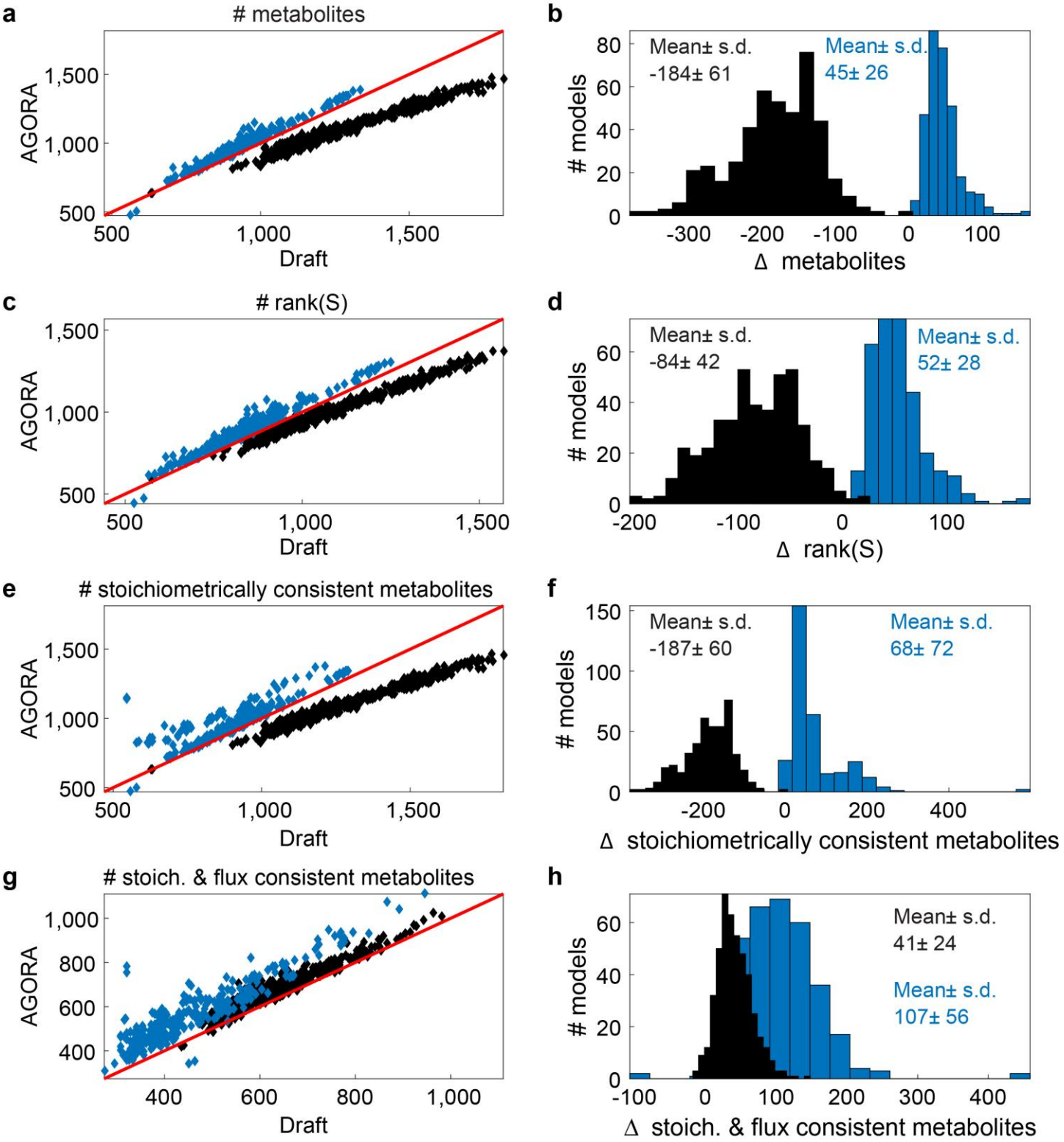


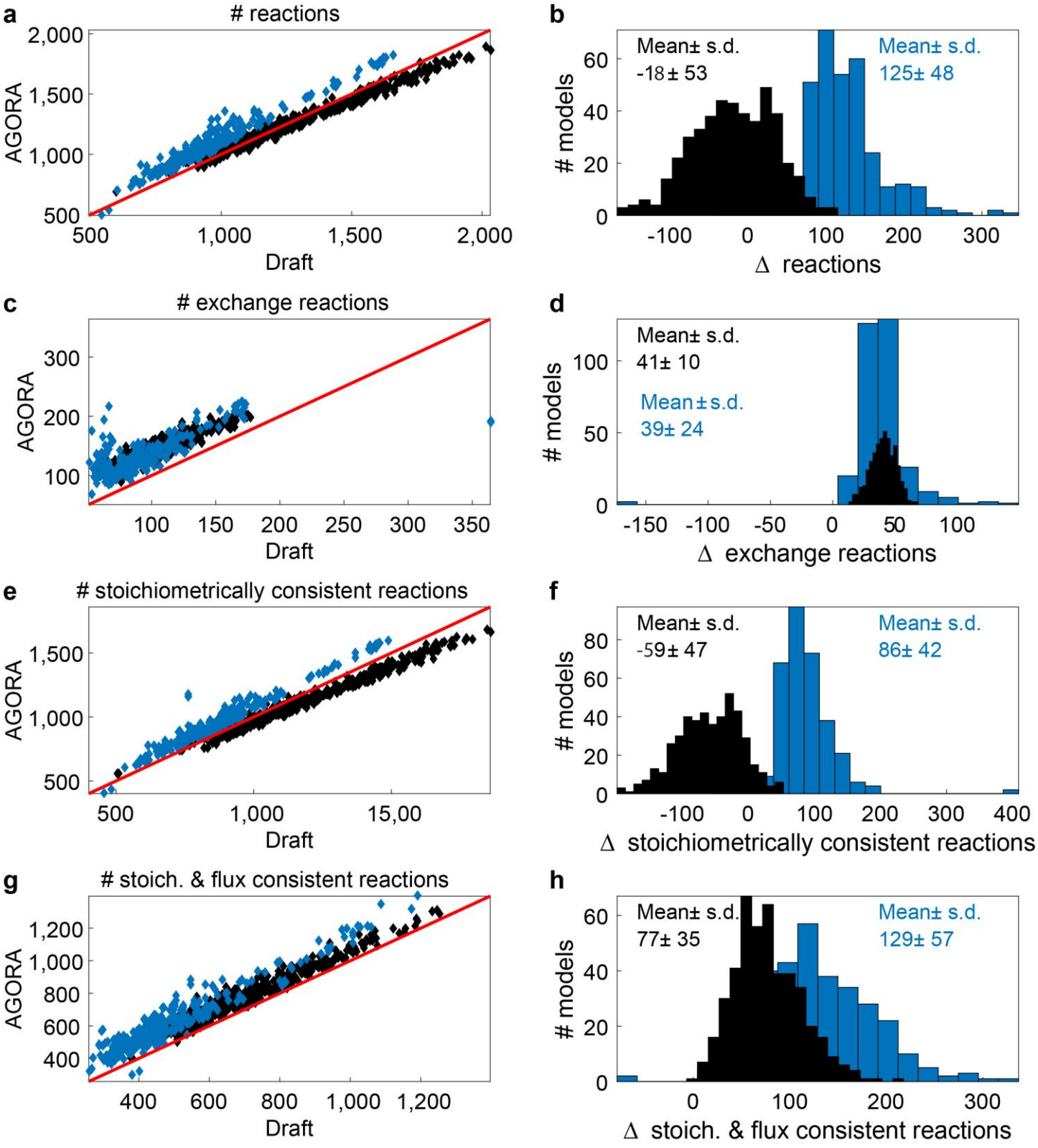
Supplementary Figure 1

Comparison of metabolite stoichiometric and flux consistency of draft and AGORA reconstructions.

Comparison of metabolite stoichiometric and flux consistency (Fleming, et al., J. Theor. Biol., 2016) of draft and AGORA

reconstructions (Supplementary Table 3). The figures highlight two groups in the draft reconstruction set. We identified the time of download as the key separating factor. Those reconstructions that have been obtained from Model SEED before summer 2015 had a smaller reaction (and metabolite) content, than those ones downloaded from model SEED or KBase afterwards. Model SEED/KBase updated and expanded their underlying database substantially in 2015 to include secondary metabolite, xenobiotics, and plant metabolism. Consequently, the average size of the draft reconstructions increased. These metabolic pathways were outside the scope of the current reconstruction effort and thus were not retained in AGORA reconstructions of the black group. Further experimental data and comparative genomic efforts will be required to establish that those out-of-scope reactions do indeed occur in the respective gut microbes. All reconstructions of the black group thus fall below the red line, meaning that the AGORA counterpart is smaller in terms of metabolites. *However, it is notable and thanks to the QC/QA effort applied to all AGORA reconstructions that the there is no observable difference between the two groups when comparing the stoichiometric and flux consistent metabolites. In all cases, we improved the quality of the reconstructions when considering this measure over the draft reconstructions.*

(a) The number of metabolites in each AGORA reconstruction versus the corresponding draft reconstruction. The red line shows the line y=x, where the number of metabolites in the AGORA reconstruction is the same as the number of metabolites in the draft reconstruction. (b) Histograms showing the change in number of metabolites after the curation of the draft reconstructions of the two groups. The mean and standard deviation of the change in number of metabolites is shown for both groups. A similar separation was observed for (c-d) the rank of the draft and AGORA stoichiometric matrices and (e-f) the number of stoichiometrically consistent metabolites in AGORA versus the draft reconstructions. (g-h) In most organisms of both groups, the number of stoichiometrically and flux consistent metabolites was increased by the AGORA curation process.

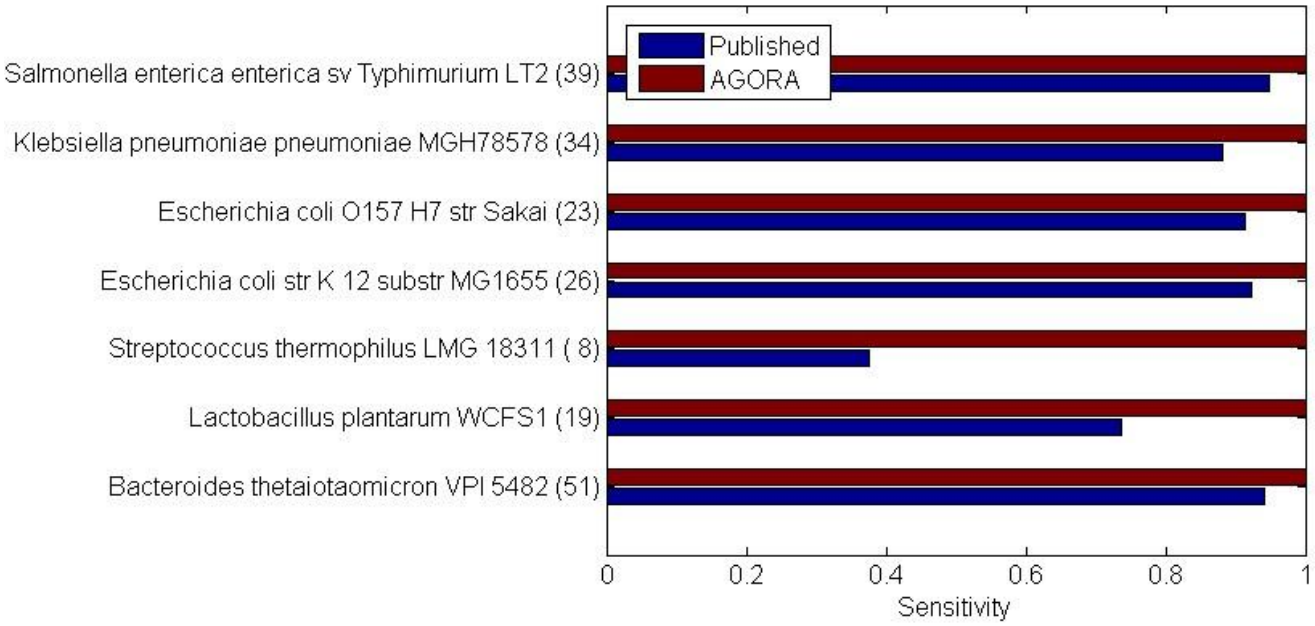


Supplementary Figure 2

Comparison of reaction stoichiometric and flux consistency of draft and AGORA reconstructions.

