

From the Institute of Human Nutrition and Food Science  
at Kiel University

# **Ecology of amino acid auxotrophic bacteria in the human gastrointestinal tract**

Dissertation  
in fulfillment of the requirements for the doctoral degree  
of the Faculty of Agricultural and Nutritional Sciences  
at Kiel University

submitted by

**M.Sc. Svenja Karen Starke**

from Berlin

Kiel, 2025

---

Dean: Prof. Dr. Tim Diekötter

1<sup>st</sup> examiner: Jun.-Prof. Dr. Silvio Waschina

2<sup>nd</sup> examiner: Prof. Dr. Manuel Liebeke

Date of the oral examination: 21<sup>st</sup> May 2025

Veröffentlicht mit Genehmigung der Agrar- und Ernährungswissenschaftlichen  
Fakultät

## Abstract

The pathophysiology and pathogenesis of several metabolic and inflammatory diseases is associated with specific taxa within the human gut microbiome and the microbial amino acid metabolism. However, these ecological changes remain still poorly understood. Studying amino acid auxotrophic bacteria, which are incapable of the *de novo* synthesis of vital nutrients like amino acids and known to influence the ecology of microbial systems, could improve the understanding of ecological imbalances of the human gut microbiome in diseases. So far, the distribution and impact of amino acid auxotrophies on the human gut microbiome is not well studied. In this thesis, the distribution of amino acid auxotrophies is characterized in small and large intestinal microbiomes in non-dysbiotic and dysbiotic states. In addition, the influence of auxotrophies on the ecology of the human gut microbiome, particularly the diversity and long-term stability, is analyzed.

This thesis revealed that auxotrophies in the human gut microbiome are ubiquitous. Auxotrophies for human-essential amino acids, especially tryptophan, are most common. In non-dysbiotic states, auxotrophies are more abundant in the small intestine than the large intestine but decreased in small intestinal bacterial overgrowth.

Analysing the ecological impact of auxotrophies on the human gut microbiome showed that auxotrophies positively influence the diversity and long-term stability of the human gut microbiome. Cross-feeding of amino acids was further positively associated with diversity and long-term stability.

Auxotrophies were predicted with genome-scale metabolic modeling in this thesis. Discrepancies are known for *in silico* predictions of auxotrophies when compared to experimentally verified auxotrophies results. Therefore, genome-scale metabolic modeling and Gapmind, which uses genome annotation for auxotrophies prediction, are validated with genomes from experimentally verified auxotrophic bacteria. Genome-scale metabolic modeling and GapMind showed an accuracy of over 90%.

Conclusively, this thesis shows that amino acid auxotrophies are common in the human gut microbiome. Significantly, auxotrophies contribute to both the diversity and long-term stability of the gut microbiome. Further, the accuracy of genome-scale metabolic modeling for predicting auxotrophies was confirmed underlining

its potential as a valuable tool in auxotrophic predictions. These findings have implications for understanding microbial ecology in gut health and may inform future strategies to modulate the gut microbiome in diseases.

## Zusammenfassung

In vielen entzündlichen und metabolischen Erkrankungen sind die Zusammensetzung und Funktion des Darmmikrobioms verändert. Dabei spielt der mikrobielle Aminosäure-Stoffwechsel eine Rolle bei der Pathophysiologie dieser Erkrankungen. Aminosäure-auxotrophe Bakterien beeinflussen die Ökologie in mikrobiellen Systemen. Die Untersuchung ihrer Verteilung im Darmmikrobiom könnte helfen, Dysbiosen in Krankheitszuständen besser zu verstehen. Auxotrophe Bakterien sind nicht in der Lage, bestimmte für ihr Wachstum benötigte Nährstoffe, wie zum Beispiel Aminosäuren, selber herzustellen, wohingegen prototrophe Bakterien selber Aminosäuren herstellen können. Auxotrophe Bakterien sind daher auf die Aufnahme von Aminosäuren aus ihrer Nährstoffumgebung angewiesen. Obwohl bekannt ist, dass Auxotrophien die Ökologie mikrobieller Systeme beeinflussen, sind ihre Verteilung und ihr Einfluss auf das Darmmikrobiom im Menschen bisher wenig untersucht. In dieser Dissertation wird die Verteilung von Aminosäure-Auxotrophien im Mikrobiom des Dünndarms und Dickdarms im dysbiotischen und nicht dysbiotischen Zustand analysiert. Weiter wird der Einfluss von auxotrophen Bakterien auf die Diversität und Langzeitstabilität des Darmmikrobioms untersucht.

Die Analyse der Verteilung von Auxotrophien im Darmmikrobiom zeigte, dass Auxotrophien im Darmmikrobiom besonders für die vom Menschen essentiellen Aminosäuren verbreitet sind. Im Mikrobiom des Dünndarms sind mehr Auxotrophien zu finden als im Dickdarm im nicht dysbiotischen Zustand. Jedoch sind sie im dysbiotischen Zustand der Dünndarmfehlbesiedelung reduziert.

Es ist bereits bekannt, dass auxotrophe Bakterien einen Einfluss auf die Ökologie von mikrobiellen Systemen haben, daher wurde ein Zusammenhang zwischen auxotrophen Bakterien im Dickdarm und der Diversität und Langzeit-Stabilität des Darmmikrobioms untersucht. Auxotrophe Bakterien haben einen positiven Einfluss auf die Diversität und Langzeitstabilität des Darmmikrobioms. Ebenso wurde ein positiver Zusammenhang zwischen dem potentiellen Austausch von Aminosäuren zwischen auxotrophen und prototrophen Bakterien und der Diversität und Langzeitstabilität beobachtet.

In dieser Dissertation werden Auxotrophien mit Genom-skaliertem metabolischer Modellierung vorhergesagt. Hierbei handelt es sich um eine *in silico* Methode,

deren Vorhersagen jedoch in anderen Studien starke Abweichungen zu experimentell validierten Auxotrophien gezeigt haben. Daher ist ein wichtiger Bestandteil dieser Dissertation die Validierung von mit metabolischer Modellierung vorhergesagten Auxotrophien. Dafür wurden metabolische Modelle mit Genomen von experimentell bestätigten auxotrophen Bakterien rekonstruiert und die Auxotrophievorhersagen daraus noch mit einer zweiten *in silico* Methode, GapMind, verglichen. Beide Methoden zeigten eine Genauigkeit von über 90 %. Zusammenfassend wird in dieser Dissertation zum ersten Mal die Verteilung von Auxotrophien im Darmmikrobiom des Menschen beschrieben sowie ihr positiver Einfluss auf die Diversität und Langzeitstabilität. Auxotrophe Bakterien sind im Dünndarmmikrobiom erhöht im Vergleich zum Dickdarmmikrobiom im nicht dysbiotischen Zustand, aber reduziert in der Dünndarmfehlbesiedelung. Weiter zeigte die Validierung der Vorhersage von Auxotrophien durch metabolische Modellierung eine Genauigkeit über 90 %. Die Ergebnisse dieser Arbeit tragen dazu bei, das Verständnis der mikrobiellen Ökologie im Darm zu vertiefen und könnten neue therapeutische Ansätze zur Modulation des Mikrobioms in Krankheitszuständen aufzeigen.

# Inhaltsverzeichnis

<b>Abstract</b> .....	<b>II</b>
<b>Zusammenfassung</b> .....	<b>IV</b>
<b>Publications</b> .....	<b>IX</b>
<b>Figures</b> .....	<b>XI</b>
<b>Tables</b> .....	<b>XIII</b>
<b>Chapter I: General Introduction</b> .....	<b>1</b>
<b>1 The microbiomes of the human gastrointestinal tract</b> .....	<b>1</b>
1.1 Protein and amino acid metabolism in small and large intestinal microbiomes .....	2
1.2 Involvement of the microbial amino acid metabolism in the gut in inflammatory and metabolic diseases .....	5
1.3 Ecology of the human gut microbiome in health and disease .....	9
<b>2 Auxotrophic genotypes in microbial communities</b> .....	<b>11</b>
2.1 Distribution of auxotrophic bacteria in the human microbiome in health and disease .....	12
2.2 Origin of auxotrophies.....	13
2.3 Influence of auxotrophies on microbial ecology .....	15
<b>3 Methodical approaches for the prediction of auxotrophies</b> .....	<b>17</b>
3.1 <i>In silico</i> prediction of auxotrophies.....	18
3.2 <i>In vitro</i> prediction of auxotrophies .....	20
<b>4 Objectives</b> .....	<b>21</b>
<b>Chapter II: Amino acid auxotrophies in human gut bacteria are linked to higher microbiome diversity and long-term stability</b> .....	<b>23</b>
<b>1 Abstract</b> .....	<b>24</b>
<b>2 Background</b> .....	<b>24</b>
<b>3 Results</b> .....	<b>26</b>
<b>4 Discussion</b> .....	<b>36</b>
<b>5 Conclusion</b> .....	<b>41</b>
<b>6 Material and Methods</b> .....	<b>42</b>
<b>Chapter III: Comparative Analysis of Amino Acid Auxotrophies and Peptidase Profiles in Non-Dysbiotic and Dysbiotic Small Intestinal Microbiomes</b> .....	<b>49</b>
<b>1 Abstract</b> .....	<b>50</b>

2	Introduction.....	50
3	Results.....	52
4	Discussion.....	60
5	Conclusion .....	65
6	Material and Methods .....	65
<b>Chapter IV: Reliability of <i>in silico</i> auxotrophy prediction methods.....</b>		<b>69</b>
1	Abstract .....	70
2	Introduction.....	70
3	Material and Methods .....	71
4	Results.....	72
5	Discussion.....	74
6	Conclusions .....	75
<b>Chapter V: General Discussion.....</b>		<b>77</b>
1	Distribution of amino acid auxotrophies in small and large intestinal microbiomes .....	77
1.1	Gene loss of human-essential and metabolically expensive amino acids in small and large intestinal microbiomes.....	78
2	The influence of auxotrophic bacteria on the diversity and stability of small and large intestinal microbiomes .....	79
3	Nutritional ecology of amino acid auxotrophic bacteria in small and large intestinal microbiomes .....	82
3.1	Potential ways of amino acids acquisition by auxotrophic bacteria in small and large intestinal microbiomes .....	83
3.2	Ubiquity of amino acid auxotrophies in the human gut microbiome indicates a shift towards microbial nutritional ecology .....	85
4	Auxotrophic bacteria as biomarkers and therapeutic targets to restore the human gut microbiome in dysbiotic states .....	88
5	Amino acid auxotrophies in the gut microbiome: putative peptidolysis, SCFA synthesis, and their impact on human health.....	90
6	Limitation of <i>in silico</i> prediction of auxotrophies .....	93
<b>Chapter VI: Conclusion.....</b>		<b>97</b>
<b>Acknowledgements.....</b>		<b>98</b>
<b>Supplementary information for Chapter II .....</b>		<b>99</b>
Supplementary Material and Methods.....		100
Supplementary Figures.....		108



Supplementary Tables .....	116
Supplementary information for Chapter III .....	133
Supplementary Tables .....	134
References .....	135

## Publications

### Peer-reviewed papers

#### Published

Starke, S., Harris, D.M.M., Zimmermann, J. et al. Amino acid auxotrophies in human gut bacteria are linked to higher microbiome diversity and long-term stability. *ISME J* **17**, 2370–2380 (2023).

#### In revision for *Computational and Structural Biotechnology Journal*

Starke, S., Harris, D.M.M., Paulay, A. et al. Comparative Analysis of Amino Acid Auxotrophies and Peptidase Profiles in Non-Dysbiotic and Dysbiotic Small Intestinal Microbiomes

#### In preparation

Starke, S., Waschina, S. Reliability of *in silico* auxotrophy predictions

## Oral presentations and posters

Svenja Busche<sup>1</sup>, Danielle Harris, Konrad Aden, Silvio Waschina. Amino acid auxotrophies are ubiquitous in the human gut microbiome. Poster presented at the Summer School „Frontiers in Microbiome Research,” 12-15 June 2022, Ebberup, Denmark

Svenja Busche<sup>1</sup>, Danielle Harris, Konrad Aden, Silvio Waschina. Amino acid auxotrophies are ubiquitous in the human gut microbiome. Oral presentation at the 18<sup>th</sup> International Symposium on Microbial Ecology, 14-19 August 2022, Lausanne, Switzerland

Svenja Busche<sup>1</sup>, Danielle Harris, Konrad Aden, Silvio Waschina. Amino acid auxotrophies are ubiquitous in the human gut microbiome. Oral presentation at the workshop „Mathematical modelling of microbiomes”, 14-16 September 2022, Plön, Germany

Svenja Busche<sup>1</sup>, Danielle Harris, Konrad Aden, Silvio Waschina. Auxotrophies in human gut bacteria are indicative of microbiome stability. Poster and short talk at 13th International Gut Microbiology Symposium 2023, 13-15 June 2023, Aberdeen, Scotland

## Figures

<b>Figure 1: Protein and amino acid sources for bacteria in the small and large intestine.....</b>	<b>3</b>
<b>Figure 2: Microbial protein degradation by the human gut microbiome [adapted from 38] .....</b>	<b>5</b>
<b>Figure 3: Microbial amino acid metabolism associated with metabolic and inflammatory diseases .....</b>	<b>7</b>
<b>Figure 4: Forms of cross-feeding .....</b>	<b>16</b>
<b>Figure 5: Workflow for the prediction of auxotrophies with genome-scale metabolic modelling.....</b>	<b>27</b>
<b>Figure 6: Abundances of auxotrophies in 3 687 genomes.....</b>	<b>29</b>
<b>Figure 7: Associations of auxotrophies and fermentation products. ....</b>	<b>30</b>
<b>Figure 8: Distribution of auxotrophies in human gut microbiomes from 185 healthy adults, their association with diversity, and serum metabolite levels. ....</b>	<b>32</b>
<b>Figure 9: Influence of auxotrophies on long-term stability of the human gut microbiome .....</b>	<b>36</b>
<b>Figure 10: Flowchart of the study .....</b>	<b>53</b>
<b>Figure 11: Comparison of auxotrophy profiles in small and large intestinal microbiomes .....</b>	<b>55</b>
<b>Figure 12: Comparison of peptidase profiles in small and large intestinal microbiomes .....</b>	<b>56</b>
<b>Figure 13: Associations between peptidases and auxotrophies in duodenal (SI-Duodenum) and large intestinal (LI-stool) microbiomes.....</b>	<b>57</b>
<b>Figure 14: Auxotrophy profiles in the duodenal microbiomes of non-SIBO and SIBO samples .....</b>	<b>59</b>
<b>Figure 15: Microbial peptidase profiles in duodenal microbiomes of SIBO and non-SIBO samples .....</b>	<b>59</b>
<b>Figure 16: Comparison of <i>in silico</i> methods to predict auxotrophies.....</b>	<b>74</b>
<b>Figure 17: Nutritional ecology of amino acid auxotrophic bacteria in the gut..</b>	<b>83</b>
<b>Figure 18: Comparison of the microenvironments of the gut in non-SIBO and SIBO subjects .....</b>	<b>87</b>

<b>Supplementary Figure 1: Association between the number of auxotrophies and the genome completeness .....</b>	<b>108</b>
<b>Supplementary Figure 2: Validation of in silico predicted auxotrophies with experimental auxotrophies .....</b>	<b>109</b>
<b>Supplementary Figure 3: Overview of the proportions of auxotrophy to prototrophy genomes per phylum from the HRGM catalog [146].....</b>	<b>110</b>
<b>Supplementary Figure 4: Number of auxotrophies for every phylum.....</b>	<b>111</b>
<b>Supplementary Figure 5: Missing enzymes in amino acid biosynthesis pathways .....</b>	<b>112</b>
<b>Supplementary Figure 6: Amino acid auxotrophies predicted with varying genome completeness levels .....</b>	<b>113</b>
<b>Supplementary Figure 7: Spearman correlation between the dietary intake of amino acids and the frequency of amino acid auxotrophic bacteria in the gut microbiomes .....</b>	<b>114</b>
<b>Supplementary Figure 8: Percentage of quality-controlled metagenomic reads from three cohorts (this study, Troci <i>et al.</i> 2022 [150], and Chen <i>et al.</i> 2021 [151]) to reference genomes from the HRGM catalog.....</b>	<b>115</b>

## Tables

<b>Table 2: Cohort characteristics</b> .....	44
<b>Supplementary Table 1: Auxotropy prediction for HRGM models</b> .....	116
<b>Supplementary Table 2: Validation of <i>in silico</i> predicted auxotrophies with experimentally verified auxotrophies in human gut bacteria.</b> .....	117
<b>Supplementary Table 3: Auxotrophy prediction for AGORA2 models</b> .....	121
<b>Supplementary Table 4: Auxotrophy prediction for 124 prototrophic genotypes</b> .....	122
<b>Supplementary Table 5: Abbreviations used in Figure 8</b> .....	127
<b>Supplementary Table 6: Logical expression used as rules to decide the absence/presence of specific compounds in a genome-specific gap filling medium*</b> .....	129
<b>Supplementary Table 7: Relative abundance of HRGM genomes in samples from the Reimagine study</b> .....	134
<b>Supplementary Table 8: Relative abundance of HRGM genomes in samples from the SIBO study</b> .....	134
<b>Supplementary Table 9: Presence of peptidases in microbial samples from the duodenum, jejunum, farthest distance and stool</b> .....	134
<b>Supplementary Table 10: Statistical results for association of peptidases with SI(Dudenum) or LI(Stool)</b> .....	134
<b>Supplementary Table 11: Presence of peptidases in duodenal samples from SIBO and non-SIBO patients</b> .....	134
<b>Supplementary Table 12: Statistical results for association of peptidases with SIBO or non-SIBO</b> .....	134

# Chapter I: General Introduction

## 1 The microbiomes of the human gastrointestinal tract

The human gastrointestinal tract is colonized by microorganisms influencing human health [1]. The microorganisms in the human gastrointestinal tract are referred to as the human gut microbiome and include bacteria, viruses, archaea, and fungi [2]. The gut microbiome influences biological processes in humans [1], [2]. Alterations in the gut microbiome composition are related to the pathophysiology of several metabolic and inflammatory diseases [1]. This section gives an overview of the diversity of bacteria found in the human gastrointestinal tract and how their metabolisms are related to the pathophysiology of diseases. Since this thesis analyses amino acid auxotrophies in human gut bacteria, this introduction focuses on the involvement of microbial amino acid metabolism in metabolic and inflammatory diseases.

Phyla found in the microbiomes from the oral cavity to the large intestine are Firmicutes, Actinobacteria, Proteobacteria, Verucomicrobiota, and Bacteroidetes, but their abundances vary in gastrointestinal regions [3]. The oral cavity is colonized by aerobic microorganisms, mainly by *Streptococcus* (Firmicutes) [3]. As a connection between the oral cavity and stomach, the esophagus is found to have a similar microbiome to the oral cavity, with *Streptococcus* being the most abundant genus [4]. With its acidic environment, the stomach harbors only a few bacteria, which belong to *Helicobacter*, Firmicutes, and Proteobacteria [5], [6]. The lower part of the gastrointestinal tract is divided into the small and large intestine. Firmicutes and Proteobacteria are most abundant in the small intestinal microbiome, but Bacteroidetes, a prominent colonizer of the large intestine, are less abundant [7].

Many studies showed the involvement of the gut microbiota in human health. The oral microbiome is involved in the development of caries, periodontitis, and other diseases located outside the oral cavity, like inflammatory bowel disease (IBD) [8], [9]. Further, the oral microbiome is associated with the risk of cancer in the esophagus [10]. *Helicobacter pylori* is known to be involved in stomach cancer [11]. Small and large intestinal microbiomes are further involved in various inflammatory and metabolic diseases [1], [12]. Thus, the microbe-host relationship is important for human health. Microbes interact with the host by the synthesis of various metabolites. A major focus has been on the bacterial synthesis of short-chain fatty acids (SCFA), namely acetate, butyrate, and propionate in the gut, which is known to be involved in inflammatory (e.g.,

IBD) and metabolic diseases (e.g., type 2 diabetes, obesity) [13], [14]. The gut microbiota produces metabolites from dietary compounds or endogenous compounds. Some metabolites act directly in the gut (e.g., on the epithelium barrier), and others are absorbed across the mucosa and transported into the bloodstream [15]. Moreover, extensively studied metabolites are tryptophan and its derived metabolites in IBD [16], [17]. Other microbiota-produced metabolites, like branched-chain amino acids, bile acids, and their derivatives, are also well studied for their impact on human health [18]. In conclusion, the gut microbiota is an important factor that influences human health. Since people with metabolic and inflammatory diseases are rising worldwide [20], [21], research about the host-microbe relationship promises new potential therapeutic approaches.

### 1.1 Protein and amino acid metabolism in small and large intestinal microbiomes

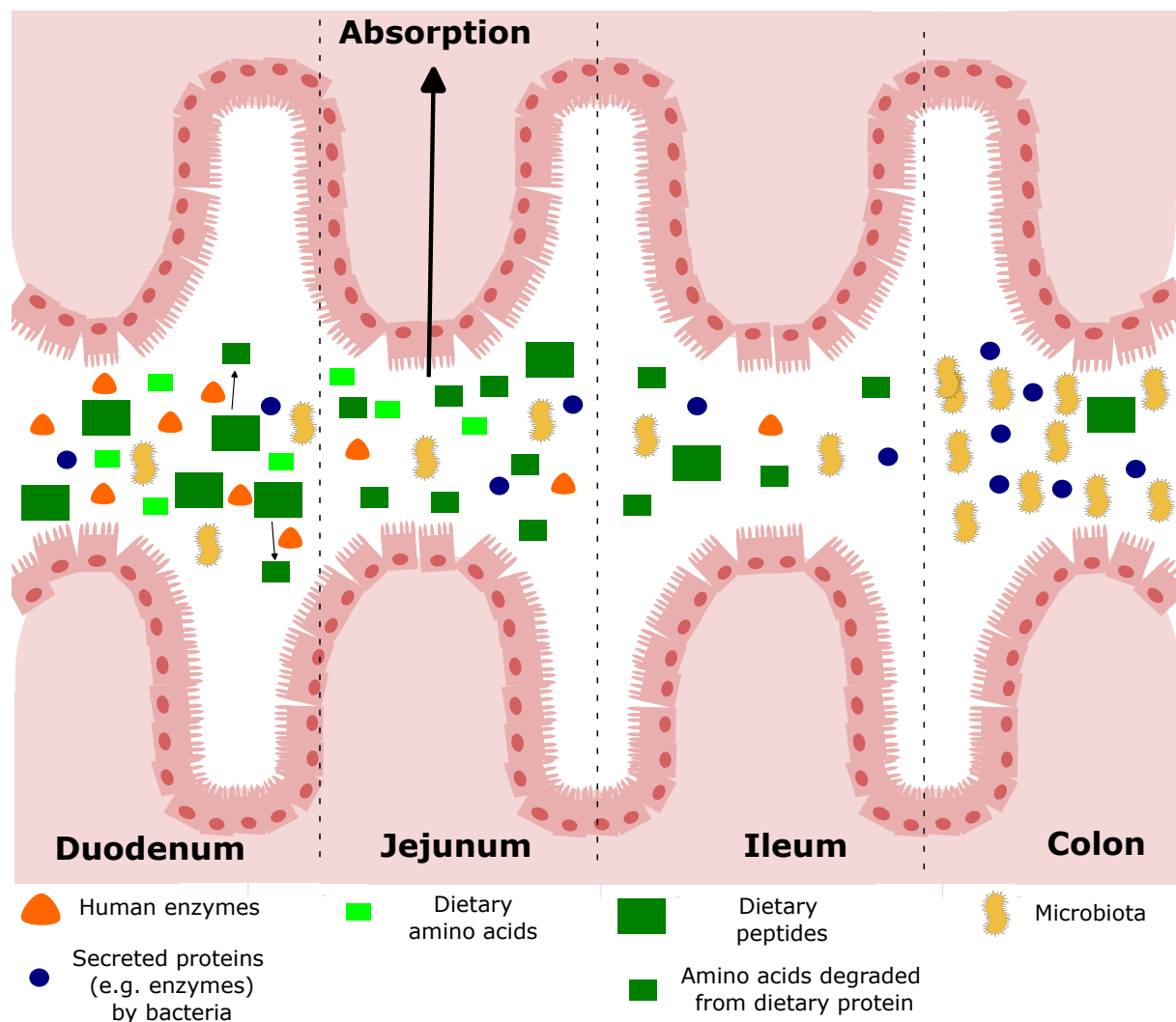
Proteins, peptides, and amino acids are energy and carbon sources for gut bacteria [19]. Auxotrophic bacteria are incapable for the *de novo* biosynthesis of amino acids, so they must acquire amino acids from their environment (see definition in subchapter II) [20]. Therefore, the availability of amino acids is crucial in the environment of auxotrophic bacteria. So far, sources for amino acids and ways for nutrient acquisition (e.g., peptidolysis) are not known for auxotrophic bacteria in the small and large intestines but will be a topic of this thesis. Therefore, this subchapter gives information about the current knowledge of protein and amino acid metabolism in small and large intestinal bacteria in general.

The nutritional landscape of proteins, peptides, and amino acids in the small and large intestine is influenced by dietary protein digestion by human enzymes, secretion of host and bacterial proteins, and peptidolysis by bacteria in small and large intestinal microbiomes (Fig.1) [21], [22], [23], [24]. Free amino acids are available for bacteria in the small intestine but inaccessible for bacteria in the large intestine [21]. They do not reach the colon in substantial amounts as they are fully absorbed in the small intestine; therefore, they are particularly a source of small intestinal bacteria [21]. Indeed, the metabolism of amino acids and proteolytic activity is strongly present in the small intestinal microbiome [25].

Peptidolysis is another way bacteria acquire amino acids in the gut, especially in the large intestine [26], [27]. Small and large intestinal microbiomes differ quantitatively



and qualitatively in proteolytic behavior [24]. The small intestinal microbiome generally has higher proteolytic activity than the colonic microbiome [27]. More soluble proteins were observed in the small intestine [28], which are more prone to be degraded by gut bacteria [22]. Bacteroidetes, more abundant in the colon, prefer host proteins (e.g., mucus), whereas Firmicutes, strongly present in the small intestine, prefer dietary proteins, which are more available in the small intestine [21]. Bacteria in the small intestine can degrade dietary proteins like casein and proteins in soybean flour [24]. Access to nutritional proteins is limited for bacteria in the lower gastrointestinal tract because of human enzyme protein degradation and amino acid absorption in the small intestine [29], which explains the proteolytic behaviour of bacteria in the small and large intestine.



**Figure 1: Protein and amino acid sources for bacteria in the small and large intestine**  
The proteinaceous environment in the small and large intestine.

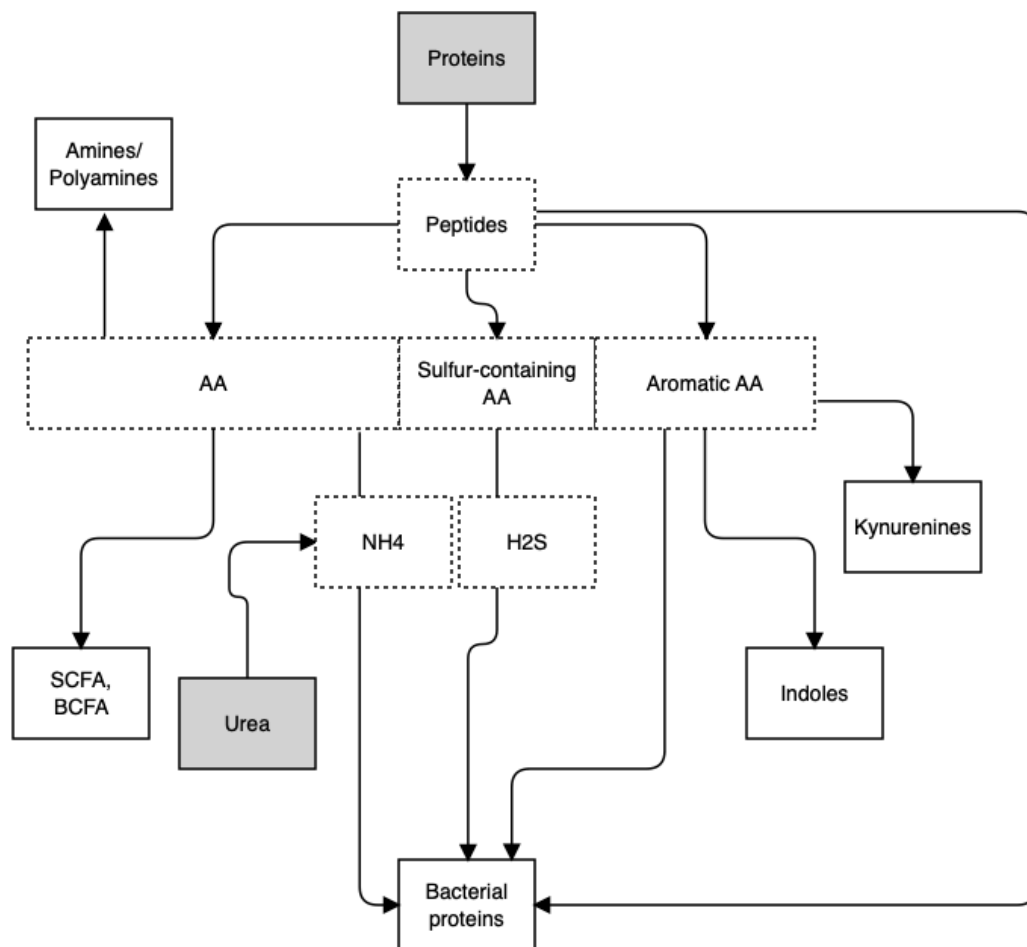
However, the colonic microbiota also has proteolytic potential towards dietary proteins [28]. Undigested dietary proteins can reach the large intestine and can be degraded by the residing bacteria [21].

Several other protein sources were identified as either preferred by small or large intestinal bacteria. *Akkermansia muciniphila* and *Bacteroides acidifaciens*, common colonizers of the colon, were identified as two bacterial species that can degrade host proteins [30]. Serum albumin was only degraded by colonic bacteria and not by small intestinal bacteria [24]. Pancreatic enzymes are also a favorable protein source of large intestinal bacteria [31]. Additionally, large intestinal bacteria degrade mucin [32]. Collagen was less degraded by bacterial proteases from the small and large intestinal microbiome [24]. Bacterial proteins like proteases are another protein source for gut bacteria. Bacteria secrete proteins in the large and small intestines, whereas more proteins (e.g., enzymes) are secreted by bacteria in the large intestines [23].

Consequences of the bacterial protein and amino acid metabolism are the production of various metabolites. Proteins, peptides, and amino acids are degraded into organic acids (e.g., SCFAs), kynurenines and indoles, and several gases like ammonia, CO<sub>2</sub>, H<sub>2</sub>, and H<sub>2</sub>S (Fig. 2) [19]. Quantitatively, the compounds derived by proteolysis are based on the protein source [22]. One study compared the synthesis of SCFA using different protein substrates. While acetate is abundantly produced across casein, bovine serum albumin, and collagen as substrates, butyrate is more readily produced using casein and bovine serum albumin. Meanwhile, propionate and succinate production are preferentially produced with casein. Further, several phenolic and indolic compounds were found in higher quantities after the proteolysis of casein [24]. Many phenolic and indolic compounds are formed by the microbial degradation of tryptophan and include indole, tryptamine, indole ethanol, indole propionic acid, indoleacetic acid, skatole, indole aldehyde, and indoleacetic acid [17]. Several strains belonging to *Bacteroidetes* spp. could degrade tryptophan into indolic compounds [33]. Further, sulfur-containing amino acids are microbially degraded to H<sub>2</sub>S. Human gut bacteria can degrade cysteine and methionine to H<sub>2</sub>S [34], [35].

In sum, the nutritional landscape of proteins, peptides, and amino acids in the intestines is shaped by several factors, including proteolysis. However, for auxotrophic bacteria in both the small and large intestines, the specific amino acids they rely on, as well as the presence of peptidase genes, remain unclear. Additionally, while amino

acids can be degraded into SCFAs and other compounds, the potential biosynthesis of SCFA of auxotrophic bacteria compared to prototrophic bacteria is not yet known.



**Figure 2: Microbial protein degradation by the human gut microbiome [adapted from 38]**  
 Pathways of the protein degradation by the human gut microbiome, grey-colored boxes are the substrates; dotted boxes include intermediate metabolites, and white colored the product; AA abbreviates amino acid.

## 1.2 Involvement of the microbial amino acid metabolism in the gut in inflammatory and metabolic diseases

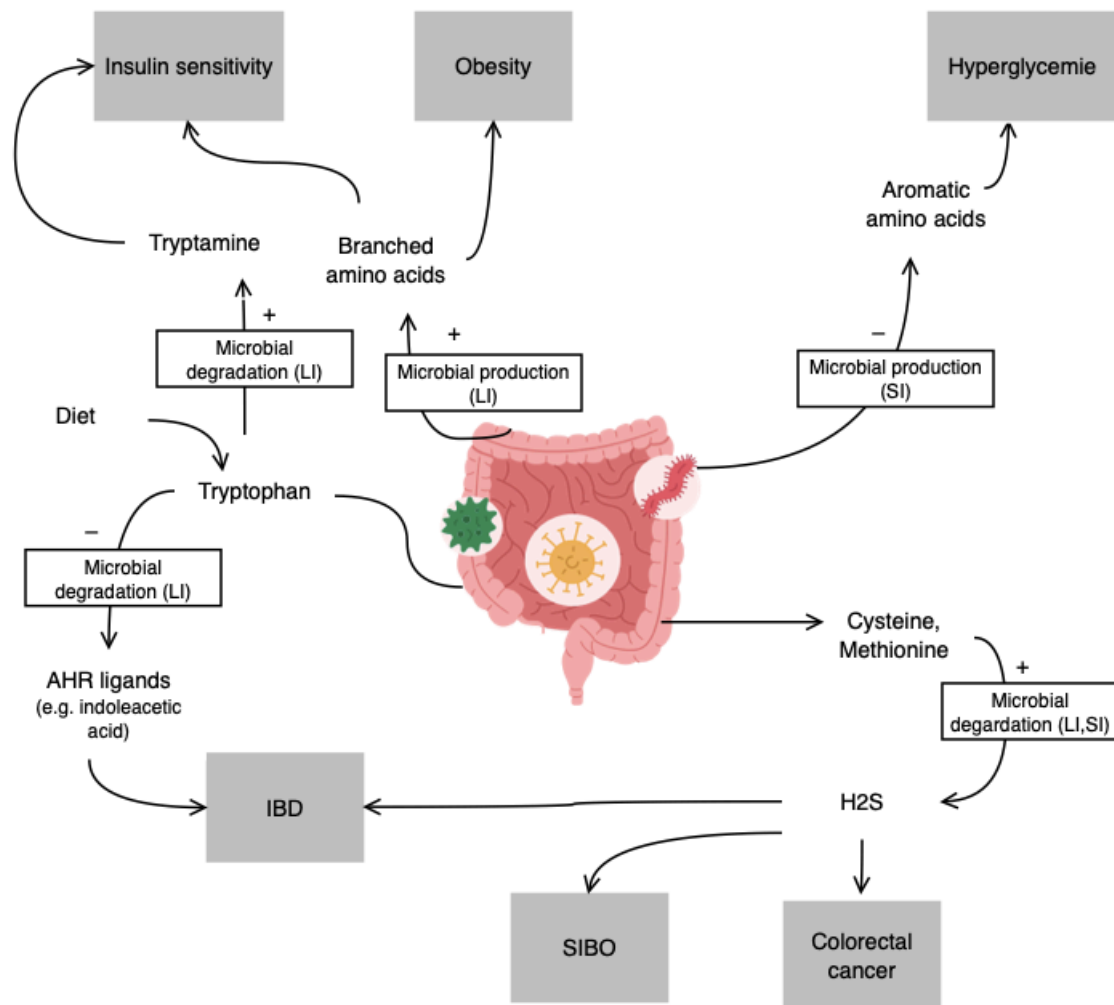
Amino acids are not only a fuel source for bacteria in the gut but for the host. Some amino acids are essential for humans because humans cannot synthesize them. Dietary proteins are digested, and essential and non-essential amino acids are absorbed into the bloodstream. Further amino acids exhibit directly effects in the gut by being a fuel source for the gut mucosa helping to maintain its integrity which is disrupted in diseases [36]. Bacteria in the human gut microbiome influence circulating levels of amino acids, which affects the host homeostasis [37]. The microbial amino

acid metabolism in the gastrointestinal tract is associated with disease progression and severity (Fig. 3) [38], [39], [40]. The biosynthesis of amino acids in the human gut microbiome is related to inflammatory and metabolic diseases and their pathophysiology [39]. Since this thesis is about amino acid auxotrophies in the human gut microbiome, this section gives an understanding of why studying the amino acid metabolism in the human gut microbiome is important and describes the current knowledge about microbial amino acid metabolisms in metabolic and inflammatory diseases. This section covers both the small and large intestinal microbiome. However, the main focus of research has been on the large intestinal microbiome and its association with diseases, as the microbial sampling process of the small intestine is still challenging [12]. Research on the metabolic functions of the small intestinal microbiome remains limited due to the challenges associated with sampling from this region.

One of the most extensively studied amino acids related to diseases is tryptophan. Several studies showed that tryptophan and its degradation products are involved in the pathophysiology of IBD. Tryptophan can be microbially degraded to kynurenine and indole derivatives [41]. Some tryptophan-derived metabolites like tryptamine and indoleacetic acid are aryl hydrocarbon receptor ligands (Ahr). Activation of the Ahr is involved in the immune response [17] and plays an important role in inflammatory diseases like IBD. A study showed that the gut microbiota in IBD has a reduced potential for the production of ligands of Ahr [42]. In addition, serum levels of tryptophan in IBD patients were found to be significantly decreased, and several of the serum levels of kynurenine pathway metabolites were altered in comparison to those of the control group [43] [44]. The microbial composition was statistically significantly different in high tryptophan serum level patients in comparison to low tryptophan serum level patients [43]. Tryptophan-derived metabolites also play a role in metabolic diseases. For example, tryptamine is involved in diabetes, and a recent study showed its effect on insulin sensitivity [38].

Additionally, tryptophan, the small intestinal microbiome and mood disorders are linked. In small intestinal bacterial overgrowth (SIBO), characterized by over colonization of  $10^5$  bacteria/ml and associated with several diseases, lower tryptophan in urine as well as higher levels of the kynurenine to tryptophan ratio were found together with anxiety and depression. After antibiotic treatment, the lactulose hydrogen breath test indicated a decrease in bacterial load in the small intestine, urinary tryptophan levels increased,

the kynurenine/tryptophan ratio decreased, and mental symptoms improved [45]. Also, other levels of tryptophan derivatives were affected by SIBO [46].



**Figure 3: Microbial amino acid metabolism associated with metabolic and inflammatory diseases**

The amino acid metabolism of small and large intestinal microbiomes is associated with several diseases. The gut microbiome shows a higher (+) or reduced (-) capacity for synthesis/degradation of the focal amino acid and its derivatives in diseases compared to healthy individuals. The icon was taken from flaticon (creator: Nortch), abbreviations: aryl hydrocarbon receptor – AHR, IBD – inflammatory bowel disease, LI – large intestine, SI –small intestine, SIBO – small intestinal bacterial overgrowth

Other amino acids involved in diseases are branched-chain amino acids (BCAAs). The microbial metabolism of BCAAs in the gut is associated with metabolic diseases and their pathophysiology. Insulin sensitivity is associated with higher serum levels of

BCAAs, and bacterial species in the gut are suggested to have an impact on serum BCAA levels [39]. Some species are the main drivers for this observed effect. *Prevotella copri* and *Bacteroides vulgatus* were presumably responsible for the biosynthesis of BCAAs. In a mouse model, the increase in BCAAs in serum and lower insulin sensitivity indicated the involvement of these species in the disease [39]. Interestingly, another study shed light on a bacterial species that positively affects insulin sensitivity. The genus *Faecalibacterium*, especially *F. prausnitzii*, showed promising potential on regulating BCAA serum levels as it negatively correlated to BCAA serum levels and the marker HOMA-IR in humans [47]. Not only the microbiome in the large intestine is involved in the pathophysiology of diabetes but the microbiome of the small intestine. A study comparing the duodenal and stool microbiota in hyperglycemia and normoglycemic groups found no difference in the stool microbiota but in the duodenal microbiota. The duodenal microbiota of hyperglycemic patients showed a reduced capacity for synthesizing aromatic amino acids, but the microbial synthesis of aromatic amino acids was negatively associated with the glycemic marker HbA1c [48].

In addition, the human gut microbiome might be related to the pathogenesis of obesity. In obese humans, associations between the human gut microbiome and BCAA serum levels were found [49]. A higher capability for producing BCAAs was found in the gut microbiome of patients with obesity. *Bacteroides thetaiotaomicron* in the gut microbiome affected amino acids in serum [50]. After the patients' weight loss, the gut microbiome showed a microbial amino acid profile similar to the lean control group [49]. However, another study in patients with obesity found no statistically significant association regarding amino acid metabolism [50].

Recent studies also showed the involvement of sulfur-containing amino acids in diseases [34], [51]. Sulfur-containing amino acids like cysteine and methionine can be microbially degraded to H<sub>2</sub>S (Fig. 2) [35]. The degradation of cysteine to H<sub>2</sub>S is more abundant in gut microbiomes of patients with IBD and colorectal cancer than in healthy individuals [34]. Indeed, measuring the exhaled amount of H<sub>2</sub>S was successfully tested as a diagnostic marker for colorectal cancer [52]. In high amounts, H<sub>2</sub>S can have pro-inflammatory effects in the gut and even promote tumorigenesis; however, in lower amounts, it can benefit the gut microbiome. It can protect the microbiome while administering antibiotics, stabilize microbial biofilms, and help maintain the gut mucus integrity [53]. Further, H<sub>2</sub>S in breath is elevated in patients with SIBO and was indicated

as a new biomarker for improving the diagnosis of SIBO in irritable bowel syndrome [54].

In conclusion, the microbial amino acid metabolism in the small and large intestines is involved in the pathophysiology of several diseases. Research about the human gut microbiome in health and disease has focused only on the presence of amino acid biosynthesis pathways. This thesis focus not on the amino acid biosynthesis but on the nutritional requirements (amino acids) of the human gut microbiome by studying amino acid auxotrophies in human gut bacteria in non-dysbiotic and dysbiotic states. Additionally, a direct influence of amino acid auxotrophies in SIBO is not known even if altered metabolomics of tryptophan and derivatives were observed. By analyzing auxotrophies in SIBO, our knowledge can be expanded on the etiology of SIBO.

### 1.3 Ecology of the human gut microbiome in health and disease

In the previous subchapters, it was highlighted that the microbial composition and metabolic profile in the gut is statistically significant different in diseases [39], [40], [49], [55]. In particular, the amino acid metabolisms is related to disease pathophysiology [39], [49]. Other characteristics observed in the human gut microbiome of patients in inflammatory and metabolic diseases is a reduced diversity and stability (Box 1) [56]. In the human gut microbiome, perturbations are antibiotics, dietary changes, and medications [57]. Additionally, less exchange of nutrients (termed cross-feeding) between microorganisms in the gut describes the ecology in the gut in disease states [58]. The following section will explore whether it is possible to define a “healthy” human gut microbiome and describe the current knowledge of the ecology of the human gut microbiome in diseases. Studying the ecology of the human gut microbiome can help expand knowledge about compositional and metabolic functional changes in diseases.

Researchers are trying to evaluate if there is a “healthy” human gut microbiome. So far, a “healthy” human gut microbiome can be characterised as having a stable core microbiome (Box 1), high taxonomic diversity and gene richness (Box 1) but there is not yet a healthy human gut microbiome defined on taxonomic level [59]. This kind of definition might never arise because of the high variability of the human gut microbiome and niche specialization of the human gut microbiome in healthy individuals, though metabolic functions are conserved across individuals [3].

The human gut microbiome in non-healthy individuals is often described as dysbiotic. A dysbiotic microbiome can be described as less adapted to its environment [60].

Recently, a study showed that the gut environment in diseases selects for specific bacteria, leading to a shift in the microbial composition. Bacteria are favored with more complete pathways for the biosynthesis of vitamins and amino acids, acting therefore independently [40]. Further, one characteristic of the human gut microbiome in dysbiotic states across several metabolic and inflammatory diseases (e.g., IBD) is a reduced diversity (Box 1) of the human gut microbiome [49], [61]. A highly diverse gut microbiome protects against the invasion of pathogens and is more resilient towards perturbations [62]. Recent studies showed that microbial communities in the gut are more stable with higher diversity [59].

Another contributing factor to gut ecology is cross-feeding between microorganisms in the gut (Box 1, subchapter 2.3. Cross-feeding is established in the human gut microbiome and one study estimated that approximately 90 % of the produced metabolites by the human gut microbiome are secreted in the intestinal environment some of which are utilised by other microorganisms [63]. Cross-feeding is reduced in the human gut microbiome in disease states [58].

**Box 1: Definitions of terms describing the ecology of the human gut microbiome**

**Stability:** Stability refers to the ability of a microbial systems to maintain the microbial composition or restore it [64]. The stability of the human gut microbiome is measured by the dissimilarity index between samples over time referred to as beta-diversity [65]. There are two terms describing stability in microbial systems. Resistance refers to the microbiome to survive external stress and resilience to survive perturbations by restoring the microbial composition to its previous states [65].

**Diversity:** The diversity refers to the number of different species found in the human gut microbiome [56]. The diversity of the human gut microbiome is analysed by calculating the alpha-diversity. The alpha-diversity refers to the number of taxonomic units present in a sample (richness) or their distribution (evenness) [65].

**Cross-feeding:** The exchange of nutrients between organism is determined as cross-feeding. Microorganisms in the human gut consume and produce metabolites [63].

**Core microbiome:** The core microbiome refers to the taxonomic composition of the microbiome shared between two and more samples and is a characteristic of a host [66].

In conclusion, a diverse human gut microbiome is intrinsically more stable, with higher resilience to perturbations and greater protection against pathogen invasion. In disease states, both diversity and cross-feeding are reduced. It is already known that auxotrophies influence the ecology in microbial communities (Chapter I 2.3). However,



it remains unclear how amino acid auxotrophies influence gut ecology, particularly in terms of diversity and stability. This presents an area for future research to deepen our understanding of the human gut microbiome in health and disease and will be the main topic of Chapter III.

