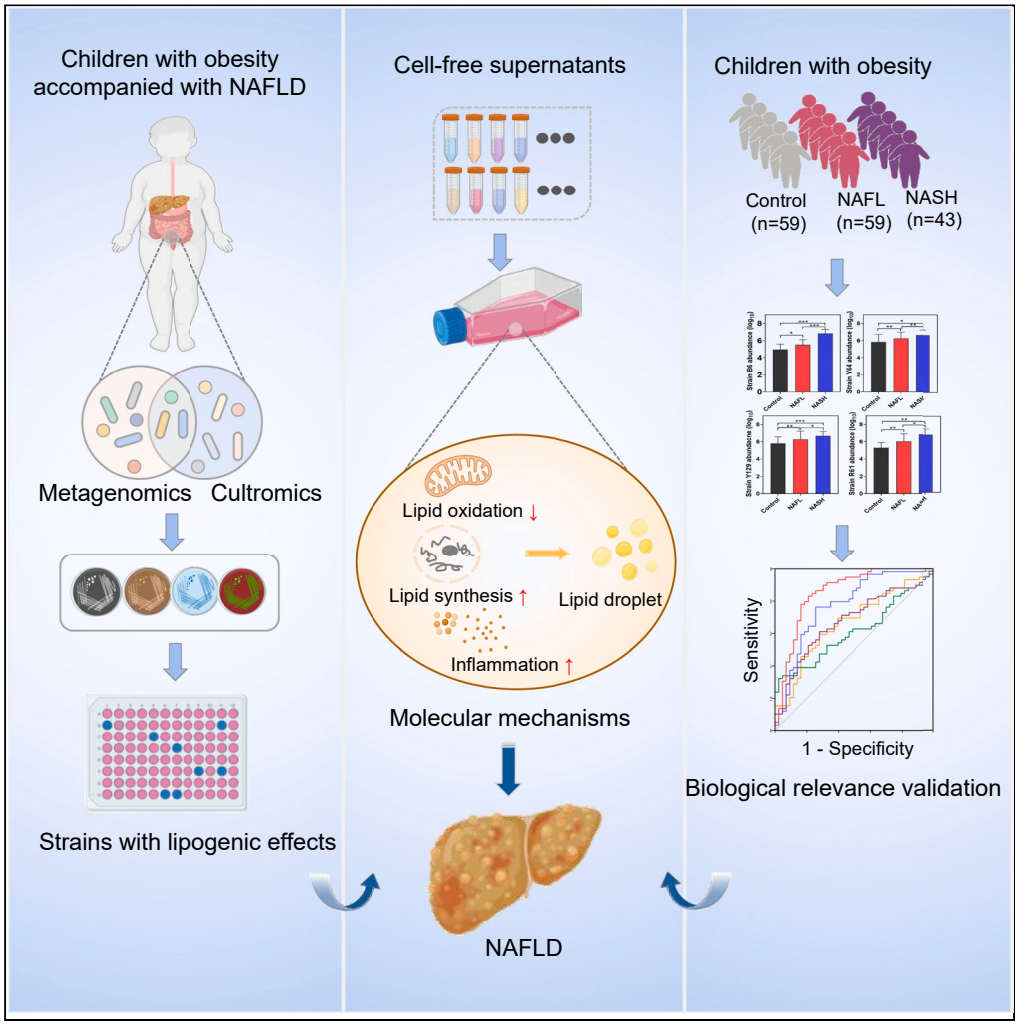
iScience

CellPress
OPEN ACCESS

Article

Identification of commensal gut bacterial strains with lipogenic effects contributing to NAFLD in children



Jia Wei, Jiayou Luo, Fei Yang, Wen Dai, Xiongfeng Pan, Miyang Luo

miyangluo@csu.edu.cn

Highlights

Dominant gut strains in children with obesity and NAFLD were identified

Four new isolates with lipogenic effects were discovered

Mechanisms of the four specific strains in NAFLD development were first revealed

A discriminative model for NAFLD based on these specific strains was established

Wei et al., iScience 27, 108861
February 16, 2024
https://doi.org/10.1016/j.isci.2024.108861

iScience

CellPress
OPEN ACCESS

Article

Identification of commensal gut bacterial strains with lipogenic effects contributing to NAFLD in children

Jia Wei,[1] Jiayou Luo,[1] Fei Yang,[2] Wen Dai,[1] Xiongfeng Pan,[1] and Miyang Luo[1,3,*]

SUMMARY

Gut microbiota is known to have a significant impact on nonalcoholic fatty liver disease (NAFLD), particularly in children with obesity. However, the specific functions of microbiota at the strain level in this population have not been fully elucidated. In this study, we successfully isolated and identified several commensal gut bacterial strains that were dominant in children with obesity and NAFLD. Among these, four novel isolates were found to have significant lipogenic effects *in vitro*. These strains exhibited a potential link to hepatocyte steatosis by regulating the expression of genes involved in lipid metabolism and inflammation. Moreover, a larger cohort analysis confirmed that these identified bacterial strains were enriched in the NAFLD group. The integrated analysis of these strains effectively distinguished NASH from NAFL. These four strains might serve as potential biomarkers in children with NAFLD. These findings provided new insights into the exploration of therapeutic targets for NAFLD.

INTRODUCTION

Nonalcoholic fatty liver disease (NAFLD) has been declared as the primary cause of chronic liver disease worldwide,[1] and more than 30% of children with obesity are suffering from NAFLD,[2] representing a significant health burden to society.[3,4] The pathogenesis of NAFLD in children with obesity remains unclear, and there is no clarity on pharmacotherapy for children with NAFLD.[5,6] Therefore, it is essential to elucidate the underlying mechanisms of NAFLD progression and to develop novel noninvasive biomarkers and therapeutic methods.

The gut microbiota is a diverse ecology that spans the length of the gastrointestinal tract and includes a variety of microorganisms such as bacteria, archaea, bacteriophages, eukaryotic viruses, and fungi. Approximately 100 trillion bacteria comprise over 99% of the gut microbiota.[7] It is now widely acknowledged that the gut microbiota influences physiological processes, including aging, digestion, absorption, metabolism, and immune system function.[8,9] Furthermore, this intricate community of microorganisms significantly influences the pathophysiology of various diseases, damaging the liver, gastrointestinal system, respiratory system, cardiovascular health, endocrine functions, and more.[10,11]

The gut-liver axis has become a pivotal transducer for NAFLD over the last ten years.[12,13] High throughput sequencing analysis (e.g., metagenomic analysis), a culture-independent technology, has greatly clarified the relationship between gut microbiota and NAFLD at the population level,[14,15] and gut microbiota has emerged as a potential source of noninvasive biomarkers for NAFLD, especially in high-risk populations.[16,17] For example, enrichment of genera such as *Eubacterium*, *Bacteroides*, *Escherichia*, *Dorea*, *Anaerococcus*, *Peptoniphilus*, and *Novosphingobium* have been correlated with NAFLD in children.[18–21] However, several factors continue to impact the accuracy of sequencing technology, including different DNA extraction techniques, variations in primer usage, sequencing depth, and analysis methods, and it is challenging to assign the sequencing results to specific live purified bacteria.[22,23] Recent studies have used microbial culturomics, a culturing approach that uses multiple culture settings and sequencing technologies for bacterial identification.[24–26] The cultures obtained from the samples represent the "alive" disease-associated bacteria, while even the most advanced sequencing techniques are unable to differentiate between viable and "dead" bacteria.[27–29] Indeed, the combination of culture techniques and 16S ribosomal RNA (16S rRNA) gene amplification for bacterial identification may generate a more comprehensive profile of the microbial community. This also provides opportunities to study the mechanisms of certain microbial strains.[24,30]

Several studies have identified specific strains that probably contributed to metabolic dysfunctions or hepatic steatosis. The publicly available strain of *Bilophila wadsworthia* ATCC 49260 and *Clostridium ramosum* DSM 1402 may exacerbate obesity or metabolic impairments.[31,32] *Klebsiella pneumonia* strains W14 and TH1 identified in adults with auto-brewing syndrome likely induced hepatic steatosis by producing ethanol.[33] Generally, gut microbiota and derived metabolites influence the hepatic lipid metabolism and inflammatory response, which promote lipid degeneration in the liver and drive the pathogenesis of NAFLD.[34,35] In addition, strains from different sources may result in

[1]Hunan Provincial Key Laboratory of Clinical Epidemiology, Xiangya School of Public Health, Central South University, Changsha 410078, Hunan, China
[2]Hunan Province Key Laboratory of Typical Environmental Pollution and Health Hazards, School of Public Health, University of South China, Hengyang 421001, Hunan, China
[3]Lead contact
*Correspondence: miyangluo@csu.edu.cn
https://doi.org/10.1016/j.isci.2024.108861

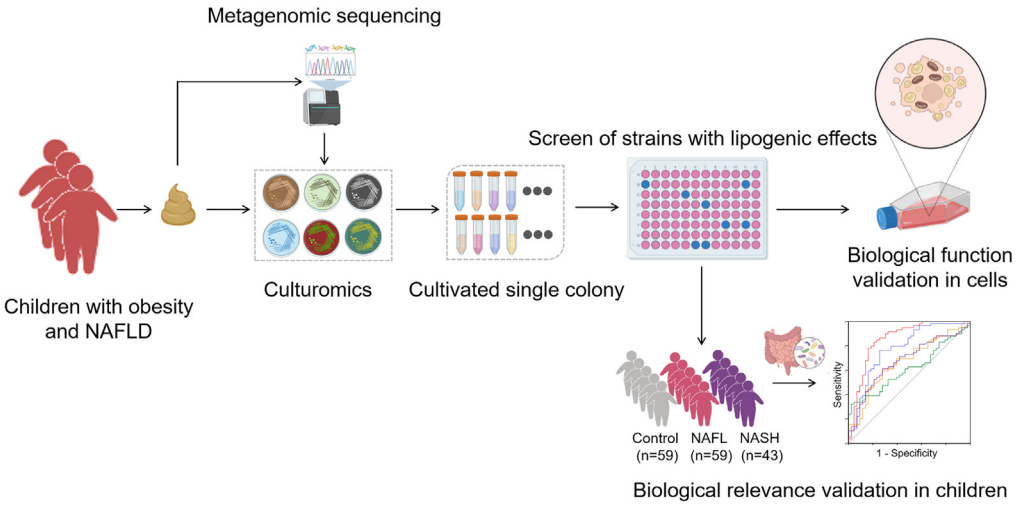


Figure 1. Overview of the workflow

inconsistent outcomes, and their diagnostic efficacy in NAFLD remains unknown. While most studies have focused on the association of gut microbiota and NAFLD in adults, there has been little research in children. Moreover, the vital pathogenic strains associated with NAFLD in children with obesity have not yet been identified and cultured. Therefore, it is necessary to isolate specific gut bacterial strains and explore their lipogenic effects and mechanisms on the development of NAFLD in children.

In this study, we aimed to isolate and characterize the dominant gut bacterial strains from children with obesity and NAFLD using metagenomic and culturomic techniques. The commensal bacterial strains were extensively screened for lipogenic effects, and their potential mechanisms contributing to NAFLD *in vitro* were investigated. Additionally, the associations of the identified bacterial strains and NAFLD in children with obesity were validated in a larger cohort, and a discriminative model based on these identified strains was established. This study may help reveal the pathogenic microbial profiles at the strain level and identify novel microbial biomarkers for NAFLD in children. The workflow is shown in Figure 1.