Comparison of reaction stoichiometric and flux consistency (Fleming, et al., J. Theor. Biol., 2016) of draft and AGORA reconstructions (Supplementary Table 3). The figures highlight two groups in the draft reconstruction set. We identified the time of download as the key separating factor. Those reconstructions that have been obtained from Model SEED before summer 2015 had a smaller reaction (and metabolite) content, than those ones downloaded from model SEED or KBase afterwards. Model SEED/KBase updated and expanded their underlying database substantially in 2015 to include secondary metabolite, xenobiotics, and plant metabolism. Consequently, the average size of the draft reconstructions increased. These metabolic pathways were outside the scope of the current reconstruction effort and thus were not retained in AGORA reconstructions of the black group. Further experimental data and comparative genomic efforts will be required to establish that those out-of-scope reactions do indeed occur in the respective gut microbes. All reconstructions of the black group thus fall below the red line, meaning that the AGORA counterpart is smaller in terms of reactions. *However, it is notable and thanks to the QC/QA effort applied to all AGORA reconstructions that the there is no observable difference between the two groups when comparing the stoichiometric and flux consistent reactions. In all cases, we improved the quality of the reconstructions when considering this measure over the draft reconstructions.*

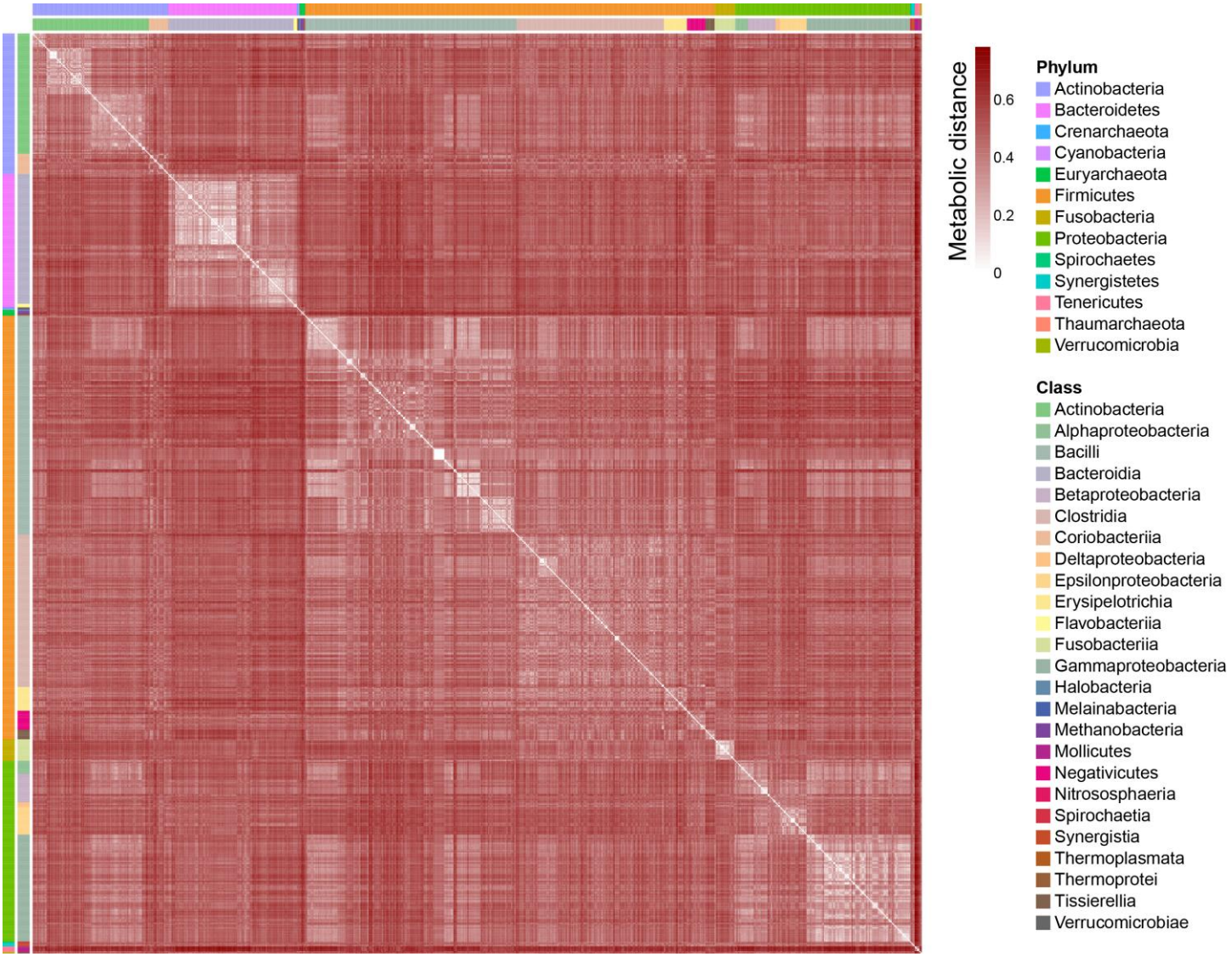
(a) The number of reactions in each AGORA reconstruction versus the corresponding draft reconstruction. The red line shows the line y=x, where the number of reactions in the AGORA reconstruction is the same as the number of metabolites in the draft reconstruction. (b) Histograms showing the change in number of reactions after the curation of the draft reconstructions of the two groups. The mean and standard deviation of the change in number of reactions is shown for both groups. A similar separation was observed for (c-d) the number of exchange reactions and (e-f) the number of stoichiometrically consistent reactions in AGORA versus the draft reconstructions. (g-h) In most organisms of both groups, reaction stoichiometric and flux consistency was improved by the AGORA curation process.



Supplementary Figure 3

Sensitivity of carbon source uptake and fermentation product secretion of seven published models and the corresponding AGORA models.

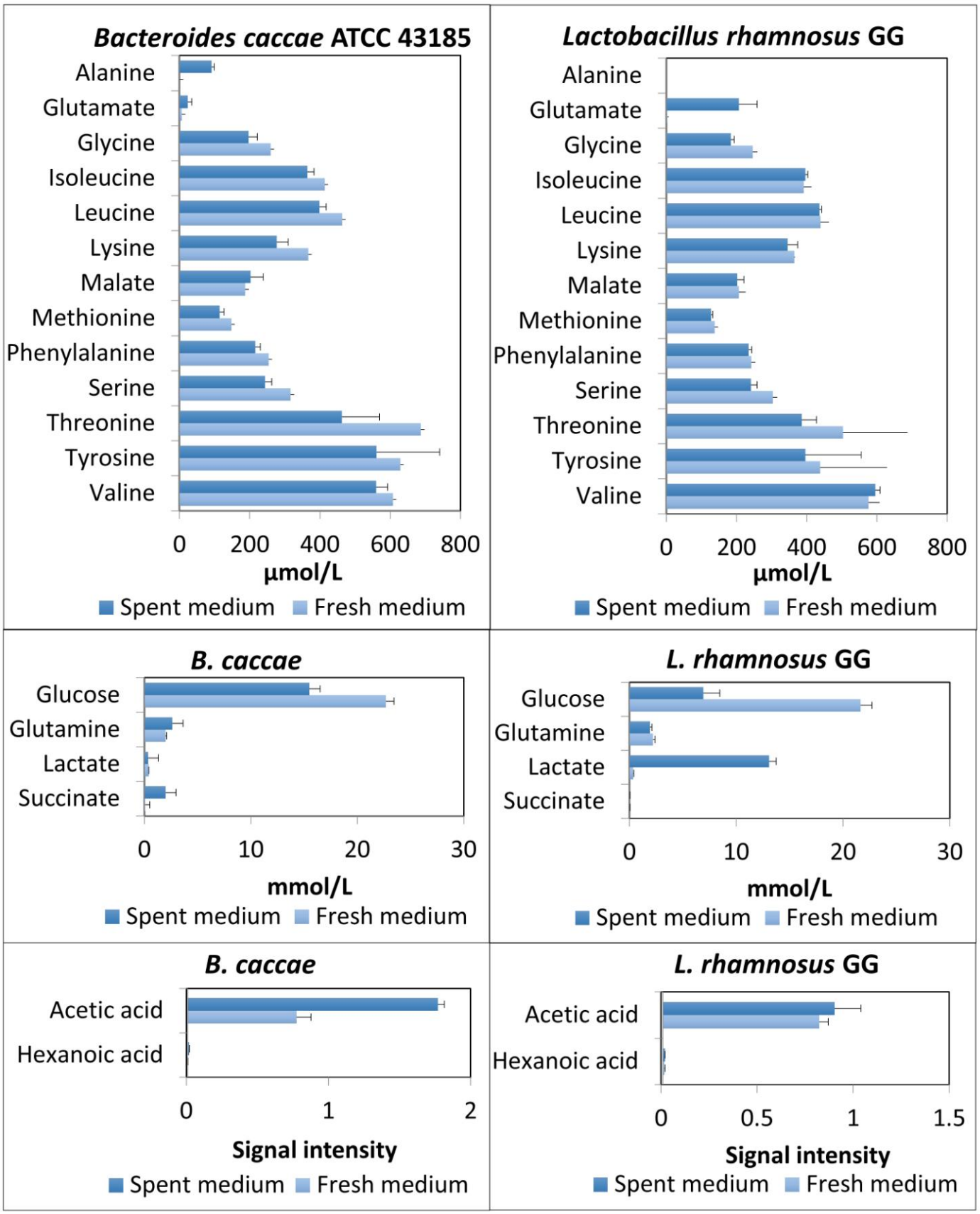
In the parentheses are the number of carbon sources or fermentation products that the models should take up or secrete, respectively, according to data from literature (Supplementary Table 1). Uptake and secretion capabilities were determined using flux variability analysis. All exchange reactions had unlimited upper and lower bounds and a minimum flux of 0.001 h^{-1} through the biomass objective function enforced. The seven published models can be found in 1) Pastink et al., Appl. Environ. Microbiol., 2009, 2) Teusink et al., J. Biol. Chem., 2006, 3) Thiele et al., BMC Syst. Biol., 2011, 4) Baumler et al., BMC Syst. Biol., 2011, 5) Liao et al., J. Bacteriol., 2011.



Supplementary Figure 4

Metabolic distances between the 773 AGORA reconstructions.

Reconstructions with identical reaction content have a metabolic distance of zero, while reconstructions having no overlap have a metabolic distance of 1 (Supplementary Table 9). Reconstructions are ordered based on phyla and taxonomic classes (Supplementary Table 5).