## 2 Auxotrophic genotypes in microbial communities

Auxotrophy defines an incapability to *de novo* synthesise essential nutrients like amino acids and B-vitamins due to a loss of biosynthetic genes [20]. This thesis focuses on amino acid auxotrophies in bacteria. Amino acids are building blocks of proteins and an energy source for bacteria [67]. Since auxotrophic bacteria are missing the biosynthetic capability of essential nutrients like amino acids, they are dependent on the essential nutrient being available in the environment for survival. Prototrophic bacteria are able to *de novo* synthesise their essential nutrients. Those two genotypes shape microbial communities [20].

Auxotrophic bacteria are common in microbial communities [68], [69]. However, they are more prevalent in host-associated microbiomes and fermented foods. For example, bacteria in several body habitats of humans showed high number of auxotrophies [68]. In free-living habitats like soil and marine, the number of auxotrophies is reduced in bacteria [68]. Essential nutrients for auxotrophic bacteria might be more available in food-associated and host-associated environments. Habitats with Proteobacteria being the most dominating taxa, like rhizosphere and soil, had the lowest number of auxotrophic bacteria, whereas habitats like the human gut dominated by Firmicutes showed more auxotrophies [68]. Most frequently found are auxotrophies for leucine, isoleucine, valine, and less frequently, for asparagine, glutamine, and glycine auxotrophies [68].

So far, research about amino acid auxotrophies in the human microbiome and their involvement in human diseases is limited though several amino acid auxotrophies are found in pathogenic bacteria, for example auxotrophic variants of *Pseudomonas aeruginosa* are associated with patients infected with cystic fibrosis (CF) [70], [71]. However, several studies have already investigated auxotrophies for B-vitamin auxotrophies in the human gut microbiome [72], [73], [74], [75]. B vitamins are co-enzymes or precursor molecules for co-enzymes in bacteria [72].

Auxotrophic bacteria are not only associated with diseases like cystic fibrosis (CF) but also offer promising opportunities for therapeutic applications. One particularly

intriguing area of research is their potential use in cancer treatment. Cancer cells require amino acids and auxotrophic bacteria can starve cancer cells to death by competing with them for amino acids. For example, *Salmonella typhimurium*, auxotrophic for leucine and arginine, was effective in prostate cancer cells and stopped the proliferation of the amino acid-dependent cancer cells in a mouse model [76]. Successful interventions of a tryptophan auxotrophic *S. typhimurium* were further observed in breast cancer [77]. Though there is another approach for auxotrophic bacteria that is the use of auxotrophic strains as biosensors to detect amino acid concentrations. For example, *Escherichia coli* was successfully engineered as a biosensor [78].

Conclusively, auxotrophs are ubiquitous in microbial communities across different habitats. They are already related to the pathogenesis of CF, used as therapeutic agents against cancer and considered as biosensors. The following subchapter states the current knowledge of auxotrophies in the human microbiome with focus on amino acid auxotrophies. Further, theories are examined regarding the origin of auxotrophies. Finally, the current understanding of how auxotrophies influence microbial ecology is summarized. By presenting an overview of auxotrophic genotypes in microbial systems, this chapter aims to highlight the significance of studying amino acid auxotrophies in the human gut microbiome.

## 2.1 Distribution of auxotrophic bacteria in the human microbiome in health and disease

In humans, several auxotrophic bacteria were observed with some presumably colonizing disease tissue, but others being common colonizers [63], [70]. Recently, a study showed their broad distribution in the oral cavity, the human gut, and on the human skin. Bacteria living in the human gut and oral cavity had 2-3 amino acid auxotrophies on average, with the oral cavity having the highest number of auxotrophies [68]. In the human gut microbiome particularly, auxotrophies for amino acids are still understudied. One study analyzed the presence of genes in the amino acid biosynthesis pathways in human gut bacteria and a loss in genes were found in the biosynthesis pathways for tryptophan, methionine, arginine and BCAA [79]. Another study analyzed auxotrophies for amino acids in several human gut bacterial strains *in vitro* and determined several auxotrophies [72]. This study further experimentally validated B-vitamin auxotrophies. B-vitamin auxotrophies for biotin,

folate, and thiamine were observed in human gut bacterial strains [72]. In particular, auxotrophies for B-vitamins are already well studied. In a genome-scale metabolic modeling approach, over 800 gut bacterial strains were studied, and over 70% of the strains had at least two B-vitamin auxotrophies. Thiamine was the most common auxotrophy among the bacterial strains, followed by riboflavin and vitamin K. Interestingly, no auxotrophies for vitamin B12 were found [74]. Multiple studies gave a broad overview of the presence of genes involved in the biosynthesis of B-vitamin in human gut bacteria and found several missing genes in the biosynthesis pathways for B-vitamins [80], [81]. In addition, one study found that more than 20% of human gut bacteria have auxotrophies for B-vitamins using a gnotobiotic mouse model. For example, in *Faecalibacterium prausnitzii* and *Akkermansia muciniphila* several auxotrophies were detected [73].

As already mentioned in the introduction of subchapter 1.2, pathogenic bacteria, thriving in host-associated environments, have several amino acid auxotrophies [82]. One well studied example of auxotrophic bacteria in host-associated environments are the abundance of auxotrophic *Pseudomonas aeruginosa* in the lungs of patients with cystic fibrosis (CF) [82], [83]. Auxotrophic variants of *P. aeruginosa* were able to grow in the amino acid-rich sputum of CF patients [84]. In severe cases, the proteolytic activity by host enzymes produces a high amino-acid content in the lungs and auxotrophic *P. aeruginosa* thrives in this amino acid-rich environment [85]. In the genera *Klebsiella*, *Yersinia*, *Escherichia*, and *Salmonella* multiple auxotrophies were found. In several *Staphylococcus aureus* strains, amino acid auxotrophies are further confirmed [86]. *Staphylococcus aureus* is known to be arginine auxotrophic which impacts its virulence in the hosts [87].

Conclusively, auxotrophic bacteria are common in the human microbiome. So far, several amino acid auxotrophic bacteria were observed in pathogens colonizing humans. Additionally, B-vitamin auxotrophies are well studied in the human gut microbiome. However, amino acid auxotrophies are still not well characterized in the human gut microbiome.

## 2.2 Origin of auxotrophies

While the previous subchapters described the distribution of auxotrophic bacteria across various environments, they did not address the theories regarding their

development. This chapter explores what drives gene loss in bacteria and whether this process occurs randomly or is influenced by environmental factors.

Results of experimental designs showed that the environmental availability of amino acids led to an increase in auxotrophic genotypes, and the number of auxotrophies increased with the number of generations [88]. Especially functions for regulative proteins and enzymes are lost [89]. The theory behind the loss of genes as a consequence of a nutrient-rich environment is called the “Black Queen Hypothesis” which posits that organisms in microbial communities tend to lose genes that are metabolically costly to maintain, provided that the required metabolites are supplied by other community members (Box 2) [90].

**Box 2: Black Queen Hypothesis and the reduction of metabolic costs**

The hypothesis behind the reduction of metabolic costs in organisms is called the “Black Queen Hypothesis”. The name refers to the card game. In the game, the black queen card is described as an undesirable good but must be provided by some players to play the game. Some genes come with high metabolic costs but can be provided by other organisms in the community. Therefore, in nutrient-rich environments, it is beneficial for organisms to lose genes that come with high metabolic costs. However, the metabolite must be provided by other community members in order to favor these organisms and compensate for the gene loss [90].

Bacteria lose particularly biosynthetic genes for metabolites with high costs [69] because it is beneficial for them to save metabolic costs [88]. The synthesis of metabolites (e.g. amino acids) comes with costs, which are defined as consumed mol ATP per synthesized molecule. Some metabolites come with higher costs than others. For example, the biosynthesis of methionine costs 20 mol ATP whereas the synthesis of glycine is neutral [91]. Studies support the theory, that the adaptive loss of genes is one reason for the development of auxotrophies in bacteria [69], [89], [92]. Particularly, in host-associated environments, which are more nutrient-rich, amino acid auxotrophies in bacteria were found [68]. As described earlier, in severe CF patients, more auxotrophic variants of *P. aeruginosa* are observed, suggesting an adaptive behavior towards the amino acid-rich environment in the lungs due to the proteolytic activity by host enzymes [85]. High amino acid levels support the formation of biofilms in the lungs of CF patients [93]. Another example supporting this theory is the mucin-related bacterium *Akkermansia muciniphila* that was found to be auxotrophic for threonine, which is provided by the protein-rich backbone of mucin in the gut [94].

Further, many amino acid auxotrophies are found in lactic acid bacteria. The protein-rich environment in milk products favors amino acid auxotrophies in bacteria [95].

The loss of genes can confer a fitness advantage, which may explain why the loss of genes involved in amino acid biosynthesis pathways in nutrient-rich environments is driven by selection rather than genetic drift [90] [88]. For example, a fitness advantage of auxotrophic strains was found in various environments, including the intestines of animals and humans and bladder infections [82]. However, the fitness gain is dependent on which enzymes in the pathway are lost [69]. Additionally, the available carbon source influences the metabolic costs: different carbon sources lead to varying metabolic expenses for amino acid biosynthesis. For example, amino acids with precursors in glycolysis and the pentose phosphate pathway are metabolically cheaper produced with glycolytic metabolites than gluconeogenic metabolites. Thus, the growth of auxotrophs is affected not only by the availability of amino acids in the environment but also by the available carbon source [96].

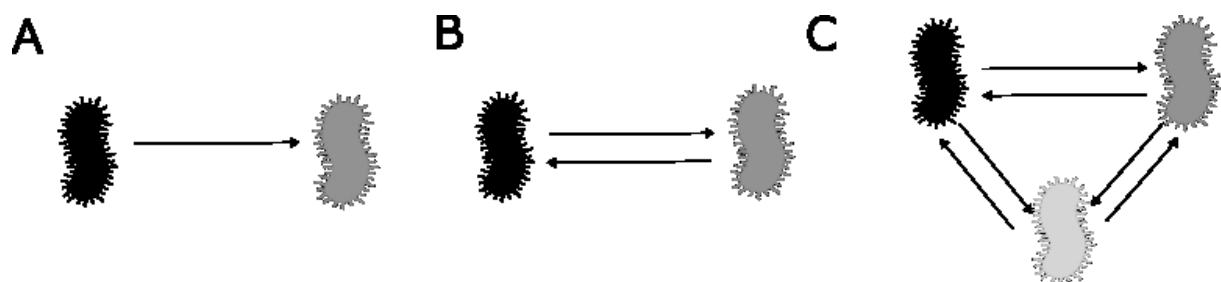
While genetic drift could contribute to gene loss in bacteria, particularly in nutrient-rich environments where random gene deletions are less detrimental [97], studies suggest that amino acid auxotrophies are more frequently driven by selection [69]. This underscores the adaptive nature of gene loss in response to environmental conditions. In summary, amino acid auxotrophies are likely driven by the loss of metabolically costly genes in nutrient-rich environments, where external sources of these metabolites reduce the selective pressure to maintain their biosynthetic pathways. However, for the human gut microbiome, it remains unclear whether auxotrophies for metabolically expensive amino acids are more prevalent in the small intestine, with its amino acid-rich environment, compared to the large intestine. To date, no studies have systematically compared the abundance of amino acid auxotrophies between these two gastrointestinal regions.

### 2.3 Influence of auxotrophies on microbial ecology

Since auxotrophies are dependent on their nutritional environment, they have to secure amino acids available in their environment. Auxotrophic bacteria can secure amino acids released by prototrophic bacteria. Indeed, the presence of auxotrophies in microbial communities drives the cross-feeding of amino acids [98]. The following section will determine how cross-feeding in auxotrophic communities influences the

ecology of microbial systems and what benefits bacteria have from engaging in cross-feeding.

There are different forms of cross-feeding: unidirectional, bidirectional, and multidirectional (Fig. 4). In co-culture experiments, all of these interactions were observed in communities with auxotrophies [99]. Unidirectional interactions between microbes can become reciprocal, and both microbes can benefit [69]. Cross-feeding in communities can decrease the number of generalists and increase the number of specialists [100]. Some microbial community members lose genes for costly metabolites, thereby becoming specialists in synthesizing metabolites for whose genes were kept. This is a strategy to save metabolic costs in microbial communities because decreasing individual costs also means that the community benefits from the labor division between the community members. Therefore, the community can benefit from gene loss.



**Figure 4: Forms of cross-feeding**

A) Unidirectional cross-feeding, B) Bidirectional cross-feeding, C) Multidirectional cross-feeding [101]

Cross-feeding from auxotrophic to prototrophic bacteria was observed in several studies [88], [102], [103]. Providing amino acids by *Saccharomyces cerevisiae* to lactobacilli promoted their growth [103]. Syntrophic interactions were observed in a 14-member community with monoauxotrophs grown together. [102]. In a fly model, cross-feeding led to the survival of an auxotrophic species in the gut despite the amino acid-lacking diet [104]. Therefore, in detrimental nutritional environments, cross-feeding protects auxotrophic species [105]. In addition, both bacteria engaging in cross-feeding benefit from fitness gain due to reduced metabolic costs by gene loss [98].

Further, cross-feeding showed an impact on the microbial ecology since it indicated a higher diversity in microbial communities over time [106]. In an experiment, it was observed that obligate cross-feeding increased the size of metabolic niches in communities which could enable the survival of more species [107]. Additionally, the

diversity is connected to the stability as an increase in diverse networks lead to more stable communities [108]. Indeed, it was observed that, cooperation positively impacted stability in microbial communities when enough nutrients for auxotrophs were provided in the microbial network [109].

Auxotrophic bacteria play a crucial role in shaping microbial ecology by driving cross-feeding interactions, which positively affect community diversity and stability. However, the influence of auxotrophic bacteria on human gut ecology, and the specific role of cross-feeding within this context, remains largely unexplored. Further studies are needed to determine how auxotrophies and cross-feeding interactions contribute to the dynamics, diversity, and stability of the human gut microbiome.

### 3 Methodical approaches for the prediction of auxotrophies

Auxotrophies are experimentally verified through bacterial cultivation but can also be predicted using *in silico* methods. Since cultivating bacteria from the human gut microbiome is challenging, *in silico* predictions offer an alternative methodological approach. Experimental and *in silico* methods for auxotrophy determination will be described in the following section.

*In vitro*, auxotrophies are predicted by comparing the bacterial growth with and without the amino acids. If the bacteria cannot grow without the amino acid, then they are considered auxotrophic for the amino acid, otherwise they are prototrophic [20]. For example, tryptophan auxotrophy is predicted if no bacterial growth is observed without tryptophan being in the growth medium. Only a few experiments have been conducted to predict auxotrophies *in vitro* in human gut bacteria, largely due to challenges in cultivating these organisms [110]. More generally, predicting auxotrophies and designing appropriate growth media remains difficult. However, alternative approaches using *in silico* methods have emerged. Due to the rise in next-generation sequencing, it is possible to predict bacterial metabolism based on genome analysis. Based on genome analysis, metabolic models can be reconstructed using automatic reconstruction software like gapseq [111] and CarveMe [112]. There is also a collection of genome-scale metabolic models reconstructed of >7000 gut microorganisms called AGORA2 [113]. With reconstructed genome-scale metabolic models, auxotrophies can be predicted with metabolic modeling. Next to metabolic modeling is genome annotation, another method used for auxotrophies prediction and is used by GapMind [114]. Auxotrophies are predicted by identifying gaps in the biosynthesis pathways of

amino acids [114]. However, several experiments showed that results of *in vitro* and *silico* auxotrophy predictions are distinct [72], [115], [116]. Therefore, validation of *in silico* auxotrophy prediction method is rather crucial to confirm predicted auxotrophies. In this thesis, auxotrophies are predicted by genome-scale metabolic modeling. The following subchapter will present the methodological background of metabolic modeling and other *in silico* methods for auxotrophy predictions. Further, the *in vitro* approach for auxotrophy verification will be explained, along with how *in silico* methods might help cultivate more bacteria.

### 3.1 *In silico* prediction of auxotrophies

*In silico*, auxotrophies can be predicted by genome annotation and genome-scale metabolic modeling. This section will focus first on genome annotation by describing the methodological approach of GapMind. Following this, genome-scale metabolic modeling will be explained, covering both the reconstruction of genome-scale metabolic models by gapseq [111] and the methodology of metabolic modeling.

Genome annotation relies on the presence/absence of genes in the query genome. GapMind predicts the confidence of genes in amino acid biosynthesis pathways [114]. Genes for amino acid biosynthesis pathways are found by searching for similarities with proteins from databases. The databases Swiss-Prot, Tiger-TIGRFam, and Pfam are used. However, some proteins were curated by the authors of GapMind based on published publications [114]. Three confidence levels are stated based on the identity thresholds and the query genome's coverage. To validate GapMind, the authors tested 35 bacteria grown in minimal media. GapMind does not state whether amino acid biosynthesis pathways are present or absent in the genome of interest but states confidence levels [114]. However, in a recent publication, that predicted auxotrophies with GapMind, auxotrophies were stated when 40 % of the reactions were missing [68]. The other *in silico* method for auxotrophy prediction is genome-scale metabolic modeling. Genome-scale metabolic modeling belongs to the constraint-based reconstruction and analysis (COBRA) framework [117]. The idea behind constrained-based metabolic modeling is to predict the growth of an organism with defined constraints and acquire knowledge about consumed and produced metabolites (Box 3). The reconstruction process of a genome-scale metabolic model starts with the genome sequence. This process will be explained using the reconstruction process of gapseq as an example since metabolic models in this thesis are reconstructed with



gapseq (Box 3). First pathways and transporters are predicted, and based on this information, a draft metabolic network is constructed [117]. Gaps in the draft network are filled to ensure the functionality of the model. Gaps are missing reactions in a metabolic network, which can prevent the model from growing. They exist because genome annotations are incomplete and lack the function of enzymes, which leads to incomplete metabolic networks [118]. Gapfilling is considered to add missing reactions to the model to enable the growth of the reconstructed model. However, the media used in the gapfilling process must be chosen carefully to avoid false implementation of reactions [119]. After the gap filling, the final model is reconstructed and can be used for metabolic modeling with flux-balance analysis. Flux-balance analysis calculates the flow of nutrients in the network and helps to predict a phenotype [120]. Mathematically, it is represented as a set of linear equations that are solved by defining an objective function (e.g., biomass formation). Seif et al. developed an algorithm named AuxoFind that uses flux-balance analysis based on the prediction of conditionally essential genes [82]. Further, reconstructed models from CarveMe can be used for auxotrophies prediction with metabolic modeling [121], [122]. In this thesis, flux-balance analysis is used on reconstructed models by gapseq to predict auxotrophies. The model's growth is compared with and without the amino acid using flux balance analysis with the objective function of biomass formation [120]. If the model cannot form biomass without the amino acid, it is auxotrophic. For each amino acid, the process is repeated to predict the auxotrophies for all twenty proteinogenic amino acids.

**Box 3: Genome-scale metabolic modeling**

Genome-scale metabolic models are reconstructed based on the genome sequence. They can be manually or automatically reconstructed. Since genome-scale metabolic models are reconstructed with gapseq in this thesis, the automatic model reconstruction process of gapseq will be explained. The model reconstruction process of gapseq consists of three steps [111]:

- 1) The genome sequence is analyzed for the presence of pathways and transporters. Blast searches for sequence evidence using the genome sequence as the query and the genome as the database. Metabolic pathways are present if 80% of the reactions are found.
- 2) A draft model is reconstructed based on the metabolic pathways and transporters found in step 1.
- 3) To ensure the model's reliability, gaps are filled, and the medium for the gap filling is predicted based on previously gained information about the absence/presence of pathways/transporters.

The final model is ready for genome-scale metabolic modeling with flux-balance analysis [120]. The metabolic network is converted into a stoichiometric matrix  $S$  with columns as the reactions and rows as the metabolites. Flux balance analysis belongs to constraint-based modeling. The constraints are defined by a steady state and upper/lower bound. The steady state means that there is always an equilibrium for every compound being consumed and produced. Linear equations define mass balance equations in a steady state ( $S \cdot v = 0$ ). The actual flux  $v$  through the metabolic networks lies between an upper and lower bound ( $a_i < v_i < b_i$ ). The flux distribution in the metabolic network is defined by an objective function  $z$  (e.g., biomass formation  $z = \text{biomass}$ ). Linear programming finds the optimal solution by calculating the fluxes that maximize the growth in the solution space defined by the constraints [120].

### 3.2 *In vitro* prediction of auxotrophies

Auxotrophies are experimentally verified by cultivating bacteria *in vitro*. Still, designing growth media for auxotrophy prediction remains challenging. In standard growth media, single components like amino acids and vitamins are not included but complex components. The presence of individual components like amino acids is, however, necessary for the prediction of auxotrophies as the growth of the bacterium is questioned with and without the amino acid. Complex components (e.g. yeast extract, peptone) lack information about the specific growth requirement of amino acids because their composition is not defined. Yeast extract, a complex compound often found in standard media, combines carbohydrates, amino acids, vitamins, trace elements, and oligomeric compounds [123]. Therefore, chemically defined media (CMD) are preferred as they contain no chemically undefined compound mixtures but only the minimal nutrients needed for the growth of the organism. In CMD, the proportions of the used pure chemicals are known, and no other undefined nutrients are included. The amino acids are as individual components included in the growth medium. The composition of CMDs can be indicated by the microbe's composition, or the complex medium can be used as a template [124]. In a study with *in vitro* investigation of amino acid auxotrophies, yeast extract was exchanged for single amino acids [72]. A method to obtain the growth medium, is for example to add and remove components from the CMD [125]. CMDs are already defined for several lactic acid bacteria that harbor several amino acid auxotrophies [126], [127].

Another strategy for designing growth media for *in vitro* verification of auxotrophies in bacteria is genome-scale metabolic modeling [117]. Under the condition that the genome is available, the nutritional requirements of the bacteria can be uncovered by

reconstructing a genome-scale metabolic model. With metabolic modeling, the essential nutrients for the growth of the bacterium can be defined. For example, a genome-scale metabolic model was reconstructed from *Campylobacter jejuni* M1cam to determine the growth medium for *in vitro* prediction of auxotrophies [128].

Conclusively, since *in vitro* verification of bacterial auxotrophies remains challenging, validating *in silico* predicted auxotrophies against experimentally verified ones could help demonstrate the reliability of auxotrophies predicted with *in silico* methods.

## 4 Objectives

Amino acid auxotrophies are found in bacteria inhabiting different environments. They influence the ecology of microbial communities. However, their distribution in the human gut microbiome and their effect on the gut ecology is not known. In this thesis, the main approach is to characterize amino acid auxotrophies in human gut bacteria and their influence on the ecology of the human gut microbiome.

In detail, the following questions are addressed in the thesis: (I) What amino acid auxotrophies can be found in the human gut microbiome? Are the abundances of amino acid auxotrophies different in small and large intestinal microbiomes? (II) How do the abundances of auxotrophies change in a dysbiosis state of the small intestinal microbiome? (III) What are the amino acid sources for auxotrophies in human gut bacteria? Is the dietary intake of amino acids the main nutrient source for auxotrophic gut bacteria? (IV) How do auxotrophic gut bacteria influence the stability and diversity of the human gut microbiome? (V) Are there associations between peptidases and amino acid auxotrophies in the human gut microbiome? (VI) How do auxotrophic gut bacteria influence the human blood metabolome? (VII) How accurate are amino acid auxotrophies predicted with genome-scale metabolic modeling? (VIII) What other *in silico* method exists for predicting auxotrophies? What differences are there in predictions compared to genome-scale metabolic modelling?

Issues are addressed by a combination of genome-scale metabolic modeling, genome-mining, and metabolomics. Auxotrophies in the human gut microbiome are predicted with genome-scale metabolic modeling using flux-balance analysis. Metabolic models are reconstructed from genomes of the human gut microbiome. Sequences of peptidases are taken from the MEROPS database. They are mapped on sequences of the human gut microbiome to study associations between auxotrophies and peptidases in small and large intestinal microbiomes. Experimentally

verified auxotrophies with available bacterial genomes were utilized for validation purposes. Metabolic models are reconstructed of genomes from experimentally verified auxotrophic bacterial strains.

First, results are presented and individually discussed in the manuscripts but later combined and argued in the general discussion. The distribution of amino acid auxotrophies and potential nutrient sources for auxotrophic bacteria inhabiting different gastrointestinal regions are reasoned. The effects behind the influence of auxotrophic gut bacteria on gut ecology are discussed. Further, the biological functions of auxotrophic gut bacteria and their potential impact on human health are argued. Genome-scale metabolic modeling as an auxotrophy prediction method is critically evaluated.

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

*Published in the ISME Journal, 17, p. 2370-2380, in October 2023 by Springer Nature under the license <http://creativecommons.org/licenses/by/4.0/>. The headlines of the supplemental figures legends were adapted.*

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

<sup>1</sup> Institute of Human Nutrition and Food Science, Nutriinformatics, Kiel University, Kiel, Germany

<sup>2</sup> Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany

<sup>3</sup> Research Group Evolutionary Ecology and Genetics, Zoological Institute, University of Kiel, Kiel, Germany

<sup>4</sup> Max Planck Institute for Evolutionary Biology, Plön, Germany

<sup>5</sup> Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM), Hanover, Germany

<sup>6</sup> Department of Internal Medicine III, University Medical Center Schleswig-Holstein, Kiel, Germany

<sup>7</sup> German Centre for Cardiovascular Research (DZHK), Partner site Hamburg, Kiel, Lübeck, Germany

<sup>8</sup> Department of Internal Medicine I, University Medical Center Schleswig-Holstein, Kiel, Germany

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

\*Corresponding authors

## 1 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 abundances were 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 contributing to microbiome ecology and host-microbiome metabolic interactions.

## 2 Background

The metabolic processes performed by the human gut microbiota have a crucial impact on human metabolism and health [14], [17], [67]. For instance, various human gut bacteria produce the short-chain fatty acid butyrate. Butyrate is a primary energy source for human colonocytes [14] and intersects with host immunological processes by mediating anti-inflammatory effects [129], [130]. 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 [131], [132]. 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 [43], [132], [133], [134].

The repertoire of molecules synthesized and eventually released by individual gut microbes comprises 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 [70], [137] or a growth-essential vitamin [72], [136]. In addition, *in vitro* growth experiments have confirmed specific amino acid and vitamin auxotrophies in common human gut bacteria [72], [94], [137].

The prevalence of auxotrophs in the human gut microbiome raises the question of the source of the required metabolites in the gastrointestinal growth environment. There are three potential sources of essential nutrients 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 [138], limiting the accessibility of diet-derived essential nutrients for the majority of the gut microbial community, which resides in the colonic region [139]. (ii) Metabolites required by auxotrophic microorganisms in the gastrointestinal tract may be host-derived, e.g., from proteins and peptides secreted by the gut epithelium into the gut lumen or from apical proteins of the host epithelial cell layer accessible to gut microorganisms [94]. (iii) Auxotrophic members of the gut microbial community might obtain essential nutrients via cross-feeding interactions with prototrophic organisms within their microbial community [140], [141].

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 [142], [143], [144], the extent of cross-feeding interactions via the exchange of essential nutrients 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 by exchanging the focal metabolites [145]. Furthermore, theoretical ecological models suggest that cross-feeding interactions between auxotrophic organisms within complex communities can increase community diversity through metabolic niche expansion [107] and community robustness to ecological

perturbance [109], 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 metagenomic 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 was associated with higher diversity of the gut bacteria and increased levels of aromatic compounds of putative microbial origin in the human serum metabolome.

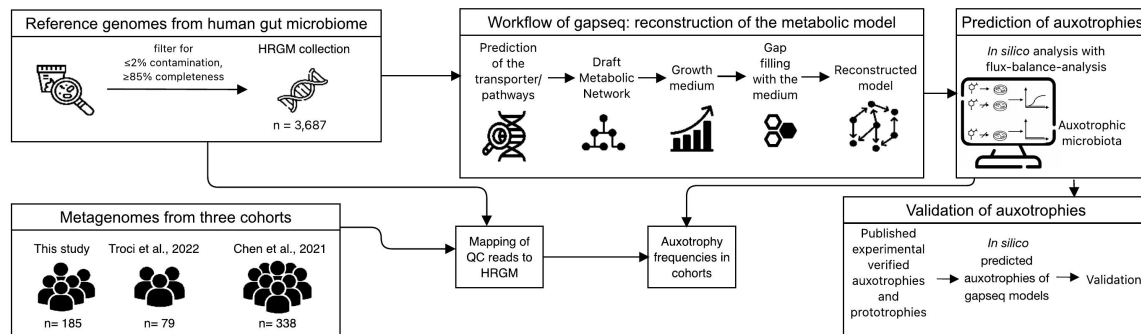
### 3 Results

#### Prediction and validation of auxotrophies with genome-scale metabolic modeling

To estimate the overall distribution of amino acid auxotrophies in the human gut microbiome, we predicted the amino acid production capacities using genome-scale metabolic modeling for all bacterial genomes ( $n=5\,414$ ) from the 'Human Reference Gut Microbiome (HRGM)' collection [146]. Auxotrophies were predicted for the 20 proteinogenic amino acids by comparing the model's growth with and without the amino acid using flux-balance analysis. If the model was not able to grow without the amino acid, then an auxotrophy was predicted (Fig. 5). To exclude an overprediction of auxotrophies due to genome incompleteness, we correlated the genome completeness and the number of auxotrophies predicted. Results showed a negative relationship between genome completeness and the number of auxotrophies per genome (Supplementary Fig. S1,  $\rho=-0.50$ ,  $p\leq 2.2e-16$ ). To combat this, the genomes were filtered for completeness  $\geq 85\%$  and contamination  $\leq 2\%$ . Only the filtered metabolic models ( $n=3\,687$ ) were used to predict auxotrophies and ongoing analysis.



All auxotrophies predicted for HRGM models are in the supplementary material (Supplementary Table S1).



**Figure 5: Workflow for the prediction of auxotrophies with genome-scale metabolic modelling**

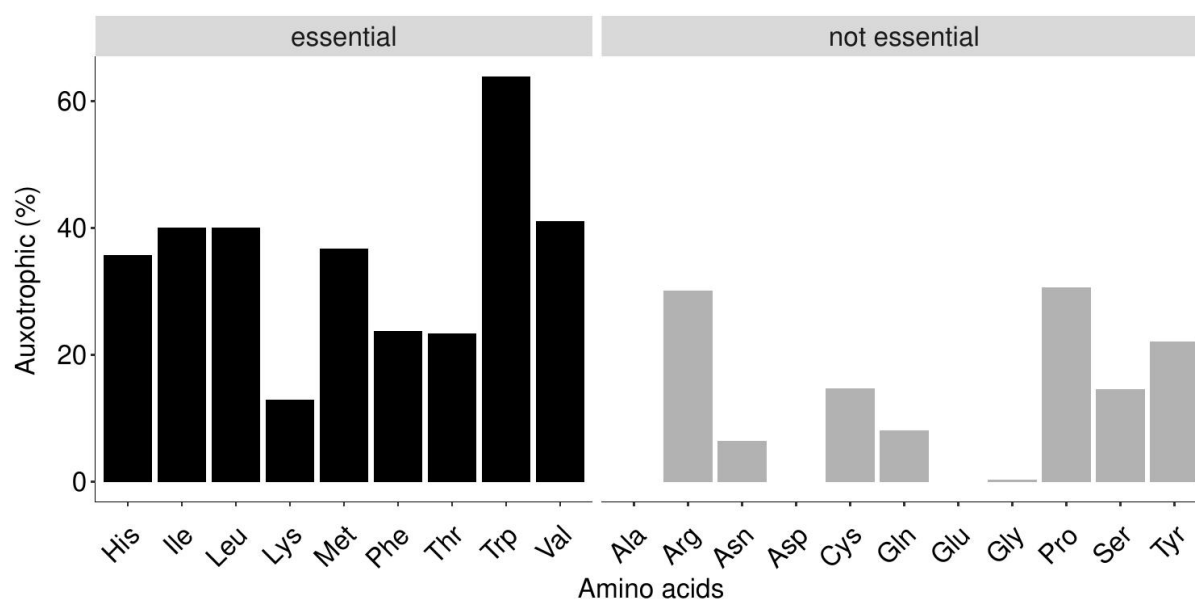
*Gapseq* was used to reconstruct genome-scale metabolic models from genomes of the Human Reference Gut Microbiome (HRGM) catalog[146]. The workflow of *gapseq* to reconstruct metabolic models consists of five steps: transporter/metabolic pathway prediction, draft metabolic network construction, growth medium prediction, gap filling, final model reconstruction. Auxotrophy prediction was performed using flux-balance analysis and validated by reconstructing *gapseq* models from experimentally verified auxotrophic strains. The predicted auxotrophies were compared on strain level from *gapseq* and AGORA2 models to experimentally verified auxotrophies. QC reads of cohorts were mapped on HRGM. Auxotrophy frequencies in cohorts were determined by mapping QC reads from the metagenomes of the cohorts to genomes from HRGM collection. Icons are from [www.flaticon.com](http://www.flaticon.com) (creators: photo3idea\_studio, Freepik, surang, Eucalyp, Voysla, juicy\_fish, smashingstocks, SBTS2018, creative\_designer).

To validate our *gapseq*-based auxotrophy predictions, we compared the predictions on strain level with *in vitro* experimentally verified auxotrophies as reported in previous studies for a total of 36 gut bacteria (Supplementary Table S2), of which most were already summarized in a recent study [79]. If a genome assembly of the experimentally tested strain was available on NCBI RefSeq, we reconstructed the genome-scale metabolic model and predicted the auxotrophies. In addition to auxotrophy predictions using our *gapseq* model collection (Supplementary Table S1), we also tested models from the AGORA2 collection (Supplementary Table S3). Auxotrophy predictions using *gapseq* models had a sensitivity of 75.5%, a specificity of 95.9%, and an accuracy of 93%. The auxotrophies predicted by the AGORA2 models showed a lower degree of

agreement with the experimental data: sensitivity (43.4%), specificity (92.3%), and accuracy (81.7%). In addition, we reconstructed genome-scale metabolic models for 124 bacterial genotypes known to be prototrophic for all 20 proteinogenic amino acids [147] to further validate our auxotrophy predictions (Supplementary Table S4). We note that the 124 prototrophic genotypes are isolates from diverse isolation sources and not from the human gut. However, the resource can be used to estimate the rate of false auxotrophy predictions [147]. In total, 99.1% of all predictions coincided with the known amino acid prototrophies of the organisms, thus suggesting a false auxotrophy prediction rate of less than 1%. 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 Fig. S2), indicating that the high frequency of auxotrophies cannot be explained by a false-positive rate associated with potential pitfalls in the model reconstruction workflow.

#### Amino acid auxotrophies are common in the human gut microbiome

Auxotrophies for tryptophan were predicted to be the most prevalent, at 63.9% of the genomes in the HRGM catalog (Fig. 6). Isoleucine, leucine, and valine (BCAA, branched-chain amino acids) auxotrophies were also estimated to be common (40.1%, 40%, 41.1%, respectively). No auxotrophies were predicted for alanine, aspartate, and glutamate. We further analyzed the observed auxotrophies at the taxonomy level by comparing the proportion and number of auxotrophies on phylum and order level (Supplementary Fig. S3). Actinobacteriota were predicted to have a higher proportion of BCAA auxotrophies compared to prototrophies (Supplementary Fig. S3). For tryptophan, a higher proportion of auxotrophic to prototrophic bacteria was estimated in Firmicutes, Actinobacteriota, and Fusobacteriota. Based on our computational analysis, Fusobacteriota had a higher auxotrophic to prototrophic ratio for almost all amino acids, whereas the opposite was predicted for Proteobacteria. This observation is further supported by the number of auxotrophies predicted per genome for Proteobacteria and Fusobacteriota (Supplementary Fig. S4). Additionally, the results suggest that auxotrophic genotypes have lost the genes for most of the enzymes required for the biosynthesis of the corresponding amino acid (Supplementary Fig. S5). Taken together, the results indicate that amino acid auxotrophies are prevalent in the human gut microbiome.

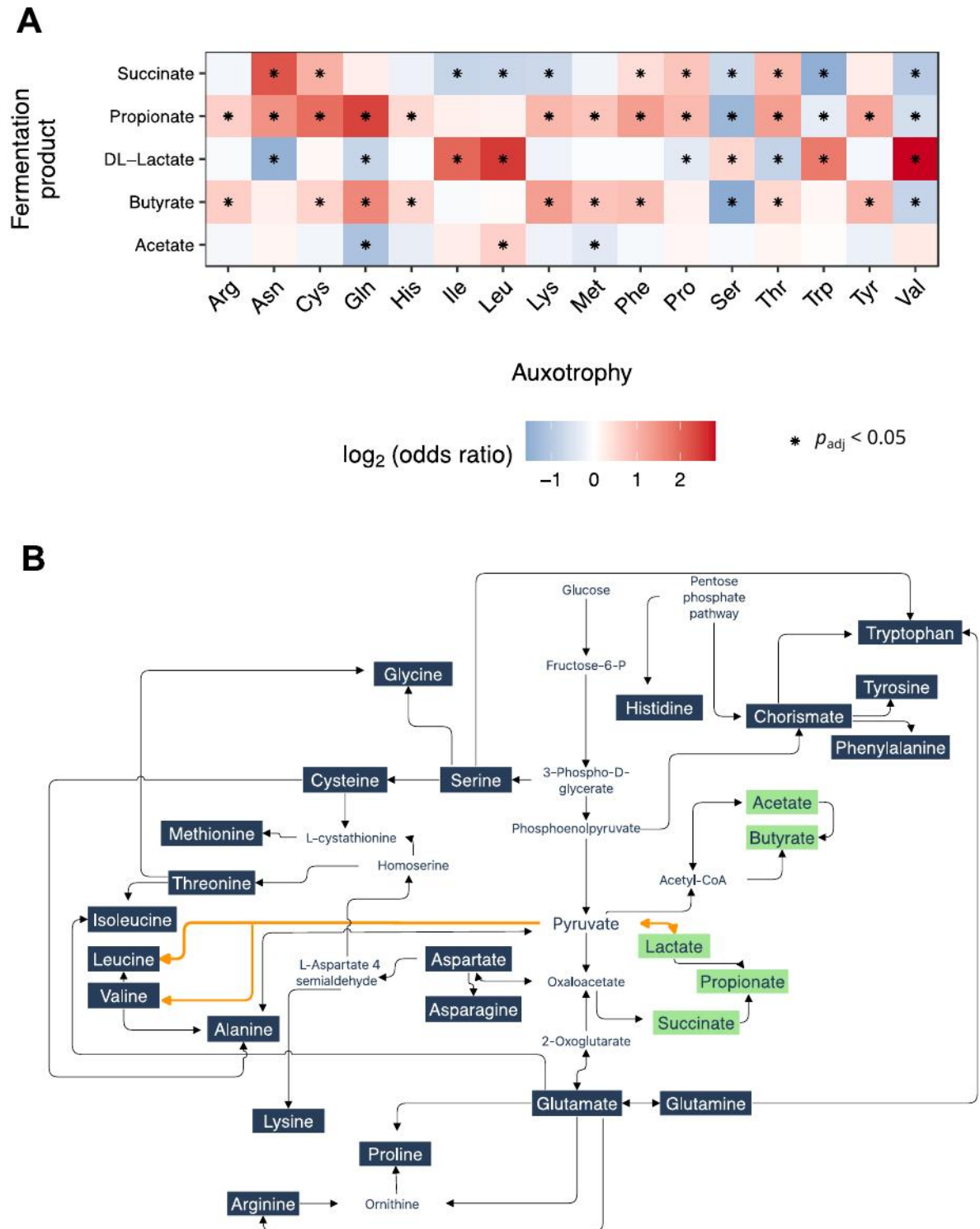


**Figure 6: Abundances of auxotrophies in 3 687 genomes**

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

Amino acid auxotrophies are associated with the profile of fermentation products. Amino acid biosynthesis pathways and pathways producing fermentation products share common precursor metabolites (Fig. 7B). For example, pyruvate is a central metabolite that is utilized for the biosynthesis of the BCAA as well as in some gut bacterial species for lactate formation, underlining the interconnection of amino acids and energy metabolism in the metabolic network.

Here, we investigated whether bacteria that are auxotrophic for specific amino acids are commonly associated with 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 corresponding organisms (Fig. 7A).



**Figure 7: 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. Asterisks denote FDR-corrected  $p$  values < 0.05. (B) Interconnection between the pathways of formation of fermentation products and amino acids,

based on MetaCyc pathways[148]. We note that not all the displayed reactions/pathways occur in every gut bacterial genotype. The metabolic network shown displays pathways commonly found in human gut metagenomes and linked to amino acid biosynthetic pathways.

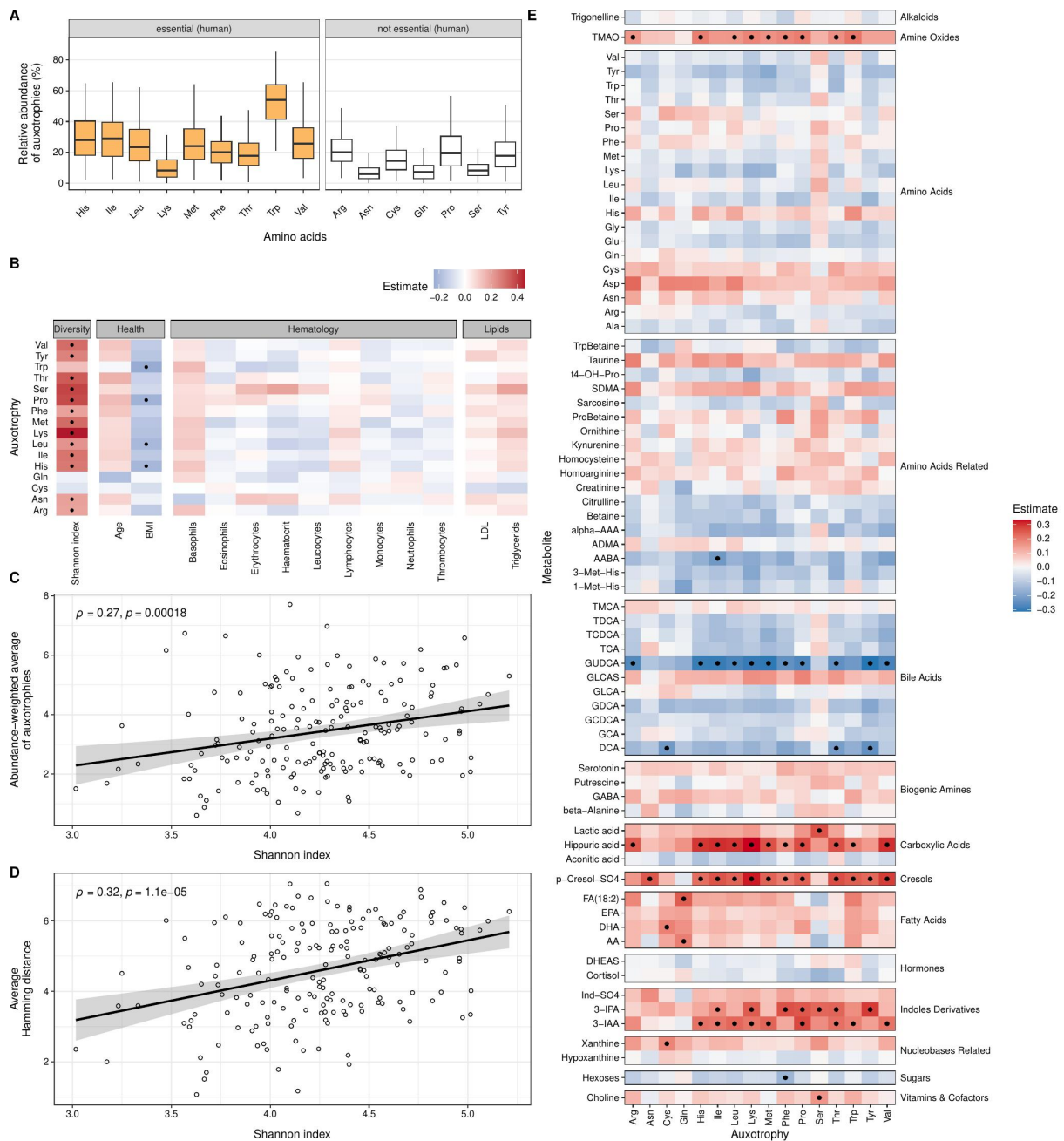
BCAA auxotrophic bacteria were more likely to produce lactate in comparison to prototrophic bacteria (Fisher's exact test for count data,  $\log_2(\text{Odds Ratio (OR)})=2.0-2.8$ , FDR-corrected  $p$  value $<0.05$ ). Propionate production was commonly predicted for glutamine auxotrophic gut bacteria ( $\log_2(\text{OR})=2.4$ , FDR-corrected  $p$  value $<0.05$ ) and by cysteine auxotrophs ( $\log_2(\text{OR})=1.9$ , FDR-corrected  $p$  value $<0.05$ ). Succinate is predominantly produced by asparagine auxotrophic gut bacteria ( $\log_2(\text{OR})=2.2$ , FDR corrected  $p$  value  $<0.05$ ). For butyrate, there was a higher association with glutamine auxotrophic bacteria ( $\log_2(\text{OR})=1.6$ , FDR-corrected  $p$  value  $< 0.05$ ).

The association of auxotrophic bacteria with the production of organic acids might be explained by the distribution of reaction 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. 7B). Pyruvate not used for BCAA biosynthesis in auxotrophic genotypes might be redirected towards lactate production. Thus, our findings suggest a plausible interplay in resource allocation between a microorganism's energy metabolism strategy and its auxotrophy profile.

## More diverse gut microbiomes are characterized by a higher auxotrophy frequency

To estimate the frequency of auxotrophies in the gut microbiome of individual persons, we quantified the relative abundance of gut bacterial genotypes from the HRGM catalog using stool metagenomes of 185 healthy adults. As mentioned above, we found a negative correlation between the number of auxotrophies and genome completeness levels (Supplementary Fig. S1). To validate that higher genome completeness levels do not affect the general pattern in the auxotrophy distribution of individual microbiomes, we determined auxotrophy frequencies with different cutoff values for completeness (80%- 95%) of the reference genomes used for quantification. Overall, the distribution of auxotrophy frequencies remained robust to increasing genome completeness levels (Supplementary Fig. S6). Therefore, we decided to keep the 85% completeness level described above.

Auxotrophies for amino acids that are essential to the human organism were more frequent than non-essential amino acids (Fig. 8A). The highest percentage of bacteria were auxotrophic for tryptophan, followed by isoleucine and histidine (median: 54%, 28.7%, 28%, respectively). Auxotrophies for leucine, methionine, phenylalanine, arginine, and valine were found with a median frequency of >20% (Fig. 8A). The lowest frequencies were detected for serine, lysine, asparagine, aspartate, alanine, and glutamate auxotrophies.



