Amino acid auxotrophies in human gut bacteria are linked to higher microbiome diversity and long-term stability

Svenja Busche[1], Danielle MM Harris[1,2], Johannes Zimmermann[3,4], Sven Schuchardt[5], Mhmd Oumari[2], Derk Frank[6,7], Corinna Bang[2], Philip Rosenstiel[2], Stefan Schreiber[2,8], Norbert Frey[6,7,#], Andre Franke[2], Konrad Aden[2], Silvio Waschina[1,*]

[1] Institute of Human Nutrition and Food Science, Nutriinformatics, Kiel University, Kiel, Germany

[2] Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany

[3] Research Group Evolutionary Ecology and Genetics, Zoological Institute, University of Kiel, Kiel, Germany

[4] Max Planck Institute for Evolutionary Biology, Plön, Germany

[5] Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM), Hanover, Germany

[6] Department of Internal Medicine III, University Medical Center Schleswig-Holstein, Kiel, Germany

[7] German Centre for Cardiovascular Research (DZHK), Partner site Hamburg, Kiel, Lübeck, Germany

[8] Department of Internal Medicine I, University Medical Center Schleswig-Holstein, Kiel, Germany

[#] Current Affiliation: Department of Internal Medicine III, University Hospital Heidelberg, Heidelberg, Germany

*Correspondence: s.waschina@nutrinf.uni-kiel.de

Authors' ORCIDs:

SB (0000-0002-8089-9700); **DH** (0000-0002-1950-399X); **JZ** (0000-0002-5041-1954); **SvS** (0000-0002-4593-9324); **MO** (0009-0008-4977-7888); **DF** (0000-0001-7561-075X); **CB** (0000-0001-6814-6151); **PR** (0000-0002-9692-8828); **StS** (0000-0003-2254-7771); **NF** (0000-0001-7611-378X); **AF** (0000-0003-1530-5811); **KA** (0000-0003-3482-7316); **SW** (0000-0002-6290-3593)

Abstract

Amino acid auxotrophies are prevalent among bacteria. They can govern ecological dynamics in microbial communities and indicate metabolic cross-feeding interactions among coexisting genotypes. Despite the ecological importance of auxotrophies, their distribution and impact on the diversity and function of the human gut microbiome remain poorly understood. This study performed the first systematic analysis of the distribution of amino acid auxotrophies in the human gut microbiome using a combined metabolomic, metagenomic, and metabolic modeling approach. Results showed that amino acid auxotrophies are ubiquitous in the colon microbiome, with tryptophan auxotrophy being the most common. Auxotrophy frequencies were found to be higher for those amino acids, that are also essential to the human host. Moreover, a higher overall abundance of auxotrophies was associated with greater microbiome diversity and stability, and the distribution of auxotrophs was found to be related to the human host's metabolome, including trimethylamine oxide, small aromatic acids, and secondary bile acids. Thus, our results suggest that amino acid auxotrophies are important factors that contribute to microbiome ecology and host-microbiome metabolic interactions.

Background

The metabolic processes performed by the human gut microbiota have a crucial impact on human metabolism and health(1–3). For instance, the short-chain fatty acid butyrate is produced by various human gut bacteria. Butyrate is a primary energy source for human colonocytes(1) and intersects with host-immunological processes by mediating anti-inflammatory effects(4,5). Another notable metabolic interaction between the human host and its gastrointestinal microbiota is the microbial transformation of aromatic amino acids into various metabolites. Recent studies suggest that aromatic amino acid-derived metabolites such as the auxins indole-3-propionic acid and indole-3-acetic acid can modulate the host immune system(6,7). Thus, these and several further studies provide evidence that gut microbial metabolites are essential factors in the pathophysiology of inflammatory diseases and the efficacy of immunomodulatory therapies(7–10).

The compounds synthesized by gut microbes are typically metabolic by-products that serve the dual purpose of energy metabolism and facilitating the biosynthesis of essential metabolites, necessary for cellular maintenance and proliferation. However, often not all metabolites required for growth and survival (i.e., nucleotides, vitamins, amino acids) can be *de-novo* synthesized by gut-dwelling microorganisms, rendering those organisms dependent (termed *auxotrophic*) on the uptake of the focal metabolite from the microbial cell's nutritional environment. Several *in-silico* studies have applied genome-mining approaches, suggesting that most analyzed gut bacteria lack biosynthetic pathways for producing at least one proteinogenic amino acid(11,12) or a growth-essential vitamin(13,14). In addition, *in-vitro* growth experiments have confirmed specific amino acid and vitamin auxotrophies in common human gut bacteria(13,15,16).

The prevalence of auxotrophs in the human gut microbiome raises the question of the source of the required metabolites in the gastrointestinal growth environment. In general, there are three potential sources of growth factors for microbial growth: (i) Required metabolites could be diet-derived. However, amino acids and vitamins are usually efficiently absorbed by the human host in the small intestine(17), limiting the accessibility of diet-derived growth factors for the majority of the gut microbial community, which resides in the

colonic region(18). (ii) Metabolites required by auxotrophic microorganisms in the gastrointestinal tract might be host-derived, e.g., from proteins and peptides released by- or attached to the gut epithelium(15). (iii) Auxotrophic members of the gut microbial community might obtain the required growth factor via cross-feeding interactions with prototrophic organisms in their vicinity(19).

While the exchange of electron donor metabolites (e.g., acetate- or lactate cross-feeding) between different microorganisms is well-documented for the human gut microbiome(20–22), the extent of cross-feeding interactions via the exchange of growth factors such as amino acids and vitamins remains still unknown. However, in vitro-experiments of synthetic microbial communities suggest that co-cultured microorganisms, which are auxotrophic for different compounds, can support each other's growth through the exchange of the focal metabolites(23). Furthermore, theoretical ecological models suggest that cross-feeding interactions between auxotrophic organisms within complex communities can increase community diversity through metabolic niche expansion(24) and community robustness to ecological perturbance(25) such as changes in the composition of the chemical environment. Thus, cross-feeding of amino acids and vitamins between different members of the human gut microbiota could be crucial determinants of microbiome dynamics, resilience, and the contribution of gut microbes to human metabolism and health.

In this study, we applied genome-scale metabolic modeling to predict the distribution and diversity of amino acid auxotrophies in the human gut microbiome. The predictions were combined with stool metagenomics sequencing and targeted serum metabolomics from observational human cohort studies to estimate auxotrophy frequencies and their impact on the human metabolome. We found that amino acids that are essential to the human host are also the most common auxotrophies in the human gut microbiome. Intriguingly, a higher frequency of auxotrophies was associated with long-term stability of the microbiome community composition. Furthermore, a higher number of auxotrophies among gut bacteria were found to be associated with higher diversity of the gut bacteria and increased levels of aromatic compounds of putative microbial origin in the human serum metabolome.

Results

Amino acid auxotrophies are common in the human gut microbiome

In order to estimate the overall distribution of amino acid auxotrophies in the human gut microbiome, we predicted the amino acid production capacities using genome-metabolic modeling for all bacterial genomes (n=5,414) from the 'Human Reference Gut Microbiome (HRGM)' collection(26). A Spearman correlation showed a negative relationship between genome completeness and the number of auxotrophies per genome (Supplementary Figure S1, ρ=-0.35, $p \leq 2.2e-16$). To combat this, the genomes were filtered for a completeness ≥85% and contamination ≤2%. For the prediction of auxotrophies and ongoing analysis only the filtered metabolic models (n=3,687) were used.

Auxotrophies for tryptophan were the most prevalent, at 64.1% of the genomes in the HRGM catalog, followed by histidine at 52.2% (Fig. 2). Isoleucine, leucine, and valine (BCAA, branched chain amino acids) auxotrophies were also detected with a high abundance (39.5%, 40.7%, 38.6%, respectively) in all genomes. No auxotrophies were detected for alanine, aspartate, and glutamate. Furthermore, to reveal auxotrophic co-occurrence, we studied the occurrence of paired auxotrophies in all genomes: auxotrophies for isoleucine, leucine, and valine occurred together with each other and with tryptophan at the highest frequencies (Fig. 3A). The Rasch sampler was used to test if the predicted co-occurrence of amino acid auxotrophies is significantly higher or lower than expected by chance. The auxotrophy pairs of tryptophan and BCAAs, as well as within the BCAA group were observed significantly more often than expected (re-sampling test, FDR-corrected $p<0.05$, Fig. 3B). We further analyzed the observed auxotrophies at the taxonomy level by comparing the proportion and number of auxotrophies on phylum and order level (Supplementary Figure S2). Actinobacteriota were shown to have a higher proportion of histidine and BCAA auxotrophies compared to prototrophies (Supplementary Figure S2). For tryptophan, a higher proportion of auxotrophic to prototrophic bacteria were observed in Firmicutes, Actinobacteriota, and Bacteroidota. Fusobacteriota have a higher auxotrophic to prototrophic ratio for almost all amino acids, whereas for Proteobacteria the opposite was predicted. This observation is further supported by the number of auxotrophies found per genome for Proteobacteria and Fusobacteriota (Supplementary Figure S3). Further, the