RESULTS

Metagenomic-based taxonomic and functional analysis

Ten children with NAFLD were included in the first cohort, including four children with simple steatosis (NAFL) and six children with steatohepatitis (NASH). The characteristics of these typical patients are presented in Table S1 (see Table S1). All of the selected children were male, aged 8–13 years, obese, and present with hepatic steatosis. Children with NASH exhibited concentrations of alanine aminotransferase (ALT), aspartate aminotransferase (AST), and triglycerides (TG) that exceeded the reference ranges. In contrast, most biochemical indices in children with NAFL were within the normal range, except for systolic blood pressure (SBP). Metagenomic analysis was performed to describe the gut microbial composition of these children. The reads number of individual samples ranged from 37,136,586 to 53,718,540. The rarefaction and species accumulation curves suggested that the estimated species richness and sample size were approaching saturation (see Figure S1). The composition of the gut microbiota in each child is shown in Figure 2 (Figure 2). At the genus level, the relative abundance of the top 20 taxonomic profiles was *Klebsiellas*, *Streptococcus*, *Enterobacter*, *Clostridium*, *Blautia*, *Bifidobacterium*, *Dorea*, and *Bacteroides*, among others (Figure 2A). Moreover, the relative abundance of the top 20 species was revealed, with *Megamonas funiformis*, *Bifidobacterium adolescentis*, *Klebsiella pneumoniae*, *Escherichia coli*, *Collinsella aerofaciens*, and *Prevotella copri* as the dominant species (Figure 2B). In addition, we explored the effect sizes (Cliff's delta) for individual features at the genus level, and the results showed that *Eubacterium*, *Enterococcus*, *Cronobacter*, *Bifidobacterium*, *Megamonas*, *Klebsiella*, *Collinsella* were enriched in the NASH group, while the *Bacteroides*, *Enterobacter*, *Clostridium*, *Dorea*, *Subdoligranulum*, *Escherichia*, *Faecalibacterium*, and *Blautia* were enriched in the NAFL group (Figure 2C). Moreover, the effect sizes (Cliff's delta) for individual features at the species level suggested that the *Bifidobacterium bifidum*, *Collinsella aerofaciens*, *Bifidobacterium longum*, *Dialister invisus*, *Klebsiella quasipneumoniae*, *Streptococcus salivarius*, *Enterococcus faecium*, were enriched in the NASH group. In contrast, *Cronobacter sakazakii*, *Blautia obeum*, *E. coli*, and *Dorea longicatena* were enriched in the NAFL group (Figure 2D). In addition, the KEGG pathway analysis of microbial genes revealed that the bacterial carbon metabolism, amino acid metabolism, membrane transport, and lipid metabolism were highly enriched (see Figure S2). Therefore, the gut microbiota might be active in energy absorption, fatty acid transport, and metabolism, all of which might affect the metabolism in the human body.

Culturomic-based bacterial analysis

To obtain gut bacterial strains for further investigation, we performed large-scale microbial cultivation of fecal samples from ten children with obesity and NAFLD. We grew bacteria from fecal samples of children with NAFLD on different solid media under anaerobic conditions. These

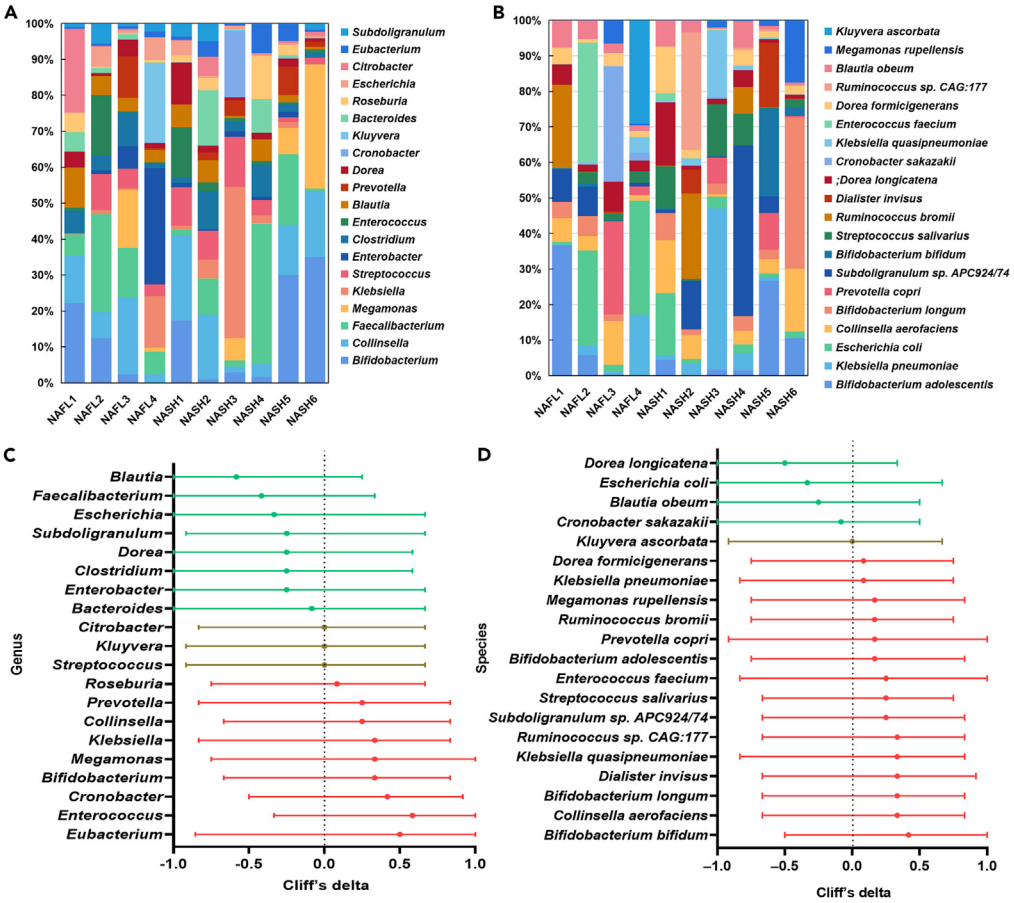


Figure 2. Metagenomic analysis of the fecal microbiota in children with obesity and NAFLD

(A) Top 20 genera identified through culture-independent sequencing.

(B) Top 20 species identified through culture-independent sequencing.

(C) Effect size of the genera estimated by Cliff's delta analysis.

(D) Effect size of species estimated by Cliff's delta analysis.

media included a wide range of substrates with good recovery capability for the numerous anaerobic gut bacteria. Every single colony with a distinctive morphology was identified according to the 16S rRNA sequences. We isolated 409 viable bacterial strains belonging to 27 genera, and analyzed the percentage of each genus in fecal samples (Figure 3A). The proportion of isolated *Bacillus, Clostridium, Dialister, Enterobacter, Enterococcus, Klebsiella,* and *Phascolarctobacterium* was higher in the NASH group, while the proportion of isolated *Bacteroides, Citrobacter, Streptococcus, Weissella* was higher in the NAFL group (Figure 3B). Furthermore, a maximum likelihood phylogenetic tree was constructed based on these commensal bacterial strains' 16S rRNA gene alignment (Figure 3C). In addition to a few unclassified strains, most bacterial strains belonged to *Streptococcus*, *Bacillus*, and *Escherichia genera*. Notably, the genus *Sutterella*, classified as *Proteobacteria* phylum, was successfully isolated and cultured in our work, although the metagenomic sequencing did not detect it. Moreover, we observed the same dominant genera, such as *Klebsiella*, *Clostridium*, *Enterococcus*, and *Bacteroides*, through culturomic and metagenomic-based analyses (Figures 2C and 3B). It is likely that the enrichment of these opportunistic pathogens probably caused gut dysbiosis and intestinal damage. These genera might represent the main risk factors through the gut-liver axis for NAFLD development.

Lipogenic effects of the isolated strains

We assessed the lipogenic effects of the isolated strains using HepG2 cell lines. Based on the dose-response experiment, we selected the concentration of cell-free supernatants from cultured bacterial strains with the maximal non-suppressive effects for further experiments. After screening the 409 isolated strains, we identified only four strains (named B6, R61, Y64, and Y129) that exhibited effects on lipid metabolism. When compared to the negative control, these four bacterial strains induced lipid droplets in HepG2 cells (Figure 4A). The relative area of lipids in the images of hepatocytes is shown in Figure S3 (see Figure S3). Among these strains, the B6 treatment group had a much higher level of lipid accumulation, followed by the R61 group (Figure 4B). Additionally, we found that the lipogenic effects of these strains were lower

