

AI-driven discovery of hypoglycemic quorum sensing peptides from the human gut microbiome

Karsten Kristiansen

kk@bio.ku.dk

University of Copenhagen <https://orcid.org/0000-0002-6024-0917>

chuang Xue

Jianwei Chen

BGI Research <https://orcid.org/0000-0003-2332-4609>

Wei Kang

Dalian University of Technology <https://orcid.org/0000-0002-2288-3540>

Ling Peng

The Hong Kong Polytechnic University

Junping Liu

Dalian University of Technology

Xiao Zhou

Dalian Institute of Chemical Physics

Yujing Liu

BGI Research

Dantong Wang

BGI Research

Xi Chen

Princeton University

Yong Yang

Dalian University of Technology

Yanbin Feng

Dalian University of Technology

Mingdong Wang

Dalian University of Technology

Zhongyi Yao

Dalian University of Technology

Hui Wan

Dalian University of Technology

Xiao Ma

Dalian University of Technology

Chao Zhu

Dalian University of Technology

Guanchen Liu

Dalian University of Technology <https://orcid.org/0009-0000-3834-6208>

Ren Wu

Dalian Institute of Chemical Physics

Hao Li

Dalian University of Technology

Sijin Wu

Xi'an Jiaotong-Liverpool University <https://orcid.org/0000-0003-4554-4610>

Ruiqi Liu

Dalian University of Technology

Simon Ming-yuen Lee

The Hong Kong Polytechnic University

Huanming Yang

BGI-Shenzhen <https://orcid.org/0000-0002-0858-3410>

Guangyi Fan

BGI Research

Ye Yin

BGI Research

Chuang Xue

Dalian University of Technology <https://orcid.org/0000-0002-3856-8457>

Article

Keywords:

Posted Date: September 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9855030/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

AI-driven discovery of hypoglycemic quorum sensing peptides from the human gut microbiome

Shijie Ren^{1,9}, Jianwei Chen^{2,3,4,9}, Wei Kang^{1,5,9}, Ling Peng⁴, Junping Liu¹, Xiaoyu Zhou^{6,7}, Yujing Liu², Dantong Wang², Xi Chen⁸, Yongliang Yang¹, Yanbin Feng¹, Mingdong Wang¹, Zhongyi Yao¹, Huihui Wan¹, Xiao Ma¹, Chao Zhu¹, Guanchen Liu¹, Ren'an Wu⁶, Haonan Li¹, Sijin Wu⁶, Ruiqi Liu¹, Simon Ming-yuen Lee⁴, Huanming Yang², Guangyi Fan^{2,4}, Ye Yin², Karsten Kristiansen^{2,3,*}, Chuang Xue^{1,5,10,*}

¹ MOE Key Laboratory of Bio-Intelligent Manufacturing, State Key Laboratory of Fine Chemicals, Frontiers Science Centre for Smart Materials Oriented Chemical Engineering, School of Bioengineering, Dalian University of Technology, Dalian 116024, China

² State Key Laboratory of Genome and Multi-omics Technologies, BGI Research, Shenzhen 518083, China

³ Laboratory of Integrative Biomedicine, Department of Biology, University of Copenhagen, Copenhagen 2100, Denmark

⁴ Department of Food Science and Nutrition, and PolyU-BGI Joint Research Centre for Genomics and Synthetic Biology in Global Ocean Resources, The Hong Kong Polytechnic University, Hong Kong, China

⁵ Ningbo Institute of Dalian University of Technology, Ningbo 315016, China

⁶ Laboratory of High-Resolution Mass Spectrometry Technologies, State Key Laboratory of Molecular Reaction Dynamics, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, 116023, China

⁷ State Key Laboratory of Molecular Reaction Dynamics, University of Chinese Academy of Sciences, Beijing 100049, China.

⁸ Princeton University, Princeton, NJ 08544, USA.

⁹ These authors contributed equally to this work.

¹⁰ Lead contact

* Correspondence: kk@bio.ku.dk, xue.1@dlut.edu.cn

Abstract

The human gut microbiome encodes a large spectrum of quorum sensing peptides (QSPs) playing a vital role in microbial communication and host-microbial interactions, yet discovery and functional characterization of novel QSPs are still limited. Here, we developed an integrated artificial intelligence (AI)-driven pipeline combining large-scale genome mining, machine learning-based screening, AlphaFold3 structural modelling, and molecular dynamics simulations to systematically identify 1,670 novel candidate QSPs (cQSPs) from 12,488 human gut microbial genomes. Subsequent experimental screening identified 29 cQSPs with α -glucosidase inhibitory activities. Notably, GIVLVPTSLVSG (GIVL), an Agr-derived QSP from *Clostridium tyrobutyricum*, not only exhibited canonical quorum-sensing functions, but also demonstrated a stable interaction with and strong inhibitory activity against α -glucosidase. In diabetic mice, administration of GIVL significantly improved glucose homeostasis and insulin sensitivity accompanied by increased production of short-chain fatty acids (SCFAs) in the gut microbiome. Our findings expand the QSP database and establish a generalizable, AI-enabled framework for the systematic discovery of microbiome-derived peptides with therapeutic potential in metabolic disease, positioning QSPs as promising sources of therapeutics for diabetes and other chronic metabolic disorders.

Main

Human microbiomes consist of trillions of microorganisms residing on and in the body, forming complex communities that communicate through chemical signaling systems to coordinate bacterial collective behaviors and adapting to dynamic environments¹⁻⁸. In Gram-positive bacteria, the communication is predominantly mediated by quorum sensing peptides (QSPs), a diverse class of secreted oligopeptides that function as key regulators of population density-dependent processes^{9,10}. Beyond their canonical roles in controlling biofilm formation, virulence, drug resistance, and interbacterial competition^{7,11}, recent studies have further revealed that QSPs also act as mediators of host-microbiome crosstalk, modulating host immune interactions, epithelial barrier function, and disease progression¹²⁻¹⁷. These findings suggest that QSPs not only act as microbial communication signals but also as bioactive molecules with the potential to directly influence host physiology.

Among the diverse families of QSPs, the autoinducing peptides (AIPs) derived from the accessory gene regulator (Agr) system represent the classic type and most extensively studied classes of QSPs¹⁸⁻²⁴ (**Extended Data Fig. 1**). In classical models such as *Staphylococcus aureus*, AIPs are generated through proteolytic processing of the AgrD precursor by the membrane-bound protease AgrB, followed by secretion and activation of the AgrC/AgrA two-component signal transduction system²⁵⁻²⁷. AIPs, also known as Agr-derived QSPs, were originally documented to regulate virulence and differentiation²⁸⁻³⁰, but have now been shown to exhibit broader biological activities, including cross-species quorum sensing

interference, antimicrobial effects, and modulation of host inflammatory responses^{13,25,30-32}. For example, AIPs secreted from the skin commensal coagulase-negative staphylococci, such as *S. simulans* and *S. saprophyticus*, can attenuate the pathogen virulence of *S. aureus* through quorum quenching and thereby reducing skin damage in a murine infection model^{32,33}. AIPs secreted from *S. aureus* have also been used to prevent abscess formation in a murine model³⁴. These findings highlight the therapeutic potential of Agr-derived QSPs in the control of infection. However, the potential roles of Agr-derived QSPs in metabolic regulation and systemic host physiology remain largely unexplored.

Despite growing recognition of their importance, together with the growing development of QSP prediction tools²⁰, large-scale discovery and functional annotation of QSPs remain limited. For example, the established QSPs database Quorumpeps (Wynendaele *et al.*, 2012), one of the few publicly available databases dedicated to experimentally validated QSPs, currently contains 359 peptide sequences, of which only 64 are derived from the Agr system³⁵. In addition, the Quorum Sensing of Human Gut Microbes (QSHGM) database provides a specialized resource for quorum sensing relevant proteins, which contains 28,567 non-redundant entries. However, it primarily focuses on QS synthases and receptors rather than signal peptides, thereby underscoring the continued scarcity of dedicated QSP resources³⁶. Furthermore, the bioactivities of these QSPs are mainly limited to the basic functions of virulence and biofilm formation in microbes that secrete these peptides, with almost no data available on cross-species or host-targeted activities. Beyond databases, computational prediction tools, such as QSPpred, iQSP, EnsembleQS, and PEPred-Suite, have been developed to facilitate QSP prediction²⁰. However, these algorithms are constrained by training on small-scale datasets with QSPs only including conventional quorum sensing effects, thus offering limited insight into the broader biological functional diversity of QSPs. Addressing these challenges requires integrative strategies that combine large-scale peptide mining with functional prediction and experimental validation.

The human gut microbiome represents a largely untapped reservoir of bioactive peptides with diverse biological functions³⁷. Large-scale computational screening has predicted hundreds of antimicrobial peptides from human gut microbiome, followed by experimental validation demonstrating potent efficacy against antibiotic-resistant pathogens^{38,39}. Moreover, recent studies have revealed that gut microbial peptides may constitute attractive potential candidates for the treatment of a variety of diseases, including cancer and metabolic diseases^{40,41}. However, even though the diversity of characterized human gut microbial peptides has increased vastly, the specific subset of QSPs remains poorly characterized. In this study, we established an integrated computational-experimental framework for the systematic discovery and functional characterization of QSPs from the human gut microbiome. By combining large-scale genome mining, machine learning-based screening, AlphaFold3 structural modelling, and molecular dynamics (MD) simulations, we identified 1,670 Agr-derived candidate QSPs (cQSPs) from the human gut microbiome and revealed their potential interactions with sucrase-isomaltase (SI), the primary α -

glucosidase localized in the small intestine. *In vitro* experiments further validated that 29 of 46 chemically synthesized cQSPs exhibited α -glucosidase inhibitory bioactivity. Especially GIVL (cQSP_1089, GIVLVPTSLVSG) derived from *Clostridium tyrobutyricum*, exhibited interesting α -glucosidase inhibitory properties. Administration of GIVL significantly improved glucose tolerance and insulin sensitivity in a diabetic mouse model and increased SCFAs production in the gut microbiome. Our study not only expands the functional landscape of QSPs and uncovers a previously unrecognized role of QSPs in host metabolic regulation but also provides a new perspective for a deeper understanding of the interaction between intestinal bacterial communication systems and host metabolism of importance for the development of peptide-based therapeutics targeting metabolic diseases.

Results

Identification and bioactivity prediction of candidate QSPs from the human gut microbiome

We first gathered a total of 12,488 human gut bacterial genomes from the Fecal Microbiota Transplantation (FMT) database (956 genomes)⁴², the Culturable Genome Reference (CGR2) database (3,324 genomes)⁴³, the Unified Human Gastrointestinal Genome (UHGG2) database⁴⁴ (4,744 genomes) and the Broad Institute-OpenBiome Microbiome Library (BIO-ML) repositories⁴⁵ (3,464 genomes) (**Fig. 1a**). Previous studies have demonstrated that the *agrBDCA* cluster is responsible for the production of the quorum sensing signal molecules in Gram-positive bacteria^{23,46}. While *AgrD* encodes the pro-peptide sequences of QSPs, the other genes in the cluster are responsible for processing, secretion, and regulation of the QSPs^{29,30,47-51}. Based on gene annotation, we detected 2,776 *agrD* gene sequences from 1,655 bacterial genomes, mainly distributed in the genera *Enterococcus_B* (179 genomes), *Blautia_A* (126 genomes), *Agathobacter* (73 genomes), *Fecalibacillus* (71 genomes), and *Clostridium* (56 genomes) belonging to the phylum *Bacillota* (**Fig. 1a** and **Supplementary Table 1**).

Since the *AgrD*-derived QSPs maintain strict compositional rules as proteolytic products of the pro-peptide sequences containing 5 to 17 amino acids (aa) featuring a highly conserved proline residue located at residue 6 from the C-terminus^{46,52}, we first anchored the C-terminal processing region of the *AgrD* pro-peptides at the conserved proline residues using multiple sequence alignment. To capture all potential bioactive forms, we applied a sliding window strategy to systematically extract all continuous fragments ranging from 5 to 17 amino acids within this region, generating an initial candidate quorum sensing peptides (cQSPs) library comprising 11,767 fragment peptides (**Fig. 1a**). To further identify high-confidence cQSPs, we used machine-learning-based tools PSRQSP⁵³ and the deep-learning-based PreTP-2L⁵⁴ for QSP sequence prediction and then used ToxinPred3⁵⁵ to predict the cytotoxicity of the predicted cQSP sequences. Finally, a total of 1,670 peptides were predicted to be non-toxic and high-confidence

cQSPs (**Fig. 1a** and **Supplementary Table 2**). Compared to the Quorumpeps database³⁵, all the 1,670 cQSPs represented novel sequences. Subsequently, the 11,767 fragment peptide sequences were clustered with the 1,670 cQSPs based on 75% similarity using CD-Hit⁵⁶ to identify the prevalence frequency of the cQSPs, revealing that 317 cQSP families were prevalent in the human gut microbiome, representing at least three fragment peptides. Finally, we selected 82 cQSPs for subsequent screening for biological activity from the 317 cQSP families with a lower predicted cytotoxicity score (< 0.3), and folding ability predicted by AlphaFold2 (**Fig. 1a**).

Based on the physiological hypothesis that secreted bioactive peptides derived from the gut microbiome may influence nutrient absorption by regulating digestive enzyme activities, we initially focused on two key extracellular enzyme families, oligosaccharidases and lipases, both closely linked to nutrient digestion in the human intestine. We used AlphaFold3 to systematically predict the binding interactions between all 82 cQSPs and 12 representative proteins from these enzyme families. This analysis revealed that 58 of the 82 cQSPs exhibited potential binding affinity specifically for sucrase-isomaltase (SI), while interactions with other digestive enzymes were unstable or unreliable (**Supplementary Table 3**). SI is the primary α -glucosidase localized to the small intestinal brush border and is responsible for hydrolyzing dietary disaccharides such as sucrose and maltose, making it a key pharmacological target for postprandial glycemic control⁵⁷. Therefore, the above binding predictions suggest that the peptides identified in this study may exert their function by specifically inhibiting α -glucosidase activity.

Computational and experimental validations of the α -glucosidase inhibitory activity of cQSPs

To validate this predicted function, we successfully synthesized 46 of the 82 cQSPs using solid-phase peptide synthesis and assessed their α -glucosidase inhibitory ability *in vitro*. Consistent with AlphaFold3 binding results, we found that 29 of 46 synthesized cQSPs showed α -glucosidase inhibitory activity, and of these, 23 cQSPs inhibited α -glucosidase with IC₅₀ values that were less than half of that of acarbose, an α -glucosidase inhibitor used for the treatment of type 2 diabetes for more than 40 years⁵⁸⁻⁶² (**Supplementary Table 4**). To assess the potential toxicity of cQSPs to mammalian cells, we performed hemolysis assays for all cQSPs and determined the 50% hemolytic concentration (HC₅₀, 50% hemolysis) toxicity values (**Fig. 1b** and **Supplementary Table 4**). These results showed that 30.43% (14/46) peptides exhibited negligible hemolytic activity against mouse erythrocytes at a concentration of 20 mg mL⁻¹. Notably, the large HC₅₀ to IC₅₀ ratios (>500) indicated that 15 tested peptides exhibited broad potential therapeutic windows, i.e., they did not show significant hemolytic activity within the effective dose range.

Among these 15 cQSPs, we found that three (cQSP_1089, cQSP_1105, and cQSP_1119) not only possessed non-cytotoxicity but also exhibited potent α -glucosidase inhibitory activity (**Fig. 1c** and **Supplementary Table 4**). The AlphaFold3 results showed that all three peptides bound within the same

active pocket of SI, whereas acarbose occupied a distinct site, suggesting a different mode of α -glucosidase binding inhibition (**Fig. 1d**). Compared with the acarbose-SI complex, the three peptide-SI complexes showed no changes in α -helix content, whereas a slight reduction in β -sheet was observed (**Supplementary Table 5**), indicating that binding interaction is primarily mediated through occupation of the catalytic pocket and local structural perturbations. We further performed molecular dynamics (MD) simulations analysis to determine the binding stability of these three peptides to SI. The MD simulations showed that all three peptides formed stable interactions with SI compared to acarbose, suggesting their potential as pharmaceutical α -glucosidase inhibitors (**Extended Data Fig. 2a, b** and **Supplementary Table 6**). Consistent with these computational predictions, the three peptides showed a typical dose-dependence on α -glucosidase, with comparable IC_{50} values of 3.27, 1.93, and 2.88 μ M, respectively, which were 1-2 orders of magnitude lower than that of the clinical drug acarbose ($IC_{50} > 100 \mu$ M). Kinetic analysis further revealed that the three peptides cQSP_1089 (15.3 μ M), cQSP_1119 (15.3 μ M), and cQSP_1105 (9.0 μ M) achieved maximum inhibition rates of 98.26%, 99.35%, and 87.41%, respectively, which was significantly better than that of the acarbose control group at the same mass concentration (inhibition rate of 22.22% at 27.1 μ M) (**Fig. 1c**). These findings highlight the strong inhibitory activity of these peptides and their potential therapeutic value. Considering the suitability for further testing using the bacterial species giving rise to these three QSPs (see **Methods**), cQSP_1089 (hereby referred to as GIVL) derived from *Clostridium tyrobutyricum* was ultimately chosen as the model molecule for mechanistic studies. Besides, to further validate the sequence specificity of GIVL, we synthesized two scrambled peptides with identical length and amino acid composition to GIVL. *In vitro* enzymatic assays confirmed that the scrambled peptides did not inhibit α -glucosidase activity (**Extended Data Fig. 3** and **Supplementary Table 7**).