results suggest that auxotrophic genotypes have lost the genes for most of the enzymes that would be required for the biosynthesis of the focal amino acid (Supplementary Figure S4).
A recent study has reported discrepancies between *in silico* predictions using metabolic models reconstructed with *carveme* (27) and *in-vitro* studies of amino acid auxotrophies in bacteria(13). In order to validate our *gapseq*-based auxotrophy predictions, we compared the predictions with *in-vitro* experimentally verified auxotrophies as reported in previous studies (Supplementary Table S2). In total, we found primary literature information for experimentally verified amino acid auxotrophies for six bacterial isolates from the human gut microbiome. Among those six isolates, 14 out of 18 (78%) auxotrophies were correctly predicted by our method. Out of 102 prototrophies, 97 (95%) were correctly predicted. In total, the agreement between predictions and experimental auxotrophy data was statistically significant (Fisher's exact test for count data, $p = 4.4e-11$). In addition, we reconstructed genome-scale metabolic models for 124 bacterial genomes that are known to be prototrophic for all 20 proteinogenic amino acids to further validate our auxotrophy predictions (27). In total, 2413/2480 (97.3%) of all predictions coincided with the known amino acid prototrophies of the organisms. Among the 20 amino acids, 17 amino acids prototrophies were correctly predicted for ≥96% of the genomes (Supplementary Figure S5). Higher frequencies of false-positive auxotrophy predictions were only observed for methionine (18% false positive rate), leucine (10%), and histidine (10%). In general, the frequency of auxotrophy predictions among genomes from human gut bacteria is generally higher compared to the collection of 124 prototrophic genomes (Supplementary Figure S5), indicating that the high frequency of auxotrophies cannot be explained by a false-positive rate associated to potential pitfalls in the model reconstruction workflow.
Taken together, the results indicate that amino acid auxotrophies are prevalent in the human gut microbiome.

Amino acid auxotrophies are associated with the profile of fermentation products

Amino acid biosynthesis pathways and fermentation by-product biosynthesis pathways share common precursors (Fig. 4B). For example, pyruvate is a central metabolite which is

utilized for the biosynthesis of the BCAA as well as lactate underlining the interconnection of the amino acid and fermentation by-product biosynthesis in the metabolic network. Here, we wanted to investigate whether bacteria, that are auxotrophic for specific amino acids, also display specific profiles of fermentation products. Therefore, we predicted the metabolic by-products of cell growth and compared those results with the auxotrophy predictions for the focal organisms (Fig. 4A). BCAA auxotrophic bacteria were more likely to produce lactate in comparison to prototrophic bacteria (Fisher's exact test for count data, $\log_2$(Odds Ratio (OR)) = 2.2-2.6, FDR-corrected p-value<0.05). Propionate production was commonly predicted for glutamine auxotrophic gut bacteria ($\log_2$(OR)= 3.0, FDR-corrected p-value <0.05) and by cysteine auxotrophs ($\log_2$(OR)= 1.9, FDR-corrected p-value <0.05). Succinate is predominantly produced by asparagine auxotrophic gut bacteria ($\log_2$(OR)= 1.9, FDR corrected p-value <0.05). For butyrate, there was a higher association with glutamine auxotrophic bacteria ($\log_2$(OR)= 1.3, FDR-corrected p-value <0.05).

The association of auxotrophic bacteria with the production of organic acids might be explained by in the distribution of reactions fluxes through the metabolic network. For instance, pyruvate is a metabolic precursor for the *de novo* biosynthesis pathways for BCAA but also for lactate formation (Fig. 4B). Pyruvate not used for BCAA biosynthesis in auxotrophic genotypes, might be redirected towards lactate production. Overall, our results indicate that pathways of energy metabolism and the formation of microbial fermentation products likely influence the evolution of amino acid auxotrophies in human gut bacteria.

More diverse gut microbiomes are characterized by a higher auxotrophy frequency

In order to estimate the frequency of auxotrophies in the gut microbiome of individual persons, we predicted auxotrophies of metagenome-assembled genomes (MAGs) from 185 metagenomes from adults (termed discovery cohort in this study) and took the MAGs' relative abundances into account. For this analysis, only high-quality MAGs with a completeness of ≥85% and contamination ≤2% were considered; which yielded in total 756 MAGs.

Strikingly, auxotrophies for amino acids that are essential to the human organism were more frequent than non-essential amino acids (Fig. 5A). The highest percentage of bacteria

were auxotrophic for tryptophan, followed by histidine and phenylalanine (Median: 50.8%, 32.1%, 33.7%, respectively). Auxotrophies for chorismate, BCAA, and proline were found with a median frequency of >25% (Fig. 5A). The lowest frequency was detected for serine, lysine, asparagine, aspartate, alanine, and glutamate auxotrophies.

Additionally, we were interested in the relationship between the proportion of auxotrophic bacteria in the human gut and the overall microbiome diversity calculated as the Shannon index (Fig. 5B-C). Overall, increasing frequencies of almost all amino acid auxotrophies are accompanied by increasing microbiome diversity (Spearman correlation, Fig. 5B). Further, we correlated the Shannon diversity with the abundance-weighted average of the number of auxotrophies per metagenome sample, which takes the relative abundance of each MAG and its total number of amino acid auxotrophies into account. With an increasing number of auxotrophies an increase in the diversity was observed (Fig. 5C, $\rho = 0.60$, $p < 2.2e-16$). This result may point towards a positive influence of auxotrophic bacteria on the microbial diversity in the gut, presumably via a higher degree of amino acid cross-feeding interactions between genotypes that are auxotrophic for different amino acids. To test this, we calculated the pairwise dissimilarity (Hamming distance) between the binary auxotrophy profiles of MAGs and the means of those differences per metagenome sample as an indicator for potential cross-feeding in the respective gut microbial community. An increasing average Hamming distance was positively associated with an increase in gut diversity (Fig. 5D, $\rho = 0.62$, $p < 2.2e-16$).

Overall, a higher number of auxotrophies in the gut community are positively correlated with a higher diversity. Especially auxotrophies for essential amino acids can be found in the human gut microbiome.

Associations of gut bacterial auxotrophies for amino acids with host health markers and the serum metabolome

The involvement of microbial metabolism in host health has been examined in several other studies (28,29) but not yet for frequency of gut microbial amino acid auxotrophies. Our results showed that several amino acid auxotrophic bacteria are inversely associated with the stool donor's BMI (Fig. 5B, partial Spearman correlation). No further statistically significant associations with health serum markers were found. Additionally, we correlated

targeted metabolomics data from serum samples with the frequencies of specific amino acid auxotrophies (Fig. 5E, partial Spearman correlation). Positive correlations were found between the tryptophan-derived 3-indolepropionic acid (3-IPA) as well as 3-indoleacetic acid (3-IAA) and tryptophan auxotrophic gut bacteria. Additionally, several other amino acid auxotrophies showed positive correlations with these metabolites. Serine and asparagine auxotrophic bacteria were positively associated with indoxyl sulfate (Ind-SO_4). P-cresol sulfate was positively correlated with many amino acid auxotrophies. Further, several significant associations were detected with metabolites from bile acid metabolism. Negative correlations were observed for glycoursodeoxycholic acid (GUDCA), a conjugated secondary bile acid metabolite, and several amino acid auxotrophies. Further, correlations with the bile acid metabolites glycodeoxycholic acid (GDCA) and deoxycholic acid (DCA) were found for the frequencies of tyrosine and cysteine auxotrophies as well as glycolithocholic acid sulphate (GLCAS) and proline auxotrophy frequency. Strong positive associations were also observed for hippuric acid and TMAO with several amino acid auxotrophies. Interestingly, no significant associations were found for serum levels of amino acids and amino acid related compounds (Fig. 5E).

Taken together, the frequency of auxotrophic bacteria is related to serum levels of several metabolites. The gut microbial contribution to so serum metabolite levels were predominantly found for metabolites, that were previously reported to be of microbial origin (e.g., 3-IPA, 3-IAA) or are derived from gut microbially-produced compounds (e.g., TMAO).

Analysis of longitudinal microbial composition data suggests a positive influence of auxotrophies on gut microbiome stability

So far, our results suggest an involvement of auxotrophic bacteria on the gut microbial diversity. Furthermore, higher alpha diversity was positively correlated with increased dissimilarity of the auxotrophy profiles of co-existing genotypes. Based on the observation that a more diverse gut microbiome was associated with the distribution of auxotrophies, we further wanted to analyze whether the frequency of auxotrophies also has an impact on long-term stability of the microbiome using data from a longitudinal study. Therefore, we re-analyzed recently published metagenomic data from a human cohort study that involves

two stool metagenomes from 72 healthy individuals each, which where three years apart(30). Microbiome stability over this three-year period was assessed by calculating the Bray-Curtis distance for the microbial composition between the two time points for each participant. The dissimilarity value, which is an inverse proxy for microbiome stability, was then tested for statistical association with the MAG abundance-weighted average number of amino acid auxotrophies per genotype at baseline (Fig. 6A). The abundance-weighted average of auxotrophies per genotype was significantly and negatively correlated with the Bray-Curtis distance (Fig. 6A, Spearman rank sum correlation test, ρ=-0.24, p=0.0403). Thus, communities with increasing frequencies of auxotrophies were characterized by higher long-term stability. Next, the Bray-Curtis distance between paired samples were also tested for a statistical association with the average Hamming distance with samples, which represents a measure of the dissimilarity between the auxotrophy profile of co-existing genotypes and an indicator for the degree of amino acid cross-feeding in the focal community. A significant and negative correlation was observed for the average Hamming distance per sample with the Bray-Curtis-Distance (Fig. 6B, ρ=-0.30, p=0.01072), indicating a positive influence of amino acid cross-feeding between auxotrophy genotypes with long-term stability of microbiome composition.