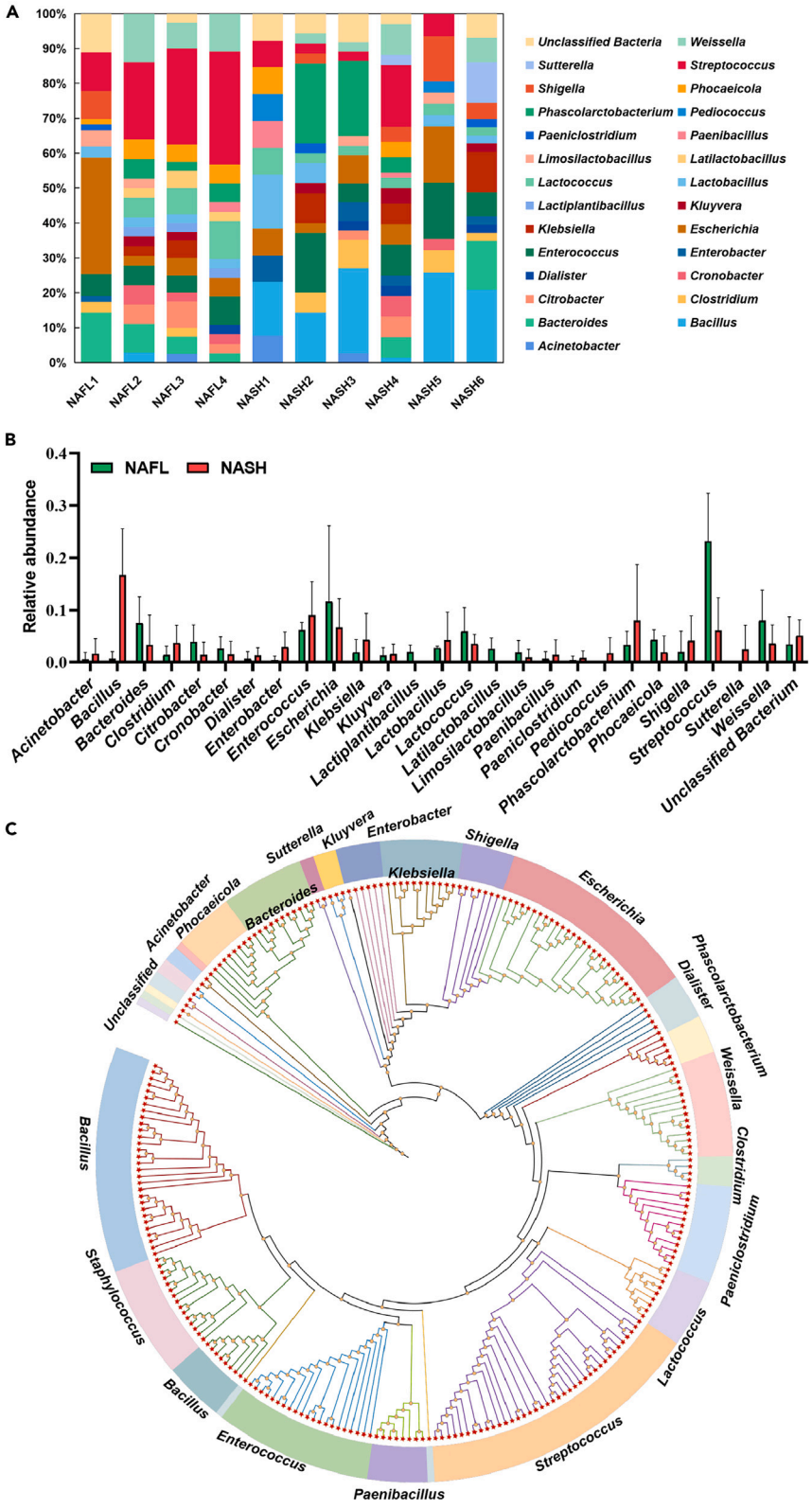

A
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
NAFL1
NAFL2
NAFL3
NAFL4
NASH1
NASH2
NASH3
NASH4
NASH5
NASH6
Unclassified Bacteria
Sutterella
Shigella
Phascolarctobacterium
Paeniclostridium
Limosilactobacillus
Lactococcus
Lactiplantibacillus
Klebsiella
Enterococcus
Dialister
Citrobacter
Bacteroides
Acinetobacter
Weissella
Streptococcus
Phocaeicola
Pediococcus
Paenibacillus
Latilactobacillus
Lactobacillus
Kluyvera
Escherichia
Enterobacter
Cronobacter
Clostridium
Bacillus
B
NAFL
NASH
Relative abundance
0.4
0.3
0.2
0.1
0.0
Acinetobacter
Bacillus
Bacteroides
Clostridium
Citrobacter
Cronobacter
Dialister
Enterobacter
Enterococcus
Escherichia
Klebsiella
Kluyvera
Lactiplantibacillus
Lactobacillus
Lactococcus
Latilactobacillus
Limosilactobacillus
Paenibacillus
Paeniclostridium
Pediococcus
Phascolarctobacterium
Phocaeicola
Shigella
Streptococcus
Sutterella
Weissella
Unclassified Bacterium
C
Sutterella
Kluyvera
Enterobacter
Shigella
Klebsiella
Escherichia
Phascolarctobacterium
Dialister
Weissella
Clostridium
Paeniclostridium
Lactococcus
Streptococcus
Paenibacillus
Enterococcus
Bacillus
Staphylococcus
Bacillus
Unclassified
Acinetobacter
Phocaeicola
Bacteroides

Figure 3. Culturomic analysis of fecal microbiota in children with NAFLD

(A) Relative abundance of bacteria isolated from children with NAFLD.

(B) Proportion of isolates at the genus level in the NAFL and NASH groups.

(C) Phylogenetic tree analyzing all isolated bacterial strains.

than those of oleic acid (OA) and phenylacetic acid (PAA), which served as positive controls. The morphology of the lipid droplets induced by these strains more closely resembled that of PAA. Furthermore, TG and total cholesterol (TC) levels in the bacterial cell-free supernatants treatment group were significantly higher than those in the corresponding control group. In particular, the B6 and R61 treatment groups exhibited higher levels of TG and TC than other groups, which was consistent with the results of Oil red O staining (Figure 4C).

Effects of the four specific bacterial strains on metabolism

To explore the mechanisms of the lipogenic strains, we investigated the expression of lipid synthesis genes (*SREBP1*, *SCD*, *ACC1*, and *FASN*), lipid oxidation genes (*ACOX1*, *CPT1A*, and *PPAR-α*), and inflammatory genes (*TNF-α*, *IL-6*) in HepG2 cells after stimulation with bacterial cell-free supernatants. Overall, the expression of these genes showed a dose-dependent manner. The expression of lipid synthesis and inflammatory genes increased with the concentrations of bacterial strains (Figure 5). More specifically, strains B6 and Y64 appeared to have a great ability to activate the expression of lipid synthesis genes and suppress the expression of lipid oxidation genes (Figures 5A and 5B). Most of the bacterial strains inhibited the expression of lipid oxidation genes, except for strain Y129. When the HepG2 cells were treated with strain Y129, the expression of some lipid oxidation genes (*CPT1A* and *PPAR-α*) and inflammatory genes (*IL-6*) were up-regulated, while the expression of *ACOX1* and *TNF-α* was down-regulated (Figure 5C). In addition, the treatment with strain B6, Y64, and R61 cell-free supernatants increased the expression of *TNF-α* and *IL-6* (Figures 5A, 5B, and 5D).

Phylogenetic analysis of the four bacterial strains with lipogenic effects

The four specific bacterial strains with lipogenic effects were successfully identified through phylogenetic analysis based on their 16S rRNA sequences. Strain B6 was classified as *Enterococcus* genus and closely resembled *Enterococcus faecium* HBUAS66133. Strain R61 was identified as the *Streptococcus* genus, with a sequence similar to *Streptococcus oralis* subsp. Tigurinus 7117668. Strain Y64 was classified as the *Escherichia* genus, similar to *Escherichia coli* ATCC 8739. Additionally, strain Y129 was affiliated with the *Klebsiella* genus and closely resembled *K. pneumoniae* F42 (Figure 6). Therefore, the four bacterial strains with lipogenic effects were named *Enterococcus* sp. B6 (Figure 6A), *Streptococcus* sp. R61 (Figure 6B), *Escherichia* sp. Y64 (Figure 6C), and *Klebsiella* sp. Y129 (Figure 6D), respectively. The *Enterococcus* sp. B6 and *Streptococcus* sp. R61 were gram-positive bacteria, while *Klebsiella* sp. Y129 and *Escherichia* sp. Y64 were gram-negative bacteria.

Associations of the four specific bacterial strains and children with NAFLD

We further investigated the biological relevance of these specific strains and childhood NAFLD using a second validation cohort to validate our findings. A total of 161 children with obesity were included, comprising 59 children with simple obesity (termed as Control), 59 children with obesity and NAFL (termed as NAFL), and 43 children with obesity and NASH (termed as NASH). The characteristics of these participants are shown in Table S2 (see Table S2). The levels of NAFLD-associated indicators, including ALT, AST, AST/ALT, TG, TC, high-density lipoprotein cholesterol (HDL-c), low-density lipoprotein cholesterol (LDL-c), C-peptide, and insulin, were statistically significant among the groups, while age, gender, BMI, body fat, SBP, and diastolic blood pressure (DBP) did not significantly differ among the groups. Moreover, the abundance of fecal *Enterococcus* sp. B6, *Streptococcus* sp. R61, *Escherichia* sp. Y64, and *Klebsiella* sp. Y129 was significantly higher in the NAFL and NASH groups than in the control group ($p < 0.05$, Figure 7A). Correlation analysis was performed between the abundance of gut bacterial strains and typical clinical indicators (Figure 7B). The abundance of *Enterococcus* sp. B6, *Streptococcus* sp. R61, *Escherichia* sp. Y64, and *Klebsiella* sp. Y129 was positively correlated with serum ALT, AST, and TG levels, while being negatively correlated with AST/ALT ($p < 0.05$). Moreover, the abundance of *Enterococcus* sp. B6 was positively correlated with serum C-peptide and insulin levels ($p < 0.05$). The abundance of *Klebsiella* sp. Y129 was positively correlated with glycosylated hemoglobin (HbA1c) and LDL-c levels ($p < 0.05$, Figure 7B).

Performance of the four specific bacterial strains in distinguishing NAFLD and controls

The diagnostic performance of the specific bacterial strains was confirmed in the second cohort. We assessed the diagnostic efficiency of fecal *Enterococcus* sp. B6, *Streptococcus* sp. R61, *Escherichia* sp. Y64, and *Klebsiella* sp. Y129 to distinguish between NAFL and controls, NASH and controls, as well as NAFL and NASH. Prediction models were constructed using the abundance of the four bacterial strains with lipogenic effects, and the area under curve (AUC) values for NAFL vs. controls, NASH vs. controls, and NAFL vs. NASH were 0.728 (95%CI 0.637–0.819), 0.950 (95%CI 0.909–0.991), and 0.856 (95%CI 0.780–0.940) respectively. As shown in Figure 8, the sensitivity and accuracy of the integrated analysis of the four bacterial strains were higher than those of the single strain signature (Figure 8). Moreover, the predictive power of the four strains for NAFL and NASH diagnosis exhibited a sensitivity and specificity of 0.831 and 0.791, respectively.

DISCUSSION

This study explored the dominant commensal gut microbiota in children with obesity and NAFLD by using the culture-dependent and culture-independent techniques. A series of commensal gut bacterial strains were successfully isolated, and four novel strains with lipogenic effects

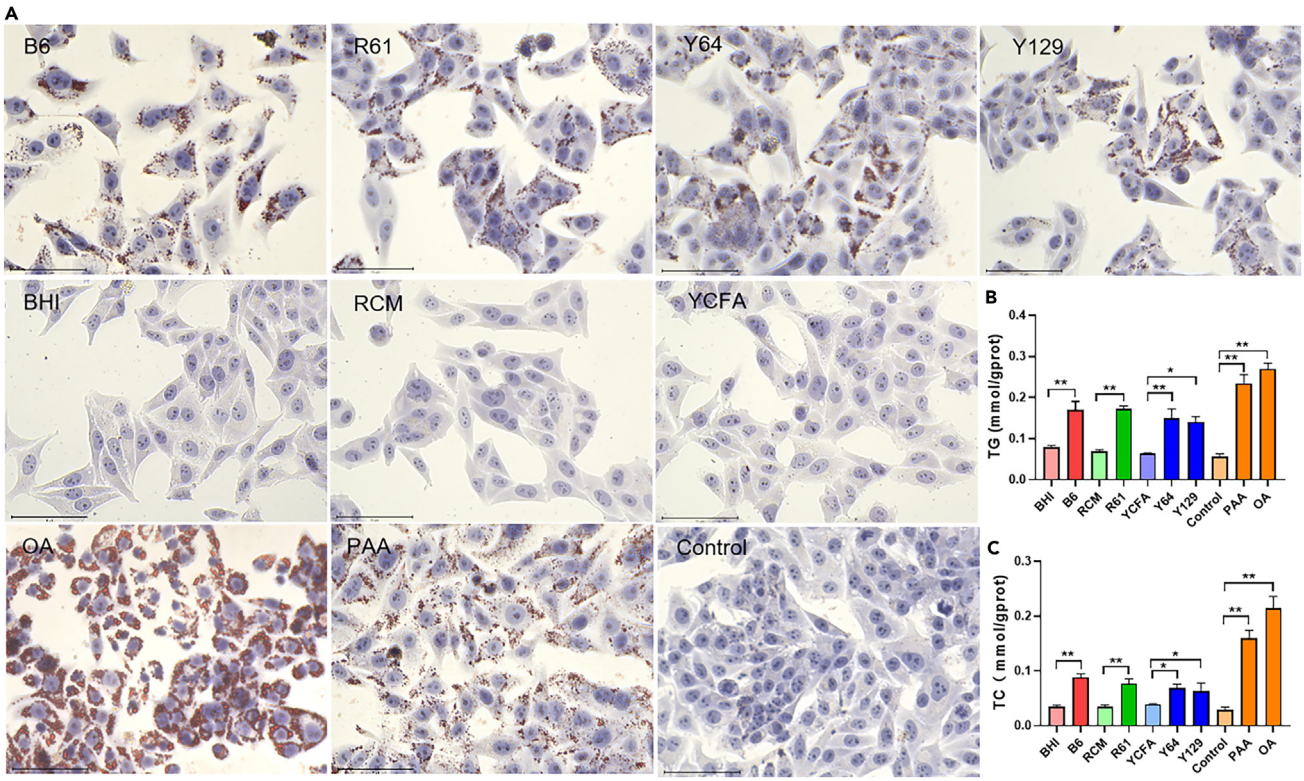


Figure 4. The four bacterial strains with lipogenic effects (B6, R61, Y64, and Y129)

(A) Representative Oil red O staining in HepG2 cells (magnification: ×400, the scale bar: 75 μm).

(B) Levels of TG influenced by the four bacterial strains.

