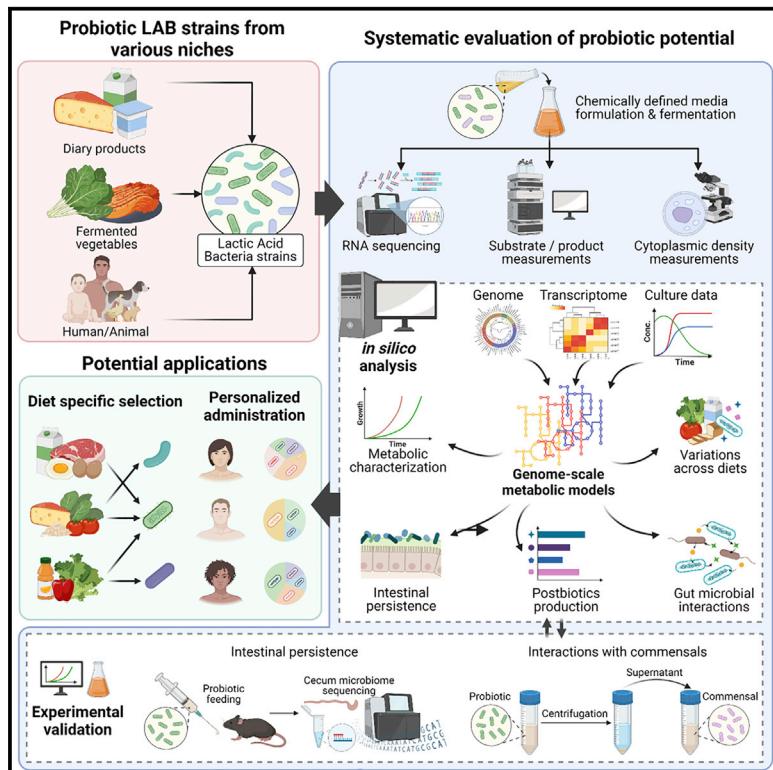


Systematic evaluation of genome-wide metabolic landscapes in lactic acid bacteria reveals diet- and strain-specific probiotic idiosyncrasies

Graphical abstract



Authors

Lokanand Koduru,
Meiyappan Lakshmanan, Yi Qing Lee, ...,
Mazlina Banu, Dave Siak Wei Ow,
Dong-Yup Lee

Correspondence

dave_ow@btি.a-star.edu.sg (D.S.W.O.),
dongyuplee@skku.edu (D.-Y.L.)

In brief

Koduru et al. present a systems biology framework where multi-omics data analysis is combined with mechanistic model-driven approaches to evaluate probiotic capabilities based on growth and postbiotics production under various dietary regimes, intestinal persistence, and potential interactions with gut commensals. We highlight the importance of personalized probiotic formulations.

Highlights

- Presents a rational framework for personalized probiotic selection and design
- Integrates multi-omics data and metabolic modeling to evaluate probiotic capability
- Unravels diet-specific patterns in probiotic colonization and postbiotic potential
- Finds *L. casei* exhibits more desirable probiotic traits among investigated LAB



Article

Systematic evaluation of genome-wide metabolic landscapes in lactic acid bacteria reveals diet- and strain-specific probiotic idiosyncrasies

Lokanand Koduru,^{1,5} Meiyappan Lakshmanan,^{2,5} Yi Qing Lee,³ Pooi-Leng Ho,² Pei-Yu Lim,² Wei Xuan Ler,² Say Kong Ng,² Dongseok Kim,³ Doo-Sang Park,⁴ Mazlina Banu,² Dave Siak Wei Ow,^{2,*} and Dong-Yup Lee^{3,6,*}

¹Institute of Molecular and Cell Biology, Agency for Science, Technology and Research (A*STAR), 61 Biopolis Drive, Proteos, Singapore 138673, Singapore

²Bioprocessing Technology Institute, Agency for Science, Technology and Research (A*STAR), 20 Biopolis Way, #06-01, Centros, Singapore 138668, Singapore

³School of Chemical Engineering, Sungkyunkwan University, 2066, Seobu-ro, Jangan-gu, Suwon 16419, Republic of Korea

⁴Korean Collection for Type Cultures (KCTC), Korea Research Institute of Bioscience and Biotechnology (KRIBB), 181 Ipsin-gil, Jeongeup 56212, Republic of Korea

⁵These authors contributed equally

⁶Lead contact

*Correspondence: dave_ow@bt.a-star.edu.sg (D.S.W.O.), dongyuplee@skku.edu (D.-Y.L.)

<https://doi.org/10.1016/j.celrep.2022.111735>

SUMMARY

Lactic acid bacteria (LAB) are well known to elicit health benefits in humans, but their functional metabolic landscapes remain unexplored. Here, we analyze differences in growth, intestinal persistence, and postbiotic biosynthesis of six representative LAB and their interactions with 15 gut bacteria under 11 dietary regimes by combining multi-omics and *in silico* modeling. We confirmed predictions on short-term persistence of LAB and their interactions with commensals using cecal microbiome abundance and spent-medium experiments. Our analyses indicate that probiotic attributes are both diet and species specific and cannot be solely explained using genomics. For example, although both *Lacticaseibacillus casei* and *Lactiplantibacillus plantarum* encode similarly sized genomes with diverse capabilities, *L. casei* exhibits a more desirable phenotype. In addition, “high-fat/low-carb” diets more likely lead to detrimental outcomes for most LAB. Collectively, our results highlight that probiotics are not “one size fits all” health supplements and lay the foundation for personalized probiotic design.

INTRODUCTION

Lactic acid bacteria (LAB) refer to a group of microaerophilic, Gram-positive bacteria that primarily ferment the hexose sugars into lactate.² LAB include several genera, *Enterococcus*, *Lactobacillus* (this genus has been recently reclassified into 25 others³), *Lactococcus*, *Leuconostoc*, *Oenococcus*, *Pediococcus*, and *Streptococcus*, with functional classification into three major fermentative groups: homolactic, facultative heterolactic and obligate heterolactic, depending on the fermentation products that result from carbohydrate catabolism.⁴ LAB are indigenous species in human food habitat and are ubiquitously found in various nutrient-rich niches such as milk environments, vegetables, and meats.⁵ Moreover, they also constitute a part of the human microbiome in the large and small intestines and colon mucosal layers, exhibiting complex molecular cross-talk with the host and other microbiota to confer several beneficial health effects, e.g., antibacterial,^{6–9} immunomodulatory and cytokine stimulatory,^{10,11} free radical scavenging,^{12,13} and antitumor activities.^{14,15} LAB have thus been a popular choice to maintain

good gut homeostasis and/or to ameliorate microbiota dysbiosis as it has been comprehensively reviewed elsewhere.^{16,17}

Although certain LAB such as *Lactiplantibacillus plantarum* (formerly *Lactobacillus plantarum*), *Ligilactobacillus salivarius* (formerly *Lactobacillus salivarius*), and *Lacticaseibacillus casei* (formerly *Lactobacillus casei*) are generally used as probiotics, their beneficial effects could be highly variable across different species, genera, and strains. While several studies have extensively demonstrated the variations via both *in vitro* and *in vivo* experiments,^{18–25} some have also reported a few commensurate health-promoting effects across various strains or even species.^{26–29} Such conflicting observations possibly stem from a myriad of factors including the differences in evaluation methodology and the distinct mechanisms of action: while some beneficial effects, e.g. bile resistance³⁰ or mucosal adhesion,^{31,32} could be common across even genera,³³ others might be niche or strain specific, e.g. synthesis of bioactive compounds, i.e. postbiotics. Therefore, it is of high importance to comprehensively examine the unique metabolic capabilities of LAB and uncover their nutritional preferences and postbiotic biosynthetic



functions in a context-specific manner. In this regard, the current availability of complete genome sequences for more than 1,000 LAB strains,^{34,35} facilitated only a limited number of comparative genomic analyses hinting at the plausible diversity in nutrient uptake and the ability to produce various anti-microbial substances.^{2,36} Here, to our knowledge, for the first time, we systematically evaluate the genome-wide functional metabolic capabilities of six representative LAB from different genera, species and fermentative groups, through an integrative framework based on comparative genomics, transcriptomics, and *in silico* modeling, together with their growth and biochemical profiles obtained from a newly formulated chemically defined medium (CDM). We subsequently validated the *in silico* predictions using short-term intestinal persistence data obtained from probiotic-fed mice and LAB-gut commensal co-culture experiments, and we observed them to be highly consistent. Such comparative analyses allow us to elucidate the probiotic determinants and probable interactions of LAB with common gut commensal and pathogenic microbes under various dietary regimes.

RESULTS

Comparative genomic analysis of LAB reveals their pan-genomic basis of probiotic capabilities

To understand how the probiotic characteristics vary across LAB, we selected six representative strains, *L. plantarum* WCFS1 (LbPt), *L. casei* subsp. *casei* ATCC 393 (LbCs), *L. salivarius* ATCC 11741 (LbSv), *Limosilactobacillus fermentum* (formerly *Lactobacillus fermentum*) ATCC 14931 (LbFm), *Lactococcus lactis* subsp. *cremoris* NZ9000 (LcLt), and *Leuconostoc mesenteroides* subsp. *mesenteroides* ATCC 8293 (LeMt), from different genera in order to cover a wide range of LAB across different species as well as diverse fermentative groups. Note that although many members of former *Lactobacillus* genera are well established as probiotics, here, we purposely chose representatives from other well-studied genera of LAB, *Leuconostoc* and *Lactococcus*, to provide a good balance in terms of species diversity (Figure 1A). We first constructed a phylogenetic tree using the single-copy core gene families of all six species considered here with four other LAB and an outgroup member, *Bacillus subtilis* (BcSt) (see STAR Methods). Not surprisingly, while three of the *Lactobacillus* strains are clustered together, LbCs is placed within another sub-clade (Figure 1B). This confirms the previous observation that *Lactobacillus* genus is paraphyletic and underlines the remarkable genetic diversity of LAB species,² enabling their adaptation to various nutrient-rich niches. Due to such diversity, recently all the species from *Lactobacillus* genus have been reclassified into 25 genera, which include 23 novel ones.³

The sizes of the six LAB genomes considered here vary widely (2.55 ± 0.84 Mb), where LbPt and LbFm have the largest (3.84 Mb) and smallest (1.86 Mb) genomes, respectively (Table S1), signifying the potential variations in probiotic traits. To comprehensively characterize how the LAB genomic catalogs differ from each other, we next quantified the genetic redundancy based on orthologous genes (see STAR Methods). Among the total 13,412 genes from all LAB, 11,068 genes (82.5%) have orthologs in other species and can be grouped into 2,386 orthologous

gene family groups, or simply known as “orthogroups” (Table S1). A total of 691 “core” gene families, including 549 single-copy genes, are present in all LAB. 1,455 “shell” gene families that are shared in a few LAB and 2,585 “cloud” gene families were also identified, emphasizing the extraordinary genetic diversity despite their small genome sizes (Figure 1C). We also noted that as many as 78% of the different orthogroups identified were present in LbPt, suggesting this strain to be the most versatile (Table S1). A significantly high number of duplicate genes were also identified in LbPt (34%), indicating that its genome is highly redundant compared with all other LAB (18%–25%). Functional annotation of the core, cloud, and shell genome (see STAR Methods) highlighted that housekeeping processes such as translation, ribosomal structure and biogenesis, replication (19% compared with 2% and 1%) and post-translational modification, protein turnover, and chaperones (4% compared with 2% and 1%) are highly enriched in the core genome compared with the cloud and shell genome (Figure 1D). In addition, a few metabolic pathways, such as energy, lipid, and nucleotide metabolism, were highly conserved across LAB species. On the other hand, functionalities more enriched in cloud and shell genome include carbohydrate, amino acid, inorganic ion, coenzyme, and other secondary metabolism and some non-metabolic categories such as “transcription” and “genes with unknown function.” Significant enrichment of metabolic genes in cloud and shell genome, particularly carbohydrate and amino acid metabolism, pinpoints the differences in their preferential adaptation to nutrient environments such as milk, plant, and animals^{37–39} as well as the variations in their probiotic capabilities.⁴⁰ Moreover, the high proportion of gene families in “genes with unknown function” category (806/900) in LAB cloud and shell genome hints at the gaps in the current understanding and the need for a comprehensive functional annotation of such genes.

Apart from the metabolic flexibility, LAB also synthesize a variety of proteins for resisting acid, oxidative, and bile stresses, conferring anti-microbial activities, protein lysing proteases, and to anchor cells on host mucosal layer surfaces. Since all these proteins are critical for a probiotic organism to colonize and sustain the harsh environments of human gut, we mined the genomes of all six LAB species to identify relevant proteins (see STAR Methods). Notably, all LAB encode for multiple cell adhesion factors, which could help them attach in the mucosal layers. In the stress resistance category, while all LAB encode at least one protein, some of them have more than one (Figure 1E). For example, LbPt genome contains five different bile resistance genes and four unique acid resistance genes. Many LAB also produce and secrete anti-microbial peptides, known as bacteriocins, which can be broadly classified into three classes: (1) class I bacteriocins, small heat stable posttranslationally modified peptides; (2) class II bacteriocins, small heat stable posttranslationally unmodified peptides; and (3) class III bacteriocins, slightly large thermo-labile stable peptides.⁴¹ We identified eight genes encoding for potential class II bacteriocins in LbPt and one each in LeMt and LbCs (Figure 1E). We also identified a potential gene encoding for the class III bacteriocin in LbSv. The ability to release cell envelope proteases (CEPs) that cleave the casein peptides in milk environment into shorter

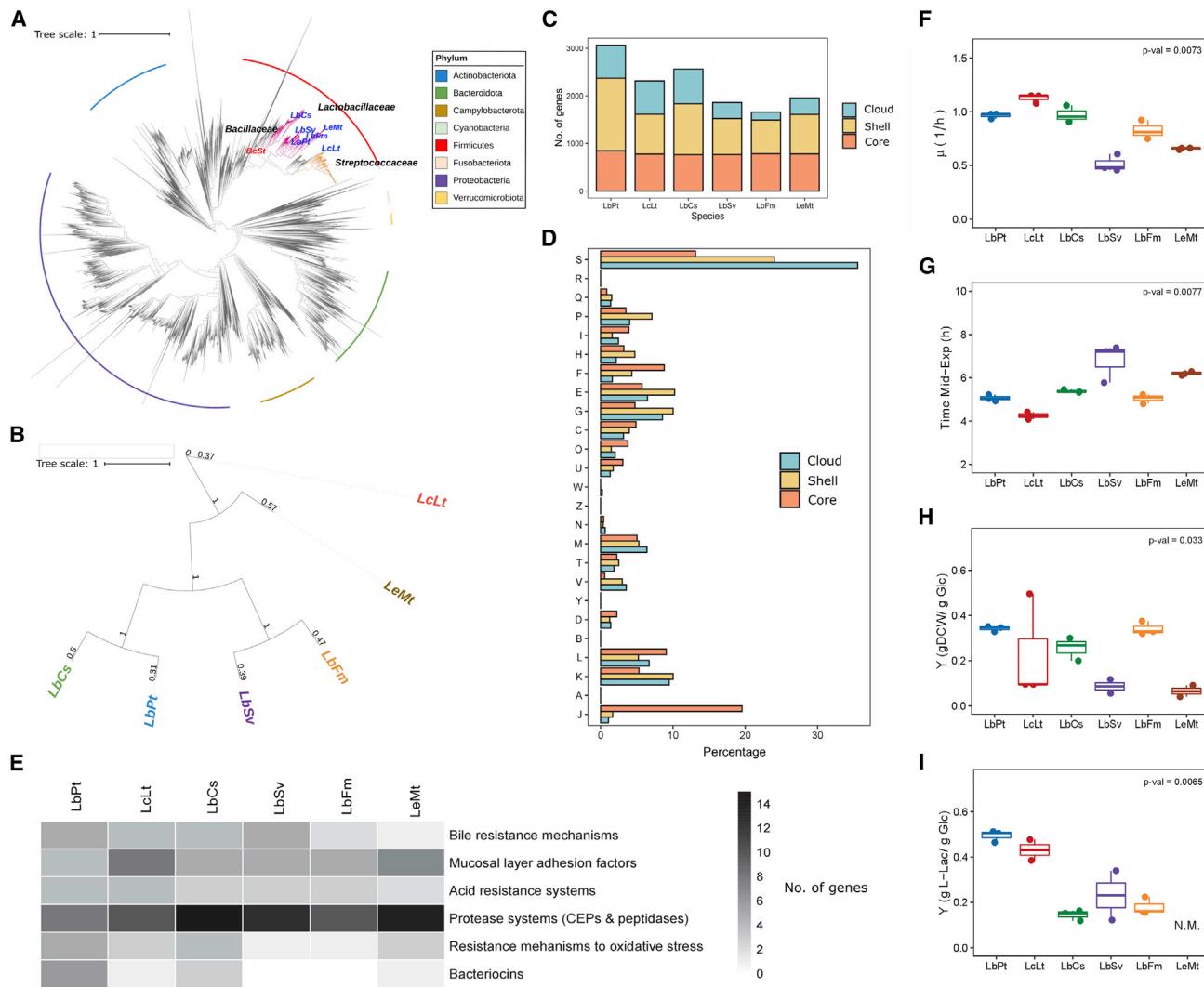


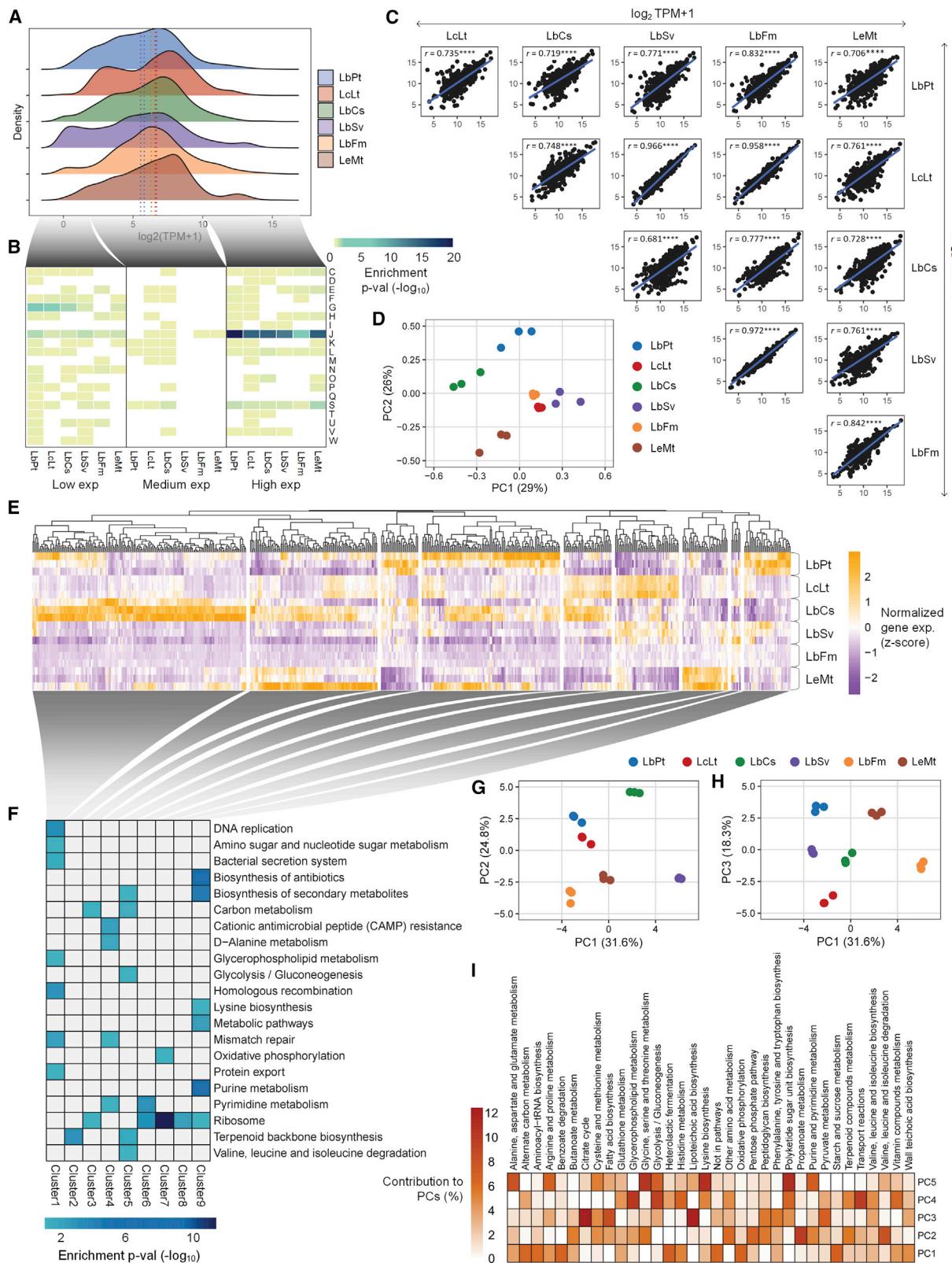
Figure 1. Genomic characteristics, growth characteristics, and fermentative phenotypes of LAB strains

(A) Genome Taxonomy Database (GTDB) bacteria phylogenetic tree showing the location of selected six LAB species, (B) phylogenetic tree showing the relationship between several LAB species with an outgroup member *B. subtilis* (BcSt), (C) the number of core, shell, and cloud genes in LAB strains, (D) the functional annotation of core, shell, and cloud genes using NOG mappings, and (E) comparison of major non-metabolic genetic determinants influencing probiotic efficacy of LAB. (F) Maximum growth rates, (G) doubling times, (H) yield of biomass over glucose, and (I) yield of lactate over glucose of LAB grown in LABDM cultures. Core genes in (B) are the genes that are present in all six LAB. Shell genes include the ones that have ortholog only in a few, and cloud genes are the ones present only in one organism. NOG categories in (C) is as follows: A, RNA processing and modification; B, chromatin structure and dynamics; C, energy production and conversion; D, cell cycle control, cell division, and chromosome partitioning; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H, coenzyme transport and metabolism; I, lipid transport and metabolism; J, translation, ribosomal structure, and biogenesis; K, transcription; L, replication, recombination, and repair; M, cell wall/membrane/envelope biogenesis; N, cell motility; O, post-translational modification, protein turnover, and chaperones; P, inorganic ion transport and metabolism; Q, secondary metabolites biosynthesis, transport, and catabolism; R, general function prediction only; S, genes with unknown function; T, signal transduction mechanisms; U, intracellular trafficking, secretion, and vesicular transport; V, defense mechanisms; W, extracellular structures; Y, nuclear structure; and Z, cytoskeleton. Three samples were used per LAB in (F)–(I). The p values indicated in the figure were estimated using Kruskal-Wallis test. L-lactate concentration in LeMt cultures was not measured because LeMt does not produce L-lactate. In (F) to (I), the middle line in each box represents the mean value, while error bars denote the minimum and maximum values.

peptides/amino acids for utilizing them as nutrient sources is also a primary attribute of most LAB. Additionally, some LAB CEPs are shown to alter inflammatory bowel disease by acting upon the inflammatory mediators. So, we also mined all the LAB genomes and identified that all of them could potentially produce at least one such CEP (Figure 1E).

