levels in each dataset based on microbiome composition, and leveraged classic meta-analysis techniques to identify metabolites that are consistently well-predicted by the microbiome. Overall we found 97 such “robustly well-predicted” metabolites, spanning several known microbiome-related metabolic pathways such as bile acids [55], tryptophan metabolites [80], polyamines [81], and polyunsaturated fatty acids [82]. Several other metabolites, such as certain MCFAs

and dicarboxylic acids, however, were barely reported in the context of the gut microbiota to date, and may suggest interesting and clinically relevant future research directions.

We further analyzed how similar are metabolite models across different datasets, both by comparing significant microbiota contributors in each model and by examining cross-predictability. We found that for some metabolites models were highly similar, yet for others, models exhibited low similarity across datasets and poor transferability. Lastly, we performed a similar analysis comparing the healthy datasets to 7 additional disease datasets and found that the overall predictability of most metabolites remained similar in disease.

Notably, the datasets included in this meta-analysis differ from one another in multiple aspects, both technical and biological. Technical factors include sample handling, sequencing/metabolomics methods, and bioinformatic methods for processing raw data, all of which were previously shown to contribute to artefactual differences in detected species or metabolites [2, 50, 51, 83, 84]. Metabolome profiling is particularly sensitive to technical variation in the profiling procedure, where the number of quantified metabolites can range from a few dozens to tens of thousands depending on whether a targeted or untargeted approach was taken and the exact instrumental setup used [2]. In our meta-analysis, the biases of each processing pipeline are potentially reflected in microbes and metabolites missing from the data, mislabeled, or in skewed abundance values. This sort of “noise”, however, is expected to mostly conceal associations, rather than to produce false ones. The second class of inter-study differences, i.e., those that stem from biological factors, include cohort characteristics such as age, geography, gender, medical background, and diet, all of which potentially introduce additional variation between datasets both in microbiome and metabolome compositions, and in microbiome-metabolome-host interactions [28, 51, 53, 85, 86]. Still, the goal of our study was to quantify variation in microbiome-metabolite associations that arise from all types of heterogeneity, and accordingly, to highlight those specific associations that seem to be consistent over all such heterogeneity-introducing factors. With the collection and publication of many additional microbiome-metabolome datasets, stratified analyses efforts could potentially reveal more robust associations while controlling for a specific factor of interest.

Beyond the findings reported in this study, our work calls attention to several important topics concerning microbiome-metabolome studies. First, the difficulties in processing and converting the different datasets into a unified format highlights a challenging first hurdle in any attempt to generalize findings across microbiome-metabolome studies. These difficulties arise both from

the large variation in data formats, especially for metabolomics data, and from the substantial differences in detected microbiome/metabolome features in each study as discussed above. Indeed, previous meta-analyses studies have discussed these challenges in the context of microbiome or metabolome independently [47, 87–89] and non-surprisingly these challenges further exacerbate here in this multi-omic meta-analysis. Improved standardization in collecting, processing, cataloging, and storing both microbiome and metabolome data (and maybe even specific standards for microbiome-metabolome datasets) is therefore key for future progress in this field.

We also note that in the context of case-control studies, where researchers attempt to detect microbiome-metabolite links characteristic of a specific disease, there may be a few different types of potential disease-relevant patterns. First, a metabolite may be strongly linked to the overall microbial composition (e.g., well-predicted) in both the case and the control study groups, but enriched or depleted in one of the groups. Such a pattern (especially if this microbiome-metabolite link is found also in many other studies) may suggest that this specific association reflects a part of the basic “house-keeping” metabolic machinery of the microbiome and that the shift in the metabolite level in disease may be attributed to a dysbiotic microbiome. Second, a metabolite may be well-predicted by the microbiome in both groups but with different taxa contributors in each group, suggesting again an altered composition or an altered metabolic activity of community members in disease. Lastly, a metabolite may be well-predicted in one group (e.g., control), but not the other. In this case, that metabolite’s level in individuals with the disease is potentially controlled more by other disease-related factors. Overall, this suggests that the robustness of microbiome-metabolite links, both across studies and across study groups, should be taken into consideration when interpreting disease-associated shifts and calls for a better theory and methodology for distinguishing between clinically relevant and non-relevant cross-omic interactions.