**Figure 8: Distribution of auxotrophies in human gut microbiomes from 185 healthy adults, their association with diversity, and serum metabolite levels.**

(A) Boxplots display 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 index. Dots indicate significant associations (FDR-corrected  $p$  values < 0.05, adjusted for the potential confounders age, sex, and BMI). (C) The abundance-weighted average of auxotrophies was calculated and correlated with the Shannon index (Spearman correlation,  $\rho=0.27$ ,  $p=0.00018$ ). (D) The 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 index in the gut was estimated ( $\rho=0.32$ ,  $p=0.00001$ ). (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 S5. Dots indicate significant associations (FDR-corrected  $p$  values < 0.05, adjusted for confounders age, sex, and BMI).

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. 8B-C). Overall, increasing frequencies of almost all amino acid auxotrophies are accompanied by increasing microbiome diversity (Spearman correlation, Fig. 8B). Further, we correlated the Shannon index with the abundance-weighted average of the number of auxotrophies per metagenome sample, which takes the relative abundance of each genome and its total number of amino acid auxotrophies into account. With an increasing number of auxotrophies, an increase in the diversity was observed (Fig. 8C,  $\rho=0.27$ ,  $p=0.00018$ ). 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 genomes 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 increased gut diversity (Fig. 8D,  $\rho=0.32$ ,  $p=0.00001$ ). Overall, a higher number of auxotrophies in the gut community is positively correlated with a higher diversity.

## 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 [49], [149] but not yet for the frequency of gut microbial amino acid auxotrophies. Our results showed that the summed relative abundance of bacteria that are auxotrophic for certain amino acids are inversely associated with the stool donor's BMI (Fig. 8B, partial Spearman correlation). No statistically significant associations with blood cell counts were found (Fig. 8B). Additionally, we correlated targeted metabolomics data from serum samples with the aggregate frequencies of bacteria auxotrophic for specific amino acids (Fig. 8E, partial Spearman correlation). Positive correlations were found between the tryptophan-derived 3-indoleacetic acid (3-IAA) as well as 3-indolepropionic acid (3-IPA) and tryptophan auxotrophic gut bacteria. Additionally, several other amino acid auxotrophies showed positive correlations with these metabolites. The total frequencies of bacteria auxotrophic for specific amino acids showed a positive correlation with P-cresol sulfate. Further, several significant associations were detected with metabolites from bile acid metabolism. Negative correlations were observed for glyoursodeoxycholic acid (GUDCA), a conjugated secondary bile acid metabolite, and several amino acid auxotrophies. Further, negative correlations with the bile acid metabolite deoxycholic acid (DCA) were found for the aggregate frequencies of bacteria auxotrophic for tyrosine, threonine, and cysteine. Positive associations were also observed for hippuric acid and TMAO with several amino acid auxotrophies. No significant associations were found for serum levels of amino acids and amino acid-related compounds (Fig. 8E).

Taken together, the frequency of auxotrophic bacteria is related to serum levels of several metabolites. The gut microbial contribution to serum metabolite levels was predominantly found for metabolites previously reported to be of microbial origin (e.g., 3-IAA) or 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. Based on this observation, we further wanted to analyze whether the frequency of auxotrophies also impacts the microbiome's long-term stability using data from two longitudinal studies. Therefore, we re-analyzed recently published

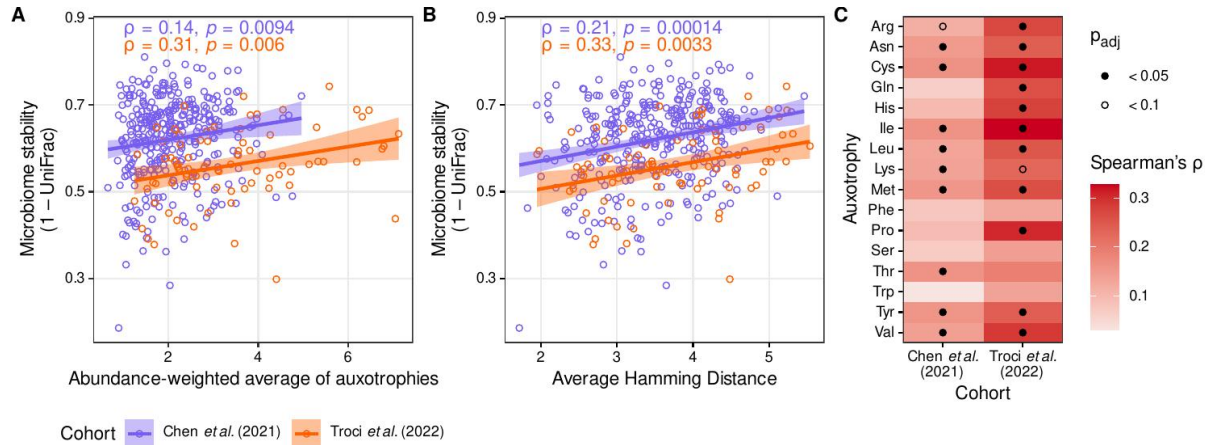


metagenomic data from two human cohort studies [150], [151]. Troci et al. included two stool metagenomes from 79 healthy individuals each where stool samples were three years apart[150]. The longitudinal study of Chen et al. involved two stool metagenomes from 338 individuals with a time difference between samples of four years[135]. Microbiome stability over the time periods was assessed by calculating the UniFrac distance for the microbial composition between the two time points for each participant. Since the UniFrac distance ranges between 0 (lowest possible dissimilarity) and 1 (highest dissimilarity), we calculated the inverse values (1-UniFrac) as a microbiome stability measure. The abundance-weighted average of auxotrophies per genotype was positively correlated with microbiome stability in both cohorts (Fig. 9A, Spearman rank sum correlation test, Troci *et al.*:  $\rho=0.31$ ,  $p=0.006$ ,  $n=79$ ; Chen *et al.*:  $\rho=0.14$ ,  $p=0.0094$ ,  $n=338$ ). We also correlated individual amino acid auxotrophy frequencies with microbiome stability to understand the impact of individual amino acid auxotrophies on long-term stability. A statistically significant positive correlation was found in both cohorts for many amino acid auxotrophies, while no negative correlation was observed (Fig. 9C).

Long-term microbiome stability was 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 a potential indicator for the degree of amino acid cross-feeding in the microbial community. A positive correlation was observed for the average Hamming distance with microbiome stability in both cohorts (Fig. 9B, Troci *et al.*:  $\rho=0.33$ ,  $p=0.0033$ ,  $n=79$ ; Chen *et al.*:  $\rho=0.21$ ,  $p=0.00014$ ,  $n=338$ ), suggesting a potential positive impact of amino acid cross-feeding among auxotrophy genotypes on the long-term stability of microbiome composition.

Auxotrophic bacteria have a high dependence on their nutritional environment. Here, we wanted to test if a higher dietary intake of amino acids affects the relative abundance of amino acid auxotrophic bacteria in the gut. Therefore, we used the dietary intake data obtained from food frequency questionnaires from Troci et al. [150]. For both study time points, the intake of amino acids 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 Fig. S7).

In sum, our findings hint at a potentially positive impact of auxotrophies on the stability of the gut microbiome. Further, the data suggest that amino acid cross-feeding may contribute to the compositional stability of the gut microbiome. We found no evidence of diet's effect on auxotrophy frequencies.



**Figure 9: Influence of auxotrophies on long-term stability of the human gut microbiome**

(A) The stability of the human gut microbiome was calculated as 1 minus the UniFrac distance between the two time points in the longitudinal studies and correlated with the abundance-weighted average of auxotrophies at the first time point to study a potential influence of auxotrophies on the long-term stability of the human gut microbiome. (B) The average Hamming distance was calculated for the first time point and then correlated with the 1-UniFrac value to investigate the influence of potential cross-feeding on long-term stability. (C) The contribution of individual amino acid auxotrophies on the stability was calculated with the Spearman correlation between the 1-UniFrac values and individual amino acid auxotrophy frequencies.

## 4 Discussion

Auxotrophies are widespread among microorganisms [69], [104]. The obligate nutritional requirements can have far-reaching consequences for the auxotrophic strains and the entire microbial community in the ecosystem [152]. Each auxotrophy for a specific essential nutrient (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 [152]. Conversely, if the focal metabolite is available, auxotrophic genotypes might gain a selective advantage over prototrophic genotypes by saving metabolic costs [96]. 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 [90]. In addition, organisms that are auxotrophic for different metabolites could engage in cooperative cross-feeding interactions [98], [153], [154]. Despite the ecological relevance of auxotrophies, their role in the human gut microbiome is largely unknown. More specifically, previous work showed that several human gut bacterial isolates are indeed amino acid auxotrophs using genome analysis and a comprehensive literature review of experimentally determined auxotrophies and prototrophies [79]. Still, the overall distribution and variation of auxotrophies in the human gut microbiome remain elusive. Here, we systematically analyzed the distribution of amino acid auxotrophies in the human gut microbiome using genome-scale metabolic modeling. Moreover, we statistically assessed the associations of inferred auxotrophy frequencies with overall microbiome diversity, long-term stability, and microbial contribution to the human metabolome.

#### *Ubiquity of auxotrophies indicates a high prevalence of cross-feeding*

Overall, high frequencies of auxotrophies were found in the human gut microbiome. For instance, we found that 54%(median) of organisms in the gut microbial communities of healthy adults are auxotrophic for tryptophan. The most frequent auxotrophies for amino acids in the human gut microbiome are also essential nutrients for the human host (Fig. 8A). While auxotrophies in human gut bacteria were reported before, the sources of amino acids for auxotrophic genotypes remain unknown. One potential source of amino acids might be dietary proteins [155]. However, most 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 most of the gut microbiome resides [155]. While most dietary free amino acids do not reach the colon, some dietary proteins that escape digestion in the small intestine can provide a nutrient source for the auxotrophic colonic microbiome [21]. 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 is unlikely to be explained by dietary sources of amino acids alone. Plus, we did not find any statistical associations between the dietary intake of amino acids of 79 adults and the frequency of auxotrophies in the microbiome (Supplementary Fig. S7), which further indicates that auxotrophic genotypes acquire their amino acids from other

sources. Another study supports our conclusion, as varying dietary concentrations of essential nutrients did not alter the frequency of auxotrophy in the gut [73].

A second potential source of amino acids might be cross-feeding interactions between auxotrophic and prototrophic genotypes. Cross-feeding between strains that are auxotrophic for different amino acids has been demonstrated in synthetic [154] and naturally occurring microbial communities [104]. Furthermore, a recent study showed that amino acids synthesized by the colonic microbiome stay in the gut and are not absorbed via the mucosa [21]. Cross-feeding as a potential source of amino acids for auxotrophic bacteria requires that prototrophic bacteria in the microbial community secrete the respective amino acids. In fact, amino acid biosynthesis and the release into their growth environment have been reported for several gut bacterial species, including members of the genus *Bacteroides* [156] and the species *Bifidobacterium longum* [157]. Thus, cross-feeding enables the growth of auxotrophic organisms even in environments where the focal nutrient is unavailable. Our results suggest a wide diversity of auxotrophic profiles between coexisting genotypes (Fig. 8D), indicating metabolic complementarity and amino acid cross-feeding in gut microbial communities.

Host-derived metabolites are the third potential source of amino acids for auxotrophic gut microbes. Yet, evidence reported in the scientific literature for gut microbial uptake of host-derived amino acids is scarce [21], [31]. 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 [94]. *A. muciniphila* is a known degrader of host mucins and resides in the mucus layer. Besides glycans, mucin consists of a core protein scaffold rich in proline, threonine, and serine [158]. Thus, the threonine auxotrophy of *A. muciniphila* may indicate that this species also utilizes host-derived threonine.

#### *Auxotrophies might promote ecological diversity and microbiome stability*

A major result of our study is the positive associations between auxotrophies and diversity of the human gut microbiome. 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 [90], [107]. Thus, we conclude that in communities with more

auxotrophic members, more cross-feeding may take place, which could promote diversity. Our results support this theory since we observed a positive association between microbiome diversity and auxotrophic profile differences among coexisting genotypes.

Microbe-microbe interactions via metabolite exchanges may also promote microbiome stability [57]. 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 to four years. Indeed, our findings from two independent cohorts 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 studies (Fig. 9). 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. This result is in line with a recent theoretical study, which demonstrates that cross-feeding between auxotrophs can facilitate that the community structure returns to equilibrium after ecological perturbation [109]. Moreover, it was previously reported that B-vitamin auxotrophies in the human microbiome are prevalent and suggest that cross-feeding B-vitamins between prototrophic and auxotrophic genotypes contributes to gut bacterial population dynamics [73]. 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 higher auxotrophy frequencies contribute to microbiome stability (Fig 9AB).

Since a reduction in gut microbiome diversity has been reported for several chronic diseases [61], [159], [160], 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 [161].

There is an ongoing debate about how different types of cell interactions (i.e., cooperation and competition) contribute to the stability of multi-species communities [109], [141], [162], [163], [164]. We want to emphasize that we do not claim that cooperative interactions are stronger than competitive interactions in stabilizing microbiomes, also because we focused in this study on one type of interaction (amino acid cross-feeding) and not on the prevalence of other kinds of interactions or the exchange of other metabolites. Instead, we argue that our results provide evidence that auxotrophies and potential amino acid cross-feeding contribute to maintaining microbiome composition.

#### *Auxotrophy associations with the human metabolome*

Pathways of amino acid biosynthesis and fermentation by-product biosynthesis share common precursors. Therefore, the loss of biosynthetic genes for amino acids might affect the flux distribution in the metabolic network [96]. 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. 7A), which may be due to the structure of the metabolic network. 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. 8E). In particular, the frequencies of several auxotrophies were correlated with phenolic and indolic metabolites, namely hippuric acid, p-cresol 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 [33]. 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 [165]. P-cresol is known to be produced by gut bacteria that metabolize tyrosine [166], and we found an association with tyrosine auxotrophic gut bacteria. Moreover, the tryptophan-derived IAA is a known agonist of the epithelial human aryl hydrocarbon receptor, an important regulator of intestinal immunity [167]. In summary, our results suggest that the contribution of phenyllic 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 *in vitro* assessments have been reported and discussed previously [72, S.], [114], [147]. 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 36 gut bacterial strains and found a sensitivity of 75% for auxotrophy predictions with *gapseq*-reconstructed genome-scale metabolic models. In addition, we performed auxotrophy prediction for 124 genomes from bacterial 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 99.1% of our prototrophy predictions are in line with the experimental data, suggesting that the high prevalence of predicted auxotrophies among the human gut bacterial genomes is not due to a potential technical bias in the *in silico* approach.

## 5 Conclusion

Our study demonstrates the prevalence and potential impact of auxotrophs in the human gut microbiome. Auxotrophies are common in the human gut microbiome, and interestingly, amino acids essential to the human host are also commonly essential for large fractions of the gut microbiome. Furthermore, human gut microbiomes with high frequencies of auxotrophies were characterized by higher alpha diversity and were more stable over time. Since gut microbial communities commonly display reduced diversity during chronic diseases, auxotrophy frequencies in the human gut microbiome could indicate a healthy gut microbiome. In addition, our results suggest that metabolite cross-feeding networks in gut bacterial communities may be an

important factor for stability and maintaining diversity. From a more technical point of view, previous studies have suggested a cautious interpretation of *in silico*-predicted auxotrophies. Therefore, we validated our *in silico* results with experimentally determined auxotrophies reported in scientific literature. This validation indicated the high predictive performance of our method, which used automatic genome-scale metabolic network reconstruction without the need for manual curation of individual genotypes. Thus, the approach can also be applied to microbial communities other than the human gut microbiome.

## 6 Material and Methods

### Reconstruction of genome-scale metabolic models

Genome-scale metabolic models were reconstructed for bacterial genomes from the Human Reference Gut Microbiome (HRGM) genome collection [146], [168]. 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 [111]. *gapseq* is a command-line tool for automated reconstructions of genome-scale metabolic networks, that utilizes pathway, reaction, and metabolite entries from the databases *MetaCyc* [148] and *ModelSEED* [125]. A detailed description of the genome-scale metabolic model reconstruction workflow can be found in the Supplementary Information and Supplementary Table S6.

### Prediction of amino acid auxotrophies

Amino acid auxotrophies were predicted with flux balance analysis [120], 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 predicted with *gapseq* (see Supplementary Information), and second, with the same medium but where the amino acid of interest was removed. An organism was defined as auxotrophic for a specific amino acid if the organism was able to form biomass in the original medium but not in the medium without the amino acid of interest. Flux balance analysis was performed in R (v4.1.2), the R package *sybil* v2.2.0 [170], and IBM ILOG CPLEX optimizer as linear programming solver. We validated our auxotrophy predictions for 150 organisms (36



from the human gut, 124 known prototrophs from different environments), for which experimental data for amino acid auxotrophies and prototrophies were available in scientific literature (see Supplementary Information for details).

When assessing the distribution of amino acid auxotrophies in sampled individual microbiomes, it is important to consider the relative abundance of different genotypes. To this end, we combined the estimated relative abundances of reference genomes (see ‘Metagenome data processing’) and predicted auxotrophies in the corresponding genomes to calculate the relative auxotrophy abundance  $y_{j,k}$  of amino acid  $k$  in sample  $j$  using the equation:

$$y_{j,k} = \sum_{i \in M} p_{i,j} b_{i,k}$$

Where  $M$  is the set of all reference genomes,  $p_{i,j}$  the relative abundance of genome  $i$  in sample  $j$ , and  $b_{i,k}$  the auxotrophy prediction with “1” if genotype  $i$  is auxotrophic for amino acid  $k$  and “0” otherwise.

### Prediction of metabolic by-products

For comparison of auxotrophic to prototrophic bacteria, the production rates of fermentation by-product formation were predicted. We undertook this analysis based on the demonstrated accuracy of gapseq in predicting fermentation products of anaerobically cultured gut bacteria [111]. Given the potential correlation between auxotrophies and the generation of metabolic by-products, investigating auxotrophy distributions could offer new insights into gut microbial metabolism and ecology. Metabolic by-products were predicted with flux-balance-analysis [120] using the flux through the biomass reaction as objective function (i.e., maximization) and subsequently analyzing the fluxes through exchange reactions. Metabolite production rates ( $\text{mmol} \cdot \text{gDW}^{-1} \cdot \text{hr}^{-1}$ ) were normalized by growth rates ( $\text{hr}^{-1}$ ), resulting in the unit  $\text{mmol/gDW}$ . Production rates  $> 1 \text{ mmol/gDW}$  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.

### Cohorts

Data from three human population cohorts were analyzed for the present study. The first cohort comprised paired stool metagenomes and serum metabolomes from 185 participants (Table 1). This cohort was recruited at the University Hospital Schleswig

Holstein, Campus Kiel 2016, and included detailed phenotypic and health-related data. The study was approved by the local ethics committee in Kiel (D441). None of the participants had received antibiotics or other medication two months before inclusion. The second cohort (Troci *et al.*, 2022) comprised longitudinal stool metagenomes from 79 study participants. Data from this cohort were already part of a previous study [150], which were reanalyzed in the present study. For each participant from this cohort, two metagenomes were sequenced from stool samples that were three years apart. In addition, for each sampling time point, data from food frequency questionnaires were available. In brief, the questionnaire, originally designed and validated for use in the German EPIC study [171], comprised 112 food items and aimed to collect the intake frequency and amount of various types of foods. The average energy intake and other nutrients per day were calculated with data from the German Food Code and Nutrient Data Base (BLS version II.3 (66)). Further information about the sampling method, study design, and sequencing method of this cohort study can be found in the original publication [150].

The third cohort integrates fecal metagenomes from the 2021 publication by Chen *et al.*, involving 338 Dutch study participants [151]. Like the second cohort, the Chen *et al.* cohort is designed longitudinally, incorporating two fecal metagenomic samples per participant over a four-year interval.

**Table 1: Cohort characteristics**

Age and BMI are given as the median and interquartile range.

Cohort	Age (years)	BMI	Female (%)	Study participants
This study	47 [40-52]	24.5 [22.2-26.4]	44.9	185
Troci <i>et al.</i> , 2022	53 [45.75-57.25]*	25.7 [23.5 - 27.5]*	37.5*	79
Chen <i>et al.</i> , 2021	47.5 [40-56]	—	55.6	338

\* Missing values/information: 7.

## 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 described earlier [172]. DNA extracts were used for metagenomic library preparation as described previously [150] 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 quality filtered using the 'qc' workflow from the metagenome-atlas pipeline tool v2.9.0 [173] with default parametrization if not stated otherwise in the Supplementary Information. Quality-controlled (QC) reads were used to estimate the relative abundance of genomes from the HRGM catalog [146] using coverM v0.6.1 [174]. Across all three analyzed metagenome data sets, a median of 76% QC reads mapped to HRGM reference genomes (Supplementary Fig. S8).

## 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 blood sample preparation and metabolite quantification details.

## Statistical data analysis

All data analysis steps and statistical tests were performed using R (v4.1.2). Flow charts (Fig. 1. and 3A) were created and rendered using Flowchart Designer 3. *p* values were corrected for multiple testing using the Benjamini and Hochberg method [175]. In all statistical tests, an adjusted *p* value of <0.05 was considered as significance threshold. UniFrac distances [176] were calculated using relative abundances of genomes using the R-package abdiv, v0.2.0 [177].

Alpha diversity was calculated using the Shannon index as implemented in the R-package 'vegan' v2.6-2 [178]. The average pairwise Hamming distance between auxotrophic profiles of co-occurring genomes was calculated per sample to study the effect of metabolic dissimilarity on diversity. 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 genomes,  $a_i$  the number of auxotrophies in genome  $i$ , and  $p_{ij}$  the relative abundance of genome  $i$  in sample  $j$ .

For the longitudinal cohorts, the UniFrac 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 to determine the association between the UniFrac distance and the Hamming distance. With 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. The Spearman correlation was used to study an association between the total dietary intake of amino acids relative to the total consumed energy (E%) and the frequency of amino acid auxotrophic bacteria. The correlation between the intake of amino acids and frequencies of amino acid auxotrophic bacteria was studied separately for both time points.

## Data availability

The reconstructed genome-scale metabolic models from the HRGM catalog are available via *Zenodo* [179]. Further, metabolic model reconstructions for 124 prototrophic genotypes and 36 gut bacterial genotypes with amino acid auxotrophy/prototrophy status known from laboratory experiments are available via *Zenodo* [180], [181]. Metagenome sequencing data are provided via the European Nucleotide Archive 'ENA' for our study cohort and the cohort from Troci *et al.* (this study accession: PRJEB60573, Troci *et al.*: PRJEB48605). Metagenome sequencing data from Chen *et al.* 2021 [151] are available upon request via the European Genome-Phenome Archive (accession: EGAD00001006959).

## Code availability

The code for analysis of the data can be found in the GitHub repositories [https://github.com/SvBusche/Auxo\\_manuscript\\_2023](https://github.com/SvBusche/Auxo_manuscript_2023) (main results) and

[https://github.com/Waschina/AGORA2\\_auxotrophies](https://github.com/Waschina/AGORA2_auxotrophies) (for auxotrophy predictions from AGORA2 metabolic models).

### Competing interests

The authors declare 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., P.R.), 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., P.R.), 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).



# Chapter III: Comparative Analysis of Amino Acid Auxotrophies and Peptidase Profiles in Non-Dysbiotic and Dysbiotic Small Intestinal Microbiomes

*In revision for the Computational and Structural Biotechnology Journal*

Svenja Starke<sup>1,\*</sup>, Danielle MM Harris<sup>1,2</sup>, Amandine Paulay<sup>3,4,5</sup>, Konrad Aden<sup>2,6</sup>, Silvio Waschina<sup>1,\*</sup>

<sup>1</sup> Institute of Human Nutrition and Food Science, Kiel University, Department of Nutriinformatics, Kiel, Germany

<sup>2</sup> Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany

<sup>3</sup> Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, 78350, Jouy-en-Josas, France

<sup>4</sup> Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France

<sup>5</sup> Biomathematica, 20000 Ajaccio, France

<sup>6</sup> Department of Internal Medicine I, Kiel University and University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany

## 1 Abstract

Small Intestinal Bacterial Overgrowth (SIBO) is linked to various gut diseases and has been associated with altered serum amino acid levels, potentially influencing mood disorders. However, the direct role of the gut microbiome in these changes remains unconfirmed. In this study, we predict amino acid auxotrophy and peptidase gene profiles in the small intestinal microbiomes of SIBO and non-SIBO subjects. Further auxotrophy and peptidase gene profiles were analyzed in the large intestinal microbiome in non-dysbiotic states to compare the resemblance of the microbial SIBO profile.

Our results revealed that auxotrophy abundances, particularly for essential amino acids, are higher in the microbiota of the small intestine than in the large intestine in non-dysbiotic states. In SIBO subjects, the abundance of auxotrophies decreased. Peptidase gene profiles were distinct between small and large intestinal microbiomes, with fewer extracellular peptidase genes in the small intestine microbiomes. However, in SIBO, extracellular peptidase genes increased. Further, more statistically significant associations were found between the abundance of auxotrophies and peptidase genes in microbiomes of the small intestine than in the large intestine.

In conclusion, the auxotrophy and peptidase profiles of the small and large intestinal microbiomes are distinct. In SIBO, the small intestinal microbiome shifts towards a metabolic state resembling that of the large intestine, particularly in its increased abundance of extracellular peptidases.

## 2 Introduction

The microbiota colonizes the entire intestinal tract, but its composition and metabolism are distinct between the small and large intestines [15]. A primary factor contributing to these differences is the greater availability of simple nutrients, including mono- and disaccharides, free amino acids, and dietary proteins and peptides, for microorganisms in the small intestine compared to the large intestine [182]. Many studies have demonstrated the involvement of gut microbial metabolism in human health [59][183], but the focus of research has tended to be on the large intestinal microbiome. The small intestinal microbiome is often referred to as the “*forgotten microbiome*” due to its role in gastrointestinal disorders [184], despite its well-documented role in the metabolic transformations of dietary molecules. Studies in recent years have shown a



link between the small intestinal microbiota and human health, demonstrating the importance of the small intestinal microbiota [50][185].

Digestion of dietary nutrients results in an altered nutritional environment in the small intestine compared to the large intestine. In the small intestine, dietary proteins are hydrolyzed into smaller peptides and amino acids, which are mainly absorbed via mucosa [186]. Given that a substantial amount of dietary proteins are degraded in the small intestine, the microbiota residing in the lower intestinal tract has limited access to proteins, peptides, and amino acids except for some proteins that escape digestion [187][188][21]. In contrast, proteins and peptides of both the human host and bacterial origin were shown to be critical sources of amino acids in the lower gastrointestinal tract [31], [94][188]. Moreover, bacteria in the large intestine were shown to secrete more proteins than in the small intestine [23].

Dietary proteins and peptides are hydrolyzed not only by human proteases but also by proteases and peptidases produced by the human gut microbiome [27]. Proteolysis releases amino acids and peptides, which are essential nutrients for many gut-dwelling microorganisms [20]. The decreasing availability of free amino acids along the gastrointestinal tract highlights the importance of proteolytic activity for protein degradation [21], [26]. In microorganisms, two phenotypes regarding a specific amino acid requirement can be identified: auxotrophic and prototrophic. Auxotrophic organisms cannot independently synthesize growth-essential nutrients, such as amino acids, and must obtain them from their environment. In contrast, prototrophic bacteria can synthesize these metabolites *de novo* [20]. Theoretical and *in-vitro* studies have highlighted the importance of auxotrophies for complex microbial communities through contributing to community composition [107] and resilience to resource fluctuations [109], [189].

Since determining amino acid auxotrophies experimentally is laborious and limited to microorganisms that can be cultivated in isolation under laboratory conditions [72], computational tools were developed that aim to predict auxotrophies from genome sequences [82], [114], [190], [191]. The application of these computational auxotrophy prediction tools has indicated that amino acid auxotrophies are prevalent among human gut bacteria [68], [190] and are positively associated with gut microbiome diversity and long-term stability [190]. However, these studies primarily focused on the large intestinal microbiome. The auxotrophy profile of the small intestinal microbiome might differ due to the higher availability of amino acids compared to the colon

ecosystem. Moreover, microbial proteolytic activity might play an important part in releasing free amino acids for auxotrophic genotypes.

Many diseases are characterized by altered amino acid metabolism of the small intestinal microbiota, for example, hyperglycemia and obesity [48], [192]. In liver cirrhosis patients, next to an alteration of the amino acid metabolism, there was also an enrichment in peptidases seen in the duodenal microbiota [193]. Microbial proteolytic activity may impact host-microbiome interactions in certain diseases. For instance, microbial peptidases involved in gluten metabolism were shown to be reduced in the small intestinal microbiota of celiac disease patients [194]. Moreover, Small Intestinal Bacterial Overgrowth (SIBO) is a dysbiosis of the small intestinal microbiota characterized by an abnormal density of bacteria  $\geq 10^5$  CFU/ml [195]. SIBO is associated with several diseases, including nonalcoholic fatty liver disease [196] and obesity [197], and is often accompanied by protein malabsorption [195], [198]. A recent study showed that levels of tryptophan and its derived metabolites are altered in urine and reflected by mental disorders in patients with SIBO [45]. Symptoms were reduced after antibiotic treatment targeting the small intestinal microbiota, suggesting that bacteria in the small intestine contribute to symptoms associated with SIBO. However, whether there is an imbalance in the small intestinal microbiota's protein and amino acid metabolism remains unclear in patients with SIBO.

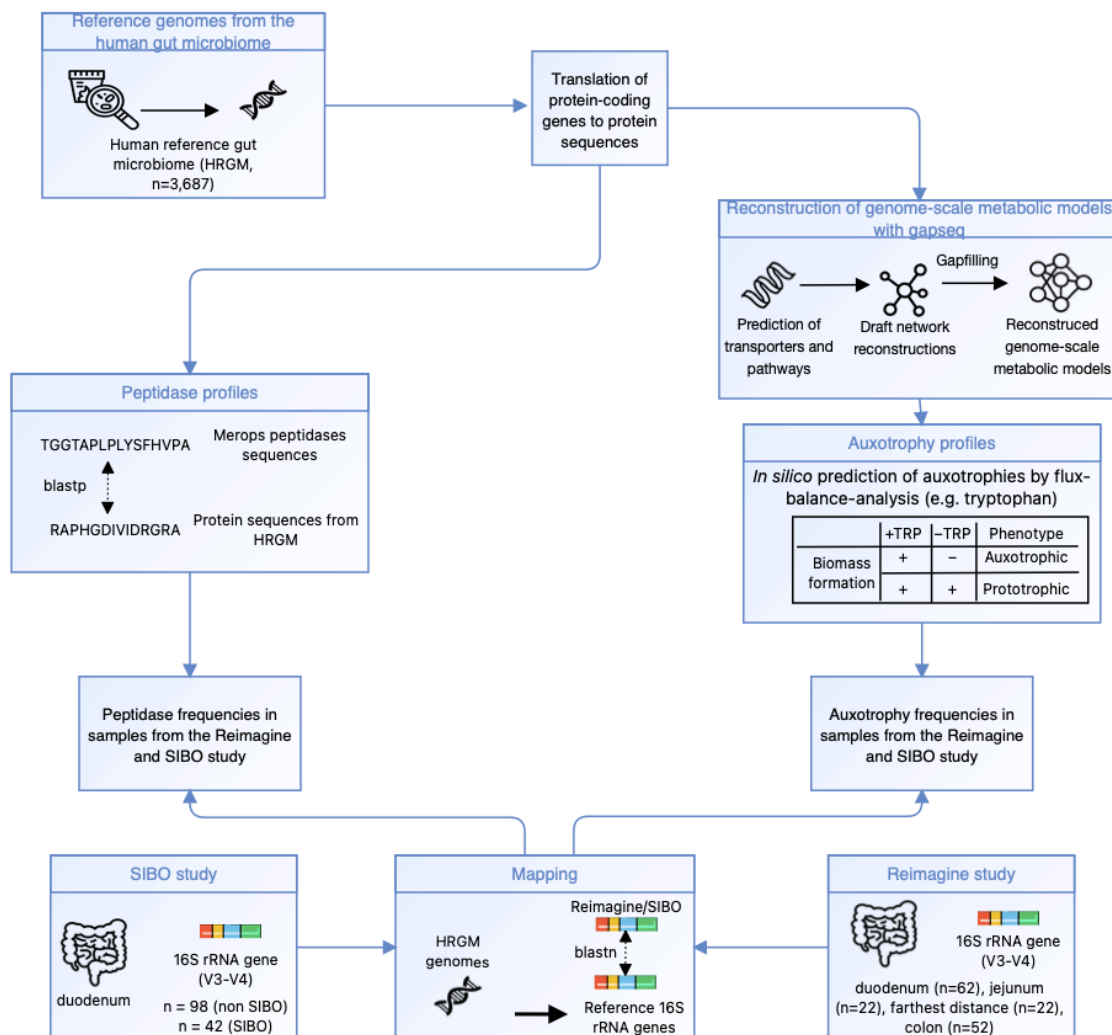
This study aimed to investigate the potential influence of varying nutritional conditions along the gastrointestinal tract on the distribution of auxotrophies by analyzing publicly available microbiome data obtained from the small intestine regions (i.e., duodenum, jejunum, and ileum) and the large intestine in a non-dysbiotic state. Auxotrophies were predicted with genome-scale metabolic modeling (Fig. 10). In addition, we compared the predicted peptidase gene profiles in microbiomes of the small intestine regions and the large intestine and assessed the potential relationship between the abundances of auxotrophies and peptidases. Furthermore, we investigated the peptidase and auxotrophy profiles of the microbiome in SIBO.

### 3 Results

#### Cohorts and microbial genome quantification

For this study, publicly available microbiome sequencing data of two cohorts ("Reimagine study" [7] and "SIBO study" [199]) were re-analyzed. The "Reimagine study" collected microbiome samples from the duodenum, jejunum, farthest distance,

and stool. By “farthest distance” the distance, that could be reached during sampling, is meant. The “SIBO study” took samples from the duodenum of patients with SIBO and a non SIBO control group [199]. Both studies were conducted in the United States of America at the Cedars-Sinai Medical Center. In both studies, 16S rRNA gene sequencing was used for microbiome profiling. For our re-analysis, Amplicon sequence variants (ASVs) were inferred from the paired-end 16S rRNA gene sequencing data in FASTQ format. Representative sequences of ASVs were mapped to reference genomes of the Human Reference Gut Microbiome (HRGM) catalog [146] to predict auxotrophies and peptidase genes. The amino acid auxotrophy profiles were predicted by genome-scale metabolic modeling, as described recently[190]. The inferred relative abundances of individual genomes from the HRGM catalogue per sample can be found in the Supplementary Table S7 and S8.



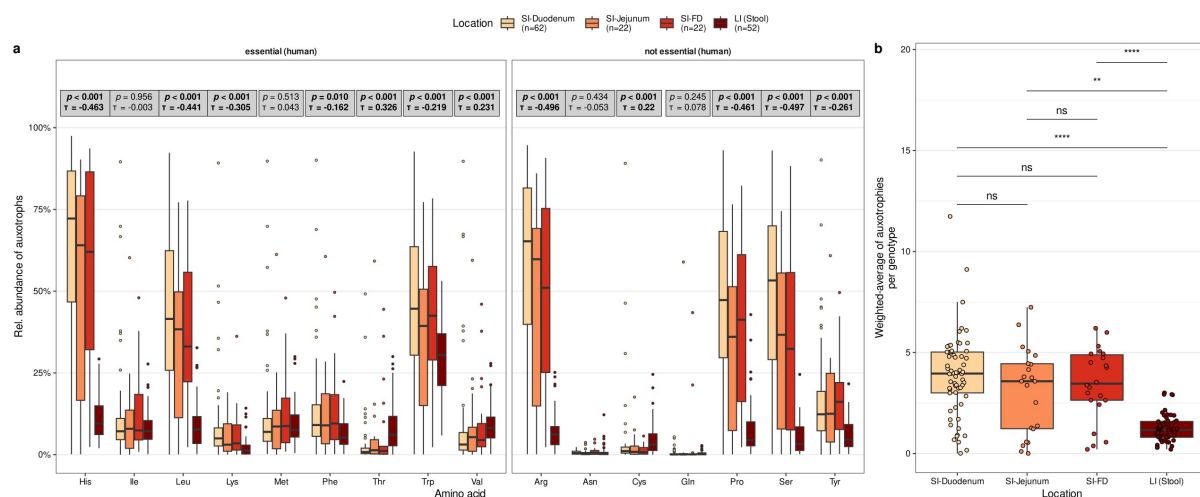
**Figure 10: Flowchart of the study**

Genome-scale metabolic models were reconstructed from genomes of the human reference gut microbiome [146] (HRGM) catalog with gapseq [111]. Auxotrophies were predicted with metabolic modeling using flux-balance analysis. The peptidase profiles were determined by scanning the peptidase sequences of the HRGM catalog for peptidases from the MEROPS database [200]. The distribution of peptidases and auxotrophies was predicted in the SIBO and Reimagine study cohorts after mapping on HRGM genomes. The farthest distance refers to the distance that could be reached during the sampling. Free available icons were taken from [www.flaticon.com](http://www.flaticon.com) (creators: photo3idea\_studio, surang, Icon home Eucalyp, Kiranshastry, Becris, Nadiinko)

### Auxotrophies are more abundant in the small intestinal microbiome than in the large intestinal microbiome

Auxotrophy abundances were compared between gastrointestinal locations: duodenum, jejunum, farthest distance, and large intestine (Fig. 11). Overall, the results show that microbiomes in the small intestine have higher auxotrophy abundances than microbiomes in the large intestine (Fig. 11b). The abundance-weighted average of the number of auxotrophies per genotype is higher in the microbiomes of all three small intestinal locations (duodenum, jejunum, farthest distance) compared to large intestinal microbiomes (Wilcoxon rank sum test, FDR-corrected  $p$  values  $< 0.05$ ).

9 out of 16 tested amino acid auxotrophies showed significantly decreasing relative abundances along the intestinal tract with higher levels in the small intestine and lower relative abundance in the large intestine (Kendall's rank correlation test with intestinal locations treated as ordered factors,  $\tau < 0$ ,  $p_{\text{adj}} < 0.05$ , Fig. 11a). These amino acids included the human essential amino acids His, Leu, Lys, Phe, and Trp and the non-essential amino acids Arg, Pro, Ser, and Tyr. In contrast, three amino acid auxotrophies displayed increasing relative abundance along the intestinal tract with the highest levels in the large intestine: Thr, Val, and Cys (Kendall's rank correlation test,  $\tau > 0$ ,  $p_{\text{adj}} < 0.05$ , Fig. 11a). No statistically significant differences were observed for the abundance-weighted average of auxotrophies between the microbiomes of the different small intestinal regions (duodenum, jejunum, farthest distance, Fig. 11b).



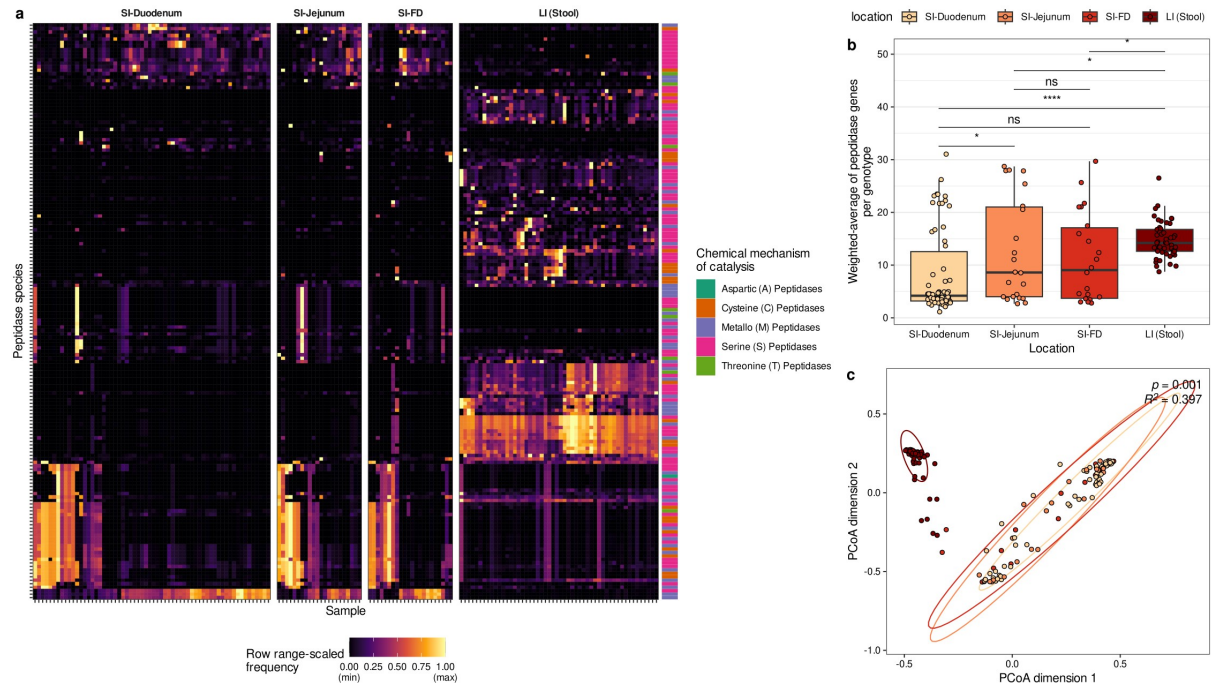
**Figure 11: Comparison of auxotrophy profiles in small and large intestinal microbiomes**

(a) Comparison of individual microbial amino acid auxotrophy abundances in the SI (Small intestine)-duodenum, SI-jejunum, SI-FD (farthest distance reached in the small intestine), and large intestine (LI). Grey boxes report  $p$  values and Kendall's  $\tau$  (tau) from Kendall rank sum correlation tests between the relative abundance of auxotrophies (y-axis) and the intestinal regions (color-coded), which is treated as an ordered factor. Reported  $p$  values are FDR-adjusted. (b) Abundance-weighted average of auxotrophies per genotype in duodenal, jejunal, farthest distance, and colonic microbiomes. Horizontal bars denote pairwise tests for statistical differences using Wilcoxon rank sum tests. Labels above bars indicate significance levels: ns – not significant ( $p$  value  $> 0.05$ ), \*\* ( $p < 0.01$ ), \*\*\*\* ( $p < 0.0001$ ).

## Distinct peptidase gene profiles of the small and large intestinal microbiomes

Next, we compared the distribution of genes encoding putatively extracellular peptidases within the microbiome inhabiting various gastrointestinal regions (duodenum, jejunum, farthest distance, large intestine; Fig. 12, Supplementary Table S9, S10). Overall, the distribution of peptidases was clearly distinct between the large and small intestinal microbiomes (Fig. 12a). The differences between gene profiles of extracellular peptidases were also evident in a Principal Coordinate Analysis (PCoA) using the Bray-Curtis Distance as metric to quantify distances between gene profiles (PERMANOVA,  $R^2 = 0.397$ ,  $p = 0.001$ , Fig. 12c). Moreover, the genotype abundance-weighted averages of the number of genes encoding extracellular peptidases indicated that these peptidase genes are less common to microbiomes from each small intestine region compared to the large intestine microbiome (Wilcoxon rank sum test,  $p < 0.05$ ; Fig. 12b). When comparing extracellular peptidase gene profiles between microbiomes of the small intestine, we observed that the microbiomes in the jejunum harbor more peptidase genes than in the duodenum ( $p = 0.02985$ , Fig. 12b). In sum, the results

indicate that the peptidase gene profile of the small intestine microbiome is distinct from that of the large intestine, and the large intestinal microbiomes harbor more genes encoding putatively extracellular peptidases.



**Figure 12: Comparison of peptidase profiles in small and large intestinal microbiomes**

(a) Heatmap of individual peptidase abundances (range-scaled) in the microbiomes of the small intestine (SI)-duodenum, SI-jejunum, SI-farthest distance (SI-FD), and large intestine (LI) microbiomes. (b) Abundance-weighted average of peptidases. Horizontal bars denote pairwise tests for statistical differences using Wilcoxon rank sum tests. Labels above bars indicate significance levels: ns – not significant ( $p$  value  $> 0.05$ ), \* ( $p < 0.05$ ), \*\*\*\* ( $p < 0.0001$ ). (c) PCoA plot of peptidase profiles of the microbiomes of the four different intestinal regions (PERMANOVA based on Bray-Curtis Distances of pairwise peptidase profiles,  $R^2 = 0.397$ ,  $p = 0.001$ ).

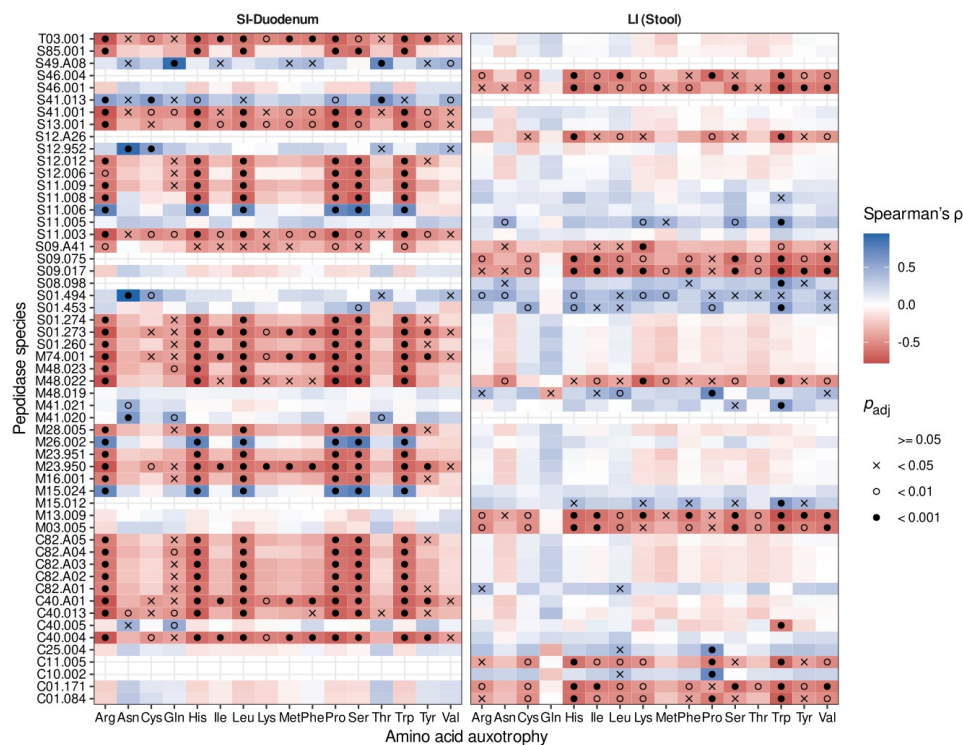
Abundances of auxotrophies are often negatively associated with abundances of peptidases in the microbiomes of the small and large intestine

Hydrolysis of extracellular accessible peptides could be a mechanism for auxotrophic bacteria to cover their demand for amino acids. Therefore, we tested for the potential association between the relative abundance of individual amino acid auxotrophies and the relative abundance of bacteria that harbor genes for specific putatively extracellular peptidases (Fig. 13). This correlation analysis was performed separately for large intestine microbiome samples and duodenal microbiome samples because the



duodenum was the small intestinal region with the largest sample size. More statistically significant correlations between auxotrophies and peptidase genes were found in the duodenum than in the large intestine (Spearman's rank correlation, Fig. 13). Additionally, most correlations were negative between abundances of auxotrophies and peptidase genes. Auxotrophies for arginine, histidine, leucine, proline, serine, and tryptophan had the most statistically significant correlations with peptidase genes in the duodenum. In the duodenum, several auxotrophies were negatively associated with peptidases involved in antibiotic resistance (S11, S12, S13 peptidase families), cell wall synthesis (S11, S12, C82, C40 peptidase families), and peptidoglycan synthesis (C82 peptidase families). Further, auxotrophies were negatively correlated with peptidase genes involved in glutathione degradation (T03.001), dietary protein degradation (S01.260), insulin and beta-amyloid degradation (M16.001), and the lysis of cell walls (M23.950, M23.951). In large intestinal microbiomes, peptidases involved in the degradation of oligopeptides degradation (S09.075) and anti-microbial peptides (C01.171) were negatively associated with several amino acid auxotrophies (Fig. 13).