Development of a novel chemically defined medium for the unbiased evaluation of LAB growth characteristics and other phenotypes

We newly formulated a CDM that supports the growth of all LAB tested. Design of such a novel CDM is highly necessary to enable an unbiased and comprehensive metabolic characterization of



(legend on next page)

all six LAB strains, which are generally auxotrophic to several nutrients, particularly amino acids and vitamins. Here, it should be emphasized that although rich nutrient media such as M17 or *Lactobacillus* broth might support their growth, the chemically undefined components in such media may introduce batch-to-batch variations and uncertainties to substrate/product measurements. We initially cultivated all six LAB in a CDM that was originally developed to grow *Lactococcus* and *Streptococcus* species.⁴² However, this medium was only able to support the growth of LcLt and LbSv. Therefore, we modified it iteratively by including several key metabolites based on the known growth requirements of each LAB, hereafter referred as LAB-defined medium (LABDM), for their unrestricted growth. Importantly, we eliminated the need for preparing multiple micronutrient component stocks of vitamins and minor salts in LABDM by replacing them with yeast nitrogen base (YNB), which is a defined commercial formulation of micronutrients widely used for yeast cultivation. The optimal supplementation of YNB ensured proper growth with minimum possible lag phase for most LAB strains (see *STAR Methods* and *Table S2*).

We next cultivated the LAB strains in LABDM and evaluated their batch growth characteristics under anaerobic conditions. While most strains grew at similar rates, LbSv and LeMt grew much slower than other LAB (*Figures 1F–1I* and *S1A*; fold change <0.75; p value = 0.007, Kruskal-Wallis test). We also monitored the concentrations of major nutrients, such as glucose and amino acids, and primary by-products, L-lactate and acetate, in the LABDM over the exponential phase (*Figures S1B* and *S1C*). No significant difference was observed in the biomass yield from glucose across LAB, although LcLt, LbSv, and LeMt showed slightly lower yields (*Figure 1H*; p value = 0.033, Kruskal-Wallis test). These findings are highly consistent with our previous study where we reported the limited ability of LeMt to catabolize amino acids and other nutrients for energy production.⁴³ In terms of L-lactate yield from glucose, LcLt and LbPt produced almost a 2-fold higher amount of lactate than other LAB (*Figure 1I*; p value = 0.0065, Kruskal-Wallis test).

Genome-wide transcriptome sequencing of LAB reveals unique gene expression profiles irrespective of their taxonomic groups

To delve deeper insights into the genome-wide transcriptional regulation among various taxonomic groups, we collected RNA at the mid-exponential phase from all strains grown in LABDM and performed RNA-seq. At the outset, we compared the global gene expression profiles by analyzing the distribution of genes that fall under various expression ranges. LbSv and LbFm have

a large number of genes in the low expression level, while LbPt, LbCs, and LeMt have a gene expression distribution slightly skewed toward the high expression range (*Figure 2A*, *Table S3*, and *Figure S2*). Subsequent enrichment analysis of non-supervised orthologous groups (NOG) terms in these gene expression categories (see *STAR Methods*) identified housekeeping functions, i.e., translation, ribosomal structure, and biogenesis (false discovery rate [FDR] adjusted p value = 8.830×10^{-30} – 1.805×10^{-17} , Fisher's exact test), in high expression range (*Figure 2B*). We also noted that the carbohydrate transport and metabolism is significantly enriched (FDR adjusted p value = 1.341×10^{-10} – 3.128×10^{-7} , Fisher's exact test) in low expression range particularly for LbPt, LbCs, and LcLt, indicating that the metabolism of LAB with larger genomes could be under significant transcriptional control. Noticeably, we found a widespread expression of genes with unknown function, highlighting the need to further characterize their role in governing LAB phenotype.

In addition to the global transcriptome analysis, we compared the interspecies transcriptomic landscapes and their predisposition to various metabolic and probiotic capabilities by first identifying the 549 one-one orthologs (including 195 metabolic genes) and examining their normalized expression values using a set of housekeeping genes to correct for the species-specific bias (see *STAR Methods* and *Figure S3*). The expression profiles of these genes were analyzed using two individual and complementary metrics: (1) principal component analysis (PCA) and (2) correlation in a pairwise manner. Surprisingly, the gene expression patterns are markedly different between phylogenetically closer species such as LbSv and LbCs (Spearman correlation, $r = 0.673$), while they are highly similar in distant species (*Figures 2C* and *2D*), e.g., LbFm and LcLt (Spearman correlation, $r = 0.96$). Hierarchical clustering of differentially expressed orthologous genes resulted in nine major groups where each cluster is mainly represented by genes that are upregulated in a particular LAB (*Figure 2E*). Enrichment of functional categories in these genes further revealed that the carbon metabolism and several growth-related biological processes such as DNA replication (FDR adjusted p value = 3.162×10^{-17} , modified Fisher's exact test) and ribosome and DNA mismatch repair (FDR adjusted p value = 0.0031, modified Fisher's exact test) were upregulated in LAB with larger genomes such as LbPt, LbCs, and LcLt, thus providing further evidences to why these species grow much faster than others (*Figures 2F* and *S4*). LeMt shows high gene expression of anaerobic respiration (FDR adjusted p value = 0.0034, Fisher's exact test), i.e., F1ATPase and menaquinone biosynthesis, indicating that redox-controlled metabolism governs the fastidious growth as reported earlier.⁴³

Figure 2. Transcriptomic characteristics of exponentially growing LAB in LABDM

(A) Distribution of gene expression values (n = 3), (B) enrichment of NOG categories in three distinct gene expression ranges, i.e. low, medium, and high, (C) plots showing pairwise correlation of gene expression of orthologous genes across various LAB, (D) principal component analysis (PCA) of the orthologous gene expression, (E) hierarchical clustering of differentially expressed orthologous genes, (F) enrichment of KEGG pathways in differentially expressed genes and (G–I) PCA of metabolic genes in various LAB strains. Numbers in (C) represent the Pearson correlation coefficient, r , between values plotted on x and y axes, and * indicates the level of significance measured using p value, one-tailed, 95% confidence interval ($p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$). The NOG term abbreviations in (B) are same as *Figure 1*. Negative log₁₀ transformations of enrichment p value are provided in (B) and (F). Clustering was done with Euclidean distances of normalized gene expressions (Z scores) across LAB in (E). PCA of metabolic genes was performed based on gene assignments to various pathways in newly reconstructed GEMs.

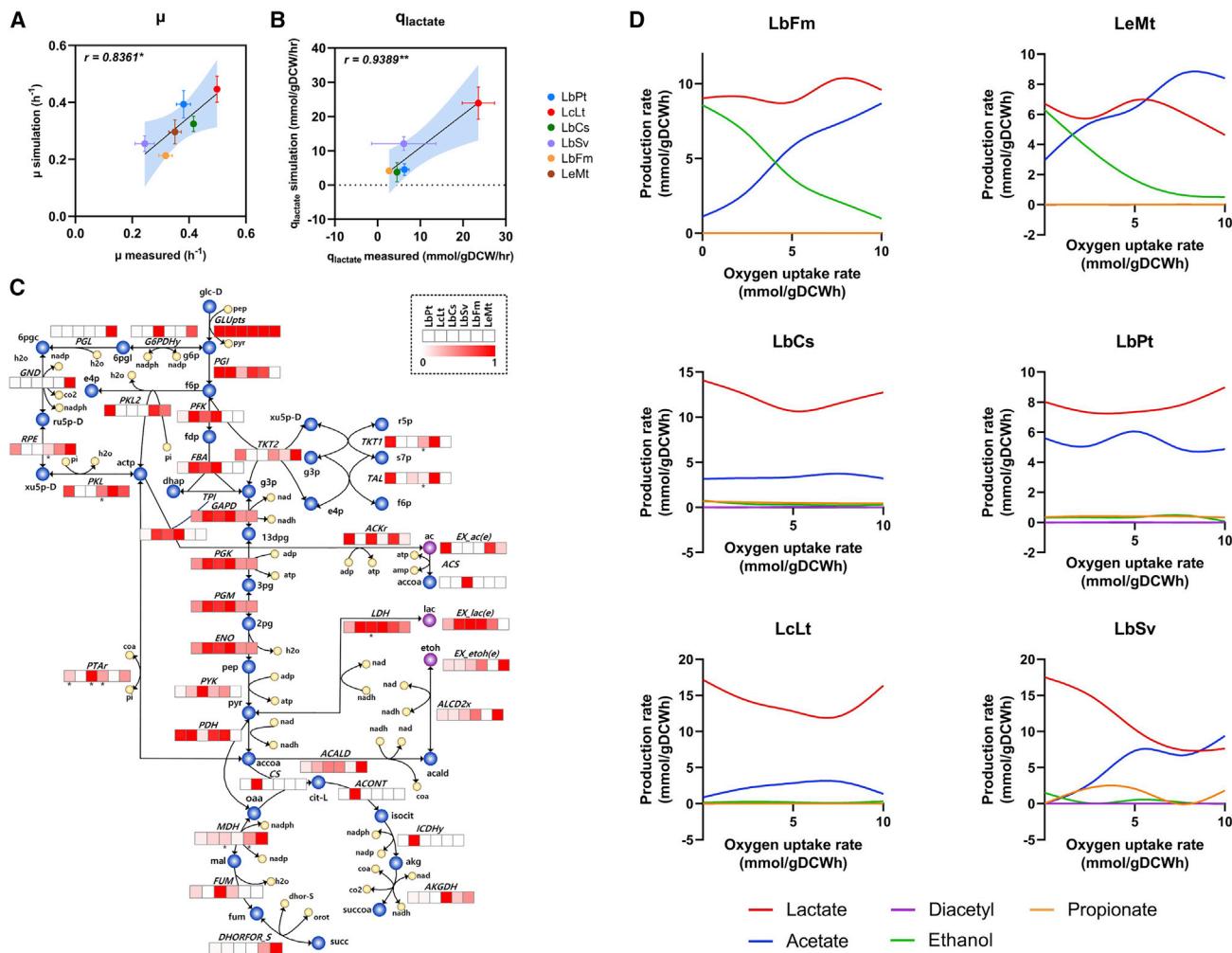


Figure 3. Phenotype predictions of LAB GEMs

(A) Comparison of GEM predicted and experimentally measured growth rates, (B) comparison of GEM predicted and experimentally measured lactate production rates, (C) flux map showing the intracellular flux distribution across glycolysis and TCA cycle reactions, and (D) qualitative profiles simulated by LAB GEMs as a function of varying oxygen uptake rates. Three samples were used per LAB for experimental measurements in (A and B). The horizontal and vertical error bars in (A) and (B) denote the standard deviation of triplicate measurements and the variations in predicted rates over multiple α_i parameters in 5,000 flux solutions, respectively. Numbers in (A) and (B) represent the Pearson correlation coefficient, r , between values plotted on x and y axes, and * indicates the level of significance measured using p value, one-tailed, 95% confidence interval ($p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$). ** in (C) indicates flux flow in the opposite direction.

Since the transcriptional regulation of each LAB is significantly different even among the phylogenetically closer species, we further performed a PCA focusing on metabolic gene expression. The first and second PCs account for ~32% and 25% variations, respectively, with major contributions from the functional categories relevant to growth (alternate carbon metabolism, aminoacyl-tRNA biosynthesis, heterolactic fermentation, oxidative phosphorylation, and starch and sucrose metabolism) and biosynthesis of bioactive compounds (protonate metabolism, butanoate metabolism, lipoteichoic acid biosynthesis, and vitamin metabolism) (Figures 2G–2I). These results clearly suggest that gene expression variations among LAB can be better captured by their ability to adapt to various environments and produce relevant bioactive compounds rather than their taxonomic differences.

LAB genome-scale metabolic models highlight the use of low-energy-yielding pathways as primary metabolic fates

As the comparative genomic analyses revealed an unprecedented divergence in metabolic and physiological capabilities of LAB, we first updated the existing genome-scale metabolic models (GEMs) for LbPt,⁴⁴ LcLt,⁴⁵ and LeMt⁴³ based on updated genome annotations and improved their scope and coverage significantly (Figure S5). Additionally, we reconstructed new GEMs for LbCs, LbFm, and LbSv, following the standard procedure⁴⁶ (see STAR Methods), to further probe the functional metabolic states of six LAB. The resulting GEMs contained 686 ± 139 genes, $1,042 \pm 65$ reactions, and 906 ± 29 unique metabolites on average, where LbFm and

LbPt represented the smallest and largest models, respectively ([Table S5](#)).

Following the reconstruction of GEMs, we then validated the model predictions using the batch culture data of LAB grown in LABDM (see [STAR Methods](#) and [Table S4](#)). The initial *in silico* growth rates predicted by flux balance analysis (FBA) were comparable to experimentally measured ones. However, most LAB models failed to simulate the lactate secretion, releasing other fermentative products instead (e.g., acetate, ethanol, formate, acetoin, and diacetyl). This motivated us to incorporate additional kinetic constraints for accurately predicting the LAB phenotype. In this regard, implementation of FBA with additional macromolecular crowding constraints (FBAwMC) has been reported to successfully describe the low-yielding acetate overflow in *Escherichia coli*,⁴⁷ the Warburg effect in cancer cells,^{48,49} and the role of redox cofactors in the switch between low- and high-yield metabolism in *Saccharomyces cerevisiae* and LcLt.⁵⁰ Therefore, we subsequently formulated an approach based on FBAwMC employing experimentally measured crowding coefficients to investigate the functional metabolic states of each LAB (see [STAR Methods](#), [Table S6](#); [Figures S6](#) and [S7](#)). Remarkably, the growth rates and lactate secretion rates predicted were highly consistent with the culture experiments (Pearson correlation, $r > 0.85$; [Figures 3A](#) and [3B](#)), thus confirming that the predominant use of low-energy-yielding pathways is a hallmark of LAB.^{44,51}

We then used the validated LAB models to characterize key differences in central metabolism by examining intracellular flux distributions obtained from LABDM conditions ([Figure 3C](#)). Expectedly, Embden-Meyerhof pathway (EMP) is the only glycolytic route found to be active in LcLt, which belongs to the homofermentative group where the primary carbon in the form of glucose-6-phosphate is broken down entirely into 2 mol of lactate. Obligate heterofermentative LAB (LeMt and LbFm) showed active use of phosphoketolase pathway (PKP) instead of EMP for glucose metabolism. This pathway metabolizes 1 mol of glucose-6-phosphate into equal amounts of acetyl-phosphate and glyceraldehyde-3-phosphate, resulting in other fermentative by-products such as acetate and ethanol along with lactate. The facultative heterofermentative LAB, LbPt, carried flux through both EMP and PKP and thus produced mixed acid products; the proportion of lactate to other products depends on the ratio of flux through EMP to that of PKP. The tricarboxylic acid (TCA) cycle flux is largely non-existent in all LAB due to the absence of multiple enzymes essential to the pathway activity ([Figure 3C](#)), conforming to the well-known fact that LAB do not aerobically respire.⁵² However, we observed exceptions to the pathway activity in LbCs and LbSv, which according to their taxonomic assignments should belong to facultative heterofermentative and homofermentative groups, respectively. In order to explore these discrepancies further, we evaluated the ability of LAB GEMs to simulate the known fermentative profiles as a function of oxygen availability in the environment depending on their functional classification into three major groups, i.e., homo-lactic, facultative heterolactic, and obligate heterolactic. LeMt and LbFm models were able to successfully recapitulate the well-known fermentative profiles of obligate heterofermentative group⁴³ that switched from high lactate and ethanol in anoxic

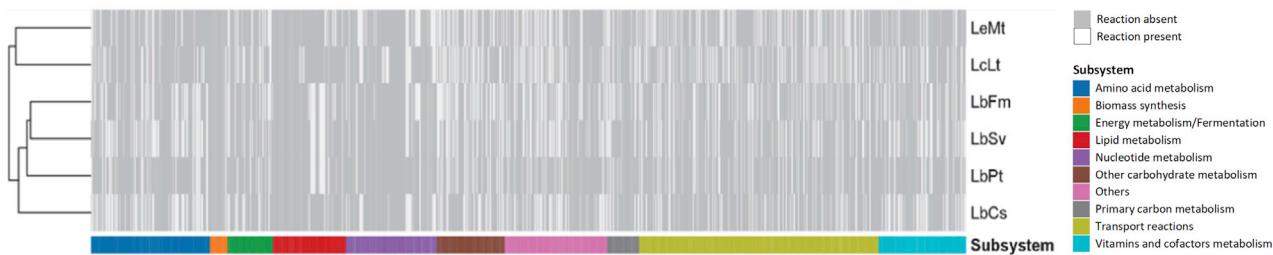
condition to high lactate and acetate during high oxygenation. LcLt model, however, resulted in a fermentative profile that produced lactate exclusively, consistent with its sole dependence on the EMP, i.e., obligate homofermentative metabolism. Simulations from LbPt GEM showed a mixed lactate-acetate profile independent of the oxygen availability, which is mainly attributed to its flexibility to use both EMP and PKP ([Figure 3D](#)). Interestingly, we observed LbCs to display a homofermentative metabolic profile as observed in LcLt although *L. casei* belongs to the facultative heterofermentative group. A further literature survey showed that the particular strain used in this study, i.e., ATCC 393, exhibits a homofermentative phenotype although it encodes PKP,⁵³ which is consistent with our predictions. Similarly, even though LbSv is known to belong to homofermentative group of LAB, our simulations indicated significant amounts of lactate and acetate due to excess pentose phosphate pathway (PPP) fluxes mediated by transketolase reaction. This is in contrary to the well-known fact that LAB do not carry high PPP flux. Therefore, we compared the gene expression levels of transketolase among different LAB using the RNA-seq data and noted that LbSv and LbFm showed negligible levels of transketolase expression. Subsequently, we inactivated the transketolase reactions, i.e., TKT1 and TKT2, in LbSv and LbFm GEMs during simulations. While LbFm profile was unaffected by this deactivation, LbSv exhibited a homofermentative metabolism in anoxic condition and a mixed fermentative profile with an increase in oxygen supply, which is consistent with earlier reports.^{54,55}

Comparison of LAB GEMs delineates their diverse metabolic repertoire and differential nutrient auxotrophy requirements

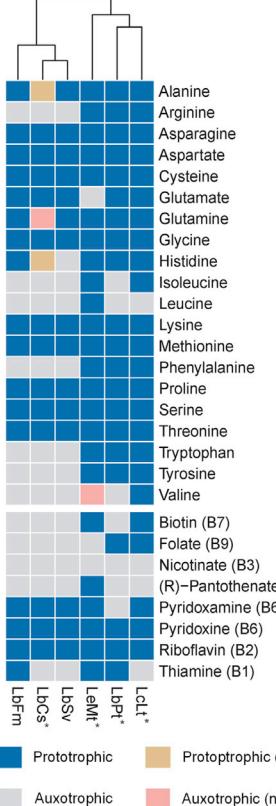
We compared the LAB GEMs using the metabolic distances calculated between pairs of any two GEMs based on their genome-wide reactome and observed that all *Lactobacillus* strains shared more reactions than taxonomically distant LeMt and LcLt ([Figure 4A](#)). We further identified the “core” (reactions present in all six GEMs) and “cloud” (reactions present either in a single GEM or in a few but not in all) metabolic repertoire, as such finding 49% of the total reactions conserved. The highly conserved metabolic subsystems in the reactome are aminoacyl-tRNA biosynthesis (100%), lysine biosynthesis (84%), peptidoglycan biosynthesis (78%), glycerophospholipid metabolism (72%), and fatty acid biosynthesis (68%). The cloud reactome, on the other hand, was largely represented by carbohydrate and amino acid metabolism, confirming the high diversity of LAB in their ability to catabolize various nutrients as highlighted by comparative genomic analysis.