To identify the residues contributing to the stability and binding affinity of the SI-GIVL complex, we performed *in silico* alanine-scanning mutagenesis on peptide GIVL against SI using AlphaFold3. The structural modelling analysis showed that Gly1, Val5, and Pro6 of GIVL contribute significantly to binding to SI and may play important roles in α -glucosidase inhibition (**Extended Data Fig. 4a**). AlphaFold3 modelling of version of GIVL with alanine substitution of Gly1, Val5, and Pro6 in GIVL revealed fewer ideal hydrogen bonds (yellow) and more electrostatic clashes (red) (**Extended Data Fig. 4b-e** and **Supplementary Table 8**), indicating that Gly1, Val5, and Pro6 play essential roles in mediating the interaction with SI. We next investigated the interaction between GIVL and α -glucosidase using fluorescence spectroscopy, circular dichroism (CD), and Fourier-transform infrared spectroscopy (FTIR). Fluorescence spectroscopy demonstrated that the intrinsic fluorescence emission peak of α -glucosidase (~334 nm) decreased gradually following the addition of GIVL, indicating their binding interaction (**Extended Data Fig. 5a**). CD spectroscopy revealed dose-dependent conformational changes, with increased α -helix content (3.7% to 8.5%) and decreased β -sheet content (32.7% to 30.4%) upon binding

of GIVL (**Extended Data Fig. 5b**). In parallel, FTIR showed that the characteristic peak (1666 cm^{-1}) corresponding to the vibration of C=O of free α -glucosidase shifted slightly to 1649 cm^{-1} with reduced intensity, indicating rearrangement of the carbonyl hydrogen bonds (**Extended Data Fig. 5c**). Surface plasmon resonance (SPR) further confirmed the binding of GIVL with an estimated dissociation constant (KD) of $7.6\text{ }\mu\text{M}$ (**Extended Data Fig. 5d**). Finally, Lineweaver-Burk analysis revealed that GIVL exhibited a mixed inhibition pattern against α -glucosidase (**Extended Data Fig. 5e**), thereby reducing the activity of the enzyme. Collectively, these results demonstrate that GIVL binds to and inhibits α -glucosidase *in vitro*.

GIVL ameliorates glucose intolerance and insulin resistance

Having established that GIVL exhibited both binding affinity and inhibitory activity against α -glucosidase, we next gavaged mice with PBS, fluorescein isothiocyanate (FITC), and GIVL conjugated with FITC (FITC-GIVL) to determine if GIVL was able to locate with SI *in vivo*. Previous studies have shown that SI is located in the brush border of the small intestinal epithelium, serving as a major target for glycemic control⁵⁷. An intense fluorescence signal was observed in the intestinal tract of mice gavaged with FITC-GIVL (**Fig. 2a**), suggesting the binding of GIVL to SI. In contrast, the scrambled peptides did not exhibit the intestinal brush-border enrichment observed with GIVL, indicating that GIVL specifically localized to SI *in vivo* in a sequence-dependent manner (**Extended Data Fig. 6**). In addition, no increase in fluorescence signals was observed in heart, liver, spleen, lung, and kidneys (**Extended Data Fig. 7**), indicating that GIVL is primarily localized to the intestinal tract. Further time-dependent distribution studies showed that the fluorescence signal of FITC-GIVL was clearly visible in the small intestine tissue during the early post-administration period (approximately 2-8 h), then gradually diminished over time, and showed no significant residual signal after approximately 10 h (**Extended Data Fig. 8**). These findings indicate that GIVL did not persist in the intestine and is largely cleared within 24 hours. Additionally, we investigated the *in vivo* location of GIVL using fluorescently labelled anti-glucosidase (SI and maltase-glucoamylase, MGAM, another α -glucosidase localized in small intestine) antibodies. Immunofluorescence analysis further revealed colocalization of GIVL with SI in the small intestine, but not with MGAM (**Fig. 2b, c**). These results suggest that GIVL administration directly inhibits SI in the small intestinal mucosa, which may delay or inhibit the enzymatic hydrolysis of sucrose and hence absorption of glucose, leading to a lower blood glucose level and passage of sucrose to the colon.

To assess whether GIVL affected glucose tolerance and insulin sensitivity, normal male mice and streptozotocin-induced T2D mice were divided into five groups: NCD: normal male mice fed a chow diet; HFD: T2D mice fed high-fat diet; HFD+LG: T2D mice fed a high-fat diet gavaged with GIVL at $50\text{ mg kg}^{-1}\text{ BW/day}$; HFD+HG: T2D mice fed a high-fat diet and gavaged with GIVL at $200\text{ mg kg}^{-1}\text{ BW/day}$; HFD+Acar: T2D mice fed a high-fat diet and gavaged with acarbose at $20\text{ mg kg}^{-1}\text{ BW/day}$

(**Fig. 3a** and **Supplementary Table 9**). After 10 weeks of GIVL intervention, no toxicity-related abnormalities were observed in body weight (**Extended Data Fig. 9a**), food intake (**Extended Data Fig. 9b**), general behavior, or survival at any tested dose. Notably, histological analysis showed that compared with the HFD group, no noticeable abnormality appeared in epididymal adipose, liver, and kidneys of mice in the HFD+LG, HFD+HG and HFD+Acar groups. Still, clear initial therapeutic effects were observed in mice receiving GIVL, indicating that both low- and high-dose GIVL could exert protective effects against multi-organ pathological changes, including reducing adipocyte size and fat accumulation in liver and kidney (**Extended Data Fig. 10**). Further, administration of a low or high dose of GIVL significantly improved serum liver enzyme profiles and lipid parameters. Specifically, low-dose GIVL markedly reduced plasma alanine aminotransferase (ALT) and aspartate aminotransferase (AST) levels, while increasing high-density lipoprotein cholesterol (HDL) and decreasing triglyceride (TG) levels compared to the HFD group (**Extended Data Fig. 11a-d**). In addition, intestinal barrier integrity was significantly enhanced in the HFD+LG group, as evidenced by increased intestinal wall thickness and a higher number of goblet cells (**Extended Data Fig. 11e, f**).

To further assess whether GIVL treatment affected glucose tolerance and insulin sensitivity, we measured the levels of fasting blood glucose (FBG) and fasting insulin (FINS) and performed an oral glucose tolerance test (OGTT) and i.p. insulin tolerance test (ITT). Compared to the HFD group, mice in the HFD+LG group showed significantly lower FBG levels (**Fig. 3b**), improved homeostasis assessment of insulin resistance (HOMA-IR) values (**Extended Data Fig. 11g**), showed a higher insulin sensitivity in the ITT (**Fig. 3c**), and an increased ability to dispose of the oral glucose load from circulation during an OGTT (**Fig. 3d**). When the GIVL dose was increased to 200 mg/kg (HFD+HG), its metabolic-improving effects were further enhanced and extended to a broader range of indicators (**Fig. 3a-e**). High-dose GIVL not only produced more pronounced effects across all the aforementioned indices, but also led to significant improvements in several indicators that did not reach statistical significance at the lower dose, including fasting insulin (FINS) levels (**Fig. 3e**), low-density lipoprotein cholesterol (LDL), total cholesterol (TC), intestinal villus height, and plasma LPS and IL-6 levels (**Extended Data Fig. 11h-l**).

Impacts of oral gavage with GIVL on gut microbiome

Previous studies have shown that acarbose lowers blood glucose not only by targeting α -glucosidase but also by impeding the hydrolysis and absorption of carbohydrates in the small intestine, which in turn alters the availability of fermentation substrates for the gut microbiome, thereby promoting the growth of beneficial bacteria that contribute to the production of short-chain fatty acids (SCFAs)⁶³⁻⁶⁶. Therefore, after confirming that GIVL targeted SI in the small intestine and improved glucose clearance and insulin sensitivity in mice, we further investigated whether GIVL administration would similarly enhance gut health in diabetic mice, analogous to the effects of acarbose. We used shotgun metagenomics sequencing

to investigate the gut microbiome composition in the different treatment groups (**Supplementary Table 10**). We found that *Bacillota*, *Bacteroidota*, and *Actinomycetota* were the predominant bacterial phyla across all groups (**Extended Data Fig. 12a**). Compared with the HFD group, GIVL-treated mice (HFD+HG) exhibited a distinct shift in microbial community structure, characterized by a reduced relative abundance of *Bacteroidales* and *Verrucomicrobial*, with an increased abundance in *Lactobacillales* (**Extended Data Fig. 12b**). The reduction in *Bacteroidales* and *Verrucomicrobiales* was also observed following acarbose treatment⁶⁶, whereas the increase in *Lactobacillales* appeared to be a specific feature of GIVL intervention.

To further define GIVL treatment-associated changes in the gut microbiome, we performed Linear Discriminant Analysis Effect Size (LEfSe) analysis to compare microbial features across the different treatments. We found that the abundance of *Fecalibaculum rodentium*, a short-chain fatty acid-producing bacterium⁶⁷, was significantly increased in all HFD-related groups (LDA score > 4.0) (**Extended Data Fig. 13**). Notably, GIVL treatment further increased the abundance of this bacterium, increasing its fecal abundance to approximately two-fold that of the HFD group (**Fig. 4a**). These results suggest that the modulation of *F. rodentium* is not solely dependent on diet but may also be specifically regulated by GIVL. In addition to *F. rodentium*, several other SCFA-producing taxa were enriched in the GIVL-treated group compared with the HFD group. For example, the relative abundance of a Lachnospiraceae bacterium, a new species associated with butyrate production^{68,69}, increased from 0.44% in the HFD group to 2.24% in the HFD+HG group. Similarly, *Bifidobacterium pseudolongum*, which is linked to acetate production⁷⁰, increased from 5.97% to 8.96%, and several *Lactobacillus* species⁷¹⁻⁷³, including *L. reuteri*, *L. johnsonii*, and *L. taiwanensis*, also showed marked increases (**Supplementary Table 11**).

Given the enrichment of SCFAs-associated taxa, we further quantified fecal SCFAs among the different groups and found that the levels of acetate, butyrate and propionate were significantly higher in the HFD+HG group compared to those in the HFD group (**Fig. 4b-d**). The increase in SCFAs was associated with a lower fecal pH with a strong inverse correlation between total SCFAs and fecal pH (7.32 ± 0.16 in the HFD+HG group and 8.14 ± 0.22 in the HFD group). Intake of an HFD diet is known to increase fecal pH⁷⁴, while GIVL or Acarbose treatment reduced fecal pH significantly (**Fig. 4e-g**). SCFAs are known to stimulate glucagon-like peptide 1 (GLP-1) secretion by binding to G-protein-coupled receptors (GPCRs), thereby contributing to the improvement of metabolic syndrome^{75,76}. In line with this, elevated fecal SCFAs in the HFD+HG group were accompanied by higher plasma GLP-1 levels compared to the HFD group following the intervention (**Fig. 4h**). However, ELISA and Western blot analyses based on experiments using NCI-H716 cells, a well-established model for studying GLP-1 secretion⁷⁷, indicated that GIVL did not significantly promote GLP-1 secretion (**Fig. 4i, j**), arguing against a direct effect on enteroendocrine cells in modulating the regulation of glucose metabolism contributing to improved glucose tolerance. Instead, these results suggest that GIVL may exert its glucose-lowering effects by

targeting SI to alter intestinal substrate composition, while concurrently modulating the proliferation of selected bacteria in the gut microbiota. The coordinated action of these mechanisms may enhance the production of SCFAs, which subsequently promotes GLP-1 secretion and contributes to improved glucose tolerance.

283

284 **GIVL acts as a secreted quorum sensing peptide from *Clostridium tyrobutyricum***

285 The peptide GIVL was predicted to originate from the *agrD* encoded pro-peptide of *C. tyrobutyricum*, a
286 nonpathogenic bacterium, but the presence of QSPs in *C. tyrobutyricum* has not been reported⁷⁸⁻⁸⁰. To
287 determine whether the peptide GIVL, in addition to functioning as a strong inhibitor of SI activity, also
288 possesses quorum-sensing (QS) activity modulating gut microbiota, which might indicate additional
289 potential activities of *C. tyrobutyricum*, we first investigated whether GIVL is an endogenous signaling
290 peptide produced by *C. tyrobutyricum* and regulated by the *agr* system. To this end, we analysed growth
291 medium from cultures of *C. tyrobutyricum* by high-resolution Orbitrap liquid chromatography-mass
292 spectrometry (LC-HRMS). We observed a dominant peak with a retention time of 7.02 min detected for
293 the wild-type strain, corresponding to a mass-to-charge ratio (*m/z*) of 1141.68360 Da (**Fig. 5a**). The *bn*
294 and *yn* ions produced by induced fragmentation of the parent ions matched exactly the chemically
295 synthesized standard of GIVL (**Fig. 5b**), confirming that GIVL is a naturally secreted product of this
296 bacterium. Further genetic evidence showed that GIVL was not detectable in the Δ *agrD* mutant strain
297 (**Fig. 5c-e**, **Extended Data Fig. 14a** and **Supplementary Table 12-15**). By contrast, in the complemented
298 strain Δ *agrD*-*agrD* (**Fig. 5c**, **Extended Data Fig. 14b** and **Supplementary Table 13, 14**), we observed a
299 dominant peak corresponding to GIVL, with a level comparable to that of the wild-type (**Fig. 5d, e**).
300 Moreover, we deleted *agrB*, responsible for the proteolytic processing of *AgrD* pro-peptide and secretion
301 of QSPs⁸¹, generating the strain Δ *agrB* (**Fig. 5c**, **Extended Data Fig. 14c** and **Supplementary Table 12-**
302 **15**). Similarly, GIVL was undetectable in the supernatant of this strain (**Fig. 5d, e**). Collectively, these
303 data revealed that the *agr* system is responsible for the production of quorum sensing peptide GIVL in *C.*
304 *tyrobutyricum*.

305 We next evaluated the potential of GIVL to function as a quorum sensing peptide. Previous studies
306 have shown that a high cell density of microbes induces the secretion of QSPs^{19,82}, and therefore, the
307 concentration of GIVL might be increased during high-density fermentation. To this end, we monitored
308 the concentration of GIVL of the wild-type *C. tyrobutyricum* strain over time. As expected, we found that
309 GIVL accumulated slowly in the early growth stage, but increased rapidly after the cell density OD₆₀₀
310 reached a threshold of 3.8 (**Extended Data Fig. 14d, e**), exhibiting a typical cell density-dependent
311 secretion pattern. It is also known that quorum sensing peptides in certain cases may function as an
312 inhibitor of biofilm formation and sporulation (QSi)⁸³⁻⁸⁵. Thus, by employing the Δ *agrD* strain to exclude
313 interference from endogenously synthesized GIVL, we observed that biofilm formation was significantly

inhibited at concentrations $\geq 100 \mu\text{M}$ (**Fig. 5f**), and sporulation was further reduced at concentrations $\geq 500 \mu\text{M}$ (**Extended Data Fig. 15**). Moreover, consistent with previous findings in *C. acetobutylicum*, where QSP modulate bioproducts synthesis⁸¹, GIVL treatment specifically enhanced butyrate production in *C. tyrobutyricum* ΔagrD strain (**Fig. 5g**).

Finally, to further explore potential mechanisms underlying GIVL-induced enrichment of *F. rodentium* in mice gut microbiota, we assessed the effects of various GIVL concentrations on the growth, QS modulation, and SCFAs production of *F. rodentium*. The results demonstrated that GIVL significantly promoted bacterial growth at concentrations $\geq 200 \mu\text{M}$, suggesting a direct growth-stimulatory effect (**Fig. 5h**). Notably, GIVL, as a QS molecule, significantly inhibited biofilm formation of *F. rodentium* at concentrations exceeding $50 \mu\text{M}$ (**Extended Data Fig. 16a**). Furthermore, when the concentrations of GIVL were $\geq 100 \mu\text{M}$, the production of SCFAs significantly increased (up to 10.6% at $200 \mu\text{M}$), particularly butyrate (**Extended Data Fig. 16b, c**). Collectively, these results suggest that GIVL-mediated enrichment of *F. rodentium* likely results from a combination of direct growth promotion and modulation of QS activity. These findings demonstrate that GIVL exhibits clear QS activity and is a naturally secreted peptide derived from the *C. tyrobutyricum* *agr* system, further indicating that *C. tyrobutyricum* may hold potential as a next-generation probiotics.