In conclusion, auxotrophies are mostly negatively correlated with peptidase genes in small and large intestinal microbiomes.



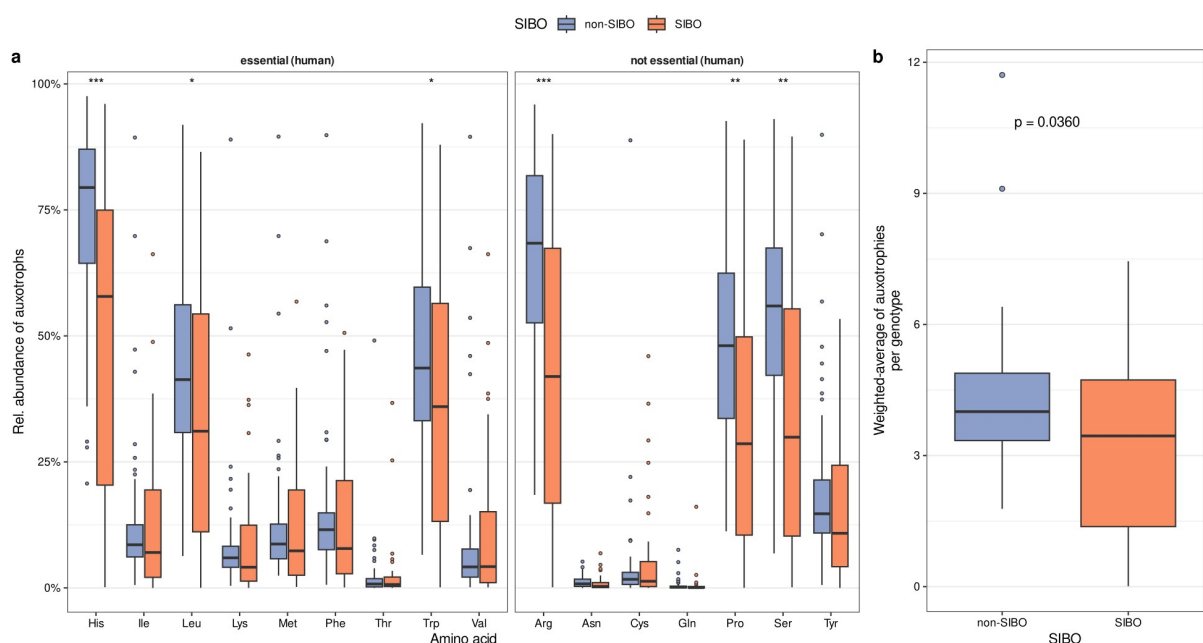
**Figure 13: Associations between peptidases and auxotrophies in duodenal (SI-Duodenum) and large intestinal (LI-stool) microbiomes**

Spearman rank sum correlations were calculated between the relative abundance of the respective amino acid auxotrophy (x-axis) and the relative abundance of specific peptidases (y-axis). p values were FDR-adjusted.

Auxotrophies are reduced, but extracellular peptidase gene abundances are increased in the small intestinal microbiomes of SIBO patients

SIBO is a dysbiotic state of the small intestine and is associated with several gastrointestinal disorders. Serum and urine levels of tryptophan and its derivatives were altered in SIBO patients, but direct involvement of the gut microbial amino acid metabolism was not studied. Therefore, we reanalyzed duodenal samples from a SIBO study[199] and compared auxotrophy and peptidase gene profiles in the duodenum of SIBO and non-SIBO subjects.

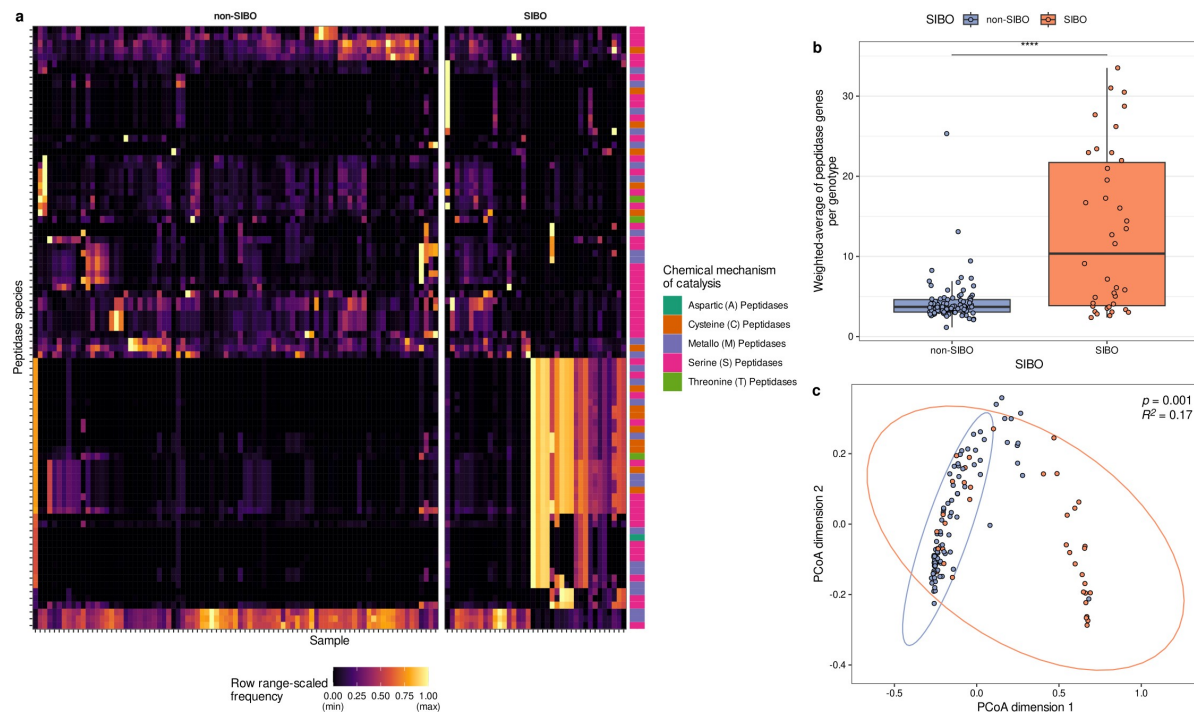
So far, we have observed that the auxotrophy abundances in the non-dysbiotic duodenal microbiome are higher than in the large intestinal microbiome (Fig. 11a). In the dysbiotic state of the duodenal microbiome of SIBO subjects, we observed that auxotrophy abundances are decreased compared to non-SIBO subjects. The abundance-weighted average of auxotrophies was decreased in the microbiome of SIBO patients (Fig. 14b, Wilcoxon rank sum test,  $p = 0.0360$ ). Auxotrophy abundances of the essential amino acid histidine, leucine, and tryptophan, as well as for the non-essential amino acids proline, arginine, and serine, were lower in the microbiomes of SIBO subjects compared to non-SIBO subjects (Fig. 14a, Wilcoxon rank sum test, FDR corrected p value  $< 0.05$ ).





**Figure 14: Auxotrophy profiles in the duodenal microbiomes of non-SIBO and SIBO samples.** (a) Individual amino acid auxotrophy abundances in the duodenal microbiome of SIBO and non-SIBO patients. Asterisks denote the statistical significance of differences in the relative abundance of the respective auxotrophy between the SIBO and the non-SIBO group (Wilcoxon rank sum test, \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ ).  $p$  values are FDR-adjusted. (b) Abundance-weighted average of the number of auxotrophies per gut microbial genotype in SIBO and non-SIBO (Wilcoxon rank sum test,  $p = 0.036$ ).

Further, the peptidase gene profiles were compared with microbiomes of non-SIBO and SIBO subjects. Some peptidases occurred either in the human gut microbiome of SIBO or non-SIBO (Fig 15a, Supplementary Table S11, S12). The PCoA showed that the peptidase profiles are distinct from the microbiomes of non-SIBO and SIBO subjects (Fig. 15c, PERMANOVA,  $R^2 = 0.17$ ,  $p = 0.001$ ). Further, the abundance-weighted average of putatively extracellular peptidases was higher in the microbiomes of SIBO patients compared to non-SIBO samples (Wilcoxon rank sum test,  $p = 0.028$ , Fig. 15b). In brief, auxotrophy abundances are reduced, but extracellular peptidase genes are more abundant in the microbiomes of SIBO patients.



**Figure 15: Microbial peptidase profiles in duodenal microbiomes of SIBO and non-SIBO samples**

(a) Heatmap of individual peptidase abundances (range-scaled) in small intestinal microbiomes from SIBO and non-SIBO samples. (b) Abundance-weighted average of

peptidases in the microbiomes of SIBO and non-SIBO samples (Wilcoxon rank sum test,  $p$  value  $< 0.0001$ ). (c) PCoA plot of peptidase profiles of the microbiomes of SIBO and non-SIBO samples (PERMANOVA based on Bray-Curtis Distances of pairwise peptidase profiles,  $R^2=0.17$ ,  $p=0.001$ ).

## 4 Discussion

Here, we compared the auxotrophy and extracellular peptidase gene profiles of the microbiomes in the small and large intestine. Moreover, we investigated whether the profiles exhibited distinct patterns in the dysbiotic state of SIBO compared to non-SIBO subjects. Auxotrophy abundances were higher in the microbiomes of the small intestine than the large intestine, which aligns with the hypothesis that the small intestinal microbiome secures more amino acids from host dietary sources than the large intestinal microbiome. In the dysbiotic state SIBO, auxotrophy abundances were reduced in the duodenal microbiome. Peptidase gene were either more abundant in small or large intestinal microbiomes, suggesting that available protein sources along the digestive tract influence the microbial peptidase gene profile. The abundance-weighted average of extracellular peptidase genes was reduced in the small intestinal microbiomes than the large intestinal microbiomes but increased in the microbiomes of SIBO patients.

Our study provides a comprehensive overview of the distribution of peptidases and amino acid auxotrophies in the small and large intestine microbiome in non-dysbiotic and dysbiotic states.

The amino acid-rich environment in the small intestine might promote higher auxotrophy abundances in non-dysbiotic microbiomes

The higher availability of free amino acids in the small intestine resulting from the digestion of dietary proteins [187], [201] suggests that the microbiome's dependence on extracellularly accessible amino acids is higher in the small than in the large intestine. Our results support this hypothesis, as we observed higher auxotrophy frequencies for several essential (tryptophan, leucine, histidine, lysine, phenylalanine) and non-essential amino acids (proline, arginine, tyrosine, serine) in microbiomes of the small intestine compared to the large intestine (Fig. 11a). This suggests that the growth conditions provided by the amino acid-rich environment in the small intestine might confer a fitness advantage for auxotrophic bacteria over prototrophic bacteria.

That auxotrophic bacteria can have a fitness advantage over prototrophic bacteria in environments providing the respective metabolites was previously reported [69]. Furthermore, the higher auxotrophy abundances for large amino acids (e.g., tryptophan, histidine) may reflect the more significant costs of biosynthesizing these compounds [96]. Additionally, tryptophan and histidine are slowly absorbed via the small intestine mucosa compared to other proteinogenic amino acids and might be available longer for uptake by the small intestine microbiome [202]. Interestingly, amino acid auxotrophies for threonine, methionine, phenylalanine, and cysteine were higher in large intestinal microbiomes. Threonine is available in the large intestine in the backbone of mucin glycoprotein. For example, *Akkermansia muciniphila*, auxotrophic for threonine, can degrade the mucin glycoproteins and is reported to be threonine auxotrophic [94]. Cysteine and methionine are present in the colonic mucosa as intermediates of the homocysteine cycle [51], and by their secretion to the lumen, they could be utilized by cysteine and methionine auxotrophic bacteria. The secretion of cysteine from the mucosa to the lumen was shown in the small intestine [203]. Further, the release of amino acids by prototrophic bacteria is another potential amino acid source and could be a source for phenylalanine. In summary, the heterogeneous auxotrophy profiles of the small and large intestines are primarily attributed to their distinct microenvironments, which result in differential amino acid availability.

Peptidase gene profiles are distinct between small or large intestinal microbiomes

Our results indicated distinct profiles of extracellular peptidase genes in small and large intestinal microbiomes (Fig. 12). The gastrointestinal tract anatomy provides diverse nutritional conditions, leading to numerous ecological niches [204]. Occupation of these niches by different bacterial species might contribute to the observed distinct peptidase profiles of the small and large intestine microbiomes. Enzyme repertoire studies on human gut microbiome isolates suggest microbial specialization for protein sources in several different microbial species [205], [206]. For example, *Akkermansia muciniphila*, a known colonizer of the large intestine, harbors a repertoire of peptidases for mucin glycoprotein degradation [205]. Similarly, *Paraprevotella* species are recognized as trypsin degraders in the large intestine [206]. Specific strains from various bacterial species are further recognized in gluten degradation in the duodenum [207]. Thus, bacteria in the human gut might specialize in protein sources tailored to

their microenvironment, which explains the repertoire of specific peptidases along the gastrointestinal tract.

Abundances of auxotrophic bacteria are negatively correlated with several peptidase genes in small and large intestinal microbiomes

The microbial peptidase activity is linked to the composition of the gut microbiome [208]. It might play an important role in degrading available proteins and peptides, thereby releasing amino acids that can serve as essential nutrients for auxotrophic bacteria. Thus, proteolytic activity contributes to the availability of amino acids for the intestinal microbiome, and we hypothesized that the abundance of genes encoding extracellular peptidases in the microbiome would be positively associated with the abundance of auxotrophies. Therefore, we correlated auxotrophy abundances with peptidase gene abundances in duodenal and large intestinal microbiomes. Against our hypothesis, we observed predominantly negative correlations between the relative abundance of amino acid auxotrophies and peptidase genes in small and large intestinal microbiomes (Fig. 13). Auxotrophic bacteria in the small intestine might save metabolic costs by potentially not engaging in proteolysis since they thrive in the nutrient-rich environment of the small intestine with accessible free amino acids derived by dietary protein digestion facilitated by peptidases secreted by the human host. However, this does not explain the large number of negative correlations between peptidase genes and auxotrophic bacteria in the large intestine. Amino acid cross-feeding between prototrophic and auxotrophic genotypes might be one contributing factor, as this potential source of amino acids for auxotrophic genotypes does not necessarily require proteolytic activity.

Our study shows that the abundance of auxotrophies in small and large intestinal microbiomes is negatively correlated with the abundance of many peptidase genes, which indicates that auxotrophic bacteria in the gut might commonly utilize different mechanisms of amino acid acquisition than proteolysis.

Increased abundances of prototrophic bacteria and extracellular peptidase genes are characteristics of the microbiome of SIBO patients

The small intestinal microbiome is still understudied compared to the large intestinal microbiome. Several studies indicated that the small intestinal microbiome is involved in the pathophysiology of specific diseases [48], [50], [185], [199]. Small intestinal bacterial overgrowth (SIBO) is characterized by an abnormal growth ( $>10^5$  organisms/ml) of bacteria [12] and is associated with several diseases (e.g., IBD). However, the ecology of bacteria within the gut microbiome of patients with SIBO remains largely unknown [198], [209]. Investigating the nutritional dependencies of small intestinal bacteria through the study of amino acid auxotrophies could provide new insights into the ecology of the small intestinal microbiome in SIBO, and improve our understanding of its role in the development of SIBO.

Here, we compared the amino acid auxotrophy profiles of patients with SIBO and non-SIBO controls. The abundance-weighted average of auxotrophies was reduced in the microbiome of patients with SIBO (Fig. 14b). This is in line with previous findings demonstrating increased small intestine colonization with colonic bacteria in SIBO [210], since we also observed that bacteria of the large intestinal microbiome commonly have less amino acid auxotrophies compared to the small intestine (Fig. 11b). There could be several potential causes for the colonization of the small intestine with atypical bacteria from the colon environment in SIBO. SIBO is frequently seen in patients with reduced gastric acid, which could lead to an increased pH in the small intestine compared to healthy individuals and creates a growth environment closer to the one found in the large intestine. Therefore, more colonic bacteria might be able to survive in this less acidic environment. Further, the motility of the gut of SIBO patients is commonly reduced [211]. This might lead to less bacteria transport from the small intestine to the large intestine [212]. Stasis is frequently observed in patients with SIBO [213], [214], further suggesting a slower transit time of food through the gastrointestinal tract, allowing nutritional resources to remain available for bacteria in the small intestine for a longer period. They could promote the overgrowth of atypical bacteria in the small intestine, which is in a non-dysbiotic state characterized by an acidic environment with quickly changing nutritional conditions [15].

Several factors may promote the colonization of colonic bacteria in the microbiome of SIBO patients. However, the connection between these changes and auxotrophies, as well as why auxotrophies are higher in the non-dysbiotic small intestinal microbiome

but lower in a dysbiotic state, has not yet been explored. Auxotrophic bacteria likely utilize free amino acids released during dietary protein digestion by proteolytic enzymes secreted into the gut lumen by the human host. While prototrophic bacteria can also access these amino acids, auxotrophic bacteria may have a fitness advantage in the nutrient-rich conditions [69] of the small intestine, which could explain their higher abundance there. Moreover, the overgrowth of bacteria in SIBO could lead to limited nutritional resources through increased nutrient competition between bacteria. Therefore, the gut environment in SIBO patients might favor prototrophic bacteria because they can synthesize the specific growth-essential nutrients (i.e., amino acids) and are less dependent on their nutritional environment compared to auxotrophic bacteria. In sum, while free amino acids available by dietary protein digestion may support the survival of auxotrophic bacteria in the small intestine in a non-dysbiotic state, the shifts in nutrient availability and competition in SIBO favor prototrophic bacteria. Thus, increased prototrophic and reduced auxotrophic bacteria in SIBO can be explained by the changed nutritional landscape in this dysbiotic state.

Our analysis indicates that the small intestinal microbiome in patients with SIBO has a higher abundance of genes encoding potential extracellular-acting peptidases compared to healthy controls. This increase suggests that SIBO-associated bacteria may engage in more aggressive nutrient acquisition strategies, potentially "competing" with the host for nutrients, such as by the uptake and catabolism of amino acids. Supporting this, Chojnacki et al. (2022) reported that urinary tryptophan levels increased in patients with SIBO following antibiotic treatment [45]. Therefore, the elevated abundance of extracellular peptidase genes may indicate enhanced bacterial peptidolysis, potentially leading to protein deficiency symptoms in patients with SIBO. Conclusively, in a non-dysbiotic state, auxotrophic bacteria are increased in small intestinal microbiomes compared to large intestinal microbiomes, but they are decreased in a dysbiotic SIBO. The microenvironment in the microbiome of SIBO patients favors prototrophic bacteria and increased extracellular peptidase genes in bacteria.

## Limitations

There are some limitations in our study. Predictions of auxotrophy and peptidase profiles are based on the mapping of 16S rRNA sequences to human gut reference microbiome sequences. Since whole genome sequences of small intestinal

microbiomes are still scarce, we decided on this approach. Therefore, the predictions are based on taxonomic information. In addition, the overall number of microbes and diversity are another contributing factor. The small intestinal microbiome harbors a lower density of microbial cells than the large intestinal microbiome [12]. Further, in SIBO, the number of microorganisms is higher than in a non-dysbiotic state [195]. Additionally, the diversity is most commonly reported to be lower in the small compared to the large intestinal microbiome [215] and could have influenced our results. Also, the presence of peptidase genes was analyzed in small and large intestinal microbiomes in dysbiotic and non-dysbiotic states, but the approach cannot conclude the actual peptidolytic activity.

## 5 Conclusion

The nutritional requirements in the small intestinal microbiome are still understudied despite its association with the pathophysiology of several diseases. This study showed that the small intestinal microbiomes have higher nutrient requirements, as indicated by increased amino acid auxotrophies, than the large intestinal microbiome in non-dysbiotic states. However, amino acid auxotrophies are reduced in the small intestinal microbiome in small intestinal bacterial overgrowth (SIBO), underlining different nutrient requirement patterns in dysbiotic states. Our study further suggests a disrupted microbial ecology in SIBO patients, highlighted by increased extracellular peptidase genes in their microbiome. As alternatives to antibiotics are increasingly considered for SIBO treatment, understanding the specific nutritional requirements of the small intestinal microbiome becomes crucial. Therefore, the results of this study could advance the knowledge to restore microbial balance and improve therapeutic outcomes for patients with SIBO.

## 6 Material and Methods

### 16S rRNA gene sequencing analysis

16S rRNA gene sequencing data were re-analyzed from two previous studies. The first study took samples from four gastrointestinal sites: the duodenum, jejunum, farthest distance, and stool [7]. The farthest distance refers to the location in the gastrointestinal tract that could be reached during sampling. The second study investigated the duodenal microbiome in SIBO [199]. Information about the sampling method is provided by the authors of the studies [7], [199]. The 16S rRNA gene

sequencing data from the first study were downloaded with the SRA toolkit [216] (version 3.0.0) from NCBI (BioProject ID: PRJNA590519) [7]. 16S rRNA gene sequencing data from the SIBO study were taken from the ENA Server (BioProject ID: PRJNA525828) [199]. Primer sequences were detected and removed from read sequences using the software Cutadapt [217]. Sequences were only kept for further analysis if the forward and reversed primer were found and removed. After trimming, the sequence length was shortened to 275. The package DADA2 (version 1.24.0) was used to generate the ASVs [218]. After checking the quality of the reads, forward reads with a length of 265 and reverse reads with a length of 245 were kept (Tolerated maximal errors (maxEE = 2) from the first study of the sample taken from four gastrointestinal sites. For the SIBO study, the forward reads were filtered for the length of 240 and reverse reads for 230 (maxEE = 2). Mapping to representative genomes of the human gut microbiome.

Representative nucleotide sequences of the calculated ASVs were mapped to the 16S rRNA genes within genomes of the Human Reference Gut Microbiome (HRGM) catalog, which consists of 5414 distinct prokaryotic species [168][146]. The mapping was based on pairwise sequence alignments using BLASTN (version 2.9.0+, REF) with a minimum query (ASV) coverage of 95% and a minimum sequence identity of 97 %. In the case of multiple hits, only hits with the maximum identity were retained for further analysis.

### Prediction of peptidases

The MEROPS peptidase unit database harbors information about peptidase units [200]. All protein coding genes and their translation to amino acid sequences within the genomes of the HRGM collections were predicted using PRODIGAL (version 2.6.3) [219]. For the prediction of peptidases in the HRGM sequences, a BLASTP search was performed with the MEROPS scan sequences (e-value < 0.01, query coverage > 85 %) [220]. After the blast search, only hits with a bitscore higher than 100 and a coverage of the MEROPS peptidase unit of 95% or higher were considered as putative peptidases in the HRGM genomes. Next, we wanted to identify those peptidase genes in the HRGM genomes, that likely encode peptidases that act on extracellular peptides (e.g., via enzyme secretion or cell surface-attached enzymes). Therefore, we used SignalP (v6.0) [221] to predict signal peptides in the peptidase amino acid sequences that indicate enzymes that are putatively secreted or translocated across the



cytoplasmic membrane. All signal peptides included in SignalP 6.0 were included in the analysis.

### Prediction of auxotrophies

Genome-scale metabolic models of individual genomes from the HRGM collection were reconstructed using gapseq (version 1.2, commit 13d88a68) [111]. All metabolic models as well as the exact gapseq commands for the reconstructions are available at Zenodo [179]. Only genomes with a completeness  $\geq 85\%$  were considered for further analysis. Auxotrophies were predicted in R (4.3.0) with the R package sybil [170] (2.2.1) using metabolic modeling as described previously [190]. In brief, flux balance analysis was performed, with the objective function defined as the flux through the biomass reaction [120]. The growth of the models was tested with and without the focal amino acid in the *in silico* growth medium. The growth medium was predicted by gapseq [111]. When no growth of the model without the focal amino acid in the medium was determined compared to the growth rate with the medium containing the amino acids, it was classified as auxotrophic for the amino acid tested.

### Statistical analysis

All statistical analysis was performed with R (4.3.0). The correlation of auxotrophy abundances with the gastrointestinal region (ordered factors: 1 - duodenum, 2 - jejunum, 3 - farthest distance reached in the small intestine, and 3 - large intestine) was determined with the Kendal Rank Correlation, and  $p$  values were adjusted for multiple testing using the false discovery rate (FDR) method. Differences in auxotrophy abundances of the microbiome in SIBO and non-SIBO were statistically evaluated with the Wilcoxon rank sum test, and  $p$  values were FDR-adjusted. Differences in the abundance-weighted average of auxotrophies and peptidases in small and large intestinal microbiomes in a non-dysbiotic and dysbiotic state were statistically determined with the Wilcoxon rank sum test. Differences in the abundance of peptidase in the duodenum and large intestine were statistically compared with the Wilcoxon signed rank test, and  $p$  values were adjusted using the FDR method. An FDR-adjusted  $p$  value of  $<0.05$  was considered statistically significant in all tests.

Peptidase profiles between small and large intestine microbiomes and microbiomes of SIBO and non-SIBO subjects were compared with PCoA analysis using the Bray-Curtis distance matrix. A permutational multivariate analysis of variances

(PERMANOVA) as implemented in the R-package 'vegan' (v2.6-4) [178] was used to statistically compare the peptidase profiles between SIBO and non-SIBO samples. The pairwise Bray-Curtis distance between peptidase profiles was used as a metric for the distance matrix.

Associations between auxotrophies and peptidases were tested using Spearman's rank correlation.

### Data availability

16S rRNA amplicon sequencing data of the Reimagine study (BioProject ID: PRJNA590519) [7] and SIBO study (BioProject ID: PRJNA525828) [199] are available on NCBI as the data was submitted by the authors of the respective studies. Genome-scale metabolic models of the HRGM catalogue are available via Zenodo [179]. Auxotrophy predictions for all HRGM genomes can be found in the supplemental material of the published article by Starke et al., 2023 [190].

### Code availability

The R code with information about the statistical data analysis and visualization is uploaded on GitHub: [https://github.com/SvStarke/Auxos\\_SI\\_LI\\_microbiomes](https://github.com/SvStarke/Auxos_SI_LI_microbiomes)

### Acknowledgments

This work was supported by the DFG Cluster of Excellence 2167 "Precision medicine in chronic inflammation" (SWa, KA), the BMBF (e:Med Juniorverbund "Try-IBD" 01ZX1915A and 01ZX2215, DFG RU5042 (DH, KA)). A.P. was the recipient of the CIFRE fellowship number 2019/0659 co-funded by Biomathematica and was supported through the Micalis team by the French National Research Agency (RestorPro no 13001827 and ANR-19-CE14-0041-01 HEPADIMAIT, ANR-18-CE18-0019 TITAN) and the European Union's Horizon 2020 research and innovation programme IHMCSA under grant agreement N°964590 and DOMINO under grant agreement N°101060218.

### Additional information

All authors declare no financial or non-financial competing interests.

## Chapter IV: Reliability of *in silico* auxotrophy prediction methods

*In preparation*

Svenja Starke<sup>1</sup>, Silvio Waschina<sup>1</sup>

<sup>1</sup>Institute of Human Nutrition and Food Science, Kiel University, Department of Nutriinformatics, Heinrich-Hecht-Platz 10, 24118 Kiel, Germany

## 1 Abstract

Auxotrophic bacteria cannot synthesize certain essential metabolites, relying instead on external sources for growth. Identifying auxotrophies in microbial genotypes is crucial for understanding microbial ecology, as interactions between auxotrophic and prototrophic organisms can influence community dynamics. Computational methods have been developed to predict amino acid auxotrophies from genome sequences, using either genome-scale metabolic modeling or genome mining of predefined biosynthetic pathways. However, the accuracy of predictions based on automated genome-scale metabolic network reconstructions has recently been questioned. In this study, we compare two state-of-the-art automated auxotrophy prediction tools and assess their accuracy using known amino acid auxotrophies in 36 human gut microbiome bacteria. Our results show that both *in silico* methods are reliable methods for auxotrophy prediction.

## 2 Introduction

The composition and metabolic activity of microbiomes play a critical role in shaping ecosystems in host-associated and free-living environments with potential impact on the host phenotype and environmental dynamics [49], [61], [222], [223], [224]. For instance, rumen microorganisms are being studied for their potential to reduce greenhouse gas emissions [222], while the amino acid metabolism of the human gut microbiome is linked to inflammatory and metabolic diseases [39], [49]. Cross-feeding plays a crucial role in shaping ecosystem structure and stability. For example, cross-feeding promotes the stability of microbial communities, particularly those with auxotrophic bacteria. Auxotrophy, the inability of bacteria to synthesize certain essential nutrients like amino acids, is increasingly recognized as a key factor in microbial ecology [20]. Recent research has highlighted their ubiquity across microbial systems, emphasizing their ecological importance [68]. Auxotrophic bacteria are particularly prevalent in host-associated environments, such as the human gut microbiome [68], [79], [190]. Recently, a study showed that they are positively associated with the diversity and long-term stability of the human gut microbiome [190]. Auxotrophies are experimentally verified by observing the growth with and without the amino acid in question [72]. However, experimental validation of auxotrophies in bacteria is challenging but the rise in high-throughput sequencing offers new possibilities for predicting auxotrophies *in silico*. In this context, genome-scale

metabolic modeling and genome annotation has gained interest as tools for the prediction of auxotrophies [68], [79], [190]. Metabolic modeling is a computational approach used to analyze the metabolic networks of microorganisms, enabling predictions about their growth requirements for certain nutrients. Genome annotation, on the other hand, is the process of identifying genes within a genomic sequence based on known databases or algorithms. Nevertheless, discrepancies between computational predictions and experimentally verified auxotrophies remain a significant challenge [72]. This highlights the importance of validating *in silico* methods, such as genome-scale metabolic modeling, to ensure their reliability.

So far, genome-scale metabolic modeling has been criticized as an unreliable method for predicting auxotrophies [119], [225]. In this study, we aim to validate auxotrophies predicted using genome-scale metabolic modeling against experimentally verified auxotrophies from bacterial strains. Additionally, we evaluate the reliability of genome-scale metabolic modeling in predicting amino acid auxotrophies. To provide a comprehensive assessment, we compare these validation results with auxotrophy predictions generated by GapMind, offering a detailed evaluation of the predictive capabilities of these *in silico* approaches.

### 3 Material and Methods

#### Experimentally verified auxotrophies

Ashniev et al. enumerated a list with experimentally verified amino acid auxotrophies [79]. The list includes 36 human gut bacterial strains. For some bacteria, auxotrophies were not predicted for all twenty proteinogenic amino acids. In the supplemental material, information about the references of the 36 experimentally verified auxotrophies can be found.

#### Reconstruction of genome-scale metabolic models

Genome-scale metabolic models were reconstructed with gapseq [111] (version 1.2 2dfa8c80). The reconstruction process takes several steps: (I) Pathways and transporters prediction, (II) Draft network reconstruction, (III) Gapfilling. (I) Reactions and transporters are predicted based on the genome sequence to state the presence of pathways in the genome of interest. Definitions of pathways are taken from MetaCyc [148]. If enough reactions based on sequence evidence are found, the pathway is considered as present. (II) the draft model network is constructed based on the found

pathways and transporters. Further, a biomass reaction is added to the draft model. (III) Gap's in the model are filled by an algorithm which adds especially reactions with a already found high sequence evidence close to the cut-off value. Reactions are added in order to enable the growth (biomass production) by the reconstructed model on a growth medium which was automatically chosen by gapseq based on found reactions. More detailed information can be found in our previous publication[190]. The reconstructed models can be found on Zenodo [181].S

### Prediction of auxotrophies with genome-scale metabolic modeling

The reconstructed metabolic models were metabolically modeled. Flux-balance analysis [120] was used to predict the ability of models to grow (form biomass) with and without the twenty proteinogenic amino acids. Amino acid auxotrophies are predicted if a model is not able to grow without an amino acid.

### Prediction of auxotrophies with GapMind

Genome sequences of the 36 bacterial strains were annotated with GapMind [114]. The detailed information of the GapMind analysis are in the Supplementary. The underlying query databases was built on 25<sup>th</sup> July 2024. GapMind uses ublast or Hmmer to identify genes in the query genome. We predicted auxotrophies based on the definition by Ramoneda et al. [68] Therefore, reactions in a pathway were considered as present with at least a medium confidence meaning at least 40% identity and 70% coverage or at least 30% identity and 80% coverage to the known protein in the database. Auxotrophies were predicted if more than 40% of the reactions were missing in the pathway. Alanine, glutamate, and aspartame are not included. All predictions were made on April 9, 2024.

## 4 Results

### *In silico* prediction of amino acid auxotrophies

For 36 human gut bacterial strains, amino acid auxotrophies are experimentally verified and their genome sequences are available via NCBI [79]. These bacterial strains were used to investigate the reliability of in silico prediction of amino acid auxotrophies with GapMind [114] and metabolic modeling of genome-scale reconstructed models from gapseq [111].

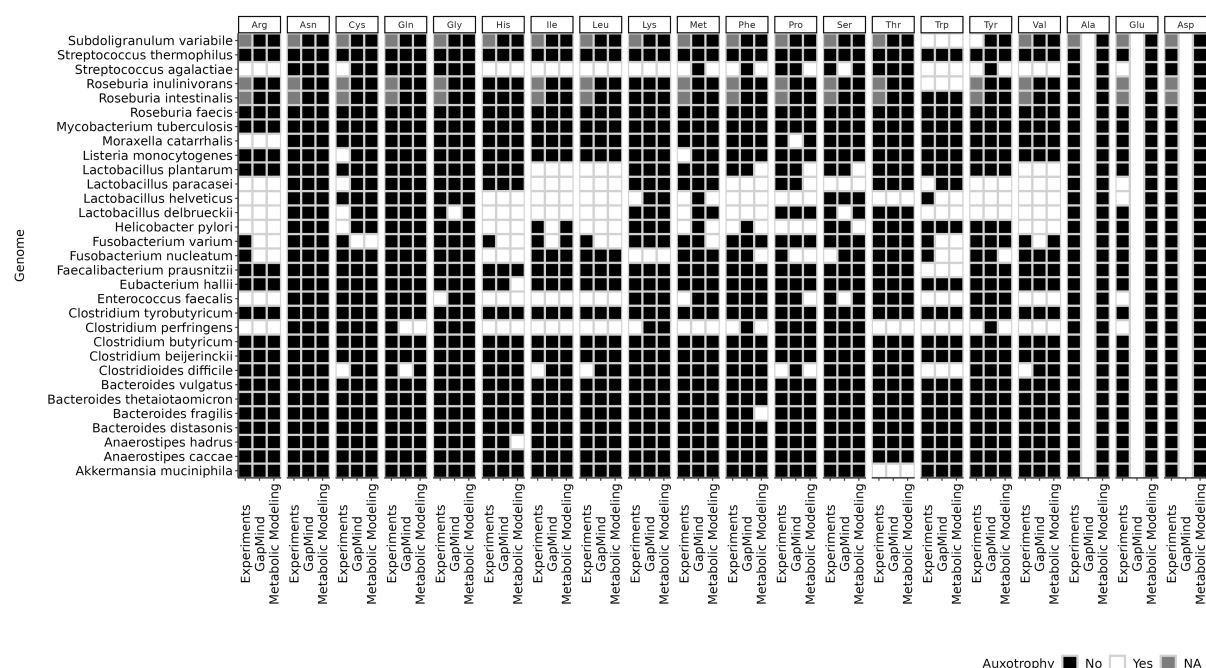
GapMind [114] predicts amino acid auxotrophies via genome annotation. The presence of an amino acid pathway is stated by identifying candidates for each step of a pathway and assign them with confidence levels. Reactions are present in a pathway with at least a medium confidence. If more than 40% of the reactions are missing in an amino acid pathway, the organism is predicted as auxotrophic for it. For the prediction of auxotrophies with metabolic modeling, genome-scale metabolic models were reconstructed with gapseq [111]. Flux-Balance analysis was used to predict the growth of the model with and without the proteinogenic amino acid [120]. If a model was not able to grow without the amino acid, an auxotrophy was predicted.

### High accuracy of *in silico* methods for auxotrophy prediction

Comparison of predicted auxotrophies via GapMind [114] and metabolic modeling with gapseq models to experimentally verified auxotrophy predictions suggests a high reliability of *in silico* methods (Fig. 16). In sum, 89 auxotrophies and 466 prototrophies were experimentally verified. Metabolic modeling predicted 94 auxotrophies and 461 prototrophies and GapMind predicted 81 auxotrophies and 474 prototrophies.

Auxotrophies predicted with GapMind had a sensitivity of 94.3%, specificity of 76.5%, and accuracy of 91.7% whereas auxotrophies predicted with metabolic modeling had a sensitivity of 96.1%, specificity of 75.5%, and accuracy of 92.6%. Especially, differences were seen in *in silico* vs. experimentally predicted auxotrophies for cysteine, proline, serine and methionine. 7 out of 36 cysteine auxotrophies were wrong predicted in both *in silico* methods. 6 out of 36 methionine and proline auxotrophies were wrong predicted with genome-scale metabolic modeling.

Overall, our results show that genome-scale metabolic modeling as well as GapMind have a accuracy > 90%.



**Figure 16: Comparison of *in silico* methods to predict auxotrophies**

Comparison of amino acid auxotrophies experimentally validated auxotrophies of 36 human gut bacterial strains with predicted auxotrophies by GapMind and genome-wide metabolic modeling.

## 5 Discussion

The results of our study show that *in silico* methods are a reliable method to predict amino acid auxotrophies in bacteria. We used bacterial genomes with experimentally verified amino acid auxotrophies for validation. *In silico* methods for amino acid auxotrophy predictions are genome annotation (GapMind) and genome-scale metabolic modeling. Both methods showed >90% sensitivity and accuracy. However, specificity was lower with 75.5% (Metabolic Modeling) and 76.5% (GapMind).

So far, not all amino acid biosynthesis pathways are known [226]. Therefore, *in silico* methods might not be able to predict all amino acid auxotrophies correctly. Our findings regarding false predicted cysteine auxotrophies were also described by Ramoneda et al. [68] Furthermore, in GapMind and in metabolic model reconstruction process the absence/presence of genes in amino acid biosynthesis pathways is stated by comparing the query genome to a database via BLAST [111], [114]. Missing organisms in the database are another reason for the observed incompatibility of *in silico* methods. However, the accuracy with >90% indicate a high reliability of both GapMind and genome- scale metabolic modeling for auxotrophy predictions in bacteria.



## 6 Conclusions

Overall, our study shows that despite the known discrepancies between experimentally verified and *in silico* predicted auxotrophies, metabolic modeling with gapseq reconstructed models and GapMind are both reliable tools for autotrophy predictions.



## Chapter V: General Discussion

### 1 Distribution of amino acid auxotrophies in small and large intestinal microbiomes

In this thesis, the distribution of amino acid auxotrophies in small and large intestinal microbiomes was analyzed. So far, it was known that amino acid auxotrophies are more prevalent in bacteria living in host-associated environments [121] but their distribution particularly in the human gut microbiome remained understudied. Therefore, this thesis's first major finding was that amino acid auxotrophies are common in the human gut microbiome (Fig. 6). One other study investigated the absence of genes in amino acid biosynthesis pathways in the human gut microbiome and showed that the human gut microbiome lost genes for the biosynthesis of many amino acids [79]. Another recent study confirmed the presence of amino acid auxotrophies in the human gut microbiome. Bacteria living in the human gut had on average >2 auxotrophies for amino acids [68]. Conclusively, auxotrophies in human gut bacteria are common, but what is the reason for the abundance of auxotrophies in the human gut microbiome?

In host-associated environments like the human gut, more auxotrophies are found in bacteria because the gut environment might be more nutrient-rich than environments of free-living bacteria. Auxotrophies can develop by an adaption of organisms to their nutrient-rich environment (Chapter I 2.2). One study showed that in a amino acid-rich environment with more generations more auxotrophies developed in bacteria [88]. Therefore, auxotrophies might be more common in the human gut microbiome because the gut environment provides a nutrient-rich landscape and natural selection leads to the loss of genes in gut bacteria. In particular, the distribution of auxotrophies in the large intestinal microbiome could be led back to the theory of the Black Queen Hypothesis (see Box 2 in Chapter I 2.2), which argues that organisms lose biosynthetic genes for amino acids that are provided by other members of the community. Since cross-feeding might be an important nutrient source of amino acids for auxotrophic bacteria in the large intestine (Chapter II), some bacteria lose genes that are provided by others. Though, the abundance of auxotrophies for some amino acids were increased but decreased for others both in small and large intestinal microbiomes. In the following section it will be reasoned why some genes of amino acid biosynthesis pathways are more often lost than others.

## 1.1 Gene loss of human-essential and metabolically expensive amino acids in small and large intestinal microbiomes

Amino acid auxotrophies were present in the microbiomes of the small and large intestines in particular for human essential amino acids (Chapter III). However, they were higher in small than in large intestinal microbiomes. The environment in the small intestine is more nutrient-rich, especially of free amino acids, due to protein digestion, and could explain the higher abundance of amino acid auxotrophies in the small intestinal microbiome (Fig. 11) but it does not explain why particularly auxotrophies for human essential amino acids higher. Humans take up their essential amino acids by diet because they lost genes for the biosynthesis of many amino acids when they started to consume other organisms [227]. Consequently, dietary protein digestion not only meets human metabolic needs but also supplies amino acid auxotrophic bacteria in the small intestine with essential amino acids, potentially giving them a fitness advantage in this environment [23]. It might be expected that human-essential amino acids are present in higher abundances in food relative to non-essential amino acids to meet this demand [228]. However, interestingly, human-essential amino acids are often found in lower relative concentrations in food. For example, amino acids such as tryptophan and histidine have the lowest abundances (respectively, median = 0.012 g/g total amino acid content, median = 0.028 g/g total amino acid content) in human food [228]. However, this aligns with the WHO's daily recommendations, which indicate that tryptophan and histidine are required in smaller amounts (mg/g protein) compared to other essential amino acids [229]. Although humans can synthesize non-essential amino acids, these are still required in larger quantities, while essential amino acids—though critical—are needed in smaller proportions [230]. This reduced proportional requirement ensures that dietary intake might suffice to meet metabolic needs, both for humans and their gut bacteria. Populations consuming significant amounts of animal-derived proteins (e.g., meat, eggs, dairy) further stabilize the dietary supply of essential amino acids in the gut lumen [228], [231]. Nevertheless, the abundance of auxotrophies in the small intestinal microbiome may vary depending on dietary patterns and the amino acid content of food [228].

However, this argumentation might explain auxotrophy abundances in small intestinal microbiomes but not in large intestinal microbiomes since access to dietary protein is limited [21]. Additionally, some auxotrophies for essential amino acids were higher in the large intestinal microbiome. In this context, auxotrophies may be sustained by other

amino acid sources like cross-feeding which will be discussed in Chapter V 3.1. Therefore, the abundance of auxotrophies for human-essential amino acids might be explained particularly in large intestinal microbiomes by the higher metabolic costs of synthesising essential amino acids. Essential amino acids such as tryptophan, histidine, and BCAAs require energy-intensive pathways for *de novo* synthesis, involving complex intermediates that are not derived from central metabolic pathways like glycolysis or the citric acid cycle [232]. By losing the genes for these biosynthetic pathways, bacteria can save energy and resources, which could provide a selective advantage [69]. In comparison, the biosynthesis of non-essential amino acids, such as glycine, is less costly and often derived from simpler metabolic intermediates, making it less likely for bacteria to lose the associated genes [232].

In conclusion, amino acid auxotrophies are more abundant in both the small and large intestinal microbiomes for human-essential and metabolically expensive amino acids. Their prevalence reflects a combination of factors, including the high cost of biosynthesis, the reliable dietary supply of essential amino acids, and microbial interactions within the gut.

## 2 The influence of auxotrophic bacteria on the diversity and stability of small and large intestinal microbiomes

This thesis showed as a major result the positive association between auxotrophic gut bacteria and the diversity and long-term stability of the human gut microbiome (Fig. 8C, 9A). Further, a positive association was found between the cross-feeding of amino acids and the diversity and long-term stability of the human gut microbiome (Fig. 8D, 9B). So far, it was known that auxotrophies influence the ecology of microbial systems [109], [145]. However, their influence on the human gut microbiome has not yet been analyzed. Therefore, these results are particularly intriguing because studies are still trying to understand the mechanisms behind the stability of the human gut microbiome, as its microbial alterations are reflected in the host phenotype and serum levels [151]. Additionally, in several diseases, the diversity is reduced in the human gut microbiome [49], [61], [199], [233]. Communities with high species richness tend to be more stable over time [151]. In this chapter, it is reasoned why the human gut microbiome's higher abundances of auxotrophic bacteria, which are more dependent on their environment and potentially more susceptible to environmental changes, are more stable and what

the positive link is between cross-feeding in the human gut microbiome and diversity and long-term stability.

The result of the positive association between auxotrophic gut bacteria and the long-term stability might be surprising because auxotrophic organisms have a higher vulnerability than prototrophic organisms due to their dependency on the nutritional environment. So far, one theoretical study analyzing the robustness of auxotrophic communities showed that auxotrophic communities are less stable when facing perturbations. However, when cross-feeding was observed in auxotrophic communities, they were more stable towards external disruptions [109]. Also, network patterns were discussed regarding the influence of auxotrophies on community stability found in aquatic environments and revealed that cross-feeding interactions in communities with auxotrophies could lead to shifts in the microbiome's stability. Indeed, this thesis observed a positive association between cross-feeding and long-term stability.