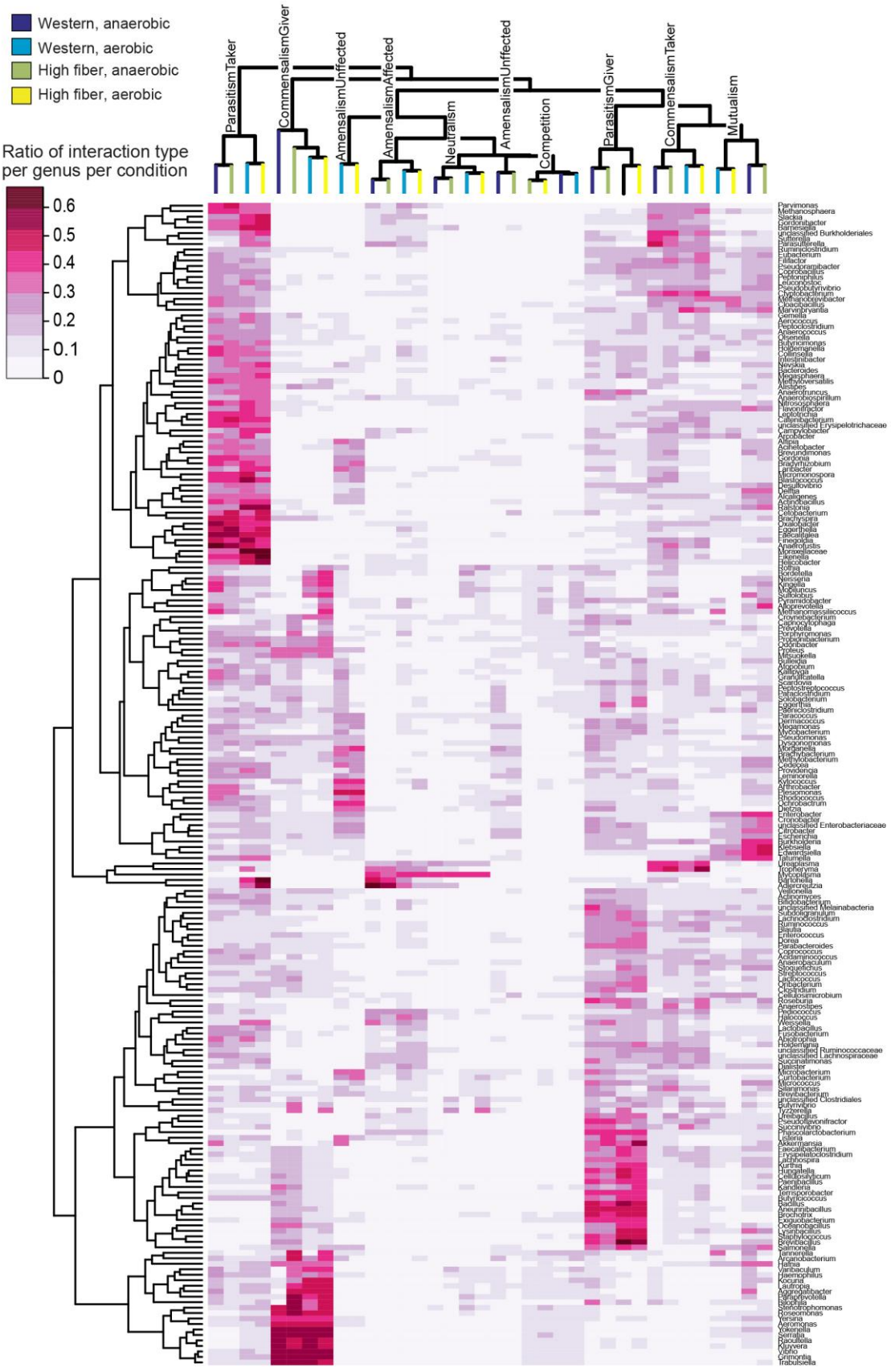


Nature Biotechnology: doi:10.1038/nbt.3703

Supplementary Figure 5

Metabolomic measurements for two bacterial strains grown *in vitro*.

Bar graphs showing the average metabolomic measurements determined for *Bacteroides caccae* ATCC 43185 and *Lactobacillus rhamnosus* GG ATCC 53103 during growth on DMEM 6429 medium supplemented with arabinogalactan (Supplementary Note 3). Error bars show the standard deviation. Statistically significant uptake and secreted is shown and compared with *in silico* predictions in Fig. 4a.

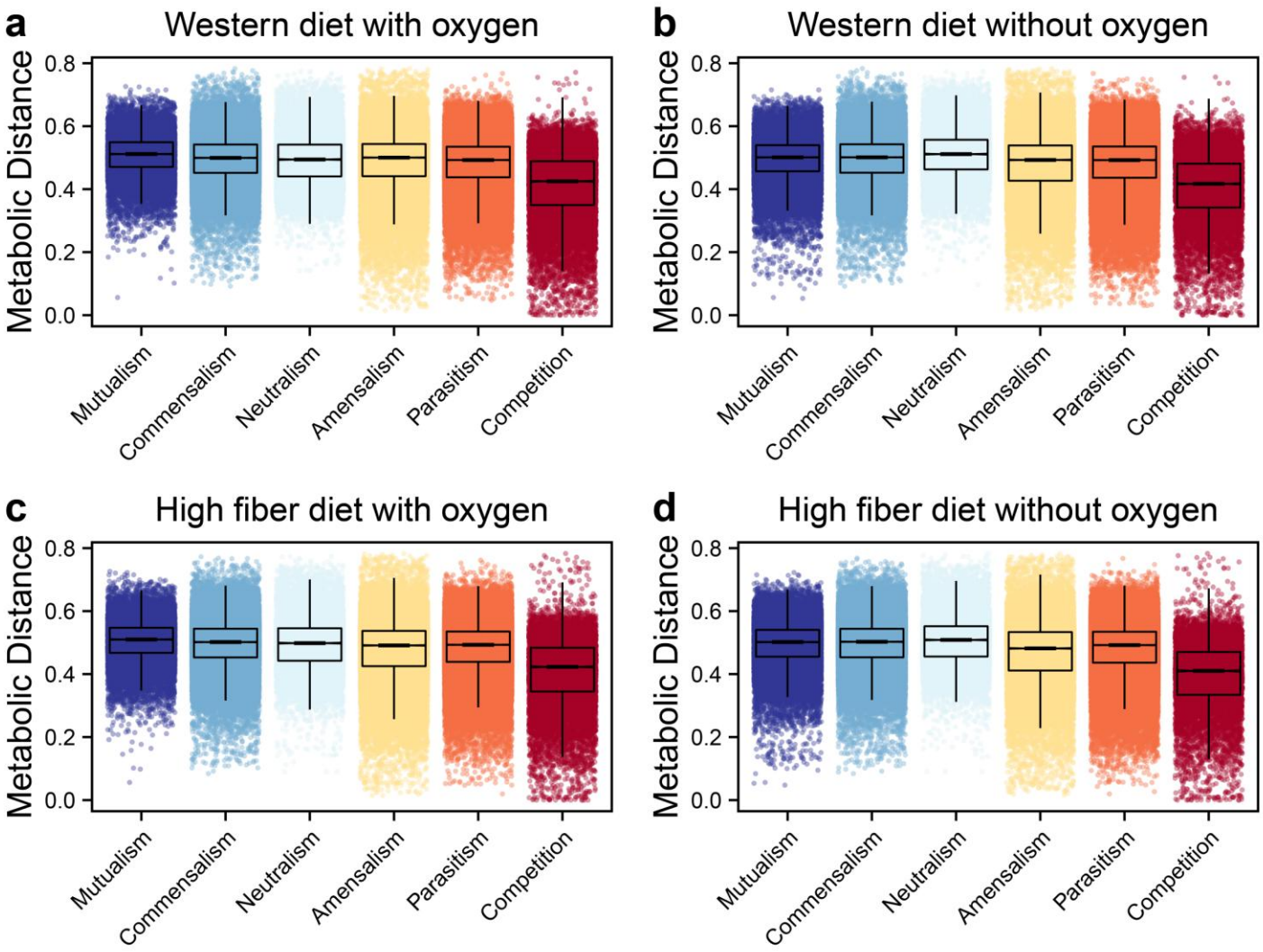


Nature Biotechnology: doi:10.1038/nbt.3703

Supplementary Figure 6

Clustering of the ratio of pairwise interaction types on the genus level per growth condition.

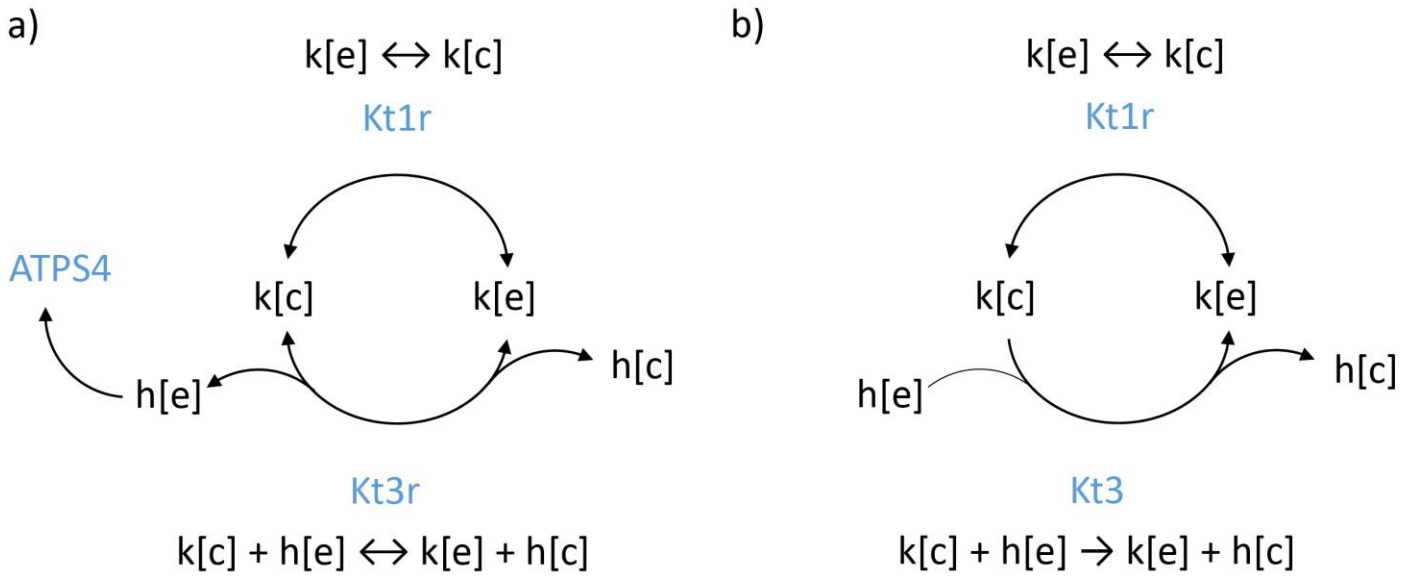
The growth of all AGORA microbe pairs (298,378 pairs) was simulated under four different conditions; using either Western diet or a high fiber diet (Supplementary Table 12) with and without oxygen. Depending on whether each microbe grew faster or slower in the co-culture simulation compared with by itself under the same condition (Supplementary Tables 6 and 9), the interaction between the two microbes was categorized as i) mutualism: both microbes grow faster, ii) commensalism: one microbe grows faster (Taker) while the other's growth rate is not affected (Giver), iii) neutralism: neither microbe's growth rate is affected, iv) amensalism: one microbe grows slower (Affected) while the other's growth rate is not affected (Unaffected), v) parasitism: one microbe grows faster (Taker) while the other grows slower (Giver), and vi) competition: both microbes grow slower. The heatmap shows the ratio of microbes per genus that have each interaction type per condition, e.g., the cell in the first row and first column shows that about 40% of microbes belonging to the Parvimonas genus have the interaction type "ParasitismTaker" on Western diet under anaerobic conditions.



Supplementary Figure 7

Metabolic distances plotted by the six types of interactions between the 298,378 microbe-microbe pairs by diet.

The growth of all AGORA microbe pairs (298,378 pairs) was simulated under four different conditions; using either Western diet or a high fiber diet (Supplementary Table 12) with and without oxygen. Depending on whether each microbe grew faster or slower in the co-culture simulation compared with by itself under the same condition (Supplementary Tables 6 and 9), the interaction between the two microbes was categorized as: mutualism (both microbes grow faster), commensalism (one microbe grows faster while the other's growth rate is not affected), neutralism (neither microbe's growth rate is affected), amensalism (one microbe grows slower while the other's growth rate is not affected), parasitism (one microbe grows faster while the other grows slower), and competition (both microbes grow slower). Based on the Jaccard index between the reaction content of the reconstructions of the two microbes, the metabolic distance between them was calculated (Online Methods, Supplementary Table 9). Reconstructions with identical reaction content have a metabolic distance of zero, whereas reconstructions sharing no reactions have a metabolic distance of 1.



Supplementary Figure 8

Example of a typical futile cycle resolved during curation.

a) When the ATP demand reaction is optimized, flux in the reverse direction through the potassium uniporter (Kt1r) and the potassium antiporter (Kt3r) leads to an unfeasibly high secretion flux of proton into the extracellular space. This leads to a corresponding flux through the ATP synthase (ATPS4), resulting in unfeasible ATP production. b) The futile cycle is resolved by replacing Kt3r with the irreversible version Kt3. This prevents infeasible ATP production while still allowing potassium transport in both directions.

Supplementary Note 1. Description of QC/QA and data-driven curation efforts.

We have compared and evaluated the predictive accuracy of AGORA against numerous resources, including genome-scale single gene deletion data (Fig. 1d), which are generally accepted as gold standard for the evaluation of the predictive potential of microbial metabolic models. The resources, which we used to assess and demonstrate the prediction accuracy, include (i) Genome-scale gene essentiality data (taken from http://www.essentialgene.org/). (ii) Experimental data on the utilization of all main groups of carbon sources. (iii) Experimental data on the major fermentation pathways in the human gut microbiota. (iv) Experimentally determined defined media/ nutrient requirements. The latter three data have been collected from more than 170 primary research papers (Supplementary Table 1) and the books: "Bergey's manual of systematic bacteriology" and "The prokaryotes: a handbook on the biology of bacteria". As such, we demonstrated the predictive sensitivity of AGORA using information generally accepted as gold standards by the field. The results can be found in Figures 1 and 3 in the main text, Supplementary Table 4, and Supplementary Fig. 3. We would like to highlight that inter-organism propagations were performed on the species level and in limited cases on the genus level. Hence, the metabolic functions were not inferred from distant organisms. It is to be noted that molecular systems biology is characterized by an iterative cycle of model predictions and experimental validation. It is natural in such a process for either prediction or validation to be one step ahead of the other. Compared to other biological domains, our experimental knowledge of the gut microbiome lags behind. Genome-scale models, such as those in AGORA, can be used to provide novel insight into the biology of the considered gut microbe and to drive the design of experimental projects.