Most microbiome-metabolome studies conduct some sort of an association-based analysis, using either correlations or simple linear regressions to identify specific microbes whose abundances across samples strongly correlate with the concentration of a specific metabolite. A smaller subset of such studies further apply machine learning techniques for this purpose, and new machine learning-based tools have been recently introduced [20, 33, 34, 41, 90, 91]. In this work, we intentionally opted for this less common, machine learning-based approach, treating the ability to predict a metabolite level based on the composition of the microbiome as a quantifiable proxy for metabolite-microbiota associations. The motivation for this choice was two-fold: first, machine learning enables to detect non-linear and complex

associations that may involve multiple different factors and that may be missed in a simpler, correlation-based approach; second, by considering the entire microbial community, rather than a specific taxon abundance, we were able to identify metabolites consistently associated with the microbiota (or a sub-population of it) as a whole while overcoming the substantial differences between studies in terms of taxonomic resolution and accuracy that potentially prevent meaningful taxon-specific comparisons. Still, as both machine learning and univariate methods ultimately capture statistical dependencies in the data, we believe that many of the insights reported in our study are relevant for other statistical methods such as correlation.

Importantly, though, high predictability of a metabolite does not necessarily indicate a direct mechanistic interaction between the microbes that contribute toward this prediction and the metabolite, just as a microbe-metabolite correlation does not necessarily indicate a direct relation of consumption or production [36]. Nonetheless, these statistical associations reveal potentially intriguing structures and dependencies in the data and so their robustness and generalizability are key. We hypothesize that the metabolites found here as robustly well-predicted are strongly tied to the gut microbiota composition even in the presence of other influencing factors, and hence that changes in the microbial composition will most likely cause changes in metabolite levels, making such metabolites a prime target for microbiome-based interventions.

Our analysis clearly faced several important caveats. One major challenge of this study and of the need to unify data from multiple different studies is that we had to limit our analysis to metabolites with a shared annotation (here, HMDB identifiers) and that are present in multiple studies (using potentially different metabolomic approaches). With these limitations, we effectively focused on 273 metabolites out of approximately 97,000 (non-unique) raw metabolites identified in the 9 healthy datasets. These 273 metabolites constitute only a tiny fraction of gut metabolites, as most gut metabolites, and specifically those of microbial origin, are largely unidentified and poorly represented in metabolic databases [3, 17, 92]. Moreover, though for the main part of the analysis, we did not directly compare genus-level statistics between cohorts, a similar limitation is applicable for the microbiome data. Collapsing the taxonomic data into genus level profiles somewhat mitigated this problem, but also constrained our ability to capture associations at the species and strain level [93–95]. Our observation that some genera are consistently associated with metabolites while others demonstrate inconsistent patterns, for example, may be affected by the extent of metabolic diversification within each genera, which is known to vary

dramatically [96]. This meta-analysis was also clearly limited by the number of cohorts included. Future work including significantly more cohorts could not only substantially expand the pool of analyzed metabolites but could also better account for technical and biological differences between studies, as noted above. Furthermore, expanding the analysis to multiple “disease” cohorts, perhaps even of the same disease, could uncover additional insights related to consistent shifts in metabolic activity of dysbiotic microbiota. This calls for both additional high-quality microbiome-metabolome studies and improved availability and standardization of such datasets.

Conclusions

The network of host-microbiota-metabolome interactions in the human gut is extremely complex and requires multiple research efforts in multiple complementary directions in order to be fully deciphered and characterized. Here, we identified metabolites consistently associated with the microbiota across diverse studies and cohorts, thus distinguishing between study-specific and robust or universal links. Our findings provide a better understanding of microbiome-metabolome interactions and allow researchers to put newly identified microbially associated metabolites within the context of other studies.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-021-01149-z>.

Additional file 1: Supplementary Table S1. Description of studies included in the analysis (expansion of Table 1). **Table S2.** 16S/WGSS data processing notes per dataset. **Table S3.** Unified genera processing notes. **Table S4.** Metabolomics processing notes per dataset. **Table S5.** Metabolite predictability per dataset. **Table S6.** Random-effects models results. **Table S7.** Genus feature contributions by permutation analysis. **Table S8.** Consistent genera contributors. **Table S9.** Subgroup meta-analysis results.

Additional file 2: Supplementary Note 1. Alternative machine learning pipelines and effect on results. **Supplementary Note 2.** Comparisons to previous studies and validation. **Supplementary Note 3.** Metabolic pathways-oriented analysis of robustly well-predicted metabolites. **Supplementary Note 4.** Effect of excluding the infants dataset on robustness results.