Discussion

In this study, we used an integrated AI-driven approach to identify 1,670 candidate QSPs from the human gut microbiome, greatly expanding the diversity of potential QSPs. Although there is emerging interest in harnessing the biomedical benefits of QSPs with possible impact on host infection, immunity, and metabolism^{13,32,33,86}, the discovery of novel QSPs and their therapeutic potential interpretation has been limited. Our AlphaFold3 structural modelling indicated that several cQSPs have the potential to target SI, a primary α -glucosidase in the small intestine. Guided by computational prediction, we successfully synthesized 46 cQSPs and found that 29 cQSPs exhibited α -glucosidase inhibitory activity.

Among them, three peptides showed potent SI inhibition without cytotoxicity *in vitro*. GIVL, a QSP produced by *C. tyrobutyricum*, emerged as the most prominent candidate. Structural modeling and binding analyses suggested that GIVL engages α -glucosidase through a binding mode distinct from that of acarbose, a clinically used bacterially derived α -glucosidase inhibitor for type 2 diabetes. Under matched assay conditions, GIVL exhibited a lower IC_{50} and stronger apparent binding to SI than acarbose, supporting its potential as a compact peptide scaffold for intestinal SI inhibition. The localization of GIVL to the small intestinal brush border further supports a mode of action in which GIVL acts locally at the epithelial interface. In diabetic mice, oral GIVL administration significantly reduced blood glucose levels and improved glucose homeostasis, insulin sensitivity, and diabetes-associated tissue injury, without overt toxicity. Mechanistically, our data supports inhibition of intestinal SI as a primary mode of action.

Beyond direct SI inhibition, GIVL also remodeled the gut microbial ecosystem. GIVL treatment enriched *F. rodentium*, a butyrate-producing commensal previously associated with improved lipid metabolism⁸⁷, regulation of glucose homeostasis⁸⁸, and delayed progression of metabolic dysfunction⁸⁹. Further mechanistic investigation revealed that higher concentrations of GIVL significantly affected growth, biofilm formation, and butyrate production of *F. rodentium in vitro*, suggesting that GIVL may promote its selective proliferation by interfering with QS regulatory and metabolic strategies in this bacterium. We observed an association between GIVL administration and elevated fecal SCFA levels, which coincided with the enrichment of *F. rodentium* and an increase in plasma GLP-1 levels, collectively underscoring a potential multi-faceted action of GIVL.

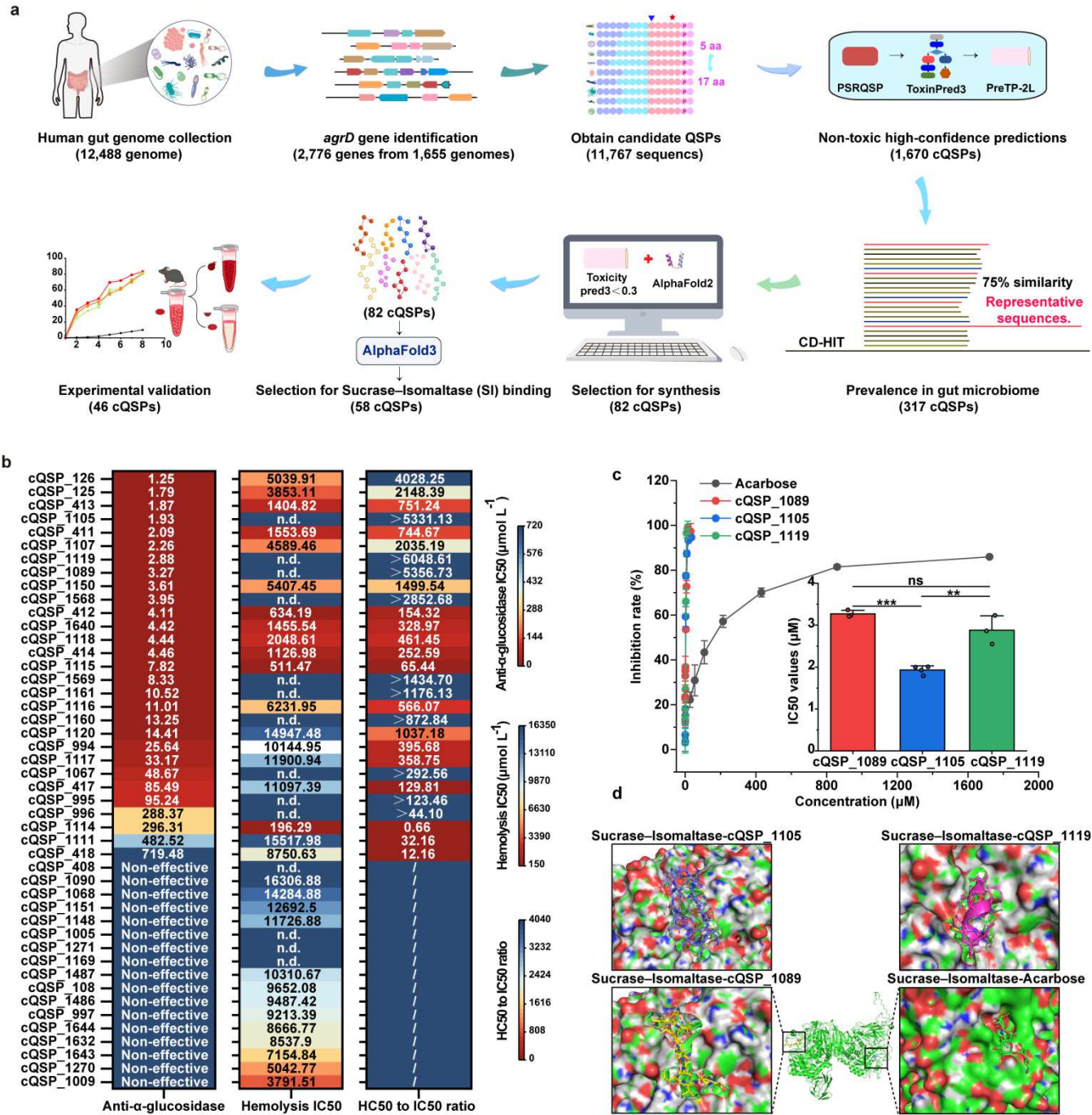
These observations support a two-tiered mechanism: GIVL first inhibits SI activity in the small intestine, increasing carbohydrate flux to the distal gut; this altered nutrient landscape then favors SCFA-associated bacteria, such as *F. rodentium*, thereby promoting SCFA production and engagement of the SCFA-GLP-1 axis. This mechanism is conceptually reminiscent of acarbose-mediated carbohydrate rerouting, although the initiating molecule, structure, and microbial regulatory effects differ. Finally, the contribution of additional bacterial taxa including *L. bacterium*, *B. pseudolongum*, and *Lactobacillus*, or higher-order microbial networks to SCFA restoration remains likely and warrants further investigation (**Extended Data Fig. 17**). We also acknowledge that, beyond *F. rodentium*, other bacterial taxa and/or more complex microbial networks may contribute to the observed changes in SCFAs levels, which warrants further investigation.

This study also clarifies the native form of GIVL. In canonical Agr systems, pro-peptides often undergo post-translational modifications processing, including AgrB-dependent modification, before becoming mature signaling peptides. For GIVL, mass spectrometry of desalted *C. tyrobutyricum* fermentation supernatants detected a peptide with the same molecular mass and MS/MS fragmentation pattern as chemically synthesized GIVL, supporting the conclusion that the naturally expressed peptide corresponds to the same linear, unmodified structure used in our functional studies. Together with genetic and functional data, these results support GIVL as an authentic QSP in *C. tyrobutyricum*.

The therapeutic implications of GIVL are notable but should be interpreted within the limits of the current study. Existing antidiabetic therapies, including α -glucosidase inhibitors such as acarbose and incretin-based drugs, can be constrained by tolerability and safety liabilities, particularly gastrointestinal adverse effects^{90,91}. In our mouse model, GIVL showed favorable short-term biocompatibility and no obvious toxicity during the 10-week intervention, while improving metabolic and histological outcomes. However, longer-term studies, dose-escalation experiments, pharmacokinetic analyses, immunogenicity assessment, and evaluation in additional preclinical models are required before translational potential can be established. It will also be important to determine whether GIVL acts exclusively through SI inhibition

383 and microbiota remodeling or whether additional epithelial, immune, or enteroendocrine pathways
384 contribute to its activity.

385 To the best of our knowledge, our study is the first to show the application potential of natural QSPs
386 for the treatment of metabolic diseases. More broadly, our findings suggest that gut microbiome-derived
387 QSPs can be leveraged as functional molecules for metabolic disease intervention. Because GIVL
388 originates from *C. tyrobutyricum*, future strategies could include engineered strains with enhanced GIVL
389 production, live biotherapeutic delivery systems, or optimized peptide analogues with improved stability
390 and potency. The cQSP library generated here may also contain additional peptides with activities beyond
391 α -glucosidase inhibition. Thus, this work provides both a candidate antidiabetic peptide and a
392 generalizable framework for systematically mining the gut microbiome for functional small peptides.
393 Further validation across disease targets, gut microbiomes, and translational models will determine the
394 broader utility of this discovery paradigm.



cQSPs from the 317 families for subsequent biological activity screening. Using AlphaFold3, we systematically predicted the binding complexes between all 82 cQSPs and representative proteins from these enzyme families, revealing that only sucrase-isomaltase (SI) exhibited high-confidence binding potential with 58 cQSPs, while the complex structures with other proteins were generally unstable or unreliable. Finally, we selected 46 cQSPs for chemical synthesis and *in vitro* validation. **b**, Comprehensive evaluation of *in vitro* activity and safety of cQSPs. This figure summarizes the multidimensional functional screening results of the candidate peptides. The left heatmap represents the IC₅₀ values (μmol⁻¹) of candidate QSPs for inhibiting α-glucosidase. “Non-effective” means that the peptide exhibited no α-glucosidase inhibitory activity *in vitro*. The values (μmol⁻¹) in the middle heatmap represent the cQSPs concentration that causes 50% hemolysis of mouse erythrocytes (HC₅₀). n.d. stands for not determined; no hemolytic activity was detected at a final concentration of 20 mg/mL. The column on the right displays the key therapeutic index HC₅₀/IC₅₀ ratio, which provides a visual representation of the safety window for cQSPs. All cQSPs are ordered primarily by IC₅₀ (ascending) and secondarily by HC₅₀, thereby grouping peptides with high inhibitory potency and favorable safety profiles together. The measured data in the heatmap are the mean value, and the experiments were performed in three independent replicates. **c**, Three cQSPs (cQSP_1089, cQSP_1105, and cQSP_1119) showed enhanced inhibitory effects against α-glucosidase relative to that of acarbose. **d**, Structural binding interaction of cQSPs (cQSP_1089, cQSP_1105, and cQSP_1119) with α-glucosidase. The binding interfaces of the protein-peptide or protein-acarbose complex were visualized using PyMOL. The relative docking positions of each complex were initially inferred using AlphaFold3 and further subjected to molecular dynamics relaxations. In each figure, the binding sites of α-glucosidase are visualized as protein surface representation in green, while acarbose and three peptides are presented in the mesh representation of the surface of the binding site and inside the ligand structures. Data are expressed as the mean ± SD. Statistical differences (ns for P > 0.05, *P < 0.05, **P < 0.01, and ***P < 0.001) were analysed using one-way ANOVA.

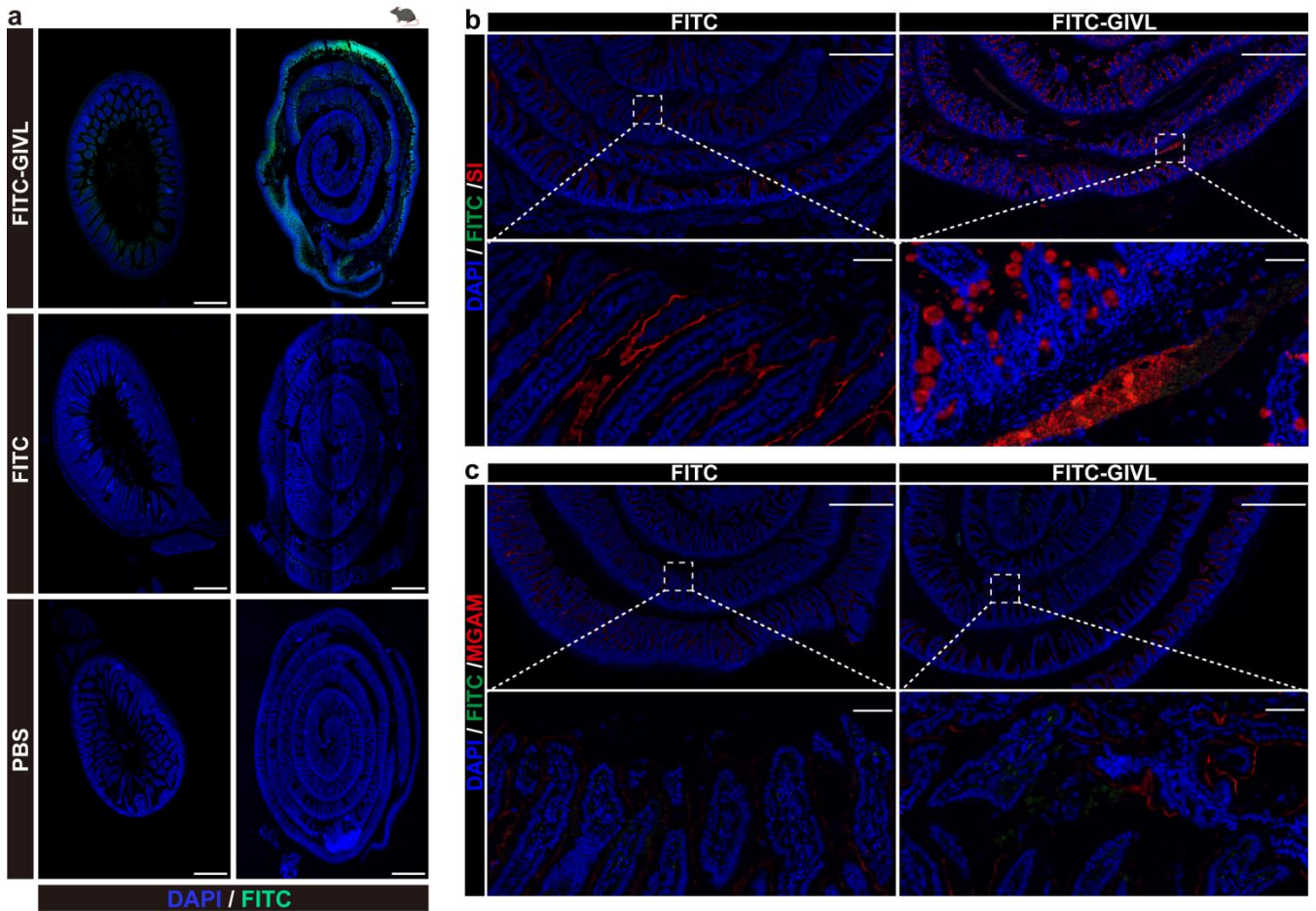


Fig. 2| GIVL specifically localizes to the small intestine and selectively binds to sucrase-isomaltase *in vivo*. **a**, Confocal microscopy showing that GIVL was specifically enriched in the small intestine of mice. The cell nucleus was stained with DAPI (blue). Intestinal cross sections (left, scale bars: 200 μ m), swiss rolls (right, scale bars: 500 μ m) and SPF mice were obtained 4 hours after oral gavage. **b-c**, Colocalization analysis of GIVL with sucrase-isomaltase (SI) (**b**) and maltase-glucoamylase (MGAM) (**c**) in the small intestine. GIVL selectively binds to SI (**b**) but not MGAM (**c**) in the small intestine. Triple immunofluorescence of the nucleus (blue), FITC/FITC-GIVL (green) and α -glucosidase (red dots located on the brush border of intestinal mucosal cells) in jejunal tissue sections.