Reaction directionality was propagated from biochemically curated reactions in the VMH database (http://vmh.life) and each was checked for consistency with thermodynamic estimates of maximum and minimum standard transformed Gibbs energy for each reaction using our state of the art Component Contribution method[1]. As documented on the VMH database, the following conditions were assumed: (i) temperature 310.15K; (ii) cytosol pH 7.2, extracellular pH 7.4; (iii) ionic strength 0.15 mol/L; (iv) 30 mV electrical potential difference between extracellular space and cytosol; and (v) minimum and maximum concentration range from 1e-7 to 1e-2 mol/L. The transformed Gibbs energy of formation of a metabolite is the sum of standard transformed Gibbs energy of formation and an RTlog(x) term, where R is the gas constant, T is temperature and x is the absolute concentration of the metabolite. Note using this approach, the activity coefficient is absorbed into the standard transformed Gibbs energy of formation. If we assume that the conditions above do vary, let us consider them one by one. (i) Temperature varies only slightly, and there are not enough enthalpic data to make a temperature adjustment beyond what we do. (ii) Our choice of pH 7.4 is representative of the terminal ileum. According to Fallingborg J.[2] "The intraluminal pH is rapidly changed from highly acid in the stomach to about pH 6 in the duodenum. The pH gradually increases in the small intestine from pH 6 to about pH 7.4 in the terminal ileum. The pH drops to 5.7 in the caecum, but again gradually increases, reaching pH 6.7 in the rectum." This slight variation in small and large intestinal pH might change the direction of some cytoplasmic transport reactions. (iii) The ionic strength is not known accurately, for any cell. We use 0.15 mol/L as is standard in the biochemical thermodynamic literature. Experimental evidence would be required to justify the use of a different value. (iv)

The electrical potential difference between extracellular space and cytosol could indeed vary, however, then one would be modeling the same organism in two different conditions so one would expect the model predictions to be different. (v) There is insufficient data on the absolute concentrations of the metabolome in the cytoplasm in any cell. With gut microbial species, the paucity of data is even more acute. Without further data, one could not safely assume a tighter concentration range.

Supplementary Note 2. Comparison with published reconstructions.

We compared all 3,192 unique reactions included in the 773 AGORA reconstructions with the 4,608 unique reactions included in 11 published gut microbe reconstructions previously used to simulate a model gut community[3] after unifying the reaction namespaces. Additionally, the reconstructions of eight strains in AGORA overlapped with published reconstructed strains[4-10] and were compared against each other directly. Subsystems were assigned to every reaction in the AGORA and the published reconstructions based on the subsystem nomenclature in Recon 2[11]. The results are shown in Supplementary Tables 7-8.

Between 310 and 1,058 (mean 682±290) reactions overlapped between the respective AGORA and published reconstructions. Between 305 and 934 (mean 770±227) reactions were unique to the curated reconstructions, and between 229 and 1,776 (mean 956±678) reactions were unique to the published reconstructions (Supplementary Table 7). Most of the reactions that were unique to the published reconstructions belonged to exchange and transport subsystems, particularly for *Escherichia coli* K-12 substr. MG1655, *Escherichia coli* O157H7 substr. Sakai, *Klebsiella pneumoniae* MGH78578, and *Salmonella enterica* sv. typhimurium LT-2 (Supplementary Table 8). These reconstructions used the *Escherichia coli* reconstruction iAF1260[12] as a template, which includes a periplasm compartment. Thus, the majority of reaction content unique to these reconstructions could be attributed to the presence of the additional periplasm compartment, which requires additional transport reactions. For example, the *Escherichia coli* K-12 substr. MG1655 reconstruction iEco1339_MG1655[9] contains 1,573 reactions not found in the AGORA reconstruction. Of those, 913 reactions take place in the periplasmic compartment. Of the 934 reactions found only in the AGORA *Escherichia coli* K-12 substr. MG1655 reconstruction, 197 are cytosolic versions of reactions that take place in the periplasm in the published reconstruction. Another major reason for discrepancies between the AGORA and published reconstruction was the captured cell wall and lipopolysaccharide biosynthesis pathways. Cell wall and lipopolysaccharide structures are species-specific and generally poorly annotated, and hence difficult to curate. Consequently, between 30 and 299 reactions from cell wall and lipopolysaccharide biosynthesis were unique to the published reconstructions (Supplementary Table 8). Curating for cell wall and lipopolysaccharides structures and including accurate transport mechanisms would require experimental data, which is not available for most AGORA organisms. Thus, these curation steps were not performed in the present study, with the exception of reconstructions from two genera, *Mycoplasma* and *Ureoplasma* (Online Methods).

At the same time, 161 reactions from cell wall and lipopolysaccharide biosynthesis were present in the *E. coli* draft reconstruction, and thus also present in its AGORA version, but not in the published reconstruction iEco1339_MG1655 (Supplementary Table 8). These reactions are involved in fatty acid biosynthesis, as the ones catalyzed by the 3-oxoacyl-acyl-carrier-protein reductase. Since the presence of these enzymes in *E. coli* is supported by genome annotation, these reactions should also be included in the published reconstruction. Another example is dipeptide degradation, which is supported by genome annotation in the AGORA reconstruction but is absent from the published *E. coli* reconstruction.

The eight AGORA reconstructions had on average a higher number of blocked reactions that the respective published reconstructions (31 +/- 9 % compared with 13 +/- 6 %). This was mainly due to the deletion of reactions that were added during the automated gap-filling step of the draft reconstructing pipelines and were found to no longer be required for biomass production. Since the presence of these reactions is not supported by gene annotation or experimental data, corresponding to a confidence score of 1[13], their inclusion is hypothetical.
Another cause for blocked reactions in AGORA reconstructions was the adjustment of reaction reversibilities to VMH standards. This led to some pathways that were carrying flux in an infeasible direction in the draft reconstructions to be blocked in the resulting AGORA reconstructions. For example, the vitamin B12 biosynthesis pathway was reversible in many draft reconstructions and often allowed to produce downstream biomass precursors from cobalamin. Making this pathway irreversible caused it to be blocked in many reconstructions due to missing steps, in agreement with many microbes being unable to synthesize cobalamin[14] (Supplementary Table 18). Nonetheless, the QC/QA curation effort led to an increase of the overall stoichiometric and flux-consistent reactions in the AGORA reconstructions compared to the draft reconstructions (Fig. 1, Supplementary Figs. 1-2).

Supplementary Note 3. *In vitro* cell cultures and cell counting.

In vitro cell culture: Pre-cultures of *Bacteroides caccae* ATCC 43185 (*B. caccae*) and *Lactobacillus rhamnosus* GG ATCC 53103 (LGG) were prepared using Brain Heart Infusion Broth (BHIS; Sigma) supplemented with 1% hemin. The pre-cultures were run for 20 hours under anaerobic conditions and while shaking at 37 °C. The volume of the cell suspension was adjusted to a maximum optical density (OD) at 600 nm of 0.8 using sterile 0.9% w/v NaCl solution to obtain reliable OD measurements. Following pre-culturing, centrifugation was carried out at 4,700 x g for 10 min at room temperature. The resulting cell pellets were washed twice in 0.9% w/v NaCl solution and subsequently resuspended in 10 ml of 0.9% w/v NaCl solution. Subsequently, 1 ml of the cell suspensions were inoculated in DMEM 6429 supplemented with 1% hemin and 3.33% vitamin with a starting OD of 0.1 and K maintained under anaerobic conditions. Culture was carried out in media with or without the addition of arabinogalactan (Sigma; 9.4 g/l). *B. caccae* and LGG were cultured for 33 and 44 hours of culture for *B. caccae* and LGG, respectively. Cells were harvested for cell counting by centrifugation (4,700 g) and 750 µL aliquots of supernatant were removed for subsequent metabolite extraction. The aliquots were snap-frozen and placed at -80 °C until analysis. The measured ODs and pH values are listed in Supplementary Table 10.

Cell counting: Bacterial pellets were thawed and subsequently stained with the Texas Red®-X dye–labeled Wheat Germ Aglutinin component which selectively binds to the surface of gram-positive bacteria, effectively distinguishing them from gram-negative bacteria. Cells were washed, resuspended in sterile 0.9% w/v NaCl solution and quantified by flow cytometry (BD Fortessa) using negative beads (Thermo Fischer) as a standard for the volume of suspension. The resulting data were analyzed using the DIVA 8.0.1 software (BD Biosciences). The cell counts are shown in Supplementary Table 10.

Real-time PCR (RT-PCR): The presence of *B. caccae* cells in the cultures were validated using qPCR. Microbial genomic DNA was extracted using the PowerSoil DNA Isolsation Kit from MoBio accordingly to the manufacturer's protocol. qPCR was performed using *B. caccae*-specific primers (Eurogentec). The *B. caccae*-specific primers were 5'-CCC GGA GTT GGA AAA CAA TG-3' (forward) and 3'-CCT CTT CAG AAA TGA GCT TTT GC-3' (reverse). 5 ng of DNA were used in a 20 µl PCR reaction mixture containing 10 µl iQ™ SYBR® Green Supermix (Bio-Rad) and 500 nM of each primer. PCR amplifications were performed on a LightCycler® 480 Instrument (40 cycles at 95°C for 10 sec and 55°C for 20 sec; Roche).

Supplementary Note 4. Fermentation and carbon source utilization pathways captured by AGORA.

A thorough literature search was performed for the reconstruction of the main fermentation pathways in the human gut and their distribution across phyla. The main products of carbohydrate and protein fermentation by the human gut microbiota are the short-chain fatty acids (SCFAs) acetate, propionate, and butyrate[15]. The different routes leading to these products were reconstructed for all AGORA organisms reported to carry these pathways and are briefly described below.

Acetate is produced via the widespread acetate kinase[16]. Moreover, bacteria in the *Bifidobacterium* genus produce acetate via the bifid shunt[16], and the genera *Blautia* and *Marvinbryantia* convert CO_2 and hydrogen to acetate via acetogenesis[17]. Three pathways exist for the conversion of carbohydrates and amino acids to propionate, with succinate, propane-1,2-diol, and lactate, respectively, as intermediates[18]. Their distribution in the human gut microbiota has been analyzed by comparative genomics and experimentally validated[18]. Carbohydrate and amino acid fermentation to butyrate is carried out via five routes, with acetyl-CoA, glutarate, lysine, 4-aminobutyrate, and succinate, respectively, as intermediates[19]. A genomic analysis of their distribution across several phyla has been performed[19]. Moreover, gut microbes produce lactate, formate, butanol, acetoin, 2,3-butanediol, and ethanol, as well as carbon dioxide and hydrogen[16]. For most gut microbes, the capabilities to produce these acids and gases are well-described in the literature[16]. Produced hydrogen is cross-fed to other species in three ways: methanogenesis by archaeal representatives resulting in methane production[20], dissimilatory sulfate reduction to sulfide[21], and acetogenesis yielding acetate[17]. Amino acid fermentation results in the production of not only SCFAs and gases, but also the branched-chain fatty acids isobutyrate and isovalerate, as well as phenols (e.g., phenylacetate) and indoles[15, 22]. The

pathways described above, resulting in 16 fermentation products and two gases in total, were included in the respective AGORA reconstructions.

The available literature was also searched for the utilization of carbon sources by AGORA organisms. The gut microbiota utilizes a variety of diet-and host-derived carbon sources, including simple sugars, starch, fiber, host-derived polysaccharides, protein, and organic acids[23]. The potential to exploit carbon sources is species-specific. While the ability to utilize mono- and disaccharides is widely spread[16], the capability to break down diet- and host-derived polysaccharides is limited to certain genera, e.g., *Bacteroides*, *Bifidobacterium*, *Roseburia*, and *Ruminococcus*[24]. Some species utilize amino acids as carbon and energy sources, e.g., *Clostridium difficile*, *Pseudomonas aeruginosa*, and *Porphyromonas gingivalis*[16], and some utilize intermediates of central metabolism and organic acids, e.g., *Bacillus cereus*, *Citrobacter* sp., *Oxalobacter formigenes*, and *Veillonella* sp.[16]. A thorough literature search was performed for the distribution of these four groups of carbon sources utilized by the AGORA organisms: (i) simple sugars and oligosaccharides, (ii) polysaccharides and fibers, (iii) amino acids, and (iv) organic acids and other metabolic intermediates. In total, information on the utilization of 95 carbon sources was gathered and the corresponding pathways were included in the respective AGORA reconstructions.