In fact, cross-feeding could be the missing link that explains the positive effect of auxotrophies on the stability of the human gut microbiome. Due to the cross-feeding of essential nutrients in the community, the dependency on external nutritional availability might be reduced. It could indicate a more stable microbiome less sensitive to external nutrient fluctuations. But why are bacteria cross-feeding in a community, and how does the community benefit? Gene loss favors cross-feeding of essential nutrients in microbial communities [98]. Organisms lose genes for metabolite synthesis and specialize in synthesizing specific nutrients, thereby providing essential nutrients for community members. The costs for synthesizing essential nutrients are shared between the members of the community, which is described by the division of labor (Chapter I 2.3) [100].

Cross-feeding influences not only the stability of the human gut microbiome but also diversity [234]. This study showed that cross-feeding positively influenced the diversity of the human gut microbiome. Cross-feeding expands the metabolic niche for bacteria [107], and therefore, more bacteria can survive, resulting in a more diverse microbiome, which could explain the positive association between auxotrophic gut bacteria and the microbiome's diversity. Evenness is an important indicator for diversity but further for stability and is positively influenced by cross-feeding in the gut microbial community. Cross-feeding increased the evenness of the gut microbiome in a gnotobiotic mouse model [235]. In addition, more diverse microbial communities are

less susceptible to disruptions. They protect against the invasion of pathogens by exchanging the nutrients essential for pathogens between members of the gut community, making them unavailable for pathogens [62]. The protection against pathogens by exchanging nutrients in a diverse human gut microbiome and, therefore, keeping the human gut microbial composition stable promotes our observed association between cross-feeding and diversity.

Further important to mention is that stool samples were analyzed for the positive association between auxotrophies and long-term stability and diversity; therefore, the results refer to the large intestinal microbiome. This thesis did not investigate the influence of auxotrophies on the long-term stability of the small intestinal microbiome. In general, the small intestinal microbiome might be less stable than the large intestinal microbiome since the nutritional environment in the small intestine is subject to nutrient fluctuations [15]. Therefore, the abundance of auxotrophic bacteria might change more quickly in the small intestinal microbiome. While cross-feeding may play an important role in supplying amino acids for auxotrophs in the large intestine and might be linked to the long-term stability of the human gut microbiome, it may not be the primary mechanism for auxotrophs in the small intestine. The nutrient-rich environment of the small intestine might not favour cross-feeding as a study showed that supplementing the environment with amino acids led to less cross-feeding in the microbial community [106]. In the next subchapter, nutrient sources of auxotrophic bacteria in the small intestine are discussed and compared to the large intestine. Further, the influence of auxotrophic bacteria on the diversity of the small intestine might differ. Generally, the diversity of the small intestinal microbiome is decreased compared to the large intestinal microbiome [7], [215]. The diversity of the small intestinal microbiome might be especially influenced by the influx of nutrients and dietary human enzymes [12]. Though auxotrophic bacteria might thrive in the amino acid-rich environment of the small intestine, they might not directly influence the microbiome's diversity. Since the argument for a positive association between auxotrophies and microbiome diversity in stool samples is primarily based on cross-feeding—a mechanism that may not be as prevalent in the small intestinal microbiome—the impact of auxotrophies on diversity in the small intestine might not be as statistically significant as observed in the large intestine.

Overall, auxotrophies positively influence the ecology of the human gut microbiome as auxotrophies are linked to a higher diversity and long-term stability of the human large

intestinal microbiome. Cross-feeding might strengthen the microbial network towards external perturbations and could be the reason for the observed positive association between auxotrophies and the long-term stability and diversity of the human gut microbiome. However, the influence of auxotrophic bacteria on ecology of the small intestinal microbiome still needs to be determined.

### 3 Nutritional ecology of amino acid auxotrophic bacteria in small and large intestinal microbiomes

The microbe-host relationship is a key area of research, particularly in understanding how the human gut microbiome provides essential nutrients like B-vitamins and amino acids, and how its biosynthetic capabilities are linked to various diseases. While many studies focused on the biosynthetic capabilities of the human gut microbiome, less is known about the nutritional requirements of the human gut microbiome. This thesis addressed this gap by examining the nutritional requirements of the human gut microbiome, focusing on amino acid auxotrophies—a relatively novel approach that centers on the perspective of the microbes themselves.

The findings of this thesis indicate that amino acid auxotrophies are common in the human gut microbiome, which is notable given the common assumption that the gut microbiome primarily supplies nutrients to humans. For instance, studies have estimated that daily B vitamin production by the human gut microbiome could meet human requirements [80] that likely could be estimated for amino acid production by the human gut microbiome. However, the results of this work suggest that many bacteria in the gut are not nutrient providers for humans, as they lack the ability for *de novo* amino acid synthesis. In fact, auxotrophies for human essential amino acids are more common within the human gut microbiome. Other studies support the ubiquity of amino acid but also B-vitamin auxotrophies in the human gut microbiome [68], [73], [74]. Additionally, produced amino acids by prototrophic bacteria in the late intestine might not serve the host but auxotrophic members in the microbial gut community as essential nutrient sources. However, results of this thesis suggest that the host benefits from the exchange of nutrients between microbes in the human gut. As in the previous subchapter discussed, cross-feeding within the microbial community is crucial for maintaining a stable and diverse human gut microbiome. These results indicate a shift in perspective towards the microbe's needs and nutritional requirements of essential nutrients in order to expand knowledge about gut ecology in health and



disease. Understanding the nutritional environment of amino acid auxotrophic bacteria in the gut is essential, as their reliance on external sources may be key to sustaining microbial diversity and stability, particularly in dysbiotic states associated with disease. The nutritional ecology possesses a framework to discuss the influence of amino acid auxotrophic bacteria on the gut ecology (Fig.17). In particular, interactions between the gut environment, auxotrophic gut bacteria and its nutritional requirements are evaluated. In this chapter, the nutritional ecology of amino acid auxotrophic bacteria in the human gut microbiome is discussed. First, the focus lies on nutritional sources of amino acid auxotrophic bacteria in the human gut microbiome. Thereby, the nutritional ecology of auxotrophic bacteria in small and large intestinal microbiomes are compared. Then, auxotrophic bacteria and their environment in context of nutrition are discussed to provide an overview of the nutritional ecology of amino acid auxotrophic bacteria in small and large intestinal microbiomes..



**Figure 17: Nutritional ecology of amino acid auxotrophic bacteria in the gut**

### 3.1 Potential ways of amino acids acquisition by auxotrophic bacteria in small and large intestinal microbiomes

The major nutrient source for auxotrophic bacteria in the small intestine might be free amino acids released by dietary protein digestion. Amino acids degraded from dietary proteins and peptides are more readily available for auxotrophic bacteria in the small intestine. They might enhance their growth as the abundance of auxotrophic bacteria is higher in the small intestinal microbiome compared to the large intestinal microbiome (Fig. 11). Human enzymes might be the main contributor to dietary protein digestion. However, there was also the assumption in Chapter III that auxotrophic bacteria could participate in the degradation of peptides in the gut. However, overall mainly negative correlations were observed between auxotrophic bacteria in the small intestine and peptidase genes (Fig.13). For example, there was a negative correlation observed with

the chymotrypsin-like peptidase and several amino acid auxotrophic bacteria in the small intestine. Since, dietary proteins are degraded by human enzymes to amino acids, the additional degradation of peptides by auxotrophic bacteria in the small intestine is not crucial in non-dysbiotic states. Indeed, results of this thesis indicate that auxotrophic bacteria in the small intestinal microbiome, found in higher abundances than in the large intestine, acquire free amino acids.

The acquisition of free amino acids released from dietary protein digestion might be the main amino acid sources for auxotrophic bacteria in the small intestine. However, the major amino acid source for large intestinal bacteria is different. In Chapter II, an interesting outcome was the origin of the major nutrient source for auxotrophic bacteria in the large intestine: cross-feeding. Even though some dietary proteins might reach the colon, does not reach a significant fraction of dietary amino acids and proteins the colon and are, therefore, not a reasonable nutrient source for colonic gut bacteria [21]. Also, no statistically significant association was observed between autotrophic bacteria in the large intestine and the dietary amino acid intake (Fig S7). Therefore, the results indicate that auxotrophic bacteria in the large intestine acquire essential amino acids by cross-feeding in the microbial community. Based on the theory of the Black Queen hypothesis (Box 2 in Chapter I 2.2), bacteria lose genes for metabolic costly amino acids because they are provided by other members of the bacterial community [90]. Indeed, a study investigating the secretome of human gut bacteria showed that more large proteins are secreted in the large intestine than small intestine [23]. For Proteobacteria, which showed in our study the least auxotrophy counts per genotype and are a major phylum in the human gut microbiome, more open reading frames of secreted proteins for amino acid metabolism and function were found [23]. Therefore, *de novo* synthesized amino acids by prototrophic bacteria might be secreted to the environment and utilized by auxotrophic bacteria in the microbial community. As already stated in the previous subchapter, cross-feeding might not be a reasonable nutrient source for auxotrophic bacteria in the small intestine due to the nutrient-rich environment in the small intestine.

Not only secreted amino acids by prototrophic bacteria are sources for auxotrophic gut bacteria but secreted amino acids by the host. Studies showed that a substantial amount of essential and non-essential amino acids are endogenously lost in the small and large intestine [236], [237], [238]. Those endogenously lost amino acids display

further nutrient sources for auxotrophic bacteria particularly in the large intestine since dietary amino acids do not reach the colon in substantial amounts [21].

Further, Chapter II showed an association between many auxotrophic large intestinal bacteria (e.g., threonine and tyrosine) and serum levels of bile acid derivatives was found. Bile acids are conjugated with amino acids [239]. Their microbial conjugation and deconjugation could further alter amino acid availability and display another potential nutrient source for auxotrophic gut bacteria not only in the large intestine but in the small intestine. Some microbial-conjugated bile acids are not deconjugated, which could further alter the amino acid availability. The concentration of deconjugated amino acids is further different in the gastrointestinal tract compared to the stool [240]. However, the presence of enzymes capable of conjugating/ deconjugating in auxotrophic genotypes has not yet been determined.

Another nutrient source for both small and large intestinal bacteria that was so far not mentioned are human proteins. Human protein abundance differs in the upper and lower gastrointestinal tract [240]. The variation in human proteins is higher in the upper gastrointestinal tract than in stool. Overall, human proteomes are correlated with microbial composition [240]. Indeed, associations were found between auxotrophies and peptidases involved in host protein degradation (e.g., C01.171) in Chapter III.

Several amino acid sources exist for auxotrophic bacteria in the gastrointestinal tract. Since their availability varies along the gastrointestinal tract, the source of amino acid for auxotrophic gut bacteria depends on their location in the gastrointestinal tract. Meanwhile, digested dietary peptides and amino acids might feed, especially auxotrophic bacteria in the small intestine, and cross-feeding between bacteria in the large intestine is the main nutrient source for auxotrophic bacteria. Knowledge about the amino acid sources of auxotrophic bacteria could help design therapeutic approaches to feed auxotrophic bacteria and enhance their growth in dysbiotic states of the human gut microbiome, which could help restore the microbiome.

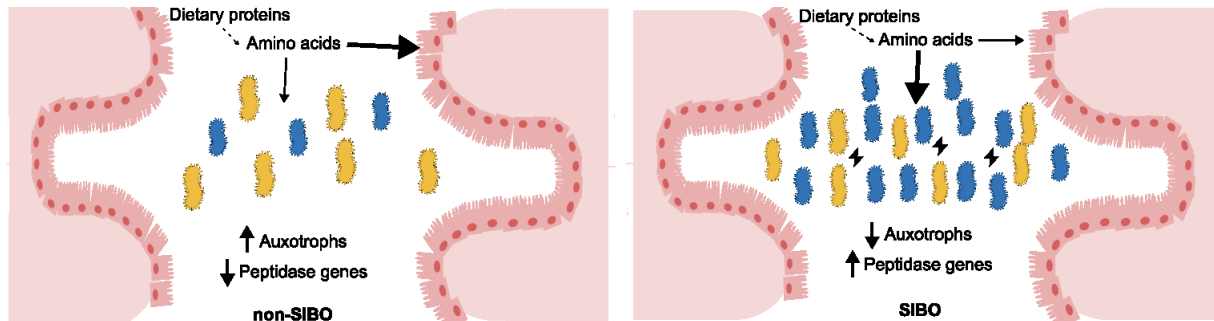
### 3.2 Ubiquity of amino acid auxotrophies in the human gut microbiome indicates a shift towards microbial nutritional ecology

Many studies showed a shift in the microbial composition in metabolic and inflammatory diseases with specific microorganism becoming more dominant in the human gut microbiome [39], [49], [233], [241]. Researchers are still questioning if the change in the microbial composition or in the environment comes first. Since amino

acid auxotrophic bacteria are widely distributed in the human gut microbiome and dependent on the nutritional environment, raises this study new discussions about the nutritional requirements of the human gut microbiome in health and disease which will be discussed in the following.

Results of Chapter III showed that amino acid auxotrophies are reduced in dysbiotic states of the small intestine. It was discussed that a change of the (nutritional) environment by reduced resource availability and increased competition might be key factors for the observed reduction of auxotrophic bacteria in the microbiome of patients with SIBO (Fig.14). Recently, another study confirmed that metabolically dependent (auxotrophic) bacteria are reduced in the gut microbial dysbiosis of IBD patients [40]. This study was particularly insightful because it involved fecal transplants from healthy individuals to IBD patients. The microbes that survived in the IBD patients' gut environment generally had larger genomes and more complete pathways for synthesizing essential nutrients, such as amino acids (prototrophic). In contrast, microorganisms with smaller genomes and fewer pathways for essential nutrients were less prevalent (auxotrophic). Even though the fecal transplants came from healthy individuals, many beneficial microbes were unable to thrive in the altered gut environment of IBD patients, indicating that changes in the environment, rather than the microbial community itself, may be the driving force behind the shift in composition [40]. This is more reasoned by the fact that auxotrophic bacteria are dependent on their environment which indicates that changes in the environment for example by a reduced availability of essential nutrients might lead to reduced auxotrophies abundances in the human gut microbiome observed in dysbiotic states. Since in Chapter II, it was shown that auxotrophic gut bacteria are positively associated with diversity and long-term stability, the reduction of auxotrophies might lead to an unstable, less diverse microbiome in the large intestine which could further support altered microbial compositions in disease states like IBD as microbial systems with less diversity are less stable. For the dysbotic state of the small intestinal microbiome, it was already discussed in Chapter III that the environment in the duodenum of the dysbiotic state SIBO characterised by overgrowth of bacteria, limited nutrient resources and competition could favor prototrophic bacteria. Further, increased peptidase genes are observed in the microbiome of SIBO patients which might be reasoned by limited resources and increased competition due to the over colonization but could further ameliorate the dysbiotic state and reduced nutrient supply for the host

(Fig. 18). Therefore, focusing on the microbial diet is not only important to potentially restore the dysbiotic state of the human gut microbiome in SIBO subjects but to supply the host with essential nutrients since in SIBO patients hypoproteinemia is seen [242], [243].



**Figure 18: Comparison of the microenvironments of the gut in non-SIBO and SIBO subjects**

In non-SIBO subjects (left figure), auxotrophs (yellow-colored bacteria) are increased, while peptidase genes are decreased in the microbiome of the small intestine. In contrast, in SIBO subjects, auxotrophs are reduced, and peptidase genes are increased (right figure) in the microbiome of the small intestine. In SIBO, bacterial overgrowth might lead to higher amino acid consumption (depicted by flashes) reasoned by an overcolonization of bacteria, which may result in nutrient limitations for both the gut microbiome and the host. As competition between bacteria increases, metabolically independent prototrophic bacteria (blue-colored bacteria) may become more abundant, as they do not rely on external amino acids.

Therefore, the change in the gut environment might come first and might lead to altered microbial compositions in the gut.

Conclusively, the environment is crucial for providing essential nutrients to auxotrophic bacteria and might be changed in disease states thereby favouring prototrophic bacteria. Thus, changes in the gut environment are likely central to shifts in microbial ecology, with significant implications for health and disease. However, influencing factors of the nutritional ecology might be different of auxotrophic bacteria living in the small and large intestine.

In summary, the nutritional ecology of amino acid auxotrophic bacteria differs significantly between the small and large intestine, driven by variations in nutrient availability and environmental conditions. In the small intestine, auxotrophic bacteria predominantly rely on free amino acids derived from dietary protein digestion, facilitated by host enzymes. In contrast, the large intestine represents a nutrient-poorer environment where auxotrophic bacteria primarily depend on cross-feeding

interactions within the microbial community. These metabolic exchanges are crucial for maintaining microbial diversity and stability, particularly in dysbiotic states. The findings underscore the importance of understanding these nutritional dynamics, as disruptions to this balance—whether through reduced nutrient availability or altered microbial composition—can exacerbate conditions such as IBD or SIBO. This highlights the potential for therapeutic strategies that target microbial diets or enhance cross-feeding interactions to restore microbial balance and support host health. Further research into the distinct ecological requirements of auxotrophic bacteria in different gut regions will deepen our understanding of the microbiome's role in health and disease.

#### 4 Auxotrophic bacteria as biomarkers and therapeutic targets to restore the human gut microbiome in dysbiotic states

The change of the microbial composition and functions in diseases is related to the etiology and pathophysiology of diseases [39], [43], [49]. So far, there are several approaches to restore the microbial composition in diseases. Probiotics, prebiotics and their combination in symbiotics are promising therapeutics. Further fecal transplants are studied to restore a dysbiotic microbiome in diseases [244]. This section discusses the therapeutic potential of amino acid auxotrophies in the human gut microbiome.

The results of this thesis indicate that auxotrophies could serve as new biomarkers for gut microbiome stability as they were positively associated with the long-term stability of the human gut microbiome. Another study proved that microbial amino acid biosynthesis is an indicator for gut microbiome fluctuations [245]. Such a biomarker could provide an early indication of ecological changes in the microbiome of the large intestine. Additionally, since auxotrophic abundances were reduced in the small intestinal microbiomes of patients with SIBO compared to a non-dysbiotic state, auxotrophy abundances in the small intestinal microbiome might further indicate ecological changes. However, there are several challenges in the development of biomarkers based on the gut microbiome. The interindividual variability of the gut microbiome must be considered when developing auxotrophies as biomarkers for the human gut microbiome in health and disease [245]. Additionally, further studies are needed to validate auxotrophies as a biomarker in the human gut microbiome and to define thresholds for distinguishing healthy and dysbiotic states on an individual basis. Though analysing and interpreting the results of microbiome studies is expensive and

complex in particular the quantification of biomarkers. Still, a recent survey came to the conclusion that biomarkers based on the human gut microbiome could be promising for example for inflammatory and metabolic diseases [246].

The distribution of auxotrophies could also play a role in restoring gut microbiome ecology in disease states. In fact, network analysis could identify bacteria with central roles in the human gut microbiome. By reconstructing networks (e.g., cross-feeding networks) of the human gut microbiome, bacteria occupying central positions within these networks could be identified. Studying these central positioned bacteria could provide insight into microbial interactions and could give further insights on gut ecology. One recent study utilised a network approach to predict interactions in microbial communities with B-vitamin auxotrophies [247]. Probiotic formulations with these key bacteria, combined with their nutritional requirements (e.g., essential amino acids), could be incorporated into synbiotics to support gut health, particularly when targeting the large intestinal microbiome. Since a study showed no difference in auxotrophic bacteria in the gut by changing dietary B-vitamin concentrations, incorporating amino acids in synbiotics could improve the growth of auxotrophic bacteria particularly when targeting the large intestinal microbiome [73]. This approach may be also relevant for gut dysbiotic states like SIBO, where auxotrophies are reduced. In SIBO, changes in the gut environment, such as limited nutrient availability and increased bacterial competition, might favor prototrophic bacteria, which have a metabolic advantage due to their independence from external amino acids (Fig. 18). Probiotics, instead of antibiotics, are already being discussed as a potential therapy for patients with SIBO. A randomized control trial already showed that probiotics, together with the FODMAP diet, characterised by a reduction of small, fermentable carbohydrates, might have the potential to restore the microbiome of SIBO patients [248]. In addition, a recently study showed promising potential targeting the small intestinal microbiome with fermented milk product. Indeed, changes were observed in the composition and function of the small intestinal microbiota after consumption of fermented milk products, in particular the microbial amino acid metabolism [249]. Further, another study showed promising results when administering probiotics to the small intestinal microbiome [250]. Therefore, a design of a synbiotic consisting of probiotic strains and essential amino acids could be a promising approach. However, the specific effects of such probiotics or synbiotics must still be thoroughly studied to confirm their potential benefits.

Overall, auxotrophic bacteria show significant potential as biomarkers in small and large intestinal microbiomes. Their therapeutic applications through probiotics, prebiotics, and synbiotics could help restore gut microbiome stability and diversity in dysbiotic states. However, these approaches require further experimental validation to fully realize their potential in clinical practice.

## 5 Amino acid auxotrophies in the gut microbiome: putative peptidolysis, SCFA synthesis, and their impact on human health

The human gut microbiome plays a vital role in health and disease through its ability to consume and synthesize metabolites (Chapter I.1). This thesis's results highlighted some metabolic characteristics of amino acid auxotrophic bacteria. Chapter II showed that auxotrophic bacteria are more likely to synthesise some SCFA and other organic acids compared to prototrophic bacteria. In addition, in Chapter III was presented that amino acid auxotrophic bacteria in small and large intestinal microbiomes are correlated with microbial peptidase genes. Therefore this chapter focuses on the potential roles of amino acid auxotrophic bacteria in the gut regarding their putative proteolytic activity, short-chain fatty acid (SCFA) synthesis, and their broader implications for human health and disease.

### Peptidolysis and its health implications

Peptidolytic activity in the gut microbiome influences host health in diverse ways, from alleviating symptoms in conditions like celiac disease through gluten degradation to contributing to inflammatory processes in inflammatory bowel disease IBD [251], [252]. While peptidase genes associated with auxotrophic bacteria remain underexplored, analyses in Chapter III revealed significant associations between these bacteria and peptidases involved in critical processes, such as glutathione degradation and antibiotic resistance.

In Chapter III, auxotrophic bacteria in the small intestinal microbiome were negatively associated with many antibiotic resistance genes when investigating the peptides gene profiles, in particular penicillin-binding proteins (S11, S13 peptidase family). The ubiquitous distribution of antibiotic resistance genes has negative consequences for the successful administration of antibiotics. Several studies showed that antibiotic resistance genes are present in the human gut microbiome [253], [254]. In several disease, the presence of antibiotic resistance genes in the human gut microbiome was associated with the pathophysiology [255], [256]. Interestingly, an experimental study



showed that nutrient competition for proline in the human gut microbiome is able to suppress the growth of an antibiotic resistant *E.coli* strain [257]. Since auxotrophic bacteria are negatively associated with peptidase genes involved in antibiotic resistance, they could have protective effects on the distribution of antibiotic resistance genes in the human gut microbiome by acquiring essential amino acids and competing with other bacteria. Therefore, the decrease in auxotrophic bacteria could have detrimental effects in SIBO and could potentially enable colonization of antibiotic resistant bacteria. Indeed, in the microbiome of SIBO patients with malabsorptive syndrome, antibiotic resistance genes were observed [258] and discussed in treatment options [259], [260]. In sum, auxotrophic bacteria in the small intestinal microbiome might positively impact the presence of antibiotic resistance genes in the human gut microbiome which could have implications in SIBO when they are reduced.

Interestingly, fewer associations between auxotrophic bacteria and peptidase genes involved in antibiotic resistance were observed in the large intestinal microbiome (Fig. 13). This contrasts with findings in the small intestine, where auxotrophic bacteria were negatively associated with many antibiotic resistance genes, particularly penicillin-binding proteins. One possible explanation for this difference lies in the nutritional environment of the large intestine, which is more stable compared to the small intestine. In the small intestine in contrast to the large intestine, nutrient fluxes are higher due to incoming enzymatic fluids and dietary digests [12]. This stability likely reduces competitive pressures for essential amino acids, which may explain why auxotrophic bacteria are less strongly associated with antibiotic resistance genes in this region. In addition, the large intestine contains a dense and diverse microbial community. This diversity allows bacteria to interact in more complex ways, such as sharing resources like amino acids. As a result, auxotrophic bacteria in the large intestinal microbiome may rely on these shared resources rather than directly competing for them (Chapter V 3). Indeed, a study showed that the microbiome's diversity might protect against the distribution of antibiotic resistance genes [261]. This interconnected environment might make it harder to detect a clear statistical link between auxotrophic bacteria and antibiotic resistance genes in the large intestine. These factors—the stable nutrient environment and the complex microbial interactions—may together explain the weaker association observed in this region. Additionally, auxotrophic bacteria could directly influence gut homeostasis in the small intestine. Negative associations with several amino acid auxotrophies were observed

for peptidase gene (T03.001) of gamma-glutamyltransferase in the duodenum. The microbial peptidase can cleave the glutamyl bonds in glutathione [262]. Glutathione is a regulator of oxidative stress in the gut. Oxidative stress is a characteristic of inflammatory conditions in IBD; reduced glutathione levels were observed [263]. Since auxotrophic bacteria are negatively associated with peptides genes involved in glutathione degradation, they could have protective effects on oxidative stress in the gut.

#### SCFA synthesis in gut dysbiosis states

In Chapter II, several auxotrophic bacteria were associated with the synthesis of organic acids such as lactate, butyrate, and succinate. SCFA and other organic acids are essential for maintaining gut health and homeostasis. The quantity of SCFA in the gut is associated with diseases and treatment with SCFA improved inflammatory conditions in the gut [264], [265]. Many studies investigating the metabolic profile of SCFA observed increased or decreased potential for the synthesis of SCFA by the human gut microbiome. For example, butyrate producing bacteria were decreased in the gut microbiome of patients with IBD [233].

In the small intestine, several organic acids are found like lactate, butyrate and formate which are associated with auxotrophic bacteria [25]. The SCFA profile in the microbiome of patients with SIBO remains understudied, likely due to the limited availability of whole genome sequence data. However, one study measured the concentrations of organic acids in the jejunal fluids of patients with SIBO and healthy controls, finding that the distribution of organic acids in SIBO patients resembled that typically found in feces. [266]. Results from Chapter III showed that the microbiome of patients with SIBO has a lower abundance of auxotrophic bacteria, similar to observations in the large intestinal microbiome. This similarity suggests that the microbiome shifts in SIBO, resembling those in the large intestine, may be linked to altered SCFA profiles, which could impact gut homeostasis and disease progression. Auxotrophic bacteria, which are associated with SCFA synthesis, might influence SCFA levels, and their reduced abundance in SIBO (Chapter III) could therefore disrupt gut homeostasis. In conclusion, peptidase genes and the synthesis of SCFA associated with amino acid auxotrophic bacteria provide critical insights into their potential roles in gut homeostasis. By influencing oxidative stress, SCFA production, and antibiotic resistance, auxotrophic bacteria may significantly impact gut health and disease states.

## 6 Limitation of *in silico* prediction of auxotrophies

In this thesis, auxotrophies are predicted via genome-scale metabolic modeling. Discrepancies are known for auxotrophies predicted with genome-scale metabolic Modeling when compared to experimentally verified auxotrophies [62], [68]. Still, in Chapters II and IV, the high reliability of auxotrophies predicted with metabolic modeling was shown. In addition, auxotrophy predictions with metabolic modeling were compared to another *in silico* method. In the following, challenges and limitations of the prediction of auxotrophies with *in silico* methods will be discussed. Additionally, an outlook for future methodological approaches are given to further improve validation of amino acid auxotrophies by genome-scale metabolic modeling.

In Chapter II, genome-scale metabolic modeling was validated and showed an accuracy of >90%. Further, in Chapter IV, another *in silico* method, called GapMind, for auxotrophic prediction based on genome annotation was evaluated and showed an accuracy of >90%. When comparing both *in silico* auxotrophy prediction methods, it can be said that gapseq reconstructed models overpredict auxotrophies and GapMind underpredict auxotrophies (Chapter IV). Both methods rely on different principles. GapMind predicts auxotrophies via genome annotation and states the presence/absence of a reaction. In particular, the percentage of absent reactions in a biosynthesis pathway is stated as an auxotrophy based on experimental data. Gapseq automatically reconstructs genome-scale metabolic models, and auxotrophies are predicted with flux-balance analysis.

Though the accuracy of GapMind and metabolic modeling was >90%, some discrepancies were observed compared to experimentally verified auxotrophies. There are several reasons for this. First, there are still some unknown gaps in amino acid pathways. Not all reactions have yet been discovered [226]. Some genes are not especially known for sulfate-reducing bacteria. Thus, for the production of serine, for example, some genes are unknown, and inconsistencies in auxotrophy predictions were found [173]. Second, the organisms in the database that determine the sequence alignment to the query genome considerably impact the prediction of the absence/presence of genes [114]. Ideally, the reactions of the pathways in the database are experimentally characterized. Gapseq and GapMind involve experimentally verified reactions in their databases [111], [114]. Third, the quality of the genomes has an influence on *in silico* auxotrophy predictions. The incompleteness

of genomes can lead to an over-prediction of auxotrophies (Figure S1). The genomes used for model reconstruction were filtered for a completeness of 85% (Chapters II, III) to avoid an overprediction of auxotrophies. One limitation of predicting auxotrophies using genome-scale metabolic modeling or experimental validation *in vitro* is that precursor amino acids for a biosynthetic pathway may be present in the growth medium. Since only one amino acid is typically removed at a time, these precursors can affect the accuracy of auxotrophy predictions. For example, L-serine is a precursor for L-tryptophan biosynthesis. If L-serine is available in the medium, tryptophan biosynthesis may still occur as long as the gene for the enzyme in the final step of the pathway is present. This could lead to a false prediction of tryptophan prototrophy (independence from tryptophan), even though the organism is actually auxotrophic for tryptophan.

There are several methodological principles in metabolic modeling that play a key role in influencing the accuracy of auxotrophy predictions. One key aspect is how nutrient utilization is predicted. Typically, a flux is set through the metabolic network based on an objective function, which, in the case of auxotrophy predictions, is often biomass formation [120]. This is a standard approach in metabolic modeling. However, bacterial cells also invest in yield not only in growth. Whether a bacterial cell increases yield or growth is dependent on growth conditions [267].

Another critical step is gap-filling during the reconstruction of genome-scale metabolic models. One study criticised gap-filling in automatically reconstructed metabolic models and emphasised the media compositions as being a crucial step [119]. The growth medium should be individually chosen for each metabolic model and relies on computational and experimental approaches to reduce inconsistencies [119]. For example, gapseq employs a gap-filling algorithm that predicts a suitable medium based on homology searches and mathematical optimization [111]. Metabolites are added to the gap-filling medium based on reactions identified in the draft metabolic networks (see Box 3 in Chapter I 3.1). These metabolites include compounds commonly found in the gut environment (see Supplementary Information for Chapter II). This individualized, algorithmic approach ensures that the growth medium is optimized to improve the accuracy of metabolic model reconstructions. In conclusion, several methodological aspects could contribute to the observed inconsistencies in auxotrophy predictions made with genome-scale metabolic modeling.

There are also methodological approaches that could help improve the validation of these predictions. Currently, auxotrophy predictions rely on genome-scale metabolic models built primarily from genome sequences without incorporating transcriptomic data. This means they lack information about actual biological activity. By integrating additional omics data, such as transcriptomic and proteomic data, gene expression could be incorporated, potentially improving the validation of these models. For example, a recent study used metatranscriptomic data to map biological activity onto genomes in a community-scale approach [268]. Using this method, the authors were able to predict amino acid auxotrophies and cross-feeding of amino acids. The findings in Chapter II indicate that cross-feeding of essential nutrients between auxotrophic and prototrophic gut species is directly linked to diversity and stability. This highlights the value of community modeling in predicting ecological interactions within microbial ecosystems. This suggests that integrating transcriptomic data, combined with a community-based approach, could improve predictions by accounting for both biological activity (via gene expression) and ecological interactions such as cross-feeding [268]. Furthermore, integrating metabolomic data could further validate the metabolic activity modeled by genome-scale metabolic networks. One study integrated genome-scale metabolic models with metabolomics and transcriptomics in a community model to study cross-feeding in a phototrophic community [269]. They experimentally validated the exchange of metabolites by the analysis of metabolomic data and RNA-sequences. Further studies could improve the validation of amino acid auxotrophies by adding transcriptomic data from human gut microbiome. In addition, to prove the exchange of amino acids, which was a positive link to diversity and stability, metabolomic data from human gut fluids could be integrated. For example, for validating an auxotrophy for tryptophan predicted by genome-scale metabolic modeling, transcriptomic data could state that genes for transports are increased and metabolomics data could show the presence of tryptophan in the gut environment of the bacterial species. .

Despite the mentioned limitations, this study has demonstrated that genome-scale metabolic modeling) is a precise method for predicting auxotrophies. Future research should focus on integrating genomic data with multi-omics approaches to enhance the biological relevance of predictions. The development of dynamic *in silico* models based on metabolic networks through community modelling and temporally resolved

omics data could further improve prediction accuracy and provide deeper insights into the microbial ecology of the gut.

## Chapter VI: Conclusion

The ecological characteristics of the human gut microbiome are linked to the pathophysiology of inflammatory and metabolic diseases. However, despite extensive research, our understanding of the complex interactions between gut microbes remains limited. Studying the ecology of the human gut microbiome and the factors that influence it could provide crucial insights into the microbe-host relationship. Auxotrophies play a significant role in shaping microbial ecosystems, yet their distribution and ecological impact within the human gut microbiome have not been thoroughly explored.

This thesis is the first to show the distribution of amino acid auxotrophies in human gut bacteria and the influence of auxotrophies on the human gut microbiome. Results showed that amino acid auxotrophies are common in human gut bacteria, with small intestinal bacteria having higher auxotrophies than large intestinal bacteria. However, in the dysbiotic state SIBO, auxotrophy abundances are reduced. These findings offer new insights into the nutritional dependencies of gut bacteria and open new ideas for manipulating microbiome composition to promote health.

Further, amino acid auxotrophies positively influence the microbiome diversity and stability, critical factors in maintaining gut homeostasis. Microbial communities with more cross-feeding were more diverse and stable over time. Given the association of many diseases with microbial imbalances and reduced diversity, these results lay the groundwork for future studies to restore dysbiotic microbiomes through targeted auxotrophy modulation. Additionally, auxotrophies hold potential as biomarkers for early detection of gut microbiome changes, with clinical applications in both diagnostics and therapeutics.

So far, auxotrophy prediction with genome-scale metabolic modeling has been criticized as unreliable. Thus, the validation of metabolic modeling and GapMind with an accuracy >90% is a massive statement for the reliability of auxotrophy prediction with *in silico* methods.

In summary, the results gained from this thesis not only advance our understanding of the human gut microbiome's ecological dynamics but also offer promising directions for developing microbiome-targeted interventions to improve human health. Future research can build upon these findings to explore the therapeutic potential of auxotrophies in restoring microbial balance and enhancing microbiome resilience in disease contexts.

## Acknowledgements

I would like to express my deepest gratitude to my first supervisor, Jun.-Prof. Silvio Waschina. Thank you for believing in me and for your continuous support throughout my PhD. I truly appreciate you always taking the time for our weekly meetings and my endless questions! I have learned so much, not only about my PhD topic but also about myself.

A special thank you to Dr. Danielle Harris for her valuable advice and inspiring ideas. Our meetings have always been a great source of motivation for me!

I also want to thank my Nutriinformatics research group – Alina Pisarevskaja, Nicola da Bernadini, and Sina Zimmermann – for their support and helpful suggestions. Your comments and discussions during our group meetings have been invaluable in preparing my manuscripts. A heartfelt thanks also to Sina Zimmermann – the best office colleague and a great listener.

I would also like to acknowledge our collaboration partners, Prof. Dr. Andre Franke, Dr. Corinna Bang, Prof. Dr. Konrad Aden, and Prof. Dr. Philip Rosenstiel.

My deepest appreciation goes to my husband, without whose support the successful completion of my PhD would not have been possible. Thank you to my parents, Susanne and Uwe Busche, for their constant encouragement and belief in me. A special thank you to my sister, Annika Busche, for always being there for me and for her unwavering support. Your encouragement and positivity have meant the world to me!



Supplementary information for Chapter II

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

## Supplementary Material and Methods

### Reconstruction of genome-scale metabolic models

Genome-scale metabolic models of prokaryotic genomes were reconstructed using *gapseq* [111]. In brief, the *gapseq* reconstruction workflow consisted of five steps: (i) Reaction and pathway prediction, (ii) prediction of metabolite cross-membrane transporters, (iii) reconstruction of a draft metabolic network based on the results from *i* and *ii*, (iv) estimation of an organism-specific growth medium-based on the predicted metabolic capabilities, and (v) gap filling of the metabolic network to enable biomass production using flux balance analysis. Model reconstructions were limited to bacterial genomes marked as representative species in the HRGM collection. Further, genomes with an estimated contamination percentage of  $\leq 2\%$  or a completion  $\geq 85\%$  were included. Based on these filters, 3 687 bacterial genomes were subject to metabolic model reconstruction. Among those genomes, 22% are from bacterial isolates, and 78% are metagenome-assembled genomes.

A recent computational study has shown in a systematic analysis of isolate- and metagenome-assembled genomes that the gap-filling medium strongly impacts the auxotrophies predicted by genome-scale metabolic modelling [119]. We note that with the reconstruction procedure used in this study, we do not rely on an arbitrarily defined gap-filling medium, which is used for every metabolic network model. Instead, for each draft network, a genome-specific gap-filling medium is predicted (see section “Prediction of a genome-specific gap-filling medium” below for details). In brief, if the medium prediction algorithm of *gapseq* finds a known biosynthetic pathway for a specific amino acid in the draft metabolic network, the amino acid will not be part of the resulting predicted medium, as the compound is likely not required from the growth environment, thus, also not necessary for subsequent gap filling. In contrast, if the medium prediction algorithm does not detect at least one complete known biosynthetic

pathway for a specific amino acid, the respective compound is added to the outcome gap-filling medium based on the rationale that this amino is a putative essential compound that needs to be obtained from the growth environment. However, it is important to note that this case does not directly imply that the organism is auxotrophic for the specific amino acid, as the subsequent gap-filling algorithm might add reactions to the model that complete a biosynthesis route from other compounds in the gap-filling medium to the amino acid. Such reactions are only added in cases where a gene was found in the query genome that displays sequence similarity to a reference gene sequence with the respective enzymatic function but where the sequence similarity was not high enough to pass the threshold of the bitscore 200 to be directly included in the draft network. For details on the gap-filling algorithm implemented in *gapseq*, please refer to the original *gapseq* publication [111].

Taken together, the model reconstruction and auxotrophy prediction that we used for the present study do not depend on one gap-filling medium composition that is defined for all organisms but adjusts the medium for each organism based on its genome information and by using the multi-step gap-filling algorithm that is implemented in *gapseq*.

### Prediction of a genome-specific gap-filling medium

The genome-scale metabolic network reconstruction process of *gapseq* requires a gap-filling medium for the final gap-filling step (see above). Here, we used a medium prediction feature (module “*gapseq medium*”) of the *gapseq* software, which can be plugged into the reconstruction workflow between the homology-based generation of a draft metabolic network and the gap-filling algorithm. We provided the additional command line option “-c cpd00007:0” to ensure that the predicted medium does not contain oxygen (compound identifier: cpd00007). The algorithm for medium prediction

tests which pathways and reactions are absent or present in the draft metabolic model. Whether a particular compound is added to the medium is decided using logical expressions that include variables for the presence (TRUE) and absence (FALSE) of pathways and reactions within the design network (see Supplementary Table S6 for all compounds and their logical expressions). For example, the disaccharide lactose (ModelSEED ID: cpd00208), has the logical expression ("LACTOSECAT-PWY" | "LACTOSEUTIL-PWY" | "BGALACT-PWY"), which means, that lactose is added to the medium if one of three known lactose degradation pathways as defined in MetaCyc[148] is already present in the draft network. In particular, for amino acids, the medium prediction module uses a similar approach to the auxotrophy prediction tool GapMind[114], which tests if a known biosynthetic pathway for a specific amino acids exists based on sequence homology. The amino acid biosynthetic pathways that are considered are also those that are defined in MetaCyc[148]. For instance, if none of the five known biosynthesis pathways in prokaryotes for lysine (<https://metacyc.org/META/NEW-IMAGE?type=PATHWAY&object=LYSINE-SYN>) is found, lysine is added to the gap-filling medium.

The medium prediction module of gapseq considered 74 compounds (Supplementary Table S6), including inorganic compounds, carbohydrates, amino acids, other carboxylic acids, and vitamins. Most of those potential nutrients are also compounds that can be found in the growth environment of colonic microorganisms, such as fibers (e.g., pectin, inulin), other dietary compounds (e.g., sulfoquinovose, daidzein), constituents of the mucins (e.g., *N*-acetylneuraminate, *N*-acetylneuraminate) and inorganics (e.g., H<sub>2</sub>, H<sub>2</sub>S, H<sub>2</sub>O).

Besides the enumeration of available nutrients, a gap-filling medium requires their individual maximum uptake rates by the microorganism. The medium prediction implemented in gapseq uses rates commonly used in manually curated genome-scale

metabolic network models, e.g.,  $0.1 \text{ mmol} \cdot \text{gDW}^{-1} \cdot \text{hr}^{-1}$  for amino acids or  $5 \text{ mmol} \cdot \text{gDW}^{-1} \cdot \text{hr}^{-1}$  for monosaccharides. In the case of oligo- and polysaccharides, the maximum uptake rates are scaled to allow the same uptake rate per subunit (e.g.,  $5 \text{ mmol} \cdot \text{gDW}^{-1} \cdot \text{hr}^{-1}$  for the monosaccharide D-glucose and  $2.5 \text{ mmol} \cdot \text{gDW}^{-1} \cdot \text{hr}^{-1}$  for the disaccharide maltose).

### Validation of auxotrophy predictions

To validate our *in silico* predicted auxotrophies, we collected the genome sequences from NCBI RefSeq for 36 bacterial strains for which experimental data were available on amino acid auxotrophies/prototrophies. The majority of these strains were already summarized previously[79]. In that study, the authors even list more than the 36 strains we analyzed here. This is because we excluded cases where we could not find a genome assembly of the exact strain used in the referenced experimental study. Moreover, we excluded the entries for species belonging to the genus *Bifidobacterium* since their auxotrophy for cysteine is ambiguous: Some studies report cysteine as an essential nutrient for growth[270], [271], while genomic analysis and genome-scale metabolic modeling indicated the presence of cysteine biosynthetic pathways[272], [273]. This ambiguity most likely stems from the fact that cysteine biosynthesis in *Bifidobacterium* species depends on the available sulfur source[272]. Genomic analysis of *Bifidobacterium bifidum* PRL2010, for instance, suggested the strain's inability to use sulfate as a sulfur source, while hydrogen sulfide or methionine could potentially serve as a sulfur source for the biosynthesis of cysteine[270], [272].

Genome-scale metabolic models for all 36 strains were reconstructed as described above. Auxotrophies were predicted with the method described in the main manuscript. In addition to the gapseq-reconstructed models, we also predicted auxotrophies for 20 of the 36 bacterial strains using genome-scale models from the AGORA2

collection[113] (Supplementary Table S2). Auxotrophies were predicted in the same manner as for *gapseq* models. In contrast to the *gapseq* models, which can contain only free amino acids and not peptides in the predicted medium, some AGORA2 models have exchange reactions with lower bounds  $< 0$  for dipeptides. In those cases, and for predicting the auxotrophy status for amino acid *x*, we changed the lower bound to 0 for all exchange reactions of dipeptides involving amino acid *x*. At the same time, we introduced a new inflow reaction of the non-*x* amino acid moiety to the model to predict only the essentiality of *x* and not of the other amino acids in the respective peptides. The sensitivity, specificity, and accuracy of auxotrophy predictions were calculated for *gapseq* models ( $n=36$ ) and AGORA2 models ( $n=20$ ).

As an additional auxotrophy prediction validation step, 124 genome-scale metabolic models were reconstructed for bacterial strains that were reported by Price, 2023, to be able to grow in a defined growth medium containing no amino acids[147]. Thus, these 124 organisms are known amino acid prototrophs and can be used to estimate the rate of false auxotrophy predictions. The original publication by Price reported 127 genomes of prototrophs; however, 3 of the corresponding genome assemblies (GCF\_000014265.1, GCF\_000020545.1, GCF\_900188395.1) were suppressed on RefSeq at the time we performed the analysis in March 2023. Auxotrophies for the 124 genome assemblies of this prototroph collection were predicted as described above, and results are summarized in Supplementary Table S3.

## Metagenome data processing

Metagenomic reads were subject to quality control and filtering using the ‘qc’ workflow from the metagenome-atlas pipeline tool v2.9.0[173]. In detail, reads were (i) deduplicated, (ii) quality filtered, and (iii) decontaminated. Modules from the BBmap suite v37.99 (BBMap - Bushnell B. - [sourceforge.net/projects/bbmap/](https://sourceforge.net/projects/bbmap/)) were used for

all three steps. In the deduplication step (i), the BBmap module *clumpify.sh* was used with the parameters “dedupe=t dupesubs=2”, which removed duplicate reads with a maximum of 2 substitutions between duplicates. The quality filter (ii) employed the BBmap module *bbduk.sh* with the parameters “hdist=1 ktrim=r mink=8 trimq=10 qtrim=rl minlength=51 maxns=-1 minbasefrequency=0.05” and otherwise default options. This quality filter trimmed reads from the right if adapter sequences were detected, trimmed reads on both sides from the first base with a quality score below 10, removed sequences that were shorter than 51 bp after trimming, removed sequences with ambiguous base calls (i.e., “N”s), and removed reads if any base had a frequency of less than 5%. Finally, reads that are likely contaminations from the human host genome or Illumina PhiX sequences were removed using the BBmap module *bbsplit.sh* using the option “maxratio=0.65” and otherwise default parametrization. This tool tested if specific reads mapped to the host genome or PhiX sequences based on sequence similarity and mapped reads were discarded from the sample’s fastq files. For the decontamination step (iii), the human reference genome assembly ‘Genome Reference Consortium Human Build 38’ (GRCh38) was used, in which low entropy regions (entropy < 0.7) were masked using the *bbmask.sh* tool within the BBmap suite. Moreover, regions that display high similarity to prokaryotic rRNA genes were additionally masked. To this end, prokaryotic small and large subunit rRNA gene sequences were retrieved from SILVA version 138.1[274] and shredded into shorter (80 bp) sequences with 40 bp overlaps using *shred.sh*. Shredded sequences were aligned to GRCh38 with a minimum identity of 85% and maximum indel length of 2 bp. Regions in GRCh38 with alignment hits were masked.