(C) Levels of TC affected by the four bacterial strains. Values are presented as mean ± standard deviation (SD). Student's *t* test, *p < 0.05, **p < 0.01.

were identified for the first time. These specific bacterial strains likely contributed to NAFLD by activating the expression of lipid synthesis genes and inflammatory genes, and suppressing the expression of lipid oxidation genes. Moreover, the abundance of the four specific gut bacterial strains (*Enterococcus* sp. B6, *Streptococcus* sp. R61, *Escherichia* sp. Y64, and *Klebsiella* sp. Y129) was significantly associated with NAFLD. These strains might serve as new noninvasive biomarkers and treatment targets in children with NAFLD. These findings provided knowledge for exploring the pathogenesis and therapeutic targets of NAFLD in children.

In this study, we identified the dominant commensal gut bacterial strains in children with NAFLD using metagenomic sequencing technology and isolated-cultured methods. Hundreds of commensal gut bacterial strains were isolated through culturing. The bacterial profile of each fecal sample was similar to the results measured using metagenomic sequencing. Moreover, the two approaches complemented each other. On the one hand, some genera and strains were determined by metagenomic analysis, while they had not been isolated and cultured, such as strains in *Blautia*, *Prevotella*, and *Eubacterium*. On the other hand, the genus *Sutterella*, which belonged to the *Proteobacteria* phylum, was successfully isolated and cultured in our work, despite not being characterized by metagenomic sequencing. Previous observations noted inconsistencies in the gut microbiota between metagenomics and bacterial culture, wherein certain gut bacteria (e.g., *Clostridium newborn*, *Clostridium gas*, and *E. coli*) related to premature necrotizing identified through culturomics did not exhibit any correlation with the disease by metagenomic analysis.[36] Li et al.[37] also suggested that only 42.17% of the large-scale culturomic-isolated species were detected by ultra-deep metagenomics. This may be because the comparative genetic databases are not robust enough; it is necessary to isolate and identify more bacteria for the reference database.[38] Moreover, isolation and culture bacteria at the strain level in children with NAFLD provided viable bacteria for causal validation, mechanism research, and further application.[39–41] Such knowledge may be helpful for further pathogen identification and probiotics development.

In recent years, a few studies have isolated and cultured pathogenic bacteria related to NAFLD in adults. Yuan et al.[33] isolated *Klebsiella pneumonia* W14 and TH1 and found that they contributed to NAFLD in adults by producing ethanol. *Enterobacter cloacae* B29 was isolated from the feces of adult patients with severe obesity, and it contributed to NALFD by utilizing the essential upstream molecular event lipopolysaccharide-toll-like-receptor-4 cross-talk.[42,43] However, few strains have been precisely isolated from humans with NAFLD, especially in children. In this study, four novel commensal bacterial strains with lipogenic effects were identified in children for the first time, including strains *Enterococcus* sp. B6, *Streptococcus* sp. R61, *Escherichia* sp. Y64, and *Klebsiella* sp. Y129. Previous research suggested that *Enterococcus* influenced host metabolism in different ways. For example, Quan et al.[44] found that *Enterococcus* reduced obesity and improved hepatic steatosis

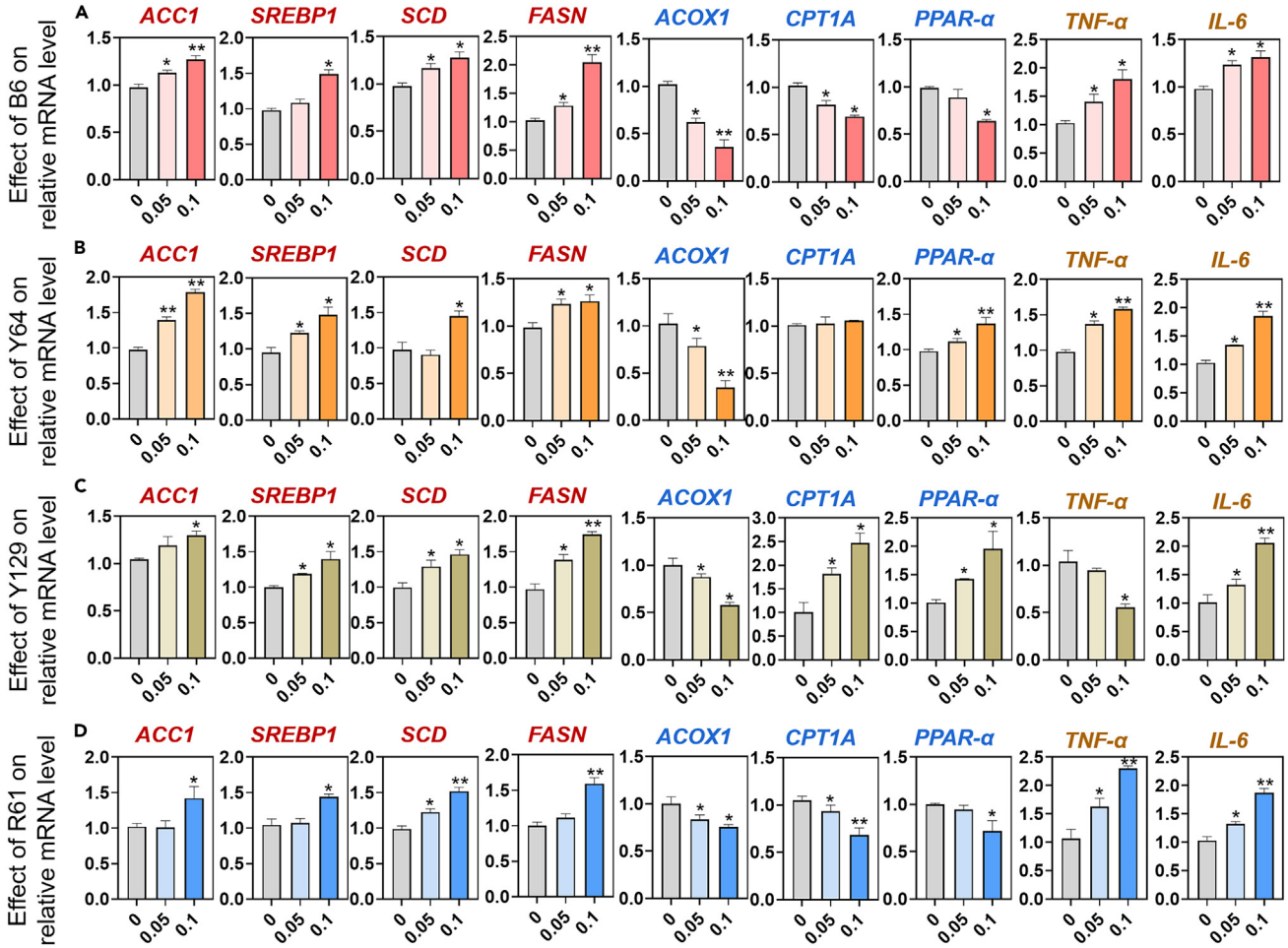


Figure 5. Expression of lipid synthesis genes (*ACC1*, *SREBP1*, *SCD*, *FASN*), lipid oxidation genes (*ACOX1*, *CPT-1A*, *PPARα*), and inflammatory genes (*TNF-α*, *IL-6*) influenced by the cell-free supernatants of bacterial strains (A) B6, (B) Y64, (C) Y129, and (D) R61 in HepG2 cells

Values are presented as mean ± SD. Student's *t* test, *p < 0.05, **p < 0.01.

by myristoleic acid. Conversely, Duan et al.[45,46] proposed that *Enterococcus* eliminated ethanol-induced liver disease by a two-subunit exotoxin. *Streptococcus mutans* might aggravate mice with NASH by collagen-binding protein and protein antigens.[47] Besides, it was hypothesized that the commercial *Escherichia fergusonii* could cause NAFLD in rats by interfering with the hepatic lipid metabolism.[48] Translocation of *E. coli* increased triglyceride synthesis in the liver of mice.[49] In addition, high alcohol-producing *K. pneumoniae* could induce liver damage via DNA methylation, mitochondrial dysfunction, and 2,3-butanediol fermentation.[50–52] These findings demonstrated that some specific bacterial strains are responsible for the development of NAFLD, and the metabolites derived from these gut bacteria might be crucial in this process. In particular, gene expression analysis showed that the four commensal bacterial strains might promote NAFLD by activating the lipid synthesis and inflammatory reaction, and suppressing the lipid oxidation process. Our findings suggest that the pathogenic bacteria in children are different from those in adults, and numerous pathogenic bacterial strains might coexist in the gastrointestinal tract of children with NAFLD. Moreover, the pathogenesis of NAFLD is also linked to an increase in lipid production and absorption that exceeded lipid oxidation and excretion, resulting in lipid accumulation and inflammatory response in the liver.[53,54] Only four commensal bacterial strains were found to be able to cause lipid accumulation in HepG2 cell lines, but many other isolated strains from the same species could not induce lipogenesis. Therefore, the lipogenic functions might be strain-specific, even though they belonged to the same species. Furthermore, the unique lipogenic effects and mechanisms of the four strains need to be investigated using comparative genomic/metabolomic analysis in the future.

Notably, our investigation extended beyond merely exploring the distribution of the gut microbiota in the discovery cohort. We delved deeper into the abundance of specific bacterial strains by validating our findings in a larger population. Despite the limited sample size of ten in the discovery cohort for metagenomics analysis, the rarefaction and species accumulation curves exhibited flattening, indicating that both sequencing depth and sample size approached the expected level. Specifically, increasing the sequencing depth and incorporating new samples did not lead to a significant rise in undiscovered species or the total number of OTUs within the community.[55,56] Moreover, our

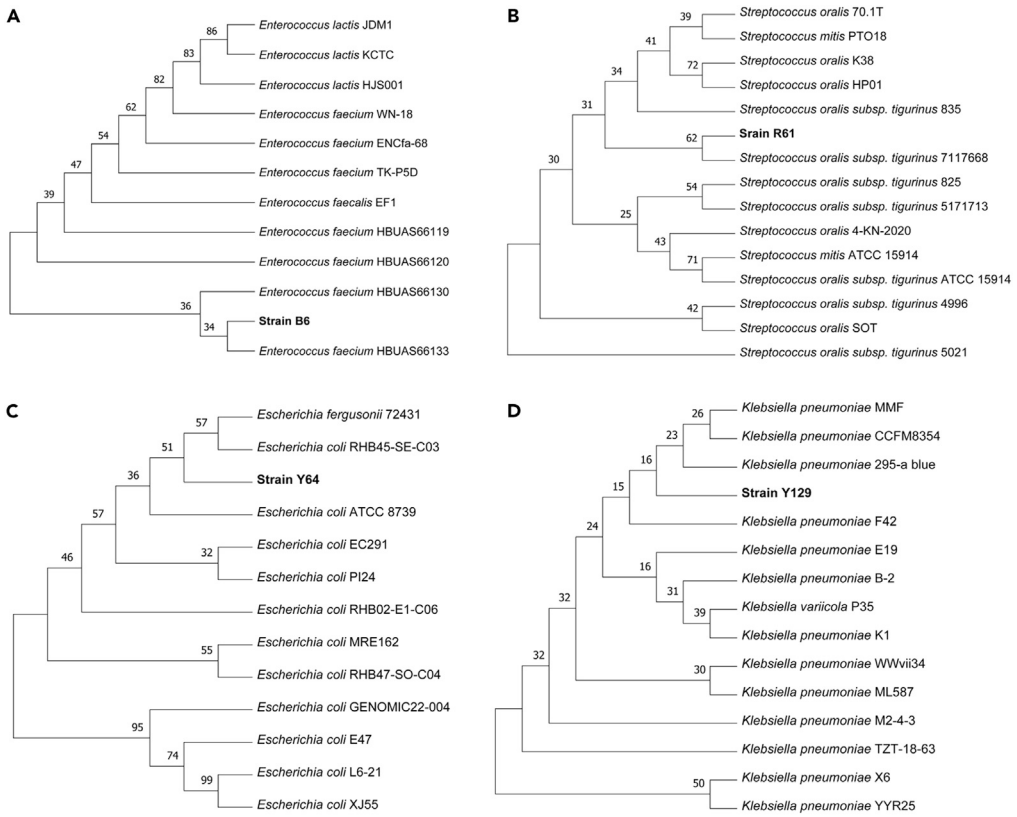


Figure 6. Phylogenetic analysis of the four bacterial strains

Neighbor-joining tree based on 16S rDNA sequences depicting the phylogenetic position of strains (A) B6, (B) R61, (C) Y64, and (D) Y129.