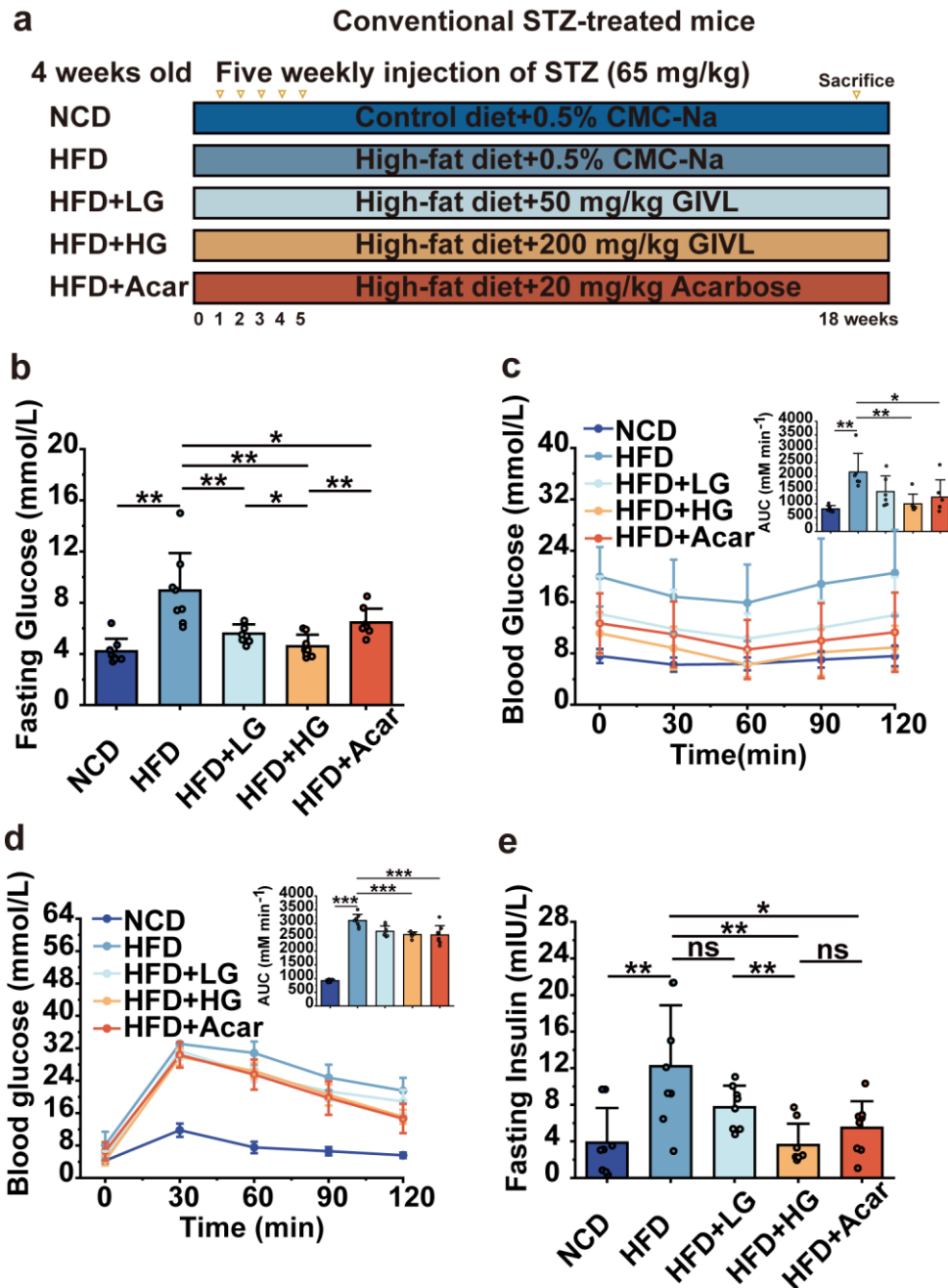


Fig. 3| GIVL treatment improves glucose homeostasis in T2D mice. a, Experimental design for experiments on streptozotocin (STZ) induced T2D mice. Mice were randomly divided into 5 groups, $n = 6-8$ in each group. (NCD: normal chow diet; HFD: T2D mice fed high-fat diet; HFD+LG: T2D mice fed a high-fat diet with GIVL at $50 \text{ mg kg}^{-1} \text{ BW/day}$; HFD+HG: T2D mice fed a high-fat diet with GIVL at $200 \text{ mg kg}^{-1} \text{ BW/day}$; HFD+Acar: T2D mice fed a high-fat diet with acarbose at $20 \text{ mg kg}^{-1} \text{ BW/day}$). Oral gavage with GIVL significantly reduced fasting blood glucose (**b**) and plasma insulin (**e**) levels in diabetic mice. GIVL gavage led to significant improvements in insulin resistance (**c**) and glucose homeostasis (**d**) in diabetic mice. Data are expressed as the mean $\pm$ SD. Statistical differences (ns for $P > 0.05$, $*P < 0.05$, $**P < 0.01$, and $***P < 0.001$) were analysed using one-way ANOVA.

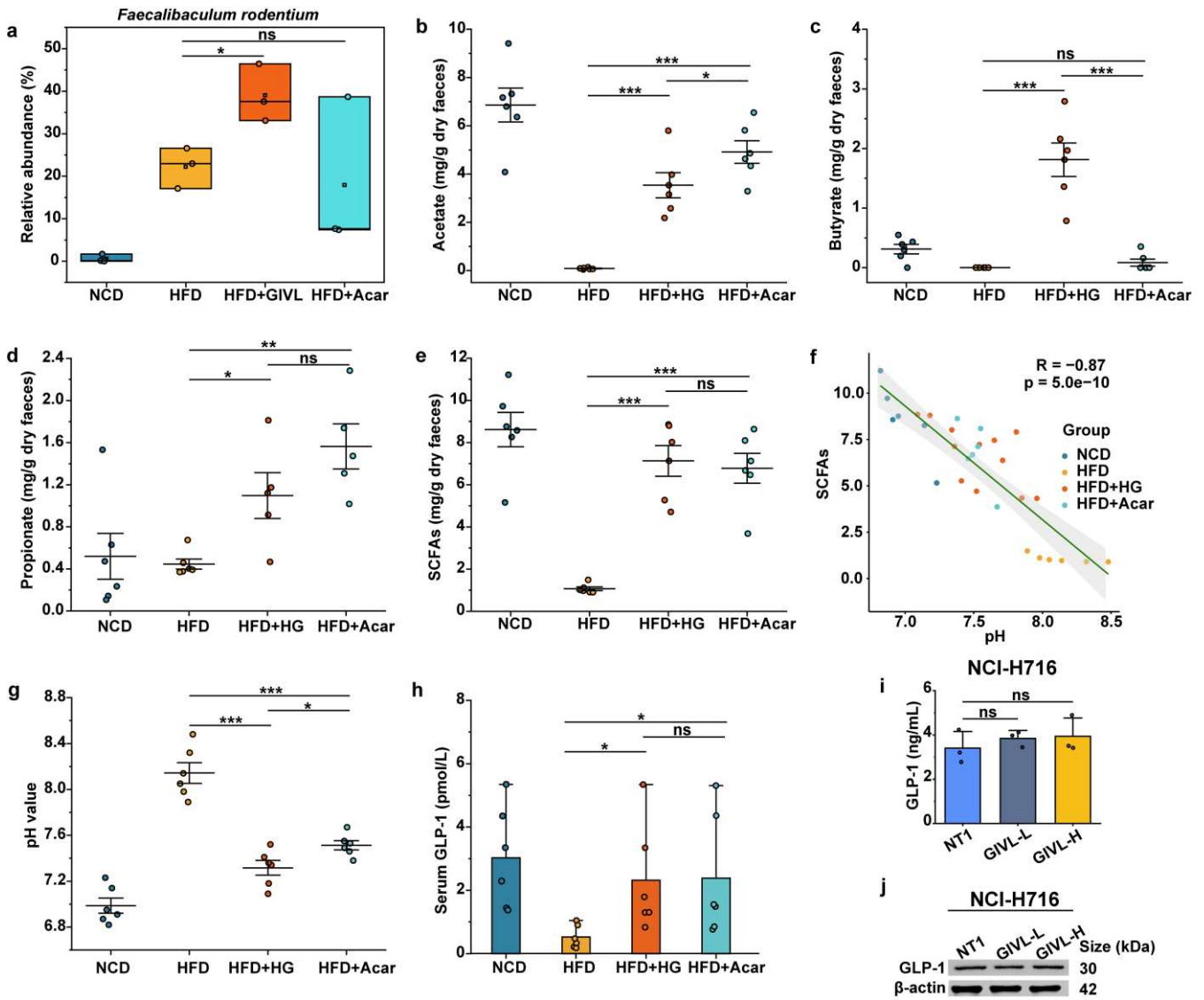


Fig. 4 | GIVL drives the selective enrichment of *F. rodentium* by reprogramming QS circuits and modulating metabolic networks. **a**, Relative abundance of *F. rodentium* in each group. **b-e**, fecal concentrations of acetate, butyrate, propionate, and SCFAs in different treatment groups were detected by gas chromatography. **f**, Correlation analysis of SCFAs and pH in the feces of each group. **g**, Comparison of fecal pH values among different treatment groups. **h**, GLP-1 levels from different treatment groups. **i, j**, Synthetic GIVL was dissolved in DMSO and added to NCI-H716 cells at the indicated final concentrations. The final DMSO concentration was maintained at 0.1% (v/v) across all groups, including the vehicle control (NT1). GIVL-L: 1 mg/mL GIVL in 0.1% DMSO; GIVL-H: 10 mg/mL GIVL in 0.1% DMSO. **i**, ELISA quantification of GLP-1 in culture supernatants after GIVL treatment at indicated concentrations. **j**, Western blot analysis of GLP-1 expression in NCI-H716 cells treated with GIVL. Data are expressed as the mean $\pm$ SD. Statistical differences (ns for $P > 0.05$, * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$) were analysed using one-way ANOVA.

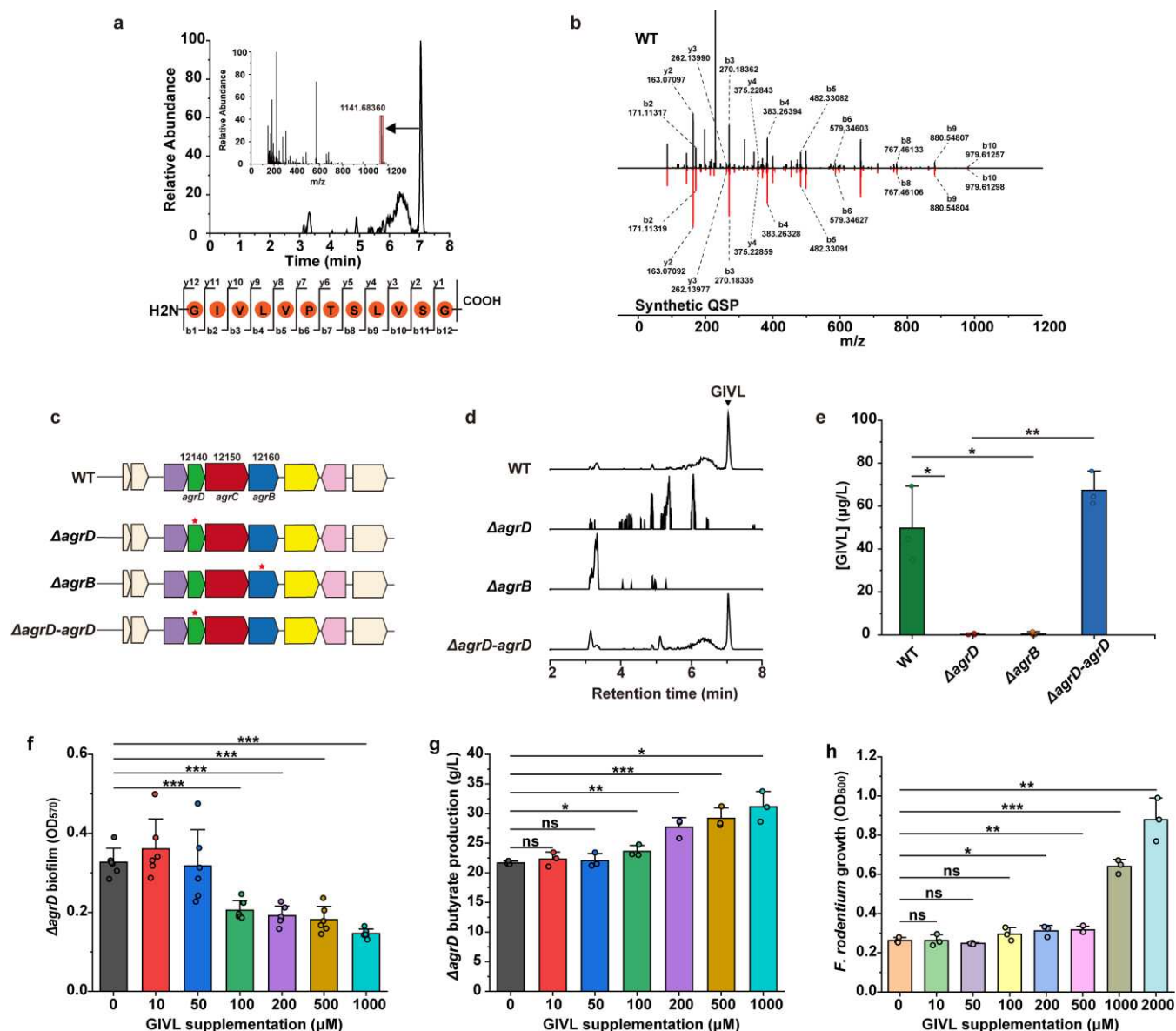


Fig. 5 | Identification of GIVL, a functional QSP secreted by *C. tyrobutyricum*. **a**, A dominant peak with a mass-to-charge ratio (m/z) of 1141.68 Da was identified in the supernatant of wild-type (WT) *C. tyrobutyricum*. **b**, LC-MS/MS analysis showed that the secondary mass spectrum of putative QSP (WT) exactly matched the QSP standard (Synthetic QSP), and the molecular formula (GIVLVPTSLVSG) was confirmed by precise mass measurements with a mass accuracy of less than 5 parts per million (ppm). **c**, Genes in the quorum-sensing pathway map to a conserved gene cluster. WT, wild-type; $\Delta agrD$, *agrD* deletion mutant; $\Delta agrB$, *agrB* deletion mutant; $\Delta agrD-agrD$, *agrD* complementary mutant. Gene designations are indicated as follows: *agrD*, hypothetical protein; *agrC*, histidine kinase; *agrB*, AgrB-like protein. Asterisks indicate genes targeted for deletion. **d**, *C. tyrobutyricum* mutants exhibit specific loss of GIVL production *in vitro*. WT and mutant strains of *C. tyrobutyricum* were cultured individually, and GIVL was analysed by LC-HRMS. Compared to WT, the $\Delta agrD$ and $\Delta agrB$ strains were deficient in GIVL production, but the dominant peak corresponding to GIVL was detected in the $\Delta agrD-agrD$ strain with a level comparable to that of WT. **e**, Quantification of GIVL in (**d**). **f-g**, Synthetic GIVL restored QS-associated phenotypes in the $\Delta agrD$ mutant in a dose-dependent manner, including biofilm formation (**f**) and butyrate production (**g**). Synthetic GIVL was dissolved in dimethyl sulfoxide (DMSO) and added to $\Delta agrD$ cultures at the indicated final concentrations prior to inoculation. The final DMSO concentration was kept at 1% (v/v) in all groups, including the 0 μM control. **h**, Synthetic GIVL promoted the growth of *F. rodentium* *in*

474 *vitro*, as assessed by OD₆₀₀, suggesting that GIVL may facilitate *F. rodentium* enrichment through direct growth
475 stimulation and QS modulation. Data are expressed as the mean $\pm$ SD. Statistical differences (ns for $P > 0.05$, * $P < 0.05$,
476 ** $P < 0.01$, and *** $P < 0.001$) were analysed using one-way ANOVA.

