

Universal adaptive evolution strategy of probiotic *Lactobacillus plantarum* to gut selection pressure from humans, mice and zebrafish

Shi Huang

University of California San Diego

Shuaiming Jiang

Hainan University

Dongxue Huo

Hainan University

Celeste Allaband

University of California, San Diego

Mehrbod Estaki

University of California San Diego <https://orcid.org/0000-0003-4556-8444>

Victor Cantu

Department of Bioengineering, University of California San Diego

Pedro Belda-Ferre

Center for Microbiome Innovation and Departments of Pediatrics, Bioengineering, and Computer Science & Engineering, University of California San Diego, La Jolla, CA <https://orcid.org/0000-0001-6532-1161>

Yoshiki Vázquez-Baeza

Center for Microbiome Innovation, Jacobs School of Engineering, University of California San Diego

Qiyun Zhu

University of California San Diego

Chenchen Ma

Hainan University

Amir Zarrinpar

University of California, San Diego

Yang-Yu Liu

Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School

Rob Knight

University of California, San Diego <https://orcid.org/0000-0002-0975-9019>

Jiachao Zhang (✉ zhjch321123@163.com)

Hainan University <https://orcid.org/0000-0001-8099-6749>

Article

Keywords: probiotic, Lactobacillus plantarum HNU082, Lp082

Posted Date: March 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-106737/v1>

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

[Read Full License](#)

Abstract

Improving the probiotic engraftment in humans requires a thorough understanding of in vivo gut-adaptive strategy of probiotics in diverse contexts. Here, we exposed the probiotic *Lactobacillus plantarum* HNU082 (Lp082) to gut microbiota of healthy humans, mice, and zebrafish for four weeks. Independent of host choice, Lp082 established and adapted the gut by acquiring highly consistent single-nucleotide mutations, which modulated carbohydrate utilization, and acid tolerance, and significantly promoted its in vivo competitive fitness. In turn, resident gut microbial strains, especially competing strains with Lp082 (e.g. *Bacteroides* spp. and *Bifidobacterium* spp.), actively respond to Lp082 engraftment by accumulating 10-70 folds more evolutionary changes than usual. Human gut microbiota exhibited a higher ecological and genetic stability than that of mice. Collectively, the highly convergent adaptation strategy of Lp082 across host environments lay the foundation for leveraging the animal model for ex vivo engineering of probiotics for better engraftment outcomes in humans.

Introduction

Probiotics have been long used as gastrointestinal therapeutics under many health conditions¹. Unlike other therapeutics, probiotics are live micro-organisms and typically leverage different strategies to adapt within the human gut under high selective pressure^{2,3}, e.g., spontaneous adaptive mutations^{4,5}. The autonomous gut-adaptive behaviors of probiotics can confer them sufficient fitness advantages to engage in interactions with gut residents and host factors^{6,7}. On the other side, it also provided novel opportunities for leveraging gut selective forces for genetic engineering of probiotics *in vivo*³. Unfortunately, for most probiotic strains, these *in vivo* genetic processes are still poorly characterized, while they can be subject to a few known factors, e.g., the individualized genomic characteristics of probiotics, the complexity or specific members in indigenous gut microbiomes, and host-filtering mechanisms^{8,9}. Therefore, more studies are urgently needed to comprehensively characterize the adaptive evolutionary behaviors of individual probiotic strains under the diverse schemes (e.g., microbiome complexity, diet and host factors^{6,7}). Among these, the impact of host factors on probiotic adaptive behaviors have been challenging to investigate^{9,10}. Until now, the interactions between probiotics and indigenous gut microbiota/host phenotypes have been mainly established using animal models rather than humans. A wide range of animal models (including mice, flies, and zebrafish) have been employed to investigate how an engrafted probiotic (such as *Lactobacillus plantarum*) adapted within the gut¹¹, and modulated host physiology¹². However, the applicability of these findings from animal models to humans has not yet been validated. Each host model has a specific selection mechanism by which maintains the resident microbial compositions¹³ and exogenous microbes in the gut. Given the characteristics of each probiotic strain is highly individualized, it is less clear that the adaptive behaviors can be always host-dependent. But, if a particular probiotic applies a universal adaptive strategy across hosts, more novel strategies on genetic engineering of probiotics can be proposed for human use. It is especially intriguing if the fitness advantages (adaptive mutations) of a probiotic gained from the mouse gut can be stably transferred to human hosts resulting in a better

probiotic engraftment outcome. Unfortunately, no study has bridged these knowledge gaps, which impede translational research and hamper the performance improvement of probiotics in human populations.