As mentioned in the main manuscript, we used the Human Reference Gut Microbiome ‘HRGM’ catalog[146] as reference genomes for quantifying representative microbial genomes in the metagenomic data sets. We had chosen this collection, as it was the

latest published human gut microorganism genome collection when we were finalizing the results of the present study. Furthermore, the HRGM collection contains 780 species-level representative genomes, which were absent in previous genome collections and assembled from metagenome samples from before under-represented Asian countries, namely Korea, Japan, and India. For each metagenome sample, the relative abundance of HRGM genomes was estimated using coverM[174] v0.6.1 with default parametrization of the module `coverm genome`.

### Targeted metabolomics of blood samples

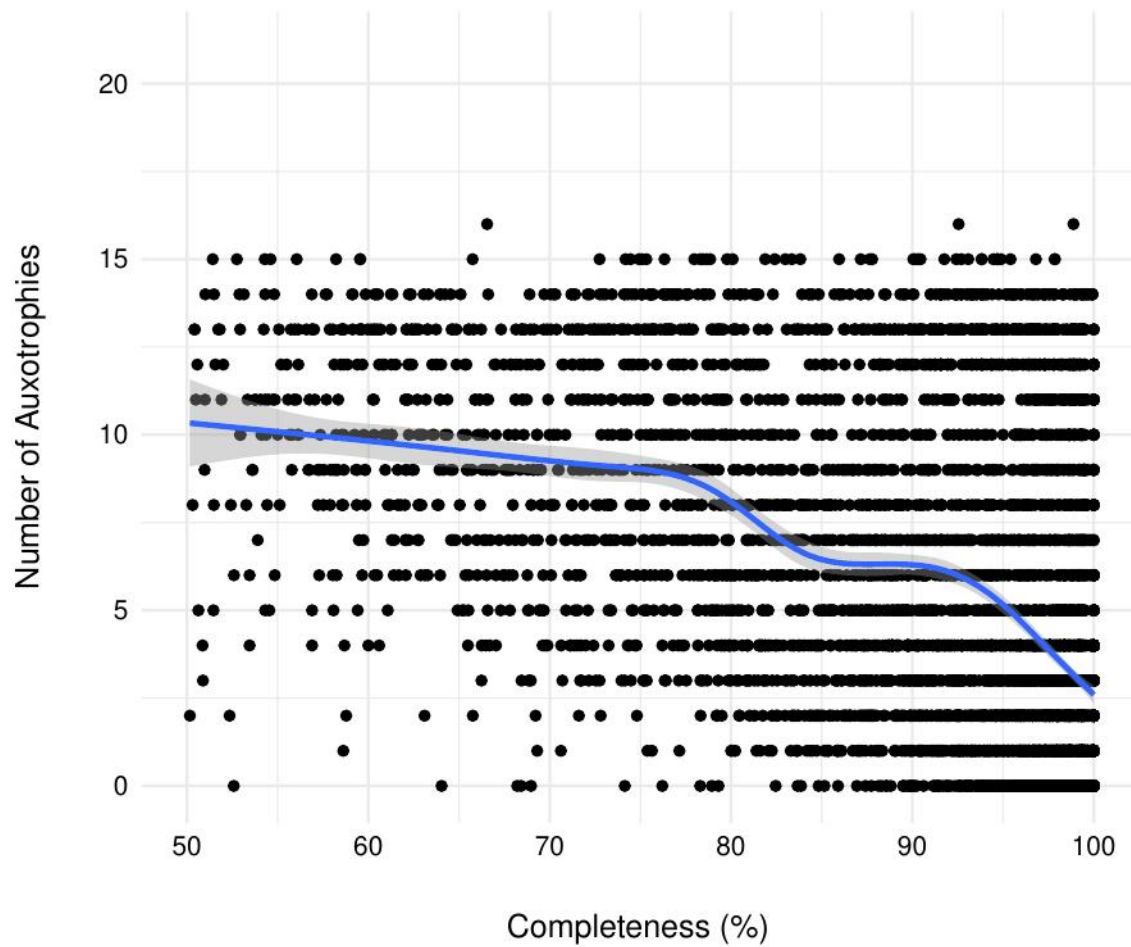
Serum samples were collected using serum s-monovette (9ml, Sarstedt, Germany). Samples were incubated upright at RT for 30 min. and centrifuged (10 min., 2000 x g). Serum was aliquoted in 500µl tubes and stored at -80°C. Metabolite quantification for serum was performed by liquid chromatography tandem mass spectrometry (LC-MS-MS) using the MxP Quant 500 kit (Biocrates Life Sciences AG, Innsbruck, Austria) according to the manufacturer's instructions. The MxP Quant 500 kit simultaneously measures 630 metabolites covering 14 small molecule and 12 different lipid classes. It combines flow injection analysis tandem mass spectrometry (FIA-MS/MS) using SCIEX 5500 QTrap mass spectrometer (SCIEX, Darmstadt, Germany) for lipids and liquid chromatography tandem mass spectrometry (LC-MS/MS) using Agilent 1290 Infinity II liquid chromatography (Santa Clara, CA, USA) coupled with a SCIEX 5500 QTrap mass spectrometer for small molecules using multiple reaction monitoring (MRM) to detect the analytes. Data evaluation for serum metabolite concentrations and quality assessment was performed with the software SCIEX Analyst software (Version 1.7.2) and the MetIDQ™ software package (Oxygen-DB110-3023), which is an integral part of the MxP Quant 500 kit.



For downstream statistical analysis, the serum metabolome data were pre-processed by imputing missing specific values using a random forests approach as implemented in the R-package 'missForest' and the function with the same name in default parametrization [275]. This imputation was limited to missing values for metabolites, which have less than 20% missing values across the data set.

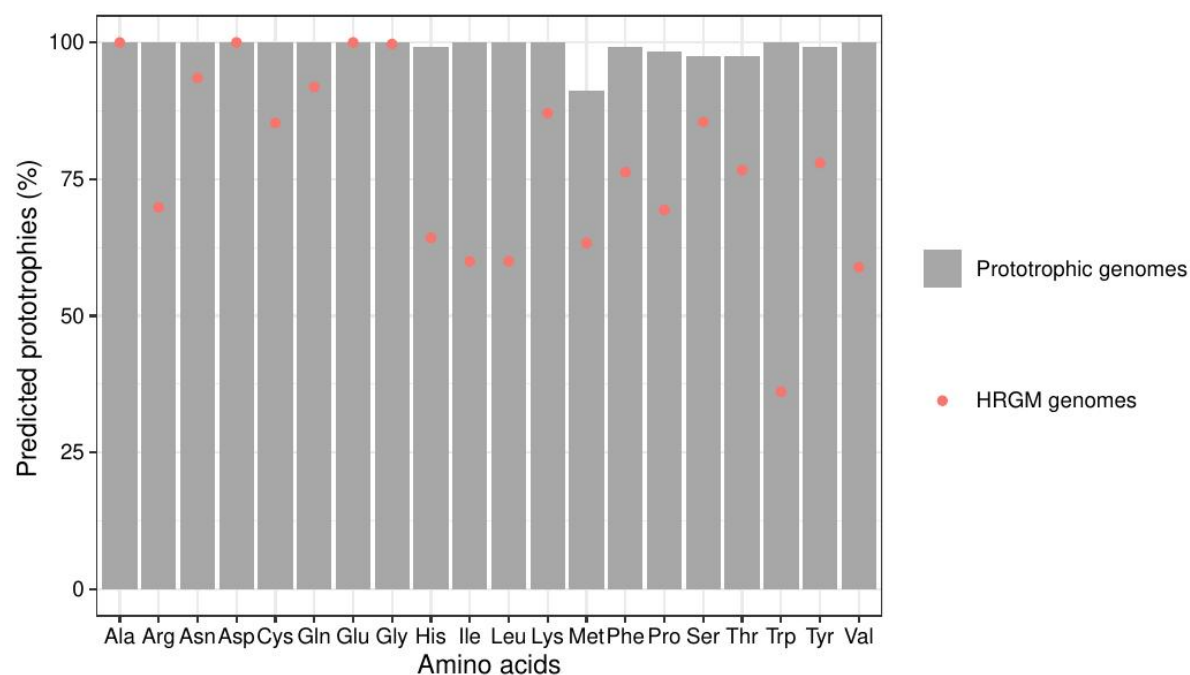
With a partial Spearman correlation, the association between the frequency of auxotrophic bacteria and serum metabolites and other hematology parameters from the DZHK cohort was evaluated [276]. We adjusted for the potential confounders sex, age, and BMI.  $p$  values were corrected for multiple testing using the False Discovery Rate (FDR) method.

## Supplementary Figures



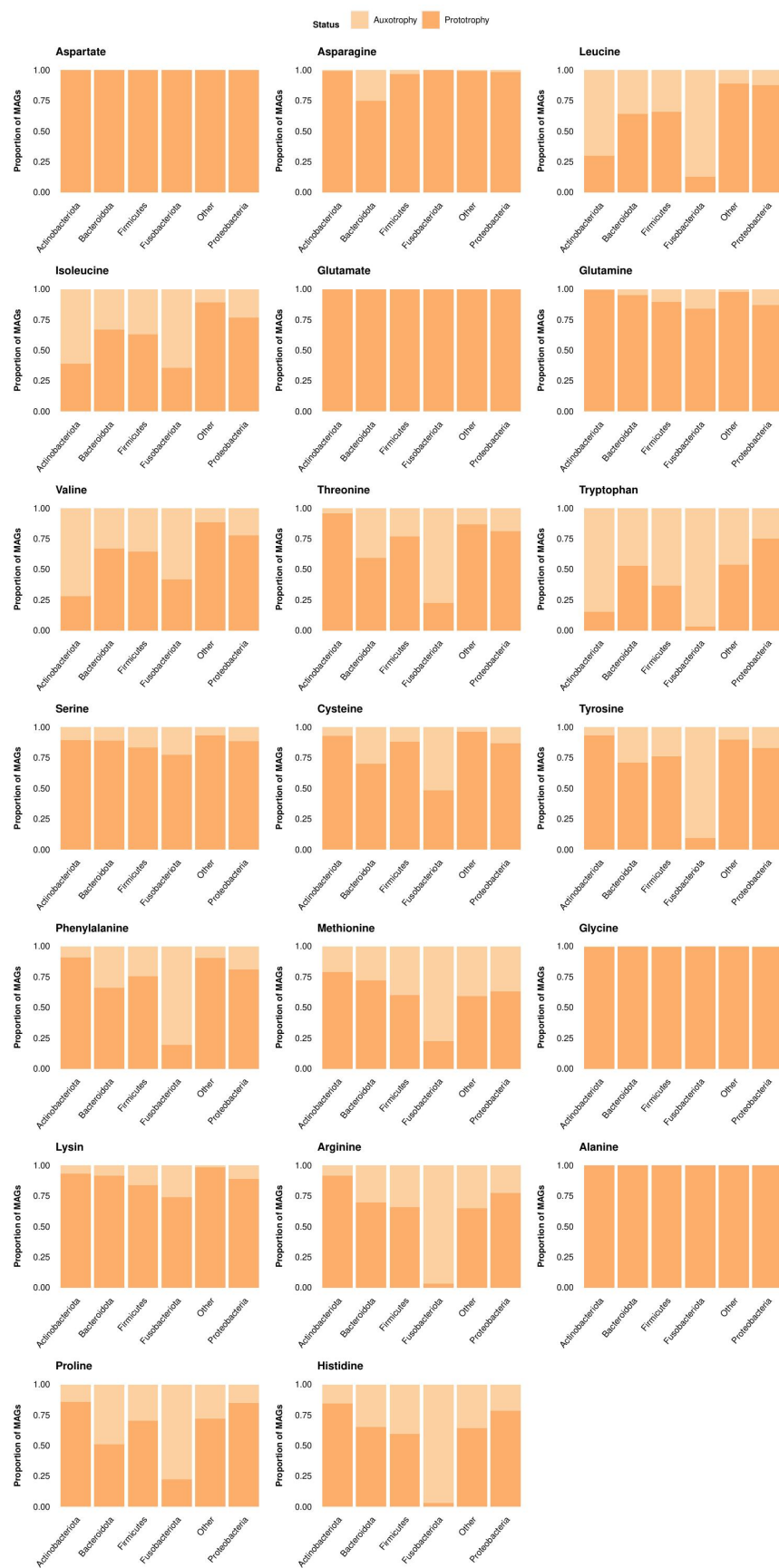
**Supplementary Figure 1: Association between the number of auxotrophies and the genome completeness**

Estimated genome completeness and the predicted number of auxotrophies for 3 687 genomes of representative species from the Human Reference Gut Microbiome (HRGM) collection [14]. The blue line shows the regression line ( $\rho = -0.50$ ,  $p \leq 2.2e-16$ ).

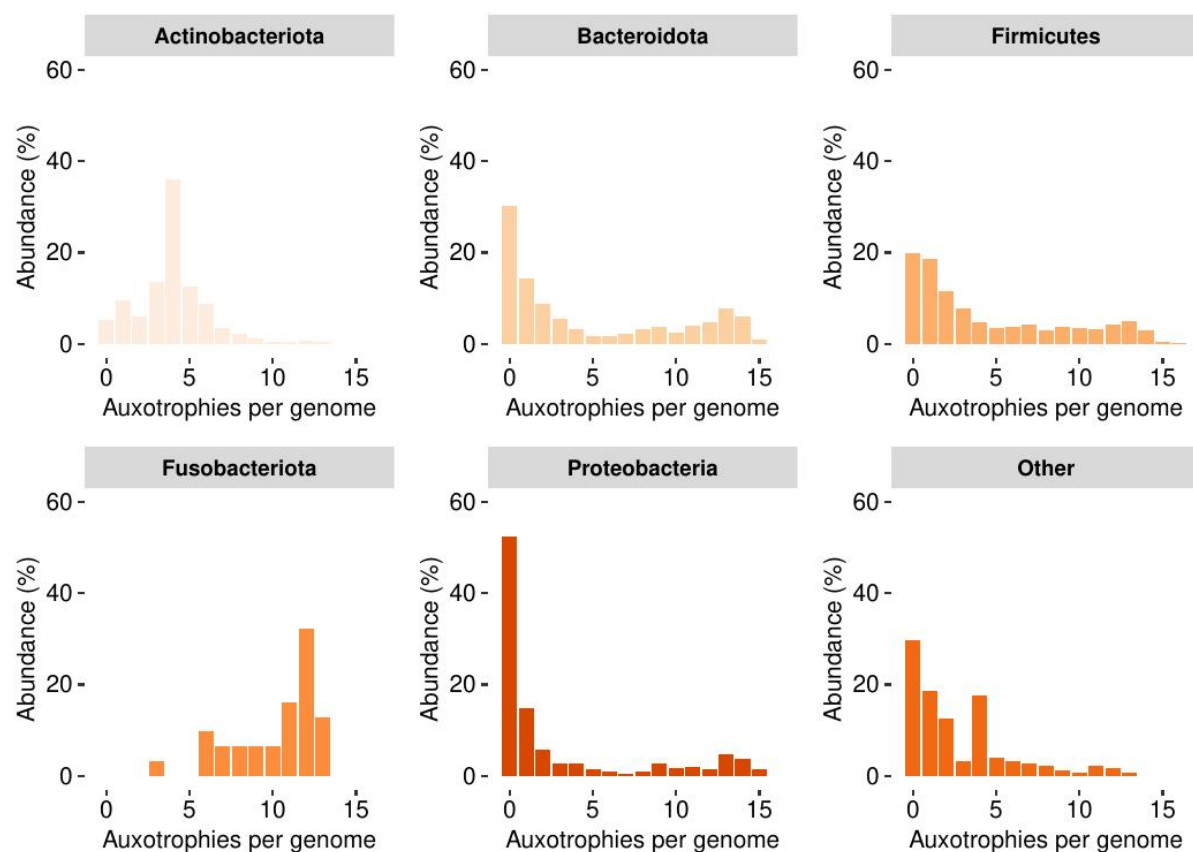


**Supplementary Figure 2: Validation of in silico predicted auxotrophies with experimental auxotrophies**

Percentage of in silico predicted prototrophies with metabolic modeling in 124 genomes known to be prototrophic [11] from laboratory experiments (grey bars). The red dots indicate the frequency of prototrophies among 3 687 genomes from human gut bacteria [14].

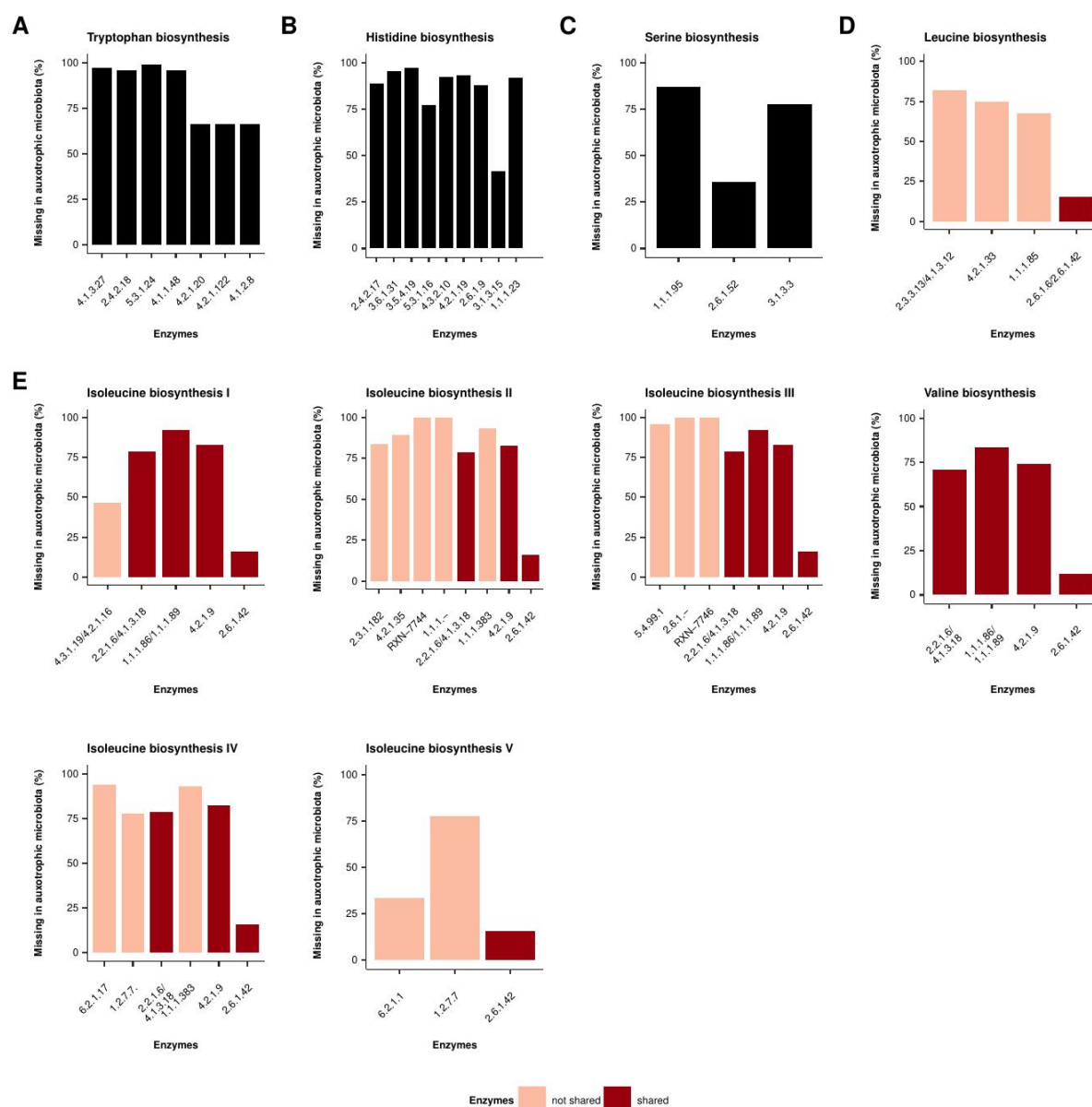


**Supplementary Figure 3: Overview of the proportions of auxotrophy to prototrophy genomes per phylum from the HRGM catalog [146]**



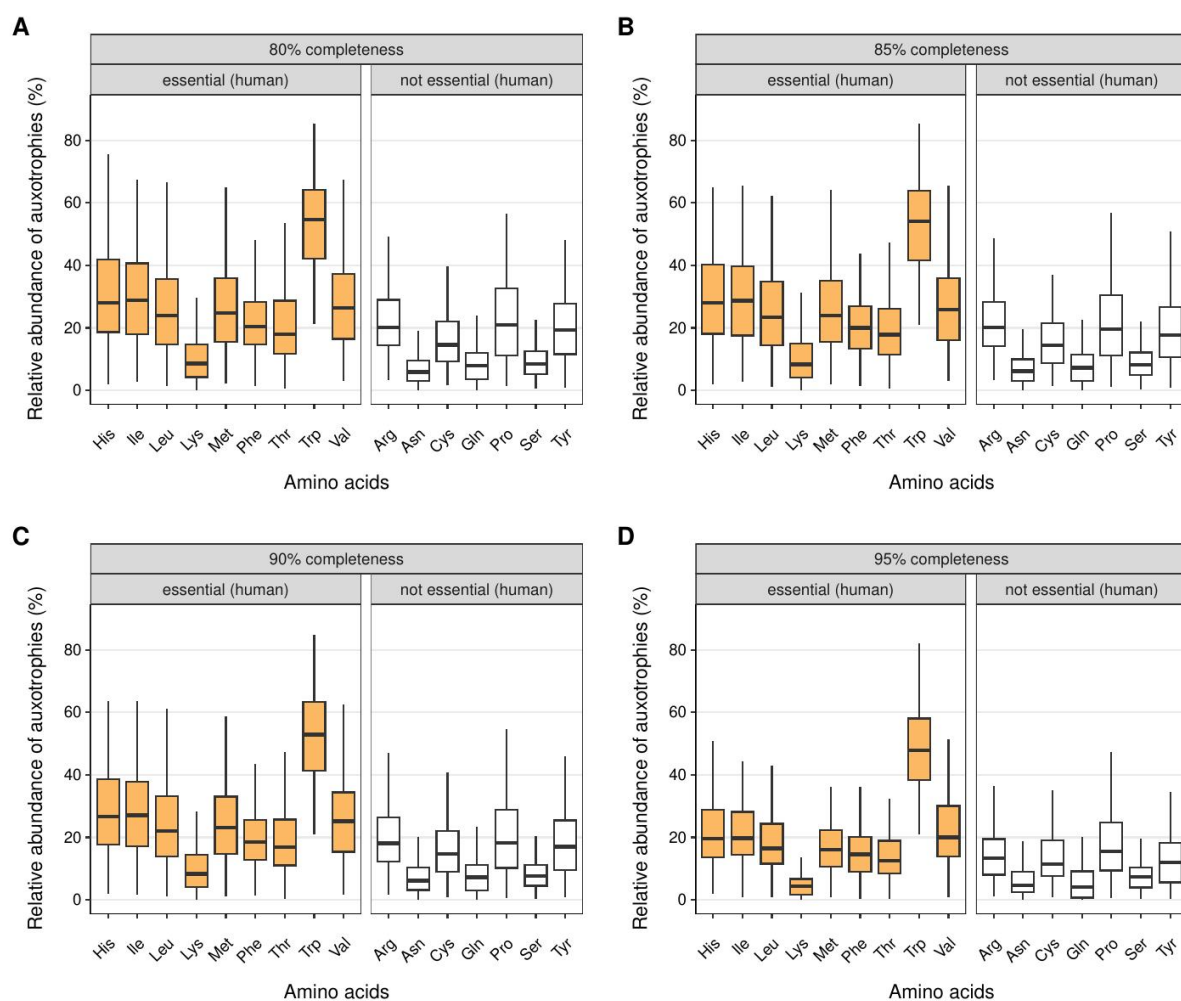
**Supplementary Figure 4: Number of auxotrophies for every phylum**

Other is a category that combines different phyla with a lower abundance in the overall HRGM catalogue [14] and for a reduction of complexity.



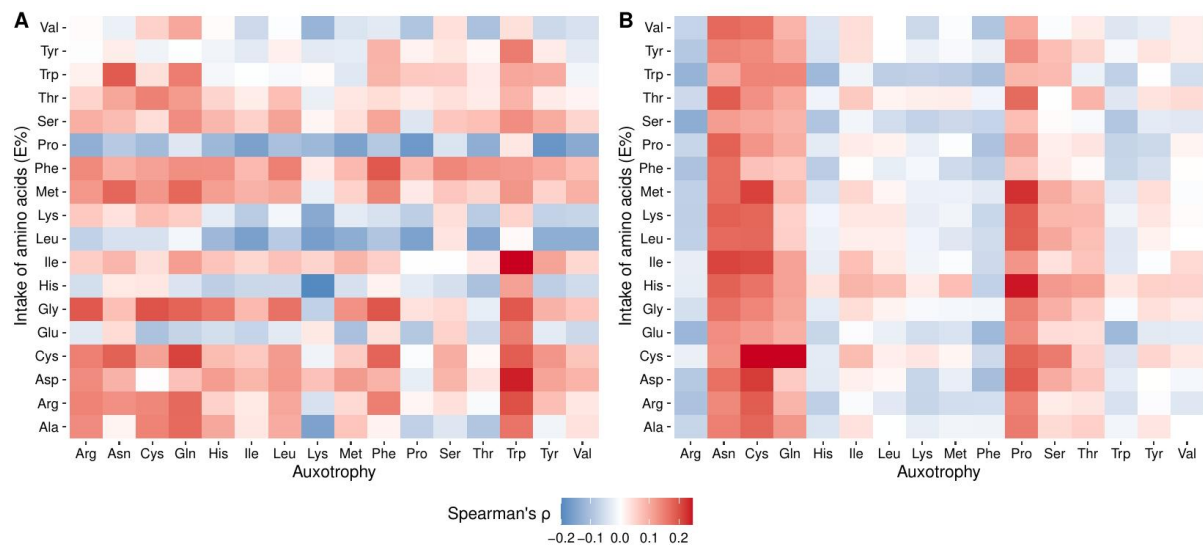
**Supplementary Figure 5: Missing enzymes in amino acid biosynthesis pathways**

The bar plots display the abundance of missing enzymes in the biosynthesis pathways for several amino acids and BCAA biosynthesis pathways in auxotrophic bacteria. The order of the enzymes in the bar plots represents the one in the pathway. (A) Missing enzymes in the tryptophan pathway of tryptophan auxotrophic bacteria, (B) Missing enzymes in the histidine pathway of histidine auxotrophic bacteria, (C) Missing enzymes in the chorismate pathway of chorismate auxotrophic bacteria, (D) Missing enzymes in the serine pathway of serine auxotrophic bacteria, (E) Comparison of the BCAA pathways of isoleucine, leucine, and valine auxotrophic bacteria, the colors indicate which enzymes are shared in the biosynthesis pathways, the definition of the pathways are based on MetaCyc.



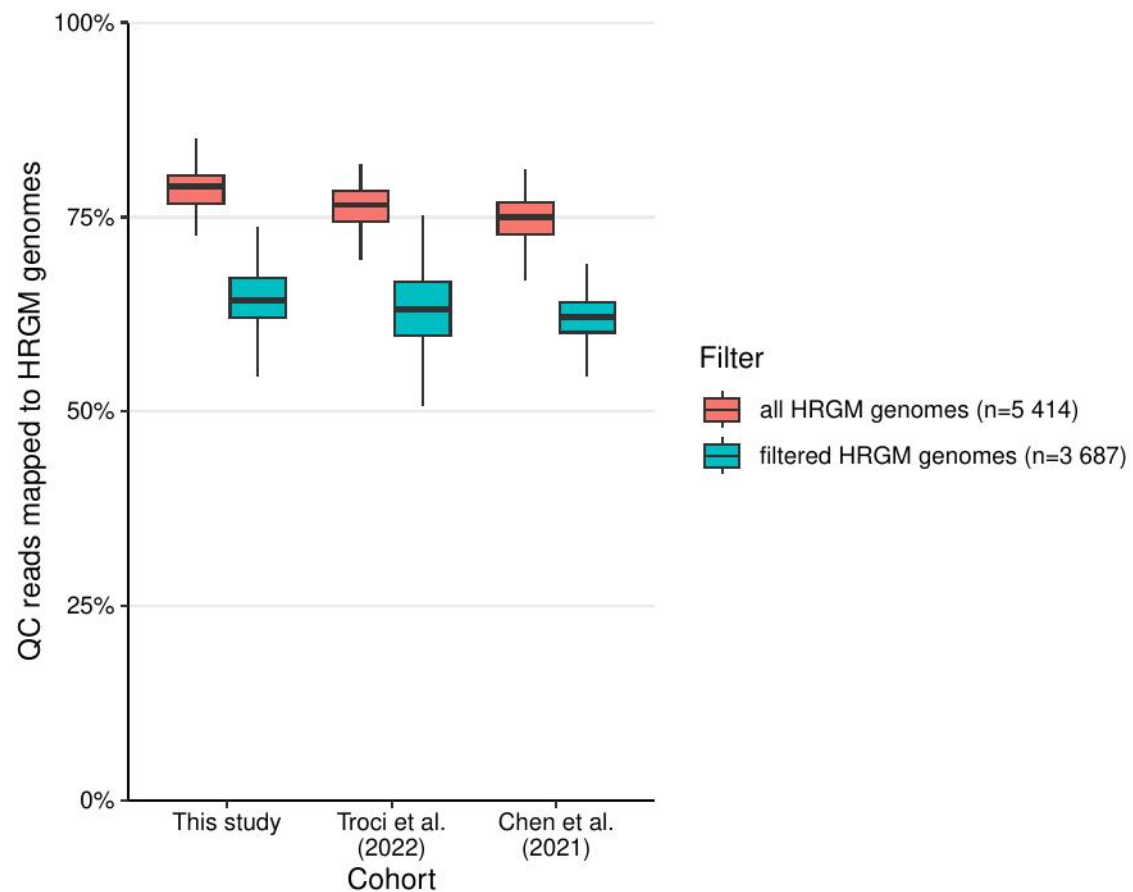
### Supplementary Figure 6: Amino acid auxotrophies predicted with varying genome completeness levels

Relative abundance of predicted amino acids auxotrophs depending on the completeness cutoff for reference genomes from the HRGM catalog. Four different genome completeness cutoffs were tested: 80% (A), 85% (B, same as Figure 4A), 90% (C), and 95% (D). The results indicate that the distribution of the relative abundance of predicted amino acid auxotrophies was stable with respect to the chosen completeness cut-off for reference genome filtering.



**Supplementary Figure 7: Spearman correlation between the dietary intake of amino acids and the frequency of amino acid auxotrophic bacteria in the gut microbiomes** (A) at the beginning of the study, (B) at the end of the study (3 years later). No statistically significant associations were found (FDR-corrected  $p$  value  $>0.05$ ).





**Supplementary Figure 8: Percentage of quality-controlled metagenomic reads from three cohorts (this study, Troci *et al.* 2022 [150], and Chen *et al.* 2021 [151]) to reference genomes from the HRGM catalog.**

## Supplementary Tables

### **Supplementary Table 1: Auxotrophy prediction for HRGM models**

*The data set about auxotrophy prediction of reconstructed genome-scale metabolic models of HRGM genomes is freely accessible via the online supporting documentation of the article Starke et al. 2023, The ISME Journal. <https://doi.org/10.1038/s41396-023-01537-3>.*

**Supplementary Table 2: Validation of *in silico* predicted auxotrophies with experimentally verified auxotrophies in human gut bacteria.**

Explanation of three columns for amino acid prediction: Auxotrophy is indicated by 0 and prototrophy is indicated by 1; first column: experimentally verified auxotrophy, second column: auxotrophy prediction with gapseq model, third column: auxotrophy prediction with AGORA2 model

Human gut bacteria	Strain	Taxo- nomy ID	AGORA2 model ID	GCF-Number	Ala			Val			Met			Leu			Ile			Pro			Trp		
<i>Akkermansia muciniphila</i>	ATCC BAA-835	349741	Akkermansia_muciniphila_ATCC_BAA_835	GCF_00002025.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
<i>Clostridioides difficile</i>	630	272563	Clostridioides_difficile_630**	GCF_000009205.2	1	1	–	0	1	–	1	1	–	0	1	–	0	1	–	0	0	–	0	0	
<i>Clostridium perfringens</i>	NCTC10240	1502	Clostridium_perfringens_ATCC_13124	GCF_900461395.1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
<i>Enterococcus faecalis</i>	OG1RF	474186	Enterococcus_faecalis_OG1RF_ATCC_47077	GCF_000172575.2	1	1	1	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1	0	0	0
<i>Faecalibacterium prausnitzii</i>	SL3/3	657322		GCF_000209855.1	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	0	0	–
<i>Fusobacterium nucleatum</i>	13_3C	1357398	Fusobacterium_nucleatum_13_3C	GCF_000523555.1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	0	1	0	0	1	0	0
<i>Fusobacterium varium</i>	ATCC 8501	856		GCF_900637705.1	1	1	–	1	1	–	1	0	–	1	0	–	1	1	–	1	1	–	1	0	–
<i>Helicobacter pylori</i>	26695	85962	Helicobacter_pylori_26695	GCF_000307795.1	1	1	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	1	1	1	1
<i>Lactobacillus paracasei</i>	LC2W	999378	Lactobacillus_paracasei_LC2W	GCF_000194785.1	1	1	1	0	0	1	1	0	1	0	0	1	0	1	1	1	0	1	0	1	1
<i>Lactobacillus delbrueckii</i>	subsp. lactis CRL581	1269760		GCF_000409675.1	1	1	–	0	0	–	0	1	–	0	0	–	0	0	–	1	1	–	0	0	–
<i>Lactobacillus helveticus</i>	CNRZ32	326425		GCF_000422165.1	1	1	–	0	0	–	0	0	–	0	0	–	0	0	–	0	0	–	1	0	–
<i>Lactobacillus plantarum</i>	WCFS1	220668	Lactobacillus_plantarum_WCFS1	GCF_000203855.3	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	1	0	1	1	1	1
<i>Listeria monocytogenes</i>	10403	393133	Listeria_monocytogenes_10403S	GCF_000168695.2	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Moraxella catarrhalis</i>	O35E	857571	Moraxella_catarrhalis_O35E	GCF_000193085.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Mycobacterium tuberculosis</i>	H37Rv	83332	Mycobacterium_tuberculosis_H37Rv	GCF_000195955.2	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
<i>Roseburia inulinivorans</i>	A2-194 (DSM 16841)	622312	Roseburia_inulinivorans_DSM_16841	GCF_020731525.1	1*	1	0	N A*	1	1	N A*	1	0	N A*	1	1	N A*	1	1	N A*	1	1	0*	0	0

Phe			Lys			Arg			His			Tyr			Thr			Glu			Gln			Gly			Ser			Cys			Asp			Asn			Ref
1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	28687644				
1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	0	1	–	1	1	–	1	1	–	2864896			
0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	2864896			
1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	8349561			
1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	1	1	1	–	25002542, 32665271			
1	0	1	0	0	0	1	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	16887630			
1	1	–	1	1	–	1	0	–	1	0	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	0	–	1	1	–	1	1	–	21547305
0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	16672389		
0	0	1	1	1	1	0	0	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	25452194		
0	0	–	1	1	–	0	0	–	0	0	–	0	0	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	0	1	–	1	1	–	1	1	–	15486708
0	0	–	0	1	–	0	0	–	0	0	–	0	0	–	0	0	–	0	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	17993552
1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	16269766	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	14602660	
1	1	1	1	1	1	0	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25156736	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	11160012		
NA*	1	1	1*	1	1	NA*	1	1	1*	1	1	NA*	1	1	NA*	1	1	NA*	1	1	NA*	1	1	NA*	1	1	NA*	1	0	NA*	1	1	NA*	1	1	NA*	1	1	32665271

Human gut bacteria	Strain	Taxonomy ID	AGORA2 model ID	GCF-Number	Ala			Val			Met			Leu			Ile			Pro			Trp		
<i>Streptococcus agalactiae</i>	A909	205921	Streptococcus_agalactiae_A909**	GCF_000012705.1	1	1	–	0	0	–	0	0	–	0	0	–	0	0	–	1	0	–	0	0	–
<i>Streptococcus thermophilus</i>	LMG 18311	264199	Streptococcus_thermophilus_LMG 18311	GCF_000011825.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Subdoligranulum variable</i>	DSM 15176	21485	Subdoligranulum_variabale_DSM_15176	GCF_025152575.1	NA*	1	0	NA*	1	1	NA*	1	0	NA*	1	0	NA*	1	1	NA*	1	1	0	0	0
<i>Anaerostipes caccae</i>	DSM 14662	411490	Anaerostipes_caccae_DSM_14662	GCF_025146005.1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1
<i>Anaerostipes hadrus</i>	SSC/2	649756	Anaerostipes_hadrus_DSM_3319	GCF_000210695.1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Clostridium beijerinckii</i>	G117	1200349	Clostridium_beijerinckii_G117	GCF_000280535.1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Clostridium butyricum</i>	DSM 10702	1316931	Clostridium_butyricum_DSM_10702	GCF_000409755.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<i>Clostridium tyrobutyricum</i>	FAM2255 2	1519		GCF_000816585.1	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–
<i>Clostridium tyrobutyricum</i>	FAM2255 3	1519		GCF_000816635.1	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–
<i>Clostridium tyrobutyricum</i>	UC7086	1212547	Clostridium_tyrobutyricum_UC7086**	GCF_000332015.1	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–
<i>Clostridium tyrobutyricum</i>	ATCC 25755	1121342	Clostridium_tyrobutyricum_DSM_2637	GCF_000359585.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
<i>Eubacterium hallii</i>	DSM 3353	411469	Eubacterium_hallii_DSM_3353	GCF_000173975.1	1	1	0	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
<i>Roseburia faecis</i>	M72/1	451638	Roseburia_faecis_M72	GCF_001406815.1	1	1	0	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
<i>Roseburia intestinalis</i>	L1-82	536231	Roseburia_intestinalis_L1_82**	GCF_900537995.1	1*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	1*	1	–
<i>Roseburia intestinalis</i>	M50/1	657315		GCF_000209995.1	1*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	1*	1	–
<i>Bacteroides fragilis</i>	NCTC 9343	272559	Bacteroides_fragilis_NCTC_9343*	GCF_000025985.1	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–
<i>Bacteroides vulgatus</i>	ATCC 8482	435590	Bacteroides_vulgatus_ATCC_8482**	GCF_000012825.1	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–
<i>Bacteroides thetaiotaomicron</i>	VPI-5482	226186	Bacteroides_thetaiotaomicron_VPI_5482**	GCF_000011065.1	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–
<i>Bacteroides distasonis</i>	ATCC 8503	435591	Parabacteroides_distasonis_ATCC_8503**	GCF_000012845.1	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–
<i>Streptococcus thermophilus</i>	ST13	1308		GCF_020404245.1	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–

Phe			Lys			Arg			His			Tyr			Thr			Glu			Gln			Gly			Ser			Cys			Asp			Asn			Ref
NA*	1	1	1*	1	1	NA*	1	1	1*	1	1	NA*	1	1	NA*	1	1	NA*	1	1	NA*	1	1	NA*	1	0	NA*	1	1	NA*	1	1	NA*	1	1	32665271			
0	0	–	0	0	–	0	0	–	0	0	–	0	0	–	1	1	–	0	1	–	1	1	–	1	1	–	1	1	–	0	1	–	1	1	–	1	1	–	342537
1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	19346354	
NA*	1		NA*	1	1	NA*	1	1	NA*	1	1	0*	1	0	NA*	1	1	NA*	1	1	NA*	1	1	NA*	1	1	NA*	1	1	NA*	1	1	NA*	1	1	NA*	1	1	32665271
1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	32665271	
1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	32665271	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	27375262	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	27375262	
1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	27375262
1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	27375262
1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	27375262
1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	27375262
1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	32665271	
1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	32665271	
NA*	1	–	1*	1	–	NA*	1	–	1*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	32665271
NA*	1	–	1*	1	–	NA*	1	–	1*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	NA*	1	–	32665271
1	0	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	4853401
1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	4853401
1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	4853401
1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	4853401
1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	1	1	–	9841662

\*: Only some amino acids were tested in the reference study. Therefore, conclusions about auxotrophies/prototrophies for other amino acids cannot be drawn. All non-tested amino acids have an NA.

\*\*\*: the model files are incorrect and therefore an auxotrophy prediction was not possible

–: No prediction for AGORA2 is possible, either because the AGORA2 collection does not contain a model for the corresponding strain or the SBML model file was not readable due to formatting issues (see \*\*).