477 **Methods**

478 **Bacterial strains and culture conditions**

479 *C. tyrobutyricum* ATCC 25755 (KCTC 5387) was obtained from the American Type Culture Collection
480 (ATCC, Manassas, VA, USA) and cultured anaerobically at 37°C in reinforced clostridial medium
481 (RCM)⁹². 25 µg/mL methyl sulfonamides (Tm), 100 µg/mL D-cycloserine, or 40 mM lactose were added
482 to the RCM medium when required for a screening for positive clonal mutants. The *Escherichia coli*
483 CA434 (HB101) was employed as a conjugating donor strain for routine culture at 37°C in LB broth
484 supplemented with chloramphenicol (30 µg/mL) to ensure maintenance of the PMTL 82151 plasmid. *E.*
485 *coli* DH5α was used for routine cloning. All media were sterilized by autoclaving at 121°C for 15 mins.
486 *F. rodentium* (strain DSM 103405) was obtained from the Deutsche Sammlung von Mikroorganismen
487 und Zellkulturen (DSMZ) and cultured anaerobically at 37°C in modified peptone–yeast extract-glucose
488 (PYG) medium supplemented with 5 µg/mL hemin, 1 µg/mL vitamin K1, and 0.5 µg/mL cysteine
489 hydrochloride. Bacterial density was monitored by measuring the optical density at 600 nm (OD₆₀₀).
490 Culture supernatants were collected from overnight cultures grown under the same conditions described
491 above, passed through 0.25 µm filters, and immediately frozen for subsequent use.

492 493 **Constructing *C. tyrobutyricum* mutants**

494 The CRISPR array expression frame was assembled with PMTL82151 using Instant Sticky-end Ligase
495 (NEB)⁹³. The promoter, two repeat sequences, the PAM sequence, and the terminator sequence were
496 assembled with PMTL82151. The PAM sequence, which consists of 38bp, is located downstream of the
497 3' end of the target gene targeted for the Cas9 protein. Next, to construct a *C. tyrobutyricum* mutant,
498 PMTL82151 carrying the CRISPR assay expression frame was introduced into wild-type *C.*
499 *tyrobutyricum* by conjugation via *E. coli* CA434 and incubated at 37°C for 48 to 96 h until *C.*
500 *tyrobutyricum*-positive transformants appeared. Genomic DNA was randomly isolated from at least 8
501 colonies, and mutations were verified by sequencing with the wild-type strain as a control.

502 503 **Phenotypes in animal models**

504 C57BL/6 mice (male, 3-4 weeks old, purchased from Changsheng Biotechnology Co., Ltd., China) were
505 used in this study and were acclimatised for 1 week in individually ventilated cages (IVC) (12 h light-
506 dark cycle, 23±1 °C temperature and 60±5% humidity) provided with sterile water and standard chow.
507 All animal experiments involved in this study were conducted in accordance with the Guide for the Care
508 and Use of Laboratory Animals published by the US National Institutes of Health (8th edition, 2011) and
509 approved by the Local Research Ethics Review Board of the Animal Ethics Committee of the Dalian
510 University of Technology (No. DUTSBE250224-07).

The mouse model of type 2 diabetes mellitus (T2DM) was constructed as follows: after one week of adaptive feeding, mice were randomly divided into two groups, namely the normal control group (NCD group, n=6~8) and the model group (HFD group, n=24~32). Mice in the NCD group were fed with a standard maintenance diet, and mice in the HFD group were fed with a high-fat diet with 60% kcal fat. Mice in the HFD group were injected intraperitoneally with streptozotocin (STZ, S0130, 65 mg/kg body weight), and mice in the NCD group were injected with an equal volume of normal saline. The blood glucose level was monitored with a portable blood glucose meter (Accu-Chek, Roche, USA), on 7~14 days after the last STZ injection, and animals with blood glucose levels >11.1 mmol/L were considered diabetic. Mice in the model group were provided with 3% to 5% sucrose aqueous solution after STZ injection and fed with a high-fat diet during the whole experiment period. Animals in the control group were given normal drinking water and standard feed throughout the experiment.

Normal mice and mice with successful modelling were randomly divided into 5 groups (n=6~8 per group): 1) NCD group: 0.5% sodium carboxymethyl cellulose (CMC-Na) by gavage+standard maintenance diet; 2) HFD group: 0.5% CMC-Na by gavage+high-fat chow; 3) HFD+LG group: 50 mg/kg GIVL (dissolved in 0.5% CMC-Na) by gavage+high-fat chow; 4) HFD+HG group: 200 mg/kg GIVL (dissolved in 0.5% CMC-Na) by gavage+high-fat chow; 5) HFD+Acar group: 20 mg/kg acarbose (dissolved in 0.5% CMC-Na) by gavage+high-fat chow. Mice in each group were orally gavaged daily for 10 weeks, during which all mice were free to drink and eat, and fasting blood glucose (FBG) values were recorded regularly. During the experiment, the bedding was changed every 2 days, the cages were kept dry and clean to avoid infection, and the mice were observed for their mental status, feeding and drinking. After the intervention, mice were euthanized by cervical dislocation. All serum and tissue samples were collected for further analysis.

Computational screening of bioactive quorum sensing peptides

We downloaded 12,488 non-redundant human gut microbial genomes from four human gut microbiome databases, including the fecal Microbiota Transplantation (FMT) database, the CGR2 database, the UHGG2 database, and the BIO-ML database. The 12,488 genomes, including 5,189 isolated genomes and 3,835 metagenome-assembled genomes (MAGs), represent the majority of high-quality genomes obtained from human gut microbiome culturomics and metagenomics. The coding sequences (CDSs) for each genome were predicted by Prokka (v1.14.6)⁹⁴. The hidden Markov model (HMM) TIGR04223.1 for the *agrD* genes was downloaded from the Protein Families (Pfam) database⁹⁵. The predicted CDSs were probed against the TIGR04223.1 module by using HMMsearch (v3.1b2) with the “--cut_ga” parameter⁹⁶. Then, multiple sequence alignment for the searched AgrD protein sequences was conducted by MAFFT (v7.407)⁹⁷, and the candidate AgrD-derived pro-peptide sequences were anchored to the 17 amino acids in front of the conserved Proline of the AgrD protein. In total, 2,776 pro-peptide sequences were obtained,

546 and then 5~17 amino acid (aa) sequences were cut off from the beginning of each pro-peptide, and 11,767
547 fragment peptides were generated.

548 *In silico* biochemical and toxicity analyses of the 11,767 peptides were conducted by ToxinPred3⁵⁵.
549 In addition, we conducted quorum sensing peptides (QSPs) prediction by using machine-learning-based
550 tools PSRQSP⁵³ for peptides with length > 7 aa and PreTP-2L⁵⁴ for peptides with length ≤ 7 aa. The
551 peptide sequences that were not identified as QSPs by these two tools were further manually evaluated
552 using QSPepPred and TPpred-PepPA. Finally, a total of 1,670 peptides with predictive cytotoxicity score
553 < 0.35 were predicted as high-confidence candidate quorum sensing peptides (cQSPs). Then all 11,767
554 fragment peptides were hierarchically clustered after alphabet reduction by CD-Hit (v4.8.1) with 75%
555 identity cutoffs as previously described⁹⁸, and clusters representing at least three fragment peptides were
556 considered as prevalent QSPs in the human gut microbiome. The potential structures of peptides were
557 determined by AlphaFold2, and the peptides with alpha helix structures were selected based on folding
558 abilities⁹⁹. Transmembrane translocation pathways were modelled, and the membrane-penetrating
559 properties were calculated by PerMM¹⁰⁰. The grand average of hydropathicity was evaluated to assess the
560 stability and solubility of peptides¹⁰¹. We selected the cQSPs which were representative sequences of
561 abundant families with low cytotoxicity scores (< 0.3), folding ability for the downstream analysis, and
562 finally, we selected 89 cQSPs for chemical synthesis.

563

564 **AlphaFold3 structural modelling**

565 To investigate potential interactions between candidate QSPs and intestinal extracellular enzymes, we
566 employed AlphaFold3¹⁰², a state-of-the-art computational method, to predict the structures of 82 peptides
567 against 12 representative intestinal enzymes from oligosaccharidases and lipases (**Table S3**). For each
568 protein-peptide pair, five independent structure predictions were performed using randomly selected
569 seeds (between 1 and 5000) to account for stochastic variation in model initialization. The best model
570 was selected based on the ipTM score and overall ranking score provided by AlphaFold3. The peptide-
571 protein complexes with an ipTM score > 0.6 and a ranking score > 0.65 were considered high-confidence
572 predictions. Furthermore, we summarized the number of peptides predicted to interact with each enzyme,
573 enabling a systematic assessment of protein-peptide interaction potential for the 12 representative
574 enzymes.

575

576 **Rosetta energy computation and interface analysis**

577 To evaluate the contribution of individual residues of GIVL to SI binding, we performed *in silico* alanine-
578 scanning mutagenesis. Each residue of GIVL was individually substituted with alanine, generating 12
579 single-point Ala mutant peptides. We used the Rosetta relax function to refine the AlphaFold3 predicted
580 protein-peptide structures by minimizing their energy by optimizing backbone and sidechain torsion

angles while keeping the overall conformation close to the starting structure. To assess the contribution of individual residues, the protein-peptide complex structures of the 12 amino acid mutants of GIVL with SI were also predicted and refined using the above method. The binding free energy and interface properties of the wild-type and 12 Ala-mutant SI-GIVL complexes were evaluated using the Rosetta InterfaceAnalyzer¹⁰³, and Rosetta energies were calculated using Rosetta score function ref2015^{104,105}. Structural interfaces for the protein-peptide relaxed complexes were analyzed and visualized using PyMol (v3.1.1)¹⁰⁶. The hydrogen bonds are defined as ideal when the distance between the donor and acceptor atoms is smaller than 3.6Å. While for non-ideal hydrogen bonds, the corresponding distance is set to 3.6-4Å. For donor-donor or acceptor-acceptor atom pairs, the interaction is considered as an electrostatic clash when the distance between them is smaller than 3.6Å. Among the 12 alanine mutants, G1A, V5A, and P6A showed the most pronounced impairment in SI binding compared with the wild-type complex. The wild-type SI-GIVL complex had a dG_{separated} score of -53.929, whereas the G1A, V5A, and P6A mutants showed substantially less favorable dG_{separated} scores of -34.735, -34.516, and -26.509, respectively. These three mutants also showed a marked reduction in the number of ideal hydrogen bonds at the binding interface, from 14 in the wild-type complex to 3, 5, and 3 in the G1A, V5A, and P6A complexes, respectively. These changes indicate weakened interfacial interactions and suggest that Gly1, Val5, and Pro6 are key residues contributing to the stability of the SI-GIVL complex.

Molecular dynamics simulations

To further evaluate the stability and binding energetics of the ligand-protein complexes, we performed molecular dynamics (MD) simulations for each complex involving different ligands and α -glucosidase. The structure of the ligand acarbose was obtained from the PubChem database¹⁰⁷, and its binding complex with SI was simulated by AutoDock Vina. Subsequently, the AlphaFold3 predicted, and Rosetta refined protein-peptide complexes for three promising peptides (cQSP_1089, cQSP_1105, and cQSP_1119) were used as initial structures for MD simulations. These simulations included processes such as solvation, neutralization, and energy minimization, which helped to obtain more stable structures. The simulations were carried out using CHARMM36 force fields¹⁰⁸. The solvated complexes were subsequently used as the initial structures for 100 ns MD simulations, and the trajectories were analyzed for the Rg and structural stability assessment. Binding free energies were calculated using the MM-PBSA method implemented in the gmx_mmpbsa package of GROMACS¹⁰⁹.

Selection of source strains for cQSP functional validation

To ensure experimental feasibility for downstream genetic, biosynthetic, and functional validation, we evaluated the source strains of the three most promising QSPs (cQSP_1089, cQSP_1105, and cQSP_1119) identified in our study. cQSP_1089 is derived from *Clostridium tyrobutyricum*, a strain that

is readily cultivable, grows stably under laboratory conditions, and is supported by a well-established genetic manipulation system¹¹⁰⁻¹¹³. In contrast, the strains producing cQSP_1105 (*Hominimerdicola aceti*) and cQSP_1119 (*Lachnospira eligens*) present substantial technical challenges. Members of the Lachnospiraceae family, including *L. eligens*, are difficult to cultivate due to limited isolate availability, poor growth in liquid media, and high sensitivity to environmental stressors such as acid stress¹¹⁴⁻¹¹⁶. Despite repeated attempts, we were unable to reliably revive the deposited strains from CGR2 isolates to obtain stable cultures suitable for genetic validation. Similarly, *H. aceti* is a recently described species⁴³. Although strains of this bacterium are available from public culture collections, no established genetic manipulation protocols or gene-editing studies have been reported to date. Consequently, while both cQSP_1105 and cQSP_1119 exhibited potent *in vitro* inhibitory activity, the inability to culture their source strains precluded experimental validation of their QS function in native hosts. Based on a comprehensive evaluation of both efficacy and experimental tractability, cQSP_1089 (GIVL) was therefore selected as the model molecule for further studies.

Peptide synthesis

The peptides used in this study were synthesized by Shanghai Gill Biochemical Co., Ltd. (Shanghai, China) using a solid phase synthesis method in a sequence from the C-terminus (carboxy-terminus) to the N-terminus (amino-terminus)¹¹⁷. The purity of all peptides was determined by high-performance liquid chromatography, and purity was greater than 95%, with the exact molecular weight determined by mass spectrometry.

α -glucosidase inhibition assay

α -glucosidase (≥ 10 units mg^{-1}) was obtained from Sigma-Aldrich Chemical Co., Ltd (Milwaukee, WI, USA). Analysis of the α -glucosidase inhibitory activity of each peptide was performed according to the previously reported method with appropriate modifications¹¹⁸. Briefly, a mixture of 20 μL α -glucosidase solution (0.2 U/mL dissolved in 0.1 M PBS, pH 6.8), 20 μL of the peptide sample solution at different concentrations (0~13.33 mg/mL in PBS), and 80 μL of PBS (0.1 M, pH 6.8) were added to 96 microwell plates and incubated at 37 °C for 10 min. Subsequently, the reaction was initiated by adding 20 μL of 10 mM pNPG (substrate), followed by incubation at 37 °C for an additional 20 min. The reaction was terminated by adding 100 μL of 0.2 M sodium carbonate. The absorbance of the solution was measured at 405 nm using a multi-function plate reader (SpectraMax M2e, Molecular Device, USA). Acarbose served as the positive control. Under experimental conditions, 50% of the peptide concentration required for α -glucosidase inhibitory activity is defined as the IC_{50} value. The inhibition degree of the peptide was calculated according to the following formula:

$$\text{Inhibition degree (\%)} = 1 - \frac{(A_s - A_b)}{A_c} \times 100\%$$

Where A_s , A_b and A_c denote the absorbance values of the peptide sample group, blank group and control group in the reaction system, respectively.

Hemolysis assay of QSPs

The hemolysis assay of cQSPs was performed according to the procedure described previously¹¹⁹. Briefly, 100 μ L of erythrocyte solution (6%, v/v) was mixed with 100 μ L of peptide solution of various concentrations (0~40 mg/mL) and incubated at 37°C for 2 h. The mixture was then centrifuged at 2000 rpm for 10 min at 4 °C, and the absorbance of the supernatant was measured at 540 nm (SpectraMax M2e, Molecular Device, USA). Erythrocytes in PBS and 1% Triton X-100 (100% hemolysis) treated erythrocytes were used as the negative and positive controls, respectively. All experiments were performed with three independent replicates. The percentage hemolysis was calculated as follows:

$$\text{Hemolysis(\%)} = \frac{\text{Absorbance of the sample} - \text{Absorbance of the negative}}{\text{Absorbance of the positive} - \text{Absorbance of the negative}} \times 100\%$$

Cell lines

NCI-H716 cells used in this study were purchased from the ATCC Cell Bank (Fuheng Biotechnology Co., Ltd., Shanghai, China, Cat: FH0013). NCI-H716 cells were cultured in 1640 basal medium (NCI-H716 CM) supplemented with 10% FBS and 1% penicillin-streptomycin. The cell line was cultured in an incubator at 37°C with 5% CO₂ and was passaged when cell confluence reached 80%-90%. To evaluate the effect of GIVL on GLP-1, cells were seeded in culture plates and stimulated after overnight incubation. The experimental groups were as follows: vehicle control 1 (0.1% DMSO, corresponding to GIVL treatment, NT1), low-concentration GIVL (1 mg/mL, GIVL-L) and high-concentration GIVL (10 mg/mL, GIVL-H). The secretion level of GLP-1 in the cell supernatant was detected using an active GLP-1 ELISA kit (JL12839, Jianglai Biotechnology, Shanghai, China) according to the manufacturer's instructions. In parallel, cell pellets were harvested for protein expression analyses. The expression level of the target protein was detected by Western blot (WB), with β -actin as an internal control.