Additional file 3: Figure S1. Statistics about genera and metabolite features shared among datasets included in this meta-analysis. **Figure S2.** Comparisons to previous studies. **Figure S3.** Characteristics of robustly well-predicted metabolites. **Figure S4.** Comparisons of metabolite models between datasets. **Figure S5.** Examples of disease-independent robustly well-predicted metabolites. **Figure S6.** Transferability of metabolite models from controls to cases within the same study. **Figure S7.** Examples of significant contributors' comparison between healthy and disease metabolite models. **Figure S8.** Comparison of different machine learning pipelines.

Acknowledgements

We would like to thank past and present Borenstein lab members for their helpful feedback and assistance throughout the research process, in particular Alex Eng, Cecilia Noecker, Yotam Cohen, Or Segal, and Ran

Armoni. We thank all authors of microbiome-metabolome studies who made their data publicly available thus enabling this meta-analysis work, and additional authors who kindly shared their data with us upon request, including Nicholas Chia, Matthew Bockol, Patricio Jeraldo, Carolyn Slupsky, Xuan He, Jonathan Jacobs, Pande Putu Erawijantari, Takuji Yamada, and Shinichi Yachida. We thank Rachel Gregor and Shoval Lagziel for their tips regarding metabolomics data, and Daniel Yekutieli, Saharon Rosset, Barak Brill, and Iman Jaljuli for their helpful inputs regarding statistical aspects of the work. We also thank the Tel Aviv University Computer Science IT team for their technical support during computational work, and Ory Pinchasy for the illustration in Fig. 3C.

Authors' contributions

EM and EB designed the research. EM identified, obtained, and processed the datasets, and performed all computational analyses. EM and YA interpreted the results. EM and EB drafted the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported in part by NIH Grants R01DK095869 and U19AG057377 and ISF Grant 2435/19. EM was supported in part by a fellowship from the Edmond J. Safra Center for Bioinformatics at Tel-Aviv University. EB is a Faculty Fellow of the Edmond J. Safra Center for Bioinformatics at Tel Aviv University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

Raw or processed sequencing and metabolomics data for each study included in the analysis can be accessed as described in Additional file 1: Table S1. Detailed predictability results, meta-analysis random-effects model results, and additional analysis results reported throughout the paper are available in the supplementary files. R scripts used for analysis are available on <https://github.com/borenstein-lab>.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author details

¹The Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel. ²Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. ³Santa Fe Institute, Santa Fe, NM, USA.