**Supplementary Table 3: Auxotrophy prediction for AGORA2 models**

*The data set about auxotrophy prediction for AGORA2 models is freely accessible via the online supporting documentation of the article Starke et al. 2023, The ISME Journal. <https://doi.org/10.1038/s41396-023-01537-3>.*

**Supplementary Table 4: Auxotrophy prediction for 124 prototrophic genotypes (Price 2022, doi: 10.1038/s41559-022-01936-3)**

Ala	Val	Met	Leu	Ile	Pro	Trp	Phe	Lys	Arg	His	Tyr	Thr	Glu	Gln	Gly	Ser	Cys	Asp	Asn	Genomes	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000006985.1	Chlorobaculum thiosulfatiphilum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000008325.1	Methylococcus capsulatus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000008765.1	Clostridium acetobutylicum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000010525.1	Azorhizobium caulinodans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000011905.1	Dehalococcoides mccartyi
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000012925.1	Geobacter metallireducens
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000012965.1	Sulfurimonas denitrificans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000013085.1	Rhodospirillum rubrum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000013325.1	Novosphingobium aromaticivorans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000015045.1	Pelobacter propionicus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000015305.1	Mycolicibacterium vanbaalenii
1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000015505.1	Polaromonas naphthalenivorans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	GCF_000015585.1	Halorhodospira halophila
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000016185.1	Rhodospirillum centenum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000016505.1	Clostridium kluyveri
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000016745.1	Geotalea uraniireducens
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000016765.1	Rhizorhabdus wittichii
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000017145.1	Sinorhizobium medicae
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000017645.1	Xanthobacter autotrophicus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000018545.1	Sinorhizobium fredii
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000019165.1	Heliomicrobium modesticaldum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000019845.1	Beijerinckia indica
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000020025.1	Nostoc punctiforme
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000020045.1	Paraburkholderia phymatum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000020385.1	Trichlorobacter lovleyi



Ala	Val	Met	Leu	Ile	Pro	Trp	Phe	Lys	Arg	His	Tyr	Thr	Glu	Gln	Gly	Ser	Cys	Asp	Asn	Genomes	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000020465.1	Chlorobium limicola
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000020505.1	Chlorobaculum parvum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000020625.1	Prosthecochloris aestuarii
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000021485.1	Acidithiobacillus ferrooxidans
1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000021545.1	Sulfurihydrogenibium azorense
1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000021565.1	Persephonella marina
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000021745.1	Methylocella silvestris
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000021925.1	Desulfitobacterium hafniense
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000022085.1	Methylobacterium nodulans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000023025.1	Teredinibacter turnerae
1	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	GCF_000023265.1	Acidimicrobium ferrooxidans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000024385.1	Thermomonospora curvata
1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	GCF_000024605.1	Ammonifex degensii
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000025485.1	Allochromatium vinosum
1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000025605.1	Thermocrinis albus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000025725.1	Denitrovibrio acetiphilus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000025945.1	Desulfotalea psychrophila
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000058485.1	Frankia alni
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000092045.1	Rhizobium etli
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000092245.1	Arcobacter nitrofigilis
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000092905.1	Kyrpidia tusciae
1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	GCF_000143965.1	Desulfarculus baarsii
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000166055.1	Rhodomicrobium vannielii
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	GCF_000170735.1	Caminibacter mediatlanticus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000183405.1	Calditerrivibrio nitroreducens
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000183725.1	Sulfuricurvum kujiense

Ala	Val	Met	Leu	Ile	Pro	Trp	Phe	Lys	Arg	His	Tyr	Thr	Glu	Gln	Gly	Ser	Cys	Asp	Asn	Genomes	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000185805.1	Thermovibrio ammonificans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000185905.1	Mesorhizobium ciceri
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000186245.1	Nitratifactor salsuginis
1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	GCF_000195295.1	Desulfobacca acetoxidans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000195775.1	Rhodopseudomonas palustris
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000204645.1	Alicyclophilus denitrificans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000214665.1	Methylomonas methanica
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000214825.1	Thiomicrospira cyclica
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000215085.1	Desulforhamulus ruminis
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000227685.2	Thioalkalivibrio paradoxus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000244975.1	Saccharomonospora cyanea
1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000284315.1	Leptospirillum ferrooxidans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000310185.1	Thauera aminoaromatica
1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000341125.1	Nocardiopsis lucentensis
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000372865.1	Methylosarcina fibrata
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000375485.1	Desulfovibrio oxycloinae
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000378965.1	Thioalkalivibrio thiocyanodenitrificans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000381085.1	Thiomicrothabodus arctica
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000384075.1	Methylovulum miyakonense
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000384235.1	Hydrogenovibrio halophilus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000385335.1	Methyloferula stellata
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000421465.1	Methylohalobius crimeensis
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000422285.1	Desulfatiglans anilini
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000423825.1	Thermithiobacillus tepidarius
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000425265.1	Maridesulfovibrio zosterae
1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000427445.1	Methylocapsa acidiphila

Ala	Val	Met	Leu	Ile	Pro	Trp	Phe	Lys	Arg	His	Tyr	Thr	Glu	Gln	Gly	Ser	Cys	Asp	Asn	Genomes	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000428045.1	Sedimenticola selenatireducens
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000429905.1	Desulfatibacillum aliphaticivorans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000429965.1	Desulfogranum mediterraneum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000429985.1	Maridesulfovibrio bastinii
1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000430725.1	Azohydromonas australica
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000482785.1	Dexia gummosa
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000483485.1	Thiomicrothrix chilensis
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000519045.1	Dechloromonas agitata
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000526715.1	Hydrogenovibrio kuenenii
1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000527155.1	Haloglycomyces albus
1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000619805.1	Sulfurihydrogenibium subterraneum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000621325.1	Thiothrix lacustris
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000711195.1	Thiomicrospira pelophila
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000711315.1	Hydrogenovibrio marinus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000745425.1	Beijerinckia mobilis
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000745855.1	Xenophilus azovorans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000745975.1	Desulfobacter vibrioformis
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000746085.1	Methylocapsa aurea
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000828635.1	Sulfuritalea hydrogenivorans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000934765.1	Thiomicrospira microaerophila
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000967895.1	Oleispira antarctica
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000968535.2	Methylovibrio alcaliphilum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_000969705.1	Nitrospirillum alkaliphilum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_001189915.1	Herbaspirillum autotrophicum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_001431535.1	Stenotrophomonas chelatiphaga
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_001484605.1	Pseudarthrobacter sulfonivorans

Ala	Val	Met	Leu	Ile	Pro	Trp	Phe	Lys	Arg	His	Tyr	Thr	Glu	Gln	Gly	Ser	Cys	Asp	Asn	Genomes	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_001541235.1	Hyphomicrobium sulfonivorans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_001579945.1	Steroidobacter denitrificans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_001592305.1	Hydrogenophaga taeniospiralis
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_001995255.1	Thioalkalivibrio halophilus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_002000365.1	Thioalkalivibrio denitrificans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_002019605.1	Alkalihalobacterium alkalinitrilicum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_002117405.1	Methylocystis bryophila
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_002151265.1	Halomonas desiderata
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900100495.1	Pseudomonas benzenivorans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900102855.1	Thiohalomonas denitrificans
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900107645.1	Desulfuromusa kysingii
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900111775.1	Desulfuromonas acetexigens
1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	GCF_900114975.1	Thermodesulforhabdus norvegica
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900115975.1	Desulfoscipio geothermicus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900129305.1	Desulfacinum infernum
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900141795.1	Sulfurivirga caldicuralii
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900142125.1	Malonomonas rubra
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900167125.1	Paucidesulfovibrio gracilis
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900176045.1	Nitratiruptor tergaricus
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900176285.1	Desulfacinum hydrothermale
1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	GCF_900188115.1	Haloechothrix alba

**Supplementary Table 5: Abbreviations used in Figure 8**

Abbreviation	Definition
TMAO	Trimethylamine N-oxide
TRPBetaine	Tryptophan betaine
t4-OH-Pro	trans-4-Hydroxyproline
SDMA	Symmetric dimethylarginine
ProBetaine	Proline betaine
alpha-AAA	alpha-Aminoadipic acid
ADMA	Asymmetric dimethylarginine
AABA	alpha-Aminobutyric acid
3-Met-His	3-Methylhistidine
1-Met-His	1-Methylhistidine
TMCA	Tauromurocholic acid
TDCA	Taurodeoxycholic acid
TCDCA	Taurochenodeoxycholic acid
TCA	Taurocholic acid
GUDCA	Glycoursodeoxycholic acid
GLCAS	Glycolithocholic acid sulfate
GLCA	Glycolithocholic acid
GDCA	Glycodeoxycholic acid
GCDCA	Glycochenodeoxycholic acid
GCA	Glycocholic acid
DCA	Deoxycholic acid
GABA	Gamma-Aminobutyric acid
p-Cresol-SO <sub>4</sub>	p-Cresol sulfate
FA(18:2)	Octadecadienoic acid
EPA	Eicosapentaenoic acid (FA(20:5 Omega3))

Abbreviation	Definition
DHA	Docosahexaenoic acid (FA(22:6 Omega3)
AA	Arachidonic acid (FA(20:4 Omega6)
DHEAS	Dehydroepiandrosterone sulfate
Ind-SO <sub>4</sub>	Indoxyl sulfate
3-IPA	3-Indolepropionic acid
3-IAA	3-Indoleacetic acid

**Supplementary Table 6: Logical expression used as rules to decide the absence/presence of specific compounds in a genome-specific gap filling medium\***

Compound name	Compound Id	Uptake bound [mmol/(gDW*hr)]	Rule
<i>Inorganic compounds</i>			
Water	cpd00001	100	TRUE
Sodium	cpd00971	10	TRUE
K <sup>+</sup>	cpd00205	10	TRUE
Cl <sup>-</sup>	cpd00099	10	TRUE
Fe <sup>2+</sup>	cpd10515	10	TRUE
Fe <sup>3+</sup>	cpd10516	0.1	TRUE
Phosphate	cpd00009	10	TRUE
Co <sup>2+</sup>	cpd00149	10	TRUE
Mg <sup>2+</sup>	cpd00254	10	TRUE
Ni <sup>2+</sup>	cpd00244	1	TRUE
Sulfate	cpd00048	10	TRUE
H <sub>2</sub> S	cpd00239	1	TRUE
Zn <sup>2+</sup>	cpd00034	10	TRUE
Cu <sup>2+</sup>	cpd00058	10	TRUE
Mn <sup>2+</sup>	cpd00030	10	TRUE
Ca <sup>2+</sup>	cpd00063	10	TRUE
Ammonium	cpd00013	10	!("N2FIX-PWY"   "PWY-7576")
N <sub>2</sub>	cpd00528	25	"N2FIX-PWY"   "PWY-7576"
O <sub>2</sub>	cpd00007	12.5	"PWY-3781"   "PWY1YI0-3"   "PWY-7279"   "PWY1YI0-6"
O <sub>2</sub>	cpd00007	12.5	"PWY0-1565"   "PWY0-1561"   "PWY0-1544"   "PWY-7545"   "PWY-7544"   "PWY0-1353"   "PWY0-1329"
O <sub>2</sub>	cpd00007	12.5	"PWY0-1334"   "PWY0-1568"   "PWY0-1335"   "PWY0-1567"
CO <sub>2</sub>	cpd00011	25	"CALVIN-PWY"   "CODH-PWY"   "PWY-5789"   "PWY-7784"   "LWP-GS"   "METHANOGENESIS-PWY"

Compound name	Compound Id	Uptake bound [mmol/(gDW*hr)]	Rule
H2	cpd11640	100	"CODH-PWY"   "LWP-GS"   "METHANOGENESIS-PWY"   "PWY-7867"
Photons	cpd11632	50	"PWY-101"

<i>Carbohydrates</i>			
D-Glucose	cpd00027	5	rxn05226   rxn08612
D-Glucose	cpd00027	5	rxn05573   rxn08616
D-Glucose	cpd00027	5	rxn05147   rxn08606
D-Fructose	cpd00082	5	rxn05560   rxn08536
Lactose	cpd00208	2.5	"LACTOSECAT-PWY"   "LACTOSEUTIL-PWY"   "BGALACT-PWY"
Starch (n=27)	cpd90003	0.18	rxn90043 & rxn90034
Maltose	cpd00179	2.5	"MALTOSECAT-PWY"
2'-Fucosyllactose	cpd90001	1.67	"2FLHYD"
Pectin	cpd27519	1.67	"PWY-7246"   "PWY-7248"
Melitose/Raffinose	cpd00382	1.67	rxn05643 & (rxn00818   rxn34781)
Sucrose	cpd00076	2.5	"SUCUTIL-PWY"   "PWY-621"   "PWY-5384"   "SUCROSEUTIL2-PWY"
Inulin	cpd28763	0.42	"INULIN-DEGRADATION-EXO"
N-acetylneuraminate	cpd00232	5	"PWY0-1324" & "GLUAMCAT-PWY"
N-Acetyl-D-glucosamine	cpd00122	5	"GLUAMCAT-PWY" & (rxn05485   rxn08041)
Xylan-b-1-4	cpd90021	0.01	"PWY-6717"

<i>Amino acids and Chorismate</i>			
L-Tryptophan	cpd00065	0.1	!("TRPSYN-PWY2"   "TRPSYN-PWY")
L-Histidine	cpd00119	0.1	!"HISTSYN-PWY"
L-Serine	cpd00054	0.1	!"SERSYN-PWY"
L-Arginine	cpd00051	0.1	!("ARGSYN-PWY"   "ARGSYNBSUB-PWY"   "PWY-5154"   "PWY-7400")
L-Asparagine	cpd00132	0.1	!("ASPARAGINE-BIOSYNTHESIS"   "ASPARAGINESYN-PWY"   "PWY490-4")



Compound name	Compound Id	Uptake bound [mmol/(gDW*hr)]	Rule
L-Glutamine	cpd00053	0.1	!"GLNSYN-PWY"
Glycine	cpd00033	0.1	!("GLYSYN-PWY"   "GLYSYN-ALA-PWY"   "GLYSYN-THR-PWY")
Chorismate	cpd00216	0.1	rxn02213 + (rxn01740   rxn01741) + (rxn01739   rxn38708) + rxn02476 + rxn01255 < 3
L-Cysteine	cpd00084	0.1	!("CYSTSYN-PWY"   "PWY-6308"   "HOMOCYSDEGR-PWY"   "PWY-7289"   "PWY-8009"   "PWY-8010")
L-Isoleucine	cpd00322	0.1	!("ILEUSYN-PWY"   "PWY-5101"   "PWY-5104")
L-Leucine	cpd00107	0.1	!("LEUSYN-PWY" & (rxn00003   rxn33164) & rxn02186 & rxn00898)
L-Lysine	cpd00039	0.1	!("DAPLYSINESYN-PWY"   "PWY-2941"   "PWY-2942"   "PWY-3081"   "PWY-5097")
L-Methionine	cpd00060	0.1	!("HOMOSER-METSYN-PWY"   "PWY-702"   "HSERMETANA-PWY"   "PWY-7977")
L-Phenylalanine	cpd00066	0.1	!((rxn01256 & rxn01000 & (rxn00493   rxn21134   rxn28060   rxn38254))   "PWY-3462"   "PWY-7432")
L-Proline	cpd00129	0.1	!("PROSYN-PWY"   "PWY-3341")
L-Threonine	cpd00161	0.1	!("HOMOSER-THRESYN-PWY"   "THRBIOSYN-GS")
L-Tyrosine	cpd00069	0.1	!("TYRSYN"   "PWY-3461"   "PWY-6120"   "PWY-6134")
L-Valine	cpd00156	0.1	!"VALSYN-PWY"
<i>Vitamins</i>			
Pantothenic acid	cpd00644	0.025	!(rxn01791   rxn16218)
Thiamin	cpd00305	0.025	!(rxn02305   rxn23042)
<i>Other carboxylic acids and alcohols</i>			
L-Lactate	cpd00159	15	rxn90054
L-Lactate	cpd00159	15	"PWY-5494"   "PWY-8086"   "PWY-8086-GS"
Acetate	cpd00029	20	rxn00875 & rxn00225 & rxn00173 & rxn90001
Acetate	cpd00029	25	"METH-ACETATE-PWY"
Succinate	cpd00036	10	"PWY-5677"
Succinate	cpd00036	10	(rxn00602   rxn01996) & rxn90096 & rxn09174

Compound name	Compound Id	Uptake bound [mmol/(gDW*hr)]	Rule
Methanol	cpd00116	20	"CO2FORM-PWY"   "PWY-6966"   "PWY-6510"   "PWY-6509"   "PWY-5506"
Formate	cpd00047	20	"PWY-7868"
<i>Other compounds</i>			
sulfoquinovose	cpd22614	10	"PWY-7446"   "PWY-7722"
Gallate	cpd01020	10	"GALLATE-DEGRADATION-II-PWY"   "GALLATE-DEGRADATION-I-PWY"   "P3-PWY"
Quercetin	cpd00313	10	"QUERCETIN-ANAERO-DEG"
Genistein	cpd03971	10	rxn90097
Daidzein	cpd07097	10	"PWY-6996"
Dimethylsulfide	cpd00450	10	"PWY-5258"
Methanethiol	cpd00324	10	"PWY-5259"

\* For column "Rule", terms in quotes refer to pathway identifiers, and terms with the prefix "rxn" refer to specific single metabolic reactions. The individual terms within the logical expression would translate to TRUE if the pathway/reaction was found by gapseq in the reconstruction steps 'find' or 'find-transport.' Finally, if the full logical expression in column "Rule" translates to TRUE, the corresponding compound is added to the final gap filling medium with a maximum uptake bound of the value stated in column "Uptake bound".

Supplementary information for Chapter III

Comparative Analysis of Amino Acid Auxotrophies and  
Peptidase Profiles in Non-Dysbiotic and Dysbiotic Small  
Intestinal Microbiomes

## Supplementary Tables

**Supplementary Table 7: Relative abundance of HRGM genomes in samples from the Reimagine study**

**Supplementary Table 8: Relative abundance of HRGM genomes in samples from the SIBO study**

**Supplementary Table 9: Presence of peptidases in microbial samples from the duodenum, jejunum, farthest distance and stool**

**Supplementary Table 10: Statistical results for association of peptidases with SI(Duodenum) or LI(Stool)**

**Supplementary Table 11: Presence of peptidases in duodenal samples from SIBO and non-SIBO patients**

**Supplementary Table 12: Statistical results for association of peptidases with SIBO or non-SIBO**

*All Supplementary tables for Chapter III can be found on Zenodo under the following link:  
10.5281/zenodo.14562773*

## References

- [1] W. M. De Vos, H. Tilg, M. Van Hul, und P. D. Cani, „Gut microbiome and health: mechanistic insights“, *Gut*, Bd. 71, Nr. 5, S. 1020–1032, Mai 2022, doi: 10.1136/gutjnl-2021-326789.
- [2] K. Hou u. a., „Microbiota in health and diseases“, *Sig Transduct Target Ther*, Bd. 7, Nr. 1, Art. Nr. 1, Apr. 2022, doi: 10.1038/s41392-022-00974-4.
- [3] The Human Microbiome Project Consortium, „Structure, function and diversity of the healthy human microbiome“, *Nature*, Bd. 486, Nr. 7402, S. 207–214, Juni 2012, doi: 10.1038/nature11234.
- [4] S. A. Fillon u. a., „Novel Device to Sample the Esophageal Microbiome—The Esophageal String Test“, *PLoS ONE*, Bd. 7, Nr. 9, S. e42938, Sep. 2012, doi: 10.1371/journal.pone.0042938.
- [5] D. Liu, R. Zhang, S. Chen, B. Sun, und K. Zhang, „Analysis of gastric microbiome reveals three distinctive microbial communities associated with the occurrence of gastric cancer“, *BMC Microbiol*, Bd. 22, Nr. 1, S. 184, Dez. 2022, doi: 10.1186/s12866-022-02594-y.
- [6] W. M. Wu, Y. S. Yang, und L. H. Peng, „Microbiota in the stomach: new insights“, *J Dig Dis*, Bd. 15, Nr. 2, S. 54–61, Feb. 2014, doi: 10.1111/1751-2980.12116.
- [7] G. G. S. Leite u. a., „Mapping the Segmental Microbiomes in the Human Small Bowel in Comparison with Stool: A REIMAGINE Study“, *Dig Dis Sci*, Bd. 65, Nr. 9, S. 2595–2604, Sep. 2020, doi: 10.1007/s10620-020-06173-x.
- [8] J. L. Baker, J. L. Mark Welch, K. M. Kauffman, J. S. McLean, und X. He, „The oral microbiome: diversity, biogeography and human health“, *Nat Rev Microbiol*, Sep. 2023, doi: 10.1038/s41579-023-00963-6.
- [9] X. Peng u. a., „Oral microbiota in human systematic diseases“, *Int J Oral Sci*, Bd. 14, Nr. 1, S. 14, Dez. 2022, doi: 10.1038/s41368-022-00163-7.
- [10] B. A. Peters u. a., „Oral Microbiome Composition Reflects Prospective Risk for Esophageal Cancers“, *Cancer Research*, Bd. 77, Nr. 23, S. 6777–6787, Dez. 2017, doi: 10.1158/0008-5472.CAN-17-1296.
- [11] L. E. Wroblewski, R. M. Peek, und K. T. Wilson, „*Helicobacter pylori* and Gastric Cancer: Factors That Modulate Disease Risk“, *Clin Microbiol Rev*, Bd. 23, Nr. 4, S. 713–739, Okt. 2010, doi: 10.1128/CMR.00011-10.
- [12] A. J. Kastl, N. A. Terry, G. D. Wu, und L. G. Albenberg, „The Structure and Function of the Human Small Intestinal Microbiota: Current Understanding and Future Directions“, *Cellular and Molecular Gastroenterology and Hepatology*, Bd. 9, Nr. 1, S. 33–45, 2020, doi: 10.1016/j.jcmgh.2019.07.006.
- [13] A. Nogal, A. M. Valdes, und C. Menni, „The role of short-chain fatty acids in the interplay between gut microbiota and diet in cardio-metabolic health“, *Gut Microbes*, Bd. 13, Nr. 1, S. 1897212, Jan. 2021, doi: 10.1080/19490976.2021.1897212.
- [14] D. Parada Venegas u. a., „Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases“, *Front. Immunol.*, Bd. 10, S. 277, März 2019, doi: 10.3389/fimmu.2019.00277.
- [15] A. M. Seekatz u. a., „Spatial and Temporal Analysis of the Stomach and Small-Intestinal Microbiota in Fasted Healthy Humans“, *mSphere*, Bd. 4, Nr. 2, S. e00126-19, Apr. 2019, doi: 10.1128/mSphere.00126-19.
- [16] S. A. Scott, J. Fu, und P. V. Chang, „Microbial tryptophan metabolites regulate gut barrier function via the aryl hydrocarbon receptor“, *Proc. Natl. Acad. Sci.*

- U.S.A., Bd. 117, Nr. 32, S. 19376–19387, Aug. 2020, doi: 10.1073/pnas.2000047117.
- [17] H. M. Roager und T. R. Licht, „Microbial tryptophan catabolites in health and disease“, *Nat Commun*, Bd. 9, Nr. 1, S. 3294, Dez. 2018, doi: 10.1038/s41467-018-05470-4.
  - [18] A. Agus, K. Clément, und H. Sokol, „Gut microbiota-derived metabolites as central regulators in metabolic disorders“, *Gut*, Bd. 70, Nr. 6, S. 1174–1182, Juni 2021, doi: 10.1136/gutjnl-2020-323071.
  - [19] A.-M. Davila u. a., „Intestinal luminal nitrogen metabolism: Role of the gut microbiota and consequences for the host“, *Pharmacological Research*, Bd. 68, Nr. 1, S. 95–107, Feb. 2013, doi: 10.1016/j.phrs.2012.11.005.
  - [20] K. Zengler und L. S. Zaramela, „The social network of microorganisms — how auxotrophies shape complex communities“, *Nat Rev Microbiol*, Bd. 16, Nr. 6, S. 383–390, Juni 2018, doi: 10.1038/s41579-018-0004-5.
  - [21] X. Zeng u. a., „Gut bacterial nutrient preferences quantified in vivo“, *Cell*, Bd. 185, Nr. 18, S. 3441–3456.e19, Sep. 2022, doi: 10.1016/j.cell.2022.07.020.
  - [22] G. T. Macfarlane und C. Allison, „Utilisation of protein by human gut bacteria“, *FEMS Microbiology Letters*, Bd. 38, Nr. 1, S. 19–24, Mai 1986, doi: 10.1111/j.1574-6968.1986.tb01934.x.
  - [23] F. Velez-Cortes und H. Wang, „Characterization and Spatial Mapping of the Human Gut Metasecretome“, *mSystems*, Bd. 7, Nr. 6, S. e00717-22, Dez. 2022, doi: 10.1128/msystems.00717-22.
  - [24] S. A. Gibson, C. McFarlan, S. Hay, und G. T. MacFarlane, „Significance of microflora in proteolysis in the colon“, *Appl Environ Microbiol*, Bd. 55, Nr. 3, S. 679–683, März 1989, doi: 10.1128/aem.55.3.679-683.1989.
  - [25] E. G. Zoetendal u. a., „The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates“, *ISME J*, Bd. 6, Nr. 7, S. 1415–1426, Juli 2012, doi: 10.1038/ismej.2011.212.
  - [26] B. Ahlman, C.-E. Leijonmarck, C. Lind, E. Vinnars, und J. Wernerman, „Free Amino Acids in Biopsy Specimens from the Human Colonic Mucosa“, *Journal of Surgical Research*, Bd. 55, Nr. 6, S. 647–653, Dez. 1993, doi: 10.1006/jsre.1993.1198.
  - [27] G. T. Macfarlane, C. Allison, S. A. W. Gibson, und J. H. Cummings, „Contribution of the microflora to proteolysis in the human large intestine“, *Journal of Applied Bacteriology*, Bd. 64, Nr. 1, S. 37–46, Jan. 1988, doi: 10.1111/j.1365-2672.1988.tb02427.x.
  - [28] G. T. Macfarlane, J. H. Cummings, und C. Allison, „Protein Degradation by Human Intestinal Bacteria“, *Microbiology*, Bd. 132, Nr. 6, S. 1647–1656, Juni 1986, doi: 10.1099/00221287-132-6-1647.
  - [29] A. T. Reese u. a., „Microbial nitrogen limitation in the mammalian large intestine“, *Nat Microbiol*, Bd. 3, Nr. 12, S. 1441–1450, Okt. 2018, doi: 10.1038/s41564-018-0267-7.
  - [30] D. Berry u. a., „Host-compound foraging by intestinal microbiota revealed by single-cell stable isotope probing“, *Proc. Natl. Acad. Sci. U.S.A.*, Bd. 110, Nr. 12, S. 4720–4725, März 2013, doi: 10.1073/pnas.1219247110.
  - [31] G. T. Macfarlane, J. H. Cummings, S. Macfarlane, und G. R. Gibson, „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*, Bd. 67, Nr. 5, S. 521–527, Nov. 1989, doi: 10.1111/j.1365-2672.1989.tb02524.x.

- [32] G. T. Macfarlane, S. Hay, und G. R. Gibson, „Influence of mucin on glycosidase, protease and arylamidase activities of human gut bacteria grown in a 3-stage continuous culture system“, *Journal of Applied Bacteriology*, Bd. 66, Nr. 5, S. 407–417, Mai 1989, doi: 10.1111/j.1365-2672.1989.tb05110.x.
- [33] W. R. Russell u. a., „Major phenylpropanoid-derived metabolites in the human gut can arise from microbial fermentation of protein“, *Mol. Nutr. Food Res.*, Bd. 57, Nr. 3, S. 523–535, März 2013, doi: 10.1002/mnfr.201200594.
- [34] D. J. Braccia, X. Jiang, M. Pop, und A. B. Hall, „The Capacity to Produce Hydrogen Sulfide (H<sub>2</sub>S) via Cysteine Degradation Is Ubiquitous in the Human Gut Microbiome“, *Front. Microbiol.*, Bd. 12, S. 705583, Okt. 2021, doi: 10.3389/fmicb.2021.705583.
- [35] P. G. Wolf u. a., „Diversity and distribution of sulfur metabolic genes in the human gut microbiome and their association with colorectal cancer“, *Microbiome*, Bd. 10, Nr. 1, S. 64, Dez. 2022, doi: 10.1186/s40168-022-01242-x.
- [36] L. Wu u. a., „Mutual interaction between gut microbiota and protein/amino acid metabolism for host mucosal immunity and health“, *Animal Nutrition*, Bd. 7, Nr. 1, Art. Nr. 1, März 2021, doi: 10.1016/j.aninu.2020.11.003.
- [37] T.-T. Li u. a., „Microbiota metabolism of intestinal amino acids impacts host nutrient homeostasis and physiology“, *Cell Host & Microbe*, Bd. 32, Nr. 5, Art. Nr. 5, Mai 2024, doi: 10.1016/j.chom.2024.04.004.
- [38] L. Zhai u. a., „Gut microbiota-derived tryptamine and phenethylamine impair insulin sensitivity in metabolic syndrome and irritable bowel syndrome“, *Nat Commun*, Bd. 14, Nr. 1, S. 4986, Aug. 2023, doi: 10.1038/s41467-023-40552-y.
- [39] H. K. Pedersen u. a., „Human gut microbes impact host serum metabolome and insulin sensitivity“, *Nature*, Bd. 535, Nr. 7612, S. 376–381, Juli 2016, doi: 10.1038/nature18646.
- [40] A. R. Watson u. a., „Metabolic independence drives gut microbial colonization and resilience in health and disease“, *Genome Biol*, Bd. 24, Nr. 1, S. 78, Apr. 2023, doi: 10.1186/s13059-023-02924-x.
- [41] A. Agus, J. Planchais, und H. Sokol, „Gut Microbiota Regulation of Tryptophan Metabolism in Health and Disease“, *Cell Host & Microbe*, Bd. 23, Nr. 6, S. 716–724, Juni 2018, doi: 10.1016/j.chom.2018.05.003.
- [42] B. Lamas u. a., „CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands“, *Nat Med*, Bd. 22, Nr. 6, S. 598–605, Juni 2016, doi: 10.1038/nm.4102.
- [43] S. Nikolaus u. a., „Increased Tryptophan Metabolism Is Associated With Activity of Inflammatory Bowel Diseases“, *Gastroenterology*, Bd. 153, Nr. 6, S. 1504–1516.e2, Dez. 2017, doi: 10.1053/j.gastro.2017.08.028.
- [44] D. M. Harris u. a., „Tryptophan wasting and disease activity as a systems phenomenon in inflammation – an analysis across 13 chronic inflammatory diseases“, *Immunology*, preprint, Aug. 2023. doi: 10.1101/2023.08.25.554383.
- [45] C. Chojnacki, T. Popławski, P. Konrad, M. Fila, J. Błasiak, und J. Chojnacki, „Antimicrobial treatment improves tryptophan metabolism and mood of patients with small intestinal bacterial overgrowth“, *Nutr Metab (Lond)*, Bd. 19, Nr. 1, S. 66, Sep. 2022, doi: 10.1186/s12986-022-00700-5.
- [46] C. Chojnacki u. a., „Altered Tryptophan Metabolism on the Kynurenine Pathway in Depressive Patients with Small Intestinal Bacterial Overgrowth“, *Nutrients*, Bd. 14, Nr. 15, S. 3217, Aug. 2022, doi: 10.3390/nu14153217.
- [47] S. Moran-Ramos u. a., „A higher bacterial inward BCAA transport driven by *Faecalibacterium prausnitzii* is associated with lower serum levels of BCAA in

- early adolescents“, *Mol Med*, Bd. 27, Nr. 1, S. 108, Dez. 2021, doi: 10.1186/s10020-021-00371-7.
- [48] A. Darra u. a., „Hyperglycemia is associated with duodenal dysbiosis and altered duodenal microenvironment“, *Sci Rep*, Bd. 13, Nr. 1, S. 11038, Juli 2023, doi: 10.1038/s41598-023-37720-x.
- [49] R. Liu u. a., „Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention“, *Nat Med*, Bd. 23, Nr. 7, S. 859–868, Juli 2017, doi: 10.1038/nm.4358.
- [50] E. Angelakis u. a., „A Metagenomic Investigation of the Duodenal Microbiota Reveals Links with Obesity“, *PLoS ONE*, Bd. 10, Nr. 9, S. e0137784, Sep. 2015, doi: 10.1371/journal.pone.0137784.
- [51] I. Morgenstern, M. T. M. Raijmakers, W. H. M. Peters, H. Hoensch, und W. Kirch, „Homocysteine, Cysteine, and Glutathione in Human Colonic Mucosa: Elevated Levels of Homocysteine in Patients with Inflammatory Bowel Disease“, *Digestive Diseases and Sciences*, Bd. 48, Nr. 10, Art. Nr. 10, 2003, doi: 10.1023/A:1026338812708.
- [52] P. Du u. a., „Role of exhaled hydrogen sulfide in the diagnosis of colorectal cancer“, *BMJ Open Gastroenterol*, Bd. 11, Nr. 1, S. e001229, Feb. 2024, doi: 10.1136/bmjgast-2023-001229.
- [53] A. G. Buret, T. Allain, J.-P. Motta, und J. L. Wallace, „Effects of Hydrogen Sulfide on the Microbiome: From Toxicity to Therapy“, *Antioxidants & Redox Signaling*, Bd. 36, Nr. 4–6, S. 211–219, Feb. 2022, doi: 10.1089/ars.2021.0004.
- [54] M. J. Villanueva-Millan u. a., „Methanogens and Hydrogen Sulfide Producing Bacteria Guide Distinct Gut Microbe Profiles and Irritable Bowel Syndrome Subtypes“, *Am J Gastroenterol*, Bd. 117, Nr. 12, S. 2055–2066, Dez. 2022, doi: 10.14309/ajg.0000000000001997.
- [55] X. C. Morgan u. a., „Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment“, *Genome Biol*, Bd. 13, Nr. 9, Art. Nr. 9, 2012, doi: 10.1186/gb-2012-13-9-r79.
- [56] C. A. Lozupone, J. I. Stombaugh, J. I. Gordon, J. K. Jansson, und R. Knight, „Diversity, stability and resilience of the human gut microbiota“, *Nature*, Bd. 489, Nr. 7415, Art. Nr. 7415, Sep. 2012, doi: 10.1038/nature11550.
- [57] M. Fassarella, E. E. Blaak, J. Penders, A. Nauta, H. Smidt, und E. G. Zoetendal, „Gut microbiome stability and resilience: elucidating the response to perturbations in order to modulate gut health“, *Gut*, Bd. 70, Nr. 3, S. 595–605, März 2021, doi: 10.1136/gutjnl-2020-321747.
- [58] V. R. Marcelino u. a., „Disease-specific loss of microbial cross-feeding interactions in the human gut“, *Nat Commun*, Bd. 14, Nr. 1, Art. Nr. 1, Okt. 2023, doi: 10.1038/s41467-023-42112-w.
- [59] Y. Fan und O. Pedersen, „Gut microbiota in human metabolic health and disease“, *Nat Rev Microbiol*, Bd. 19, Nr. 1, Art. Nr. 1, Jan. 2021, doi: 10.1038/s41579-020-0433-9.
- [60] B. A. Daisley u. a., „Emerging connections between gut microbiome bioenergetics and chronic metabolic diseases“, *Cell Reports*, Bd. 37, Nr. 10, S. 110087, Dez. 2021, doi: 10.1016/j.celrep.2021.110087.
- [61] K. Aden u. a., „Metabolic Functions of Gut Microbes Associate With Efficacy of Tumor Necrosis Factor Antagonists in Patients With Inflammatory Bowel Diseases“, *Gastroenterology*, Bd. 157, Nr. 5, S. 1279-1292.e11, Nov. 2019, doi: 10.1053/j.gastro.2019.07.025.



- [62] F. Spragge *u. a.*, „Microbiome diversity protects against pathogens by nutrient blocking“, *Science*, Bd. 382, Nr. 6676, S. eadj3502, Dez. 2023, doi: 10.1126/science.adj3502.
- [63] T. Wang, A. Goyal, V. Dubinkina, und S. Maslov, „Evidence for a multi-level trophic organization of the human gut microbiome“, *PLoS Comput Biol*, Bd. 15, Nr. 12, S. e1007524, Dez. 2019, doi: 10.1371/journal.pcbi.1007524.
- [64] F. Bäckhed *u. a.*, „Defining a Healthy Human Gut Microbiome: Current Concepts, Future Directions, and Clinical Applications“, *Cell Host & Microbe*, Bd. 12, Nr. 5, S. 611–622, Nov. 2012, doi: 10.1016/j.chom.2012.10.012.
- [65] J. A. Gilbert und S. V. Lynch, „Community ecology as a framework for human microbiome research“, *Nat Med*, Bd. 25, Nr. 6, S. 884–889, Juni 2019, doi: 10.1038/s41591-019-0464-9.
- [66] A. T. Neu  
, E. E. Allen, und K. Roy, „Defining and quantifying the core microbiome: Challenges and prospects“, *Proc. Natl. Acad. Sci. U.S.A.*, Bd. 118, Nr. 51, S. e2104429118, Dez. 2021, doi: 10.1073/pnas.2104429118.
- [67] E. Neis, C. Dejong, und S. Rensen, „The Role of Microbial Amino Acid Metabolism in Host Metabolism“, *Nutrients*, Bd. 7, Nr. 4, S. 2930–2946, Apr. 2015, doi: 10.3390/nu7042930.
- [68] J. Ramoneda, T. B. N. Jensen, M. N. Price, E. O. Casamayor, und N. Fierer, „Taxonomic and environmental distribution of bacterial amino acid auxotrophies“, *Nat Commun*, Bd. 14, Nr. 1, S. 7608, Nov. 2023, doi: 10.1038/s41467-023-43435-4.
- [69] G. D’Souza, S. Waschina, S. Pande, K. Bohl, C. Kaleta, und C. Kost, „LESS IS MORE: SELECTIVE ADVANTAGES CAN EXPLAIN THE PREVALENT LOSS OF BIOSYNTHETIC GENES IN BACTERIA: ADAPTIVE LOSS OF BIOSYNTHETIC GENES IN BACTERIA“, *Evolution*, Bd. 68, Nr. 9, S. 2559–2570, Sep. 2014, doi: 10.1111/evo.12468.
- [70] R. F. Taylor, M. E. Hodson, und T. L. Pitt, „Adult cystic fibrosis: association of acute pulmonary exacerbations and increasing severity of lung disease with auxotrophic mutants of *Pseudomonas aeruginosa*.“, *Thorax*, Bd. 48, Nr. 10, Art. Nr. 10, Okt. 1993, doi: 10.1136/thx.48.10.1002.
- [71] S. R. Thomas, „Increased sputum amino acid concentrations and auxotrophy of *Pseudomonas aeruginosa* in severe cystic fibrosis lung disease“, *Thorax*, Bd. 55, Nr. 9, Art. Nr. 9, Sep. 2000, doi: 10.1136/thorax.55.9.795.
- [72] E. C. Soto-Martin *u. a.*, „Vitamin Biosynthesis by Human Gut Butyrate-Producing Bacteria and Cross-Feeding in Synthetic Microbial Communities“, *mBio*, Bd. 11, Nr. 4, S. e00886-20, Aug. 2020, doi: 10.1128/mBio.00886-20.
- [73] V. Sharma *u. a.*, „B-Vitamin Sharing Promotes Stability of Gut Microbial Communities“, *Front. Microbiol.*, Bd. 10, S. 1485, Juli 2019, doi: 10.3389/fmicb.2019.01485.
- [74] J. P. Molina Ortiz, M. N. Read, D. D. McClure, A. Holmes, F. Dehghani, und E. R. Shanahan, „High throughput genome scale modeling predicts microbial vitamin requirements contribute to gut microbiome community structure“, *Gut Microbes*, Bd. 14, Nr. 1, S. 2118831, Dez. 2022, doi: 10.1080/19490976.2022.2118831.
- [75] D. A. Rodionov *u. a.*, „Micronutrient Requirements and Sharing Capabilities of the Human Gut Microbiome“, *Front. Microbiol.*, Bd. 10, S. 1316, Juni 2019, doi: 10.3389/fmicb.2019.01316.

- [76] M. Zhao *u. a.*, „Tumor-targeting bacterial therapy with amino acid auxotrophs of GFP-expressing *Salmonella typhimurium*“, *Proc. Natl. Acad. Sci. U.S.A.*, Bd. 102, Nr. 3, S. 755–760, Jan. 2005, doi: 10.1073/pnas.0408422102.
- [77] V. Jawalagatti, P. Kirthika, und J. H. Lee, „Targeting primary and metastatic tumor growth in an aggressive breast cancer by engineered tryptophan auxotrophic *Salmonella Typhimurium*“, *Molecular Therapy - Oncolytics*, Bd. 25, S. 350–363, Juni 2022, doi: 10.1016/j.omto.2022.05.004.
- [78] F. Bertels, H. Merker, und C. Kost, „Design and Characterization of Auxotrophy-Based Amino Acid Biosensors“, *PLoS ONE*, Bd. 7, Nr. 7, S. e41349, Juli 2012, doi: 10.1371/journal.pone.0041349.
- [79] G. A. Ashniev, S. N. Petrov, S. N. Iablokov, und D. A. Rodionov, „Genomics-Based Reconstruction and Predictive Profiling of Amino Acid Biosynthesis in the Human Gut Microbiome“, *Microorganisms*, Bd. 10, Nr. 4, Art. Nr. 4, März 2022, doi: 10.3390/microorganisms10040740.
- [80] S. Magnúsdóttir, D. Ravcheev, V. De Crécy-Lagard, und I. Thiele, „Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes“, *Front. Genet.*, Bd. 6, Apr. 2015, doi: 10.3389/fgene.2015.00148.
- [81] C. Tarracchini *u. a.*, „Exploring the vitamin biosynthesis landscape of the human gut microbiota“, *mSystems*, Bd. 9, Nr. 10, S. e0092924, Okt. 2024, doi: 10.1128/msystems.00929-24.
- [82] Y. Seif, K. S. Choudhary, Y. Hefner, A. Anand, L. Yang, und B. O. Palsson, „Metabolic and genetic basis for auxotrophies in Gram-negative species“, *Proc Natl Acad Sci USA*, Bd. 117, Nr. 11, S. 6264–6273, März 2020, doi: 10.1073/pnas.1910499117.
- [83] A. L. Barth und T. L. Pitt, „Auxotrophic variants of *Pseudomonas aeruginosa* are selected from prototrophic wild-type strains in respiratory infections in patients with cystic fibrosis“, *J Clin Microbiol*, Bd. 33, Nr. 1, S. 37–40, Jan. 1995, doi: 10.1128/jcm.33.1.37-40.1995.
- [84] A. L. Barth und T. L. Pitt, „The high amino-acid content of sputum from cystic fibrosis patients promotes growth of auxotrophic *Pseudomonas aeruginosa*“, *Journal of Medical Microbiology*, Bd. 45, Nr. 2, S. 110–119, Aug. 1996, doi: 10.1099/00222615-45-2-110.
- [85] R. A. Quinn *u. a.*, „Neutrophilic proteolysis in the cystic fibrosis lung correlates with a pathogenic microbiome“, *Microbiome*, Bd. 7, Nr. 1, S. 23, Dez. 2019, doi: 10.1186/s40168-019-0636-3.
- [86] E. Bosi, J. M. Monk, R. K. Aziz, M. Fondi, V. Nizet, und B. Ø. Palsson, „Comparative genome-scale modelling of *Staphylococcus aureus* strains identifies strain-specific metabolic capabilities linked to pathogenicity“, *Proc. Natl. Acad. Sci. U.S.A.*, Bd. 113, Nr. 26, Juni 2016, doi: 10.1073/pnas.1523199113.
- [87] A. S. Nuxoll *u. a.*, „CcpA Regulates Arginine Biosynthesis in *Staphylococcus aureus* through Repression of Proline Catabolism“, *PLoS Pathog*, Bd. 8, Nr. 11, S. e1003033, Nov. 2012, doi: 10.1371/journal.ppat.1003033.
- [88] G. D’Souza und C. Kost, „Experimental Evolution of Metabolic Dependency in Bacteria“, *PLoS Genet*, Bd. 12, Nr. 11, S. e1006364, Nov. 2016, doi: 10.1371/journal.pgen.1006364.
- [89] A. K. Hottes, P. L. Freddolino, A. Khare, Z. N. Donnell, J. C. Liu, und S. Tavazoie, „Bacterial Adaptation through Loss of Function“, *PLoS Genet*, Bd. 9, Nr. 7, S. e1003617, Juli 2013, doi: 10.1371/journal.pgen.1003617.

- [90] J. J. Morris, R. E. Lenski, und E. R. Zinser, „The Black Queen Hypothesis: Evolution of Dependencies through Adaptive Gene Loss“, *mBio*, Bd. 3, Nr. 2, S. e00036-12, Mai 2012, doi: 10.1128/mBio.00036-12.
- [91] C. Kaleta, S. Schäuble, U. Rinas, und S. Schuster, „Metabolic costs of amino acid and protein production in *Escherichia coli*“, *Biotechnology Journal*, Bd. 8, Nr. 9, S. 1105–1114, Sep. 2013, doi: 10.1002/biot.201200267.
- [92] G. D’Souza, S. Waschina, C. Kaleta, und C. Kost, „Plasticity and epistasis strongly affect bacterial fitness after losing multiple metabolic genes: PLASTICITY AND EPISTASIS IN AUXOTROPHIC BACTERIA“, *Evolution*, Bd. 69, Nr. 5, S. 1244–1254, Mai 2015, doi: 10.1111/evo.12640.
- [93] D. D. Sriramulu, „Amino Acids Enhance Adaptive Behaviour of *Pseudomonas Aeruginosa* in the Cystic Fibrosis Lung Environment“, *Microbiol Insights*, Bd. 3, S. MBI.S4694, Jan. 2010, doi: 10.4137/MBI.S4694.
- [94] K. C. H. van der Ark, S. Aalvink, M. Suarez-Diez, P. J. Schaap, W. M. de Vos, und C. Belzer, „Model-driven design of a minimal medium for *Akkermansia muciniphila* confirms mucus adaptation“, *Microb. Biotechnol.*, Bd. 11, Nr. 3, S. 476–485, Mai 2018, doi: 10.1111/1751-7915.13033.
- [95] B. Teusink und D. Molenaar, „Systems biology of lactic acid bacteria: For food and thought“, *Current Opinion in Systems Biology*, Bd. 6, S. 7–13, Dez. 2017, doi: 10.1016/j.coisb.2017.07.005.
- [96] S. Waschina, G. D’Souza, C. Kost, und C. Kaleta, „Metabolic network architecture and carbon source determine metabolite production costs“, *FEBS J*, Bd. 283, Nr. 11, S. 2149–2163, Juni 2016, doi: 10.1111/febs.13727.
- [97] C.-H. Kuo, N. A. Moran, und H. Ochman, „The consequences of genetic drift for bacterial genome complexity“, *Genome Res.*, Bd. 19, Nr. 8, S. 1450–1454, Aug. 2009, doi: 10.1101/gr.091785.109.
- [98] S. Pande *u. a.*, „Fitness and stability of obligate cross-feeding interactions that emerge upon gene loss in bacteria“, *ISME J*, Bd. 8, Nr. 5, S. 953–962, Mai 2014, doi: 10.1038/ismej.2013.211.
- [99] R. K. Fritts, A. L. McCully, und J. B. McKinlay, „Extracellular Metabolism Sets the Table for Microbial Cross-Feeding“, *Microbiol Mol Biol Rev*, Bd. 85, Nr. 1, S. e00135-20, Feb. 2021, doi: 10.1128/MMBR.00135-20.
- [100] S. Giri, S. Waschina, C. Kaleta, und C. Kost, „Defining Division of Labor in Microbial Communities“, *Journal of Molecular Biology*, Bd. 431, Nr. 23, S. 4712–4731, Nov. 2019, doi: 10.1016/j.jmb.2019.06.023.
- [101] Y. Liang, A. Ma, und G. Zhuang, „Construction of Environmental Synthetic Microbial Consortia: Based on Engineering and Ecological Principles“, *Front. Microbiol.*, Bd. 13, S. 829717, Feb. 2022, doi: 10.3389/fmicb.2022.829717.
- [102] M. T. Mee, J. J. Collins, G. M. Church, und H. H. Wang, „Syntrophic exchange in synthetic microbial communities“, *Proc. Natl. Acad. Sci. U.S.A.*, Bd. 111, Nr. 20, Mai 2014, doi: 10.1073/pnas.1405641111.
- [103] L. Lin, R. Du, Y. Wang, Q. Wu, und Y. Xu, „Regulation of auxotrophic lactobacilli growth by amino acid cross-feeding interaction“, *International Journal of Food Microbiology*, Bd. 377, S. 109769, Sep. 2022, doi: 10.1016/j.ijfoodmicro.2022.109769.
- [104] S. F. Henriques *u. a.*, „Metabolic cross-feeding in imbalanced diets allows gut microbes to improve reproduction and alter host behaviour“, *Nat Commun*, Bd. 11, Nr. 1, S. 4236, Dez. 2020, doi: 10.1038/s41467-020-18049-9.
- [105] A. Zelezniak, S. Andrejev, O. Ponomarova, D. R. Mende, P. Bork, und K. R. Patil, „Metabolic dependencies drive species co-occurrence in diverse microbial

- communities“, *Proc. Natl. Acad. Sci. U.S.A.*, Bd. 112, Nr. 20, S. 6449–6454, Mai 2015, doi: 10.1073/pnas.1421834112.
- [106] S. Germerodt u. a., „Pervasive Selection for Cooperative Cross-Feeding in Bacterial Communities“, *PLoS Comput Biol*, Bd. 12, Nr. 6, S. e1004986, Juni 2016, doi: 10.1371/journal.pcbi.1004986.
- [107] L. Oña, S. Giri, N. Avermann, M. Kreienbaum, K. M. Thormann, und C. Kost, „Obligate cross-feeding expands the metabolic niche of bacteria“, *Nat Ecol Evol*, Bd. 5, Nr. 9, S. 1224–1232, Sep. 2021, doi: 10.1038/s41559-021-01505-0.
- [108] K. S. McCann, „The diversity–stability debate“, *Nature*, Bd. 405, Nr. 6783, S. 228–233, Mai 2000, doi: 10.1038/35012234.
- [109] L. Oña und C. Kost, „Cooperation increases robustness to ecological disturbance in microbial cross-feeding networks“, *Ecology Letters*, Bd. 25, Nr. 6, S. 1410–1420, Juni 2022, doi: 10.1111/ele.14006.
- [110] X. Wan, Q. Yang, X. Wang, Y. Bai, und Z. Liu, „Isolation and Cultivation of Human Gut Microorganisms: A Review“, *Microorganisms*, Bd. 11, Nr. 4, S. 1080, Apr. 2023, doi: 10.3390/microorganisms11041080.
- [111] J. Zimmermann, C. Kaleta, und S. Waschina, „gapseq: informed prediction of bacterial metabolic pathways and reconstruction of accurate metabolic models“, *Genome Biol*, Bd. 22, Nr. 1, S. 81, Dez. 2021, doi: 10.1186/s13059-021-02295-1.
- [112] D. Machado, S. Andrejev, M. Tramontano, und K. R. Patil, „Fast automated reconstruction of genome-scale metabolic models for microbial species and communities“, *Nucleic Acids Research*, Bd. 46, Nr. 15, S. 7542–7553, Sep. 2018, doi: 10.1093/nar/gky537.
- [113] A. Heinken u. a., „Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine“, *Nat Biotechnol*, Jan. 2023, doi: 10.1038/s41587-022-01628-0.
- [114] M. N. Price, A. M. Deutschbauer, und A. P. Arkin, „GapMind: Automated Annotation of Amino Acid Biosynthesis“, *mSystems*, Bd. 5, Nr. 3, S. e00291-20, Juni 2020, doi: 10.1128/mSystems.00291-20.
- [115] J. M. Monk u. a., „Genome-scale metabolic reconstructions of multiple *Escherichia coli* strains highlight strain-specific adaptations to nutritional environments“, *Proc. Natl. Acad. Sci. U.S.A.*, Bd. 110, Nr. 50, S. 20338–20343, Dez. 2013, doi: 10.1073/pnas.1307797110.
- [116] N. Veith u. a., „Using a Genome-Scale Metabolic Model of *Enterococcus faecalis* V583 To Assess Amino Acid Uptake and Its Impact on Central Metabolism“, *Appl Environ Microbiol*, Bd. 81, Nr. 5, S. 1622–1633, März 2015, doi: 10.1128/AEM.03279-14.
- [117] L. Heirendt u. a., „Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0“, *Nat Protoc*, Bd. 14, Nr. 3, S. 639–702, März 2019, doi: 10.1038/s41596-018-0098-2.
- [118] P. D. Karp, D. Weaver, und M. Latendresse, „How accurate is automated gap filling of metabolic models?“, *BMC Syst Biol*, Bd. 12, Nr. 1, S. 73, Dez. 2018, doi: 10.1186/s12918-018-0593-7.
- [119] B. Borer und S. Magnúsdóttir, „The media composition as a crucial element in high-throughput metabolic network reconstruction“, *Interface Focus*, Bd. 13, Nr. 2, S. 20220070, Apr. 2023, doi: 10.1098/rsfs.2022.0070.
- [120] J. D. Orth, I. Thiele, und B. Ø. Palsson, „What is flux balance analysis?“, *Nat Biotechnol*, Bd. 28, Nr. 3, S. 245–248, März 2010, doi: 10.1038/nbt.1614.