Next, we investigated the availability of pertinent biosynthetic enzymes to produce various amino acids and vitamins in LAB to reason out why some of them are auxotrophic to certain nutrients ([Figure 4B](#)). Among all LAB, LeMt is not auxotrophic to all amino acids, except glutamate, as it contains the biosynthetic enzymes of almost all amino acids. It is well known that many LAB are auxotrophic to all three branched chain amino acids, i.e., isoleucine, leucine, and valine.^{56–58} Correspondingly, we noted the key gene (*ilvC*) encoding the enzyme ketol-acid reductoisomerase for the conversion of 2-(S)-acetolactate into

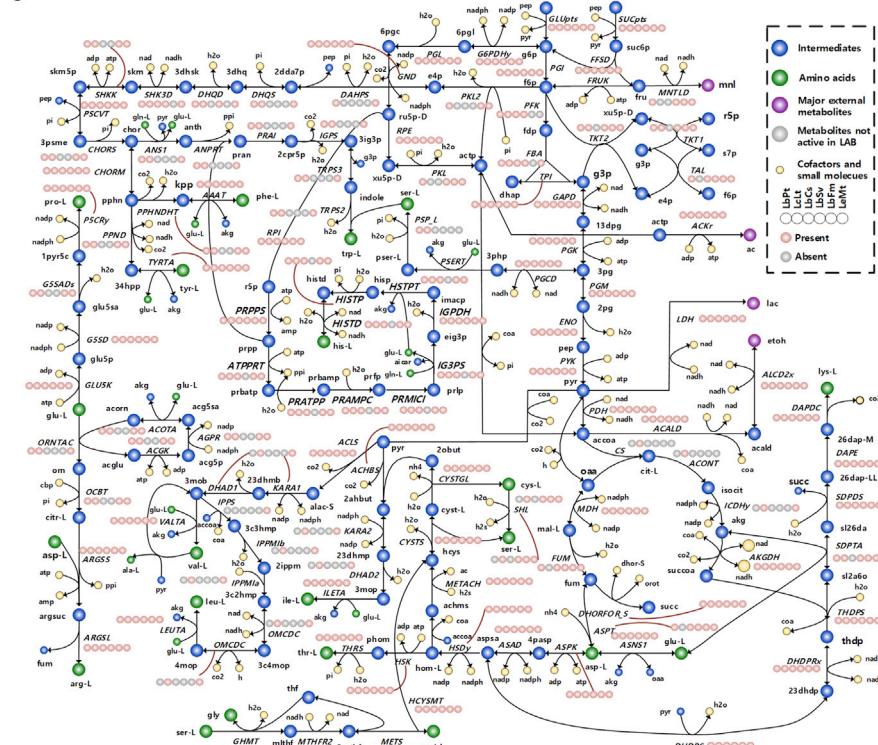
A



B



C



D

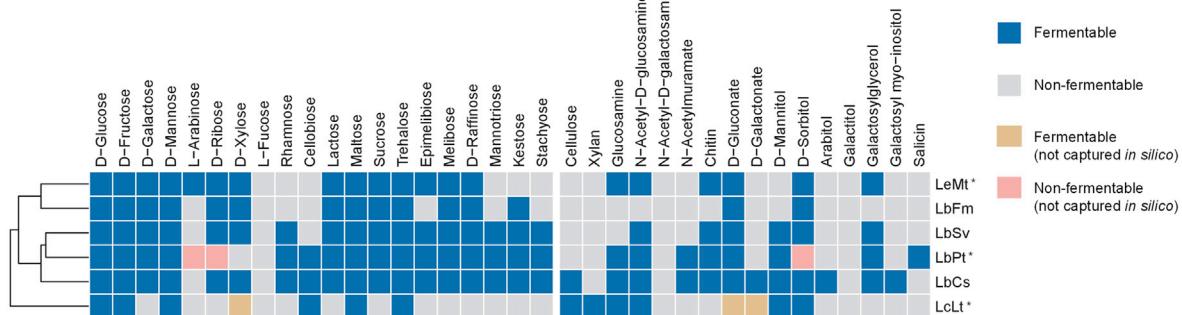
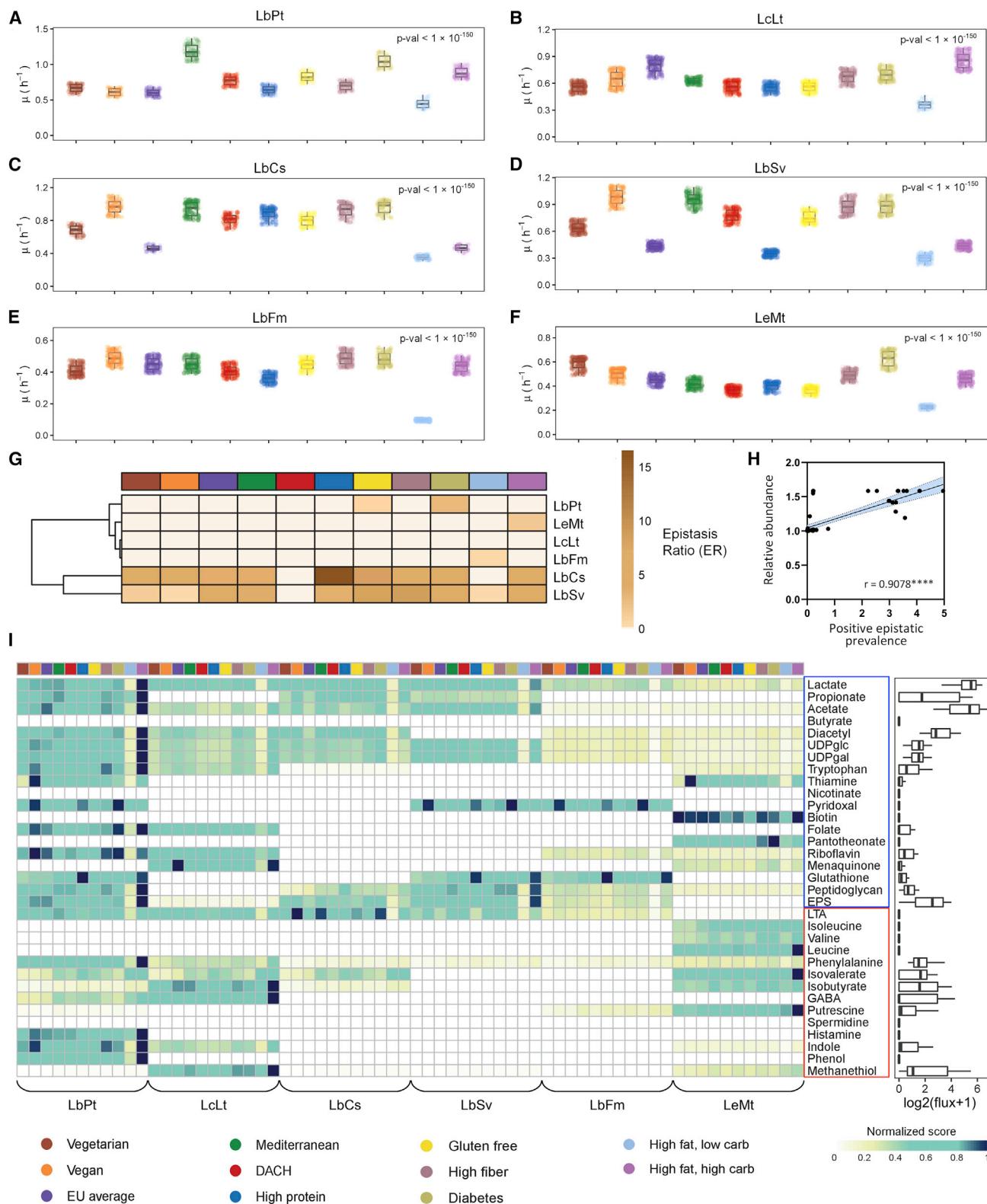


Figure 4. GEM characterized core and cloud metabolic capabilities of LAB

(A) Heatmap showing presence/absence of reactions in different LAB species and subsystems across LAB, (B) amino acid auxotrophy as revealed by LAB GEMs and their comparison with literature data, (C) metabolic network showing the presence and absence of different reactions in central carbon metabolism and the amino acid biosynthesis pathways. “**” in (B) and (D) indicates that these species had corresponding literature data on amino acid auxotrophy and carbon substrate phenotyping for comparisons with model predictions. Amino acid auxotrophic simulations were performed using flux balance analysis, whereby each amino acid uptake rate was set to zero, while maximizing the growth cellular objective. The LAB were classified as auxotrophic to the amino acid if there was >90% growth reduction due to exclusion of amino acid from the media. In fermentable carbon substrate phenotyping, the ability of LAB to consume the corresponding sugar and to produce lactate or other use of fermentative by-products was tested.



(legend on next page)

(R)-dihydroxyisovalerate was missing in all LAB, except LeMt and LcLt (Figure 4C). Even though all biosynthetic genes are present, LcLt is still auxotrophic to valine and isoleucine, indicating the potential inexpression or inactivity of some of these genes. Some of the LAB were also additionally auxotrophic to aromatic amino acids (AAAs), phenylalanine, tryptophan, and tyrosine, as the key gene in chorismate (AAA precursor) synthesis, *aroH*, encoding the enzyme 3-deoxy-7-phosphoheptulonate synthase is missing. In terms of vitamins, all six LAB were unable to synthesize nicotinate, a key precursor in NAD(P) biosynthesis, and need to be consumed from the media since it is essential for cell growth. Folate is another key vitamin required for *de novo* nucleotides biosynthesis; several LAB were unable to synthesize this vitamin due to the absence of *folP* gene that encodes the dihydropteroate synthase (Figure S8). Interestingly, all LAB, except LeMt, also lacked the *panC* and *bioD* genes, which encode the enzymes responsible for the final step in pantothenate (CoA precursor) and biotin (acetyl-CoA carboxylase cofactor) biosynthesis, respectively, so they are unable to synthesize these vitamins that are critical in fatty acid metabolism. Furthermore, we used the gene expression data to understand why some LAB require lesser concentrations of vitamins to grow than others. To do so, we used a metric that combines the vitamin auxotrophy data and the average Z score of the corresponding gene expression of vitamin biosynthetic pathway, and we found that LbCs, LbFm, and LbSv have overall low auxotrophy-weighted Z scores compared with LbPt, LcLt, and LeMt for vitamin biosynthesis (Table S7). This is consistent with the observations made during CDM optimization iterations (Table S2), where LbCs, LbFm, and LbSv generally required higher vitamin concentrations for optimal growth compared with LbPt, LcLt, and LeMt. We also investigated the ability of LAB to ferment various carbohydrates, showing that all *Lactobacillus* strains can metabolize a diverse range of sugars including hexoses, pentoses, disaccharides, and trisaccharides, while LeMt and LcLt have limited capability (Figure 4D).

In silico analysis of LAB GEMs highlights the diet-specific patterns in probiotic persistence and postbiotic synthesis

Previously, several cohort-based studies have reported that the gut microbiome composition is more sensitive to the dietary habits than the ethnic or geographical makeup.^{59–61} Similarly, we hypothesize that some probiotics may potentially survive and produce beneficial compounds much better in certain diets, which could be tested *in silico*. We explored the probiotic capability of LAB under 11 different diet conditions including vegetarian, vegan, EU average, Mediterranean, DACH, high-protein,

gluten-free, high-fiber, diabetes, high-fat/low-carb, and high-fat/high-carb diets by resorting to model-driven analysis (see STAR Methods). The resulting growth patterns signified very distinct diet-specific phenotypes: most LAB, especially obligate homofermentative, preferred diabetes and high-fiber diets, whereas all of them exhibited very low growth rates in high-fat diets (Figures 5A–5F and S9; p value $< 1 \times 10^{-150}$, Kruskal-Wallis test). The high-fiber and diabetes diets are usually rich in inulin-type fructooligosaccharides that are catabolized via high ATP yielding pathway involving the (phospho-)fructo-furanosidase (*BfrA* or *SacA/Scrb*) enzyme, thus giving rise to the increased growth rates.⁶² Interestingly, we noted that most LAB failed to grow well in high-fat/low-carb diets. We further assessed how different diets influence LAB growth by computing the reduced costs⁶³ of each exchange reaction in all diets across different LAB (see STAR Methods). This analysis indicated that the reduced costs and uptake fluxes of nutrients present in diets such as high-fiber and diabetes diets were more favorable for the growth of most LAB, while they were most unfavorable in high-fat/low-carb diet (Table S8). It should be noted that the high-fat/low-carb diet contained higher fatty acid content, which LbCs, LcLt, and LeMt were able to consume in very low quantities through the fatty acid transporters and incorporated into their cell membranes. In contrast, LbSv, LbPt, and LbFm did not consume fatty acids due to the lack of such transporters. Moreover, all LAB lack fatty acid catabolizing pathways, such as the β-oxidation, so they are unable derive energy from fatty acids.⁶⁴ Taken together, the lower proportion of the readily metabolizable energy source, i.e., carbohydrates, in a high-fat/low-carb diet (Table S4) contributes to the poor growth of LAB. We also observed minimal variations in growth rates of heterofermentative LAB (LbFm and LeMt) across different diets, mainly due to their usage of the rigid and low-yielding phosphoketolase pathway that generates similar levels of ATP from diverse carbohydrates and their limited capacity to use amino acids as alternative energy sources.⁴³ In summary, our *in silico* simulation results indicate that while high-fiber diets promote the probiotic growth and functionality, high-fat diets lead to poor performance, which is in high concordance with previously published clinical data.^{65–67}

Although the essentiality of probiotic persistence in the gut to impart beneficial effects has been still debated,⁶⁸ the increased residence and fitness can provide sufficient time for desirable immunomodulation and pathogen exclusion. It has been earlier shown that pigs and mice colonized with probiotic LAB displayed better immunomodulatory effects⁶⁹ and cytokine stimulation,⁷⁰ respectively. Among multiple factors, the genetic determinants of probiotic strains are reported to be one of the key

Figure 5. Diet-dependent probiotic characteristics of LAB simulated *in silico*

(A–F) Variations in simulated growth rates across various diets in each LAB species (p value $< 1 \times 10^{-150}$, Kruskal-Wallis test), (G) ratio of positive to negative epistasis, (H) correlation between epistasis ratio and experimentally observed fold change of LAB between the probiotic-fed and control mouse cecum, and (I) postbiotic production capabilities of LAB in various diets. Growth rates in (A)–(F) are simulated by constraining each LAB GEM with respective diet conditions over 5,000 samples. Number in (H) represents the Spearman's Rank correlation coefficient, r , between values plotted on x and y axes, and * indicates the level of significance measured using p value, one-tailed, 95% confidence interval ($p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$). In (A) to (F), the middle line in each box represents the mean value, while error bars denote the minimum and maximum values. In (I), the values in the heatmap are production potential of each compound normalized to the maximum value across all diets and species. The boxplot in (I) shows the overall variation in production levels across all diets and species. Compounds colored in blue and red denote the beneficial and detrimental postbiotics, respectively.

contributing factors for their stress resistance, thus influencing their cellular fitness in the gut.^{71–74} In fact, several probiotics capable of colonizing and persisting in the gut environment are intrinsically stress resistant, e.g., antibiotic resistant.^{75–79} Moreover, it has been extensively reported that the ability to resist environmental stress is associated with the presence of dominant-positive epistatic landscapes in the strains.^{80–84} Kryazhimskiy (2021)⁸⁵ has recently demonstrated that epistatic interactions at lower levels of network hierarchy (e.g., those found among metabolic genes) can propagate to higher levels influencing the phenotype of strains. Taking cues from these, we examined the prevalence of positive epistasis between all reaction pairs in the LAB GEMs under different diet conditions (see **STAR Methods**). These simulations indicated that LbCs and LbSv, both of which are originally isolated from the human environment, have higher prevalence of positive epistasis (Figure 5G). In order to assess if the *in silico* estimated prevalence of positive epistasis serves as metric to predict probiotic gut survival, we carried out relevant mice experiments (see **STAR Methods**). Samples were collected from the intestinal cecum of mice fed with each of the probiotic LAB strains as well from the control group (i.e., no probiotic supplementation) and were subjected to microbiome sequencing. Differential abundance analysis of operational strain units annotated from the 16S rRNA sequencing revealed significantly higher amounts of corresponding LAB in the probiotic-fed mice (compared with that of the control; Table S6), clearly indicating that they have persisted in the gut environment during the 2-week intervention. We further observed a significant positive correlation (Spearman's rank correlation, $r = 0.9078$; Figure 5H) between the fold change of LAB species abundance in the probiotic LAB-fed mice to that of control mice and simulated prevalence of positive epistasis, highlighting the utility of the proposed metric to infer short-term intestinal survival of the probiotic species. Particularly, we observed that mice fed with LbSv and LbCs strains were able to persist in the mouse gut better than other LAB strains, while LcLt displayed limited ability to persist in the gut ($\log_2 FC(LbSv) = 8.89$, $\log_2 FC(LbCs) = 7.67$, $\log_2 FC(LcLt) = 0.76$), an observation that is consistent with our simulations. When we further analyzed the metabolic subsystems of these epistatic reaction pairs, we noticed that the reactions essential for biomass formation were significantly enriched (Table S9), similar to an earlier study that reported that positive epistasis usually occurs between essential reactions.⁸⁶ Furthermore, the relevant pathways to probiotic adhesion, survival, and antibiotic resistance in the gut environment, such as peptidoglycan biosynthesis and polyketide sugar unit biosynthesis,⁸⁷ were also significantly enriched, possibly explaining how the presence of positive epistatic landscapes could confer strains with the ability to persist in the gut environment.

Next, we evaluated the capacities of LAB to produce relevant bioactive compounds (postbiotics) that are beneficial to health, such as the short-chain fatty acids (SCFAs) including butyrate, propionate, and acetate. To do so, we computed the theoretical maximum yield of various postbiotic compounds while constraining the biomass flux at 50% of its maximum achievable value (see **STAR Methods** for details). It should be noted that some metabolites secreted by LAB are also known to aggravate health conditions. For example, lipoteichoic acid (LTA) promotes

pro-inflammatory responses such as TNF- α induction.⁸⁸ Thus, their overall benefit to the host depends on the balance between the production of beneficial and detrimental metabolites. LbPt and LcLt have a high amount of growth-coupled production of various beneficial postbiotic compounds such as SCFAs, amino acids, and vitamins across all diets comparatively (Figure 5I). LbCs and LbSv also secrete similar levels of SCFAs across various diets, although they are incapable of synthesizing vitamins and amino acids as high as LbPt and LcLt. On the other hand, LeMt and LbFm synthesized relatively low levels of SCFAs. We also found that certain postbiotics biosynthetic pathways, including SCFAs, diacetyl, UDP-glucose, UDP-galactose, and tryptophan, were highly expressed in LbCs and LbPt, while they were downregulated in LbFm and LbSv (Table S8). Furthermore, similar to the growth patterns observed in different diets, the production level of beneficial metabolites is higher in carbohydrate-rich diets, e.g., EU average, DACH, high-fiber, and diabetes diets, while it is low in a high-fat/low-carb diet. Interestingly, plant-based diets such as vegan, Mediterranean, and diabetes diets consistently enabled several LAB to produce higher amounts of vitamins. Since a unique pattern of gene expression in fermentation pathways was observed for each LAB, we again predicted their postbiotic production capacities with additional constraints based on transcriptome data. These new simulations show that while the production of primary metabolites such as lactate and acetate is not significantly affected by gene expression, others may be partially regulated at the transcript level in a few species including LbCs and LeMt (Figure S10). Here, it should be noted that while we analyzed the biosynthetic capacities of LAB, some of them could also consume/degrade the postbiotics including certain amino acids and vitamins synthesized by other LAB, as suggested by the auxotrophic analysis (Figure 4B), and this may partly compromise the beneficial effects.