Auxotrophic bacteria have a high dependence on their nutritional environment. Therefore, we wanted to see if a higher dietary intake of amino acids affects the frequency of amino acid auxotrophic bacteria in the gut. Therefore, we made use of the dietary intake data obtained from food frequency questionnaires from the longitudinal cohort. The intake of amino acids at the beginning of the study was tested for correlation with the frequency of amino acid auxotrophies in the microbiomes. No significant correlations between the frequency of auxotrophic bacteria and the dietary intake of amino acids were observed (Supplementary Figure S6).

In sum, our results suggest a positive effect of auxotrophies on the gut microbiome stability. Further, the data suggests that amino acid cross-feeding may contribute to compositional stability of the gut microbiome. Surprisingly, we found no evidence for the effect of diet on auxotrophy frequencies as a higher dietary intake of amino acids was not significantly correlated with higher frequencies of auxotrophic bacteria.

Discussion

Auxotrophies are widespread among microorganisms (11,31). The obligate nutritional requirements can have far-reaching consequences for the auxotrophic strains and the entire microbial community in the ecosystem (32). On the one hand, each auxotrophy for a specific growth factor (e.g., amino acids) increases the organism's dependence on the nutritional environment, coupling the organism's survival and proliferation to the availability of the specific compound (32). On the other hand, if the focal metabolite is available, auxotrophic genotypes might gain a selective advantage over prototrophic genotypes by saving metabolic costs (33). In microbial communities, auxotrophies can affect the interactions between microorganisms and their hosts, where auxotrophs could act as *recyclers* of metabolites that other community members release as by-products of their metabolism (34). In addition, organisms that are auxotrophic for different metabolites could engage in cooperative cross-feeding interactions (35–37). Despite the ecological relevance of auxotrophies, their role in the human gut microbiome is largely unknown. More specifically, several auxotrophies were confirmed by laboratory cultivation experiments of isolates from the human gut microbiota (13,15). Still, the overall distribution and variation of auxotrophies in the human gut microbiome remain elusive. Here, we performed a systematic analysis of the distribution of amino acid auxotrophies in the human gut microbiome using genome-scale metabolic modeling. Moreover, we statistically assessed the associations of auxotrophy frequencies with overall microbiome diversity, long-term stability, and microbial contribution to the human metabolome.

Ubiquity of auxotrophies indicates high prevalence of cross-feeding

Overall, high frequencies of auxotrophies were found in the human gut microbiome. For instance, we found that 50.8% (median) of organisms in the gut microbial communities of healthy adults are auxotrophic for tryptophan (Fig. 5A). Interestingly, the most frequent auxotrophies for amino acids in the human gut microbiome are also essential nutrients for the human host (Fig. 5A). While auxotrophies in human gut bacteria were reported before, the sources of amino acids for auxotrophic genotypes remain unknown. There are three potential sources of amino acids of auxotrophic members of the gut microbiome:

First, amino acids might be acquired from dietary proteins (38). However, most of the diet-derived protein is broken down in the upper gastrointestinal tract and amino acids are absorbed by the human host, limiting protein and amino acid passage to the colon, where the majority of the gut microbiome resides (38). Our predictions are based on genomes from stool samples, which predominantly reflect the microbiome composition in the large intestine. Therefore, we argue that the high frequency of amino acid auxotrophies predicted for the colon microbiome in this study are unlikely to be explained by dietary sources of amino acids alone. Plus, we did not find any statistical associations of the dietary intake of amino acids of 75 adults and the frequency of auxotrophies in the microbiome (Supplementary Figure S6), which further indicates that auxotrophic genotypes acquire their amino acids also from other sources.

Second, auxotrophs might obtain their essential amino acids by engaging in cross-feeding interactions with prototrophic genotypes. Cross-feeding between strains that are auxotrophic for different amino acids has been demonstrated in synthetic (37) and naturally occurring microbial communities (31). Thus, cross-feeding enables the growth of auxotrophic organisms even in environments where the focal nutrient is not available. Our results suggest a wide diversity of auxotrophic profiles between coexisting genotypes (Fig. 5D), indicating metabolic complementarity and amino acid cross-feeding in gut microbial communities.

The third potential source of amino acids for auxotrophic gut microbes are host-derived metabolites. Yet, evidence reported in scientific literature for gut microbial uptake of host-derived amino acids is scarce(39,40). An interesting case where an auxotrophic gut bacterium covers its demand for the focal amino acid might be *Akkermansia muciniphila*. Our predictions show that this bacterium is auxotrophic for threonine, which is in agreement with previous cultivation experiments (15). *A. muciniphila* is a known degrader of host mucins and resides in the mucus layer. Besides glycans, mucin consists of a core protein scaffold that is rich in proline, threonine, and serine (41). Thus, the threonine auxotrophy of *A. muciniphila* may indicate that this species also utilizes host-derived threonine.

Auxotrophies are associated with alpha diversity

A major result of our study is the positive associations between auxotrophies and diversity of the human gut microbiome (Fig. 5C). Earlier studies that used theoretical approaches suggested that auxotrophies can increase and maintain diversity in microbial communities by creating niches for different organisms to occupy through metabolite cross-feeding (24,34). Thus, we conclude that in communities with more auxotrophic members, more cross-feeding may take place, which could promote diversity. Our results further support this theory, since we observed a positive association between microbiome diversity and the dissimilarity in auxotrophic profiles among coexisting genotypes (Fig. 5D).

Auxotrophies and microbiome stability

Microbe-microbe interactions via metabolite exchanges may promote diversity and contribute to microbiome stability (42). Here, we tested if having more auxotrophies as an indicator for metabolite cross-feeding in the gut microbiome is linked to greater stability in healthy adults over three years. Indeed, our findings indicate that microbiomes with a higher average frequency of auxotrophies at the beginning of the study period remained more stable throughout the duration of the study (Fig. 6A). The association of auxotrophies with microbiome stability was even more pronounced when considering the dissimilarity of auxotrophy profiles of coexisting genotypes as a proxy for amino acid cross-feeding (Fig. 6B). This result is in line with a theoretical study by Oña and Kost, which demonstrates that cross-feeding between auxotrophs can facilitate that the community structure returns to equilibrium after ecological perturbance (25). Moreover, Sharma *et al.* (2019) reported that B-vitamin auxotrophies in the human microbiome are prevalent and suggest that cross-feeding B-vitamins between prototrophic and auxotrophic genotypes contribute to gut bacterial population dynamics. The authors also base their conclusion on experimental results, where gnotobiotic mice were colonized by a human fecal microbial community. In these experiments, varying dietary B vitamin intake in mice did not result in appreciable changes in gut microbial community structure, including the proportion of B vitamin-auxotrophic subpopulations, which further suggests cross-feeding as a source of essential nutrients for auxotrophic bacteria in the gut environment and supports our hypothesis that a higher auxotrophy frequencies contribute to microbiome stability (Fig 6AB).

Since a reduction in gut microbiome diversity has been reported for several chronic diseases(43–45), our results and the methodology to predict auxotrophy frequencies may

guide the development of novel personalized treatment strategies by targeting ecological interactions between coexisting gut microorganisms. For instance, oral administration of microencapsulated amino acids with delayed content release could be used to specifically promote the growth of beneficial subpopulations of the large intestine microbial community, which are auxotrophic for the focal compound (46).

Auxotrophy associations with the human metabolome

Pathways of amino acid biosynthesis and fermentation by-product biosynthesis share common precursors (Fig. 4B). Therefore, the loss of biosynthetic genes for amino acids might affect the flux distribution in the metabolic network (33). Fermentation by-products such as the organic acids butyrate, acetate, and propionate have implications for human physiology (1). Hence, we wanted to investigate whether specific amino acid auxotrophies are associated with the profile of fermentation products released by gut bacteria. Comparison of the fermentation by-product profile of auxotrophic and prototrophic bacteria revealed statistically significant associations (Fig. 4A), which may be due to the structure of the metabolic network (Fig. 4B). For example, BCAA auxotrophic bacteria are more likely to be lactate producers, which might be attributed to the fact that the common precursor of BCAA synthesis and lactate synthesis, pyruvate, is no longer used for BCAA synthesis in BCAA auxotrophic bacteria but can be used for lactate formation. The altered fermentation profile in auxotrophic bacteria may therefore indicate the importance of the nutritional requirements of gut bacteria for the microbiome's contribution to the human metabolome.

Indeed, when we tested for associations of the relative abundance of amino acid auxotrophs with compounds of the human metabolome, we found several significant correlations (Fig. 5E). In particular, the frequencies of several auxotrophies were correlated with phenylic and indolic metabolites, namely hippuric acid, p-cresol sulfate, indoxyl sulfate, 3-indole acetic acid (IAA), and 3-indole propionic acid (IPA). These compounds were previously reported to be of microbial origin or are derived from gut microbially-produced metabolites (47). For instance, hippuric acid and p-cresol sulfate levels were reported to strongly correlate with the microbiome alpha diversity in a large human cohort study (48). Moreover, the tryptophan-derived IAA is a known agonist of the epithelial human aryl hydrocarbon receptor, an important regulator of intestinal immunity(49). p-cresol is known

to be produced by gut bacteria that metabolize tyrosine (50). In summary, our results suggest that the contribution of phenylic and indolic compounds to the human metabolome is linked to metabolic processes performed by amino acid auxotrophic gut bacteria.