On the other hand, the probiotics engraftment can turn to be ecological and evolutionary forces reshaping the indigenous microbial communities. Many metagenomics studies explored the ecological impact of probiotic ingestion on gut microbiota, yet marginal changes in the composition of gut microbiota have been noted^{10,14-16}. But it is under-reported that the genetic compositions in the human gut microbiome are constantly evolving under intrinsic (such as aging) or external environmental disturbance (such as diet)^{8,17,18} regardless of observable ecological changes. It is likely that evolutionary responses buffered a variety of environmental changes and further exerted a long-term effect on the population dynamics of evolving species^{17,19}. While probiotics were proposed to modulate the gut ecosystem for digestive health^{3,6,20}, they can also have a notable impact on the *in vivo* evolutionary trajectories or functions of those residents regardless of whether probiotics colonize or transiently pass through the gut¹⁴. However, few studies have systematically assessed risks in the evolution of native gut microbiota under the selection of ingested probiotics.

To address these questions, we employed *Lactobacillus plantarum* HNU082 (Lp082), which was isolated from traditional fermented food²¹ and whole-genome sequenced (PRJCA000348, PRJNA637783), as a model probiotic strain²². It has gained increasing attention as its conventional characteristics of probiotic *Lactobacillus plantarum*^{12,22} and specific functions such as hyperlipidemia prevention²³ and regulation of neurotransmitter secretion disorder. Here, we explored the effects of host-derived selection pressures (humans, mice, and zebrafish) on the genetic stability of probiotics and, in turn, its ecological and evolutionary impact on the indigenous gut microbiota using shotgun metagenomic sequencing and isolate resequencing methods.

Results

The adaptive evolution of probiotics within the gut of distinct hosts

We employed the standard reference-based approach to explore the genetic changes of the consumed probiotics Lp082 under the *in vivo* natural selection in the gut of multiple hosts (humans, mice and zebrafish) (**Fig. 1A**). We sequenced the complete genome of this reference probiotic strain, including one chromosome and four plasmids. Next, we isolated this strain from the feces of hosts at different time points for whole-genome sequencing. To identify and quantify the putative genetic mutations (such as single nucleotide variants) in the host-adapted strains, we compared the genome of all isolates with the reference genome. In total, 109 bacterial strains of Lp082 were isolated from feces or intestine content of three hosts, out of which 77 isolates were from humans, 25 from mice, and only 7 from zebrafish for the duration of the whole experiment (**Fig. 1B, 1C, Table S1**). A total of 71 putative single nucleotide variants (SNPs) and 2 mobile genetic elements (**Fig. S1B**) were identified and annotated from genome sequencing data of *in vivo*-adapted probiotic strains in human, mouse, and zebrafish models, out of which only 22

SNPs could be experimentally verified using PCR (**Fig. 1D**). By contrast, under the *in vitro* condition, no SNP was annotated in Lp082 incubated in the de Man, Rogosa and Sharpe (MRS) agar within 28 days (**Fig. S1A**), suggesting these adaptive mutations only occurred during the gut passage of this probiotic.

We next sought to characterize *in vivo* adaptive mutations of Lp082 across host models. First, despite the radically distinct host selection pressures, we did not identify any host-specific SNPs in the genome of Lp082. Furthermore, no between-host difference in the accumulated mutations, that is, average number of SNPs at the last time point was observed (human=9.93; mouse=10; zebrafish=9; $p=0.926$, Kruskal-Wallis test) as well as mutation types (**Fig. 1C**).