***In silico* evaluation of LAB ability to cross-feed common gut microbes unravels their positive associations with commensal species**

Probiotics have also shown to selectively interact with the gut flora and alter their abundance,^{89,90} particularly in the context of their use as live biotherapeutics. Here, we investigated the ability of LAB to cross-feed common gut microbes such as probiotics (*Bifidobacterium adolescentis*), commensals (*Akkermansia muciniphila*, *Bacteroides thetaiotaomicron*, *Escherichia coli* W3110, *Faecalibacterium prausnitzii*, *Roseburia hominis*), opportunistic pathogens (*Klebsiella pneumoniae*, *Pseudomonas aeruginosa*), and pathogens (*Salmonella enterica* ser. *Typhimurium*, *Shigella flexneri*, *Staphylococcus aureus*) (see **STAR Methods**). *In silico* simulations indicated that most LAB were able to enhance the growth of commensals such as *F. prausnitzii* and *R. hominis* (Figure 6A), which are widely recognized for their protective role in many metabolic and inflammatory disorders. However, the cross-feeding interactions of LAB with pathogens indicated mixed trends: while they did not show any appreciable growth-promoting effects in some pathogens such as *P. aeruginosa* and *S. aureus*, they promote the growth of some pathogens, e.g., *K. pneumoniae* and *S. enterica*, by cross-feeding nucleosides and vitamins. Interestingly, among all LAB,

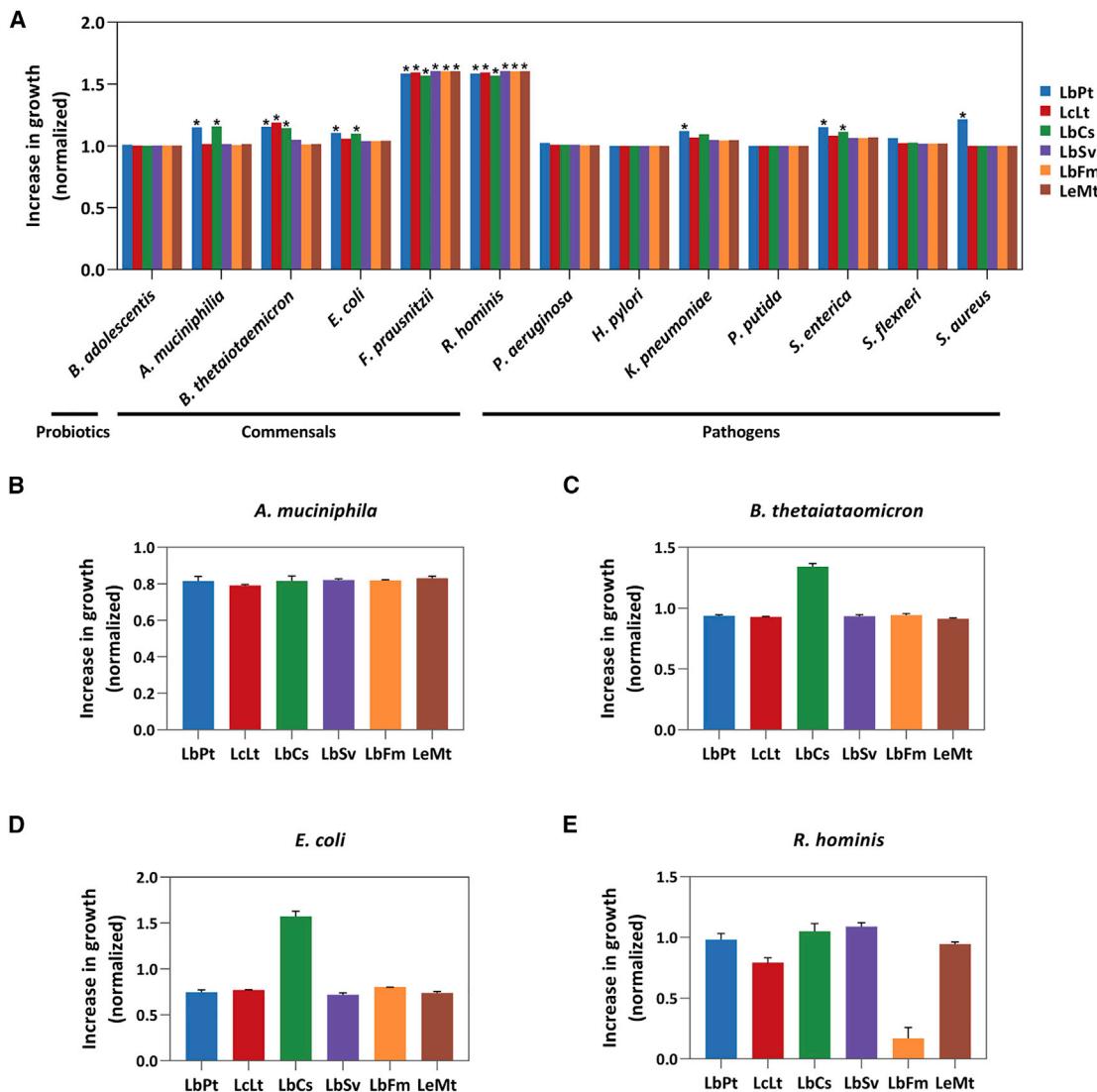


Figure 6. Potential cross-feeding effects of LAB on other gut microbes

(A) Model-simulated increase in growth of various gut microbes based on the cross-feeding of LAB-produced compounds and (B–E) experimentally observed increase/decrease in the optical density (OD) (measured after 12 h) of each commensal when it grows separately with the supernatants from LAB cell culture. In (A), “**” indicates the LAB have a cross-feeding effect based on the criteria that the growth rate of the commensal is increased by $\geq 10\%$ when supplemented with LAB supernatants. Three samples were used per LAB for experimental measurements in (B)–(E). The error bars in (B)–(E) indicate standard deviations.

LbCs possess the superior ability to promote the growth of several commensals such as *B. thetaiotaomicron*, *E. coli*, and *R. hominis*. To further validate these simulation results, we first conducted an extensive literature survey that allowed us to confirm that LAB indeed have positive interactions with certain commensal strains. For example, a previous study reported that the culture supernatants from LcLt and *Lactobacillus paracasei* stimulated *F. prausnitzii* growth by supplementing B vitamins and lactate/acetate.⁹¹ However, we could not find any evidence of similar interactions between LAB and other keystone gut commensals such as *A. muciniphilia*, *B. thetaiotaomicron*, and *R. hominis*. Therefore, we newly performed cell culture experiments to evaluate such cross-feeding between LAB and

A. muciniphilia, *B. thetaiotaomicron*, *E. coli*, and *R. hominis* (see STAR Methods, Figures 6B–6E). Of the total 24 combinations experimentally evaluated (six LAB \times four commensal strains), 17 of them agreed well with model predictions (Matthews correlation coefficient = 0.40242). Consistent with experimental observations, simulations indicated LbCs has the best growth-promoting effects on various commensal microbiota, including *B. thetaiotaomicron*, *E. coli*, and *R. hominis*. Detailed analysis of the metabolites that were cross-fed between LAB and these species indicated that pyruvate, glycine, lactate, and N-glucosamine secreted by LbCs positively influenced the commensals (Table S10). While simulations also showed a marginal positive effect of LbCs on *A. muciniphilia*, experiments showed

an inhibitory effect. Such discrepancy between model predictions and experimental observations could be attributed to at least two reasons: (1) the metabolite-toxicity on bacterial cells is not accounted by metabolic models, and (2) production of anti-microbial compounds such as bacteriocins by LAB is not captured by metabolic models. Similarly, although simulations indicated LbPt can promote *E. coli* growth, our experiments as well as a previous study⁹² suggested LbPt inhibits *E. coli* growth, which could be due to the fact that LbPt is one of the LAB producing multiple bacteriocins.

DISCUSSION

While conventional probiotics have been generally developed as food supplements to maintain a good gastrointestinal health, next-generation probiotics are expected to provide preventive and therapeutic means to intervene with various diseased states.⁹³ In this regard, a key challenge is to design combinatorial probiotic formulae based on their synergistic interactions within and with the microbiome. Strain-specific differences and influence of environmental factors including diet could further confound their application.⁶⁸ Moreover, there exists no rational framework for screening and functionally classifying probiotics targeting human health, although the Food and Agriculture Organization of the United Nations has currently placed certain guidelines that highly rely on various experimental methods, such as those involving animal models.⁹⁴ However, it is impractical to test the efficacy of all candidates and their combinations at the preclinical stage. Therefore, herein we provide a first-pass rational screening procedure to earmark promising candidates using an *in silico* model-guided approach that could be further verified *in vivo*.

Our comparative genomic and transcriptomic analyses suggested LbPt and LbCs are the common LAB with a higher proportion of functionally relevant genes supporting gastrointestinal survival via the resistance to acid, bile, and oxidative stress, production of bacteriocins and CEPs, and mucosal adhesion. Interestingly, both LAB have relatively larger genomes and coincidentally belong to the facultative heterofermentative group. Another interesting finding is the extraordinary diversity of LAB cloud and shell genomes, especially enriched on carbohydrate and amino acid metabolism, which motivated us to explore the functional metabolic capabilities as well as the cellular fitness across different diet conditions. Indeed, our model-driven analysis confirmed diet-specific patterns where high-fat/low-carb (ketogenic) diets lead to poor probiotic performance and diets rich in inulin-type fructooligosaccharides are preferable, in general. Consistent with our results, recent clinical studies have also shown that high-fat/low-carb diets cause unfavorable changes in gut microbiome and reduction in fecal SCFAs, which could be due to the elimination of probiotic bacteria.^{65–67} These observations clearly demonstrate the need to select appropriate probiotics depending on the dietary habit of a person or administer them as synbiotics, i.e., probiotics with relevant “prebiotics,” to enable their efficient functioning in the gut. Such rational administration of probiotics becomes even more significant for their use as live biotherapeutics. In this regard, although our results indicate LbPt, LcLt, and LeMt have the highest capabilities for

the postbiotic production, they also synthesize increased levels of LTA, which is highly implicated in the stimulation of pro-inflammatory cytokines,⁸⁸ so caution must be exercised while recommending them to patients with inflammatory bowel disease (IBD). Our simulations also hinted that LbCs could be a promising probiotic candidate to treat/manage IBD as it selectively promotes the growth of two commensals (*A. muciniphilia* and *F. prausnitzii*) that are reported to have a strong inverse correlation with IBD,⁹⁵ and it exhibits superior gut-persisting abilities *in vivo*. Numerous earlier studies have also reported that LbCs strains consistently confer anti-inflammatory and protective effect against IBD and colorectal cancer.^{96–101} However, LbPt and LeMt displayed positive interactions with both pathogens and commensals growth, and thus may preferably be avoided for treating immunocompromised patients whose natural ability to eliminate gastrointestinal pathogens is generally limited.¹⁰² These observations clearly demonstrate the advantage of adopting a multi-omics-based systems approach compared with the widely used comparative genomics analysis. Particularly, although LbCs and LbPt encode larger genomes with diverse capabilities, our analysis highlights that LbCs has better functional traits of a probiotic as it selectively synthesizes beneficial postbiotic compounds, promotes the growth of beneficial commensals, and possesses the superior ability to persist in the gut environment. Here, it should be noted that the postbiotic yields provide theoretical estimates of potential benefits to the host and can be used for strain prioritization. However, these estimates may vary depending on different experimental/clinical contexts.

It is a common narrative that the probiotic capabilities of various strains of the same species are often different.^{21,103–105} Therefore, here, we evaluated the metabolic and probiotic capabilities of three strains of three LAB species, LbPt, LcLt, and LbCs, to assess whether such differences are as high as the ones observed between the different species. For this purpose, we first reconstructed the corresponding strain-specific GEMs from the representative strains using orthologous genes. We further curated each strain-specific model by accounting for its unique genome content (see **STAR Methods** for detailed reconstruction procedure). Following the reconstruction of strain-specific GEMs, we first compared the metabolic capabilities of different strains by examining their ability to ferment 32 different carbon substrates, and auxotrophy requirements to 20 amino acids and eight vitamins. These results clearly indicated that while there exist some differences across the strains, they were more similar to other strains in same species than that of other species (Figure 7A). We also compared how the different strains perform in the “EU average” diet in terms of cellular fitness and metabolic interactions with other representative species in the human gut. The maximum achievable growth rates of different strains of the same species were very similar (Figure 7B), possibly due to the use of the same biomass coefficients across strains due to lack of strain-specific macromolecular composition data availability. In terms of postbiotic producing abilities under the average EU diet, while the strain-specific differences are minimal when compared across species (Figure 7F), we noticed a certain level of dependence on fermentative groups: the obligate heterofermentative LAB, i.e. LeMt and LbFm, exhibit

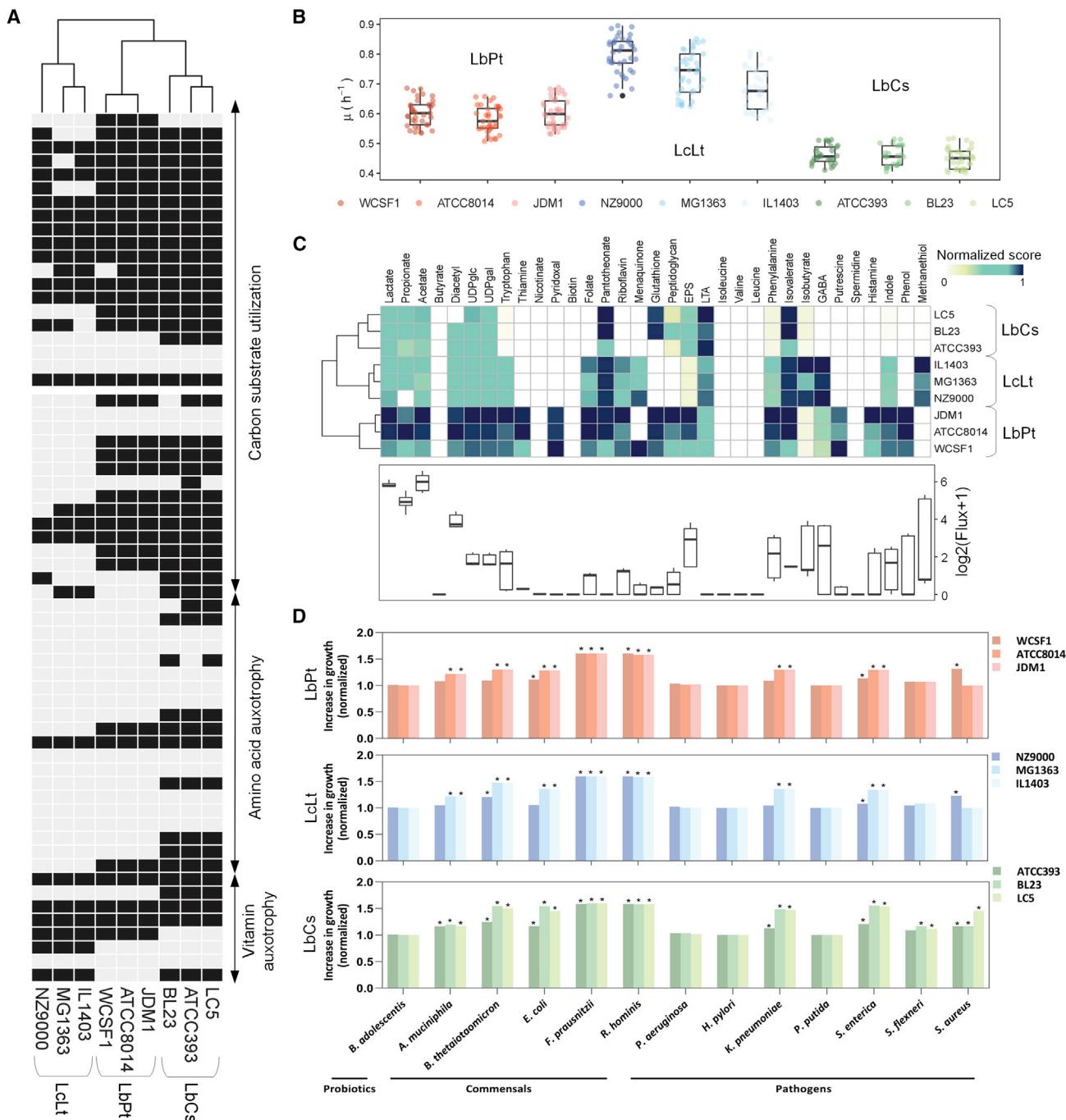


Figure 7. Strain-specific variations in metabolic and postbiotic capabilities of LAB

Strain-specific GEM predicted metabolic capabilities (A), growth rates (B), postbiotic production capabilities (C), model-simulated increase in growth of various gut microbes based on the cross-feeding of LAB-produced compounds (D). p values reported in (A) are determined by Kruskal-Wallis test. In (C), the values in the heatmap are production potential of each compound normalized to the maximum value across all diets and species. The boxplot in (C) shows the overall variation in production levels across all diets and species. In (D), “**” indicates the LAB to have a cross-feeding effect based on the criteria that the growth rate of the commensal is increased by $\geq 10\%$ when supplemented with LAB supernatants. Representative species used to evaluate pairwise interactions are the same as Figure 6.

comparable postbiotic production capabilities although they belong to different genera, suggesting the dominant role of carbon-central metabolism in governing such functionalities.

Limitations of the study

There are two main limitations of our systematic framework: (1) metabolic model-based analyses do not account for

confounding variables associated with host genetics, lifestyle, and environment, all of which affect the performance of probiotic strains. (2) Genome quality of probiotic strains used here vary considerably. In addition, some of the probiotic evaluation aspects, e.g., the influence of transient and persisting strains on host health, are still uncertain.⁶⁸ Therefore, we urge caution in extrapolating inferences from the presented framework and recommend further experimental studies and clinical investigation.

Despite such limitations, this work seeds the foundation for model-guided selection of relevant probiotics and can be further extended by accounting confounding factors in clinical settings, such as host genetics and actual diets, through community-wide efforts to foresee well-controlled clinical trials. Overall, the proposed approach is poised to accelerate the characterization of probiotic capabilities of multiple candidates beyond the conventional *Lactobacillus* and *Bifidobacterium* species, thus speeding up the rational design of personalized smart probiotics in the near future.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE

- RESOURCE AVAILABILITY

- Lead contact
- Materials availability
- Data and code availability

- EXPERIMENTAL MODEL AND SUBJECT DETAILS

- LAB strains, media and growth conditions
- Commensal culture with LAB supernatants
- LAB culture for mice supplementation
- Mouse study design

- METHOD DETAILS

- Growth and biochemical analysis of culture supernatants
- Cell volume and cell number measurements
- Determination of total protein content
- Total RNA extraction and purification
- Preparation of cDNA libraries and RNA-sequencing
- Microbiome DNA extraction and sequencing library construction

- QUANTIFICATION AND STATISTICAL ANALYSIS

- Statistical analysis of LAB growth characteristics
- Comparative genomic analysis
- Transcriptome data analysis
- Microbiome data analysis
- Reconstruction of LAB genome-scale models
- Auxotrophy-weighted transcriptome Z score metric
- Constraint-based flux analysis with macromolecular crowding constraints
- Estimation of cell crowding coefficients (a)
- Simulation of LAB fermentation profiles as a function of oxygen
- Evaluation of positive and negative epistasis in LAB GEMs

- *In silico* analysis of LAB phenotypes in various diets
- *In silico* examination of LAB ability to cross-feed common gut microbes

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2022.111735>.