Limitations

The method of our study is subject to certain limitations. In our study, auxotrophies were predicted with reconstructed genome-scale metabolic models. Discrepancies between metabolic modelling-based predictions and results from vitro assessments have been reported and discussed previously (13,27,51). Thus, it is crucial to validate *in silico* prediction with *in vitro* results of auxotrophies. Here we compared our *in-silico* results with *in vitro* results for gut bacterial strains and found a significant agreement of auxotrophy predictions and experimental data (Fisher's exact test for count data, p= 4.4e-11). However, it should be considered that it is still often difficult to cultivate human gut bacteria. Here, we were able to find data for experimentally determined auxotrophs for only six bacterial isolates from the human gut microbiome with a total of 19 amino acid auxotrophies (Supplementary Table S2). In addition, we performed auxotrophy prediction for 124 genomes from strains, that are not human gut bacteria but known from cultivation experiments to be prototrophic for all 20 proteinogenic amino acids. This test showed that 97% of our prototrophy predictions are in line with experimental data, suggesting that the high prevalence of predicted auxotrophies among the human gut bacterial genomes is indeed realistic and not due to a potential technical bias in the *in-silico* approach.

Conclusion

In conclusion, amino acid auxotrophies are common in the human gut microbiome. An increasing abundance of auxotrophic members is positively correlated with a higher diversity and a more stable gut community. In sum, the results indicate a potential impact of auxotrophic bacteria on the microbial ecology in the gut and may guide the development of novel intervention strategies to restore microbiome diversity.

Material and Methods

Cohorts

Data from two human population cohorts were analyzed for the present study. The first cohort, here named 'discovery cohort', comprised paired stool metagenomes and serum metabolomes from 185 participants. This cohort was recruited at the University Hospital Schleswig Holstein, Campus Kiel 2016 and comprised detailed phenotypic, disease related and dietary information. The study was approved by the local ethic committee in Kiel (D441). None of the participants had received any antibiotics or other medication 2 months prior to inclusion.

The second cohort comprised longitudinal stool metagenomes from 72 study participants. Data from this cohort were already part of a previous study (30), which were reanalyzed in the present study. For each participant from the longitudinal cohort, two metagenomes were sequenced from stool samples that were 3 years apart. In addition, for each sampling time point data from food frequency questionnaires were available. Detailed information about the sampling method, study design and sequencing method can be found in the publication of Troci *et al.*, 2022 (30).

Metagenome sequencing

DNA of stool samples was extracted using the QIAamp DNA fast stool mini kit automated on the QIAcube (Qiagen, Hilden, Germany) with a prior bead-beating step as has been described in detail earlier (52). DNA extracts were used for metagenomic library preparation as described previously (30) using Illumina Nextera DNA Library Preparation Kit (Illumina, San Diego, CA) and sequenced with 2x150 bp paired-end reads on a NovaSeq platform (Illumina).

Metagenome data processing

Metagenomic reads were processed using the metagenome-atlas workflow v2.9.0(53) with default parametrization if not stated otherwise in the following description. In brief, the applied workflow consisted of four main steps: (1) Quality control and filtering, (2) read assembly, (3) binning of contigs, and (4) bin clustering to sub-species level metagenome-

assembled genomes and quantification. Further details for each step are provided in the Supplementary Information.

Reconstruction of genome-scale metabolic models

Genome-scale metabolic models were reconstructed for bacterial genomes from the Human Reference Gut Microbiome (HRGM) genome collection(26,54). The HRGM collection combines isolate and metagenome-assembled genomes (MAGs) from several data sources to summarize genome sequences obtained from human fecal samples. Metabolic models were reconstructed using gapseq version 1.2(55). The model files as well as the documentation of the reconstruction workflow are publicly available at *Zenodo*(56). A detailed description for the genome-scale metabolic model reconstruction workflow can be found in the Supplementary Information.

Prediction of amino acid auxotrophies

Amino acid auxotrophies were predicted with flux balance analysis(57), where the objective function was set to the flux through the biomass formation reaction. In detail, each model was tested for its ability to form biomass under two different environmental conditions: First with the growth medium that was predicted with *gapseq* (see above) and, second, with the same medium but without the focal amino acid. An organism was defined as auxotrophic for a specific amino acid, if the predicted growth rate without the focal amino acid was less than 5% of the growth rate with the original medium. Flux balance analysis was performed in R (version 4.1.2) and R package *sybil* version 2.2.0 (58). We validated our auxotrophy predictions for 130 organisms, for which experimental data for amino acid auxotrophies and prototrophies were available in scientific literature (see Supplementary Information for details).

Prediction of metabolic by-products

For comparison of auxotrophic to prototrophic bacteria the production rates of fermentation by-product formation were predicted. Metabolic by-products were predicted with flux-balance-analysis(57) using the flux through the biomass reaction as objective function (i.e. maximization). Rates of metabolite production ($mmol*gDW^{-1}*hr^{-1}$) were

normalized by growth rates (hr^{-1}), which results in the unit mmol/gDW. Production rates > 1 mmol/gDW were considered as microbial production. The production of the two enantiomers D- and L-lactate were combined since their production rates were interchangeable in the FBA solution. Flux balance analysis was performed in R (version 4.1.2) and R package sybil (58) (2.2.0).

Targeted metabolomics of blood samples

Metabolite quantification for serum was performed by liquid chromatography tandem mass spectrometry (LC-MS-MS) using the MxP Quant 500 kit (Biocrates Life Sciences AG, Innsbruck, Austria) according to the manufacturer's instructions. Please refer to the Supplementary Information document for details on blood sample preparation and metabolite quantification.

Statistical data analysis

All data analysis steps and statistical tests were performed using R (version 4.1.2). Flow charts (Fig 1 and 3A) were created and rendered using Flowchart Designer 3.

The Spearman correlation analysis was used for the completeness of the genomes and number of auxotrophies. p-values were corrected for multiple testing using the Benjamini and Hochberg method (59). In all statistical tests, a p-value of <0.05 was considered as significance threshold. Bray-Curtis distances were calculated using relative abundances of MAGs using the R-package vegan, version 2.6-2 (60).

In order to address if specific auxotrophies occur more often or less often together, than expected by chance we employed a re-sampling approach using the Rasch Sampler version 0.8-8 (61) (see Supplementary Information for details).

Alpha diversity was calculate using the Shannon index as implemented in the R-package 'vegan' (60). To study the effect of the metabolic dissimilarity on diversity, the average pairwise Hamming distance between auxotrophic profiles of co-occurring MAGs was calculated per sample. In other words, the Hamming distance is the number of amino acids for which the two genotypes had different auxotrophy predictions. In addition to the Hamming distance, we also calculated the abundance-weighted average of auxotrophies per genome y_j for each sample j using the equation:

$$y_j = \sum_{i \in M} a_i p_{ij}$$

Where M is the set of all MAGs, a_i the number of auxotrophies in MAG i, and p_{ij} the relative abundance of MAG i in sample j.

For the longitudinal cohort, the Bray-Curtis distance was correlated with the abundance-weighted average of auxotrophies per genome at the first time point using the spearman correlation. Further, the Spearman correlation was used for determining the association between the Bray-Curtis distance and the hamming distance. With the food frequency questionnaires, the total dietary intake of amino acids per day was summed up for every individual and the energy percentage was then calculated based on the total energy intake per day. For studying an association between the total dietary intake of amino acids relative the total consumed energy (E%) and the frequency of amino acid auxotrophic bacteria, the Spearman correlation was used. The correlation between the intake of amino acids and frequencies of amino acid auxotrophic bacteria was studied separately for both time points (Supplementary Figure S6).

Data availability

The reconstructed genome-scale metabolic models from the HRGM catalogue are available via *Zenodo* (56). MAGs and their metabolic model reconstructions for the discovery- and the longitudinal cohort are also available via *Zenodo* (62). Metabolic model reconstructions for 124 prototrophic genotypes are available via *Zenodo* (63). Metagenome sequencing data are provided with open access via the European Nucleotide Archive 'ENA' for both cohorts (discovery cohort accession: PRJEB60573, longitudinal cohort accession: PRJEB48605).

Code availability

The code for analysis of the data can be found in the github repository: https://github.com/SvBusche/Auxo_manuscript_2023

Competing interests

The authors declare that there are no conflicts of interest related to this work.

Funding

This work was supported by the DFG Cluster of excellence "Precision medicine in chronic inflammation (PMI)" RTF V and TI-1 (K.A, S.W.), the PMI Miniproposal funding (S.W.), the EKFS (Clinician Scientist Professorship, K.A.), the BMBF (eMED Juniorverbund "Try-IBD" 01ZX1915A, K.A.), the DFG RU5042 (A.F., K.A.), DZHK (N.F.), and DFG FR 1289/17-1 (N.F.). This research was further supported through high-performance computing resources available at the Kiel University Computing Centre, which received funding from the German Research Foundation (DFG project number 440395346).