We next profiled and compared the temporal evolutionary dynamics of Lp082 over a four-weeks sampling period between the humans and mice (**Fig 1E, 1F, S1C-1G**). As for all hosts, the mutational frequency of Lp082 steadily increased in the initial two weeks of gut colonization, suggesting the fitness advantages of lineages were carrying them relative to their ancestors. Since relatively few new SNPs occurred, the mutation rate tends to stabilize after the third week of colonization (**Fig. 1C**). Based on these adaptive mutations, we employed the phylogeny to identify five lineages of the probiotic Lp082, which commonly emerged in humans and mice (**Fig. 1E**). We further sought to investigate how similar the evolutionary history of these co-existing probiotic lineages can be between two hosts (**Fig. 1E, F**). Overall, the temporal dynamics pattern of the probiotic lineages throughout the sampling period was highly conservative between humans and mice ($p=1e-4$, $R=0.86$, Mantel test based on Jaccard distance). Despite its occupation in the ecologically and spatially distinct niches, Lp082 acquired highly similar mutations in almost the same order and timing for the gut adaptation. At the initial stage of colonization, the sub-lineage E1 firstly emerged by acquiring a mutation (SNP41) on the gene encoding IgG binding protein (Gene 2601). In the following sampling period, this lineage co-existed with other descendants with a consistently high frequency in the Lp082 population (>50%) (**Fig. 1E, F, Table 1**), suggesting that Lp082 should persistently resist the host immunity for establishing and maintaining its ecological niche in the gut. Next, the lineages E1-A and E1-B appeared and maintained their proportion in the Lp082 population in the gut throughout the sampling period. The most notable adaptive mutations (SNP10, 47, and 32) from these lineages were involved in the genes (Gene1717, 0658 and 2804) encoding inner membrane protein response to acid pH, transcriptional activator for 3-phenylpropionic acid catabolism, and transpose, indicating Lp082 was actively developing the acid-tolerance capability *in vivo* by adaptive mutations. Around 7 days after probiotic ingestion, E1-B-1 emerged and carried the representative mutations (e.g., SNP 23) that can enhance the capability of rhamnose utilization (Gene 1257). Interestingly, this sub-lineage vanished at the end of our sampling period, and survived slightly longer in mice (Day 28) than in human gut (Day 21), which is the only observable difference in the probiotic evolutionary history in the gut between two hosts.

We next wondered if or what genes in the probiotic genome had a similar pattern of adaptive mutations across independent lineages or host subjects. Firstly, we focused on those genes that have at least two SNPs (and their distance on the genome is no more than 2000bp) in the independent lineages from distinct host subjects or host species, indicating that they underwent parallel evolution. Such adaptive

mutations identified from a gene are not likely a consequence of the genetic drifts, indicating these genes are under natural selection. Among all 12 mutated genes, we identified five such genes from all 109 isolates (encoding trehalose 6-phosphate phosphorylase, transposase, and ABC transporter ATP-binding protein) (**Fig. 2A**). Remarkably, these genes carried consistent multiple mutations across most host subjects of humans, mice and zebrafish (**Fig. 2A, Table S3**). Next, we identified the other seven genes as single-mutation genes involved in bacterial alpha-L-rhamnosidase 6 hairpin glycosidase, immunoglobulin G-binding protein, glutathione reductase, inner membrane protein response to acidic pH, and transcriptional activator for 3-phenylpropionic acid catabolism (**Fig. 2B**). Collectively, the presence/absence patterns of adaptive mutations in each of 12 genes are highly consistent within and between host species (**Fig. 2B**), strongly demonstrating a universal adaptive strategy of this probiotic strain under varying selection pressures.

We next tested if those SNPs conferred the fitness advantage of Lp082 in the host gut using the *in vitro* model. We selectively validated the phenotypic changes of the host-adapted isolates related to the rhamnose utilization and acid tolerance (**Fig. 2C, S2B**). Firstly, we focused on a non-synonymous SNP in Gene 1257 (annotated as bacterial alpha-L-rhamnosidase 6 hairpin glycosidase) which was identified in 21 host-adapted isolates. Intriguingly, all these mutants exhibited a higher capability in utilizing rhamnose *in vitro* than the original strain. While those mutants rapidly responded to the rhamnose substrate in the medium, the reference strain did not grow in the early 3 hours after its inoculation, suggesting this SNP had conferred a significant fitness improvement of Lp082 after the gut passage. Furthermore, mutational isolates from humans and mice exhibited a highly consistent growth pattern. The human-derived strains just have a slightly longer period of exponential phase than did mouse-derived strains. Next, we tested the acid tolerance of 78 mutational isolates with a synonymous SNP in Gene 1717 (encoding inner membrane protein response to acidic pH) independently. Among those 78 mutants, 16 (20.5%) exhibited a significant higher survival rate under each of low pH conditions ($p < 0.05$, pH=3.5, 2.5, 2.0 respectively) than the reference strain. Notably, the *in vitro* experiments of these isolates were conducted 9 month after their passage through the gut of hosts, suggesting that these gut-adaptive mutations and related phenotypic changes can be stably inherited and preserved in a relatively long time.