ACKNOWLEDGMENTS

This work was supported by Biomedical Research Council of A*STAR (Agency for Science, Technology and Research), Singapore, Industry Alignment Fund (PRE-POSITIONING) Program H18/01/a0/C11 (Food Structure Engineering for Nutrition and Health), A*STAR, the National Research Foundation of Korea (NRF) grant (2020R1A2C2007192) and Korea Innovation Foundation grant (2021-DD-UP-0369) funded by the MSIT, and the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries (IPIET) funded by the MAFRA (32136-05-1-HD050). The authors thank Shawn Hoon and Fong Tian Wong for advice on RNA-seq library preparation steps, Elena Heng for assistance in transcriptome sequencing experiments, and Tessa Tan for the advice and support on HPLC analysis.

AUTHOR CONTRIBUTIONS

L.K., M.L., D.S.-W.O., and D.-Y.L. conceived the project. L.K., P.-Y.L., M.B., and W.X.L. performed LAB experiments, and they profiled culture supernatants with inputs from S.K.N. P.L.H. and D.-S.P. performed commensal culture experiments. L.K. formulated the new LABDM with inputs from D.S.-W.O. L.K., P.-Y.L., and M.L. were involved in transcriptome sequencing and relevant bioinformatics analysis. L.K. and M.L. performed the comparative genomic analyses. Y.Q.L. and D.K. analyzed the microbiome data with inputs from M.L. L.K. and M.L. reconstructed the genome-scale models, developed *in silico* methods, and implemented them with assistance from Y.Q.L. L.K. and M.L. drafted the initial manuscript. L.K., M.L., D.S.-W.O., and D.-Y.L. were involved in editing and revising the manuscript. D.S.-W.O. and D.-Y.L. supervised and coordinated the project.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 9, 2022

Revised: June 24, 2022

Accepted: November 7, 2022

Published: December 6, 2022

REFERENCES

1. Glont, M., Nguyen, T.V.N., Graesslin, M., Hälke, R., Ali, R., Schramm, J., Wimalaratne, S.M., Kothamachu, V.B., Rodriguez, N., Swat, M.J., et al. (2018). BioModels: expanding horizons to include more modelling approaches and formats. *Nucleic Acids Res.* **46**, 1248–1253.
2. Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E., Pavlov, A., Pavlova, N., Karamychev, V., Polouchine, N., et al. (2006). Comparative genomics of the lactic acid bacteria. *Proc. Natl. Acad. Sci. USA* **103**, 15611–15616.
3. Zheng, J., Wittouck, S., Salvetti, E., Franz, C.M.A.P., Harris, H.M.B., Mattarelli, P., O’toole, P.W., Pot, B., Vandamme, P., Walter, J., et al. (2020). A taxonomic note on the genus *Lactobacillus*: description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *Int. J. Syst. Evol. Microbiol.* **70**, 2782–2858.
4. Gänzle, M.G. (2015). Lactic metabolism revisited: metabolism of lactic acid bacteria in food fermentations and food spoilage. *Curr. Opin. Food Sci.* **2**, 106–117.

5. Carr, F.J., Chill, D., and Maida, N. (2002). The lactic acid bacteria: a literature survey. *Crit. Rev. Microbiol.* 28, 281–370.
6. Golić, N., Veljović, K., Popović, N., Djokić, J., Strahinić, I., Mrvaljević, I., and Terzić-Vidojević, A. (2017). In vitro and in vivo antagonistic activity of new probiotic culture against *Clostridium difficile* and *Clostridium perfringens*. *BMC Microbiol.* 17, 108.
7. Shelby, R.D., Janzow, G.E., Mashburn-Warren, L., Galley, J., Tengberg, N., Navarro, J., Conces, M., Bailey, M.T., Goodman, S.D., and Besner, G.E. (2020). A novel probiotic therapeutic in a murine model of *Clostridioides difficile* colitis. *Gut Microb.* 12, 1814119.
8. Ma, Y., Yang, J.Y., Peng, X., Xiao, K.Y., Xu, Q., and Wang, C. (2020). Which probiotic has the best effect on preventing *Clostridium difficile*-associated diarrhea? A systematic review and network meta-analysis. *J. Dig. Dis.* 21, 69–80.
9. Nakazato, G., Paganelli, F.L., Lago, J.C., Aoki, F.H., Mobilon, C., Brocchi, M., Stehling, E.G., and Silveira, W.D. (2011). *Lactobacillus acidophilus* decreases *Salmonella typhimurium* invasion in vivo. *J. Food Saf.* 31, 284–289.
10. Galdeano, C.M., and Perdigón, G. (2006). The probiotic bacterium *Lactobacillus casei* induces activation of the gut mucosal immune system through innate immunity. *Clin. Vaccine Immunol.* 13, 219–226.
11. Garcia-Castillo, V., Komatsu, R., Clua, P., Indo, Y., Takagi, M., Salva, S., Islam, M.A., Alvarez, S., Takahashi, H., Garcia-Cancino, A., et al. (2019). Evaluation of the immunomodulatory activities of the probiotic strain *Lactobacillus fermentum* UCO-979C. *Front. Immunol.* 10, 1376.
12. Zhang, D.I., Li, C., Shi, R., Zhao, F., and Yang, Z. (2020). *Lactobacillus fermentum* JX306 restrain d-galactose-induced oxidative stress of mice through its antioxidant activity. *Pol. J. Microbiol.* 69, 205–215.
13. Zhao, J., Tian, F., Yan, S., Zhai, Q., Zhang, H., and Chen, W. (2018). *Lactobacillus plantarum* CCFM10 alleviating oxidative stress and restoring the gut microbiota in d-galactose-induced aging mice. *Food Funct.* 9, 917–924.
14. Toi, M., Hirota, S., Tomotaki, A., Sato, N., Hozumi, Y., Anan, K., Nagashima, T., Tokuda, Y., Masuda, N., Ohsumi, S., et al. (2013). Probiotic beverage with soy isoflavone consumption for breast cancer prevention: a case-control study. *Curr. Nutr. Food Sci.* 9, 194–200.
15. El-Nezami, H.S., Polychronaki, N.N., Ma, J., Zhu, H., Ling, W., Salminen, E.K., Juvonen, R.O., Salminen, S.J., Poussa, T., and Mykkänen, H.M. (2006). Probiotic supplementation reduces a biomarker for increased risk of liver cancer in young men from Southern China. *Am. J. Clin. Nutr.* 83, 1199–1203.
16. Hardy, H., Harris, J., Lyon, E., Beal, J., and Foey, A.D. (2013). Probiotics, prebiotics and immunomodulation of gut mucosal defences: homeostasis and immunopathology. *Nutrients* 5, 1869–1912.
17. van Baarlen, P., Wells, J.M., and Kleerebezem, M. (2013). Regulation of intestinal homeostasis and immunity with probiotic lactobacilli. *Trends Immunol.* 34, 208–215.
18. Azas-Braesco, V., Bresson, J.L., Guarner, F., and Cortier, G. (2010). Not all lactic acid bacteria are probiotics, but some are. *Br. J. Nutr.* 103, 1079–1081.
19. Ibou-Zekri, N., Blum, S., Schiffrin, E.J., and Von der Weid, T. (2003). Divergent patterns of colonization and immune response elicited from two intestinal *Lactobacillus* strains that display similar properties in vitro. *Infect. Immun.* 71, 428–436.
20. Campana, R., van Hemert, S., and Baffone, W. (2017). Strain-specific probiotic properties of lactic acid bacteria and their interference with human intestinal pathogens invasion. *Gut Pathog.* 9, 12.
21. Liu, J., Hu, D., Chen, Y., Huang, H., Zhang, H., Zhao, J., Gu, Z., and Chen, W. (2018). Strain-specific properties of *Lactobacillus plantarum* for prevention of *Salmonella* infection. *Food Funct.* 9, 3673–3682.
22. Dietrich, C.G., Kottmann, T., and Alavi, M. (2014). Commercially available probiotic drinks containing *Lactobacillus casei* DN-114001 reduce antibiotic-associated diarrhea. *World J. Gastroenterol.* 20, 15837–15844.
23. De Roock, S., Van Elk, M., Van Dijk, M.E.A., Timmerman, H.M., Rijkers, G.T., Prakken, B.J., Hoekstra, M.O., and De Kleer, I.M. (2010). Lactic acid bacteria differ in their ability to induce functional regulatory T cells in humans. *Clin. Exp. Allergy* 40, 103–110.
24. Ashraf, R., Vasiljevic, T., Day, S.L., Smith, S.C., and Donkor, O.N. (2014). Lactic acid bacteria and probiotic organisms induce different cytokine profile and regulatory T cells mechanisms. *J. Funct. Foods* 6, 395–409.
25. Xia, Y., Lu, M., Chen, G., Cao, J., Gao, F., Wang, M., Liu, Z., Zhang, D., Zhu, H., and Yi, M. (2018). Effects of dietary *Lactobacillus rhamnosus* JCM1136 and *Lactococcus lactis* subsp. *lactis* JCM5805 on the growth, intestinal microbiota, morphology, immune response and disease resistance of juvenile Nile tilapia, *Oreochromis niloticus*. *Fish Shellfish Immunol.* 76, 368–379.
26. Möller, P.L., Paerregaard, A., Gad, M., Kristensen, N.N., and Claesson, M.H. (2005). Colitic scid mice fed *Lactobacillus* spp. show an ameliorated gut histopathology and an altered cytokine profile by local T cells. *Inflamm. Bowel Dis.* 11, 814–819.
27. D'Arienzo, R., Bozzella, G., Rossi, M., De Bellis, P., Lavermicocca, P., and Sisto, A. (2011). Distinct immunomodulatory properties of *Lactobacillus paracasei* strains. *J. Appl. Microbiol.* 111, 1482–1491.
28. Nerstedt, A., Nilsson, E.C., Ohlson, K., Håkansson, J., Thomas Svensson, L., Löwenadler, B., Svensson, U.K., and Mahlapuu, M. (2007). Administration of *Lactobacillus* evokes coordinated changes in the intestinal expression profile of genes regulating energy homeostasis and immune phenotype in mice. *Br. J. Nutr.* 97, 1117–1127.
29. Martín, R., Olivares, M., Marín, M.L., Fernández, L., Xaus, J., and Rodríguez, J.M. (2005). Probiotic potential of 3 lactobacilli strains isolated from breast milk. *J. Hum. Lact.* 21, 8–17; quiz 18–21, 41.
30. Ruiz, L., Margolles, A., and Sánchez, B. (2013). Bile resistance mechanisms in *lactobacillus* and *Bifidobacterium*. *Front. Microbiol.* 4, 396.
31. Zhou, Z., Wang, W., Liu, W., Gatlin, D.M., Zhang, Y., Yao, B., and Ringø, E. (2012). Identification of highly-adhesive gut *Lactobacillus* strains in zebrafish (*Danio rerio*) by partial *rpoB* gene sequence analysis. *Aquaculture* 370–371, 150–157.
32. Pagnini, C., Corleto, V.D., Martorelli, M., Lanini, C., D'Ambra, G., Di Giulio, E., and Delle Fave, G. (2018). Mucosal adhesion and anti-inflammatory effects of *Lactobacillus rhamnosus* GG in the human colonic mucosa: a proof-of-concept study. *World J. Gastroenterol.* 24, 4652–4662.
33. Turpin, W., Humbot, C., Noordine, M.L., Thomas, M., and Guyot, J.P. (2012). *Lactobacillaceae* and cell adhesion: genomic and functional screening. *PLoS One* 7, e38034.
34. Wu, C., Huang, J., and Zhou, R. (2017). Genomics of lactic acid bacteria: current status and potential applications. *Crit. Rev. Microbiol.* 43, 393–404.
35. De Filippis, F., Pasolli, E., and Ercolini, D. (2020). The food-gut axis: lactic acid bacteria and their link to food, the gut microbiome and human health. *FEMS Microbiol. Rev.* 44, 454–489.
36. Sun, Z., Harris, H.M.B., McCann, A., Guo, C., Argimón, S., Zhang, W., Yang, X., Jeffery, I.B., Cooney, J.C., Kagawa, T.F., et al. (2015). Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. *Nat. Commun.* 6, 8322.
37. Pasolli, E., De Filippis, F., Mauriello, I.E., Cumbo, F., Walsh, A.M., Leech, J., Cotter, P.D., Segata, N., and Ercolini, D. (2020). Large-scale genome-wide analysis links lactic acid bacteria from food with the gut microbiome. *Nat. Commun.* 11, 2610.
38. Valeriano, V.D.V., Oh, J.K., Bagon, B.B., Kim, H., and Kang, D.K. (2019). Comparative genomic analysis of *Lactobacillus mucosae* LM1 identifies potential niche-specific genes and pathways for gastrointestinal adaptation. *Genomics* 111, 24–33.
39. Drissi, F., Merhej, V., Angelakis, E., El Kaoutari, A., Carrière, F., Henrissat, B., and Raoult, D. (2014). Comparative genomics analysis of *Lactobacillus* species associated with weight gain or weight protection. *Nutr. Diabetes* 4, e109.

40. Klaenhammer, T.R., Barrangou, R., Buck, B.L., Azcarate-Peril, M.A., and Altermann, E. (2005). Genomic features of lactic acid bacteria effecting bioprocessing and health. *FEMS Microbiol. Rev.* 29, 393–409.
41. Alvarez-Sieiro, P., Montalbán-López, M., Mu, D., and Kuipers, O.P. (2016). Bacteriocins of lactic acid bacteria: extending the family. *Appl. Microbiol. Biotechnol.* 100, 2939–2951.
42. Zhang, G., Mills, D.A., and Block, D.E. (2009). Development of chemically defined media supporting high-cell-density growth of lactococci, enterococci, and streptococci. *Appl. Environ. Microbiol.* 75, 1080–1087.
43. Koduru, L., Kim, Y., Bang, J., Lakshmanan, M., Han, N.S., and Lee, D.-Y. (2017). Genome-scale modeling and transcriptome analysis of *Leuconostoc mesenteroides* unravel the redox governed metabolic states in obligate heterofermentative lactic acid bacteria. *Sci. Rep.* 7, 15721.
44. Teusink, B., Wiersma, A., Molenaar, D., Francke, C., De Vos, W.M., Siezen, R.J., and Smid, E.J. (2006). Analysis of growth of *Lactobacillus plantarum* WCFS1 on a complex medium using a genome-scale metabolic model. *J. Biol. Chem.* 281, 40041–40048.
45. Flahaut, N.A.L., Wiersma, A., Van De Bunt, B., Martens, D.E., Schaap, P.J., Sijtsma, L., Dos Santos, V.A.M., and De Vos, W.M. (2013). Genome-scale metabolic model for *Lactococcus lactis* MG1363 and its application to the analysis of flavor formation. *Appl. Microbiol. Biotechnol.* 97, 8729–8739.
46. Thiele, I., and Palsson, B.O. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* 5, 93–121.
47. Beg, Q.K., Vazquez, A., Ernst, J., de Menezes, M.A., Bar-Joseph, Z., Barabási, A.L., and Oltvai, Z.N. (2007). Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc. Natl. Acad. Sci. USA* 104, 12663–12668.
48. Vazquez, A., Liu, J., Zhou, Y., and Oltvai, Z.N. (2010). Catabolic efficiency of aerobic glycolysis: the Warburg effect revisited. *BMC Syst. Biol.* 4, 58.
49. Shlomi, T., Benyaminini, T., Gottlieb, E., Sharan, R., and Ruppin, E. (2011). Genome-scale metabolic modeling elucidates the role of proliferative adaptation in causing the warburg effect. *PLoS Comput. Biol.* 7, e1002018.
50. van Hoek, M.J.A., and Merks, R.M.H. (2012). Redox balance is key to explaining full vs. partial switching to low-yield metabolism. *BMC Syst. Biol.* 6, 22.
51. Oliveira, A.P., Nielsen, J., and Förster, J. (2005). Modeling *Lactococcus lactis* using a genome-scale flux model. *BMC Microbiol.* 5, 39.
52. Morishita, T., and Yajima, M. (1995). Incomplete operation of biosynthetic and bioenergetic functions of the citric acid cycle in multiple auxotrophic lactobacilli. *Biosci. Biotechnol. Biochem.* 59, 251–255.
53. Costa, S., Summa, D., Semeraro, B., Zappaterra, F., Rugiero, I., and Tamburini, E. (2020). Fermentation as a strategy for bio-transforming waste into resources: lactic acid production from agri-food residues. *Fermentation* 7, 3.
54. Martín, R., Jiménez, E., Olivares, M., Marín, M.L., Fernández, L., Xaus, J., and Rodríguez, J.M. (2006). *Lactobacillus salivarius* CECT 5713, a potential probiotic strain isolated from infant feces and breast milk of a mother-child pair. *Int. J. Food Microbiol.* 112, 35–43.
55. Claesson, M.J., Li, Y., Leahy, S., Canchaya, C., Van Pijkeren, J.P., Cerdeno-Tarraga, A.M., Parkhill, J., Flynn, S., O'Sullivan, G.C., Collins, J.K., et al. (2006). Multireplicon genome architecture of *Lactobacillus salivarius*. *Proc. Natl. Acad. Sci. USA* 103, 6718–6723.
56. Teusink, B., van Enckevort, F.H.J., Francke, C., Wiersma, A., Wegkamp, A., Smid, E.J., and Siezen, R.J. (2005). In silico reconstruction of the metabolic pathways of *Lactobacillus plantarum*: comparing predictions of nutrient requirements with those from growth experiments. *Appl. Environ. Microbiol.* 71, 7253–7262.
57. Pastink, M.I., Teusink, B., Hols, P., Visser, S., de Vos, W.M., and Hugenholtz, J. (2009). Genome-scale model of *Streptococcus thermophilus* LMG18311 for metabolic comparison of lactic acid bacteria. *Appl. Environ. Microbiol.* 75, 3627–3633.
58. Vinay-Lara, E., Hamilton, J.J., Stahl, B., Broadbent, J.R., Reed, J.L., and Steele, J.L. (2014). Genome-scale reconstruction of metabolic networks of *Lactobacillus casei* ATCC 334 and 12A. *PLoS One* 9, e110785.
59. Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.Y., Keilbaugh, S.A., Bewtra, M., Knights, D., Walters, W.A., Knight, R., et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334, 105–108.
60. David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varma, Y., Fischbach, M.A., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505, 559–563.
61. Zarrinpar, A., Chaix, A., Yoosoph, S., and Panda, S. (2014). Diet and feeding pattern affect the diurnal dynamics of the gut microbiome. *Cell Metab.* 20, 1006–1017.
62. Gänzle, M.G., and Follador, R. (2012). Metabolism of oligosaccharides and starch in lactobacilli: a review. *Front. Microbiol.* 3, 1–15.
63. Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S.N., Richelle, A., Heinken, A., Haraldsdóttir, H.S., Wachowiak, J., Keating, S.M., Vlasov, V., et al. (2019). Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protoc.* 14, 639–702.
64. Yao, J., and Rock, C.O. (2017). Exogenous fatty acid metabolism in bacteria. *Biochimie* 141, 30–39.
65. Wan, Y., Wang, F., Yuan, J., Li, J., Jiang, D., Zhang, J., Li, H., Wang, R., Tang, J., Huang, T., et al. (2019). Effects of dietary fat on gut microbiota and faecal metabolites, and their relationship with cardiometabolic risk factors: a 6-month randomised controlled-feeding trial. *Gut* 68, 1417–1429.
66. Swidsinski, A., Dörrfel, Y., Loening-Baucke, V., Gille, C., Göktas, Ö., Reihsauer, A., Neuhäus, J., Weylandt, K.H., Guschin, A., and Bock, M. (2017). Reduced mass and diversity of the colonic microbiome in patients with multiple sclerosis and their improvement with ketogenic diet. *Front. Microbiol.* 8, 1141.
67. Brinkworth, G.D., Noakes, M., Clifton, P.M., and Bird, A.R. (2009). Comparative effects of very low-carbohydrate, high-fat and high-carbohydrate, low-fat weight-loss diets on bowel habit and faecal short-chain fatty acids and bacterial populations. *Br. J. Nutr.* 101, 1493–1502.
68. Suez, J., Zmora, N., Segal, E., and Elinav, E. (2019). The pros, cons, and many unknowns of probiotics. *Nat. Med.* 25, 716–729.
69. Vlasova, A.N., Chattha, K.S., Kandasamy, S., Liu, Z., Esseili, M., Shao, L., Rajashekara, G., and Saif, L.J. (2013). Lactobacilli and Bifidobacteria promote immune homeostasis by modulating innate immune responses to human Rotavirus in neonatal gnotobiotic pigs. *PLoS One* 8, e76962.
70. Pavan, S., Desreumaux, P., and Mercenier, A. (2003). Use of mouse models to evaluate the persistence, safety, and immune modulation capacities of lactic acid bacteria. *Clin. Diagn. Lab. Immunol.* 10, 696–701.
71. Wu, M., McNulty, N.P., Rodionov, D.A., Khoroshkin, M.S., Griffin, N.W., Cheng, J., Latreille, P., Kerstetter, R.A., Terrapon, N., Henrissat, B., et al. (2015). Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut *Bacteroides*. *Science* 350, aac5992.
72. Xiao, Y., Zhao, J., Zhang, H., Zhai, Q., and Chen, W. (2021). Mining genome traits that determine the different gut colonization potential of *Lactobacillus* and *Bifidobacterium* species. *Microb. Genom.* 7, 000581.
73. Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., and Gordon, J.I. (2009). Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe* 6, 279–289.
74. Watson, A.R., Füssel, J., Veseli, I., Delongchamp, J.Z., Silva, M., Trigodet, F., Lolans, K., Shaiber, A., Fogarty, E., Runde, J.M., et al. (2022). Metabolic independence drives gut microbial colonization and resilience in health and disease. Preprint at bioRxiv. <https://doi.org/10.1101/2021.03.02.433653>.