References

1. Parada Venegas D, De la Fuente MK, Landskron G, González MJ, Quera R, Dijkstra G, et al. Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. Front Immunol. 2019 Mar 11;10:277.

2. Neis E, Dejong C, Rensen S. The Role of Microbial Amino Acid Metabolism in Host Metabolism. Nutrients. 2015 Apr 16;7(4):2930–46.

3. Roager HM, Licht TR. Microbial tryptophan catabolites in health and disease. Nat Commun. 2018 Dec;9(1):3294.

4. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. Nat Rev Microbiol. 2014 Oct;12(10):661–72.

5. Hinrichsen F, Hamm J, Westermann M, Schröder L, Shima K, Mishra N, et al. Microbial regulation of hexokinase 2 links mitochondrial metabolism and cell death in colitis. Cell Metabolism. 2021 Dec;33(12):2355-2366.e8.

6. Dodd D, Spitzer MH, Van Treuren W, Merrill BD, Hryckowian AJ, Higginbottom SK, et al. A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. Nature. 2017 Nov;551(7682):648–52.

7. Tintelnot J, Xu Y, Lesker TR, Schönlein M, Konczalla L, Giannou AD, et al. Microbiota-derived 3-IAA influences chemotherapy efficacy in pancreatic cancer. Nature. 2023 Mar 2;615(7950):168–74.

8. Nikolaus S, Schulte B, Al-Massad N, Thieme F, Schulte DM, Bethge J, et al. Increased Tryptophan Metabolism Is Associated With Activity of Inflammatory Bowel Diseases. Gastroenterology. 2017 Dec;153(6):1504-1516.e2.

9. Effenberger M, Reider S, Waschina S, Bronowski C, Enrich B, Adolph TE, et al. Microbial Butyrate Synthesis Indicates Therapeutic Efficacy of Azathioprine in IBD Patients. Journal of Crohn's and Colitis. 2021 Jan 13;15(1):88–98.

10. Pryor R, Norvaisas P, Marinos G, Best L, Thingholm LB, Quintaneiro LM, et al. Host-Microbe-Drug-Nutrient Screen Identifies Bacterial Effectors of Metformin Therapy. Cell. 2019 Sep;178(6):1299-1312.e29.

11. D'Souza G, Waschina S, Pande S, Bohl K, Kaleta C, Kost C. LESS IS MORE: SELECTIVE ADVANTAGES CAN EXPLAIN THE PREVALENT LOSS OF BIOSYNTHETIC GENES IN BACTERIA: ADAPTIVE LOSS OF BIOSYNTHETIC GENES IN BACTERIA. Evolution. 2014 Sep;68(9):2559–70.

12. Heinken A, Khan MT, Paglia G, Rodionov DA, Harmsen HJM, Thiele I. Functional Metabolic Map of Faecalibacterium prausnitzii, a Beneficial Human Gut Microbe. J Bacteriol. 2014 Sep 15;196(18):3289–302.

13. Soto-Martin EC, Warnke I, Farquharson FM, Christodoulou M, Horgan G, Derrien M, et al. Vitamin Biosynthesis by Human Gut Butyrate-Producing Bacteria and Cross-Feeding in Synthetic Microbial Communities. Relman DA, editor. mBio. 2020 Aug 25;11(4):e00886-20.

14. Xavier JC, Patil KR, Rocha I. Integration of Biomass Formulations of Genome-Scale Metabolic Models with Experimental Data Reveals Universally Essential Cofactors in Prokaryotes. Metabolic Engineering. 2017 Jan;39:200–8.

15. van der Ark KCH, Aalvink S, Suarez-Diez M, Schaap PJ, de Vos WM, Belzer C. Model-driven design of a minimal medium for *Akkermansia muciniphila* confirms mucus adaptation. Microb Biotechnol. 2018 May;11(3):476–85.

16. Tramontano M, Andrejev S, Pruteanu M, Klünemann M, Kuhn M, Galardini M, et al. Nutritional preferences of human gut bacteria reveal their metabolic idiosyncrasies. Nat Microbiol. 2018 Apr;3(4):514–22.

17. Kiela PR, Ghishan FK. Physiology of Intestinal Absorption and Secretion. Best Practice & Research Clinical Gastroenterology. 2016 Apr;30(2):145–59.

18. Dethlefsen L, McFall-Ngai M, Relman DA. An ecological and evolutionary perspective on human–microbe mutualism and disease. Nature. 2007 Oct;449(7164):811–8.

19. Weiss AS, Burrichter AG, Durai Raj AC, von Strempel A, Meng C, Kleigrewe K, et al. In vitro interaction network of a synthetic gut bacterial community. ISME J. 2021 Dec 2;1–15.

20. Morris BEL, Henneberger R, Huber H, Moissl-Eichinger C. Microbial syntrophy: interaction for the common good. FEMS Microbiol Rev. 2013 May;37(3):384–406.

21. Rivière A, Gagnon M, Weckx S, Roy D, De Vuyst L. Mutual Cross-Feeding Interactions between Bifidobacterium longum subsp. longum NCC2705 and Eubacterium rectale ATCC 33656 Explain the Bifidogenic and Butyrogenic Effects of Arabinoxylan Oligosaccharides. Schloss PD, editor. Appl Environ Microbiol. 2015 Nov 15;81(22):7767–81.

22. Shetty SA, Kostopoulos I, Geerlings SY, Smidt H, de Vos WM, Belzer C. Dynamic metabolic interactions and trophic roles of human gut microbes identified using a minimal microbiome exhibiting ecological properties. ISME J. 2022 Sep;16(9):2144–59.

23. Giri S, Oña L, Waschina S, Shitut S, Yousif G, Kaleta C, et al. Metabolic dissimilarity determines the establishment of cross-feeding interactions in bacteria. Current Biology. 2021 Dec;31(24):5547-5557.e6.

24. Oña L, Giri S, Avermann N, Kreienbaum M, Thormann KM, Kost C. Obligate cross-feeding expands the metabolic niche of bacteria. Nat Ecol Evol. 2021 Sep;5(9):1224–32.

25. Oña L, Kost C. Cooperation increases robustness to ecological disturbance in microbial cross-feeding networks. Ecology Letters. 2022 Jun;25(6):1410–20.

26. Kim CY, Lee M, Yang S, Kim K, Yong D, Kim HR, et al. Human reference gut microbiome catalog including newly assembled genomes from under-represented Asian metagenomes. Genome Med. 2021 Dec;13(1):134.

27. Price M. Erroneous predictions of auxotrophies by CarveMe. Nat Ecol Evol [Internet]. 2022 Dec 5 [cited 2023 Mar 7]; Available from: https://www.nature.com/articles/s41559-022-01936-3

28. Liu R, Hong J, Xu X, Feng Q, Zhang D, Gu Y, et al. Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. Nat Med. 2017 Jul;23(7):859–68.

29. Davenport M, Poles J, Leung JM, Wolff MJ, Abidi WM, Ullman T, et al. Metabolic Alterations to the Mucosal Microbiota in Inflammatory Bowel Disease: Inflammatory Bowel Diseases. 2014 Apr;20(4):723–31.

30. Troci A, Rausch P, Waschina S, Lieb W, Franke A, Bang C. Long-Term Dietary Effects on Human Gut Microbiota Composition Employing Shotgun Metagenomics Data Analysis. Molecular Nutrition Food Res. 2022 Jul 5;2101098.

31. Henriques SF, Dhakan DB, Serra L, Francisco AP, Carvalho-Santos Z, Baltazar C, et al. Metabolic cross-feeding in imbalanced diets allows gut microbes to improve reproduction and alter host behaviour. Nat Commun. 2020 Dec;11(1):4236.

32. D'Souza G, Shitut S, Preussger D, Yousif G, Waschina S, Kost C. Ecology and evolution of metabolic cross-feeding interactions in bacteria. Nat Prod Rep. 2018;35(5):455–88.

33. Waschina S, D'Souza G, Kost C, Kaleta C. Metabolic network architecture and carbon source determine metabolite production costs. FEBS J. 2016 Jun;283(11):2149–63.

34. Morris JJ, Lenski RE, Zinser ER. The Black Queen Hypothesis: Evolution of Dependencies through Adaptive Gene Loss. mBio. 2012 May 2;3(2):e00036-12.

35. Garcia SL, Buck M, McMahon KD, Grossart HP, Eiler A, Warnecke F. Auxotrophy and intrapopulation complementary in the 'interactome' of a cultivated freshwater model community. Mol Ecol. 2015 Sep;24(17):4449–59.

36. Pande S, Merker H, Bohl K, Reichelt M, Schuster S, de Figueiredo LF, et al. Fitness and stability of obligate cross-feeding interactions that emerge upon gene loss in bacteria. ISME J. 2014 May;8(5):953–62.

37. Pande S, Shitut S, Freund L, Westermann M, Bertels F, Colesie C, et al. Metabolic cross-feeding via intercellular nanotubes among bacteria. Nat Commun. 2015 Feb 23;6(1):6238.

38. van der Wielen N, Moughan PJ, Mensink M. Amino Acid Absorption in the Large Intestine of Humans and Porcine Models. J Nutr. 2017 Aug;147(8):1493–8.