Received: 24 January 2021 Accepted: 18 August 2021

Published online: 12 October 2021

References

- Sharon G, Garg N, Debelius J, Knight R, Dorrestein PC, Mazmanian SK. Specialized metabolites from the microbiome in health and disease. *Cell Metab.* 2014;20(5):719–30. <https://doi.org/10.1016/j.cmet.2014.10.016>.
- Smirnov KS, Maier TV, Walker A, Heinzmann SS, Forcisi S, Martinez I, et al. Challenges of metabolomics in human gut microbiota research. *Int J Med Microbiol.* 2016;306(5):266–79. <https://doi.org/10.1016/j.ijmm.2016.03.006>.
- Dorrestein PC, Mazmanian SK, Knight R. Finding the missing links among metabolites, microbes, and the host. *Immunity.* 2014;40(6):824–32. <https://doi.org/10.1016/j.immuni.2014.05.015>.
- Cryan JF, Dinan TG. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat Rev Neurosci.* 2012;13(10):701–12. <https://doi.org/10.1038/nrn3346>.
- Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med.* 2013;19(5):576–85. <https://doi.org/10.1038/nm.3145>.
- Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature.* 2011;472(7341):57–65. <https://doi.org/10.1038/nature09922>.
- Koh A, Molinaro A, Ståhlman M, Khan MT, Schmidt C, Mannerås-Holm L, et al. Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell.* 2018;175:947–961.e17.
- Blacher E, Bashiardes S, Shapiro H, Rothschild D, Mor U, Dori-Bachash M, et al. Potential roles of gut microbiome and metabolites in modulating ALS in mice. *Nature.* 2019;572(7770):474–80. <https://doi.org/10.1038/s41586-019-1443-5>.
- Smith PM, Howitt MR, Panikov N, Michaud M, Gallini CA, Bohlooly-Y M, et al. The microbial metabolites, short-chain fatty acids, regulate colonic T reg cell homeostasis. *Science.* 2013;341(6145):569–73. <https://doi.org/10.1126/science.1241165>.
- Steed AL, Christophi GP, Kaiko GE, Sun L, Goodwin VM, Jain U, et al. The microbial metabolite desaminotyrosine protects from influenza through type I interferon. *Science.* 2017;357(6350):498–502. <https://doi.org/10.1126/science.aam5336>.
- Valles-Colomer M, Falony G, Darzi Y, Tigchelaar EF, Wang J, Tito RY, et al. The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat Microbiol.* 2019;4(4):623–32. <https://doi.org/10.1038/s41564-018-0337-x>.
- Wong AC, Levy M. New Approaches to Microbiome-Based Therapies. *mSystems.* 2019;4(3):e00122–19. <https://doi.org/10.1128/mSystems.00122-19>.
- Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, et al. Host-gut microbiota metabolic interactions. *Science.* 2012;336(6086):1262–7. <https://doi.org/10.1126/science.1223813>.
- Summer MOA. Advancing gut microbiome research using cultivation. *Curr Opin Microbiol.* 2015;27:127–32. <https://doi.org/10.1016/j.mib.2015.08.004>.
- Marcobal A, Kashyap PC, Nelson TA, Aronov PA, Donia MS, Spormann A, et al. A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. *ISME J.* 2013;7(10):1933–43. <https://doi.org/10.1038/ismej.2013.89>.
- Magnúsdóttir S, Thiele I. Modeling metabolism of the human gut microbiome. *Curr Opin Biotechnol.* 2018;51:90–6. <https://doi.org/10.1016/j.copbio.2017.12.005>.
- Van Treuren W, Dodd D. Microbial Contribution to the Human Metabolome: Implications for Health and Disease. *Annu Rev Pathol Mech Dis.* 2020;15(1):345–69. <https://doi.org/10.1146/annurev-pathol-020117-043559>.
- McHardy IH, Goudarzi M, Tong M, Ruegger PM, Schwager E, Weger JR, et al. Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. *Microbiome.* 2013;1(1):17. <https://doi.org/10.1186/2049-2618-1-17>.
- Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature.* 2019;569(7758):655–62. <https://doi.org/10.1038/s41586-019-1237-9>.
- Zierer J, Jackson MA, Kastenmüller G, Mangino M, Long T, Telenti A, et al. The fecal metabolome as a functional readout of the gut microbiome. *Nat Genet.* 2018;50(6):790–5. <https://doi.org/10.1038/s41588-018-0135-7>.
- Jacobs JP, Goudarzi M, Singh N, Tong M, McHardy IH, Ruegger P, et al. A Disease-Associated Microbial and Metabolomics State in Relatives of Pediatric Inflammatory Bowel Disease Patients. *Cell Mol Gastroenterol Hepatol.* 2016;2(6):750–66. <https://doi.org/10.1016/j.jcmgh.2016.06.004>.
- Franzosa EA, Sirota-Madi A, Avila-Pacheco J, Fornoel N, Haiser HJ, Reinker S, et al. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol.* 2019;4(2):293–305. <https://doi.org/10.1038/s41586-018-0306-4>.
- Visconti A, Le Roy CI, Rosa F, Rossi N, Martin TC, Mohny RP, et al. Interplay between the human gut microbiome and host metabolism. *Nat Commun.* 2019;10(1):4505. <https://doi.org/10.1038/s41467-019-12476-z>.
- Wilmanski T, Rappaport N, Earls JC, Magis AT, Manor O, Lovejoy J, et al. Blood metabolome predicts gut microbiome α -diversity in humans. *Nat Biotechnol.* 2019;37(10):1217–28. <https://doi.org/10.1038/s41587-019-0233-9>.
- Thomas AM, Manghi P, Asnicar F, Pasolli E, Armanini F, Zolfo M, et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat Med.* 2019;25(4):667–78. <https://doi.org/10.1038/s41591-019-0405-7>.
- Poyet M, Groussin M, Gibbons SM, Avila-Pacheco J, Jiang X, Kearney SM, et al. A library of human gut bacterial isolates paired with longitudinal