75. De Castro, J.A., Kesavelu, D., Lahiri, K.R., Chaijitraruch, N., Chongsrisawat, V., Jog, P.P., Liaw, Y.H., Nguyen, G.K., Nguyen, T.V.H., Pai, U.A., et al. (2020). Recommendations for the adjuvant use of the poly-antibiotic-resistant probiotic *Bacillus clausii* (O/C, SIN, N/R, T) in acute, chronic, and antibiotic-associated diarrhea in children: consensus from Asian experts. *Trop. Dis. Travel Med. Vaccines* 6, 21–15.
76. Kelesidis, T., and Pothoulakis, C. (2012). Efficacy and safety of the probiotic *Saccharomyces boulardii* for the prevention and therapy of gastrointestinal disorders. *Therap. Adv. Gastroenterol.* 5, 111–125.
77. Waśko, A., Skrzypczak, K., Polak-Berecka, M., and Kuzdralski, A. (2012). Genetic mechanisms of variation in erythromycin resistance in *Lactobacillus rhamnosus* strains. *J. Antibiot.* 65, 583–586.
78. Lokesh, D., Parkesh, R., and kammara, R. (2018). *Bifidobacterium adolescentis* is intrinsically resistant to antitubercular drugs. *Sci. Rep.* 8, 11897.
79. Duranti, S., Lugli, G.A., Mancabelli, L., Turroni, F., Milani, C., Mangifesta, M., Ferrario, C., Anzalone, R., Viappiani, A., van Sinderen, D., and Ventura, M. (2017). Prevalence of antibiotic resistance genes among human gut-derived bifidobacteria. *Appl. Environ. Microbiol.* 83, e02894–16.
80. Trindade, S., Sousa, A., Xavier, K.B., Dionisia, F., Ferreira, M.G., and Gordo, I. (2009). Positive epistasis drives the acquisition of multidrug resistance. *PLoS Genet.* 5, e1000578.
81. Zhang, T.H., Dai, L., Barton, J.P., Du, Y., Tan, Y., Pang, W., Chakraborty, A.K., Lloyd-Smith, J.O., and Sun, R. (2020). Predominance of positive epistasis among drug resistance-associated mutations in HIV-1 protease. *PLoS Genet.* 16, e1009009.
82. Hall, A.R., and Maclean, R.C. (2011). Epistasis buffers the fitness effects of Rifampicin- resistance mutations in *Pseudomonas aeruginosa*. *Evolution* 65, 2370–2379.
83. Jasnos, L., Tomala, K., Paczesniak, D., and Korona, R. (2008). Interactions between stressful environment and gene deletions alleviate the expected average loss of fitness in yeast. *Genetics* 178, 2105–2111.
84. Zhang, Z., Bendixsen, D.P., Janzen, T., Nolte, A.W., Greig, D., and Stelkens, R. (2020). Recombining your way out of trouble: the genetic architecture of hybrid fitness under environmental stress. *Mol. Biol. Evol.* 37, 167–182.
85. Kryazhimskiy, S. (2021). Emergence and propagation of epistasis in metabolic networks. *Elife* 10, e60200.
86. He, X., Qian, W., Wang, Z., Li, Y., and Zhang, J. (2010). Prevalent positive epistasis in *Escherichia coli* and *Saccharomyces cerevisiae* metabolic networks. *Nat. Genet.* 42, 272–276.
87. Martínez, B., Rodríguez, A., Kulakauskas, S., and Chapot-Chartier, M.P. (2020). Cell wall homeostasis in lactic acid bacteria: threats and defences. *FEMS Microbiol. Rev.* 44, 538–564.
88. Zadeh, M., Khan, M.W., Goh, Y.J., Selle, K., Owen, J.L., Klaenhammer, T., and Mohamadzadeh, M. (2012). Induction of intestinal pro-inflammatory immune responses by lipoteichoic acid. *J. Inflamm.* 9, 7.
89. Mendes, M.C.S., Paulino, D.S., Brambilla, S.R., Camargo, J.A., Persinoti, G.F., and Carvalheira, J.B.C. (2018). Microbiota modification by probiotic supplementation reduces colitis associated colon cancer in mice. *World J. Gastroenterol.* 24, 1995–2008.
90. Shin, D., Chang, S.Y., Bogere, P., Won, K., Choi, J.Y., Choi, Y.J., Lee, H.K., Hur, J., Park, B.Y., Kim, Y., and Heo, J. (2019). Beneficial roles of probiotics on the modulation of gut microbiota and immune response in pigs. *PLoS One* 14, e0220843.
91. Lebas, M., Garault, P., Carrillo, D., Codoñer, F.M., and Derrien, M. (2020). Metabolic response of *Faecalibacterium prausnitzii* to cell-free supernatants from lactic acid bacteria. *Microorganisms* 8, 1528.
92. Bu, Y., Liu, Y., Liu, Y., Wang, S., Liu, Q., Hao, H., and Yi, H. (2022). Screening and probiotic potential evaluation of bacteriocin-producing *lactiplantibacillus plantarum* in vitro. *Foods* 11, 1575.
93. El Hage, R., Hernandez-Sanabria, E., and Van de Wiele, T. (2017). Emerging trends in “smart probiotics”: functional consideration for the development of novel health and industrial applications. *Front. Microbiol.* 8, 1889.
94. de Melo Pereira, G.V., de Oliveira Coelho, B., Magalhães Júnior, A.I., Thomaz-Soccol, V., and Soccol, C.R. (2018). How to select a probiotic? A review and update of methods and criteria. *Biotechnol. Adv.* 36, 2060–2076.
95. Lopez-Siles, M., Enrich-Capó, N., Aldeguer, X., Sabat-Mir, M., Duncan, S.H., Garcia-Gil, L.J., and Martinez-Medina, M. (2018). Alterations in the abundance and co-occurrence of *Akkermansia muciniphila* and *Faecalibacterium prausnitzii* in the colonic mucosa of inflammatory bowel disease subjects. *Front. Cell. Infect. Microbiol.* 8, 281.
96. Lenoir, M., del Carmen, S., Cortes-Perez, N.G., Lozano-Ojalvo, D., Muñoz-Provencio, D., Chain, F., Langella, P., de Moreno de LeBlanc, A., LeBlanc, J.G., and Bermúdez-Humarán, L.G. (2016). *Lactobacillus casei* BL23 regulates Treg and Th17 T-cell populations and reduces DMH-associated colorectal cancer. *J. Gastroenterol.* 51, 862–873.
97. Yan, F., Cao, H., Cover, T.L., Whitehead, R., Washington, M.K., and Polk, D.B. (2007). Soluble proteins produced by probiotic bacteria regulate intestinal epithelial cell survival and growth. *Gastroenterology* 132, 562–575.
98. Hor, Y.Y., Lew, L.C., Lau, A.S.Y., Ong, J.S., Chuah, L.O., Lee, Y.Y., Choi, S.B., Rashid, F., Wahid, N., Sun, Z., et al. (2018). Probiotic *Lactobacillus casei* Zhang (LCZ) alleviates respiratory, gastrointestinal & RBC abnormality via immuno-modulatory, anti-inflammatory & anti-oxidative actions. *J. Funct. Foods* 44, 235–245.
99. Tien, M.-T., Girardin, S.E., Regnault, B., Le Bourhis, L., Dillies, M.-A., Coppée, J.Y., Bourdet-Sicard, R., Sansonetti, P.J., and Pédrón, T. (2006). Anti-inflammatory effect of *Lactobacillus casei* on *Shigella*-infected human intestinal epithelial cells. *J. Immunol.* 176, 1228–1237.
100. Vaghef-Mehraban, E., Alipour, B., Homayouni-Rad, A., Sharif, S.K., Asghari-Jafarabadi, M., and Zavvari, S. (2014). Probiotic supplementation improves inflammatory status in patients with rheumatoid arthritis. *Nutrition* 30, 430–435.
101. Schiffer, C., Lalanne, A.I., Cassard, L., Mancardi, D.A., Malbec, O., Bruhn, P., Dif, F., and Daëron, M. (2011). A strain of *Lactobacillus casei* inhibits the effector phase of immune inflammation. *J. Immunol.* 187, 2646–2655.
102. Taur, Y., and Pamer, E.G. (2013). The intestinal microbiota and susceptibility to infection in immunocompromised patients. *Curr. Opin. Infect. Dis.* 26, 332–337.
103. Aktas, B., De Wolfe, T.J., Safdar, N., Darien, B.J., and Steele, J.L. (2016). The impact of *Lactobacillus casei* on the composition of the cecal microbiota and innate immune system is strain specific. *PLoS One* 11, e0156374.
104. Lee, I.C., Caggianiello, G., van Swam, I.I., Taverne, N., Meijerink, M., Bron, P.A., Spano, G., and Kleerebezem, M. (2016). Strain-specific features of extracellular polysaccharides and their impact on *Lactobacillus plantarum*-host interactions. *Appl. Environ. Microbiol.* 82, 3959–3970.
105. Kolling, Y., Salva, S., Villena, J., and Alvarez, S. (2018). Are the immuno-modulatory properties of *Lactobacillus rhamnosus* CRL1505 peptidoglycan common for all *Lactobacilli* during respiratory infection in malnourished mice? *PLoS One* 13, e0194034.
106. Sprouffske, K., and Wagner, A. (2016). Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinf.* 17, 172.
107. Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20, 238.
108. Katoh, K., and Standley, D.M. (2013). MAFFT: multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
109. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313.

110. Emms, D.M., and Kelly, S. (2018). STAG: species tree inference from all genes. Preprint at bioRxiv. <https://doi.org/10.1101/267914>.
111. Emms, D.M., and Kelly, S. (2017). STRIDE: species tree root inference from gene duplication events. *Mol. Biol. Evol.* **34**, 3267–3278.
112. Letunic, I., and Bork, P. (2016). Interactive Tree of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **44**, 242–245.
113. Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forsslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L., et al. (2019). EggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, 309–314.
114. de Jong, A., van Heel, A.J., Kok, J., and Kuipers, O.P. (2010). BAGEL2: mining for bacteriocins in genomic data. *Nucleic Acids Res.* **38**, 647–651.
115. Andrews, S. (2015). FASTQC: A Quality Control Tool for High Throughput Sequence Data (Babraham Bioinformatics).
116. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.
117. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.
118. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* **12**, 323.
119. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300.
120. Peixoto, L., Risso, D., Poplawski, S.G., Wimmer, M.E., Speed, T.P., Wood, M.A., and Abel, T. (2015). How data analysis affects power, reproducibility and biological insight of RNA-seq studies in complex datasets. *Nucleic Acids Res.* **43**, 7664–7674.
121. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550.
122. Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* **4**, P3.
123. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583.
124. Segota, I., and Long, T. (2019). A high-resolution pipeline for 16S-sequencing identifies bacterial strains in human microbiome. Preprint at bioRxiv. <https://doi.org/10.1101/565572>.
125. Kumar, V.S., and Maranas, C.D. (2009). GrowMatch: an automated method for reconciling *in silico/in vivo* growth predictions. *PLoS Comput. Biol.* **5**, e1000308.
126. Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., Busk, P.K., Xu, Y., and Yin, Y. (2018). dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **46**, 95–101.
127. Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, 29–37.
128. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60.
129. Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **37**, 233–238.
130. Lieven, C., Beber, M.E., Olivier, B.G., Bergmann, F.T., Ataman, M., Babaei, P., Bartell, J.A., Blank, L.M., Chauhan, S., Correia, K., et al. (2020). MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.* **38**, 272–276.
131. Sonnhammer, E.L.L., and Östlund, G. (2015). InParanoid 8: orthology analysis between 273 proteomes, mostly eukaryotic. *Nucleic Acids Res.* **43**, 234–239.
132. Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.
133. Moreno-Hagelsieb, G., and Latimer, K. (2008). Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* **24**, 319–324.
134. Kim, Y.J., Eom, H.J., Seo, E.Y., Lee, D.Y., Kim, J.H., and Han, N.S. (2012). Development of a chemically defined minimal medium for the exponential growth of *Leuconostoc mesenteroides* ATCC8293. *J. Microbiol. Biotechnol.* **22**, 1518–1522.
135. Stevens, M.J.A. (2008). Transcriptome Response of *Lactobacillus Plantarum* to Global Regulator Deficiency, Stress and Other Environmental Conditions (Wageningen University and Research).
136. Bounaix, M.S., Gabriel, V., Robert, H., Morel, S., Remaud-Siméon, M., Gabriel, B., and Fontagné-Faucher, C. (2010). Characterization of glucan-producing *Leuconostoc* strains isolated from sourdough. *Int. J. Food Microbiol.* **144**, 1–9.
137. Wegmann, U., O'Connell-Motherway, M., Zomer, A., Buist, G., Shearman, C., Canchaya, C., Ventura, M., Goessmann, A., Gasson, M.J., Kuipers, O.P., et al. (2007). Complete genome sequence of the prototype lactic acid bacterium *Lactococcus lactis* subsp. *cremoris* MG1363. *J. Bacteriol.* **189**, 3256–3270.
138. Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **27**, 29–34.
139. Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C.A., Holland, T.A., Keseler, I.M., Kothari, A., Kubo, A., et al. (2014). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* **42**, D459–D471.
140. Feist, A.M., and Palsson, B.O. (2010). The biomass objective function. *Curr. Opin. Microbiol.* **13**, 344–349.
141. Lee, B. (1983). Calculation of volume fluctuation for globular protein models. *Proc. Natl. Acad. Sci. USA* **80**, 622–626.
142. Bengtsson, M., Ståhlberg, A., Rorsman, P., and Kubista, M. (2005). Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. *Genome Res.* **15**, 1388–1392.
143. Schellenberger, J., Que, R., Fleming, R.M.T., Thiele, I., Orth, J.D., Feist, A.M., Zielinski, D.C., Bordbar, A., Lewis, N.E., Rahamanian, S., et al. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* **6**, 1290–1307.
144. Schomburg, I., Chang, A., Ebeling, C., Gremse, M., Heldt, C., Huhn, G., and Schomburg, D. (2004). BRENDA, the enzyme database: updates and major new developments. *Nucleic Acids Res.* **32**, 431–433.
145. Segrè, D., Deluna, A., Church, G.M., and Kishony, R. (2005). Modular epistasis in yeast metabolism. *Nat. Genet.* **37**, 77–83.
146. Noronha, A., Modamio, J., Jarosz, Y., Guerard, E., Sompairac, N., Preciat, G., Danielsdóttir, A.D., Krecke, M., Merten, D., Haraldsdóttir, H.S., et al. (2019). The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic Acids Res.* **47**, 614–624.
147. Leblanc, J.C., Yoon, H., Kombadjian, A., and Verger, P. (2000). Nutritional intakes of vegetarian populations in France. *Eur. J. Clin. Nutr.* **54**, 443–449.
148. Elmada, I., Weichselbaum, E., König, J., de Winter A-M, R., Trolle, E., Haapala, I., Uusitalo, U., Mennen, L., Hercberg, S., Wolfram, G., et al. (2005). European nutrition and health report 2004. *Forum Nutr.* **48**, 1–220.

149. Waldmann, A., Koschizke, J.W., Leitzmann, C., and Hahn, A. (2003). Dietary intakes and lifestyle factors of a vegan population in Germany: results from the German Vegan Study. *Eur. J. Clin. Nutr.* **57**, 947–955.
150. Neal, E.G., Chaffe, H., Schwartz, R.H., Lawson, M.S., Edwards, N., Fitzsimmons, G., Whitney, A., and Cross, J.H. (2008). The ketogenic diet for the treatment of childhood epilepsy: a randomised controlled trial. *Lancet Neurol.* **7**, 500–506.
151. Elmadfa, I., and Meyer, A.L. (2012). Diet quality, a term subject to change over time. *Int. J. Vitam. Nutr. Res.* **82**, 144–147.
152. Willett, W.C., Sacks, F., Trichopoulou, A., Drescher, G., Ferro-Luzzi, A., Helsing, E., and Trichopoulos, D. (1995). Mediterranean diet pyramid: a cultural model for healthy eating. *Am. J. Clin. Nutr.* **61**, 1402–1406.
153. Berendsen, A., Santoro, A., Pini, E., Cevenini, E., Ostan, R., Pietruszka, B., Rolf, K., Cano, N., Caille, A., Lyon-Belgy, N., et al. (2013). A parallel randomized trial on the effect of a healthful diet on inflammaging and its consequences in European elderly people: design of the NU-AGE dietary intervention study. *Mech. Ageing Dev.* **134**, 523–530.
154. Nelson, K.M., Reiber, G., and Boyko, E.J.; NHANES III (2002). Diet and exercise among adults with type 2 diabetes: findings from the third national health and nutrition examination survey (NHANES III). *Diabetes Care* **25**, 1722–1728.
155. Orth, J.D., Thiele, I., and Palsson, B.Ø. (2010). What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248.
156. Monk, J.M., Charusanti, P., Aziz, R.K., Lerman, J.A., Premyodhin, N., Orth, J.D., Feist, A.M., and Palsson, B.Ø. (2013). Genome-scale metabolic reconstructions of multiple *Escherichia coli* strains highlight strain-specific adaptations to nutritional environments. *Proc. Natl. Acad. Sci. USA* **110**, 20338–20343.
157. Bosi, E., Monk, J.M., Aziz, R.K., Fondi, M., Nizet, V., and Palsson, B.Ø. (2016). Comparative genome-scale modelling of *Staphylococcus aureus* strains identifies strain-specific metabolic capabilities linked to pathogenicity. *Proc. Natl. Acad. Sci. USA* **113**, 3801–3809.
158. Raghunathan, A., Reed, J., Shin, S., Palsson, B., and Daefler, S. (2009). Constraint-based analysis of metabolic capacity of *Salmonella typhimurium* during host-pathogen interaction. *BMC Syst. Biol.* **3**, 38–16.
159. Thiele, I., Vo, T.D., Price, N.D., and Palsson, B.Ø. (2005). Expanded metabolic reconstruction of *Helicobacter pylori* (iT341 GSM/GPR): an in silico genome-scale characterization of single- and double-deletion mutants. *J. Bacteriol.* **187**, 5818–5830.
160. Nogales, J., Palsson, B., and Thiele, I. (2008). A genome-scale metabolic reconstruction of *Pseudomonas putida* KT2440: iJN746 as a cell factory. *BMC Syst. Biol.* **2**, 79.
161. Oberhardt, M.A., Puchałka, J., Fryer, K.E., Martins dos Santos, V.A.P., and Papin, J.A. (2008). Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **190**, 2790–2803.
162. Veith, N., Solheim, M., van Grinsven, K.W.A., Olivier, B.G., Levering, J., Grossholz, R., Hugenholtz, J., Holo, H., Nes, I., Teusink, B., and Kummer, U. (2015). Using a genome-scale metabolic model of *Enterococcus faecalis* V583 to assess amino acid uptake and its impact on central metabolism. *Appl. Environ. Microbiol.* **81**, 1622–1633.
163. Ottman, N., Davids, M., Suarez-Diez, M., Boeren, S., Schaap, P.J., Martins Dos Santos, V.A.P., Smidt, H., Belzer, C., and de Vos, W.M. (2017). Genome-scale model and omics analysis of metabolic capacities of *Akkermansia muciniphila* reveal a preferential mucin-degrading lifestyle. *Appl. Environ. Microbiol.* **83**, e01014–17.
164. El-Semman, I.E., Karlsson, F.H., Shoae, S., Nookaew, I., Soliman, T.H., and Nielsen, J. (2014). Genome-scale metabolic reconstructions of *Bifidobacterium adolescentis* L2-32 and *Faecalibacterium prausnitzii* A2-165 and their interaction. *BMC Syst. Biol.* **8**, 41.
165. Heinken, A., Sahoo, S., Fleming, R.M.T., and Thiele, I. (2013). Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microb.* **4**, 28–40.
166. Liao, Y.-C., Huang, T.-W., Chen, F.-C., Charusanti, P., Hong, J.S.J., Chang, H.-Y., Tsai, S.-F., Palsson, B.O., and Hsiung, C.A. (2011). An experimentally validated genome-scale metabolic reconstruction of *Klebsiella pneumoniae* MGH 78578, iYL1228. *J. Bacteriol.* **193**, 1710–1717.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
<i>Leuconostoc mesenteroides</i> subsp. <i>mesenteroides</i> ATCC 8293	ATCC	ATCC 8293
<i>Lacticaseibacillus casei</i> ATCC 393	ATCC	ATCC 393
<i>Lactiplantibacillus plantarum</i> WCSF1	ATCC	ATCC BAA-793
<i>Ligilactobacillus salivarius</i> ATCC 11741	ATCC	ATCC 11741
<i>Limosilactobacillus fermentum</i> ATCC 14931	ATCC	ATCC 14931
<i>Lactococcus lactis</i> subsp. <i>cremoris</i> NZ9000	Boca Scientific	ELS09000-01
<i>Akkermansia muciniphila</i> ATCC BAA-835	Korean Collection for Type Cultures	KCTC 15667
<i>Bacteroides thetaiotaomicron</i> VPI-5482	Korean Collection for Type Cultures	KCTC 5723
<i>Escherichia coli</i> W3110	Korean Collection for Type Cultures	KCTC 2223
<i>Roseburia hominis</i> A2-183	Korean Collection for Type Cultures	KCTC 5845
Chemicals, peptides, and recombinant proteins		
Glucose (Gluc)	Sigma Aldrich	Cat#G7021
L-Histidine (His)	Sigma Aldrich	Cat#H6034
L-Isoleucine (Ile)	Sigma Aldrich	Cat#I7403
L-Leucine (Leu)	Sigma Aldrich	Cat#L8912
L-Methionine (Met)	Sigma Aldrich	Cat#M5308
L-Valine (Val)	Sigma Aldrich	Cat#V0513
L-Arginine (Arg)	Sigma Aldrich	Cat#A8094
KH2PO4	Sigma Aldrich	Cat#P5655
K2HPO4	Sigma Aldrich	Cat#P5655
L-Glutamic acid (Glu)	Sigma Aldrich	Cat#G8415
L-Phenylalanine (Phe)	Sigma Aldrich	Cat#P5482
L-Proline (Pro)	Sigma Aldrich	Cat#P5607
L-Asparagine (Asn)	Sigma Aldrich	Cat#A4159
L-Aspartic acid (Asp)	Sigma Aldrich	Cat#A7219
L-Glutamine (Glx)	Sigma Aldrich	Cat#G8540
L-Serine (Ser)	Sigma Aldrich	Cat#S4311
L-Threonine (Thr)	Sigma Aldrich	Cat#T8441
L-Cysteine (Cys) HCl	Sigma Aldrich	Cat#C7352
L-Alanine (Ala)	Sigma Aldrich	Cat#A7469
Glycine (Gly)	Sigma Aldrich	Cat#G8790
L-Lysine (Lys) HCl	Sigma Aldrich	Cat#L8662
L-Tryptophan (Trp)	Sigma Aldrich	Cat#T8941
L-Tyrosine (Tyr)	Sigma Aldrich	Cat#T8566
lipoic acid	Sigma Aldrich	Cat#T1395
Tween 80	Sigma Aldrich	Cat#P4780
Adenine	Sigma Aldrich	Cat#A2786
Guanine	Sigma Aldrich	Cat#G6779
Uracil	Sigma Aldrich	Cat#U1128
Xanthine	Sigma Aldrich	Cat#X3627
MOPS	Sigma Aldrich	Cat#M3183
Tricine	Sigma Aldrich	Cat#T5816