Kinetics of α -glucosidase inhibition

The inhibition kinetics of GIVL against α -glucosidase were determined by the Lineweaver-Burk method. Different concentrations of GIVL (0, 110, 220, and 440 μ M) were mixed with different concentrations of the reaction substrate pNPG (0.5, 1, 2, 2.5, 3, and 4 mM) and allowed to react; α -glucosidase (0.2 U/mL) was used as a control to determine the inhibition mode on α -glucosidase. The maximum reaction rate (V_{max}) and Michaelis–Menten (K_m) were obtained according to the following Michaelis–Menten equation.

$$\frac{1}{V} = \frac{K_m}{V_{max}} \left(1 + \frac{[I]}{K_i} \right) \frac{1}{[S]} + \frac{1}{V_{max}} \left(1 + \frac{[I]}{K_{is}} \right) \quad (1)$$

The attached drawings refer to the following formula:

$$\text{Slope} = \frac{K_m}{V_{max}} + \frac{K_m[I]}{V_{max}K_i} \quad (2)$$

and

$$\text{Y-intercept} = \frac{1}{V_{max}} + \frac{[I]}{V_{max}K_{is}} \quad (3)$$

Measurement of fluorescence spectra

The endogenous emission fluorescence of GIVL on α -glucosidase was studied using a steady-state fluorescence spectrometer (FLS1000, Edinburgh, UK). The mixed system of α -glucosidase with different concentrations of GIVL (0, 1.32, 2.63, 5.26, 9.64, 19.28, and 38.55 μ M) was first co-incubated at 25 °C in a constant temperature water bath for 10 min, and then 300 μ L of the reaction solution was quickly transferred into a micro quartz cuvette for assay. The emission wavelength was set from 290 to 390 nm, and the excitation wavelength was 280 nm, with a scan step of 60 nm/min and a slit width of 2.0/1.5 nm.

Circular dichroism (CD) spectroscopy

CD measurements were carried out on a Chirascan spectrometer (MOS 500, Bio-Logic, Claix, France) using quartz cells with a path length of 0.1 cm at 25 °C. Scans were conducted between wavelengths of 190 and 250 nm with a bandwidth of 2.0 nm and 1 nm/s scan speed. For the preparation of samples, briefly, a constant concentration of α -glucosidase (4.0 μ M) was maintained while the peptide GIVL concentration was varied from 0 to 48 μ M ([GIVL]/[α -glucosidase] = 0:1, 3:1, 6:1 and 12:1). All CD spectra were signal corrected with PBS (0.1 M pH=6.8) buffer under a constant flush with nitrogen and the content of the secondary structures of α -glucosidase was estimated with the BeStSel software¹²⁰.

FT-IR spectrum analysis, Infrared spectrum (FTIR) scanning

The FTIR scan of the peptide GIVL refers to the method of Wang et al¹¹⁸ with modifications. Briefly, the peptide sample was lyophilized with α -glucosidase in a certain proportion for FT-IR measurement with a scanning wavenumber range of 4000 to 400 cm^{-1} at the resolution of 4 cm^{-1} , with 32 scans per spectrum.

Surface plasmon resonance

The CM5 chip-coupled protein assay was performed in triplicate at 25°C on a Biacore T200, with curve fitting and data analysis using the Biacore T200 evaluation software (GE Healthcare). Briefly, the CM5 chip surface was activated with a mixture of EDC and NHS (from the amino coupling kit) at a flow rate of 10 μ L/min for 20 min. The remaining carboxyl groups on the chip surface were then blocked with

ethanolamine, further flushing out the poorly bound ligands. GIVL stock solution (8.76 mM) was diluted to a range of concentrations (27.38, 13.69, 3.42, 0.43, and 0.21 μ M (all in 5% DMSO) and applied using a flow of 30 μ L min⁻¹ for 120 seconds in each run. Data from the sample channels were collected using the Biacore T200 evaluation software (version 3.1, GE Healthcare) and subtracted from the reference channel. Dissociation constant (KD) was obtained by globally fitting the data to a steady-state affinity model.

722

723 **Localization of GIVL after oral gavage**

724 To trace the transit of GIVL through the gastrointestinal tract, we synthesized fluorescence-labeled FITC-
725 GIVL. Male C57 mice received $\approx$ 200 μ L free PBS, free FITC, or FITC-GIVL (1mg/kg) by gavage. 4h
726 after oral gavage, the mice were sacrificed, and major organs were harvested. The fluorescence images
727 of the organs were obtained by using an LSM 980 (Zeiss, Oberkochen, Germany).

728

729 **Immunofluorescence of the small intestines**

730 Tissue segments were immediately removed, washed with PBS, and placed in 4% paraformaldehyde.
731 Immunofluorescence was performed using an MGAM antibody (1:500; 22195-1-AP; Proteintech) and an
732 SI antibody (sc-393424; Santa Cruz), respectively. Detection and labelling were performed using a
733 secondary antibody conjugated to CY3 goat anti-rabbit IgG (H+L) (1:300; GB21303; Servicebio) and
734 Goat Anti-Mouse IgG H&L (HRP) (1:5000; ab6789; Abcam). Cell nuclei were stained with 4',6-
735 diamidino-2-phenylindole (DAPI), and the fluorescence images were measured using a microscope slide
736 scanner (Nikon, Tokyo, Japan).

737

738 **Biochemical analysis**

739 Plasma was prepared using sodium heparin anticoagulant tubes manufactured by BD (Cat. No. 366480).
740 The plasma levels of TG, TC, LDLC, HDLC, ALT, and AST were measured using a commercial assay
741 kit (Nanjing Jian Cheng Institute of Biological Engineering Ltd., Nanjing, China) according to the
742 manufacturer's instructions. The plasma insulin, GLP-1, LPS and IL-6 were measured by using
743 commercial ELISA kits (Biotech, Shanghai, China). After overnight fasting, mice were gavaged with
744 glucose (2 g/kg body weight) for OGTT or intraperitoneally injected with insulin (0.75 U/kg body weight)
745 for ITT, then blood glucose was collected at 0, 30, 60, 90 and 120 minutes with glucose levels measured
746 using a glucose meter (Accu-Chek, Roche, USA). The area under the curve (AUC) of postprandial blood
747 glucose was calculated using the following formula: $AUC = 0.5 \times (PBG_0 + 2 \times PBG_{30} + 2 \times PBG_{60} +$
748 $2 \times PBG_{90} + PBG_{120}) \times 30$, where PBG_x represents the blood glucose concentration measured at x minutes
749 after oral gavage. The insulin resistance index (HOMA-IR) was evaluated using a homeostatic model
750 defined as fasting blood glucose (FBG, mmol/L) $\times$ fasting insulin (FINS, mIU/L)/22.5.

751 **Histological analysis**

752 Major organs (including the heart, liver, spleen, kidneys, and epididymis) were collected and fixed in 4%
753 paraformaldehyde (PFA). Hematoxylin and eosin (H&E) staining was performed and analysed using
754 CaseViewer 3.2.1.

755
756 **Metagenomic sequencing and bioinformatics analysis**

757 DNA was extracted from mouse fecal samples using the QIAamp-DNA Mini Kit (QSK) according to the
758 manufacturer's instructions and was sequenced using the Illumina Novaseq 6000 at Wekemo Tech Group
759 Co., Ltd. (Shenzhen, China). Next, we constructed a DNA-paired library with an insert size of 350 bp
760 from the tested DNA samples by using the NEB Next®Ultra™ DNA Library Prep Kit for Illumina (NEB,
761 USA). Finally, qualified libraries were performed for sequencing with 150 bp paired-end reads generated
762 through the Illumina Novaseq 6000 sequencing platform.

763 Raw data contains a certain percentage of low-quality data, which needs to be pre-processed to
764 ensure the accuracy and reliability of the results of subsequent analyses. The adaptor contamination and
765 low-quality reads were discarded from the raw data using *Trimmomatic* software, and the remaining reads
766 were filtered against the human genome to obtain the high-quality clean reads with Bowtie2^{121,122}. The
767 clean reads were annotated and classified using MetaPhlAn software (v4.1.0), and the taxonomic
768 abundance profiles were generated. Specifically, linear discriminant analysis (LEfSe) was performed to
769 compare the differences in species composition between samples¹²³.

770
771 **Fecal short-chain fatty acids (SCFAs) profiling**

772 Fecal SCFAs were quantified using gas chromatography/mass spectrometry (GC-MS) with an Agilent
773 DB-FFAP platform (G2790A). Briefly, 200 mg of fecal sample was suspended in a 1 mL hydrochloric
774 acid (1 M) for acidification, then vortexed and left for 10 min while shaking and mixing thoroughly. The
775 extract was ultrasonicated (40 kHz) at 4°C for 10 min and centrifuged at 12,000 rpm for 30 min. The
776 supernatant was extracted by adding 500 µL ethyl acetate and was filtered through a 0.22 µm nylon filter
777 (Millipore PVDF).

778
779 **Quantification of QSP secreted by *C. tyrobutyricum* using LC-HRMS**

780 The wild-type or mutant *C. tyrobutyricum* was inoculated in a 250 mL CGM anaerobic culture bottle at
781 37°C and 150 rpm. Samples were regularly collected and centrifuged at 12,000 rpm for 10 min. The
782 supernatant was desalted using Pierce C18 Tips (Thermo Fisher Scientific Inc., Massachusetts, USA),
783 and approximately 500 µL of each sample was filtered through a Millipore 0.22 µm membrane. The
784 compounds were then separated on a Vanquish UPLC system (Thermo Scientific, Germany) equipped
785 with a Hypersil GOLD VANQUISH column C18 (175 Å, 100 × 2.1 mm i.d, 1.9 µm), and mass

spectrometry analysis was performed on a Q Exactive plus mass spectrometer (Thermo Fisher Scientific, Germany) equipped with an electrospray ionization (ESI) source. Solvent A was 0.1% formic acid (v/v) in an aqueous solution, and solvent B was 0.1% formic acid (v/v) in acetonitrile. The gradients were set at a flow rate of 0.2 ml/min: 10% solvent B (0~1 min), 20% solvent B (1~5 min), 80% solvent B (5~7 min), and 10% solvent B (7~12 min). The area was normalized using the GIVL internal standard, and the supernatant GIVL concentration of the fermentation broth was quantified by comparison with a standard curve prepared from a series of GIVL standard dilutions.

Sporulation assay

Liquid sporulation assays were performed according to a previously described method with minor modifications^{124,125}. To induce sporulation, *C. tyrobutyricum* was inoculated into CBM-S broth (CBM supplemented with 0.5% w/v CaCO₃ and 5% w/v glucose) and incubated at 37°C for 5 days. After incubation, samples were collected from biological triplicate liquid pre-cultures. The 20 µL samples were heat-shocked at 80°C for 10 min, inoculated with serial dilutions on CBM plates, and colonies were counted after 24~48 h of 37°C incubation to calculate the number of heat-resistant colony-forming units per mL (cfu/mL). To evaluate the effect of GIVL supplementation on the liquid sporulation ability of the *ΔagrD* strain, Synthetic GIVL was dissolved in dimethyl sulfoxide (DMSO) and added to the CBM-S medium of *ΔagrD* prior to inoculation to achieve final concentrations ranging from 0 to 1000 µM. Additionally, the final DMSO concentration was maintained at 1% (v/v) in all treatment groups (including the 0 µM control) to exclude solvent effects.

Growth and SCFA analysis

The *Clostridium tyrobutyricum ΔagrD* mutant strain and *Faecalibaculum rodentium* were used to evaluate the effects of GIVL on bacterial growth and SCFA production under shake-flask culture conditions. Briefly, overnight cultures of the *C. tyrobutyricum ΔagrD* mutant and *F. rodentium* were inoculated at 5% (v/v) into 100 mL of fresh medium in shake flasks. The *C. tyrobutyricum ΔagrD* mutant was cultured in RCM or CGM medium, whereas *F. rodentium* was cultured in PYG medium. All cultures were incubated under anaerobic conditions at 37°C with shaking at 150 rpm. Synthetic GIVL was dissolved in dimethyl sulfoxide (DMSO) and added to the culture medium before inoculation at the indicated final concentrations. For the *ΔagrD* mutant, GIVL was added at final concentrations of 0–1000 µM. For *F. rodentium*, GIVL was added at final concentrations of 0–2000 µM. The final concentration of DMSO was maintained at 1% (v/v) in all groups, including the 0 µM control group. Bacterial growth was monitored by measuring the optical density at 600 nm (OD₆₀₀) using a spectrophotometer (Thermo Spectronic, USA). Cells in the fermentation broth were removed by centrifugation at 12 000×g for 10 min,

and supernatants were filtered through a 0.22 µm PVDF syringe filter. Filtered samples were then analysed for acids (acetic and butyric) by gas chromatography (Agilent G2790A, USA) following the procedures described elsewhere¹²⁶. Results of replicates were expressed as mean ± standard deviation (SD). One-way ANOVA was used for statistical comparison between groups and a *p*-value < 0.05 was considered significant.

Biofilm formation assay

Biofilm formation by *Clostridium tyrobutyricum* *ΔagrD* mutant strain and *Fecalibaculum rodentium* (strain DSM 103405) was assessed using a previously described method with minor modifications^{127,128}. A 40 mM GIVL stock solution was prepared in dimethyl sulfoxide (DMSO). Before inoculation, GIVL was added to the corresponding fresh medium at the indicated final concentrations, following the same concentration ranges used in the above growth and SCFA analysis. The final DMSO concentration was maintained at 1% (v/v) in all groups, including the 0 µM control group. Briefly, overnight cultures were diluted by inoculating 20 µL of bacterial culture into 180 µL of fresh medium containing the indicated concentrations of GIVL in each well of a 96-well plate, resulting in an initial cell density (OD₆₀₀) of approximately 0.2. The plates were then incubated statically at 37°C in an anaerobic chamber. After 20 h of incubation, the culture supernatant was carefully removed to preserve the attached biofilm. To fix the biofilm biomass, 200 µL of methanol was added to each well and incubated at room temperature for 15 min before discarding the methanol. Then, 200 µL of 0.3% crystal violet solution was added to each well, and the plates were stained at room temperature for 10 min. The staining solution was discarded, and the plates were washed three times with distilled water to remove unbound dye. Finally, 200 µL of 33% acetic acid was added to each well and incubated at room temperature for 10 min to solubilize the bound dye. Biofilm formation was then quantified by measuring the OD₅₇₀.

Data availability

Raw sequencing data of the mouse metagenome is available from the SRA database under accession number PRJNA1202381.

Code availability

All software and code sources used in this study are listed in Methods. The code used for the analyses in this study is available on GitHub (<https://github.com/BGI-Qingdao/Peptides>).