Compositional and functional alterations of resident gut microbiota in response to probiotic engraftment

We next examined how the resident gut microbiota ecologically responded to probiotic engraftment. Compared to the placebo group, the ingestion of Lp082 introduced a notable fluctuation in the taxonomic structure of indigenous microbial communities in both humans and mice (**Fig. 3A, S3A, S3B**). As for humans, the probiotic intake did not significantly alter the overall taxonomic and functional composition of gut microbiome in any time points compared to baseline ($R^2=0.088$, $p=0.979$; PERMANOVA, **Fig. 3A**). Such a low probiotic-derived impact on the microbial compositions of the human gut may result from a higher level of individuality or resilience in the human gut microbiota prior to ingestion. By contrast, we observed a strong impact of probiotics on the structure and function of mouse gut microbiota ($R^2=0.28$, $p=0.005$; PERMANOVA, **Fig. 3A, S3C, S3D**). Compared to the baseline, the structure of the mouse gut microbiome significantly changed on Day 7, 14 and 28 after probiotic ingestion ($R^2=0.28$, $p=0.005$ on

Day 14; $R^2=0.19$, $p=0.039$ at Day 21; $R^2=0.59$, $p=0.004$ at Day 28; PERMANOVA). All these observations were confirmed by both Aitchison (**Fig. 3**) and Bray-Curtis distance metrics (**Fig. S3A, S3B**).

Next, we sought to identify metagenomic markers associated with probiotic intake. Specifically, we defined a metagenomic marker according to the following conditions: 1) it significantly changed over time (e.g., from Day 0 to 14) in the probiotic group; 2) it maintains stable abundance during the experiment in the placebo group; 3) it was significantly different in abundance at a given time point (such Day 14) between the two groups. In the human population, no organismal or functional markers were identified based on these criteria which may be due to large inter-individual variations in the gut microbiome. In the mouse model, *Faecalibaculum rodentium* and 10 species from *Bacteroides* (including *Bacteroides caccae*, *Bacteroides ovatus*, *Bacteroides vulgatus*, *Bacteroides xylanisolvens*, etc.) increased significantly, while other 12 species including *Prevotella dentalis*, *Prevotella oris* and *Pseudomonas fluorescens* decreased sharply after probiotic ingestion (**Fig. S3C**). Meanwhile, we also identified a total of 21 metabolic pathways (including D-galacturonate degradation, pyruvate fermentation to acetate and lactate, starch degradation, tetrapyrrole biosynthesis, L-citrulline biosynthesis, glycogen biosynthesis and L-isoleucine biosynthesis) enriched in the time points during probiotic colonization (**Fig. S3D**).

In order to compare microbiome responses to probiotic colonization across host species, we here focused on 19 microbial species commonly found in both mammalian hosts, which taxonomically belong to the genus of *Bacteroides* or *Bifidobacterium* (**Fig. 3C**). We found that these 19 microbial species residing the human and mouse gut had distinct temporal dynamics in response to probiotics ingestion. This suggests that the ecological response of resident microbes to probiotic engraftment highly depends on the host or gut microbiota context (**Fig. 3B**).