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CoCl ₂ .6 H ₂ O	Sigma Aldrich	Cat#C8661
EDTA	Sigma Aldrich	Cat#E6758
Sodium thioglycollate	Sigma Aldrich	Cat#T3758
(NH ₄) ₂ SO ₄	Sigma Aldrich	Cat#A2939
Biotin	Sigma Aldrich	Cat#B4639
Folic acid	Sigma Aldrich	Cat#F8758
Yeast nitrogen base without amino acids and ammonium sulfate	Sigma Aldrich	Cat#Y1251
Critical commercial assays		
RNAprotect	Qiagen	Cat#76104
RNaseZAP	Thermo Fisher Scientific	Cat#AM9780
RNeasy Mini kit	Qiagen	Cat#74106
Turbo DNA-free kit	Thermo Fisher Scientific	Cat#AM1907
Bacterial Ribo-Zero Magnetic kit	Illumina	Cat#MRZB12424
TruSeq Stranded mRNA Library Preparation kit	Illumina	Cat#20020595
SuperScript II reverse transcriptase	Thermo Fisher Scientific	Cat#18064022
QIAamp PowerFecal Pro DNA Kit	Qiagen	Cat#51804
Nextera XT DNA Library Prep Kit	Illumina	Cat#FC-131-1096
Nextera Index Kit	Illumina	Cat#FC-131-2001
Acetate Assay Kit (Colorimetric)	Abcam	Cat#ab204719
Bradford Protein Assay Kit	Bio-Rad	Cat#5000201
Deposited data		
Raw sequence reads (transcriptome)	This paper	NCBI BioProject: PRJNA574885
Raw sequence reads (microbiome)	This paper	NCBI BioProject: PRJNA882996
Experimental models: Organisms/strains		
Mouse: C57BL/6	OrientBio	C57BL/6NCrlOri
Software and algorithms		
MATLAB	MathWorks	N/A
R	N/A	www.r-project.org
GraphPad	N/A	GraphPad Software, Inc.
COBRA toolbox	Heirendt et al. ⁶³	https://opencobra.github.io/cobratoolbox/stable/
Growthcurver	Sprouffske and Wagner ¹⁰⁶	https://cran.r-project.org/web/packages/growthcurver/vignettes/Growthcurver-vignette.html
Orthofinder2	Emms and Kelly ¹⁰⁷	https://github.com/davidemms/OrthoFinder
MAFFT	Katoh and Standley ¹⁰⁸	https://mafft.cbrc.jp/alignment/software/
RAxML	Stamatakis ¹⁰⁹	https://cme.h-its.org/exelixis/web/software/raxml/
STAG	Emms et al. ¹¹⁰	https://github.com/davidemms/STAG
STRIDE	Emms and Kelly ¹¹¹	http://www.stevekellylab.com/stride
iTOL	Letunic and Bork ¹¹²	https://itol.embl.de/
eggNOG mapper	Huerta-Cepas et al. ¹¹³	http://eggnog-mapper.embl.de/
BAGEL	De Jong et al. ¹¹⁴	http://bagel4.molgenrug.nl/
FastQC	Andrews ¹¹⁵	https://github.com/s-andrews/FastQC
Trimmomatic v0.32	Bolger et al. ¹¹⁶	https://github.com/usadellab/Trimmomatic
STAR v 2.5.3a	Dobin et al. ¹¹⁷	https://github.com/alexdobin/STAR

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RSEM v1.3.0	Li and Dewey ¹¹⁸	https://deweylab.github.io/RSEM/
False discovery rate adjustment	Benjamini and Hochberg ¹¹⁹	https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
RUVseq	Peixoto et al. ¹²⁰	https://bioconductor.org/packages/release/bioc/html/RUVSeq.html
DEseq2	Love et al. ¹²¹	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
DAVID	Dennis et al. ¹²²	https://david.ncifcrf.gov/
DADA2	Callahan et al. ¹²³	https://benjineb.github.io/dada2/tutorial.html
HiMAP2	Segota and Long ¹²⁴	https://doi.org/10.1101/565572
gapFind	Kumar and Maranas ¹²⁵	https://doi.org/10.1371/journal.pcbi.1000308
dbCAN2	Zhang et al. ¹²⁶	https://github.com/linnabrown/run_dbcan
HMMER	Finn et al. ¹²⁷	http://hmmer.org
DIAMOND	Buchfink et al. ¹²⁸	https://github.com/bbuchfink/diamond
CAZy	Cantarel et al. ¹²⁹	http://www.cazy.org/
MEMOTE	Lieven et al. ¹³⁰	https://memote.io/
InParanoid	Sonnhammer and Östlund ¹³¹	https://inparanoid.sbc.su.se/cgi-bin/index.cgi
<hr/>		
Other		
Genome-scale metabolic models	This study	MODEL2210190001; MODEL2210190002; MODEL2210190003; MODEL2210190004; MODEL2210190005; MODEL2210190006; MODEL2210190007; MODEL2210190008; MODEL2210190009; MODEL2210190010; MODEL2210190011; MODEL2210190012
LAB Probiotic Evaluation Framework	This study	https://github.com/skku-pdse/LAB-probiotic-evaluation-framework

RESOURCE AVAILABILITY

Lead contact

Further information and requests for datasets and/or protocols may be directed to, and will be fulfilled by the Lead Contact Dong-Yup Lee (dongyuplee@skku.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The newly sequenced transcriptome and microbiome reads are available in NCBI SRA (<https://www.ncbi.nlm.nih.gov/sra>): PRJNA574885 and PRJNA882996.
- The FBAwMC in silico analysis codes used in this work is available online at <https://github.com/skku-pdse/LAB-probiotic-evaluation-framework>. All the LAB GEMs are deposited to BioModels¹ database and available under following accessions: *L. casei* subsp. *casei* ATCC393: MODEL2210190001, *L. casei* BL23: MODEL2210190002, *L. casei* LC5: MODEL2210190003, *L. fermentum* ATCC14931: MODEL2210190004, *L. plantarum* WCFS1: MODEL2210190005, *L. plantarum* ATCC8014: MODEL2210190006, *L. plantarum* JDM1: MODEL2210190007, *L. salivarius* ATCC11741: MODEL2210190008, *L. lactis* subsp. *cremoris* NZ9000: MODEL2210190009, *L. lactis* subsp. *lactis* IL1403: MODEL2210190010, *L. lactis* subsp. *cremoris* MG1363: MODEL2210190011, *L. mesenteroides* subsp. *mesenteroides* ATCC 8293: MODEL2210190012. The SBML files of all the LAB models are also available in GitHub repository: <https://github.com/skku-pdse/LAB-probiotic-evaluation-framework>.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

LAB strains, media and growth conditions

LAB strains, *Leuconostoc mesenteroides* subsp. *mesenteroides* ATCC 8293, *Lacticaseibacillus casei* ATCC 393 (formerly *Lactobacillus casei* subsp. *casei* ATCC 393), *Lactiplantibacillus plantarum* WCSF1 (formerly *Lactobacillus plantarum* WCSF1), *Ligilactobacillus salivarius* ATCC 11741 (formerly *Lactobacillus salivarius* ATCC 11741), *Limosilactobacillus fermentum* ATCC 11741 (formerly *Lactobacillus fermentum* ATCC 14931) were from ATCC, and *Lactococcus lactis* subsp. *cremoris* NZ9000 was from Boca Scientific. The lyophilized cultures were first revived by aseptically transferring the contents of the ampoule into 10mL each of MRS and GM17 media, equally. The revived cultures were then sub-cultured using the same media to ensure stability. A chemically defined medium capable of supporting all six LAB strains (LABDM) was developed by iteratively modifying components of a previously established CDM⁴² (see Table S2). Seed LAB cultures for fermentation experiments were prepared by transferring the 10% (v/v) cultures to LABDM and incubating them overnight at 30°C. Biological triplicates of the LAB were grown in 50 mL conical tissue culture tubes (Thermo Fisher Scientific) as static cultures with added reducing agent, sodium thioglycolate (1g/L), and no head-space to maintain anaerobic conditions. A separate culture tube was maintained for each time point, to avoid aeration during sampling.

Commensal culture with LAB supernatants

Commensal strains, *A. muciniphila* ATCC BAA-835, *B. thetaiotaomicron* VPI-5482, *E. coli* W3110 and *R. hominis* A2-183, were obtained from the Korean Collection for Type Cultures (KCTC, <https://www.kctc.kribb.re.kr>). LABDM supplemented with heme, N-acetylglucosamine was used to culture all four commensal strains. Media were purged with N₂ gas prior to use, ensuring that they are devoid of oxygen. LAB strains were first statically grown in Man Rogosa Sharpe (MRS) broth (BD, Sparks, MD, USA) overnight at 30°C. Simultaneously, commensal bacteria were inoculated to LABDM medium and allowed them to grow overnight at 37°C. Supernatants from the LAB culture 10% v/v was then transferred to freshly prepared LABDM in a culture tube. The overnight-grown commensal culture was seeded to this tube to a final OD600 of 0.1–0.2. All experiments were conducted in triplicates and were done in an anaerobic chamber. The culture tube was placed inside anaerobic jar and purged with N₂ for 10 min before including the GazPak EZ sachets (Becton-Dickinson). After sealing, the jar was placed in a 37°C incubator. Growth of the commensal bacteria was recorded by measuring OD after 12 h of incubation.

LAB culture for mice supplementation

All six LAB strains were first cultured in MRS agar (BD, Sparks, MD, USA). A single colony was obtained after streaking and incubating for 48 h at 37°C. Each colony was incubated with MRS broth at 37°C. *L. plantarum*, *L. salivarius*, *L. fermentum*, *L. casei* were incubated for 24 h, while *L. lactis* and *L. mesenteroides* were incubated for 48 h. Cultured strains were washed, centrifuged, and suspended in phosphate-buffered saline (PBS) in the concentration of 1 × 10⁹ CFU/100 µL.

Mouse study design

Six-week-old C57B/6 male mice supplied by OrientBio (Seoul, Korea) were housed in sterilized cages controlled at 22 ± 2°C, 45 ± 10% humidity, 12 h light/dark cycle and were freely fed with normal feed (4.5 Kcal %, Research Diet, Purina, Seongnam, Korea) and filtered water. Mice were acclimatized for one week in said conditions. Next, acclimatized mice were divided into 7 groups (n = 10), and fed with high-fat diet (60 Kcal % fat, Research Diet, New Brunswick, NJ, USA). 100 µL of LAB solution (one strain for each group, excluding control group) was orally administered every day for one week. Cecum tissue was collected one week after cessation of probiotic LAB administration and was used for microbiome analysis. Ethics approval for the mouse study was provided by the Institutional Animal Care and Use Committee (IACUC) of AtoGen Co., Ltd., registration number (ATG-IACUC-SKK-200713 and ATG-IACUC-SKK-200904).

METHOD DETAILS

Growth and biochemical analysis of culture supernatants

Cells and supernatants were harvested from the biological triplicate LAB cultures cultivated in LABDM for every one-hour time interval until the stationary phase for various growth and other biochemical assays. Optical density at 600nm (OD600) was measured using Shimadzu UV-1700 spectrophotometer. Dry cell weights (DCW) were obtained by centrifuging 2 mL of cell cultures, subsequently drying the pellet at 100°C overnight, and weighing them using a balance. Standard curves of OD600 vs gDCW were plotted to estimate the conversion factors. Glucose and L-lactate profiles at different time intervals spanning exponential phase were analyzed using YSI® biochemistry analyzer. Amino acid profiles were measured using Waters ACQUITY-UPLC system, AccQ•Tag™ Ultra Column (2.1 × 100 mm) and AccQ-Tag derivatization kit, following the manufacturer instructions. Acetate in cell culture supernatant was measured using Acetate Assay Kit (Colorimetric) (abcam®) based on the recommended procedures from manufacturer. Note that biological triplicates were employed for all sample measurements.

Cell volume and cell number measurements

LAB single cell volumes were estimated by approximating the shapes of LcLt and LeMt to spherical, and LbPt, LbCs, LbSv and LbFm to cylindrical dimensions. Average cell diameters and lengths of at least a hundred cells measured using an optical microscope at 100X magnification under oil immersion were used to estimate single cell volumes. Bacterial cultures were diluted 1000X before microscopic measurements to ensure proper dispersal of cells. Cell numbers in a given volume of the same diluted culture were enumerated using plating and automatic CFU counter (Scan 1200, Interscience, Saint Nom, France).

Determination of total protein content

Total protein content was measured using Bradford assay.¹³² Briefly, cell pellets obtained from 10 mL of exponentially growing cells were suspended in 1 mL ice-cold lysis buffer containing 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, and 100 μ L of lysozyme (10 mg/mL), and incubated on ice for 15 min. Next, the cells were sonicated at a frequency of 20 kHz for 10 \times 30 s with 1 min interval between each sonication cycle. Ten rounds of sonication resulted in “close to complete” protein release, after which the protein concentration saturated. Bradford protein assay kit (Bio-Rad®) with standard Bovine Serum Albumin (BSA) solution was used to estimate the protein concentration in 100 μ L of the sonicated, which were then converted to total protein content per gram DCW by accounting for dilution and OD600-to-gDCW conversion factors (Table S6).

Total RNA extraction and purification

To extract the total RNA, 10 mL aliquots were harvested from the biological triplicate LAB cultures cultivated in LABDM at time points that correspond to their respective mid-exponential phases. It was first mixed with 10 mL of Qiagen RNAProtect reagent and incubated for 10 min at ambient temperature to stabilize RNA. Bench and biosafety cabinet surfaces used for RNA work were decontaminated with RNaseZAP (Thermo Fisher Scientific) to minimize contact of samples with RNase. Half of the sample aliquotes (10 mL) containing RNAProtect reagent were centrifuged at 6000 \times g for 10 min. The remaining samples were stored at –80°C up to 4 weeks for any later use. Prior to RNA extraction, the cell pellets were treated with 1 mL ice-cold lysis buffer containing 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, and 100 μ L of lysozyme (10 mg/mL), all prepared using RNase-free water, and incubated for 30 min. The lysed solution was further subjected to mechanical cell disruption involving nuclease-free glass beads and thermomixer to ensure maximum release of RNA. Qiagen RNeasy Mini kit was used to extract total RNA from the cell pellets following manufacturer instructions. DNase treatment was performed using Turbo DNA-free kit (Thermo Fisher Scientific) to remove any genomic DNA contamination. Integrity and quality of the extracted RNA was initially checked by the presence of 16s and 23s rRNA in samples subjected to agarose gel electrophoresis and later confirmed using Agilent Bioanalyzer. Bacterial Ribo-Zero Magnetic kit (Illumina) was used to deplete rRNA and enrich mRNA in the samples, following manufacturer’s instructions. Agilent Bioanalyzer was used to ensure the depletion of rRNA from all samples before the preparation of cDNA libraries.

Preparation of cDNA libraries and RNA-sequencing

Fragmentation of rRNA depleted mRNA samples and subsequent preparation of cDNA libraries were performed using Illumina TruSeq Stranded mRNA Library Preparation kit (Low Sample Protocol), following manufacturer’s instructions. First cDNA strand synthesis in the reverse transcription polymerase chain reaction (RT-PCR) was performed using the reagents supplied by the manufacturer along with the SuperScript II reverse transcriptase (Thermo Fisher Scientific). Indexed cDNA libraries were pooled and single-end sequenced on Illumina HiSeq 2500 Rapid V2 platform to 51 bp read length.

Microbiome DNA extraction and sequencing library construction

DNA was extracted from the cecum of mice using QIAamp PowerFecal Pro DNA Kit (Qiagen, Hilden, Germany). Then, Nextera XT DNA Library Prep Kit and Nextera Index Kit (Illumina, San Diego, CA, USA) was used to amplify and index the V4 region of the 16S rRNA gene to create a paired end sequencing library. Subsequently, sequencing was performed using Illumina iSeq 100 (Illumina).