comprehensive approach included metagenomic and culturomic analyses, enabling us to identify bacterial profiles in children with NAFLD and isolate live bacterial strains for further causality investigation. Besides, we explored the lipogenic effects of these strains *in vitro*, which indicated the important roles of these strains in NAFLD. Furthermore, we demonstrated that the abundance of strains with lipogenic effects, including *Enterococcus* sp. B6, *Streptococcus* sp. R61, *Escherichia* sp. Y64, and *Klebsiella* sp. Y129, was significantly higher in children with NAFLD than in controls in a larger validation cohort. This multi-faceted investigation and validation process enhanced the robustness of our findings, reinforcing the relevance and credibility of our results. These data supported the hypothesis that these bacterial strains might serve as a critical microbiological marker in children with NAFLD. Some earlier studies also indicated a link between gut microbiota and NAFLD.[19,57,58] Nevertheless, gut microbiota was also associated with weight and obesity status.[59,60] This study is the first to focus on the differences in gut microbes between children with NAFLD and those with simple obesity. Moreover, integrated analysis of the four specific bacterial strains showed an excellent ability to distinguish between NAFLD and controls, as well as between NASH and NAFL. Microbiota data for disease prediction provides great potential in understanding the microbial roles in disease diagnosis and prognosis.[61,62] The accuracy of the strain-based model was higher than the model with the serological markers and morphological phenotypes alone.[14,63] As the application of liver biopsy, the gold standard for NAFLD diagnosis, is limited mainly by its invasive nature,[64] many studies have also investigated the prediction power of gut microbiota in NAFLD diagnosis.[65,66] For example, Hoyles et al.[65] constructed a model using fecal metagenomics and other molecular phenomic signatures, and reported an AUC of 0.87 in the NAFLD diagnosis among obese women. However, the power of the model in distinguishing different stages of NAFLD (such as NAFL vs. NASH) was not investigated. Our study paves the way for further investigations into the possible value of a noninvasive examination to discriminate between NAFL and NASH in the future.

This study possesses several advantages. We explored the gut microbiota involved in children with NAFLD at the strain level using metagenomic combined with culturomic technology. Subsequently, through rigorous cell experiments, we identified four novel commensal bacterial strains that probably participate in NAFLD development in children with obesity for the first time. Moreover, we conducted both *in vitro* experiments to explore related molecular mechanisms and a validation study on the population level. In brief, this study consumed time and labor, and it was innovative and challenging to isolate and evaluate the functional properties of such enormous isolates at the strain level.

For clinical implication, the fecal abundance of these strains might serve as a novel noninvasive biomarker for NAFLD. In the future, we may estimate and forecast the risk condition of the host metabolism by determining the abundance of these strains in the fecal environment. Additionally, we could develop medications, probiotics, and bacteriophages that precisely inhibit the growth of these pathogenic strains and restore intestinal balance. Moreover, identifying harmful small molecules (e.g., microbial-derived metabolites, toxins, enzymes) and the

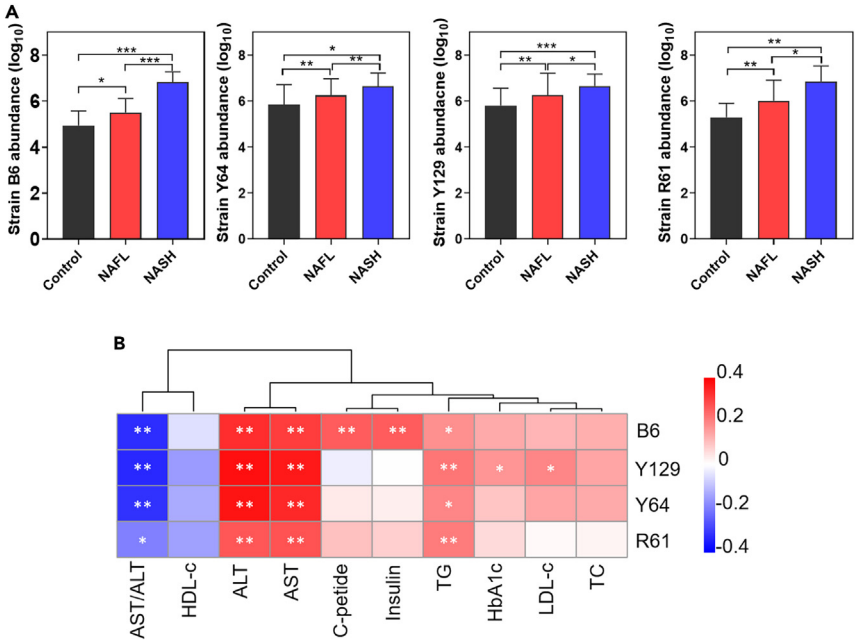


Figure 7. Association of the specific bacterial strains and children NAFLD

(A) The abundance of strain B6, Y64, Y129, and R61 in the control (n = 59), NAFL (n = 59), and NASH (n = 43) groups. Kruskal-Wallis test for pairwise multiple comparisons.

(B) Spearman's rank correlation revealed the relationship between clinical indicators and strains with lipogenic effects. The abundance of strains with lipogenic effects was determined by qPCR and defined as the $\log_{10}$ (16s rRNA gene copies/mL). *p < 0.05, **p < 0.01, ***p < 0.001.

encoding genes may assist us in modifying the functionalities of these pathogenic strains. Finally, revealing the underlying mechanisms of these strains that might aggravate NAFLD development would allow us to effectively treat it with selective inhibitors or activators that specifically hinder NAFLD development. These approaches are valuable as they provide the possibility for precise NAFLD prevention and treatment in young patients.

In conclusion, this study was the first endeavor to reveal the dominant gut microbial profile of children with NAFLD using metagenomics and culturomics. Four novel commensal bacterial strains with lipogenic effects were identified successfully, which might regulate hepatic lipid metabolism and contribute to hepatocyte steatosis. Notably, compared to the controls, the abundance of the four novel bacterial strains was higher in children with NAFLD, suggesting that they might represent a novel noninvasive biomarker for the disease. These findings provide new insights to identify new biomarkers and therapeutic targets that interrupt NAFLD development.

Limitations of the study

However, there were also some limitations in this work. First, we employed limited media and culture conditions to isolate commensal bacterial strains in this study; thus, some bacterial strains, both with and without lipogenic effects, might not have been entirely isolated. Further studies should aim to enhance culture conditions and comprehensively capture the gut microbial profiles of children with obesity. Second, it is necessary to validate the role of the four specific bacterial strains in various cell lines. Finally, we focused on bacterial strains with lipogenic effects in this work, whereas alternative bacterial strains and their pathogenic mechanisms in NAFLD development have not yet been explored. In addition, longitudinal and longer-term follow-up studies are required to further confirm our claims.

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 STUDY PARTICIPANT DETAILS
  - Subjects enrolled and samples collection
  - Cell line and maintenance

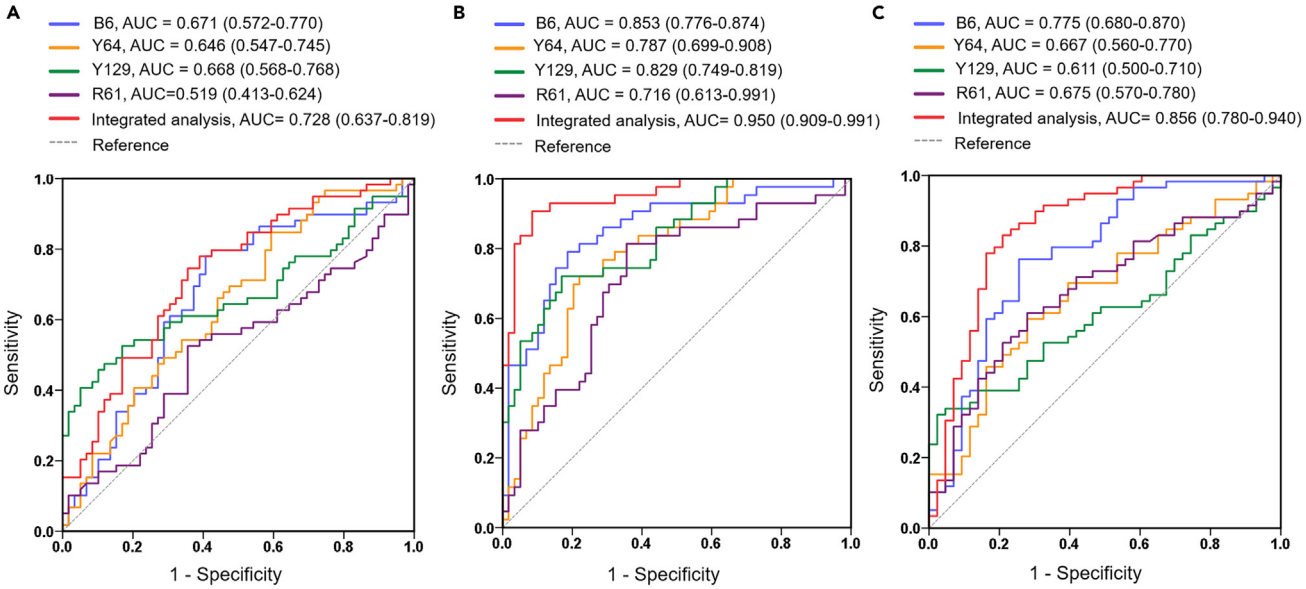


Figure 8. Receiver operating characteristic curve of the models constructed using the abundances of bacterial strains with lipogenic effects

(A) Diagnostic accuracy of the model for NAFL vs. control.

(B) Diagnostic accuracy of the model for control vs. NAFH.

(C) Diagnostic accuracy of the model for NAFL vs. NAFH.

- METHOD DETAILS
  - Total DNA isolation and metagenomic analysis
  - Culturomic analysis
  - Bacterial identification
  - Screening of candidate bacterial strains with lipogenic effects
  - Candidate bacterial strains treatment in HepG2 cells
  - Measuring the specific bacterial strains in children with NAFLD and controls
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.108861.

ACKNOWLEDGMENTS

We thank all study participants, research staff, and students who participated in this work. This work was supported by grants from the National Natural Science Foundation of China (82304171), the Hunan Province Natural Science Foundation (2022JJ40668, 2021JJ30901), and the Fundamental Research Funds for the Central Universities of Central South University (2022ZZTS0846).