Supplementary Note 5. Definition of sub-pathways.

A sub-pathway was determined as any set of reactions that converts an initial substrate of the pathway into the final product(s) of the pathway. For example, the biosynthesis of a purine nucleotide has one initial substrate, phosphoribosyl pyrophosphate, and two final products, AMP and GMP. Thus, this pathway includes two sub-pathways. A sub-pathway was considered complete if all genes required for all the reactions in the sub-pathway were present in the subsystem. A sub-pathway was gap-filled if the length of a gap in the pathway did not exceed one reaction. In this case, the gaps were filled by reactions not associated with GPRs. Sub-pathways with gaps longer than one reaction were considered incomplete and reactions for these sub-pathways were not included in reconstructions. These criteria were applied to all the reconstructed pathways except the citric acid cycle, because the presence of an incomplete citric acid cycle has been confirmed for multiple microbial genomes[25]. These incomplete versions of the citric acid cycle are used for the biosynthesis of various compounds, such as fatty and amino acids[25]. Thus, no gap-filling was performed for the citric acid cycle and reactions for this pathway were included in the reconstructions regardless of the pathway's completeness in the genome.

Supplementary Note 6. Curation of respiration and quinone biosynthesis in AGORA.

The genomes of the reconstructed gut bacteria contain multiple aerobic reductases as well as anaerobic reductases for tetrathionate, thiosulfate, polysulfide, sulfite, adenylyl sulfate, heterodisulfides, fumarate, trimethylamine N-oxide, dimethyl sulfoxide, nitrate, nitrate, nitrogen oxide, nitrous oxide, selenate, and arsenate[26]. Nonetheless, the reactions for reduction of respiratory electron acceptors should include two half-reactions, one for a reduction of the electron acceptor itself and another for an oxidation of the corresponding quinone. Since the

repertoire of synthesized quinones varies among bacterial taxa[27] and quinones demonstrate specificity for their electron acceptors[28], the inclusion of respiratory reactions into the models was preceded by the reconstruction of the quinone biosynthetic pathways.

In bacteria, one pathway has been described for ubiquinone (UQ) biosynthesis. For menaquinone (MK) biosynthesis, two different pathways are known, the first one through O-succinylbenzoate and the second one through futalosine[27]. The last steps of MK biosynthesis through futalosine are unknown, but are proposed to be catalyzed by a polyprenyltransferase, a carboxy-lyase, and a methyltransferase[29]. All three steps of polyprenylation, carboxyl elimination, and methylation are present also in UQ biosynthesis and in MK biosynthesis O-succinylbenzoate biosynthesis in the same order as listed above. So, we proposed that this reaction mechanism should be conserved in MK through futalosine biosynthesis. Thus, in analogy with the corresponding steps in the UQ and MK via O-succinylbenzoate biosynthesis pathways, we predicted the three last steps of the MK through futalosine pathway in the analyzed genomes:

1) Polyprenyltransferase
   $$\text{1,4-dihydroxy-6-naphthoate} + \text{polyprenyl-pyrophosphate} \rightarrow \text{3-polyprenyl-1,4-dihydroxy-6-naphthoate} + \text{pyrophosphate}$$
2) Carboxy-lyase
   $$\text{3-polyprenyl-1,4-dihydroxy-6-naphthoate} \rightarrow \text{2-demethyl menaquinol} + CO_2$$
3) Methyltransferase
   $$\text{2-demethyl menaquinol} + \text{S-adenosyl-L-methionine} \rightarrow \text{menaquinol} + \text{S-adenosyl-L-homocysteine} + H^+$$

The reactions for the respiratory reduction of electron acceptors were constructed in agreement with the following features, (1) the presence of the quinone biosynthetic pathways in the analyzed genome, (2) specificity of the electron acceptor to quinones, and (3) subcellular localization of an active center of the corresponding reductase. For example, the *Bacteroides thetaiotaomicron* genome contains the biosynthesis pathway for MK and 2-demethylmenaquinone (DMK), one aerobic reductase with the cytoplasmic active center, and two anaerobic reductases: a nitrite reductase with an extracellular active center and a fumarate reductase with a cytoplasmic active center. Because oxygen can be reduced with MK, whereas nitrite and fumarate can be reduced with both MK and DMK[28], the respiratory reduction of electron acceptors in *Bacteroides thetaiotaomicron* is carried out by the following five reactions:

i. Cyd: aerobic reductase
   $$0.5\ O_2[c] + \text{menaquinol}[c] + 2\ H^+[c] \rightarrow H_2O[c] + \text{menaquinone}[c] + 2\ H^+[e]$$
ii. Nrf: nitrite reductase
   $$NO_2[e] + 3\ \text{menaquinol}[c] + 2\ H^+[c] \rightarrow NH_4^+ + 3\ \text{menaquinone}[c] + H_2O[c]$$
   $$NO_2[e] + 3\ \text{2-demethyl menaquinol}[c] + 2\ H^+[c] \rightarrow NH_4^+ + 3\ \text{2-demethyl menaquinone}[c] + H_2O[c]$$
iii. Frd: fumarate reductase
   $$\text{fumarate}[c] + \text{menaquinol}[c] \rightarrow \text{succinate}[c] + \text{menaquinone}[c]$$
   $$\text{fumarate}[c] + \text{2-demethyl menaquinol}[c] \rightarrow \text{succinate}[c] + \text{2-demethyl menaquinone}[c]$$

Reactions for ATP synthesis via proton-driven ATP synthases were also added to the reconstructions. All analyzed genomes have genes for F-type or V-type ATP synthases. All added respiration, quinone biosynthesis, and ATP synthase reactions are listed in Supplementary Table 16.

Supplementary Note 7. Curation of nutrient requirements.

The *in silico* growth requirements were computed by setting the lower bounds for all exchange reactions to zero individually and predicting if the model could still produce biomass. The analysis revealed essential metabolites for *in silico* growth that are unlikely to be found in the human gut, such as end products of coenzyme biosynthesis, including CoA and NAD(P)H. Bacteria take up precursors of these coenzymes in the form of vitamins, such as pantothenic acid or nicotinic acid. Exchange and transport reactions for vitamins were added where necessary and we ensured that CoA synthesis from pantothenic acid, as well as NADH biosynthesis from nicotinic acid, were unblocked in all metabolic reconstructions. In some cases, a draft reconstructions required dimers or oligomers in the *in silico* growth medium to fulfill a monomer requirement. As it can be expected that uptake of the monomer itself would also satisfy the growth requirement, exchange and transport reactions for the corresponding monomers were added. For example, the galactose-containing oligosaccharides stachyose, melibiose, and lactose were in some cases essential for a model to satisfy its requirement for galactose. Similarly, some draft models depended on dipeptides for certain amino acids, so we enabled the uptake of amino acid monomers.

False negative predictions were resolved in the following ways: (i) for essential nutrients, if they were not yet present in the reconstruction, exchange and transport reactions were added, (ii) for metabolites not included in the biomass reaction, a demand reaction was added for metabolites required *in vitro* and a minimal flux through the demand reaction was enforced, and (iii) for metabolites that were false positives due to gap-filled reactions in the biosynthesis pathways, those gap-filled reactions were removed. False positive predictions were resolved by manually inspecting and gap-filling the corresponding pathways to enable production or consumption of dead-end metabolites. This curation was performed for 244 models.

After curation, 173 false positive predictions remained, which were due to multiple gaps in the networks that could not be resolved without further experimental and genomic evidence. For most of the 219 remaining false negative predictions, the compounds were essential *in vitro* although the biosynthesis pathways were completely annotated in the bacterial genomes. This has been previously observed, e.g., in *Lactobacillus plantarum* WCFS1, and may be explained by feedback inhibition of the biosynthesis pathways *in vitro*[30].

Supplementary Note 8. Metabolite extraction.

Short chain fatty acid extraction, derivatization, and GC-MS measurement

The extraction of short chain fatty acids was based on a protocol from Moreau *et al.*[31]. Briefly, 10 µL of the internal standard (2-Ethylbutyric acid, *c* = 200 mmol/L) were added to 190 µL of

medium. After acidification with 10 µL of hydrochloric acid (*c* = 1 mol/L), 1 mL of diethyl ether was added and the samples were vortexed for 10 min at 1400 rpm at room temperature (Eppendorf Thermomixer). The upper organic phase was separated by centrifugation (5 min, 21,000 xg) and 900 µL were collected in a new reaction tube. Again, 1 mL of diethyl ether were added to the medium, incubated and separated by centrifugation. 900 µL of the organic phase were combined with the first extract. Then, 250 µl were transferred into a GC glass vial with micro insert (5-250 µL) in triplicates. For derivatization, 25 µL of N-tert-Butyldimethylsilyl-N-methyltrifluoroacetamide with 1% tert-Butyldimethylchlorosilane (MTBSTFA + 1% TBDMSCI, Restek) was added and the samples were incubated for a minimum of 1 hour at room temperature. To determine retention times and evaluate separation efficiency, a Volatile Free Acid Mix (Sigma-Aldrich) including all compounds of interest was prepared, extracted, and derivatized as described before.

GC-MS analysis was performed by using an Agilent 7890A GC coupled to an Agilent 5975C inert XL Mass Selective Detector (Agilent Technologies). A sample volume of 1 µl was injected into a Split/Splitless inlet, operating in split mode (20:1) at 270 °C. The gas chromatograph was equipped with a 20 m (I.D. 180 µm, film 0.18 µm) DB-1MS capillary column (Agilent J&W GC Column). Helium was used as carrier gas with a constant flow rate of 1.0 ml/min. The GC oven temperature was held at 80 °C for 0.75 min and increased to 150 °C at 15 °C/min. After 2 min, the temperature was increased at 50 °C/min to 280 °C and held for 2 min. The total run time was 12.017 min. The transfer line temperature was set to 280 °C. The mass selective detector (MSD) was operating under electron ionization at 70 eV. The MS source was held at 230 °C and the quadrupole at 150 °C. The detector was switched off during elution of MTBSTFA. For quantification, measurements of the compounds of interest were performed in selected ion monitoring mode. Dwell times as well as quantification and qualification ions (m/z) are shown in Supplementary Table 20.

Absolute quantification of medium components using the YSI Biochemistry Analyzer

Prior to measurement**,** media samples were filtrated (PHENEX-RC 4mm, 0.2 µm; Phenomenex) to remove particles. Absolute quantitative values for lactic acid, glutamine, glutamic acid, and glucose were acquired using the 2950D Biochemistry Analyzer (YSI). In addition, for a precise and reliable quantification, external concentration curves for each compound of interest were prepared and measured in triplicate.

Polar metabolite extraction, derivatization, and GC-MS measurement

Extracellular metabolites from media samples were extracted in triplicate using ice-cold extraction fluid (5:1 methanol/water, v/v) containing the internal standards [U-13C]ribitol (*c* = 10 µg/mL; Omicron Biochemicals, Inc) and pentanedioic acid-D6 (*c* = 4 µg/mL; C/D/N Isotopes Inc.). 20 µL of medium was added to 180 µL ice-cold extraction fluid, vortexed for 15 min at 4 °C and 1,400 rpm (Eppendorf Thermomixer), then, centrifuged at 21,000 xg for 5 min at 4 °C. 50 µL of medium extracts were transferred to GC glass vial with micro insert (5-250 µL) and evaporated under vacuum to dry at -4 °C. For absolute metabolite quantification, a dilution series of a standard mixture containing all metabolites of interest was included in the extraction procedure and measured in triplicates. Metabolite derivatization was performed by using a multipurpose

sampler (Gerstel). Dried medium extracts were dissolved in 15 µl pyridine, containing 20 mg/ml methoxyamine hydrochloride (Sigma-Aldrich), at 55 °C for 90 min under shaking. After adding 15 µl MTBSTFA + 1% TBDMSCI (Restek), samples were incubated at 55 °C for 60 min under continuous shaking.