- multiomics data enables mechanistic microbiome research. *Nat Med.* 2019; 25(9):1442–52. <https://doi.org/10.1038/s41591-019-0559-3>.
27. Kim M, Vogtmann E, Ahlquist DA, Devens ME, Kisiel JB, Taylor WR, et al. Fecal metabolomic signatures in colorectal adenoma patients are associated with gut microbiota and early events of colorectal cancer pathogenesis. *MBio.* 2020;11(1). <https://doi.org/10.1128/mBio.03186-19>.
 28. Tang Z-Z, Chen G, Hong Q, Huang S, Smith HM, Shah RD, et al. Multi-omic analysis of the microbiome and metabolome in healthy subjects reveals microbiome-dependent relationships between diet and metabolites. *Front Genet.* 2019;10:454. <https://doi.org/10.3389/fgene.2019.00454>.
 29. Ursell LK, Haiser HJ, Van Treuren W, Garg N, Reddivari L, Vanamala J, et al. The intestinal metabolome: an intersection between microbiota and host. *Gastroenterology.* 2014;146(6):1470–6. <https://doi.org/10.1053/j.gastro.2014.03.001>.
 30. Chong J, Xia J. Computational approaches for integrative analysis of the metabolome and microbiome. *Metabolites.* 2017;7(4):62. <https://doi.org/10.3390/metabo7040062>.
 31. Kostic AD, Gevers D, Siljander H, Vatanen T, Hyötyläinen T, Hämäläinen A-M, et al. The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. *Cell Host Microbe.* 2015;17(2): 260–73. <https://doi.org/10.1016/j.chom.2015.01.001>.
 32. Lamichhane S, Sen P, Dickens AM, Orešič M, Bertram HC. Gut metabolome meets microbiome: a methodological perspective to understand the relationship between host and microbe. *Methods.* 2018;149:3–12. <https://doi.org/10.1016/j.ymeth.2018.04.029>.
 33. Mallick H, Franzosa EA, McIver LJ, Banerjee S, Sirota-Madi A, Kostic AD, et al. Predictive metabolomic profiling of microbial communities using amplicon or metagenomic sequences. *Nat Commun.* 2019;10(1):3136. <https://doi.org/10.1038/s41467-019-10927-1>.
 34. Morton JT, Aksenov AA, Nothias LF, Foulds JR, Quinn RA, Badri MH, et al. Learning representations of microbe–metabolite interactions. *Nat Methods.* 2019;16(12):1306–14. <https://doi.org/10.1038/s41592-019-0616-3>.
 35. Bar N, Korem T, Weissbrod O, Zeevi D, Rothschild D, Leviatan S, et al. A reference map of potential determinants for the human serum metabolome. *Nature.* 2020;588:135–40. <https://doi.org/10.1038/s41586-020-2896-2>.
 36. Noecker C, Chiu HC, McNally CP, Borenstein E. Defining and Evaluating Microbial Contributions to Metabolite Variation in Microbiome-Metabolome Association Studies. *mSystems.* 2019;4(6):1–28. <https://doi.org/10.1128/mSystems.00579-19>.
 37. Simons K. How Can Omic Science be Improved? *Proteomics.* 2018;18(5-6): 1800039. <https://doi.org/10.1002/prot.201800039>.
 38. Schloss PD. Identifying and overcoming threats to reproducibility, replicability, robustness, and generalizability in microbiome research. *MBio.* 2018;9(3):e00525-18. <https://doi.org/10.1128/mBio.00525-18>.
 39. Yachida S, Mizutani S, Shiroma H, Shiba S, Nakajima T, Sakamoto T, et al. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat Med.* 2019;25(6): 968–76. <https://doi.org/10.1038/s41591-019-0458-7>.
 40. He X, Parenti M, Grip T, Lönnerdal B, Timby N, Domellöf M, et al. Fecal microbiome and metabolome of infants fed bovine MFGM supplemented formula or standard formula with breast-fed infants as reference: a randomized controlled trial. *Sci Rep.* 2019;9(1):11589. <https://doi.org/10.1038/s41598-019-47953-4>.
 41. Mars RAT, Yang Y, Ward T, Houtti M, Priya S, Lekatz HR, et al. Longitudinal Multi-omics Reveals Subset-Specific Mechanisms Underlying Irritable Bowel Syndrome. *Cell.* 2020;182:1460–1473.e17. <https://doi.org/10.1016/j.cell.2020.08.007>.
 42. Sinha R, Ahn J, Sampson JN, Shi J, Yu G, Xiong X, et al. Fecal microbiota, fecal metabolome, and colorectal cancer interrelations. *PLoS One.* 2016; 11(3):e0152126. <https://doi.org/10.1371/journal.pone.0152126>.
 43. Erawijantari PP, Mizutani S, Shiroma H, Shiba S, Nakajima T, Sakamoto T, et al. Influence of gastrectomy for gastric cancer treatment on faecal microbiome and metabolome profiles. *Gut.* 2020;69(8):1404–15. <https://doi.org/10.1136/gutjnl-2019-319188>.
 44. Wang Z, Yang Y, Yan Z, Liu H, Chen B, Liang Z, et al. Multi-omic meta-analysis identifies functional signatures of airway microbiome in chronic obstructive pulmonary disease. *ISME J.* 2020:1–18. <https://doi.org/10.1038/s41396-020-0727-y>.