QUANTIFICATION AND STATISTICAL ANALYSIS**Statistical analysis of LAB growth characteristics**

LAB growth characteristics were evaluated on the basis of four parameters: (1) maximum growth rate that can occur if total population size is not limited, (2) doubling time, (3) biomass yield over glucose, and (4) L-lactate yield over glucose. Briefly, maximum growth rate is the maximum value of slope along the log transformed growth curve. Doubling time is the time required to reach half the maximum OD in the culture. Biomass yield over glucose and L-lactate yield over glucose are the ratio of LAB gDCW accumulated over the culture to that of g glucose consumed and g L-Lactate produced to that of g glucose consumed, respectively. Maximum growth rate and doubling time were calculated by fitting a logistic equation to the culture data using Growthcurver.¹⁰⁶

Comparative genomic analysis

Reconstruction of phylogenetic tree: To understand the evolutionary perspective of LAB strains examined in this study, phylogenetic tree was constructed based on a set of core proteins present in all species using Orthofinder2.¹⁰⁷ Note that we also included four

other LAB strains, *Lactobacillus reuteri* ATCC 53608, *Lactobacillus delbrueckii* subsp. *bulgaricus* ATCC 11842, *Lactobacillus acidophilus* NCFM, and *Bacillus subtilis* subsp. *subtilis* str. 168, for this particular analysis. Briefly, OrthoFinder2 first identified the single copy orthologous proteins across all species based on the orthogroups and performed multiple sequence alignments (MSAs) using the MAFFT,¹⁰⁸ with default parameters. Unrooted species phylogenetic trees were then inferred from concatenated MSA of single-copy genes using RAxML,¹⁰⁹ following the STAG (Species Tree inference from All Genes) approach.¹¹⁰ Subsequently this unrooted species tree is rooted using the STRIDE algorithm.¹¹¹ The inferred phylogenetic tree was visualized using Interactive Tree Of Life (iTOL).¹¹²

Identification and functional annotation of LAB core, shell and cloud genome: In order to identify the core, shell and cloud genome, we first established the orthogroups among all LAB six species using Orthofinder2.¹⁰⁷ Subsequently, we identified the orthogroups with at least a single gene copy in any one organism, and established it as the “core” gene families that are present in all species. The orthogroups which are exclusively present only in one species is considered as “cloud” genome. All remaining orthogroups were considered as “shell” genome. Functional enrichment of core, shell and cloud genome was performed using eggNOG mapper based on the NOG categories.¹¹³ Note that NOG categories are similar to well-known COG categories but it is identified in an unsupervised manner.

Genomic comparisons for non-metabolic probiotic determinants: BLASTp orthology search was used to mine the LAB genomes for identifying ORFs encoding proteins for resisting acid, oxidative and bile stresses, protein lysing proteases and anchor cells on host mucosal layer surfaces. To do so, we first surveyed for genes known to be associated with probiotic characteristics in various LAB. Appropriate cutoffs (Identity >50%, query coverage >70%, E-value < 1E-06)¹³³ were used to select orthologs. Potential bacteriocin operons in the LAB genome was mined using BAGEL.¹¹⁴

Transcriptome data analysis

Transcriptome alignment and gene expression quantification: First, the quality of RNA-seq reads were assessed with FastQC.¹¹⁵ The adapters and low-quality reads were subsequently trimmed using Trimmomatic v0.32¹¹⁹ and the trimmed reads were then aligned with the respective genome assemblies using STAR v 2.5.3a¹²⁰. Finally, RSEM v1.3.0¹²¹ was used to quantify expression levels in terms of counts, FPKM and TPM from the alignment files (Table S3).

Functional enrichment of gene expression categories: Initially, the genes are grouped into three groups based on TPM values: i) low expression range ($\log_2(\text{TPM}+1) < 2.5$), ii) medium expression range ($\log_2(\text{TPM}+1) > 2.5 \& < 10$) and iii) high expression range ($\log_2(\text{TPM}+1) > 10$). Subsequently, a p value is calculated to represent the significance of NOG category enrichment in different gene expression groups using the Fisher’s exact test by creating a 2-by-2 contingency table as follows: The first row contains the information of total genes in the pathway for which enrichment score is being calculated. First column of first row contains the total genes in the input which belong to a particular pathway while second column in the same row contains the remaining total number of genes in that pathway. Second row comprises the information on total genes in the input which does not belong to that pathway and the total number of genes in the entire genome which does not belong to that pathway in columns 1 and 2, respectively. The p values obtained from Fisher’s exact test is then adjusted for false-discovery rates based on previously proposed method.¹¹⁹

Comparison of gene expression in orthologous genes: The direct comparison of gene expression across LAB species in gene wise manner is not possible because their genome sizes vary as well as their genome functionalities are different. Therefore, in order to compare the gene expression levels of same gene functions across LAB, we first identified one-to-one orthologues as mentioned previously. We subsequently normalized the raw counts of these orthologues based on a set of housekeeping genes using RUV-seq.¹²⁰ The housekeeping genes were identified using the following criteria: i) genes which have expression over 95 percentile range in each LAB and ii) genes which are present in one-to-one orthologues list (Table S1). Orthologous gene expression profiles were then compared using PCA and Spearman correlation in a pairwise manner.

Identification of differentially expressed genes: DEseq2¹²¹ was used to identify differentially expressed one-to-one orthologous genes across the 6 LAB. Note that we used the gene lengths in each LAB as an additional normalization factor in this step since the gene length will be different in each LAB for the same gene. Those genes with FDR adjusted p values < 0.01 were classified as significantly differentially expressed genes.

Functional enrichment of differentially expressed genes: Differentially expressed genes were functionally enriched based on KEGG pathways and GO terms using Database for Visualization and Integrative Discovery (DAVID).¹²² Pathways and GO terms which have an FDR adjusted p value < 0.01 were only considered for further analysis.

Microbiome data analysis

First, the quality of microbiome sequencing (16S rRNA) reads were assessed with FastQC.¹¹⁵ The sequence data was then trimmed, denoised and merged through DADA2¹²³ (version 1.16) pipeline. Chimeric reads were also removed in using the same pipeline. The resulting denoised sequences or amplicon sequence variant (ASV) were annotated using a database constructed by the HiMAP2¹²⁴ pipeline, which provides high coverage annotation at the species level. DESeq2¹²¹ package 3 in R environment (<https://www.r-project.org/>) was then used to identify differentially abundant OSU ids in each probiotic administered group in comparison to the control group. To further confirm the species identity, we obtained the nucleotide sequences of OSUs annotated as our considered six LABs and did a BLASTn search against the NCBI rRNA_typestrains/16S_ribosomal_RNA database.

Reconstruction of LAB genome-scale models

We first built the GEMs of LbPt, LcLt and LeMt by expanding the previously published models which are manually curated.^{43–45} To do so, the network gaps in these GEMs were identified using gapFind algorithm,¹²⁵ followed by addition of new reactions based on bibliographic information for their resolution. Next, we included new genes, reactions and metabolites and updated the GPR for several existing reactions based on the latest genome annotations, following the standard procedures.⁴⁶ Additionally, we used the amino acid auxotrophy data available in literature^{58,134} to manually curate all three GEMs by reducing the inconsistencies in auxotrophies observed in experiments and during constraint-based simulations. We also performed a search for Carbohydrate-Active enZYmes (CAZymes) in the six LAB using dbCAN2¹²⁶ which uses HMMER,¹²⁷ DIAMOND¹²⁸ and CAZy¹²⁹ tools to search the queried genomes against the carbohydrate enzyme database. Significant hits (E-value <1e-15 and coverage >0.35 for HMMER + dbCAN2; E-value <1e-102 for DIAMOND + CAZy) with known EC number annotations were then queried against the genome of all 6 LAB using BLASTp algorithm to confirm their presence. Hits that passed the cut-off (Identity >50%, query coverage >30%, E-value <1E-06) were added to the corresponding LAB models (Table S5). Literature data on fermentable substrate phenotyping data^{135–137} was then used to test the agreement between model predictions and experiments, subsequently eliminating the inconsistencies through the addition of relevant metabolic and transport reactions.

To newly reconstruct the LbFm, LbSv and LbCs GEMs, initially the metabolic pathway information based on genome annotations of these LAB, was collected from KEGG¹³⁸ and MetaCyc¹³⁹ databases. The draft reaction networks were then assembled by combining information from these databases, along with the information based on gene orthologs of the existing LAB models (LbPt, LcLt and LeMt). Next, the biomass reaction, which is typically used as an objective function within the FBA framework to predict metabolic fluxes of exponentially growing cells,¹⁴⁰ was added to each newly developed model based on experimental and literature information (see Table S5). Note that since protein constitutes the major macromolecular component of biomass, we measured the total protein content of different LAB experimentally and used it for biomass reaction formulations (Table S5). Gaps in the these draft networks were then identified using gapFind algorithm¹²⁵ and subsequently filled by adding new reactions based on bibliographic information. We then formulated the gene-protein-reaction (GPR) relationships and rectified the model inconsistencies, including elemental, charge imbalances. We also used the amino acid auxotrophy data for LbCs⁵⁸ to manually curate the corresponding GEM to reduce the model predicted inconsistencies in auxotrophies.

In all six LAB GEMs, growth and non-growth associated ATP maintenance costs were either adopted from existing GEMs or estimated based on available culture data (Table S5). Finally, the reconstructed models' quality and consistency were evaluated using online tool MEMOTE.¹³⁰ The MEMOTE reports of all LAB models are provided in Data S1.

The strain specific models of LbCs, LbPt and LcLt were reconstructed by following a four-step approach. First, the one-to-one orthologues were identified between the base strain for which GEM is available and the strain for which GEM has to be reconstructed using InParanoid.¹³¹ Using the one-to-one orthologues, a draft GEM was then assembled for each new strain from the base strain by replacing the corresponding gene loci in the GPR, and removing the genes and associated reactions that are not present in the new strain. Next, one-to-many and many-to-one gene associations were manually curated and updated in the GEM. We again identified one-to-one orthologous genes between the new strain and all other LAB species whose GEM are available. Subsequently, we added new reactions that are present in this new strain but not in the base strain based on inParanoid orthology. Finally, this model was manually curated by adding and/or removing relevant genes, reactions and metabolites using the substrate phenotyping data of the new strain for which GEM is being reconstructed.

Auxotrophy-weighted transcriptome Z score metric

First, TPM values of RNA-seq data was mapped to each LAB GEM using the GPR mapping. Then, Z-scores across expression levels of all reactions in each model were calculated with the following equation:

$$Z = \frac{x - \mu}{\sigma}$$

where x = observed value, μ = mean of the sample, σ = standard deviation of the sample. Next, all reactions related to the biosynthetic genes of each vitamin were identified and their corresponding Z-scores were averaged across each LAB. We further refined the scores by multiplying the Z-scores with 0 of vitamins with incomplete biosynthesis pathways.

Constraint-based flux analysis with macromolecular crowding constraints

Constraint-based flux analysis with macromolecular crowding constraints⁴⁷ was used to analyze the metabolic phenotype of the LAB under various environmental conditions such as LABDM and different dietary regimes. The corresponding optimization problem can be mathematically represented as follows:

$$\max Z = \sum_j c_j f_j \quad (1)$$

$$\text{subject to } \sum_j S_{ij} f_j = 0 \forall \text{metabolite } i \quad (2)$$

$$f_j^{\min} \leq f_j \leq f_j^{\max} \forall \text{reaction } j \quad (3)$$

$$\sum_{j=1}^N a_j f_j \leq 1 \quad (4)$$

where Z is the cellular objective, c_j is the relative weights of each metabolic reaction to biomass formation. S_{ij} is the stoichiometric coefficient of metabolite i of reaction j ; f_j is the flux through the reaction j ; f_j^{\min} and f_j^{\max} are lower and upper bounds on the flux through the reaction j , respectively; and $a_j = C \cdot v_j / b_j$, is referred to as 'crowding coefficient', C is cytoplasmic density which requires to be measured experimentally, b_j is the kinetic constants which is based on enzyme turnover numbers⁴⁷ and v_j is calculated using the molar mass (M_j) and specific volume of the protein (enzyme) as follows:

$$v_j = M_j \cdot v_j^{\text{specific}} \quad (5)$$

Note that since the specific volume of the proteins (v_j^{specific}) is known to vary very minimally in different cells,¹⁴¹ we used the average value (0.73 mL/g) as reported earlier.⁴⁷

The additional constraint in FBAwMC, [Equation 4](#), can be simply implemented in the S matrix itself. To do so, consider that stoichiometric matrix S_{ij} and the flux vector f_j are represented as below:

$$S_{ij} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1j} \\ \vdots & \vdots & \ddots & \vdots \\ S_{i1} & S_{i2} & \cdots & S_{ij} \end{bmatrix}$$

$$f_j = \begin{bmatrix} f_1 \\ \vdots \\ f_j \end{bmatrix}$$

Now, the additional constraint in FBAwMC based on crowding coefficients can be conveniently included in the S matrix in the form of a dummy reaction involving a pseudo metabolite whose sink flux is allowed to vary between 0 and 1 as follows.

$$S_{(i+1)(j+1)} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1(j+1)} \\ a_{1...} & a_{2...} & \cdots & a_{(j+1)...} \\ S_{(i+1)1} & S_{(i+1)2} & \cdots & S_{(i+1)(j+1)} \end{bmatrix}$$

$$f_{j+1} = \begin{bmatrix} f_1 \\ f_{dr...} \\ f_{j+1} \end{bmatrix}$$

where, $a_1, a_2 \dots a_{(j+1)}$ are the coefficients of a single pseudo metabolite added to all reactions ranging from 1 (j+1). Note that, $a_1, a_2 \dots a_{(j+1)}$ are same as the crowding coefficients of reactions 1 (j+1). The crowding coefficients of the dummy reaction, biomass objective, non-gene associated reactions and exchange reactions are set to zero. f_{dr} represents the flux of the dummy reaction and is allowed to vary in the range of 0 and 1, satisfying enzyme capacity constraint.

In order to simulate the physiology of exponentially growing LAB in LABDM, the biomass reaction was maximized while simultaneously constraining the uptake/secretion rates of glucose and all 20 amino acids at experimentally measured values ([Table S4](#)). Note that we used 5000 different permutations in each simulation to account for reactions with unknown C values and obtained 5000 flux solutions. The geometric mean of the flux solutions surrounding the maxima of the resultant lognormal distributions accurately represent the average cellular state,¹⁴² and were then used to assess the phenotype (growth rate/byproduct secretion) of the LAB. All simulations were implemented using COBRA toolbox¹⁴³ and Gurobi7 (<http://www.gurobi.com>) optimization solver.

Estimation of cell crowding coefficients (a)

The estimation of crowding coefficients required two kinds of experimental data: enzyme kinetic parameters, i.e. turnover numbers, K_{cat} (s^{-1}) and cytoplasmic densities.⁴⁷ These are used in a specific manner to estimate the macromolecular crowding coefficients as explained in previous section. The enzyme turnover numbers were obtained from BRENDA database.¹⁴⁴ It should be noted that we used the turnover numbers specific to LAB whenever it is available. However, if it is not measured in LAB for any of the enzymes, then,

the maximum reported value from any organism was used to avoid incorporation of smaller turnover numbers (higher crowding effects), which may limit the overall flux of the corresponding reactions. Cytoplasmic density of each LAB was estimated using the previously measured DCW values, volume occupied by a single cell and total number of cells in a given culture volume. Note that since cytoplasmic density is defined as the ratio of cell mass to cell volume, it was thus calculated after accounting for the necessary dilution factors.

Simulation of LAB fermentation profiles as a function of oxygen

The fermentation profiles of LAB as a function of oxygen was simulated by implementing FBAwMC while simultaneously constraining glucose uptake at 10 mmol/gDCW/hr, amino acids at 1 mmol/gDCW/hr. Note that the inorganic compounds were allowed to exchange freely. Oxygen uptake was gradually varied from 0 to 10 mmol/gDCW/hr in a step size of 0.1 and FBAwMC was implemented to obtain 5000 flux solutions for each step. The exchange fluxes of main fermentation products including acetate, lactate, ethanol, diacetyl and propionate at each step were collected, and their average values were used to plot the fermentation profiles. A smoothing spline function was used to visualize the fermentation trends in GraphPad Prism (version 9.3.1; www.graphpad.com).

Evaluation of positive and negative epistasis in LAB GEMs

Epistasis interactions in the LAB metabolic networks were evaluated based on multiplicative formulation proposed earlier.¹⁴⁵ Mathematically, the epistatic effect (ϵ) in terms of fitness, W (e.g. *in silico* growth rate of LAB) can be represented as follows.

$$\epsilon = W_{xy} - W_x \cdot W_y$$

where W_{xy} is the fitness after double deletion of genes/reactions x and y , W_x and W_y are fitness values after the corresponding single deletions.

Here, we slightly modified the formulation in order to avoid erroneous epistatic effects for fitness values less than unity.

$$\epsilon = W_{xy} - \min(W_x, W_y)$$

The epistasis between gene/reaction x and y is said to positive/buffering if ϵ is positive (e.g. $\epsilon > 0.0001$) and negative/aggravating if ϵ is negative (e.g. $\epsilon < -0.0001$). Otherwise ($\epsilon \approx 0$), x and y are likely to show no epistasis between each other. Positive epistatic reaction pairs were then functionally enriched based on metabolic pathways described in each GEMs using the Fisher's exact test by creating a 2-by-2 contingency table as mentioned earlier in the "Functional enrichment of gene expression categories" section.

In silico analysis of LAB phenotypes in various diets

We examined LAB phenotype in 11 different diets whose compositional information is available in VMH database.¹⁴⁶ Composition of 11 diet conditions including vegetarian, vegan, EU average, Mediterranean, DACH, high protein, gluten free, high fiber, diabetes, high fat-low carb, and high fat-high carb were obtained from Virtual Metabolic Human database (<https://www.vmh.life/#nutrition>) and utilized in the simulations to evaluate probiotic capabilities. These dietary conditions were formulated based on established diet surveys and include amounts of fatty acids, amino acids, carbohydrates, dietary fibers, vitamins, minerals and trace elements of one-day consumption.¹⁴⁶ The defined diets are represented in the form of exchange reactions which can be readily constrained within the FBA framework.⁶³ Each diet is distinct in its composition, for example the vegetarian diets have dairy and egg products consumption in addition to fruits and vegetables,^{147,148} while a vegan diet is a plant based diet which does not include any animal derived product.¹⁴⁸ Similarly, the high fiber diet has a high composition of fiber but also contains animal derived products.¹⁴⁹ On the other hand, the high fat, low carb diet is similar to a ketogenic diet which contains high levels of lipid and low levels of carbohydrate,¹⁵⁰ whereas the high fat, high carb diet is high in calories as it contains low amount of dietary fibers but high amounts of simple sugars, saturated fatty acids, and cholesterol. In addition, diets also differ between countries or region, for example the EU diet designed based on an Austrian Survey of 1002 people has carbohydrates as the main energy source,¹⁵¹ while the Mediterranean diet based on 2 studies is characterized by abundance of fresh plant foods, minimally processed food, and olive oil and low amounts of meat.^{152,153} Furthermore, The DACH diet has well-balanced amounts of carbohydrates, protein and lipids to achieve healthy nutritional status, which was established by the society for Nutrition in Switzerland, Germany and Austria.¹⁴⁸ Some other diets represent the specific individual requirements, for example strength athletes or overweight people requires a high protein diet,⁶⁷ and the gluten free diet is based on a 24 h weight protocol from a subject who suffers from gluten intolerance. In addition, the diet for type 2 diabetes patients is low calorie and has high amounts of vitamins and minerals.¹⁵⁴

To simulate the cellular growth of LAB in various diets, the exchange reactions were constrained at corresponding uptake rates (Table S4). We used 5000 different permutations in each diet simulation to account for reactions with unknown C values and obtained 5000 flux solutions for each diet. The geometric mean of the flux solutions surrounding the maxima of the resultant lognormal distributions were then used to report the phenotype in that diet as it has been done in LABDM simulations. Further to understand, how each diet influences the LAB growth phenotype in a different manner, we extracted the reduced costs for the optimal solutions obtained surrounding the maxima of the resultant lognormal distributions and the average values were used for further analysis. Reduced cost of each reaction is the measure of how the objective will be reduced (attenuated) if the corresponding reaction is forced to carry a unit flux.¹⁵⁵ The mean flux and reduced costs of all exchange reaction in each diet is provided in Table S8.

To compute the theoretical yield of the postbiotic metabolites, pseudo-reactions representing their secretion was added and iteratively used as the objective function. Biomass was constrained to 50% of the optimum value. FBAwMC was implemented using COBRA toolbox⁶³ functionalities and the objective was maximized to obtain the *in silico* postbiotic yields of each LAB while constraining the exchange reactions at corresponding uptake rates provided in VMH database for each diet. We used 5000 different permutations for each FBAwMC simulation to account for reactions with unknown C values. The geometric mean of the resulting 5000 flux solutions surrounding the maxima of the lognormal distributions accurately represents the average cellular state,¹⁴² and was then used to calculate the postbiotic yield. Finally, the yield of each postbiotic compound was normalized to the maximum value across different diets and LAB.

In silico examination of LAB ability to cross-feed common gut microbes

LAB ability to cross-feed the common gut microbes were evaluated *in silico* based on the following approach. Briefly, the growth phenotype of each LAB growing in MRS media under anaerobic conditions was first simulated using FBAwMC approach described earlier (Table S4). Then, the fermentative products and other byproducts synthesized by LAB were added to the LABDM as additional metabolites available in the media and the growth phenotype of gut microbes under anaerobic conditions was simulated using the corresponding GEMs of commensal or pathogen microbes which have been previously published.^{156–166} The growth rate of each gut microbe obtained in the LABDM with supplements from LAB was then normalized with the value obtained from the simulations without the LAB supplements.