851 References

- 852 1 Wu, S. *et al.* Molecules-mediated bidirectional interactions between microbes and human cells.
853 *NPJ Biofilms Microbiomes* **11**, 38 (2025). <https://doi.org/10.1038/s41522-025-00657-2>
- 854 2 Aggarwal, N. *et al.* Microbiome and Human Health: Current Understanding, Engineering, and
855 Enabling Technologies. *Chem Rev* **123**, 31-72 (2023).
856 <https://doi.org/10.1021/acs.chemrev.2c00431>
- 857 3 Oliveira, R. A., Cabral, V., Torcato, I. & Xavier, K. B. Deciphering the quorum-sensing lexicon
858 of the gut microbiota. *Cell Host Microbe* **31**, 500-512 (2023).
859 <https://doi.org/10.1016/j.chom.2023.03.015>
- 860 4 Joos, R. *et al.* Examining the healthy human microbiome concept. *Nat Rev Microbiol* **23**, 192-
861 205 (2025). <https://doi.org/10.1038/s41579-024-01107-0>
- 862 5 Faith, J. J. *et al.* The long-term stability of the human gut microbiota. *Science* **341**, 1237439
863 (2013). <https://doi.org/10.1126/science.1237439>
- 864 6 Whiteley, M., Diggle, S. P. & Greenberg, E. P. Progress in and promise of bacterial quorum
865 sensing research. *Nature* **551**, 313-320 (2017). <https://doi.org/10.1038/nature24624>
- 866 7 Mukherjee, S. & Bassler, B. L. Bacterial quorum sensing in complex and dynamically changing
867 environments. *Nat Rev Microbiol* **17**, 371-382 (2019). [https://doi.org/10.1038/s41579-019-0186-](https://doi.org/10.1038/s41579-019-0186-5)
868 **5**
- 869 8 Ng, W. L. & Bassler, B. L. Bacterial quorum-sensing network architectures. *Annu Rev Genet* **43**,
870 197-222 (2009). <https://doi.org/10.1146/annurev-genet-102108-134304>
- 871 9 Kleerebezem, M., Quadri, L. E., Kuipers, O. P. & de Vos, W. M. Quorum sensing by peptide
872 pheromones and two-component signal-transduction systems in Gram-positive bacteria. *Mol*
873 *Microbiol* **24**, 895-904 (1997). <https://doi.org/10.1046/j.1365-2958.1997.4251782.x>
- 874 10 Miller, M. B. & Bassler, B. L. Quorum sensing in bacteria. *Annu Rev Microbiol* **55**, 165-199
875 (2001). <https://doi.org/10.1146/annurev.micro.55.1.165>
- 876 11 Milly, T. A. & Tal-Gan, Y. Targeting Peptide-Based Quorum Sensing Systems for the
877 Treatment of Gram-Positive Bacterial Infections. *Pept Sci (Hoboken)* **115** (2023).
878 <https://doi.org/10.1002/pep2.24298>
- 879 12 Janssens, Y. *et al.* PapRIV, a BV-2 microglial cell activating quorum sensing peptide. *Sci Rep*
880 **11**, 10723 (2021). <https://doi.org/10.1038/s41598-021-90030-y>
- 881 13 De Spiegeleer, A. *et al.* Bacterial Quorum-Sensing Peptides as Immune Modulators Present in
882 Systemic Circulation. *Biomolecules* **13** (2023). <https://doi.org/10.3390/biom13020296>
- 883 14 Medapati, M. R. *et al.* Bitter taste receptor T2R14 detects quorum sensing molecules from
884 cariogenic *Streptococcus mutans* and mediates innate immune responses in gingival epithelial
885 cells. *Faseb j* **35**, e21375 (2021). <https://doi.org/10.1096/fj.202000208R>
- 886 15 Williams, M. R. *et al.* Quorum sensing between bacterial species on the skin protects against
887 epidermal injury in atopic dermatitis. *Sci Transl Med* **11** (2019).
888 <https://doi.org/10.1126/scitranslmed.aat8329>
- 889 16 Wynendaele, E. *et al.* The quorum sensing peptide EntF* promotes colorectal cancer metastasis
890 in mice: a new factor in the host-microbiome interaction. *BMC Biol* **20**, 151 (2022).
891 <https://doi.org/10.1186/s12915-022-01317-z>
- 892 17 Wynendaele, E. *et al.* Quorum Sensing Peptides Selectively Penetrate the Blood-Brain Barrier.
893 *PLoS One* **10**, e0142071 (2015). <https://doi.org/10.1371/journal.pone.0142071>
- 894 18 Taga, M. E. & Bassler, B. L. Chemical communication among bacteria. *Proceedings of the*
895 *National Academy of Sciences* **100**, 14549-14554 (2003).
896 <https://doi.org/doi:10.1073/pnas.1934514100>
- 897 19 Verbeke, F. *et al.* Peptides as Quorum Sensing Molecules: Measurement Techniques and
898 Obtained Levels In vitro and In vivo. *Front Neurosci* **11**, 183 (2017).
899 <https://doi.org/10.3389/fnins.2017.00183>

900 20 Kumar, S. *et al.* Computational tools for exploring peptide-membrane interactions in gram-
901 positive bacteria. *Comput Struct Biotechnol J* **21**, 1995-2008 (2023).
902 <https://doi.org/10.1016/j.csbj.2023.02.051>

903 21 Murray, E. J. & Williams, P. Detection of Agr-Type Autoinducing Peptides Produced by
904 *Staphylococcus aureus*. *Methods Mol Biol* **1673**, 89-96 (2018). [https://doi.org/10.1007/978-1-](https://doi.org/10.1007/978-1-4939-7309-5_7)
905 [4939-7309-5_7](https://doi.org/10.1007/978-1-4939-7309-5_7)

906 22 Jenul, C. & Horswill, A. R. Regulation of *Staphylococcus aureus* Virulence. *Microbiol Spectr* **7**
907 (2019). <https://doi.org/10.1128/microbiolspec.GPP3-0031-2018>

908 23 Novick, R. P. & Geisinger, E. Quorum sensing in staphylococci. *Annu Rev Genet* **42**, 541-564
909 (2008). <https://doi.org/10.1146/annurev.genet.42.110807.091640>

910 24 Williams, P. Microbial Primer: agr-mediated quorum sensing in Gram-positive pathogens.
911 *Microbiology (Reading)* **171** (2025). <https://doi.org/10.1099/mic.0.001590>

912 25 Williams, P., Hill, P., Bonev, B. & Chan, W. C. Quorum-sensing, intra- and inter-species
913 competition in the staphylococci. *Microbiology (Reading)* **169** (2023).
914 <https://doi.org/10.1099/mic.0.001381>

915 26 Sadr, S., Zargar, B., Aucoin, M. G. & Ingalls, B. Construction and Functional Characterization
916 of a Heterologous Quorum Sensing Circuit in *Clostridium sporogenes*. *ACS Synth Biol* **14**, 4857-
917 4868 (2025). <https://doi.org/10.1021/acssynbio.5c00628>

918 27 Bronesky, D. *et al.* *Staphylococcus aureus* RNAIII and Its Regulon Link Quorum Sensing,
919 Stress Responses, Metabolic Adaptation, and Regulation of Virulence Gene Expression. *Annu*
920 *Rev Microbiol* **70**, 299-316 (2016). <https://doi.org/10.1146/annurev-micro-102215-095708>

921 28 Chen, N. *et al.* Autoinducing peptides regulate antibiotic production to potentially shape root
922 microbiome. *Sci Adv* **11**, eadw5076 (2025). <https://doi.org/10.1126/sciadv.adw5076>

923 29 Ji, G., Beavis, R. C. & Novick, R. P. Cell density control of staphylococcal virulence mediated
924 by an octapeptide pheromone. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 12055-12059 (1995).
925 <https://doi.org/10.1073/pnas.92.26.12055>

926 30 Ji, G., Beavis, R. & Novick, R. P. Bacterial interference caused by autoinducing peptide
927 variants. *Science* **276**, 2027-2030 (1997). <https://doi.org/10.1126/science.276.5321.2027>

928 31 Cheung, G. Y., Joo, H. S., Chatterjee, S. S. & Otto, M. Phenol-soluble modulins--critical
929 determinants of staphylococcal virulence. *FEMS Microbiol Rev* **38**, 698-719 (2014).
930 <https://doi.org/10.1111/1574-6976.12057>

931 32 Brown, M. M. *et al.* Novel Peptide from Commensal *Staphylococcus simulans* Blocks
932 Methicillin-Resistant *Staphylococcus aureus* Quorum Sensing and Protects Host Skin from
933 Damage. *Antimicrob Agents Chemother* **64** (2020). <https://doi.org/10.1128/aac.00172-20>

934 33 Song, L. *et al.* *Staphylococcus saprophyticus* prevents skin damage by inhibiting
935 *Staphylococcus aureus* quorum sensing. *Sci Rep* **15**, 13082 (2025).
936 <https://doi.org/10.1038/s41598-025-97044-w>

937 34 Wright, J. S., 3rd, Jin, R. & Novick, R. P. Transient interference with staphylococcal quorum
938 sensing blocks abscess formation. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 1691-1696 (2005).
939 <https://doi.org/10.1073/pnas.0407661102>

940 35 Wynendaele, E. *et al.* Quorumpeps database: chemical space, microbial origin and functionality
941 of quorum sensing peptides. *Nucleic Acids Res* **41**, D655-659 (2013).
942 <https://doi.org/10.1093/nar/gks1137>

943 36 Wu, S. *et al.* Machine learning aided construction of the quorum sensing communication
944 network for human gut microbiota. *Nat Commun* **13**, 3079 (2022).
945 <https://doi.org/10.1038/s41467-022-30741-6>

946 37 Zhang, J. *et al.* Large-scale biosynthetic analysis of human microbiomes reveals diverse
947 protective ribosomal peptides. *Nat Commun* **16**, 3054 (2025). [https://doi.org/10.1038/s41467-](https://doi.org/10.1038/s41467-025-58280-w)
948 [025-58280-w](https://doi.org/10.1038/s41467-025-58280-w)

949 38 Ma, Y. *et al.* Identification of antimicrobial peptides from the human gut microbiome using deep
950 learning. *Nature Biotechnology* **40**, 921-931 (2022). [https://doi.org/10.1038/s41587-022-01226-](https://doi.org/10.1038/s41587-022-01226-0)
951 [0](https://doi.org/10.1038/s41587-022-01226-0)

952 39 Torres, M. D. T. *et al.* Mining human microbiomes reveals an untapped source of peptide
 953 antibiotics. *Cell* **187**, 5453-5467.e5415 (2024).
 954 <https://doi.org/https://doi.org/10.1016/j.cell.2024.07.027>
 955 40 Ma, Y. *et al.* Efficient Mining of Anticancer Peptides from Gut Metagenome. *Adv Sci (Weinh)*
 956 **10**, e2300107 (2023). <https://doi.org/10.1002/adv.202300107>
 957 41 Fan, Y. *et al.* Polypeptides synthesized by common bacteria in the human gut improve rodent
 958 metabolism. *Nat Microbiol* **10**, 1918-1939 (2025). <https://doi.org/10.1038/s41564-025-02064-x>
 959 42 Aggarwala, V. *et al.* Precise quantification of bacterial strains after fecal microbiota
 960 transplantation delineates long-term engraftment and explains outcomes. *Nat Microbiol* **6**, 1309-
 961 1318 (2021). <https://doi.org/10.1038/s41564-021-00966-0>
 962 43 Lin, X. *et al.* The genomic landscape of reference genomes of cultivated human gut bacteria. *Nat*
 963 *Commun* **14**, 1663 (2023). <https://doi.org/10.1038/s41467-023-37396-x>
 964 44 Almeida, A. *et al.* A unified catalog of 204,938 reference genomes from the human gut
 965 microbiome. *Nat Biotechnol* **39**, 105-114 (2021). <https://doi.org/10.1038/s41587-020-0603-3>
 966 45 Poyet, M. *et al.* A library of human gut bacterial isolates paired with longitudinal multiomics
 967 data enables mechanistic microbiome research. *Nat Med* **25**, 1442-1452 (2019).
 968 <https://doi.org/10.1038/s41591-019-0559-3>
 969 46 Steiner, E., Scott, J., Minton, N. P. & Winzer, K. An agr quorum sensing system that regulates
 970 granule formation and sporulation in *Clostridium acetobutylicum*. *Appl Environ Microbiol* **78**,
 971 1113-1122 (2012). <https://doi.org/10.1128/aem.06376-11>
 972 47 Zhao, A. *et al.* Reconstitution of the *S. aureus* agr quorum sensing pathway reveals a direct role
 973 for the integral membrane protease MroQ in pheromone biosynthesis. *Proc. Natl. Acad. Sci. U.*
 974 *S. A.* **119**, e2202661119 (2022). <https://doi.org/10.1073/pnas.2202661119>
 975 48 Wang, B., Zhao, A., Novick, R. P. & Muir, T. W. Activation and inhibition of the receptor
 976 histidine kinase AgrC occurs through opposite helical transduction motions. *Mol Cell* **53**, 929-
 977 940 (2014). <https://doi.org/10.1016/j.molcel.2014.02.029>
 978 49 Wang, B., Zhao, A., Novick, R. P. & Muir, T. W. Key driving forces in the biosynthesis of
 979 autoinducing peptides required for staphylococcal virulence. *Proc. Natl. Acad. Sci. U. S. A.* **112**,
 980 10679-10684 (2015). <https://doi.org/10.1073/pnas.1506030112>
 981 50 Wang, B. *et al.* Functional Plasticity of the AgrC Receptor Histidine Kinase Required for
 982 Staphylococcal Virulence. *Cell Chem Biol* **24**, 76-86 (2017).
 983 <https://doi.org/10.1016/j.chembiol.2016.12.008>
 984 51 Xie, Q. *et al.* Identification of a Molecular Latch that Regulates Staphylococcal Virulence. *Cell*
 985 *Chem Biol* **26**, 548-558.e544 (2019). <https://doi.org/10.1016/j.chembiol.2019.01.006>
 986 52 Ni, N., Li, M., Wang, J. & Wang, B. Inhibitors and antagonists of bacterial quorum sensing.
 987 *Med Res Rev* **29**, 65-124 (2009). <https://doi.org/10.1002/med.20145>
 988 53 Charoenkwan, P. *et al.* PSRQSP: An effective approach for the interpretable prediction of
 989 quorum sensing peptide using propensity score representation learning. *Comput Biol Med* **158**,
 990 106784 (2023). <https://doi.org/10.1016/j.compbiomed.2023.106784>
 991 54 Yan, K., Guo, Y. & Liu, B. PreTP-2L: identification of therapeutic peptides and their types
 992 using two-layer ensemble learning framework. *Bioinformatics* **39** (2023).
 993 <https://doi.org/10.1093/bioinformatics/btad125>
 994 55 Rathore, A. S., Choudhury, S., Arora, A., Tijare, P. & Raghava, G. P. S. ToxinPred 3.0: An
 995 improved method for predicting the toxicity of peptides. *Comput Biol Med* **179**, 108926 (2024).
 996 <https://doi.org/10.1016/j.compbiomed.2024.108926>
 997 56 Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation
 998 sequencing data. *Bioinformatics* **28**, 3150-3152 (2012).
 999 <https://doi.org/10.1093/bioinformatics/bts565>
 1000 57 Lee, B. H. *et al.* Modulation of starch digestion for slow glucose release through "togglng" of
 1001 activities of mucosal α -glucosidases. *J Biol Chem* **287**, 31929-31938 (2012).
 1002 <https://doi.org/10.1074/jbc.M112.351858>

1003 58 van de Laar, F. A. Alpha-glucosidase inhibitors in the early treatment of type 2 diabetes. *Vasc*
 1004 *Health Risk Manag* **4**, 1189-1195 (2008). <https://doi.org/10.2147/vhrm.s3119>
 1005 59 Laube, H., Fouladfar, M., Aubell, R. & Schmitz, H. [Effect of glucosidase inhibitor, Bay g 5421
 1006 (acarbose), on the blood glucose in obese diabetic patients ty pe 2 (NIDDM) (author's transl)].
 1007 *Arzneimittelforschung* **30**, 1154-1157 (1980).
 1008 60 Schmidt, D. D. *et al.* alpha-Glucosidase inhibitors. New complex oligosaccharides of microbial
 1009 origin. *Naturwissenschaften* **64**, 535-536 (1977). <https://doi.org/10.1007/bf00483561>
 1010 61 Rockser, Y. & Wehmeier, U. F. The gac-gene cluster for the production of acarbose from
 1011 *Streptomyces glaucescens* GLA.O: identification, isolation and characterization. *J Biotechnol*
 1012 **140**, 114-123 (2009). <https://doi.org/10.1016/j.jbiotec.2008.10.016>
 1013 62 Wang, Y. J. *et al.* Actinoplanes utahensis ZJB-08196 fed-batch fermentation at elevated
 1014 osmolality for enhancing acarbose production. *Bioresour Technol* **103**, 337-342 (2012).
 1015 <https://doi.org/10.1016/j.biortech.2011.09.121>
 1016 63 Gu, Y. *et al.* Analyses of gut microbiota and plasma bile acids enable stratification of patients
 1017 for antidiabetic treatment. *Nat Commun* **8**, 1785 (2017). [https://doi.org/10.1038/s41467-017-](https://doi.org/10.1038/s41467-017-01682-2)
 1018 [01682-2](https://doi.org/10.1038/s41467-017-01682-2)
 1019 64 Lebovitz, H. E. ALPHA-GLUCOSIDASE INHIBITORS. *Endocrinology and Metabolism*
 1020 *Clinics of North America* **26**, 539-551 (1997). [https://doi.org/https://doi.org/10.1016/S0889-](https://doi.org/https://doi.org/10.1016/S0889-8529(05)70266-8)
 1021 [8529\(05\)70266-8](https://doi.org/https://doi.org/10.1016/S0889-8529(05)70266-8)
 1022 65 Smith, B. J. *et al.* Changes in the gut microbiome and fermentation products concurrent with
 1023 enhanced longevity in acarbose-treated mice. *BMC Microbiol* **19**, 130 (2019).
 1024 <https://doi.org/10.1186/s12866-019-1494-7>
 1025 66 Baxter, N. T., Lesniak, N. A., Sinani, H., Schloss, P. D. & Koropatkin, N. M. The Glucoamylase
 1026 Inhibitor Acarbose Has a Diet-Dependent and Reversible Effect on the Murine Gut Microbiome.
 1027 *mSphere* **4** (2019). <https://doi.org/10.1128/mSphere.00528-18>
 1028 67 Zagato, E. *et al.* Endogenous murine microbiota member Faecalibaculum rodentium and its
 1029 human homologue protect from intestinal tumour growth. *Nat Microbiol* **5**, 511-524 (2020).
 1030 <https://doi.org/10.1038/s41564-019-0649-5>
 1031 68 Garcia-Serrano, A. M. *et al.* Butyrate-producing bacteria as probiotic supplement: beneficial
 1032 effects on metabolism and modulation of behaviour in an obesity mouse model. *Benef Microbes*
 1033 **16**, 109-124 (2024). <https://doi.org/10.1163/18762891-bja00040>
 1034 69 Vacca, M. *et al.* The Controversial Role of Human Gut Lachnospiraceae. *Microorganisms* **8**
 1035 (2020). <https://doi.org/10.3390/microorganisms8040573>
 1036 70 Song, Q. *et al.* Bifidobacterium pseudolongum-generated acetate suppresses non-alcoholic fatty
 1037 liver disease-associated hepatocellular carcinoma. *J Hepatol* **79**, 1352-1365 (2023).
 1038 <https://doi.org/10.1016/j.jhep.2023.07.005>
 1039 71 Sun, L. *et al.* Limosilactobacillus reuteri CCFM1132 modulates gut microbiota and alleviates
 1040 hyperuricemia: a randomized, double-blind, placebo-controlled trial. *Food Science and Human*
 1041 *Wellness* **15**, 9250316 (2026). <https://doi.org/10.26599/FSHW.2024.9250316>
 1042 72 Chen, Y. The SCFAs Production of Syntrophic Culture of L. johnsonii SZ-YL and A.
 1043 Muciniphila in Different Macrobutrients. *Highlights in Science, Engineering and Technology*
 1044 (2023).
 1045 73 Nami, Y., Shaghaghi Ranjbar, M., Modarres Aval, M. & Haghshenas, B. Harnessing
 1046 Lactobacillus-derived SCFAs for food and health: Pathways, genes, and functional implications.
 1047 *Curr Res Microb Sci* **9**, 100496 (2025). <https://doi.org/10.1016/j.crmicr.2025.100496>
 1048 74 Chen, X., Wang, Z., Wang, D. & Kan, J. Effects of resistant starch III on the serum lipid levels
 1049 and gut microbiota of Kunming mice under high-fat diet. *Food Science and Human Wellness* **12**,
 1050 575-583 (2023). <https://doi.org/https://doi.org/10.1016/j.fshw.2022.07.060>
 1051 75 Huang, Q., Deng, H., Sun, M.-C. & Zhao, C. Food regulation of incretin axis: Molecular
 1052 mechanisms and clinical outlook in diabetes. *Trends in Food Science & Technology* **165**,
 1053 105326 (2025). <https://doi.org/https://doi.org/10.1016/j.tifs.2025.105326>