Co-occurrence relationship between Lp082 and resident microbes in the gut of humans and mice

We next constructed a co-occurrence network in each host species to identify the correlation between resident microbes and the ingested probiotic. The co-occurrence networks were constructed using SpiecEasi²⁴ based on the species-level taxonomic profiles derived from shotgun metagenomic sequencing data (Fig. 3C). In the human model, we identified 10 species whose abundances strongly correlated with that of Lp082: *Alistipes finegoldii* and *Odoribacter splanchnicus* positively correlated with Lp082, while several *Bacteroides* spp. and *Bifidobacterium* spp. (*Bacteroides intestinalis*, *Bacteroides salanitronis*, *Bacteroides uniformis*, *Bacteroides thetaiotaomicron*, *Bifidobacterium longum*, *Bacteroides xylanisolvens*, and *Bifidobacterium adolescentis*) negatively correlated with Lp082. In the mouse model, we identified 4 species (*Bacteroides caccae*, *Bacteroides vulgatus*, *Bacteroides ovatus*, and *Bacteroides xylanisolvens*) strongly correlated with Lp082, and all the 4 species negatively correlated with Lp082. We found that no common resident gut microbes across host species showed consistent correlation with Lp082. Except for *Bacteroides xylanisolvens*, there were no common negatively correlated bacterial species found between the human and mouse models (**Fig. C**). While most microbial competitors (with the exception of *Bacteroides xylanisolvens*) of Lp082 in the mouse model were also identified in the human gut, none of them exhibited the same co-occurrence relationship in the network analysis.

Profound evolutionary changes within the resident gut microbiota in response to probiotic engraftment

Probiotic intake can give rise to dramatic changes in the genetic compositions in the resident species, yet it was surprisingly underreported. Here, we studied the within-genome evolution of approximately 37 prevalent microbial strains in the human or mouse gut. Firstly, we observed a striking difference in the overall number of adaptive mutations in the resident microbiota between placebo and probiotics groups (**Fig. 4A**). The average number of mutations that occurred in the placebo group is only 1.92 (or 1.69) for humans (or mice), respectively. By contrast, probiotic intake gave rise to a dramatic change in mutation frequency in the resident gut microbes, which reached up to 16.90 (or 78.02) per species in the human (or mouse) model, respectively (**Fig. 4B, Fig. S4A**). It indicated that a resident strain can accumulate in average 4.33 (or 7.67) SNPs per day in humans (or mice), respectively, which was remarkable given the high inter-individual variation in the gut microbiota of both humans and mice. Conversely, the genome of Lp082 was relatively stable: only 8 and 10 SNPs were detected on Day 14 and 28 respectively (**Fig. 4B**). We next asked if these adaptive behaviors of residents might give rise to compositional or functional changes in the communities. In the mouse gut, the more mutations occurred in resident gut microbiota, the more pronounced shifts can be observed in the taxonomic or functional profiles after probiotic ingestion (Mantel test; **Fig. S4B**). This suggested that adaptive mutations could be a strong driving force reshaping the resident gut microbiota.

We next sought to characterize the distribution of these adaptive mutations among resident microbial strains. We found that the majority of mutations distributed in *Bacteroides* spp. and *Bifidobacterium* spp., which were mostly identified as ecological competitors of Lp082 in the resident gut microbiota of either humans or mice. We further demonstrated that the putative ecological competitors (such as *Bacteroides* spp. and *Bifidobacterium* spp.) proactively responded to dietary probiotics by accumulating significantly more adaptive mutations than non-competitors (i.e. other bystanders or facilitators) of this probiotic ($p < 0.05$, Wilcoxon rank-sum test, **Fig. 4C, S4A**). Remarkably, in the mouse model, mutations occurred in probiotic competitors were one to two orders of magnitude greater than those in the related species in the placebo group or the probiotic group over 28 days. Therefore, we provided the clue about the potentially causal links between evolutionary and population dynamics in the resident gut microbiota.

We next characterized how evolutionary dynamics in resident gut microbiota differed between humans and mice. Firstly, the resident microbes in the mouse gut accumulated much more adaptive mutations than those in the human gut (**Fig. 4C, D**). For example, *Bacteroides xylanisolvens* is one of the overlapped members in the gut microbiota of both humans and mice, being negatively correlated with Lp082 in both hosts. After 14 days of probiotic ingestion, we identified on average 596 within-host SNPs in the mouse gut microbiotas, whereas only 15 SNPs identified on average for a human individual in the gut microbiota (**Fig. 4D**). These may suggest probiotics intake imposes more intensive selection pressure to resident gut microbiota in the mouse than does in the human. Next, we observed a substantial functional difference in the adaptive-mutated genes of the resident gut microbiota between human and mouse models (**Table S9**). The human-associated mutations primarily occurred in the genes related to cAMP-binding proteins and conjugative transposon protein, whereas mouse-associated mutations involved in more diverse