GC-MS analysis was performed by using an Agilent 7890A GC coupled to an Agilent 5975C inert XL Mass Selective Detector (Agilent Technologies). A sample volume of 1 µl was injected into a Split/Splitless inlet, operating in split mode (10:1) at 270 °C. The gas chromatograph was equipped with a 30 m (I.D. 250 µm, film 0.25 µm) DB-35MS capillary column + 5 m DuraGuard column in front of the analytical column (Agilent J&W GC Column). Helium was used as carrier gas with a constant flow rate of 1.0 ml/min. The GC oven temperature was held at 100 °C for 2 min and increased to 300 °C at 10 °C/min and held for 4 min. The total run time was 26 min. The transfer line temperature was set to 280 °C. The MSD was operating under electron ionization at 70 eV. The MS source was held at 230 °C and the quadrupole at 150 °C. For precise quantification, GC-MS measurements of the derivatives of interest were performed in selected ion monitoring mode. Dwell times as well as quantification and qualification ions (m/z) are shown in Supplementary Table 20.

Data normalization and data processing

All GC-MS chromatograms were processed using MetaboliteDetector, v3.020151231Ra[32]. The software package supports automatic deconvolution of all mass spectra. Compounds were annotated by retention time and mass spectrum. The internal standards were added at the same concentration to every medium sample to correct for uncontrolled sample losses and analyte degradation during metabolite extraction. The data set was normalized by using the response ratio of the QI_Analyte and the QI_Internal Standard (peak area of the analyte divided by the peak area of the internal standard). Absolute concentrations were determined using calibration curves from external standards. To evaluate the variability of independent cultivations, mean values of three technical replicates have been calculated for each biological replicate.

Supplementary Table 4: List of tests that the AGORA reconstructions were subjected to evaluate the curation effort. The number of reconstructions that could be curated and number of reconstructions that passed the tests are shown. Several genomes were missing from the PubSEED platform[33] and could thus not be curated based on the comparative genomic analyses (Supplementary Table 2). Similarly, a few organisms were not captured in the literature-driven curation on carbon sources and fermentation pathways (Supplementary Table 1). Less than a third of the AGORA organisms had literature information on essential nutrients (Supplementary Table 1) and were curated based on the available information.

Test	**# of reconstructions that could be curated**	**# of reconstructions that passed the tests**
Reaction and metabolite nomenclature standardized with VMH database	773	773
Reaction constraints standardized with VMH database	773	773
Can grow anaerobically	773	773
No metabolites are produced from nothing	773	773
ATP production rates from the available carbon sources are feasible	773	773
Metabolite formulas are defined and mass-charge balanced	773	773
Number of gap-filling reactions that were included for modeling purposes only is minimized	773	773
In silico growth rates on the defined diets are in realistic ranges	773	773
Gene-protein-reaction associations and reactions in aerobic and anaerobic respiration, B-vitamin biosynthesis pathways, central carbon metabolism, amino acid biosynthesis and/or pyrimidine and purine biosynthesis determined by a comparative genomics approach are implemented	612	612
Carbon source utilization pathways supported by evidence from literature are present and can carry flux	732	732
Fermentation pathways supported by evidence from literature are present and can carry flux	765	765
Species' capabilities/ incapabilities to synthesize essential biomass precursors are captured	244	112 passed all tests, the remaining 132 have one or more false positive or false negative predictions.

Supplementary Table 7: Overview of the comparison of the reaction content between AGORA reconstructions and published reconstructions targeting the same strain. Also shown is the comparison between all 773 AGORA reconstructions and 11 published gut microbe reconstructions that were previously used to construct a simplified gut microbe community model[3]. Eight strains overlapped between both reconstruction collections and were compared directly. Shown is the comparison between the eight overlapping strains individually, and between the pooled reactions of all 11 published and all 773 AGORA reconstructions. recon. = reconstruction.

Reconstructed Species	**Total reactions in AGORA recon.**	**Total reactions in published recon.**	**Overlapping reactions**	**Reactions only in AGORA recon.**	**Reactions only in published recon.**	**Ref. for published recon.**
Bacteroides thetaiotaomicron VPI 5482	1,362	1,528	1058	305	470	[4]
Lactobacillus plantarum WCFS1	1,213	777	395	818	382	[6]
Streptococcus thermophilus LMG 18311	927	556	327	600	229	[5]
Escherichia coli str K 12 substr. MG1655	1,786	2,426	853	934	1,573	[9]
Escherichia coli O157 H7 str. Sakai	1,742	2,372	821	922	1,551	[9]
Helicobacter pylori 26695	1,014	555	310	705	245	[8]
Klebsiella pneumoniae MGH78578	1,801	2,262	843	959	1,419	[10]
Salmonella enterica sv. Typhimurium LT2	1,765	2,623	847	919	1,776	[7]
All reconstructions (11 vs . 773)	3,192	4,608	2,066	1,127	1,540	[4-10, 34, 35]

Supplementary Table 8: Subsystem coverage of reactions that are overlapping between curated reconstructions and published reconstructions targeting the same strain. Also shown are differences in subsystem coverage between all 773 reconstructions and 11 published gut microbe reconstructions. Shown is the comparison between the eight overlapping strains individually, and between the pooled reactions of all 11 published and all 773 AGORA reconstructions. For references for the published reconstructions, refer to Supplementary Table 7. PPP = pentose phosphate pathway.

		Amino acid metabolism	Biomass	Carbohydrate metabolism	Cell wall/ lipopolysaccharide biosynthesis	Central metabolism	Exchange/demand	Fatty acid metabolism	Nitrogen, sulfur, and iron metabolism	Other	PPP and amino sugars	Polysaccharide degradation	Purine and pyrimidine metabolism	Transport	tRNA charging	Vitamin and cofactor biosynthesis
Bacteroides thetaiotao-micron VPI 5482	Overlapping	151	0	28	118	63	181	118	2	6	46	65	100	73	0	108
	AGORA only	36	1	13	61	22	34	2	1	7	7	3	18	49	1	49
	Published only	47	1	18	59	22	107	13	0	4	12	35	14	104	0	33
Lactobacillus plantarum WCFS1	Overlapping	70	0	14	17	38	70	1	0	5	25	3	67	41	1	43
	AGORA only	82	1	23	193	43	86	119	5	7	16	6	51	116	0	70
	Published only	56	2	15	57	24	44	15	6	7	4	0	31	83	18	20
Streptococcus thermophilus LMG 18311	Overlapping	69	0	11	11	26	35	0	0	2	13	0	82	41	1	35
	AGORA only	77	1	11	135	39	73	82	2	8	8	1	47	69	0	46
	Published only	33	1	5	52	16	9	17	2	2	3	0	16	40	18	16
Escherichia coli str K 12 substr. MG1655	Overlapping	131	0	35	133	89	160	45	5	8	44	1	100	0	2	100
	AGORA only	93	1	14	161	73	59	107	16	11	23	3	58	238	0	77
	Published only	72	2	79	261	65	146	66	37	15	11	1	64	686	21	47
Escherichia coli O157 H7 str. Sakai	Overlapping	123	0	34	131	85	151	45	5	7	38	1	99	0	2	100
	AGORA only	97	1	15	160	72	57	107	16	12	18	4	59	227	0	77
	Published only	72	2	75	254	66	154	66	36	15	12	1	65	665	21	47
Helicobacter pylori 26695	Overlapping	59	0	6	20	19	55	4	0	2	13	0	47	29	2	54
	AGORA only	64	1	1	208	38	63	111	3	11	11	1	39	82	0	73
	Published only	39	1	5	30	34	23	7	2	2	5	0	21	48	0	28
	Overlapping	141	0	31	127	83	144	43	4	10	38	1	99	2	2	111

		Amino acid metabolism	Biomass	Carbohydrate metabolism	Cell wall/ lipopolysaccharide biosynthesis	Central metabolism	Exchange/demand	Fatty acid metabolism	Nitrogen, sulfur, and iron metabolism	Other	PPP and amino sugars	Polysaccharide degradation	Purine and pyrimidine metabolism	Transport	tRNA charging	Vitamin and cofactor biosynthesis
Klebsiella pneumoniae MGH78578	AGORA only	109	1	15	165	73	70	101	11	19	23	8	50	227	0	83
	Published only	73	1	63	238	52	149	61	12	12	9	1	61	629	21	44
Salmonella enterica sv. Typhimurium LT2	Overlapping	131	0	33	132	73	169	45	5	9	44	1	100	0	2	103
	AGORA only	92	1	14	160	79	39	107	17	11	23	3	59	213	0	102
	Published only	70	3	63	299	73	198	65	59	18	7	1	61	778	21	60
All reconstructions (11 vs . 773)	Overlapping	271	0	66	286	163	332	165	11	20	74	69	149	243	2	214
	AGORA only	130	1	26	85	134	91	42	38	29	30	10	50	192	0	65
	Published only	158	12	102	443	115	252	64	70	40	17	38	83	1,033	24	88

Supplementary Table 10: Cell count per ml, optical density (OD), and pH values before and after cell culture of *Bacteroides caccae* ATCC 43185 (*B. caccae*) and *Lactobacillus rhamnosus* GG ATCC 53103 (LGG). All samples were grown on DMEM 6429 medium supplemented with 1% haemin, 3.33% vitamin K, and 9.4 g/L arabinogalactan.

Sample	OD initial (time=0)	OD endpoint	Cell count at endpoint	Time (hours) of culture	pH before	pH after
B.caccae (1)	0.09	0.49	2.86E+07	31	8.05	NA
B.caccae (2)	0.05	0.56	2.79E+06	27	8.05	7.16
B.caccae (3)	0.09	0.42	3.09E+05	27	8.05	7.27
LGG (1)	0.09	0.55	5.38E+07	31	8.2	5.72
LGG (2)	0.12	0.1	1.27E+05	48	8.2	7.1
LGG (3)	0.1	0.14	1.76E+05	48	8.2	6.64

Supplementary Table 11: Uptake rates (mmol gDW^{-1} h^{-1}) implemented to simulate DMEM 6429 medium.

Metabolite ID	Exchange reaction ID	Metabolite name	Uptake rate
Amino acids			
ala_L	EX_ala_L(e)	L-alanine	1
arg_L	EX_arg_L(e)	L-arginine	1
asn_L	EX_asn_L(e)	L-asparagine	1
asp_L	EX_asp_L(e)	L-aspartate	1
cys_L	EX_cys_L(e)	L-cysteine	1
gln_L	EX_gln_L(e)	L-glutamine	1
glu_L	EX_glu_L(e)	L-glutamate	1
gly	EX_gly(e)	Glycine	1
his_L	EX_his_L(e)	L-histidine	1
ile_L	EX_ile_L(e)	L-isoleucine	1
leu_L	EX_leu_L(e)	L-leucine	1
lys_L	EX_lys_L(e)	L-lysine	1
met_L	EX_met_L(e)	L-methionine	1
phe_L	EX_phe_L(e)	L-phenylalanine	1
pro_L	EX_pro_L(e)	L-proline	1
ser_L	EX_ser_L(e)	L-serine	1
thr_L	EX_thr_L(e)	L-threonine	1
trp_L	EX_trp_L(e)	L-tryptophan	1
tyr_L	EX_tyr_L(e)	L-tyrosine	1
val_L	EX_val_L(e)	L-valine	1
Carbon sources			
glc_D	EX_glc(e)	Glucose	4.5
pyr	EX_pyr(e)	Pyruvate	1
Minerals, vitamins, other			
ca2	EX_ca2(e)	Calcium(2+)	1
chol	EX_chol(e)	Choline	1
cl	EX_cl(e)	Chloride	1
cobalt2	EX_cobalt2(e)	Co2+	1
cu2	EX_cu2(e)	Cu2+	1
fe2	EX_fe2(e)	Fe2+	1
fe3	EX_fe3(e)	Fe3+	1
fol	EX_fol(e)	Folate	1
k	EX_k(e)	Potassium	1
h2o	EX_h2o(e)	Water	10
h2s	EX_h2s(e)	Hydrogen sulfide	1
inost	EX_inost(e)	Inositol	1

Metabolite ID	Exchange reaction ID	Metabolite name	Uptake rate
mg2	EX_mg2(e)	Magnesium	1
mn2	EX_mn2(e)	Manganese	1
ncam	EX_ncam(e)	Nicotinamide	1
pi	EX_pi(e)	Hydrogenphosphate	10
pnto_R	EX_pnto_R(e)	Pantothenate	1
pydxn	EX_pydxn(e)	Pyridoxine	1
ribflv	EX_ribflv(e)	Riboflavin	1
so4	EX_so4(e)	Sulfate	1
thm	EX_thm(e)	Thiamin	1
zn2	EX_zn2(e)	Zinc	1
Supplemented nutrients (based on experimental procedure, see Supplementary Note 8)			
mqn7	EX_mqn7(e)	Menaquinone 7	1
mqn8	EX_mqn8(e)	Menaquinone 8	1
pheme	EX_pheme(e)	Protoheme	1
arabinogal	EX_arabinogal(e)	Larch arabinogalactan	0.0094
Metabolites required *in silico*			
q8	EX_q8(e)	Ubiquinone-8	1
sheme	EX_sheme(e)	Siroheme	1

Supplementary Table 12: Uptake rates (mmol gDW^{-1} h^{-1}) for dietary compounds implemented to simulate Western and high fiber diet.