AUTHOR CONTRIBUTIONS

Conceptualization, J.W. and M.L.; Methodology, J.W. and W.D.; Investigation, J.W. and X.P.; Resources, F.Y., M.L., and J.L., Visualization, J.W.; Data Curation, W.D. and J.L.; Funding Acquisition, M.L., J.L., and J.W.; Project Administration, M.L., F.Y., and J.L.; Supervision, M.L. and J.L.; Writing–Original Draft, J.W.; Writing–Review and Editing, J.W., J.L., F.Y., and M.L.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 14, 2023
Revised: November 7, 2023
Accepted: January 8, 2024
Published: January 11, 2024

REFERENCES

1. Targher, G., Tilg, H., and Byrne, C.D. (2021). Non-alcoholic fatty liver disease: a multisystem disease requiring a multidisciplinary and holistic approach. Lancet. Gastroenterol. Hepatol. *6*, 578–588.
2. Polyzos, S.A., Kountouras, J., and Mantzoros, C.S. (2019). Obesity and nonalcoholic fatty liver disease: From pathophysiology to therapeutics. Metabolism *92*, 82–97.
3. Loomba, R., Friedman, S.L., and Shulman, G.I. (2021). Mechanisms and disease consequences of nonalcoholic fatty liver disease. Cell *184*, 2537–2564.
4. Simon, T.G., Roelstraete, B., Hartjes, K., Shah, U., Khalili, H., Arnell, H., and Ludvigsson, J.F. (2021). Non-alcoholic fatty liver disease in children and young adults is associated with increased long-term mortality. J. Hepatol. *75*, 1034–1041.
5. Reimer, K.C., Wree, A., Roderburg, C., and Tacke, F. (2020). New drugs for NAFLD: lessons from basic models to the clinic. Hepatol. Int. *14*, 8–23.
6. Xie, L., Yuan, Y., Xu, S., Lu, S., Gu, J., Wang, Y., Wang, Y., Zhang, X., Chen, S., Li, J., et al. (2022). Downregulation of hepatic ceruloplasmin ameliorates NAFLD via SCO1-AMPK-LKB1 complex. Cell Rep. *41*, 111498.
7. Chassaing, B., Etienne-Mesmin, L., and Gewirtz, A.T. (2014). Microbiota-liver axis in hepatic disease. Hepatology *59*, 328–339.
8. Young, V.B. (2017). The role of the microbiome in human health and disease: an introduction for clinicians. Br. Med. J. *356*, j831.
9. Koning, M., Herrema, H., Nieuwdorp, M., and Meijnikman, A.S. (2023). Targeting nonalcoholic fatty liver disease via gut microbiome-centered therapies. Gut Microb. *15*, 2226922.
10. Tremaroli, V., and Bäckhed, F. (2012). Functional interactions between the gut microbiota and host metabolism. Nature *489*, 242–249.
11. Sommer, F., and Bäckhed, F. (2013). The gut microbiota–masters of host development and physiology. Nat. Rev. Microbiol. *11*, 227–238.
12. Sharpton, S.R., Schnabl, B., Knight, R., and Loomba, R. (2021). Current Concepts, Opportunities, and Challenges of Gut Microbiome-Based Personalized Medicine in Nonalcoholic Fatty Liver Disease. Cell Metabol. *33*, 21–32.
13. Qiao, S., Bao, L., Wang, K., Sun, S., Liao, M., Liu, C., Zhou, N., Ma, K., Zhang, Y., Chen, Y., et al. (2020). Activation of a Specific Gut Bacteroides-Folate-Liver Axis Benefits for the Alleviation of Nonalcoholic Hepatic Steatosis. Cell Rep. *32*, 108005.
14. Aron-Wisnewsky, J., Vigliotti, C., Witjes, J., Le, P., Holleboom, A.G., Verheij, J., Nieuwdorp, M., and Clément, K. (2020). Gut microbiota and human NAFLD: disentangling microbial signatures from metabolic disorders. Nat. Rev. Gastroenterol. Hepatol. *17*, 279–297.
15. Leung, C., Rivera, L., Furness, J.B., and Angus, P.W. (2016). The role of the gut microbiota in NAFLD. Nat. Rev. Gastroenterol. Hepatol. *13*, 412–425.
16. Loomba, R., Seguritan, V., Li, W., Long, T., Klitgord, N., Bhatt, A., Dulai, P.S., Caussy, C., Bettencourt, R., Highlander, S.K., et al. (2017). Gut Microbiome-Based Metagenomic Signature for Non-invasive Detection of Advanced Fibrosis in Human Nonalcoholic Fatty Liver Disease. Cell Metabol. *25*, 1054–1062.e5.
17. Leung, H., Long, X., Ni, Y., Qian, L., Nychas, E., Siliceo, S.L., Pohl, D., Hanhineva, K., Liu, Y., Xu, A., et al. (2022). Risk assessment with gut microbiome and metabolite markers in NAFLD development. Sci. Transl. Med. *14*, eabk0855.
18. Schwimmer, J.B., Johnson, J.S., Angeles, J.E., Behling, C., Belt, P.H., Borecki, I., Bross, C., Durelle, J., Goyal, N.P., Hamilton, G., et al. (2019). Microbiome Signatures Associated With Steatohepatitis and Moderate to Severe Fibrosis in Children With Nonalcoholic Fatty Liver Disease. Gastroenterology *157*, 1109–1122.
19. Zhu, L., Baker, S.S., Gill, C., Liu, W., Alkhouri, R., Baker, R.D., and Gill, S.R. (2013). Characterization of gut microbiomes in nonalcoholic steatohepatitis (NASH) patients: a connection between endogenous alcohol and NASH. Hepatology *57*, 601–609.
20. Michail, S., Lin, M., Frey, M.R., Fanter, R., Paliy, O., Hilbush, B., and Reo, N.V. (2015). Altered gut microbial energy and metabolism in children with non-alcoholic fatty liver disease. FEMS Microbiol. Ecol. *91*, 1–9.
21. Del Chierico, F., Nobili, V., Vernocchi, P., Russo, A., De Stefanis, C., Gnani, D., Furlanello, C., Zandonà, A., Paci, P., Capuani, G., et al. (2017). Gut microbiota profiling of pediatric nonalcoholic fatty liver disease and obese patients unveiled by an integrated meta-omics-based approach. Hepatology *65*, 451–464.
22. Lagier, J.C., Dubourg, G., Million, M., Cadoret, F., Bilen, M., Fenollar, F., Levasseur, A., Rolain, J.M., Fournier, P.E., and Raoult, D. (2018). Culturing the human microbiota and culturomics. Nat. Rev. Microbiol. *16*, 540–550.
23. Han, D., Gao, P., Li, R., Tan, P., Xie, J., Zhang, R., and Li, J. (2020). Multicenter assessment of microbial community profiling using 16S rRNA gene sequencing and shotgun metagenomic sequencing. J. Adv. Res. *26*, 111–121.
24. Lagier, J.C., Khelaifia, S., Alou, M.T., Ndongo, S., Dione, N., Hugon, P., Caputo, A., Cadoret, F., Traore, S.I., Seck, E.H., et al. (2016). Culture of previously uncultured members of the human gut microbiota by culturomics. Nat. Microbiol. *1*, 16203.
25. Forster, S.C., Kumar, N., Anonye, B.O., Almeida, A., Viciani, E., Stares, M.D., Dunn, M., Mkandawire, T.T., Zhu, A., Shao, Y., et al. (2019). A human gut bacterial genome and culture collection for improved metagenomic analyses. Nat. Biotechnol. *37*, 186–192.
26. Zou, Y., Xue, W., Luo, G., Deng, Z., Qin, P., Guo, R., Sun, H., Xia, Y., Liang, S., Dai, Y., et al. (2019). 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. Nat. Biotechnol. *37*, 179–185.
27. Browne, H.P., Forster, S.C., Anonye, B.O., Kumar, N., Neville, B.A., Stares, M.D., Goulding, D., and Lawley, T.D. (2016). Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. Nature *533*, 543–546.
28. Lagier, J.C., Hugon, P., Khelaifia, S., Fournier, P.E., La Scola, B., and Raoult, D. (2015). The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin. Microbiol. Rev. *28*, 237–264.
29. Mukhopadhya, I., Martin, J.C., Shaw, S., McKinley, A.J., Gratz, S.W., and Scott, K.P. (2022). Comparison of microbial signatures between paired faecal and rectal biopsy samples from healthy volunteers using next-generation sequencing and culturomics. Microbiome *10*, 171.
30. Whelan, F.J., Waddell, B., Syed, S.A., Shekarriz, S., Rabin, H.R., Parkins, M.D., and Surette, M.G. (2020). Culture-enriched metagenomic sequencing enables in-depth profiling of the cystic fibrosis lung microbiota. Nat. Microbiol. *5*, 379–390.
31. Natividad, J.M., Lamas, B., Pham, H.P., Michel, M.L., Rainteau, D., Bridonneau, C., da Costa, G., van Hylckama Vlieg, J., Sovran, B., Chamignon, C., et al. (2018). Bilophila wadsworthia aggravates high fat diet induced metabolic dysfunctions in mice. Nat. Commun. *9*, 2802.
32. Woting, A., Pfeiffer, N., Loh, G., Klaus, S., and Blaut, M. (2014). Clostridium ramosum promotes high-fat diet-induced obesity in gnotobiotic mouse models. mBio *5*, e01530-14.
33. Yuan, J., Chen, C., Cui, J., Lu, J., Yan, C., Wei, X., Zhao, X., Li, N., Li, S., Xue, G., et al. (2019). Fatty Liver Disease Caused by High-Alcohol-Producing Klebsiella pneumoniae. Cell Metabol. *30*, 675–688.e7.
34. Meijnikman, A.S., Davids, M., Herrema, H., Aydin, O., Tremaroli, V., Rios-Morales, M., Levels, H., Bruin, S., de Brauw, M., Verheij, J., et al. (2022). Microbiome-derived ethanol in nonalcoholic fatty liver disease. Nat. Med. *28*, 2100–2106.
35. Tilg, H., Adolph, T.E., and Trauner, M. (2022). Gut-liver axis: Pathophysiological concepts and clinical implications. Cell Metabol. *34*, 1700–1718.
36. Cassir, N., Benamar, S., Khalil, J.B., Croce, O., Saint-Faust, M., Jacquot, A., Million, M., Azza, S., Armstrong, N., Henry, M., et al. (2015). Clostridium butyricum Strains and Dysbiosis Linked to Necrotizing Enterocolitis in Preterm Neonates. Clin. Infect. Dis. *61*, 1107–1115.
37. Li, C., Luan, Z., Zhao, Y., Chen, J., Yang, Y., Wang, C., Jing, Y., Qi, S., Li, Z., Guo, H., et al. (2022). Deep insights into the gut microbial community of extreme longevity in south Chinese centenarians by ultra-deep metagenomics and large-scale culturomics. NPJ Biofilms Microbiomes *8*, 28.
38. Lagier, J.C., Million, M., Hugon, P., Armougom, F., and Raoult, D. (2012). Human gut microbiota: repertoire and variations. Front. Cell. Infect. Microbiol. *2*, 136.
39. Lewis, W.H., Tahon, G., Geesink, P., Sousa, D.Z., and Ettema, T.J.G. (2021). Innovations to culturing the uncultured microbial majority. Nat. Rev. Microbiol. *19*, 225–240.
40. Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., Butterfield, C.N., Hernsdorf, A.W., Amano, Y., Ise, K., et al. (2016). A new view of the tree of life. Nat. Microbiol. *1*, 16048.
41. Bilen, M., Dufour, J.C., Lagier, J.C., Cadoret, F., Daoud, Z., Dubourg, G., and Raoult, D. (2018). The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species. Microbiome *6*, 94.
42. Fei, N., Bruneau, A., Zhang, X., Wang, R., Wang, J., Rabot, S., Gérard, P., and Zhao, L. (2020). Endotoxin Producers Overgrowing in Human Gut Microbiota as the Causative Agents for Nonalcoholic Fatty Liver Disease. mBio *11*.
43. Fei, N., and Zhao, L. (2013). An opportunistic pathogen isolated from the gut of an obese