39. Macfarlane GT, Cummings JH, Macfarlane S, Gibson GR. Influence of retention time on degradation of pancreatic enzymes by human colonic bacteria grown in a 3-stage continuous culture system. Journal of Applied Bacteriology. 1989 Nov;67(5):521–7.

40. Zeng X, Xing X, Gupta M, Keber FC, Lopez JG, Lee YCJ, et al. Gut bacterial nutrient preferences quantified in vivo. Cell. 2022 Sep;185(18):3441-3456.e19.

41. Bansil R, Turner BS. The biology of mucus: Composition, synthesis and organization. Advanced Drug Delivery Reviews. 2018 Jan;124:3–15.

42. Fassarella M, Blaak EE, Penders J, Nauta A, Smidt H, Zoetendal EG. Gut microbiome stability and resilience: elucidating the response to perturbations in order to modulate gut health. Gut. 2021 Mar;70(3):595–605.

43. Scher JU, Ubeda C, Artacho A, Attur M, Isaac S, Reddy SM, et al. Decreased Bacterial Diversity Characterizes the Altered Gut Microbiota in Patients With Psoriatic Arthritis, Resembling Dysbiosis in Inflammatory Bowel Disease: Gut Microbiota in PsA. Arthritis & Rheumatology. 2015 Jan;67(1):128–39.

44. Aden K, Rehman A, Waschina S, Pan WH, Walker A, Lucio M, et al. Metabolic Functions of Gut Microbes Associate With Efficacy of Tumor Necrosis Factor Antagonists in Patients With Inflammatory Bowel Diseases. Gastroenterology. 2019 Nov;157(5):1279-1292.e11.

45. Lepage P, Häsler R, Spehlmann ME, Rehman A, Zvirbliene A, Begun A, et al. Twin Study Indicates Loss of Interaction Between Microbiota and Mucosa of Patients With Ulcerative Colitis. Gastroenterology. 2011 Jul;141(1):227–36.

46. Fangmann D, Theismann EM, Türk K, Schulte DM, Relling I, Hartmann K, et al. Targeted Microbiome Intervention by Microencapsulated Delayed-Release Niacin Beneficially Affects Insulin Sensitivity in Humans. Diabetes Care. 2018 Mar 1;41(3):398–405.

47. Russell WR, Duncan SH, Scobbie L, Duncan G, Cantlay L, Calder AG, et al. Major phenylpropanoid-derived metabolites in the human gut can arise from microbial fermentation of protein. Mol Nutr Food Res. 2013 Mar;57(3):523–35.

48. Brial F, Chilloux J, Nielsen T, Vieira-Silva S, Falony G, Andrikopoulos P, et al. Human and preclinical studies of the host–gut microbiome co-metabolite hippurate as a marker and mediator of metabolic health. Gut. 2021 Nov;70(11):2105–14.

49. Lamas B, Natividad JM, Sokol H. Aryl hydrocarbon receptor and intestinal immunity. Mucosal Immunology. 2018 Jul;11(4):1024–38.

50. Saito Y, Sato T, Nomoto K, Tsuji H. Identification of phenol- and p-cresol-producing intestinal bacteria by using media supplemented with tyrosine and its metabolites. FEMS Microbiology Ecology. 2018 Sep 1;94(9):fiy125.

51. Price MN, Deutschbauer AM, Arkin AP. GapMind: Automated Annotation of Amino Acid Biosynthesis. Hallam SJ, editor. mSystems. 2020 Jun 30;5(3):e00291-20.

52. Thingholm LB, Bang C, Rühlemann MC, Starke A, Sicks F, Kaspari V, et al. Ecology impacts the decrease of Spirochaetes and Prevotella in the fecal gut microbiota of urban humans. BMC Microbiol. 2021 Oct 11;21(1):276.

53. Kieser S, Brown J, Zdobnov EM, Trajkovski M, McCue LA. ATLAS: a Snakemake workflow for assembly, annotation, and genomic binning of metagenome sequence data. BMC Bioinformatics. 2020 Dec;21(1):257.

54. Almeida A, Nayfach S, Boland M, Strozzi F, Beracochea M, Shi ZJ, et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. Nat Biotechnol. 2021 Jan;39(1):105–14.

55. Zimmermann J, Kaleta C, Waschina S. gapseq: informed prediction of bacterial metabolic pathways and reconstruction of accurate metabolic models. Genome Biol. 2021 Dec;22(1):81.

56. Waschina, Silvio, Zimmermann, Johannes, Kaleta, Christoph. gapseq reconstructions for 5414 genomes from the HRGM collection [Internet]. Zenodo; 2023 [cited 2023 Feb 16]. Available from: https://zenodo.org/record/7645465

57. Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? Nat Biotechnol. 2010 Mar;28(3):245–8.

58. Gelius-Dietrich G, Desouki AA, Fritzemeier CJ, Lercher MJ. sybil – Efficient constraint-based modelling in R. BMC Syst Biol. 2013 Dec;7(1):125.

59. Benjamini Y, Krieger AM, Yekutieli D. Adaptive linear step-up procedures that control the false discovery rate. Biometrika. 2006 Sep 1;93(3):491–507.

60. Jari Oksanen, Gavin L. Simpson, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, et al. vegan: Community Ecology Package. R package version 26-2 [Internet]. 2022; Available from: https://CRAN.R-project.org/package=vegan

61. Verhelst ND, Hatzinger R, Mair P. The Rasch Sampler. J Stat Soft [Internet]. 2007 [cited 2022 Apr 21];20(4). Available from: http://www.jstatsoft.org/v20/i04/

62. Waschina, Silvio. MAGs and gapseq models for auxotrophy predictions in the human gut microbiome [Internet]. Zenodo; 2023 [cited 2023 Feb 17]. Available from: https://zenodo.org/record/7646601

63. Busche, Svenja, Waschina, Silvio. gapseq reconstructions for 124 prototrophic genotypes [Internet]. Zenodo; 2023 [cited 2023 Mar 13]. Available from: https://zenodo.org/record/7708385

64. Caspi R, Billington R, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, et al. The MetaCyc database of metabolic pathways and enzymes. Nucleic Acids Research. 2018 Jan 4;46(D1):D633–9.

Acknowledgements

We would like to thank the staff of the IKMB (*Institute für klinische Molekularbiologie*) microbiome laboratory and the IKMB sequencing facilities for their excellent technical support.

Author information

Authors and Affiliations

Institute of Human Nutrition and Food Science, Kiel University, Department of Nutriinformatics, Kiel, Germany

Svenja Busche, Danielle Harris, Silvio Waschina

Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany

Danielle Harris, Mhmd Oumari, Corinna Bang, Philip Rosenstiel, Stefan Schreiber, Andre Franke, Konrad Aden

Research Group Evolutionary Ecology and Genetics, Zoological Institute, University of Kiel, Kiel, Germany

Johannes Zimmermann

Max Planck Institute for Evolutionary Biology, Plön, Germany

Johannes Zimmermann

Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM), Hanover, Germany

Sven Schuchardt

Department of Internal Medicine III, University Medical Center Schleswig-Holstein, Kiel, Germany

Derk Frank , Norbert Frey

German Centre for Cardiovascular Diseases (DZHK), Partner site Hamburg, Kiel, Lübeck, Germany

Derk Frank

Department of Internal Medicine I, University Medical Center Schleswig-Holstein, Kiel, Germany

Stefan Schreiber, Konrad Aden

Author contribution

Conceptualization: S.W.

Methodology: S.B., Sv.S., S.W.

Software: S.B., D.H., S.W., J.Z.

Validation: D.H., S.W., J.Z.

Formal analysis: S.B., D.H, M.O.

Investigation: S.B.

Resources: Sv.S., D.F., N.F, A.F., K.A., S.W.

Data curation: C.B., M.O., D.F.

Writing - Original Draft: S.B., S.W.

Writing - Review & Editing: S.B., D.H., S.W.

Visualization: S.B.

Project administration: K.A., S.W.

Funding acquisition: A.F., N.F., P.R., St.S., K.A., S.W.