1054 76 Akash, M. S. H., Fiayyaz, F., Rehman, K., Sabir, S. & Rasool, M. H. Gut Microbiota and
1055 Metabolic Disorders: Advances in Therapeutic Interventions. *Crit Rev Immunol* **39**, 223-237
1056 (2019). <https://doi.org/10.1615/CritRevImmunol.2019030614>
1057 77 Reimer, R. A. *et al.* A human cellular model for studying the regulation of glucagon-like
1058 peptide-1 secretion. *Endocrinology* **142**, 4522-4528 (2001).
1059 <https://doi.org/10.1210/endo.142.10.8415>
1060 78 Liu, T., Zhu, L., Zhu, Z. & Jiang, L. Genome Sequence Analysis of *Clostridium tyrobutyricum*,
1061 a Promising Microbial Host for Human Health and Industrial Applications. *Curr Microbiol* **77**,
1062 3685-3694 (2020). <https://doi.org/10.1007/s00284-020-02175-0>
1063 79 Jiang, L. *et al.* Genome Sequence of *Clostridium tyrobutyricum* ATCC 25755, a Butyric Acid-
1064 Overproducing Strain. *Genome Announc* **1** (2013). <https://doi.org/10.1128/genomeA.00308-13>
1065 80 Ivy, R. A. & Wiedmann, M. in *Encyclopedia of Food Microbiology (Second Edition)* (eds
1066 Carl A. Batt & Mary Lou Tortorello) 468-473 (Academic Press, 2014).
1067 81 Zhu, C. *et al.* Elucidating the Biosynthesis and Function of an Autoinducing Peptide in
1068 *Clostridium acetobutylicum*. *Angew Chem Int Ed Engl* **64**, e202500904 (2025).
1069 <https://doi.org/10.1002/anie.202500904>
1070 82 Rutherford, S. T. & Bassler, B. L. Bacterial quorum sensing: its role in virulence and
1071 possibilities for its control. *Cold Spring Harb Perspect Med* **2** (2012).
1072 <https://doi.org/10.1101/cshperspect.a012427>
1073 83 Boles, B. R. & Horswill, A. R. Agr-mediated dispersal of *Staphylococcus aureus* biofilms. *PLoS*
1074 *Pathog* **4**, e1000052 (2008). <https://doi.org/10.1371/journal.ppat.1000052>
1075 84 Taha, M. N., Saafan, A. E., Ahmedy, A., El Gebaly, E. & Khairalla, A. S. Two novel synthetic
1076 peptides inhibit quorum sensing-dependent biofilm formation and some virulence factors in
1077 *Pseudomonas aeruginosa* PAO1. *J Microbiol* **57**, 618-625 (2019).
1078 <https://doi.org/10.1007/s12275-019-8548-2>
1079 85 Khan, M. A., Zhu, L. & Zhu, H. Peptide-based approaches to quorum-sensing disruption:
1080 emerging trends and applications in antimicrobial therapy. *Bioorganic & Medicinal Chemistry*
1081 **133**, 118496 (2026). <https://doi.org/https://doi.org/10.1016/j.bmc.2025.118496>
1082 86 Wynendaele, E. *et al.* The quorum sensing peptide EntF* promotes colorectal cancer metastasis
1083 in mice: a new factor in the host-microbiome interaction. *BMC Biology* **20**, 151 (2022).
1084 <https://doi.org/10.1186/s12915-022-01317-z>
1085 87 Yin, C. *et al.* *Schizophyllum commune* polysaccharide ameliorates diet-induced obesity by
1086 regulating the Faecalibaculum rodentium/indole-3-lactic acid/microRNA axis. *Carbohydr Polym*
1087 **371**, 124436 (2026). <https://doi.org/10.1016/j.carbpol.2025.124436>
1088 88 Guo, L. *et al.* Murine gut microbial interactions exert antihyperglycemic effects. *Isme j* **19**
1089 (2025). <https://doi.org/10.1093/ismejo/wraf028>
1090 89 Kawano, Y. *et al.* Microbiota imbalance induced by dietary sugar disrupts immune-mediated
1091 protection from metabolic syndrome. *Cell* **185**, 3501-3519.e3520 (2022).
1092 <https://doi.org/10.1016/j.cell.2022.08.005>
1093 90 Martin, A. E. & Montgomery, P. A. Acarbose: An α -glucosidase inhibitor. *American Journal of*
1094 *Health-System Pharmacy* **53**, 2277-2290 (1996). <https://doi.org/10.1093/ajhp/53.19.2277>
1095 91 Shurrab, N. T. & Arafa, E.-S. A. Metformin: A review of its therapeutic efficacy and adverse
1096 effects. *Obesity Medicine* **17**, 100186 (2020).
1097 <https://doi.org/https://doi.org/10.1016/j.obmed.2020.100186>
1098 92 Liu, X., Zhu, Y. & Yang, S.-T. Butyric acid and hydrogen production by *Clostridium*
1099 *tyrobutyricum* ATCC 25755 and mutants. *Enzyme and Microbial Technology* **38**, 521-528
1100 (2006). <https://doi.org/https://doi.org/10.1016/j.enzmictec.2005.07.008>
1101 93 Zhou, X. *et al.* Exploiting heterologous and endogenous CRISPR-Cas systems for genome
1102 editing in the probiotic *Clostridium butyricum*. *Biotechnol Bioeng* **118**, 2448-2459 (2021).
1103 <https://doi.org/10.1002/bit.27753>
1104 94 Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-2069
1105 (2014). <https://doi.org/10.1093/bioinformatics/btu153>

1106 95 Finn, R. D. *et al.* Pfam: the protein families database. *Nucleic Acids Research* **42**, D222-D230
1107 (2013). <https://doi.org/10.1093/nar/gkt1223>

1108 96 Eddy, S. R. Accelerated Profile HMM Searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).
1109 <https://doi.org/10.1371/journal.pcbi.1002195>

1110 97 Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-likelihood trees
1111 for large alignments. *PLoS One* **5**, e9490 (2010). <https://doi.org/10.1371/journal.pone.0009490>

1112 98 Santos-Júnior, C. D. *et al.* Discovery of antimicrobial peptides in the global microbiome with
1113 machine learning. *Cell* **187**, 3761-3778.e3716 (2024). <https://doi.org/10.1016/j.cell.2024.05.013>

1114 99 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-
1115 589 (2021). <https://doi.org/10.1038/s41586-021-03819-2>

1116 100 Lomize, A. L. *et al.* PerMM: A Web Tool and Database for Analysis of Passive Membrane
1117 Permeability and Translocation Pathways of Bioactive Molecules. *J Chem Inf Model* **59**, 3094-
1118 3099 (2019). <https://doi.org/10.1021/acs.jcim.9b00225>

1119 101 Wilkins, M. R. *et al.* Protein identification and analysis tools in the ExPASy server. *Methods*
1120 *Mol Biol* **112**, 531-552 (1999). <https://doi.org/10.1385/1-59259-584-7:531>

1121 102 Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3.
1122 *Nature* **630**, 493-500 (2024). <https://doi.org/10.1038/s41586-024-07487-w>

1123 103 Fleishman, S. J. *et al.* RosettaScripts: a scripting language interface to the Rosetta
1124 macromolecular modeling suite. *PLoS One* **6**, e20161 (2011).
1125 <https://doi.org/10.1371/journal.pone.0020161>

1126 104 Rubenstein, A. B., Blacklock, K., Nguyen, H., Case, D. A. & Khare, S. D. Systematic
1127 Comparison of Amber and Rosetta Energy Functions for Protein Structure Evaluation. *J Chem*
1128 *Theory Comput* **14**, 6015-6025 (2018). <https://doi.org/10.1021/acs.jctc.8b00303>

1129 105 Alford, R. F. *et al.* The Rosetta All-Atom Energy Function for Macromolecular Modeling and
1130 Design. *J Chem Theory Comput* **13**, 3031-3048 (2017). <https://doi.org/10.1021/acs.jctc.7b00125>

1131 106 Fakihi, Z. & Germain, H. Implication of ribosomal protein in abiotic and biotic stress. *Planta*
1132 **261**, 85 (2025). <https://doi.org/10.1007/s00425-025-04665-6>

1133 107 Kim, S. *et al.* PubChem 2025 update. *Nucleic Acids Res* **53**, D1516-d1525 (2025).
1134 <https://doi.org/10.1093/nar/gkae1059>

1135 108 Huang, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered
1136 proteins. *Nat Methods* **14**, 71-73 (2017). <https://doi.org/10.1038/nmeth.4067>

1137 109 Valdés-Tresanco, M. S., Valdés-Tresanco, M. E., Valiente, P. A. & Moreno, E. gmx_MMPBSA:
1138 A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J Chem Theory*
1139 *Comput* **17**, 6281-6291 (2021). <https://doi.org/10.1021/acs.jctc.1c00645>

1140 110 Du, J., McGraw, A. & Hestekin, J. A. Modeling of *Clostridium tyrobutyricum* for Butyric Acid
1141 Selectivity in Continuous Fermentation. *Energies* **7**, 1-15 (2014).

1142 111 Wu, Z. & Yang, S. T. Extractive fermentation for butyric acid production from glucose by
1143 *Clostridium tyrobutyricum*. *Biotechnol Bioeng* **82**, 93-102 (2003).
1144 <https://doi.org/10.1002/bit.10542>

1145 112 Lee, J., Jang, Y. S., Han, M. J., Kim, J. Y. & Lee, S. Y. Deciphering *Clostridium tyrobutyricum*
1146 Metabolism Based on the Whole-Genome Sequence and Proteome Analyses. *mBio* **7** (2016).
1147 <https://doi.org/10.1128/mBio.00743-16>

1148 113 Zhang, J., Hong, W., Guo, L., Wang, Y. & Wang, Y. Enhancing plasmid transformation
1149 efficiency and enabling CRISPR-Cas9/Cpf1-based genome editing in *Clostridium*
1150 *tyrobutyricum*. *Biotechnol Bioeng* **117**, 2911-2917 (2020). <https://doi.org/10.1002/bit.27435>

1151 114 Brownlie, E. J. E., Chaharlangi, D., Wong, E. O., Kim, D. & Navarre, W. W. Acids produced by
1152 lactobacilli inhibit the growth of commensal Lachnospiraceae and S24-7 bacteria. *Gut Microbes*
1153 **14**, 2046452 (2022). <https://doi.org/10.1080/19490976.2022.2046452>

1154 115 Abdugheni, R. *et al.* Metabolite profiling of human-originated Lachnospiraceae at the strain
1155 level. *Imeta* **1**, e58 (2022). <https://doi.org/10.1002/imt2.58>

- Sorbara, M. T. *et al.* Functional and Genomic Variation between Human-Derived Isolates of Lachnospiraceae Reveals Inter- and Intra-Species Diversity. *Cell Host Microbe* **28**, 134-146.e134 (2020). <https://doi.org/10.1016/j.chom.2020.05.005>
- Cardullo, N. *et al.* Synthesis of Rosmarinic Acid Amides as Antioxidative and Hypoglycemic Agents. *J Nat Prod* **82**, 573-582 (2019). <https://doi.org/10.1021/acs.jnatprod.8b01002>
- Wang, M. *et al.* Inhibitory mechanism of novel allosteric inhibitor, Chinese bayberry (*Myrica rubra* Sieb. et Zucc.) leaves proanthocyanidins against α -glucosidase. *Journal of Functional Foods* **56**, 286-294 (2019). <https://doi.org/https://doi.org/10.1016/j.jff.2019.03.026>
- Avrahami, D. & Shai, Y. A new group of antifungal and antibacterial lipopeptides derived from non-membrane active peptides conjugated to palmitic acid. *J Biol Chem* **279**, 12277-12285 (2004). <https://doi.org/10.1074/jbc.M312260200>
- Micsonai, A. *et al.* BeStSel: webserver for secondary structure and fold prediction for protein CD spectroscopy. *Nucleic Acids Res* **50**, W90-w98 (2022). <https://doi.org/10.1093/nar/gkac345>
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014). <https://doi.org/10.1093/bioinformatics/btu170>
- Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012). <https://doi.org/10.1038/nmeth.1923>
- Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* **12**, R60 (2011). <https://doi.org/10.1186/gb-2011-12-6-r60>
- Herman, N. A. *et al.* The industrial anaerobe *Clostridium acetobutylicum* uses polyketides to regulate cellular differentiation. *Nat Commun* **8**, 1514 (2017). <https://doi.org/10.1038/s41467-017-01809-5>
- Ehsaan, M. *et al.* Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. *Biotechnol Biofuels* **9**, 4 (2016). <https://doi.org/10.1186/s13068-015-0410-0>
- Wu, Y. D., Xue, C., Chen, L. J. & Bai, F. W. Effect of zinc supplementation on acetone-butanol-ethanol fermentation by *Clostridium acetobutylicum*. *J Biotechnol* **165**, 18-21 (2013). <https://doi.org/10.1016/j.jbiotec.2013.02.009>
- Luo, K., Guo, X., Zhang, H., Fu, H. & Wang, J. The Physiological Functions of AbrB on Sporulation, Biofilm Formation and Carbon Source Utilization in *Clostridium tyrobutyricum*. *Bioengineering (Basel)* **9** (2022). <https://doi.org/10.3390/bioengineering9100575>
- Chen, K. Y. *et al.* The Transcriptional Regulator Lrp Contributes to Toxin Expression, Sporulation, and Swimming Motility in *Clostridium difficile*. *Front Cell Infect Microbiol* **9**, 356 (2019). <https://doi.org/10.3389/fcimb.2019.00356>

Acknowledgements

The work was supported by the grants of Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XDXM509), National Natural Science Foundation of China (U22A20424 and 22378048), and Guangdong Basic and Applied Basic Research Foundation (2025B1515020038).

Author contributions

C.X. conceptualized and managed this study. J.C., L.P., Y.L., Y.F., M.W., Y.Y., H.L., and S.W. developed computational screening. S.R., X.Z., Z.Y., J.L., H.W., R.L., and X.M. carried out the experiments. S.R., J.C., L.P., D.W., Y.Y., W.K., R.W., and C.X. analysed the data. S.R., J.C., K.K., Y.Y., W.K., and C.X. drafted the manuscript. C.X. and K.K. supervised the work. S.R., J.C., W.K., X.C., C.Z., G.L., S.L., H.Y., G.F., Y.Y., K.K., and C.X. edited the manuscript.

1203
1204 **Competing interests**
1205 The authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementalTables.xlsx](#)
- [Supplementaryinformation.docx](#)