Metabolite ID	Exchange reaction ID	Metabolite name	Western Diet	High fiber Diet
Sugars				
arab_L	EX_arab_L(e)	L-arabinose	0.17878295	0.04736842
cellb	EX_cellb(e)	Cellobiose	0.07449289	0.01973684
drib	EX_drib(e)	2-deoxy-D-ribose	0.17878295	0.04736842
fru	EX_fru(e)	D-Fructose	0.14898579	0.03947368
fuc_L	EX_fuc_L(e)	L-fucose	0.14898579	0.03947368
gal	EX_gal(e)	D-Galactose	0.14898579	0.03947368
glc_D	EX_glc(e)	D-glucose	0.14898579	0.03947368
glcn	EX_glcn(e)	D-gluconate	0.14898579	0.03947368
lcts	EX_lcts(e)	Lactose	0.07449289	0.01973684
malt	EX_malt(e)	Maltose	0.07449289	0.01973684
man	EX_man(e)	D-Mannose	0.14898579	0.03947368
melib	EX_melib(e)	Melibiose	0.07449289	0.01973684
mnl	EX_mnl(e)	D-Mannitol	0.14898579	0.03947368
oxa	EX_oxa(e)	Oxalate(2-)	0.44695737	0.11842105
rib_D	EX_rib_D(e)	D-ribose	0.17878295	0.04736842
rmn	EX_rmn(e)	L-Rhamnose	0.14898579	0.03947368
sucr	EX_sucr(e)	Sucrose	0.07449289	0.01973684
tre	EX_tre(e)	Trehalose	0.07449289	0.01973684
xyl_D	EX_xyl_D(e)	D-xylose	0.17878295	0.04736842
strch1	EX_strch1(e)	Starch	0.25733909	0.06818182
Fiber				
amylopect900	EX_amylopect900(e)	Amylopectin	0.00001567	0.00034722
amylose300	EX_amylose300(e)	Amylose	0.00004702	0.00104167
arabinan101	EX_arabinan101(e)	Arabinan	0.00016628	0.00368369
arabinogal	EX_arabinogal(e)	Larch arabinogalactan	0.00002191	0.00048550
arabinoxyl	EX_arabinoxyl(e)	Arabinoxylan	0.00030665	0.00679348
bglc	EX_bglc(e)	Beta-glucan	0.00000007	0.00000156
cellul	EX_cellul(e)	Cellulose	0.00002821	0.00062500
dextran40	EX_dextran40(e)	Dextran 40, 1,6-alpha-D-Glucan	0.00017632	0.00390625
galmannan	EX_galmannan(e)	Carob galactomannan	0.00001411	0.00031250
glcmannan	EX_glcmannan(e)	Konjac glucomannan	0.00003288	0.00072844
homogal	EX_homogal(e)	Homogalacturonan	0.00012823	0.00284091
inulin	EX_inulin(e)	Chicory inulin	0.00047019	0.01041667
kestopt	EX_kestopt(e)	Kestopentaose	0.00282117	0.06250000
levan1000	EX_levan1000(e)	Levan, 1000 fructose units	0.00001411	0.00031250

Metabolite ID	Exchange reaction ID	Metabolite name	Western Diet	High fiber Diet
lichn	EX_lichn(e)	Lichenin from Icelandic moss	0.00008298	0.00183824
lmn30	EX_lmn30(e)	Laminarin	0.00047019	0.01041667
pect	EX_pect(e)	Pectin	0.00003339	0.00073964
pullulan1200	EX_pullulan1200(e)	Pullulan	0.00001175	0.00026042
raffin	EX_raffin(e)	Raffinose	0.00470194	0.10416667
rhamnogalurI	EX_rhamnogalurI(e)	Potato rhamnogalacturonan I	0.00001449	0.00032106
rhamnogalurII	EX_rhamnogalurII(e)	Wine rhamnogalacturonan II	0.00026699	0.00591483
starch1200	EX_starch1200(e)	Resistant starch	0.00001175	0.00026042
xylan	EX_xylan(e)	Oat spelt xylan	0.00003206	0.00071023
xyluglc	EX_xyluglc(e)	Xyluglucan	0.00001315	0.00029124
Fat				
arachd	EX_arachd(e)	Arachidonate	0.00332813	0.00166406
chsterol	EX_chsterol(e)	Cholesterol	0.00495795	0.00247898
glyc	EX_glyc(e)	Glycerol	1.79965486	0.89982743
hdca	EX_hdca(e)	Hexadecanoate (n-C16:0)	0.39637090	0.19818545
hdcea	EX_hdcea(e)	Hexadecenoate (n-C16:1)	0.03651697	0.01825848
lnlc	EX_lnlc(e)	Linoleate	0.35910921	0.17955461
lnlnca	EX_lnlnca(e)	Alpha-linolenate	0.01756512	0.00878256
lnlncg	EX_lnlncg(e)	Gamma-linolenate	0.01756512	0.00878256
ocdca	EX_ocdca(e)	Octadecanoate (n-C18:0)	0.16928260	0.08464130
ocdcea	EX_ocdcea(e)	Octadecenoate (n-C18:1)	0.68144465	0.34072233
octa	EX_octa(e)	Octanoate (n-C8:0)	0.01294272	0.00647136
ttdca	EX_ttdca(e)	Tetradecanoate (n-C14:0)	0.06867567	0.03433784
Protein				
ala_L	EX_ala_L(e)	L-alanine	1	1
arg_L	EX_arg_L(e)	L-arginine	0.15	0.15
asn_L	EX_asn_L(e)	L-asparagine	0.225	0.225
asp_L	EX_asp_L(e)	L-aspartate	0.225	0.225
cys_L	EX_cys_L(e)	L-cysteine	1	1
gln_L	EX_gln_L(e)	L-glutamine	0.18	0.18
glu_L	EX_glu_L(e)	L-glutamate	0.18	0.18
gly	EX_gly(e)	Glycine	0.45	0.45
his_L	EX_his_L(e)	L-histidine	0.15	0.15
ile_L	EX_ile_L(e)	L-isoleucine	0.15	0.15
leu_L	EX_leu_L(e)	L-leucine	0.15	0.15
lys_L	EX_lys_L(e)	L-lysine	0.15	0.15
met_L	EX_met_L(e)	L-methionine	0.18	0.18

Metabolite ID	Exchange reaction ID	Metabolite name	Western Diet	High fiber Diet
phe_L	EX_phe_L(e)	L-phenylalanine	1	1
pro_L	EX_pro_L(e)	L-proline	0.18	0.18
ser_L	EX_ser_L(e)	L-serine	1	1
thr_L	EX_thr_L(e)	L-threonine	0.225	0.225
trp_L	EX_trp_L(e)	L-tryptophan	0.08181818	0.08181818
tyr_L	EX_tyr_L(e)	L-tyrosine	1	1
val_L	EX_val_L(e)	L-valine	0.18	0.18
Minerals, vitamins, other				
12dgr180	EX_12dgr180(e)	1,2-Diacyl-sn-glycerol (dioctadecanoyl, n-C18:0)	1	1
26dap_M	EX_26dap_M(e)	meso-2,6-Diaminoheptanedioate	1	1
2dmmq8	EX_2dmmq8(e)	2-Demethylmenaquinone 8	1	1
2obut	EX_2obut(e)	2-Oxobutanoate	1	1
3mop	EX_3mop(e)	3-methyl-2-oxopentanoate	1	1
4abz	EX_4abz(e)	4-Aminobenzoate	1	1
4hbz	EX_4hbz(e)	4-hydroxybenzoate	1	1
ac	EX_ac(e)	Acetate	1	1
acgam	EX_acgam(e)	N-acetyl-D-glucosamine	1	1
acmana	EX_acmana(e)	N-acetyl-D-mannosamine	1	1
acnam	EX_acnam(e)	N-acetylneuraminate	1	1
ade	EX_ade(e)	Adenine	1	1
adn	EX_adn(e)	Adenosine	1	1
adocbl	EX_adocbl(e)	Adenosylcobalamin	1	1
ala_D	EX_ala_D(e)	D-alanine	1	1
amp	EX_amp(e)	AMP	1	1
arab_D	EX_arab_D(e)	D-Arabinose	1	1
btn	EX_btn(e)	Biotin	1	1
ca2	EX_ca2(e)	Calcium(2+)	1	1
cbl1	EX_cbl1(e)	Cob(I)alamin	1	1
cgly	EX_cgly(e)	L-cysteinylglycine	1	1
chol	EX_chol(e)	Choline	1	1
chor	EX_chor(e)	Chorismate	1	1
cit	EX_cit(e)	Citrate	1	1
cl	EX_cl(e)	Chloride	1	1
cobalt2	EX_cobalt2(e)	Co2+	1	1
csn	EX_csn(e)	Cytosine	1	1
cu2	EX_cu2(e)	Cu2+	1	1

Metabolite ID	Exchange reaction ID	Metabolite name	Western Diet	High fiber Diet
dad_2	EX_dad_2(e)	2-deoxyadenosine	1	1
dcyt	EX_dcyt(e)	Deoxycytidine	1	1
ddca	EX_ddca(e)	Laurate	1	1
dgsn	EX_dgsn(e)	Deoxyguanosine	1	1
fe2	EX_fe2(e)	Fe2+	1	1
fe3	EX_fe3(e)	Fe3+	1	1
fe3dcit	EX_fe3dcit(e)	Fe(III)dicitrate	1	1
fald	EX_fald(e)	Formaldehyde	1	1
fol	EX_fol(e)	Folate	1	1
for	EX_for(e)	Formate	1	1
fum	EX_fum(e)	Fumarate	1	1
gam	EX_gam(e)	D-Glucosamine	1	1
glu_D	EX_glu_D(e)	D-Glutamate	1	1
glyc3p	EX_glyc3p(e)	Glycerol 3-phosphate	1	1
gthox	EX_gthox(e)	Oxidized glutathione	1	1
gthrd	EX_gthrd(e)	Reduced glutathione	1	1
gua	EX_gua(e)	Guanine	1	1
h	EX_h(e)	Proton	1	1
H2	EX_h2(e)	Hydrogen	1	1
h2o	EX_h2o(e)	Water	10	10
h2s	EX_h2s(e)	Hydrogen sulfide	1	1
hxan	EX_hxan(e)	Hypoxanthine	1	1
indole	EX_indole(e)	Indole	1	1
k	EX_k(e)	Potassium	1	1
lanost	EX_lanost(e)	lanosterol	1	1
meoh	EX_meoh(e)	Methanol	10	10
metsox_S_L	EX_metsox_S_L(e)	L-Methionine Sulfoxide	1	1
mg2	EX_mg2(e)	Magnesium	1	1
mn2	EX_mn2(e)	Mn2+	1	1
mobd	EX_mobd(e)	Molybdate	1	1
mqn7	EX_mqn7(e)	Menaquinone 7	1	1
mqn8	EX_mqn8(e)	Menaquinone 8	1	1
na1	EX_na1(e)	Sodium	1	1
nac	EX_nac(e)	Nicotinate	1	1
ncam	EX_ncam(e)	Nicotinamide	1	1
nmn	EX_nmn(e)	NMN	1	1
no2	EX_no2(e)	Nitrite	1	1
no3	EX_no3(e)	Nitrate	1	1
orn	EX_orn(e)	Ornithine	1	1
pheme	EX_pheme(e)	Protoheme	1	1

Metabolite ID	Exchange reaction ID	Metabolite name	Western Diet	High fiber Diet
pi	EX_pi(e)	Hydrogenphosphate	1	1
pime	EX_pime(e)	Pimelate	1	1
pnto_R	EX_pnto_R(e)	(R)-Pantothenate	1	1
ptrc	EX_ptrc(e)	Putrescine	1	1
pydam	EX_pydam(e)	Pyridoxamine	1	1
pydx	EX_pydx(e)	Pyridoxal	1	1
pydx5p	EX_pydx5p(e)	Pyridoxal 5-phosphate	1	1
pydxn	EX_pydxn(e)	Pyridoxine	1	1
q8	EX_q8(e)	Ubiquinone-8	1	1
ribflv	EX_ribflv(e)	Riboflavin	1	1
sel	EX_sel(e)	Selenate	1	1
sheme	EX_sheme(e)	Siroheme	1	1
so4	EX_so4(e)	Sulfate	1	1
spmd	EX_spmd(e)	Spermidine	1	1
thm	EX_thm(e)	Thiamin	1	1
thymd	EX_thymd(e)	Thymidine	1	1
ura	EX_ura(e)	Uracil	1	1
uri	EX_uri(e)	Uridine	1	1
xan	EX_xan(e)	Xanthine	1	1
zn2	EX_zn2(e)	Zinc	1	1

Supplementary Table 17: Reactions associated with functional roles of the eight different B-vitamin biosynthesis pathways. The functional roles involved in the eight B-vitamin biosynthesis pathways are the same as described in a study by Magnusdottir et al. [14].