human causes obesity in germfree mice. ISME J. *7*, 880–884.
44. Quan, L.H., Zhang, C., Dong, M., Jiang, J., Xu, H., Yan, C., Liu, X., Zhou, H., Zhang, H., Chen, L., et al. (2020). Myristoleic acid produced by enterococci reduces obesity through brown adipose tissue activation. Gut *69*, 1239–1247.
45. Duan, Y., Llorente, C., Lang, S., Brandl, K., Chu, H., Jiang, L., White, R.C., Clarke, T.H., Nguyen, K., Torralba, M., et al. (2019). Bacteriophage targeting of gut bacterium attenuates alcoholic liver disease. Nature *575*, 505–511.
46. Nunez, N., Derré-Bobillot, A., Trainel, N., Lakisic, G., Lecomte, A., Mercier-Nomé, F., Cassard, A.M., Bierne, H., Serror, P., and Archambaud, C. (2022). The unforeseen intracellular lifestyle of Enterococcus faecalis in hepatocytes. Gut Microb. *14*, 2058851.
47. Naka, S., Hatakeyama, R., Takashima, Y., Matsumoto-Nakano, M., Nomura, R., and Nakano, K. (2016). Contributions of Streptococcus mutans Cnm and PA antigens to aggravation of non-alcoholic steatohepatitis in mice. Sci. Rep. *6*, 36886.
48. Xin, F.Z., Zhao, Z.H., Liu, X.L., Pan, Q., Wang, Z.X., Zeng, L., Zhang, Q.R., Ye, L., Wang, M.Y., Zhang, R.N., et al. (2022). Escherichia fergusonii Promotes Nonobese Nonalcoholic Fatty Liver Disease by Interfering With Host Hepatic Lipid Metabolism Through Its Own msRNA 23487. Cell. Mol. Gastroenterol. Hepatol. *13*, 827–841.
49. Shen, B., Gu, T., Shen, Z., Zhou, C., Guo, Y., Wang, J., Li, B., Xu, X., Li, F., Zhang, Q., et al. (2023). Escherichia coli Promotes Endothelial to Mesenchymal Transformation of Liver Sinusoidal Endothelial Cells and Exacerbates Nonalcoholic Fatty Liver Disease Via Its Flagellin. Cell. Mol. Gastroenterol. Hepatol. *16*, 857–879.
50. Li, N.N., Li, W., Feng, J.X., Zhang, W.W., Zhang, R., Du, S.H., Liu, S.Y., Xue, G.H., Yan, C., Cui, J.H., et al. (2021). High alcohol-producing Klebsiella pneumoniae causes fatty liver disease through 2,3-butanediol fermentation pathway *in vivo*. Gut Microb. *13*, 1979883.
51. Chen, X., Zhang, Z., Li, H., Zhao, J., Wei, X., Lin, W., Zhao, X., Jiang, A., and Yuan, J. (2020). Endogenous ethanol produced by intestinal bacteria induces mitochondrial dysfunction in non-alcoholic fatty liver disease. J. Gastroenterol. Hepatol. *35*, 2009–2019.
52. Zhang, R., Xu, Z., Xue, G., Feng, J., Du, B., Gan, L., Fan, Z., Fu, T., Feng, Y., Zhao, H., et al. (2023). Combined Methylation and Transcriptome Analysis of Liver Injury of Nonalcoholic Fatty Liver Disease Induced by High Alcohol-Producing Klebsiella pneumoniae. Microbiol. Spectr. *11*, e0532322.
53. Xu, X., Poulsen, K.L., Wu, L., Liu, S., Miyata, T., Song, Q., Wei, Q., Zhao, C., Lin, C., and Yang, J. (2022). Targeted therapeutics and novel signaling pathways in non-alcohol-associated fatty liver/steatohepatitis (NAFL/NASH). Signal Transduct. Targeted Ther. *7*, 287.
54. Wang, S., Sheng, F., Zou, L., Xiao, J., and Li, P. (2021). Hyperoside attenuates non-alcoholic fatty liver disease in rats via cholesterol metabolism and bile acid metabolism. J. Adv. Res. *34*, 109–122.
55. Kemp, P.F., and Aller, J.Y. (2004). Bacterial diversity in aquatic and other environments: what 16S rDNA libraries can tell us. FEMS Microbiol. Ecol. *47*, 161–177.
56. Chao, A., and Shen, T.J. (2004). Nonparametric Prediction in Species Sampling. J. Agric. Biol. Environ. Stat. *9*, 253–269.
57. Boursier, J., Mueller, O., Barret, M., Machado, M., Fizanne, L., Araujo-Perez, F., Guy, C.D., Seed, P.C., Rawls, J.F., David, L.A., et al. (2016). The severity of nonalcoholic fatty liver disease is associated with gut dysbiosis and shift in the metabolic function of the gut microbiota. Hepatology *63*, 764–775.
58. Pan, X., Kaminga, A.C., Liu, A., Wen, S.W., Luo, M., and Luo, J. (2021). Gut Microbiota, Glucose, Lipid, and Water-Electrolyte Metabolism in Children With Nonalcoholic Fatty Liver Disease. Front. Cell. Infect. Microbiol. *11*, 683743.
59. Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P., et al. (2009). A core gut microbiome in obese and lean twins. Nature *457*, 480–484.
60. Stanislawski, M.A., Dabelea, D., Lange, L.A., Wagner, B.D., and Lozupone, C.A. (2019). Gut microbiota phenotypes of obesity. NPJ Biofilms Microbiomes *5*, 18.
61. Wang, C., Segal, L.N., Hu, J., Zhou, B., Hayes, R.B., Ahn, J., and Li, H. (2022). Microbial risk score for capturing microbial characteristics, integrating multi-omics data, and predicting disease risk. Microbiome *10*, 121.
62. Kovács, Z., Glover, L., Reidy, F., MacSharry, J., and Saldova, R. (2021). Novel diagnostic options for endometriosis - Based on the glycome and microbiome. J. Adv. Res. *33*, 167–181.
63. Pirmoazen, A.M., Khurana, A., El Kaffas, A., and Kamaya, A. (2020). Quantitative ultrasound approaches for diagnosis and monitoring hepatic steatosis in nonalcoholic fatty liver disease. Theranostics *10*, 4277–4289.
64. Vajro, P., Lenta, S., Socha, P., Dhawan, A., McKiernan, P., Baumann, U., Durmaz, O., Lacaille, F., McLin, V., and Nobili, V. (2012). Diagnosis of nonalcoholic fatty liver disease in children and adolescents: position paper of the ESPGHAN Hepatology Committee. J. Pediatr. Gastroenterol. Nutr. *54*, 700–713.
65. Hoyles, L., Fernández-Real, J.M., Federici, M., Serino, M., Abbott, J., Charpentier, J., Heymes, C., Luque, J.L., Anthony, E., Barton, R.H., et al. (2018). Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women. Nat. Med. *24*, 1070–1080.
66. Tilg, H., Adolph, T.E., Dudek, M., and Knolle, P. (2021). Non-alcoholic fatty liver disease: the interplay between metabolism, microbes and immunity. Nat. Metab. *3*, 1596–1607.
67. Li, R., Yu, C., Li, Y., Lam, T.W., Yiu, S.M., Kristiansen, K., and Wang, J. (2009). SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics *25*, 1966–1967.
68. Zhu, W., Lomsadze, A., and Borodovsky, M. (2010). Ab initio gene identification in metagenomic sequences. Nucleic Acids Res. *38*, e132.
69. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. Nat. Methods *12*, 59–60.
70. Group of China Obesity Task Force (2004). Body mass index reference normal for screening overweight and obesity in Chinese children and adolescents. Clin. J. Epidemiol. *25*, 97–102.
71. Zhou, X., and Fu, J. (2018). Expert consensus on diagnosis and treatment of nonalcoholic fatty liver disease in children. Chin. J. Pract. Pediatr. *33*, 487–492.
72. Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics *22*, 1658–1659.
73. Karlsson, F.H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C.J., Fagerberg, B., Nielsen, J., and Bäckhed, F. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature *498*, 99–103.
74. Asnicar, F., Weingart, G., Tickle, T.L., Huttenhower, C., and Segata, N. (2015). Compact graphical representation of phylogenetic data and metadata with GraPhlAn. PeerJ *3*, e1029.
75. Wu, H.T., Lu, F.H., Ou, H.Y., Su, Y.C., Hung, H.C., Wu, J.S., Yang, Y.C., Wu, C.L., and Chang, C.J. (2013). The role of hepassocin in the development of non-alcoholic fatty liver disease. J. Hepatol. *59*, 1065–1072.
76. Wei, J., Dai, W., Pan, X., Zhong, Y., Xu, N., Ye, P., Wang, J., Li, J., Yang, F., Luo, J., and Luo, M. (2023). Identifying the Novel Gut Microbial Metabolite Contributing to Metabolic Syndrome in Children Based on Integrative Analyses of Microbiome-Metabolome Signatures. Microbiol. Spectr. *11*, e0377122.
77. Brankatschk, R., Bodenhausen, N., Zeyer, J., and Bürgmann, H. (2012). Simple absolute quantification method correcting for quantitative PCR efficiency variations for microbial community samples. Appl. Environ. Microbiol. *78*, 4481–4489.
78. Ho, J., Tumkaya, T., Aryal, S., Choi, H., and Claridge-Chang, A. (2019). Moving beyond P values: data analysis with estimation graphics. Nat. Methods *16*, 565–566.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Enterococcus sp. B6	This study	N/A
Streptococcus sp. R61	This study	N/A
Escherichia sp. Y64	This study	N/A
Klebsiella sp. Y129	This study	N/A
Biological samples		
Fecal samples of children with NAFLD	This study	N/A
Chemicals, peptides, and recombinant proteins		
Trypsin-EDTA	Gibco	Cat# 25200-056
DMEM, high glucose, no glutamine	Gibco	Cat# 11960069
Fetal bovine serum	Gibco	Cat# 10270-106
Penicillin-Streptomycin (10,000 U/mL)	Gibco	Cat# 15140-122
Oleic acid	Sigma-Aldrich	Cat# O1008, CAS: 112-80-1
Phenylacetic acid	Sigma-Aldrich	Cat# P16621, CAS: 103-82-2
Brian Heart Infusion medium	Oxoid	Cat# CM1135
Reinforced Clostridial Medium	Oxoid	Cat# CM0149
Columbia Medium	Oxoid	Cat# CM0331B
Man Rogosa Sharp medium	Oxoid	Cat# CM0361B
Critical commercial assays		
Stool DNA Isolation Kit	Qiagen	Cat# DP302
Gel Extraction Kit	Qiagen	Cat# DP204
Oil red O staining	Solarbio	Cat# G1262
Total triglycerides (TG) kits	Jiancheng	Cat# A110-1-1
Total cholesterol (TC) kits	Jiancheng	Cat# A111-1-1
Trizol reagent	Vazyme	Cat# R401-01
HiScript II Q RT SuperNix for qPCR	Vazyme	Cat# R212-01
ChamQ Universal SYBR qPCR Master Mix	Vazyme	Cat# Q711-02
Deposited data		
16S rRNA gene sequence of strain B6	This study	NCBI Accession number: OP874955
16S rRNA gene sequence of strain R61	This study	NCBI Accession number: OP875089
16S rRNA gene sequence of strain Y64	This study	NCBI Accession number: OP875090
16S rRNA gene sequence of strain Y129	This study	NCBI Accession number: OP875098
Experimental models: Cell lines		
HepG2	the Cancer Research Institute of Central South University	
Oligonucleotides		
Specific primers used for qRT-PCR of cell genes, see Table S4	This study	
Specific primers used for qPCR of bacterial genes, see Table S5	This study	

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
SOAPdenovo v2.04	Li et al.[67]	https://doi.org/10.1093/bioinformatics/btp336
MetaGeneMark v2.10	Zhu et al.[68]	https://doi.org/10.1093/nar/gkq275
DIAMOND software	Buchfink et al.[69]	https://doi.org/10.1038/nmeth.3176
Graphpad Prism v 8.0.1	GraphPad Software Inc.	https://www.graphpad.com

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Miyang Luo (miyangluo@csu.edu.cn).