 45. Tian J, Xue W, Yin H, Zhang N, Zhou J, Long Z, et al. Differential metabolic alterations and biomarkers between gastric cancer and colorectal cancer: A systematic review and meta-analysis. *Onco Targets Ther.* 2020;13:6093–108. <https://doi.org/10.2147/OTT.S247393>.
 46. Jiang Y, Zhu Z, Shi J, An Y, Zhang K, Wang Y, et al. Metabolomics in the development and progression of dementia: A systematic review. *Frontiers in Neuroscience.* 2019;13(APR):343. <https://doi.org/10.3389/fnins.2019.00343>.
 47. Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat Commun.* 2017;8(1):1784. <https://doi.org/10.1038/s41467-017-01973-8>.
 48. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol.* 2019;37(8):852–7. <https://doi.org/10.1038/s41587-019-0209-9>.
 49. David S, Wishart, Yannick Djoumbou Feunang, Ana Marcu, An Chi Guo, Kevin Liang, Rosa Vázquez-Fresno, et al. HMDB 4.0: the human metabolome database for 2018 | *Nucleic Acids Research* | Oxford Academic. *Nucleic Acids Res.* 2018;46:D608–D617. <https://academic.oup.com/nar/article/46/D1/D608/4616873>. Accessed 29 Jun 2020.
 50. Jovel J, Patterson J, Wang W, Hotte N, O'Keefe S, Mitchell T, et al. Characterization of the gut microbiome using 16S or shotgun metagenomics. *Front Microbiol.* 2016;7(APR):459. <https://doi.org/10.3389/fmicb.2016.00459>.
 51. Debelius J, Song SJ, Vazquez-Baeza Y, Xu ZZ, Gonzalez A, Knight R. Tiny microbes, enormous impacts: what matters in gut microbiome studies? *Genome Biol.* 2016;17(1):217. <https://doi.org/10.1186/s13059-016-1086-x>.
 52. Zou Y, Xue W, Luo G, Deng Z, Qin P, Guo R, et al. 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat Biotechnol.* 2019;37(2):179–85. <https://doi.org/10.1038/s41587-018-0008-8>.
 53. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. *Nature.* 2012;486(7402):222–7. <https://doi.org/10.1038/nature11053>.
 54. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis*; 2009. www.wiley.com. Accessed 16 Jul 2020. <https://doi.org/10.1002/9780470743386>.
 55. Molinero N, Ruiz L, Sánchez B, Margolles A, Delgado S. Intestinal bacteria interplay with bile and cholesterol metabolism: implications on host physiology. *Front Physiol.* 2019;10(MAR):185. <https://doi.org/10.3389/fphys.2019.00185>.
 56. Schönfeld P, Wojtczak L. Short- and medium-chain fatty acids in energy metabolism: The cellular perspective. *J Lipid Res.* 2016;57(6):943–54. <https://doi.org/10.1194/jlr.R067629>.
 57. Vujkovic-Cvijin I, Sklar J, Jiang L, Natarajan L, Knight R, Belkaid Y. Host variables confound gut microbiota studies of human disease. *Nature.* 2020; 587(7834):448–54. <https://doi.org/10.1038/s41586-020-2881-9>.
 58. Mingrone G, Castagneto-Gissey L, Macé K. Use of dicarboxylic acids in type 2 diabetes. *Br J Clin Pharmacol.* 2013;75(3):671–6. <https://doi.org/10.1111/j.1365-2125.2012.04177.x>.
 59. Nagao K, Yanagita T. Medium-chain fatty acids: Functional lipids for the prevention and treatment of the metabolic syndrome. *Pharmacol Res.* 2010; 61(3):208–12. <https://doi.org/10.1016/j.phrs.2009.11.007>.
 60. Kriaa A, Bourgin M, Potiron A, Mkaouer H, Jablaoui A, Gérard P, et al. Microbial impact on cholesterol and bile acid metabolism: current status and future prospects. *J Lipid Res.* 2019;60(2):323–32. <https://doi.org/10.1194/jlr.R088989>.
 61. Lei M, Menon R, Manteiga S, Alden N, Hunt C, Alaniz RC, et al. Environmental chemical diethylhexyl phthalate alters intestinal microbiota community structure and metabolite profile in mice. *mSystems.* 2019;4. <https://doi.org/10.1128/mSystems.00724-19>.
 62. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30. <https://www.ncbi.nlm.nih.gov/pubmed/10592173>. Accessed 11 Mar 2019. <https://doi.org/10.1093/nar/28.1.27>.
 63. Gonzalez A, Navas-Molina JA, Kosciulek T, McDonald D, Vázquez-Baeza Y, Ackermann G, et al. Qiita: rapid, web-enabled microbiome meta-analysis. *Nat Methods.* 2018;15(10):796–8. <https://doi.org/10.1038/s41592-018-0141-9>.
 64. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, et al. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis | *Nucleic Acids Research* | Oxford Academic. *Nucleic Acids Res.* 2018;46:W486–94. <https://academic.oup.com/nar/article/46/W1/W486/4995686>. Accessed 4 Aug 2020. <https://doi.org/10.1093/nar/gky310>.
 65. De Livera AM, Olshansky M, Speed TP. Statistical analysis of metabolomics data. *Methods Mol Biol.* 2013;1055:291–307. https://doi.org/10.1007/978-1-62703-577-4_20.

66. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13(7):581–3. <https://doi.org/10.1038/nmeth.3869>.
67. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*. 2018;6(1):90. <https://doi.org/10.1186/s40168-018-0470-z>.
68. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res*. 2007;35:7188–96. <https://academic.oup.com/nar/article/35/21/7188/2376260>. Accessed 6 Aug 2020. <https://doi.org/10.1093/nar/gkm864>.
69. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nature Methods*. 2015;12(10):902–3. <https://doi.org/10.1038/nmeth.3589>.
70. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. New York; 2016. p. 785–94. <https://doi.org/10.1145/2939672.2939785>.
71. Kuhn M, Wickham H. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. 2020. <https://www.tidymodels.org>.
72. Wright MN, Ziegler A. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J Stat Softw*. 2015;77. <https://doi.org/10.18637/jss.v077.i01>.
73. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33:1–22. <https://doi.org/10.18637/jss.v033.i01>.
74. Schwarzer G. meta: An R package for meta-analysis. *R news*. 2007;7:40–5. https://cran.rstudio.org/doc/Rnews/Rnews_2007-3.pdf#page=40.
75. Viechtbauer W. Conducting meta-analyses in R with the metafor Package. *J Stat Softw*. 2010;36:1–48. <https://doi.org/10.18637/jss.v036.i03>.
76. Altmann A, Tološi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. *Bioinformatics*. 2010;26(10):1340–7. <https://doi.org/10.1093/bioinformatics/btq134>.
77. Pasolli E, Truong DT, Malik F, Waldron L, Segata N. Machine learning meta-analysis of large metagenomic datasets: tools and biological insights. *PLOS Comput Biol*. 2016;12(7):e1004977. <https://doi.org/10.1371/journal.pcbi.1004977>.
78. Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat Med*. 2019;25(4):679–89. <https://doi.org/10.1038/s41591-019-0406-6>.
79. Harrer M, Cuijpers P, Furukawa T, Ebert D. Doing meta-analysis in R: a hands-on guide; 2019. <https://doi.org/10.5281/zenodo.2551802>.
80. Ma N, Ma X. Dietary amino acids and the gut-microbiome-immune axis: physiological metabolism and therapeutic prospects. *Compr Rev Food Sci Food Saf*. 2019;18(1):221–42. <https://doi.org/10.1111/1541-4337.12401>.
81. Tofalo R, Cocchi S, Suzzi G. Polyamines and gut microbiota. *Front Nutr*. 2019;6:16. <https://doi.org/10.3389/fnut.2019.00016>.
82. Kishino S, Takeuchi M, Park SB, Hirata A, Kitamura N, Kunisawa J, et al. Polyunsaturated fatty acid saturation by gut lactic acid bacteria affecting host lipid composition. *Proc Natl Acad Sci U S A*. 2013;110(44):17808–13. <https://doi.org/10.1073/pnas.1312937110>.
83. Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, et al. Best practices for analysing microbiomes. *Nat Rev Microbiol*. 2018;16(7):410–22. <https://doi.org/10.1038/s41579-018-0029-9>.
84. Liang Y, Dong T, Chen M, He L, Wang T, Liu X, et al. Systematic analysis of impact of sampling regions and storage methods on fecal gut microbiome and metabolome profiles. *mSphere*. 2020;5(1):1–13. <https://doi.org/10.1128/mSphere.00763-19>.
85. Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, et al. Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell*. 2019;176:649–662.e20. <https://doi.org/10.1016/j.cell.2019.01.001>.
86. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. *Science*. 2016;352(6285):560–4. <https://doi.org/10.1126/science.aad3503>.
87. Bisanz JE, Upadhyay V, Turnbaugh JA, Ly K, Turnbaugh PJ. Meta-analysis reveals reproducible gut microbiome alterations in response to a high-fat diet. *Cell Host Microbe*. 2019;26:265–272.e4.
88. Rocca JD, Simonin M, Blaszczyk JR, Ernakovich JG, Gibbons SM, Midani FS, et al. The microbiome stress project: toward a global meta-analysis of environmental stressors and their effects on microbial communities. *Front Microbiol*. 2019;9. <https://doi.org/10.3389/fmicb.2018.03272>.
89. Mak TD, Goudarzi M, Laiakis EC, Stein SE. Disparate metabolomics data reassembler: a novel algorithm for agglomerating incongruent LC-MS metabolomics datasets. *Anal Chem*. 2020;92(7):5231–9. <https://doi.org/10.1021/acs.analchem.9b05763>.
90. Taylor BC, Lejzerowicz F, Poirer M, Shaffer JP, Jiang L, Aksenov A, et al. Consumption of fermented foods is associated with systematic differences in the gut microbiome and metabolome. *mSystems*. 2020;5(2):1–16. <https://doi.org/10.1128/mSystems.00901-19>.
91. Raghuvanshi R, Vasco K, Vázquez-Baeza Y, Jiang L, Morton JT, Li D, et al. High-resolution longitudinal dynamics of the cystic fibrosis sputum microbiome and metabolome through antibiotic therapy. *mSystems*. 2020;5(3):1–15. <https://doi.org/10.1128/mSystems.00292-20>.
92. Melnik AV, da Silva RR, Hyde ER, Aksenov AA, Vargas F, Bouslimani A, et al. Coupling targeted and untargeted mass spectrometry for metabolome-microbiome-wide association studies of human fecal samples. *Anal Chem*. 2017;89(14):7549–59. <https://doi.org/10.1021/acs.analchem.7b01381>.
93. Zhao S, Lieberman TD, Poyet M, Kauffman KM, Gibbons SM, Groussin M, et al. Adaptive evolution within gut microbiomes of healthy people. *Cell Host Microbe*. 2019;25:656–667.e8.
94. Barboza M, Sela DA, Pirim C, LoCascio RG, Freeman SL, German JB, et al. Glycoprofiling bifidobacterial consumption of galacto-oligosaccharides by mass spectrometry reveals strain-specific, preferential consumption of glycans. *Appl Environ Microbiol*. 2009;75(23):7319–25. <https://doi.org/10.1128/AEM.00842-09>.
95. Haider HJ, Gootenberg DB, Chatman K, Sirasani G, Balskus EP, Turnbaugh PJ. Predicting and manipulating cardiac drug inactivation by the human gut bacterium *Eggerthella lenta*. *Science*. 2013;341(July):295–9.
96. Vieira-Silva S, Falony G, Darzi Y, Lima-Mendez G, Garcia Yunta R, Okuda S, et al. Species-function relationships shape ecological properties of the human gut microbiome. *Nat Microbiol*. 2016;1(8):1–8. <https://doi.org/10.1038/nmicrobiol.2016.88>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