B-vitamin	Functional role	VMH Reaction(s)
Biotin	BioH / BioG	PMACPME
	BioW	EX_pime(e), PIMEtr
	BioC	ACS, ACCOAC, MACPMT, MALCOACD, 3OAACPR1, 3HACPR1, EACPR1, GACPCD, 3OAACPR2, 3HACPR2, EACPR2
	BioF	AOXSr2
	BioA	AMAOTr
	BioD	DBTS
	BioB	BTS4
Cobalamin	CysG / CbiKX	SHCHCC
	CbiL	CPC2MT
	CbiG	CPC3MT
	CbiF	CPC4MT
	CobF	CPC5MT
	CbiJ	CPC6R
	CbiE / CbiT	CPC6MT
	CbiC	CPC8MM
	CbiA	CYRDAS
	CobAT	CYRDAR, CYRDAAT
	CbiP	ADCYRS
	CbiB	ADCPS2
	CobU	ACBIPGT
	CobS	ADOCBLS
Folate	FolE1 / FolE2	GTPCI
	folQ2 / folQ3	DNTPPA
	FolB / ptpS-III	DHNPA2
	FolK	HPPK2
	FolP	DHPS2
	pabAc / pabAb	ADCS
	pabAa	ADCL
	FolCDHFS	DHFS
	FolCFPGS / FolC2	FPGS
	Dhfr0 / Dhfr1 / Dhfr2	DHFR
Niacin	ASPOX	ASPO7
	ASPDH	ASPO2
	QSYN	QULNS
	QAPRT	NNDPR
	NaMNAT_D	NNATr
	NADS	NADS1

	NMNS	R0527
	NMNAT / NMNAT_R / NMNAT_M	NMNAT
	NADK	NADK
Pantothenate	ASPDC	ASP1DC
	KPHMT	MOHMT
	KPRED / KARED	DPR
	PBAL	PANTS
	PANK / PANK2 / PANK3	PNTK
	PPCS	PPNCL3
	PPCDC	PPCDC
	PPAT	PTPAT
	DPCK	DPCOAK
Pyridoxine	PdxT / PdxS	PLPS
	Dxs	DXPS
	gapA	E4PD
	PdxB	PERD
	PdxF / PdxA	OHPBAT
	PdxJ	PDX5PS
	PdxH	PDX5PO2, PYAM5POr
Riboflavin	GTPCH2	GTPCII2
	PyrD / PyrD-a	DHPPDA2
	PyrR	APRAUR
	PyrP	PMDPHT
	DHBPS	DB4PS
	DMRLS	RBFSa
	RSAe / RSAalpha	RBFSb
	RK	RBFK
	FMNAT	FMNAT
Thiamin	ThiG	THZPSN
	ThiC	AMPMS2
	ThiD / ThiD_alt	PMPK, HMPK1
	TMP-Pase(ThiE)	TMPPP

Supplementary Table 20: GC-MS dwell times and quantification and qualification ions (m/z) for the measured short-chain fatty acids and amino acids.

Derivatives	Quant-Ion (m/z)	Qual-Ion (m/z)	Qual-Ion (m/z)	Dwell Time (ms)
Short-chain fatty acids				
Formic acid 1TBDMS	103.0	75.0	99.0	20
Acetic acid 1TBDMS	117.0	75.0	99.0	20
Butyric acid 1TBDMS	145.1	75.0	115.1	20
Isobutyric acid 1TBDMS	145.1	75.0	115.1	20
Valeric acid 1TBDMS	159.1	75.0	201.1	20
IS 2-Ethylbutyric acid 1TBDMS	173.1	75.0	115.1	20
4-Methylvaleric acid 1TBDMS	173.1	75.0	215.1	20
Hexanoic acid 1TBDMS	173.1	75.0	131.0	20
Polar metabolites				
Alanine 2TBDMS	260.2	158.1	232.1	70
Glycine 2TBDMS	246.1	189.1	218.1	50
Valine 2TBDMS	288.6	186.1	260.2	70
Leucine 2TBDMS	302.2	200.2	274.2	70
Isoleucine 2TBDMS	302.2	200.2	274.2	70
Threonine 2TBDMS	290.2	159.1	303.2	50
Proline 2TBDMS	286.2	184.1	258.2	50
Succinic acid 2TBDMS	289.1	215.1	331.2	50
IS Glutaric acid-D6 2TBDMS	309.2	235.2	351.3	70
Serine 3TBDMS	390.2	302.2	362.2	50
Threonine 3TBDMS	404.2	376.3	417.3	70
Methionine 2TBDMS	320.2	218.1	292.2	50
Malic acid 3TBDMS	419.2	287.1	403.2	50
Phenylalanine 2TBDMS	336.2	308.2	combined	50
Aspartic acid 3TBDMS	418.2	316.2	(coeluting)	
Ornithine 3TBDMS	474.4	184.1	286.2	50
Glutamic acid_3TBDMS	432.3	330.2	404.3	50
Lysine 3TBDMS	431.3	300.2	488.4	70
Tyrosine 3TBDMS	466.3	302.2	438.3	50
Histidine 3TBDMS	440.3	338.3	196.1	70

Supplementary References

1. Noor, E., Haraldsdottir, H.S., Milo, R. & Fleming, R.M. Consistent estimation of Gibbs energy using component contributions. *PLoS Comput Biol* **9**, e1003098 (2013).
2. Fallingborg, J. Intraluminal pH of the human gastrointestinal tract. *Danish medical bulletin* **46**, 183-196 (1999).
3. Heinken, A. & Thiele, I. Systematic prediction of health-relevant human-microbial co-metabolism through a computational framework. *Gut Microbes* (2015).
4. Heinken, A., Sahoo, S., Fleming, R.M. & Thiele, I. Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut microbes* **4**, 28-40 (2013).
5. Pastink, M.I. et al. Genome-scale model of Streptococcus thermophilus LMG18311 for metabolic comparison of lactic acid bacteria. *Appl Environ Microbiol* **75**, 3627-3633 (2009).
6. Teusink, B. et al. Analysis of growth of Lactobacillus plantarum WCFS1 on a complex medium using a genome-scale metabolic model. *The Journal of biological chemistry* **281**, 40041-40048 (2006).
7. Thiele, I. et al. A community effort towards a knowledge-base and mathematical model of the human pathogen Salmonella Typhimurium LT2. *BMC systems biology* **5**, 8 (2011).
8. Thiele, I., Vo, T.D., Price, N.D. & Palsson, B.O. Expanded metabolic reconstruction of Helicobacter pylori (iIT341 GSM/GPR): an in silico genome-scale characterization of single- and double-deletion mutants. *Journal of bacteriology* **187**, 5818-5830 (2005).
9. Baumler, D.J., Peplinski, R.G., Reed, J.L., Glasner, J.D. & Perna, N.T. The evolution of metabolic networks of E. coli. *BMC systems biology* **5**, 182 (2011).
10. Liao, Y.C. et al. An experimentally validated genome-scale metabolic reconstruction of Klebsiella pneumoniae MGH 78578, iYL1228. *Journal of bacteriology* **193**, 1710-1717 (2011).
11. Thiele, I. et al. A community-driven global reconstruction of human metabolism. *Nature biotechnology* **31**, 419-425 (2013).
12. Feist, A.M. et al. A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular systems biology* **3**, 121 (2007).
13. Thiele, I. & Palsson, B.O. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature protocols* **5**, 93-121 (2010).
14. Magnusdottir, S., Ravcheev, D., de Crecy-Lagard, V. & Thiele, I. Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. *Front Genet* **6**, 148 (2015).
15. Macfarlane, G.T. & Macfarlane, S. Bacteria, colonic fermentation, and gastrointestinal health. *Journal of AOAC International* **95**, 50-60 (2012).
16. Krieg, N.R. Bergey's manual of systematic bacteriology, 2nd ed., volume 4. (Springer, New York; 2010).
17. Rey, F.E. et al. Dissecting the in vivo metabolic potential of two human gut acetogens. *The Journal of biological chemistry* **285**, 22082-22090 (2010).
18. Reichardt, N. et al. Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *The ISME journal* **8**, 1323-1335 (2014).
19. Vital, M., Howe, A.C. & Tiedje, J.M. Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. *mBio* **5**, e00889 (2014).
20. Samuel, B.S. et al. Genomic and metabolic adaptations of Methanobrevibacter smithii to the human gut. *P Natl Acad Sci USA* **104**, 10643-10648 (2007).
21. Rey, F.E. et al. Metabolic niche of a prominent sulfate-reducing human gut bacterium. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 13582-13587 (2013).

22. Davila, A.M. et al. Re-print of "Intestinal luminal nitrogen metabolism: role of the gut microbiota and consequences for the host". *Pharmacological research* **69**, 114-126 (2013).
23. Flint, H.J. The impact of nutrition on the human microbiome. *Nutrition reviews* **70 Suppl 1**, S10-13 (2012).
24. Flint, H.J., Scott, K.P., Duncan, S.H., Louis, P. & Forano, E. Microbial degradation of complex carbohydrates in the gut. *Gut microbes* **3**, 289-306 (2012).
25. Huynen, M.A., Dandekar, T. & Bork, P. Variation and evolution of the citric-acid cycle: a genomic perspective. *Trends in microbiology* **7**, 281-291 (1999).
26. Ravcheev, D.A. & Thiele, I. Systematic genomic analysis reveals the complementary aerobic and anaerobic respiration capacities of the human gut microbiota. *Front Microbiol* **5**, 674 (2014).
27. Nowicka, B. & Kruk, J. Occurrence, biosynthesis and function of isoprenoid quinones. *Biochim Biophys Acta* **1797**, 1587-1605 (2010).
28. Wissenbach, U., Ternes, D. & Unden, G. An Escherichia coli mutant containing only demethylmenaquinone, but no menaquinone: effects on fumarate, dimethylsulfoxide, trimethylamine N-oxide and nitrate respiration. *Archives of microbiology* **158**, 68-73 (1992).
29. Zhi, X.Y. et al. The futalosine pathway played an important role in menaquinone biosynthesis during early prokaryote evolution. *Genome Biol Evol* **6**, 149-160 (2014).
30. Teusink, B. et al. In silico reconstruction of the metabolic pathways of Lactobacillus plantarum: comparing predictions of nutrient requirements with those from growth experiments. *Appl Environ Microbiol* **71**, 7253-7262 (2005).
31. Moreau, N.M. et al. Simultaneous measurement of plasma concentrations and 13C-enrichment of short-chain fatty acids, lactic acid and ketone bodies by gas chromatography coupled to mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* **784**, 395-403 (2003).
32. Hiller, K. et al. MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis. *Analytical chemistry* **81**, 3429-3439 (2009).
33. Aziz, R.K. et al. SEED servers: high-performance access to the SEED genomes, annotations, and metabolic models. *PLoS One* **7**, e48053 (2012).
34. Heinken, A. et al. A functional metabolic map of Faecalibacterium prausnitzii, a beneficial human gut microbe. *J Bacteriol* (2014).
35. Flahaut, N.A. et al. Genome-scale metabolic model for Lactococcus lactis MG1363 and its application to the analysis of flavor formation. *Applied microbiology and biotechnology* **97**, 8729-8739 (2013).