Materials availability

This study didn't generate new unique reagents.

Data and code availability

- All data have been deposited at the National Center for Biotechnology Information repository (NCBI, http://www.ncbi.nlm.nih.gov/genbank/) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report the original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Subjects enrolled and samples collection

Children with obesity aged 6-18 years were recruited from the Hunan Children's Hospital (Changsha, China). The obesity was screened using BMI-for-age and BMI-for-sex according to the national standards.[70] The diagnostic criteria of NAFLD (including NAFL and NASH) refer to expert consensus on the diagnosis and treatment of NAFLD in Children.[71] Children with obesity who were first diagnosed with NAFLD were included in the NAFLD groups. Children with obesity but without NAFLD were set as the control group. In addition, patients were excluded if they had excessive alcohol consumption (alcohol consumption >140 g/week for boys and >70 g/week for girls), a history of viral hepatitis and cirrhosis complications, or received steatosis-inducing medications, probiotics or antibiotics within one month. The studies were approved by the Xiangya School of Public Health Central South University Human Research Ethics Committee (XYGW-2020-20). All processes were performed following the ethical guidelines of the Helsinki Declaration. Written and informed consent was obtained from the participant's legal guardian/next of kin. Fresh fecal samples of children were collected and stored at -80°C.

Cell line and maintenance

HepG2 cells were obtained from the Cancer Research Institute of Central South University (Changsha, China). The cells were cultured in high-glucose Dulbecco's Modified Eagle's Medium (DMEM) (Gibco BRL, Grand Island, New York, USA) supplemented with 10% (vol/vol) fetal bovine serum and 1% penicillin/streptomycin, maintained in a 5% CO_2 humidified atmosphere at 37°C (Thermo Fisher Scientific, USA).

METHOD DETAILS

Total DNA isolation and metagenomic analysis

Total gut microbiota DNA was isolated from fecal pellets using a stool DNA Isolation Kit (Qiagen, Valencia, CA). Sequencing libraries were generated using NEBNext® Ultra ™ DNA Library Prep Kit for Illumina (NEB, USA) following manufacturer's recommendations. All samples were sequenced based on an Illumina NovaSeq platform (paired-end; insert size, 350 bp). After the removal of the adaptor and low-quality reads, high-quality pair-end reads for the samples were obtained (clean data) and assembled into contigs for all samples using the assembly software SOAPdenovo v2.04.[67] Open reading frames (ORFs) were predicted using the MetaGeneMark v2.10 with default parameters.[68] All ORFs were clustered by CD-hit v4.5.8 to construct a non-redundant gene catalog using a stringent criterion of 95% identity at the nucleotide level over 90% of the length of the shorter ORFs.[72] The final non-redundant gene catalog was constructed. The sequencing depths and sample sizes were evaluated according to the rarefaction and species accumulation curves. For the taxonomic assignment, DIAMOND software (V0.9.9) was used to blast the predicted genes.[69] The number of genes and the abundance of each sample in each taxonomy hierarchy was obtained based on the LCA annotation results.[73] The phylogenetic tree was constructed using GraPhlAn.[74]

Culturomic analysis

The metagenomic analysis provided initial evidence of the gut microbiota composition and functionality in the host. To make up for the limitation of sequencing technology and obtain live bacterial strains for causal inference, we performed culturomic analysis of fecal samples from the ten children with obesity and NAFLD. The fresh fecal samples were immediately transferred into an anaerobic station (Longyue LAI-3T, Shanghai, China). All the experiments were performed in an anaerobic environment (an atmosphere of 5% hydrogen, 10% carbon dioxide, and 85% nitrogen). Each fecal sample (about 0.2 g) was suspended in a sterile phosphate-buffered solution (PBS, pH 7.0) and homogenized for 15 min at a speed of 200 rpm/min. Then, the mixtures were diluted with PBS, and the serial dilutions were used to seed in various agar plates for anaerobic culturing. The exact quantity of stools diluted mixture was plated and incubated at 37°C for 24 h, and different colonies were observed and picked out. A second round of purification was carried out by sub-culturing these re-purified colonies to ensure that pure single strains were isolated. Various media were used in this work, including the BHI (Oxoid, USA), RCM (Haibo, Qingdao, China), CoM (Haibo, Qingdao, China), MRS (Solarbio, Beijing, China), NB, and YCFA media (see Table S3). All agar plates were prepared according to the manufacturer's instructions, and maintained anaerobically overnight to remove oxygen before plating.

Bacterial identification

The bacterial 16S rRNA gene of every single colony was amplified by PCR using the two universal primers, 27F (5'-AGAGTTTGATCMTGGCTCAG-3') and 1492R (5'-TACGGYTACCTTGTTACGACTT-3'). The products were sequenced by Sangon Biotech (Shanghai, China). The evolutionary relationships of these isolated bacterial strains were evaluated using the phylogenetic reconstruction strategy. We performed the similarity-based search in the NCBI BLASTn network service to provide a glimpse of provisional recognition. The maximum-likelihood (ML) phylogenetic tree was constructed using PhyML 3.0 (http://www.atgc-montpellier.fr/phyml/), and visualized using the online tool interactive Tree Of Life (https://itol.embl.de/). In addition, the neighbor-joining phylogenetic tree of the interested strains was constructed by MEGA11 (www.megasoftware.net).

Screening of candidate bacterial strains with lipogenic effects

The growth of bacteria was measured as the optical density at the 600 nm wavelength. Single bacterial strains were cultured until the stationary phase, and the cell-free supernatants of cultured bacterial strains with 3×10^9 cfu/mL were collected by centrifuging (5000 × g, 10 min, 4°C), then filtered through a 0.22 μm Millipore Aseptic PES membrane and termed as the cell-free supernatants stock solution.

The cell viability was assessed by the CCK8 method (Vazyme, Nanjing, China). For co-culture experiments, cells were incubated in HepG2 growth medium containing 10 % of the bacterial cell-free supernatants stock solution (treatment group) for 24 h at 37°C and 5% CO_2. The HepG2 growth medium supplemented with the corresponding bacterial culture media was used as the negative control. The OA and PAA (Sigma-Aldrich, MO, USA) were used as positive controls. The final concentrations of OA and PAA were set as 0.5 mM and 10 μM, respectively.[65,75] The HepG2 cells were exposed to the bacterial cell-free supernatant solution, OA, PAA, and the bacterial culture media for 24 h, respectively. All experiments were performed in triplicate. Lipid droplets in cells were determined by Oil red O staining (Solarbio, Beijing, China) according to the manufacturer's protocol. Pictures were taken on the auto cell imaging system (Invitrogen, Thermo Fisher Scientific, USA). We quantified the red area on the images of cells stained with red oil using Image Pro Plus v 6.0 software. Comparing the Oil red O staining images among these groups, when the lipid droplets were similar to the positive control, and the area of lipid droplets (orange-red color in cells) was obviously larger in the treatment group than in the negative control group, the strain might be preliminarily considered to induce lipid accumulated in cells.

Candidate bacterial strains treatment in HepG2 cells

To further validate the lipogenic effects of these candidate strains, HepG2 cells were seeded on a 6-well plate (at 6×10^5 cells per well). The HepG2 cells incubated in the growth medium containing 5% or 10% of the four bacterial cell-free supernatants, respectively, were set as the treatment groups. The HepG2 cells incubated in the growth medium supplemented with an equal quantity of the corresponding bacterial culture media were set as the control groups. All the treatment and control groups were incubated at 37°C for 24 h. The intracellular TG and TC levels were measured using the enzymatic reagent kits (Jiancheng, Nanjing, China) according to the manufacturer's instructions. When the TG and TC levels were higher in the treatment groups than in the negative control groups, we considered the strain as candidate strain with lipogenic effects. In addition, total RNA and qRT-PCR analyses were performed as per our previous report.[76] Briefly, total RNA was extracted and purified using Trizol reagent (Vazyme, Nanjing, China) following the manufacturer's protocol. The cDNA was synthesized by HiScript II Q RT SuperNix for qPCR (+gDNA wiper) (Vazyme, Nanjing, China). qPCR was performed with a ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China) on a LightCycler 480 Real-Time PCR System (Roche Diagnostics). The primers are shown in Table S4 (see Table S4). The relative quantification values for each mRNA were calculated by the $2^{-\Delta\Delta Ct}$ method. In summary, the strain with lipogenic effects was ascertained based on the following criteria: (1) induced lipid droplets in cells, (2) increased the levels of TG and TC in cells, (3) influenced the expression of genes involved in lipid metabolism and inflammation.

Measuring the specific bacterial strains in children with NAFLD and controls

A second cohort was used to verify the abundance of bacterial strains with lipogenic effects in children. The recruiting process and the clinical indicators information collection have been described above. Finally, 59 controls (children with obesity but without NAFLD), 59 children with

obesity and NAFL, and 43 children with obesity and NASH were included. Abundance of the four bacterial strains in fecal samples was analyzed by qPCR analysis.

Specifically, the specific primers were designed according to the whole sequences of the 16S rRNA genes of the four lipogenic strains using primer 6.0 (see Table S5). Firstly, total DNA of the four lipogenic strains was extracted, then the 16S rRNA genes were amplified using PCR, and the products were purified using the Gel Extraction Kit (Tiangen, Beijing, China). The purified 16S rRNA gene fragments were prepared as standard templates. The concentration of DNA was quantified by a nano spectrophotometer (Implen, Germany). The corresponding copy number of strains was calculated using the following equation:[77] $C(copies/mL) = C_{DNA}\,(ng/mL) \times 6.02 \times 10^{23}\,(bp/mol) / L_{DNA}\,(bp) \times 660\,(g/mol)$. Where C (copies/mL) represents the concentration of target genes calculated from the equation, C_{DNA} (ng/μL) represents the DNA concentration, and L_{DNA} (bp) represents the length of the target DNA fragment. The Avogadro was 6.022×10^{23} bp/mol, and the average molecular mass of a double-stranded DNA base pair is 660 (g/mol). Ct values in each 10-fold serial dilution of the standard templates were measured through a qPCR analysis. The linear regression of $\log_{10}$ (16S rRNA gene copies/mL) versus Ct value was established to generate a standard curve for each strain. The total DNA was extracted, and the 16S rRNA gene of each lipogenic strain was amplified by qPCR to quantify the abundance of the four lipogenic strains in fecal samples. The Ct values were obtained, and the concentrations of bacterial strains were calculated according to the standard curve.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

Statistical analysis was performed using SPSS (version 24.0, IBM Corp, USA) and Graph-pad Prism (version 8, San Diego, California, USA). For the microbiome analysis, taxonomic composition between groups was assessed using non-parametric tests (Mann-Whitney U) and corrected for multiple testing using the Benjamini-Hochberg approach. Non-parametric directional standardized effect sizes were likewise taken as the Cliff's delta.[78] For the population-based analysis, data were expressed as mean ± standard deviation (SD) or median (interquartile range). Statistical analysis was determined kruskal-wallis test for pairwise multiple comparisons. Spearman correlation tests were used to evaluate the association between bacterial strains and the clinical indicators. Additionally, receiver-operating characteristic (ROC) curve analysis was performed to assess the AUC, sensitivity, and specificity of bacterial strains in screening tests. For cell experiments, data were expressed as mean ± SD. Statistical analysis was performed using a student t-test to compare the different between the treatment groups and the control groups. Values of $P < 0.05$ were considered as statistically significant.