functions, such as the degradation and utilization of the complex carbohydrates (Beta-galactosidase, Phosphoglycolate phosphatase, 3-oxoacyl-[acyl-carrier protein] reductase, and Galactoside O-acetyltransferase), Tetracycline resistance element mobilization regulatory protein, Asparaginyl-tRNA synthetase, DNA primase, conjugative transposon protein, and Transposase. There was only one gene (traG) that was consistently mutated in both the human and the mouse models, which related to the conjugative transfer of plasmid RP4 (TraG-like proteins are essential components of type IV secretion systems²⁵). Collectively, even though the engrafted probiotic applied universal strategy in the distinct hosts, the indigenous microbiome response to its colonization was highly divergent, underscoring the importance of careful interpretation of ecological and evolutionary patterns identified in the mouse model for human-related translational studies.

Discussion

Probiotics are live microorganisms that, in sufficient dose, confer a benefit to the host, and are widely used for improving digestive health. However, their genomes and functional traits can vary during administration due to the myriad selection pressures in the host luminal environment²⁶. The gastrointestinal tract harbors a large collection of microorganisms. The resident gut microbes are typically well adapted to their host, including the immune system and fluctuations in environmental resources over a relatively long period, resulting in higher competitive fitness over any potential invaders¹⁷. Changes in environmental selection pressures, such as competitions for limited carbon sources, force new microbes to dramatically reshape their genome and functional characteristics in a short-time scale.

In the present study, we demonstrated how the probiotic *Lactobacillus plantarum* HNU082 (Lp082) can evolve within the human and mouse gut by acquiring heritable changes to its genome. We found that Lp082 accumulated single-nucleotide mutations that modulate carbon utilization (i.e. rhamnose) and acid tolerance in two mammalian host models. The significant improvement in rhamnose utilization can be attributed to a nonsynonymous mutation in the gene related to bacterial alpha-L-rhamnosidase 6 hairpin glycosidase. Furthermore, the findings also indicated that rhamnose can serve as a strain-specific prebiotic²⁷ that enhances the survival of Lp082 in the host gut. More interestingly, a synonymous mutation in the gene related to the function of inner membrane protein might respond to acidic pH. 20.5% (N=16) of 78 mutational strains with this SNP had better performance in acid tolerance. Similarly, Crook *et al.* reported that probiotic *E. coli* Nissle gained *in vivo* fitness by accumulating mutations related to carbohydrate utilization and acid tolerance within the mouse gut microbiome³. As a gram-negative bacterial probiotic, *E. coli* Nissle acquired 456 (336 unique) evolutionary changes across 171 genes within 35 days, while Lp082, a gram-positive bacterial probiotic, only accumulated 22 SNPs across 12 genes within 28 days. This may suggest higher genome flexibility or instability of a gram-negative probiotic over gram-positive ones (such as conventional lactic acid bacteria) under *in vivo* selection.

Knowledge of normal genomic variations within and between host species can further allow us to develop new probiotics or modify existing strains that can engraft and colonize a human host, adding desired functions, safely, effectively, and reliably^{6,7}. For example, it is translational promising if probiotics can be genetically engineered by means of experimental evolution in the animal hosts where *in vivo* fitness can be improved under the human gut selection pressure³. While many animal models (e.g. mice, flies, and zebrafish) have been employed to study adaptive evolution of probiotics^{11,12}, very few studies have applied animal-gut-adapted probiotics to humans. There are several reasons. Firstly, it remains challenging to conduct in-deep study of *in vivo* adaptive evolutionary behaviors of individual probiotic strains. The probiotic characteristics are highly strain-specific, the *in vivo* evolutionary behaviors can be more divergent. Unfortunately, the Next, the *in vivo* adaptation can be shaped by multiple intrinsic and extrinsic factors and thus highly unpredictable. A systematic investigation is required in the diverse schemes of probiotic strains, microbiome diversity levels, diet, and host factors. Crook *et al.* compared the genome adaptation of *E. coli* Nissle (EcN) in the host gut with different microbiome complexities³. As a gut commensal, this strain accumulated fewer mutations in the “high diversity” gut microbiome, while its genome varied more in the “low diversity” microbiome. By contrast, we found that Lp082 maintained a much higher genetic stability than EcN regardless of the radical differences in the microbiome complexity between human and mouse gut. Other than that, its adaptive evolutionary changes were not differentiated by the host factors from humans and mice. We argue that the universal adaptive strategy of Lp082 against selection pressures from multiple hosts could suggest the transferability of the fitness advantages of this probiotic from mice to humans, which definitely merits the further investigation.