Figures

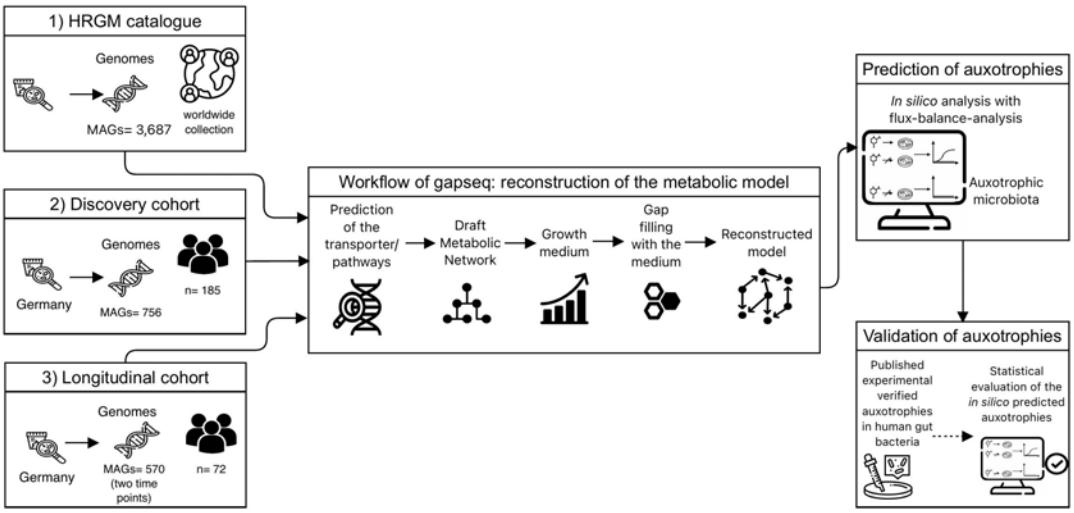

Figure 1: Workflow for the prediction of auxotrophies with genome-scale metabolic modeling

Gapseq was used to reconstruct genome-scale metabolic models from MAGs for genomes from the Human Reference Gut Microbiome (HRGM) catalogue(26) and metagenome-assembly genomes (MAGs) from the discovery cohort and longitudinal cohort. The workflow of gapseq to reconstruct metabolic models consists of five steps: transporter/metabolic pathway prediction, draft metabolic network construction, growth medium prediction, gap filling, final model reconstruction(55). Auxotrophy prediction was performed using flux-balance analysis(57). Free available icons were taken from www.flaticon.com (creators: photo3idea_studio, Freepik, surang, Eucalyp, Voysla, juicy_fish, smashingstocks, SBTS2018, creative_designer).

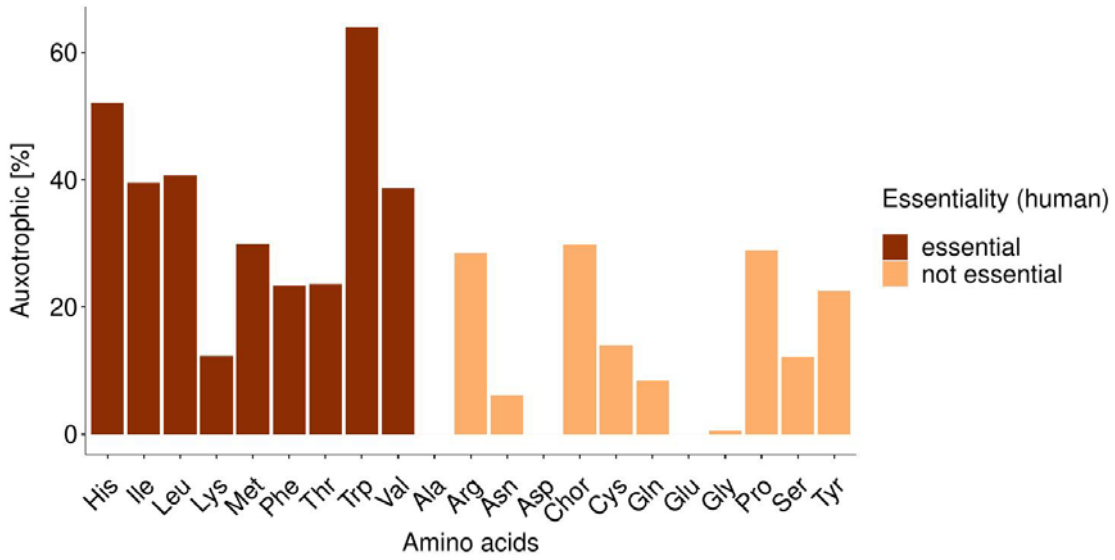

Figure 2: Abundances of auxotrophies in 3,687 genomes

The predicted amino acid auxotrophies in HRGM genomes were categorized into human essential and nonessential amino acids.

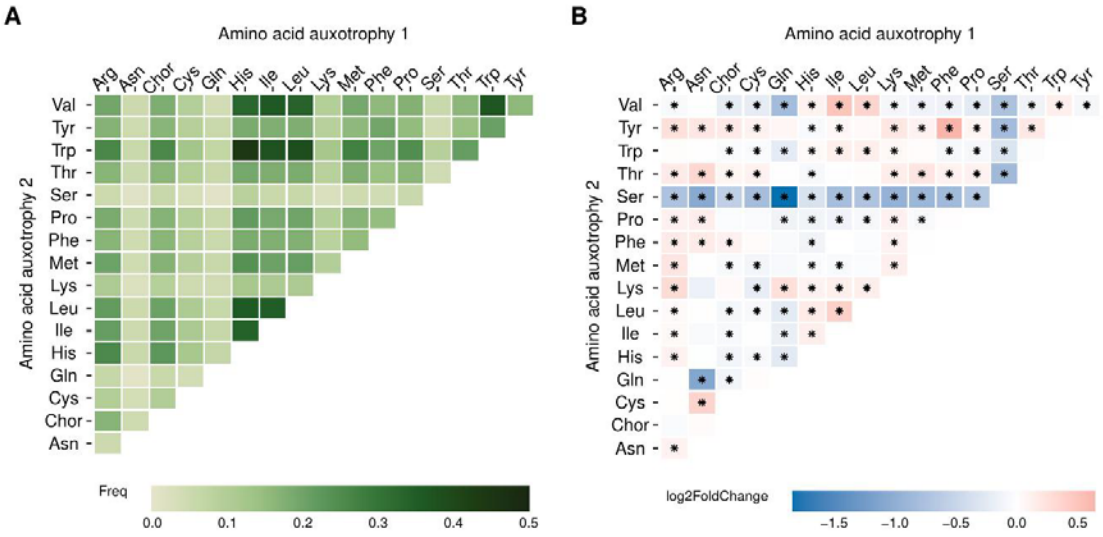

Figure 3: Co-occurrence of amino acid auxotrophies in bacterial genomes

(A) For 3,687 HRGM genomes, the frequencies of pairwise auxotrophy co-occurrence were calculated. (B) The figure represents the statistical assessment of whether the observed co-occurrence of auxotrophies is more frequent than expected by chance. For this purpose, the fold change for each co-auxotrophy was calculated by forming the quotient of the observed frequency and the median of the frequency determined by chance (resampling test (Rasch sampler) with 1,000 iterations). Asterisks indicate statistical significance (FDR-corrected $p<0.05$).

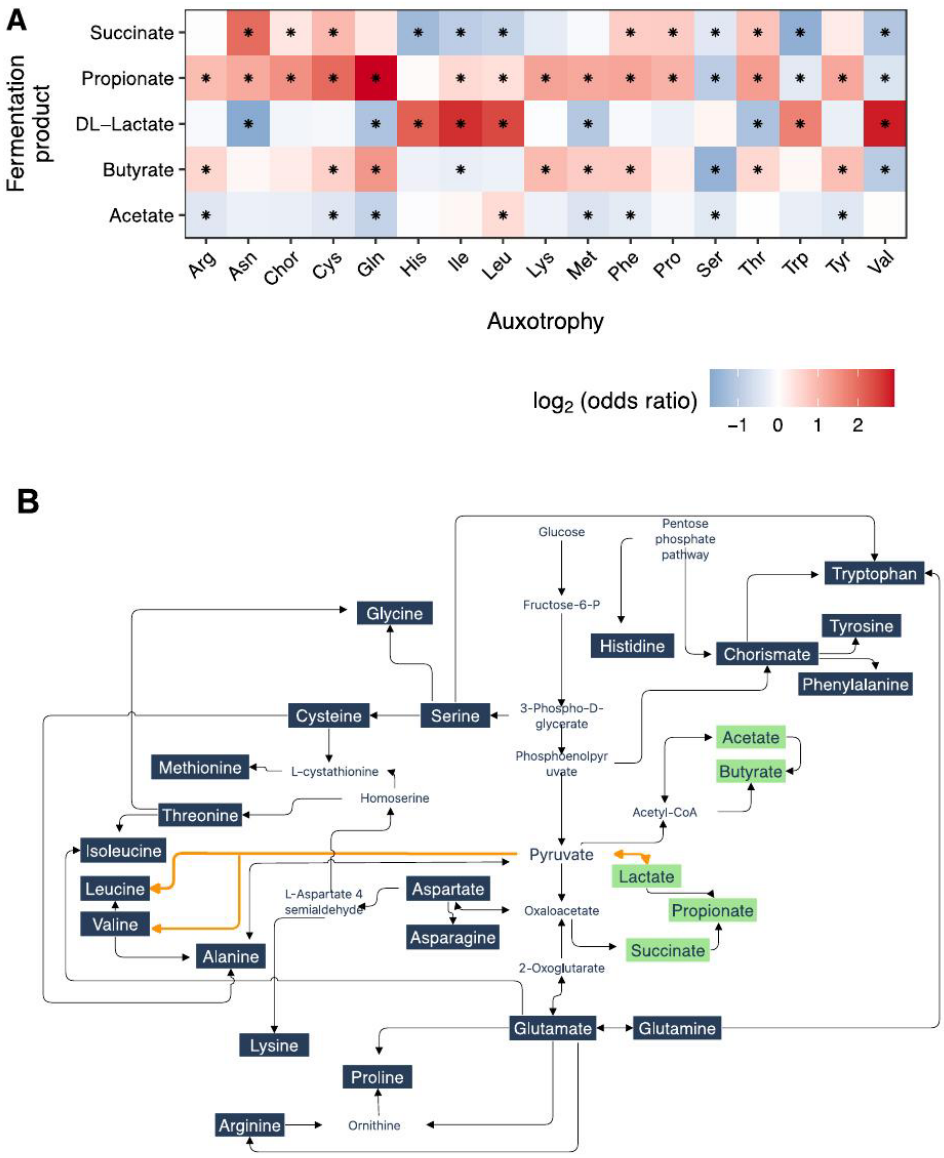

Figure 4: Associations of auxotrophies and fermentation products

(A) Comparison of fermentation product production rates in auxotrophic and prototrophic bacteria. Production rates of fermentation by-products were predicted with flux-balance analysis (cutoff-value > 1 mmol/gDW) in 3,687 HRGM genomes. The association with the auxotrophic or prototrophic phenotype was statistically evaluated with the Fisher test for exact count data by calculating odds ratios. Asterisk denote FDR-corrected p-values <0.05. (B) Interconnection between the pathways of formation of fermentation products and amino acids, based on Metacyc pathways (64).