- [121] J. S. L. Yu *u. a.*, „Microbial communities form rich extracellular metabolomes that foster metabolic interactions and promote drug tolerance“, *Nature Microbiology*, Bd. 7, Nr. 4, S. 542–555, Apr. 2022, doi: 10.1038/s41564-022-01072-5.
- [122] D. Machado *u. a.*, „Polarization of microbial communities between competitive and cooperative metabolism“, *Nat Ecol Evol*, Bd. 5, Nr. 2, S. 195–203, Jan. 2021, doi: 10.1038/s41559-020-01353-4.
- [123] J. Müller, M. Beckers, N. Mußmann, J. Bongaerts, und J. Büchs, „Elucidation of auxotrophic deficiencies of *Bacillus pumilus* DSM 18097 to develop a defined minimal medium“, *Microb Cell Fact*, Bd. 17, Nr. 1, Art. Nr. 1, Dez. 2018, doi: 10.1186/s12934-018-0956-1.
- [124] J. Zhang und R. Greasham, „Chemically defined media for commercial fermentations“, *Applied Microbiology and Biotechnology*, Bd. 51, Nr. 4, S. 407–421, Apr. 1999, doi: 10.1007/s002530051411.
- [125] G. Zhang, D. A. Mills, und D. E. Block, „Development of Chemically Defined Media Supporting High-Cell-Density Growth of Lactococci, Enterococci, and Streptococci“, *Appl Environ Microbiol*, Bd. 75, Nr. 4, S. 1080–1087, Feb. 2009, doi: 10.1128/AEM.01416-08.
- [126] E. M. H. ~~obert~~, R. R. Raya, und G. S. D. Giori, „Nutritional Requirements of *Lactobacillus delbrueckii* subsp. *lactis* in a Chemically Defined Medium“, *Curr Microbiol*, Bd. 49, Nr. 5, Art. Nr. 5, Nov. 2004, doi: 10.1007/s00284-004-4357-9.
- [127] P. R. Jensen und K. Hammer, „Minimal Requirements for Exponential Growth of *Lactococcus lactis*“, *Appl Environ Microbiol*, Bd. 59, Nr. 12, Art. Nr. 12, Dez. 1993, doi: 10.1128/aem.59.12.4363-4366.1993.
- [128] N. Tejera *u. a.*, „Genome-Scale Metabolic Model Driven Design of a Defined Medium for *Campylobacter jejuni* M1cam“, *Front. Microbiol.*, Bd. 11, S. 1072, Juni 2020, doi: 10.3389/fmicb.2020.01072.
- [129] P. Louis, G. L. Hold, und H. J. Flint, „The gut microbiota, bacterial metabolites and colorectal cancer“, *Nat Rev Microbiol*, Bd. 12, Nr. 10, S. 661–672, Okt. 2014, doi: 10.1038/nrmicro3344.
- [130] F. Hinrichsen *u. a.*, „Microbial regulation of hexokinase 2 links mitochondrial metabolism and cell death in colitis“, *Cell Metabolism*, Bd. 33, Nr. 12, S. 2355–2366.e8, Dez. 2021, doi: 10.1016/j.cmet.2021.11.004.
- [131] D. Dodd *u. a.*, „A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites“, *Nature*, Bd. 551, Nr. 7682, S. 648–652, Nov. 2017, doi: 10.1038/nature24661.
- [132] J. Tintelnot *u. a.*, „Microbiota-derived 3-IAA influences chemotherapy efficacy in pancreatic cancer“, *Nature*, Bd. 615, Nr. 7950, S. 168–174, März 2023, doi: 10.1038/s41586-023-05728-y.
- [133] M. Effenberger *u. a.*, „Microbial Butyrate Synthesis Indicates Therapeutic Efficacy of Azathioprine in IBD Patients“, *Journal of Crohn's and Colitis*, Bd. 15, Nr. 1, S. 88–98, Jan. 2021, doi: 10.1093/ecco-jcc/jjaa152.
- [134] R. Pryor *u. a.*, „Host-Microbe-Drug-Nutrient Screen Identifies Bacterial Effectors of Metformin Therapy“, *Cell*, Bd. 178, Nr. 6, S. 1299–1312.e29, Sep. 2019, doi: 10.1016/j.cell.2019.08.003.
- [135] A. Heinken, M. T. Khan, G. Paglia, D. A. Rodionov, H. J. M. Harmsen, und I. Thiele, „Functional Metabolic Map of *Faecalibacterium prausnitzii*, a Beneficial Human Gut Microbe“, *J Bacteriol*, Bd. 196, Nr. 18, S. 3289–3302, Sep. 2014, doi: 10.1128/JB.01780-14.
- [136] J. C. Xavier, K. R. Patil, und I. Rocha, „Integration of Biomass Formulations of Genome-Scale Metabolic Models with Experimental Data Reveals Universally

- Essential Cofactors in Prokaryotes“, *Metabolic Engineering*, Bd. 39, S. 200–208, Jan. 2017, doi: 10.1016/j.ymben.2016.12.002.
- [137] M. Tramontano *u. a.*, „Nutritional preferences of human gut bacteria reveal their metabolic idiosyncrasies“, *Nat Microbiol*, Bd. 3, Nr. 4, S. 514–522, Apr. 2018, doi: 10.1038/s41564-018-0123-9.
- [138] P. R. Kiela und F. K. Ghishan, „Physiology of Intestinal Absorption and Secretion“, *Best Practice & Research Clinical Gastroenterology*, Bd. 30, Nr. 2, S. 145–159, Apr. 2016, doi: 10.1016/j.bpg.2016.02.007.
- [139] L. Dethlefsen, M. McFall-Ngai, und D. A. Relman, „An ecological and evolutionary perspective on human–microbe mutualism and disease“, *Nature*, Bd. 449, Nr. 7164, S. 811–818, Okt. 2007, doi: 10.1038/nature06245.
- [140] A. S. Weiss *u. a.*, „In vitro interaction network of a synthetic gut bacterial community“, *ISME J*, S. 1–15, Dez. 2021, doi: 10.1038/s41396-021-01153-z.
- [141] S. Van Vliet, C. Hauert, K. Fridberg, M. Ackermann, und A. Dal Co, „Global dynamics of microbial communities emerge from local interaction rules“, *PLoS Comput Biol*, Bd. 18, Nr. 3, S. e1009877, März 2022, doi: 10.1371/journal.pcbi.1009877.
- [142] B. E. L. Morris, R. Henneberger, H. Huber, und C. Moissl-Eichinger, „Microbial syntrophy: interaction for the common good“, *FEMS Microbiol Rev*, Bd. 37, Nr. 3, S. 384–406, Mai 2013, doi: 10.1111/1574-6976.12019.
- [143] A. Rivière, M. Gagnon, S. Weckx, D. Roy, und L. De Vuyst, „Mutual Cross-Feeding Interactions between *Bifidobacterium longum* subsp. *longum* NCC2705 and *Eubacterium rectale* ATCC 33656 Explain the Bifidogenic and Butyrogenic Effects of Arabinoxylan Oligosaccharides“, *Appl Environ Microbiol*, Bd. 81, Nr. 22, S. 7767–7781, Nov. 2015, doi: 10.1128/AEM.02089-15.
- [144] S. A. Shetty, I. Kostopoulos, S. Y. Geerlings, H. Smidt, W. M. de Vos, und C. Belzer, „Dynamic metabolic interactions and trophic roles of human gut microbes identified using a minimal microbiome exhibiting ecological properties“, *ISME J*, Bd. 16, Nr. 9, S. 2144–2159, Sep. 2022, doi: 10.1038/s41396-022-01255-2.
- [145] S. Giri *u. a.*, „Metabolic dissimilarity determines the establishment of cross-feeding interactions in bacteria“, *Current Biology*, Bd. 31, Nr. 24, S. 5547–5557.e6, Dez. 2021, doi: 10.1016/j.cub.2021.10.019.
- [146] C. Y. Kim *u. a.*, „Human reference gut microbiome catalog including newly assembled genomes from under-represented Asian metagenomes“, *Genome Med*, Bd. 13, Nr. 1, S. 134, Dez. 2021, doi: 10.1186/s13073-021-00950-7.
- [147] M. Price, „Erroneous predictions of auxotrophies by CarveMe“, *Nat Ecol Evol*, Bd. 7, Nr. 2, Art. Nr. 2, Feb. 2023, doi: 10.1038/s41559-022-01936-3.
- [148] R. Caspi *u. a.*, „The MetaCyc database of metabolic pathways and enzymes“, *Nucleic Acids Research*, Bd. 46, Nr. D1, S. D633–D639, Jan. 2018, doi: 10.1093/nar/gkx935.
- [149] M. Davenport *u. a.*, „Metabolic Alterations to the Mucosal Microbiota in Inflammatory Bowel Disease“, *Inflammatory Bowel Diseases*, Bd. 20, Nr. 4, S. 723–731, Apr. 2014, doi: 10.1097/MIB.000000000000011.
- [150] A. Troci, P. Rausch, S. Waschina, W. Lieb, A. Franke, und C. Bang, „Long-Term Dietary Effects on Human Gut Microbiota Composition Employing Shotgun Metagenomics Data Analysis“, *Molecular Nutrition Food Res*, S. 2101098, Juli 2022, doi: 10.1002/mnfr.202101098.
- [151] L. Chen *u. a.*, „The long-term genetic stability and individual specificity of the human gut microbiome“, *Cell*, Bd. 184, Nr. 9, S. 2302–2315.e12, Apr. 2021, doi: 10.1016/j.cell.2021.03.024.

- [152] G. D'Souza, S. Shitut, D. Preussger, G. Yousif, S. Waschina, und C. Kost, „Ecology and evolution of metabolic cross-feeding interactions in bacteria“, *Nat. Prod. Rep.*, Bd. 35, Nr. 5, S. 455–488, 2018, doi: 10.1039/C8NP00009C.
- [153] S. L. Garcia, M. Buck, K. D. McMahon, H.-P. Grossart, A. Eiler, und F. Warnecke, „Auxotrophy and intrapopulation complementary in the ‘interactome’ of a cultivated freshwater model community“, *Mol Ecol*, Bd. 24, Nr. 17, S. 4449–4459, Sep. 2015, doi: 10.1111/mec.13319.
- [154] S. Pande u. a., „Metabolic cross-feeding via intercellular nanotubes among bacteria“, *Nat Commun*, Bd. 6, Nr. 1, S. 6238, Feb. 2015, doi: 10.1038/ncomms7238.
- [155] N. van der Wielen, P. J. Moughan, und M. Mensink, „Amino Acid Absorption in the Large Intestine of Humans and Porcine Models“, *J. Nutr.*, Bd. 147, Nr. 8, S. 1493–1498, Aug. 2017, doi: 10.3945/jn.117.248187.
- [156] S. Han u. a., „A metabolomics pipeline for the mechanistic interrogation of the gut microbiome“, *Nature*, Bd. 595, Nr. 7867, Art. Nr. 7867, Juli 2021, doi: 10.1038/s41586-021-03707-9.
- [157] B. Sánchez u. a., „Low-pH Adaptation and the Acid Tolerance Response of *Bifidobacterium longum* Biotype *longum*“, *Applied and Environmental Microbiology*, Bd. 73, Nr. 20, S. 6450–6459, Okt. 2007, doi: 10.1128/AEM.00886-07.
- [158] R. Bansil und B. S. Turner, „The biology of mucus: Composition, synthesis and organization“, *Advanced Drug Delivery Reviews*, Bd. 124, S. 3–15, Jan. 2018, doi: 10.1016/j.addr.2017.09.023.
- [159] J. U. Scher u. a., „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*, Bd. 67, Nr. 1, S. 128–139, Jan. 2015, doi: 10.1002/art.38892.
- [160] P. Lepage u. a., „Twin Study Indicates Loss of Interaction Between Microbiota and Mucosa of Patients With Ulcerative Colitis“, *Gastroenterology*, Bd. 141, Nr. 1, S. 227–236, Juli 2011, doi: 10.1053/j.gastro.2011.04.011.
- [161] D. Fangmann u. a., „Targeted Microbiome Intervention by Microencapsulated Delayed-Release Niacin Beneficially Affects Insulin Sensitivity in Humans“, *Diabetes Care*, Bd. 41, Nr. 3, S. 398–405, März 2018, doi: 10.2337/dc17-1967.
- [162] D. Gonze, K. Z. Coyte, L. Lahti, und K. Faust, „Microbial communities as dynamical systems“, *Current Opinion in Microbiology*, Bd. 44, S. 41–49, Aug. 2018, doi: 10.1016/j.mib.2018.07.004.
- [163] K. Z. Coyte, J. Schluter, und K. R. Foster, „The ecology of the microbiome: Networks, competition, and stability“, *Science*, Bd. 350, Nr. 6261, S. 663–666, Nov. 2015, doi: 10.1126/science.aad2602.
- [164] J. D. Palmer und K. R. Foster, „Bacterial species rarely work together“, *Science*, Bd. 376, Nr. 6593, S. 581–582, Mai 2022, doi: 10.1126/science.abn5093.
- [165] F. Brial u. a., „Human and preclinical studies of the host–gut microbiome co-metabolite hippurate as a marker and mediator of metabolic health“, *Gut*, Bd. 70, Nr. 11, S. 2105–2114, Nov. 2021, doi: 10.1136/gutjnl-2020-323314.
- [166] Y. Saito, T. Sato, K. Nomoto, und H. Tsuji, „Identification of phenol- and p-cresol-producing intestinal bacteria by using media supplemented with tyrosine and its metabolites“, *FEMS Microbiology Ecology*, Bd. 94, Nr. 9, S. fiy125, Sep. 2018, doi: 10.1093/femsec/fiy125.
- [167] B. Lamas, J. M. Natividad, und H. Sokol, „Aryl hydrocarbon receptor and intestinal immunity“, *Mucosal Immunology*, Bd. 11, Nr. 4, S. 1024–1038, Juli 2018, doi: 10.1038/s41385-018-0019-2.

- [168] A. Almeida *u. a.*, „A unified catalog of 204,938 reference genomes from the human gut microbiome“, *Nat Biotechnol*, Bd. 39, Nr. 1, S. 105–114, Jan. 2021, doi: 10.1038/s41587-020-0603-3.
- [169] C. S. Henry, M. DeJongh, A. A. Best, P. M. Frybarger, B. Lindsay, und R. L. Stevens, „High-throughput generation, optimization and analysis of genome-scale metabolic models“, *Nat Biotechnol*, Bd. 28, Nr. 9, S. 977–982, Sep. 2010, doi: 10.1038/nbt.1672.
- [170] G. Gelius-Dietrich, A. A. Desouki, C. J. Fritzemeier, und M. J. Lercher, „sybil – Efficient constraint-based modelling in R“, *BMC Syst Biol*, Bd. 7, Nr. 1, S. 125, Dez. 2013, doi: 10.1186/1752-0509-7-125.
- [171] H. Boeing, J. Wahrendorf, und N. Becker, „EPIC-Germany – A Source for Studies into Diet and Risk of Chronic Diseases“, *Annals of Nutrition and Metabolism*, Bd. 43, Nr. 4, S. 195–204, Dez. 1999, doi: 10.1159/000012786.
- [172] L. B. Thingholm *u. a.*, „Ecology impacts the decrease of Spirochaetes and Prevotella in the fecal gut microbiota of urban humans“, *BMC Microbiol*, Bd. 21, Nr. 1, S. 276, Okt. 2021, doi: 10.1186/s12866-021-02337-5.
- [173] S. Kieser, J. Brown, E. M. Zdobnov, M. Trajkovski, und L. A. McCue, „ATLAS: a Snakemake workflow for assembly, annotation, and genomic binning of metagenome sequence data“, *BMC Bioinformatics*, Bd. 21, Nr. 1, S. 257, Dez. 2020, doi: 10.1186/s12859-020-03585-4.
- [174] B. J. Woodcroft, *CoverM*. (18. August 2023). Rust. Zugegriffen: 20. August 2023. [Online]. Verfügbar unter: <https://github.com/wwood/CoverM>
- [175] Y. Benjamini, A. M. Krieger, und D. Yekutieli, „Adaptive linear step-up procedures that control the false discovery rate“, *Biometrika*, Bd. 93, Nr. 3, S. 491–507, Sep. 2006, doi: 10.1093/biomet/93.3.491.
- [176] C. Lozupone, M. E. Lladser, D. Knights, J. Stombaugh, und R. Knight, „UniFrac: an effective distance metric for microbial community comparison“, *ISME J*, Bd. 5, Nr. 2, S. 169–172, Feb. 2011, doi: 10.1038/ismej.2010.133.
- [177] K. Bittinger, *abdiv: Alpha and Beta Diversity Measures*. (20. Januar 2020). Zugegriffen: 21. August 2023. [Online]. Verfügbar unter: <https://cran.r-project.org/web/packages/abdiv/index.html>
- [178] Jari Oksanen *u. a.*, „vegan: Community Ecology Package“, *R package version 2.6-2*, 2022, [Online]. Verfügbar unter: <https://CRAN.R-project.org/package=vegan>
- [179] S. Waschina, J. Zimmermann, und C. Kaleta, „gapseq reconstructions for 5414 genomes from the HRGM collection“. Zenodo, 30. April 2023. doi: 10.5281/ZENODO.8283247.
- [180] Busche, Svenja und Waschina, Silvio, „gapseq reconstructions for 124 prototrophic genotypes“. Zenodo, 10. Mai 2023. doi: 10.5281/ZENODO.7919571.
- [181] S. Starke und S. Waschina, „gapseq reconstruction for 36 human gut bacterial strains with known amino acid auxotrophy status“. Zenodo, 21. August 2023. doi: 10.5281/ZENODO.8269533.
- [182] S. Yersin und P. Vonaesch, „Small intestinal microbiota: from taxonomic composition to metabolism“, *Trends in Microbiology*, S. S0966842X24000568, März 2024, doi: 10.1016/j.tim.2024.02.013.
- [183] E. Le Chatelier *u. a.*, „Richness of human gut microbiome correlates with metabolic markers“, *Nature*, Bd. 500, Nr. 7464, S. 541–546, Aug. 2013, doi: 10.1038/nature12506.



- [184] E. B. Chang und K. Martinez-Guryn, „Small intestinal microbiota: the neglected stepchild needed for fat digestion and absorption“, *Gut Microbes*, Bd. 10, Nr. 2, S. 235–240, März 2019, doi: 10.1080/19490976.2018.1502539.
- [185] K. Martinez-Guryn u. a., „Small Intestine Microbiota Regulate Host Digestive and Absorptive Adaptive Responses to Dietary Lipids“, *Cell Host & Microbe*, Bd. 23, Nr. 4, S. 458–469.e5, Apr. 2018, doi: 10.1016/j.chom.2018.03.011.
- [186] J. Trommelen, D. Tomé, und L. J. C. van Loon, „Gut amino acid absorption in humans: Concepts and relevance for postprandial metabolism“, *Clinical Nutrition Open Science*, Bd. 36, S. 43–55, Apr. 2021, doi: 10.1016/j.nutos.2020.12.006.
- [187] Y. C. Chung, Y. S. Kim, A. Shadchehr, A. Garrido, I. L. Macgregor, und M. H. Sleisenger, „Protein digestion and absorption in human small intestine“, *Gastroenterology*, Bd. 76, Nr. 6, S. 1415–1421, Juni 1979, doi: 10.1016/0016-5085(79)90410-4.
- [188] W. Miner-Williams, A. Deglaire, R. Benamouzig, M. F. Fuller, D. Tomé, und P. J. Moughan, „Endogenous proteins in terminal ileal digesta of adult subjects fed a casein-based diet“, *The American Journal of Clinical Nutrition*, Bd. 96, Nr. 3, S. 508–515, Sep. 2012, doi: 10.3945/ajcn.111.033472.
- [189] T. Wang, A. B. George, und S. Maslov, „Higher-Order Interactions in Auxotroph Communities Enhance Their Resilience to Resource Fluctuations“, 26. Mai 2024, *bioRxiv*. doi: 10.1101/2024.05.22.595348.
- [190] S. Starke u. a., „Amino acid auxotrophies in human gut bacteria are linked to higher microbiome diversity and long-term stability“, *ISME J*, Okt. 2023, doi: 10.1038/s41396-023-01537-3.
- [191] D. Machado u. a., „Polarization of microbial communities between competitive and cooperative metabolism“, *Nat Ecol Evol*, Bd. 5, Nr. 2, S. 195–203, Feb. 2021, doi: 10.1038/s41559-020-01353-4.
- [192] I. Granata u. a., „Duodenal Metatranscriptomics to Define Human and Microbial Functional Alterations Associated with Severe Obesity: A Pilot Study“, *Microorganisms*, Bd. 8, Nr. 11, S. 1811, Nov. 2020, doi: 10.3390/microorganisms8111811.
- [193] Y. Chen, F. Ji, J. Guo, D. Shi, D. Fang, und L. Li, „Dysbiosis of small intestinal microbiota in liver cirrhosis and its association with etiology“, *Sci Rep*, Bd. 6, Nr. 1, S. 34055, Sep. 2016, doi: 10.1038/srep34055.
- [194] R. Bodkhe u. a., „Comparison of Small Gut and Whole Gut Microbiota of First-Degree Relatives With Adult Celiac Disease Patients and Controls“, *Front. Microbiol.*, Bd. 10, S. 164, Feb. 2019, doi: 10.3389/fmicb.2019.00164.
- [195] J. Bures, „Small intestinal bacterial overgrowth syndrome“, *WJG*, Bd. 16, Nr. 24, S. 2978, 2010, doi: 10.3748/wjg.v16.i24.2978.
- [196] L. Miele u. a., „Increased intestinal permeability and tight junction alterations in nonalcoholic fatty liver disease“, *Hepatology*, Bd. 49, Nr. 6, S. 1877–1887, 2009, doi: 10.1002/hep.22848.
- [197] J.-M. Sabaté u. a., „High Prevalence of Small Intestinal Bacterial Overgrowth in Patients with Morbid Obesity: A Contributor to Severe Hepatic Steatosis“, *OBES SURG*, Bd. 18, Nr. 4, S. 371–377, Apr. 2008, doi: 10.1007/s11695-007-9398-2.
- [198] J. Klaus, U. Spaniol, G. Adler, R. A. Mason, M. Reinshagen, und C. von Tirpitz C, „Small intestinal bacterial overgrowth mimicking acute flare as a pitfall in patients with Crohn’s Disease“, *BMC Gastroenterol*, Bd. 9, Nr. 1, S. 61, Dez. 2009, doi: 10.1186/1471-230X-9-61.
- [199] G. Leite u. a., „The duodenal microbiome is altered in small intestinal bacterial overgrowth“, *PLoS ONE*, Bd. 15, Nr. 7, S. e0234906, Juli 2020, doi: 10.1371/journal.pone.0234906.

- [200] N. D. Rawlings, A. J. Barrett, P. D. Thomas, X. Huang, A. Bateman, und R. D. Finn, „The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database“, *Nucleic Acids Research*, Bd. 46, Nr. D1, S. D624–D632, Jan. 2018, doi: 10.1093/nar/gkx1134.
- [201] B. Borgstrom, A. Dahlqvist, G. Lundh, und J. Sjovall, „Studies of intestinal digestion and absorption in the human“, *J Clin Invest*, Bd. 36, Nr. 10, S. 1521–1536, Okt. 1957, doi: 10.1172/JCI103549.
- [202] S. A. Adibi, S. J. Gray, und E. Menden, „The Kinetics of Amino Acid Absorption and Alteration of Plasma Composition of Free Amino Acids After Intestinal Perfusion of Amino Acid Mixtures“, *The American Journal of Clinical Nutrition*, Bd. 20, Nr. 1, S. 24–33, Jan. 1967, doi: 10.1093/ajcn/20.1.24.
- [203] L. J. Dahm und D. P. Jones, „Secretion of cysteine and glutathione from mucosa to lumen in rat small intestine“, *American Journal of Physiology-Gastrointestinal and Liver Physiology*, Bd. 267, Nr. 2, S. G292–G300, Aug. 1994, doi: 10.1152/ajpgi.1994.267.2.G292.
- [204] F. C. Pereira und D. Berry, „Microbial nutrient niches in the gut: Microbial nutrient niches in the gut“, *Environ Microbiol*, Bd. 19, Nr. 4, S. 1366–1378, Apr. 2017, doi: 10.1111/1462-2920.13659.
- [205] B. Trastoy, A. Naegeli, I. Anso, J. Sjögren, und M. E. Guerin, „Structural basis of mammalian mucin processing by the human gut O-glycopeptidase OgpA from *Akkermansia muciniphila*“, *Nat Commun*, Bd. 11, Nr. 1, S. 4844, Sep. 2020, doi: 10.1038/s41467-020-18696-y.
- [206] Y. Li u. a., „Identification of trypsin-degrading commensals in the large intestine“, *Nature*, Bd. 609, Nr. 7927, S. 582–589, Sep. 2022, doi: 10.1038/s41586-022-05181-3.
- [207] A. Caminero u. a., „Diversity of the cultivable human gut microbiome involved in gluten metabolism: isolation of microorganisms with potential interest for coeliac disease“, *FEMS Microbiol Ecol*, Bd. 88, Nr. 2, S. 309–319, Mai 2014, doi: 10.1111/1574-6941.12295.
- [208] I. M. Carroll u. a., „Fecal Protease Activity Is Associated with Compositional Alterations in the Intestinal Microbiota“, *PLoS ONE*, Bd. 8, Nr. 10, S. e78017, Okt. 2013, doi: 10.1371/journal.pone.0078017.
- [209] S. S. C. Rao und J. Bhagatwala, „Small Intestinal Bacterial Overgrowth: Clinical Features and Therapeutic Management“, *Clin Transl Gastroenterol*, Bd. 10, Nr. 10, Art. Nr. 10, Okt. 2019, doi: 10.14309/ctg.0000000000000078.
- [210] Y. Bouhnik u. a., „Bacterial Populations Contaminating The Upper Gut in Patients With Small Intestinal Bacterial Overgrowth Syndrome“, *American Journal of Gastroenterology*, Bd. 94, Nr. 5, S. 1327–1331, Mai 1999, doi: 10.1111/j.1572-0241.1999.01016.x.
- [211] G. Vantrappen, J. Janssens, J. Hellemans, und Y. Ghooos, „The Interdigestive Motor Complex of Normal Subjects and Patients with Bacterial Overgrowth of the Small Intestine“, *J. Clin. Invest.*, Bd. 59, Nr. 6, S. 1158–1166, Juni 1977, doi: 10.1172/JCI108740.
- [212] V. B. Nieuwenhuijs u. a., „The Role of Interdigestive Small Bowel Motility in the Regulation of Gut Microflora, Bacterial Overgrowth, and Bacterial Translocation in Rats“, *Annals of Surgery*, Bd. 228, Nr. 2, S. 188–193, Aug. 1998, doi: 10.1097/00000658-199808000-00007.
- [213] J. Suri, R. Kataria, Z. Malik, H. P. Parkman, und R. Schey, „Elevated methane levels in small intestinal bacterial overgrowth suggests delayed small bowel and

- colonic transit“, *Medicine*, Bd. 97, Nr. 21, S. e10554, Mai 2018, doi: 10.1097/MD.00000000000010554.
- [214] B. C. Roland u. a., „Small Intestinal Transit Time Is Delayed in Small Intestinal Bacterial Overgrowth“, *Journal of Clinical Gastroenterology*, Bd. 49, Nr. 7, S. 571–576, Aug. 2015, doi: 10.1097/MCG.0000000000000257.
- [215] M. Wang, S. Ahrn, B. Jeppsson, und G. Molin, „Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes“, *FEMS Microbiology Ecology*, Bd. 54, Nr. 2, S. 219–231, Okt. 2005, doi: 10.1016/j.femsec.2005.03.012.
- [216] R. Leinonen, H. Sugawara, M. Shumway, und on behalf of the International Nucleotide Sequence Database Collaboration, „The Sequence Read Archive“, *Nucleic Acids Research*, Bd. 39, Nr. Database, S. D19–D21, Jan. 2011, doi: 10.1093/nar/gkq1019.
- [217] M. Martin, „Cutadapt removes adapter sequences from high-throughput sequencing reads“, *EMBnet j.*, Bd. 17, Nr. 1, S. 10, Mai 2011, doi: 10.14806/ej.17.1.200.
- [218] B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, und S. P. Holmes, „DADA2: High-resolution sample inference from Illumina amplicon data“, *Nat Methods*, Bd. 13, Nr. 7, S. 581–583, Juli 2016, doi: 10.1038/nmeth.3869.
- [219] D. Hyatt, G.-L. Chen, P. F. LoCascio, M. L. Land, F. W. Larimer, und L. J. Hauser, „Prodigal: prokaryotic gene recognition and translation initiation site identification“, *BMC Bioinformatics*, Bd. 11, Nr. 1, S. 119, Dez. 2010, doi: 10.1186/1471-2105-11-119.
- [220] N. D. Rawlings und F. R. Morton, „The MEROPS batch BLAST: A tool to detect peptidases and their non-peptidase homologues in a genome“, *Biochimie*, Bd. 90, Nr. 2, S. 243–259, Feb. 2008, doi: 10.1016/j.biochi.2007.09.014.
- [221] F. Teufel u. a., „SignalP 6.0 predicts all five types of signal peptides using protein language models“, *Nat Biotechnol*, Bd. 40, Nr. 7, S. 1023–1025, Juli 2022, doi: 10.1038/s41587-021-01156-3.
- [222] R. J. Wallace u. a., „The rumen microbial metagenome associated with high methane production in cattle“, *BMC Genomics*, Bd. 16, Nr. 1, S. 839, Dez. 2015, doi: 10.1186/s12864-015-2032-0.
- [223] M. Custodio u. a., „Microbial diversity in intensively farmed lake sediment contaminated by heavy metals and identification of microbial taxa bioindicators of environmental quality“, *Sci Rep*, Bd. 12, Nr. 1, S. 80, Jan. 2022, doi: 10.1038/s41598-021-03949-7.
- [224] Z. Zhang u. a., „Global biogeography of microbes driving ocean ecological status under climate change“, *Nat Commun*, Bd. 15, Nr. 1, S. 4657, Mai 2024, doi: 10.1038/s41467-024-49124-0.
- [225] M. Price, „Erroneous predictions of auxotrophies by CarveMe“, *Nat Ecol Evol*, Dez. 2022, doi: 10.1038/s41559-022-01936-3.
- [226] M. N. Price u. a., „Filling gaps in bacterial amino acid biosynthesis pathways with high-throughput genetics“, *PLoS Genet*, Bd. 14, Nr. 1, S. e1007147, Jan. 2018, doi: 10.1371/journal.pgen.1007147.
- [227] S. H. Payne und W. F. Loomis, „Retention and loss of amino acid biosynthetic pathways based on analysis of whole-genome sequences“, *Eukaryot Cell*, Bd. 5, Nr. 2, S. 272–276, Feb. 2006, doi: 10.1128/EC.5.2.272-276.2006.
- [228] Z. Dai, W. Zheng, und J. W. Locasale, „Amino acid variability, tradeoffs and optimality in human diet“, *Nat Commun*, Bd. 13, Nr. 1, S. 6683, Nov. 2022, doi: 10.1038/s41467-022-34486-0.

- [229] Joint WHO/FAO/UNU Expert Consultation, „Protein and amino acid requirements in human nutrition“, *World Health Organ Tech Rep Ser*, Nr. 935, S. 1–265, back cover, 2007.
- [230] G. Wu, „Dietary protein intake and human health“, *Food Funct.*, Bd. 7, Nr. 3, S. 1251–1265, 2016, doi: 10.1039/C5FO01530H.
- [231] S. Dietrich *u. a.*, „Amino acid intake and plasma concentrations and their interplay with gut microbiota in vegans and omnivores in Germany“, *Eur J Nutr*, Bd. 61, Nr. 4, S. 2103–2114, Juni 2022, doi: 10.1007/s00394-021-02790-y.
- [232] H. Akashi und T. Gojobori, „Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*“, *Proc. Natl. Acad. Sci. U.S.A.*, Bd. 99, Nr. 6, Art. Nr. 6, März 2002, doi: 10.1073/pnas.062526999.
- [233] IBDMDB Investigators *u. a.*, „Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases“, *Nature*, Bd. 569, Nr. 7758, S. 655–662, Mai 2019, doi: 10.1038/s41586-019-1237-9.
- [234] W. M. Johnson *u. a.*, „Auxotrophic interactions: a stabilizing attribute of aquatic microbial communities?“, *FEMS Microbiology Ecology*, Bd. 96, Nr. 11, S. fiaa115, Nov. 2020, doi: 10.1093/femsec/fiaa115.
- [235] M. Ziesack *u. a.*, „Engineered Interspecies Amino Acid Cross-Feeding Increases Population Evenness in a Synthetic Bacterial Consortium“, *mSystems*, Bd. 4, Nr. 4, S. e00352-19, Aug. 2019, doi: 10.1128/mSystems.00352-19.
- [236] P. J. Moughan, C. A. Butts, A. M. Rowan, und A. Deglaire, „Dietary peptides increase endogenous amino acid losses from the gut in adults“, *The American Journal of Clinical Nutrition*, Bd. 81, Nr. 6, S. 1359–1365, Juni 2005, doi: 10.1093/ajcn/81.6.1359.
- [237] C. Gaudichon *u. a.*, „Ileal losses of nitrogen and amino acids in humans and their importance to the assessment of amino acid requirements“, *Gastroenterology*, Bd. 123, Nr. 1, S. 50–59, Juli 2002, doi: 10.1053/gast.2002.34233.
- [238] C. S. Starck, R. R. Wolfe, und P. J. Moughan, „Endogenous Amino Acid Losses from the Gastrointestinal Tract of the Adult Human—A Quantitative Model“, *The Journal of Nutrition*, Bd. 148, Nr. 11, S. 1871–1881, Nov. 2018, doi: 10.1093/jn/nxy162.
- [239] D. V. Guzior und R. A. Quinn, „Review: microbial transformations of human bile acids“, *Microbiome*, Bd. 9, Nr. 1, S. 140, Dez. 2021, doi: 10.1186/s40168-021-01101-1.
- [240] D. Shalon *u. a.*, „Profiling the human intestinal environment under physiological conditions“, *Nature*, Bd. 617, Nr. 7961, S. 581–591, Mai 2023, doi: 10.1038/s41586-023-05989-7.
- [241] L. Henneke *u. a.*, „A dietary carbohydrate – gut *Parasutterella* – human fatty acid biosynthesis metabolic axis in obesity and type 2 diabetes“, *Gut Microbes*, Bd. 14, Nr. 1, Art. Nr. 1, Dez. 2022, doi: 10.1080/19490976.2022.2057778.
- [242] I. Marie, P. Ducrotte, P. Denis, J.-F. Menard, und H. Levesque, „Small intestinal bacterial overgrowth in systemic sclerosis“, *Rheumatology*, Bd. 48, Nr. 10, S. 1314–1319, Okt. 2009, doi: 10.1093/rheumatology/kep226.
- [243] E. M. M. Quigley, J. A. Murray, und M. Pimentel, „AGA Clinical Practice Update on Small Intestinal Bacterial Overgrowth: Expert Review“, *Gastroenterology*, Bd. 159, Nr. 4, S. 1526–1532, Okt. 2020, doi: 10.1053/j.gastro.2020.06.090.
- [244] M. J. Blaser, „Fecal Microbiota Transplantation for Dysbiosis — Predictable Risks“, *N Engl J Med*, Bd. 381, Nr. 21, S. 2064–2066, Nov. 2019, doi: 10.1056/NEJMe1913807.
- [245] J. Chen, S. L. Siliceo, Y. Ni, H. B. Nielsen, A. Xu, und G. Panagiotou, „Identification of robust and generalizable biomarkers for microbiome-based

- stratification in lifestyle interventions“, *Microbiome*, Bd. 11, Nr. 1, S. 178, Aug. 2023, doi: 10.1186/s40168-023-01604-z.
- [246] J. Rodriguez u. a., „State of the art and the future of microbiome-based biomarkers: a multidisciplinary Delphi consensus“, *The Lancet Microbe*, S. 100948, Sep. 2024, doi: 10.1016/j.lanmic.2024.07.011.
- [247] T. Hessler u. a., „Vitamin interdependencies predicted by metagenomics-informed network analyses and validated in microbial community microcosms“, *Nat Commun*, Bd. 14, Nr. 1, S. 4768, Aug. 2023, doi: 10.1038/s41467-023-40360-4.
- [248] L. Redondo-Cuevas u. a., „Do Herbal Supplements and Probiotics Complement Antibiotics and Diet in the Management of SIBO? A Randomized Clinical Trial“, *Nutrients*, Bd. 16, Nr. 7, S. 1083, Apr. 2024, doi: 10.3390/nu16071083.
- [249] E. Zaccaria u. a., „Endogenous small intestinal microbiome determinants of transient colonisation efficiency by bacteria from fermented dairy products: a randomised controlled trial“, *Microbiome*, Bd. 11, Nr. 1, S. 43, März 2023, doi: 10.1186/s40168-023-01491-4.
- [250] J. Jansma u. a., „Dynamic effects of probiotic formula ecologic®825 on human small intestinal ileostoma microbiota: a network theory approach“, *Gut Microbes*, Bd. 15, Nr. 1, S. 2232506, Dez. 2023, doi: 10.1080/19490976.2023.2232506.
- [251] L. Cavaletti u. a., „E40, a novel microbial protease efficiently detoxifying gluten proteins, for the dietary management of gluten intolerance“, *Sci Rep*, Bd. 9, Nr. 1, Art. Nr. 1, Sep. 2019, doi: 10.1038/s41598-019-48299-7.
- [252] A. Santiago u. a., „Crohn’s disease proteolytic microbiota enhances inflammation through PAR2 pathway in gnotobiotic mice“, *Gut Microbes*, Bd. 15, Nr. 1, Art. Nr. 1, Dez. 2023, doi: 10.1080/19490976.2023.2205425.
- [253] P. Escudeiro, J. Pothier, F. Dionisio, und T. Nogueira, „Antibiotic Resistance Gene Diversity and Virulence Gene Diversity Are Correlated in Human Gut and Environmental Microbiomes“, *mSphere*, Bd. 4, Nr. 3, S. e00135-19, Juni 2019, doi: 10.1128/mSphere.00135-19.
- [254] Y. Hu u. a., „Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota“, *Nat Commun*, Bd. 4, Nr. 1, S. 2151, Juli 2013, doi: 10.1038/ncomms3151.
- [255] S. Fredriksen, S. De Warle, P. Van Baarlen, J. Boekhorst, und J. M. Wells, „Resistome expansion in disease-associated human gut microbiomes“, *Microbiome*, Bd. 11, Nr. 1, S. 166, Juli 2023, doi: 10.1186/s40168-023-01610-1.
- [256] M. Shuai u. a., „Human Gut Antibiotic Resistome and Progression of Diabetes“, *Advanced Science*, Bd. 9, Nr. 11, S. 2104965, Apr. 2022, doi: 10.1002/advs.202104965.
- [257] Y. Momose, K. Hirayama, und K. Itoh, „Competition for proline between indigenous *Escherichia coli* and *E. coli* O157:H7 in gnotobiotic mice associated with infant intestinal microbiota and its contribution to the colonization resistance against *E. coli* O157:H7“, *Antonie van Leeuwenhoek*, Bd. 94, Nr. 2, S. 165–171, Aug. 2008, doi: 10.1007/s10482-008-9222-6.
- [258] U. Ghoshal, U. C. Ghoshal, P. Ranjan, S. R. Naik, und A. Ayyagari, „Spectrum and antibiotic sensitivity of bacteria contaminating the upper gut in patients with malabsorption syndrome from the tropics“, *BMC Gastroenterol*, Bd. 3, Nr. 1, S. 9, Mai 2003, doi: 10.1186/1471-230X-3-9.
- [259] P. Roszkowska u. a., „Small Intestinal Bacterial Overgrowth (SIBO) and Twelve Groups of Related Diseases—Current State of Knowledge“, *Biomedicines*, Bd. 12, Nr. 5, S. 1030, Mai 2024, doi: 10.3390/biomedicines12051030.

- [260] M. Pimentel, „Review of rifaximin as treatment for SIBO and IBS“, *Expert Opinion on Investigational Drugs*, Bd. 18, Nr. 3, S. 349–358, März 2009, doi: 10.1517/13543780902780175.
- [261] U. Klümper u. a., „Environmental microbiome diversity and stability is a barrier to antimicrobial resistance gene accumulation“, *Commun Biol*, Bd. 7, Nr. 1, S. 706, Juni 2024, doi: 10.1038/s42003-024-06338-8.
- [262] H. Suzuki, W. Hashimoto, und H. Kumagai, „Escherichia coli K-12 can utilize an exogenous gamma-glutamyl peptide as an amino acid source, for which gamma-glutamyltranspeptidase is essential“, *J Bacteriol*, Bd. 175, Nr. 18, Art. Nr. 18, Sep. 1993, doi: 10.1128/jb.175.18.6038-6040.1993.
- [263] S. V. Rana u. a., „Relationship of cytokines, oxidative stress and GI motility with bacterial overgrowth in ulcerative colitis patients“, *Journal of Crohn's and Colitis*, Bd. 8, Nr. 8, S. 859–865, Aug. 2014, doi: 10.1016/j.crohns.2014.01.007.
- [264] T. V. Hung und T. Suzuki, „Short-Chain Fatty Acids Suppress Inflammatory Reactions in Caco-2 Cells and Mouse Colons“, *J. Agric. Food Chem.*, Bd. 66, Nr. 1, S. 108–117, Jan. 2018, doi: 10.1021/acs.jafc.7b04233.
- [265] G. Chen u. a., „Sodium Butyrate Inhibits Inflammation and Maintains Epithelium Barrier Integrity in a TNBS-induced Inflammatory Bowel Disease Mice Model“, *EBioMedicine*, Bd. 30, S. 317–325, Apr. 2018, doi: 10.1016/j.ebiom.2018.03.030.
- [266] T. Hosverstad, A. Bjørnnekleit, O. Fausa, und T. Midtvedt, „Short-Chain Fatty Acids in the Small-Bowel Bacterial Overgrowth Syndrome“, *Scandinavian Journal of Gastroenterology*, Bd. 20, Nr. 4, S. 492–499, Jan. 1985, doi: 10.3109/00365528509089686.
- [267] M. T. Wortel, E. Noor, M. Ferris, F. J. Bruggeman, und W. Liebermeister, „Metabolic enzyme cost explains variable trade-offs between microbial growth rate and yield“, *PLoS Comput Biol*, Bd. 14, Nr. 2, S. e1006010, Feb. 2018, doi: 10.1371/journal.pcbi.1006010.
- [268] G. Zampieri, S. Campanaro, C. Angione, und L. Treu, „Metatranscriptomics-guided genome-scale metabolic modeling of microbial communities“, *Cell Reports Methods*, Bd. 3, Nr. 1, S. 100383, Jan. 2023, doi: 10.1016/j.crmeth.2022.100383.
- [269] C. Zuñiga u. a., „Environmental stimuli drive a transition from cooperation to competition in synthetic phototrophic communities“, *Nat Microbiol*, Bd. 4, Nr. 12, S. 2184–2191, Okt. 2019, doi: 10.1038/s41564-019-0567-6.
- [270] C. Ferrario u. a., „Exploring Amino Acid Auxotrophy in *Bifidobacterium bifidum* PRL2010“, *Frontiers in Microbiology*, Bd. 6, 2015, Zugegriffen: 23. August 2023. [Online]. Verfügbar unter: <https://www.frontiersin.org/articles/10.3389/fmicb.2015.01331>
- [271] M. Veda, S. Nakamoto, R. Nakai, und A. Takagi, „Establishment of a defined minimal medium and isolation of auxotrophic mutants for *Bifidobacterium bifidum* Es 5“, *The Journal of General and Applied Microbiology*, Bd. 29, Nr. 2, S. 103–114, 1983, doi: 10.2323/jgam.29.103.
- [272] M. Schöpping, P. Gaspar, A. R. Neves, C. J. Franzén, und A. A. Zeidan, „Identifying the essential nutritional requirements of the probiotic bacteria *Bifidobacterium animalis* and *Bifidobacterium longum* through genome-scale modeling“, *npj Syst Biol Appl*, Bd. 7, Nr. 1, Art. Nr. 1, Dez. 2021, doi: 10.1038/s41540-021-00207-4.
- [273] J.-H. Lee und D. J. O'Sullivan, „Genomic Insights into *Bifidobacteria*“, *Microbiology and Molecular Biology Reviews*, Bd. 74, Nr. 3, S. 378–416, Sep. 2010, doi: 10.1128/mmbr.00004-10.

- [274] C. Quast *u. a.*, „The SILVA ribosomal RNA gene database project: improved data processing and web-based tools“, *Nucleic Acids Research*, Bd. 41, Nr. D1, S. D590–D596, Nov. 2012, doi: 10.1093/nar/gks1219.
- [275] Daniel J. Stekhoven, „missForest: Nonparametric Missing Value Imputation using Random Forest“, *Bioinformatics*, Bd. 28, Nr. 1, S. 112–118, 2012.
- [276] Q. Liu, C. Li, V. Wanga, und B. E. Shepherd, „Covariate-adjusted Spearman’s rank correlation with probability-scale residuals: Covariate-Adjusted Spearman’s Rank Correlation“, *Biom*, Bd. 74, Nr. 2, S. 595–605, Juni 2018, doi: 10.1111/biom.12812.