Probiotics adapt and evolve *in vivo* in order to survive. They also influence selection pressures on resident microbial strains. Here we systematically investigated the ecological and evolutionary impacts of a probiotic *Lactobacillus plantarum* on the resident gut microbiome in humans and mice. First, we observed a minimal ecological change (either compositional or functional changes) in the human gut due to probiotic ingestion which is consistent with previous studies^{15,28}. By contrast, the resident population of gut bacteria can rapidly and extensively evolve within 3-7 days after probiotic administration in both humans and mice. Most within-genome changes persisted for 28 days in our experiment. In the placebo group, very few evolutionary changes in resident microbes were found over time. Notably, strains competing with probiotic Lp082 account for the largest proportion (Human:74.36%; Mouse:77.68%) of evolutionary events in the whole resident gut microbial community. As described by the “Red Queen hypothesis”⁶, microbial competitions within our gut ecosystems can force the acceleration of microbial evolution possibly without any observable ecological changes, resulting in apparent stability. It also suggested that the common or even daily supplement of probiotics could rapidly become a strong driving force of evolution and ecology in the resident gut microbiome within a relatively short timescale, which has been largely overlooked. Given the substantial number of adaptive mutations accumulated, it is still challenging to identify any worrying ones in the abundant resident gut microbes, which have potential to alter the gut ecological features in the long-term run. Therefore, the ecological and evolutionary impact of dietary probiotics on gut microbiota should be seriously considered prior to administration²⁹. Accordingly, future work should carefully reflect the limitations in the present evaluation system of probiotic efficacy

and safety in order to better design a framework for systematic assessment of interactions between probiotics, gut microbiome, and host over time.

Methods And Materials

The experimental design

In this study, we used Lp082 as a model probiotic strain to explore the effects of host-derived selection pressure (the mouse and human models were applied, the zebrafish model was used for further verification of adaptive mutations) on the genetic stability of the consumed probiotics and the impact of probiotic mutations on the indigenous gut microbiome of different hosts (the mouse and human model). First, we sequenced the complete genome of this model probiotic strain, including one chromosome and four plasmids. Secondly, we isolated the probiotics from the feces of hosts at different time points to identify genetic mutations using whole-genome resequencing. Simultaneously, the original strain was continuously inoculated *in vitro* and sequenced to assess genetic mutation in the absence of host selective pressure. Next, we employed the metagenomic sequencing method to characterize temporal dynamics of the abundance of probiotics strain and other gut-microbiota members and genetic variations in the community level after probiotic ingestion and confirming the impacts by comparing the results with that in the placebo group. Last but not least, the difference in genetic variations of Lp082 and its impact on the indigenous gut microbiota among hosts were highlighted in this study.

In the mice (C57BL/6, 5 weeks age) experiments, each animal was housed in a single cage. Room temperature was 26°C and the padding was replaced once every day. All mice were divided into two groups, the probiotic group ($n=6$), and the placebo group ($n=5$). For the mice in the probiotic group, Lp082 (about 4×10^8 CFU/g with the fodder) was infused daily for 7 days. Fresh feces were collected from each cage for bacterial isolation on days 3, 7, 14, 21, and 28 after stopping probiotic feeding. Three pieces of fresh excrement were placed in a 7 mL sample tube, 4.5 ml saline (0.85%) was added after sterilization and then homogenized with homogenizer. Then the fecal samples were diluted and coated for Lp082 isolation as well as for shotgun metagenomic sequencing. For the mice in the placebo group, the same feeding method was performed only without the probiotic infused. The fresh feces were collected from each cage on days 0, 14 and 28 for shotgun metagenomic sequencing.