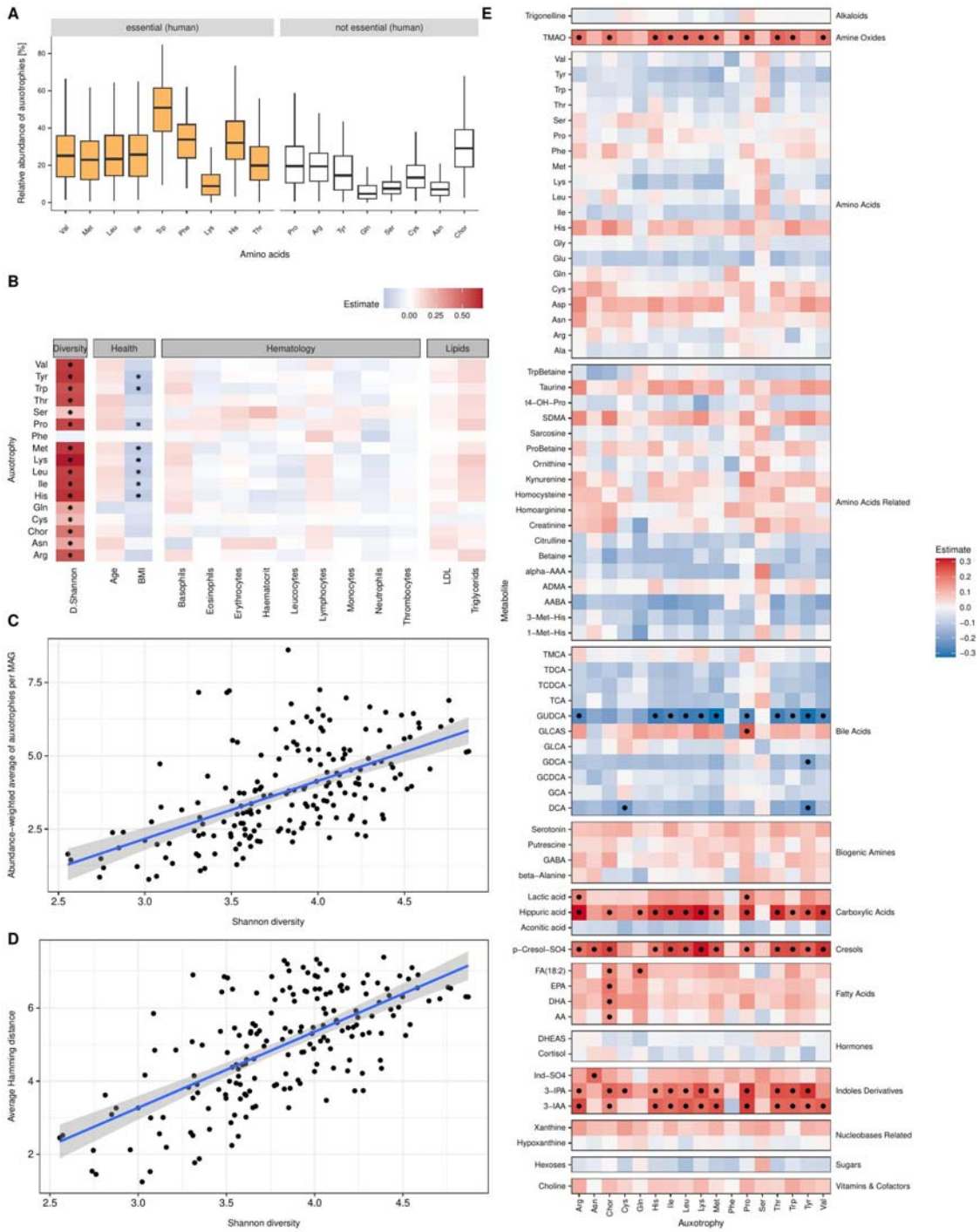

Figure 5: Distribution of auxotrophies in human gut microbiome from the discovery cohort, their association with diversity and serum metabolite levels.

(A) Boxplots displaying the abundance of amino acid auxotrophies in the human gut microbiome (n=185 samples). (B) Partial Spearman correlation between the frequency of auxotrophic gut bacteria and serum levels of health markers and microbiome Shannon diversity. Dots indicate significant associations (FDR-corrected p-values < 0.05, adjusted for the potential confounders age, sex, and BMI). (C) Abundance-

weighted average of auxotrophies per MAG was calculated and correlated with the Shannon microbiome diversity (Spearman correlation, ρ =0.60, p<2.2e-16). (D) Average hamming distance was calculated to study the metabolic dissimilarity of auxotrophy profiles of coexisting genotypes and therefore potential cross-feeding interactions within the microbial communities. With the Spearman correlation, the association between the calculated average hamming distance and the Shannon diversity in the gut was estimated (ρ =0.62, p<2.2e-16). (E) Partial Spearman correlations between the serum levels of metabolites and the frequency of auxotrophic bacteria in the gut microbiome. Abbreviations for the serum metabolite levels can be found in Supplementary Table S1. Dots indicate significant associations (FDR-corrected p-values < 0.05, adjusted for confounders age, sex, and BMI).

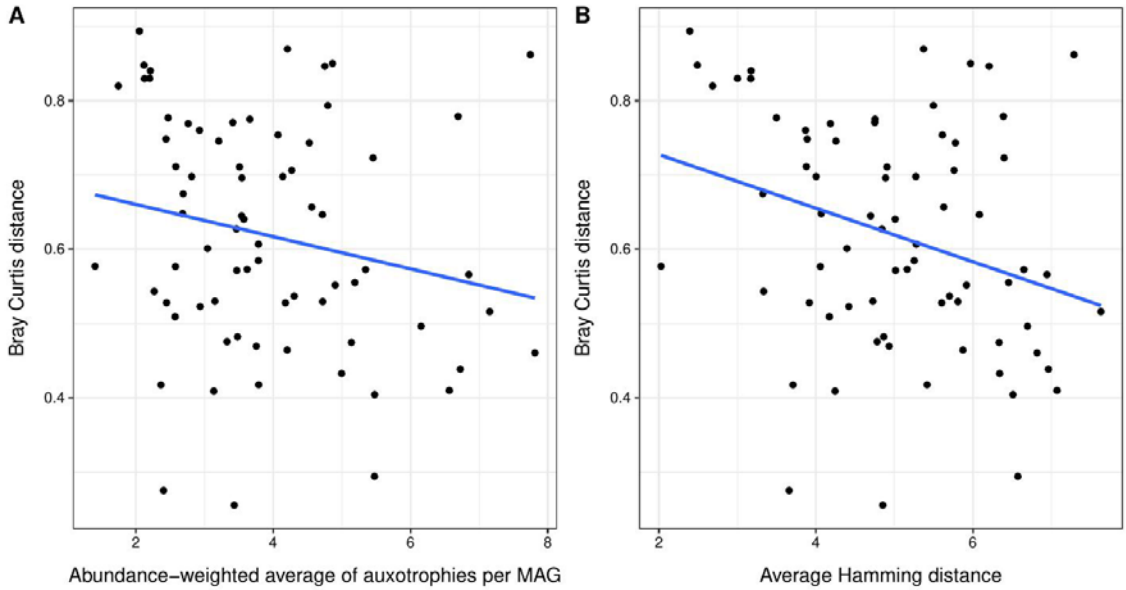

Figure 6: Influence of auxotrophies on long-term (3 years) stability of the human gut microbiome

(A) The stability of the human gut microbiome was calculated with the Bray-Curtis distance between the two time-points in the longitudinal study and correlated with the abundance-weighted average of auxotrophies per MAG at the first time-point to study a potential influence of auxotrophies on the long-term stability of the human gut microbiome (ρ =-0.24, p=0.0403, n =72), (B) The average Hamming distance was calculated for the first time-point and then correlated with the Bray Curtis distance to investigate the influence of potential cross-feeding on the long-term-stability (Spearman correlation, ρ =-0.30, p=0.01072, n =72), the blue lines display linear regression results.

Table 1: Cohort characteristics with median ± standard deviation

	Discovery cohort	Longitudinal cohort
Age	47 ± 8.3	49.9 ± 7.3
BMI	24.5 ± 3.9	25.6 ± 6
Female (%)	44.2	40
Stool samples with metagenomes	185	144
Participants with two metagenomes	–	72