For human participants, each individual was informed of the experimental guidelines and details and consent obtained; 12 volunteers agreed to participate in the experiment. They were randomly divided into 2 groups including the probiotic group ($n=7$, 4 females) and the placebo group ($n=5$, 2 females). During the experiment, the subjects were asked to avoid ingesting any probiotic product or antibiotic and to maintain their regular diet. All healthy participants finished the whole experiment including 6 females and 6 males aged from 18-20 (BMI 19.19-22.49), they did not have inflammatory bowel disease or diabetes and had not used antibiotics for at least 3 months prior to sampling. For the seven volunteers in the probiotic group, they were asked to consume vacuum freeze-drying Lp082 powder (including 7×10^9 CFU live strains) 2g every day for 7 days³⁰, and fecal samples were collected on days 3, 7, 14, 21 and 28 after

stopping probiotic consumption. Fresh feces were placed in a sterile fecal sampling tube, then the diluted feces samples spread to the MRS agar plate for Lp082 isolation as well as for shotgun metagenomic sequencing. In the placebo group, the five volunteers were requested to maintain a regular diet during the whole experiment, and their fresh feces were collected on days 0, 14, and 28 for shotgun metagenomic sequencing. The study was reviewed and approved by the Ethics Committee of Hainan University (HNU-2018037, Haikou, China), and informed consent was obtained from all volunteers before they enrolled in the study. The participants provided written informed consent to participate in the study. Sampling and all described subsequent steps were conducted in accordance with the approved guidelines.

Additionally, a zebrafish model was used to validate our findings related to the impacts of host intestinal selective pressure on the genetic stability of the ingested probiotic. A total of 20 fish (15 weeks age) tanks (15-18 zebrafish in each tank) were used in this experiment. The water was changed and the fish were fed at 9:00 am daily, the capacity of feed was calculated according to 3% of the body weight of each fish per day. After the adaptation period, *Lp082* of 10^8 CFU/g was fed for 7 days with the fodder²³. The intestines of 5 fish were taken and homogenized with 1 mL saline (0.85%) in 1.5 mL EP tube and in 3 days, 7 days, 14 days, 21 days and 28 days after stopping probiotic feeding. Then the intestine samples were diluted and coated for *Lp082* isolation.

Isolation and confirmation of the ingested probiotic strain (Lp082) in the feces

The reference strain Lp082 used in the present study was first isolated in traditional fermented food in Hainan province of China²¹ and further certified as a probiotic due to the common characteristics of probiotic *Lactobacillus* spp. and specific functions such as hyperlipidemia prevention²³ and neurotransmitter secretion disorder regulation. We obtained the complete genome of the strain with functional annotations in our previous research (PRJCA000348, PRJNA637783). By comparing with other genome sequenced *Lactobacillus plantarum* strains (Table S4), we identified strain-specific primers of Lp082, which was used for genetic confirmation of isolates in the next step. The detailed experimental steps for Lp082 isolation from feces and confirmation were as follows:

(1) Feces/intestinal contents were taken from the sample tube, and sterilized saline (0.85%) was added to the sample tube. After stirring with a homogenizer, the samples were diluted and coated. The mixture of 0.5 mL and 4.5 mL of 0.85% saline (NaCl) was recorded as 10-1, 10-2, 10-3, 10-4, 10-5, 10-6 were obtained.

(2) Each diluted gradient was selected for coating. About 100 μ L mixed liquid was taken on solid de Man, Rogosa and Sharpe (MRS) medium and 50 μ L 1280 μ g/mL vancomycin solution and 50 μ L 1280 μ g/mL norfloxacin solution was added evenly. Date, number, weeks and concentration were marked on a petri dish, which was placed in an incubator at 37 °C for 48h.

(3) The single colony cultured in the medium was selected and again cultured in the test tube of 5 mL MRS broth medium. After being shaken uniformly, the bacteria were cultured in an incubator at 37 °C for 48h.

(4) Colony PCR was used to validate the bacterial solution by strain-specific primers. Samples that could be amplified were preserved and whole-genome sequenced for further validation.

Isolate whole genome sequencing and data quality control.

Bacterial genomic DNA was extracted from the isolates for whole-genome sequencing. Using high-throughput sequencing, paired-end reads (2 x 150 bp) library of each single bacteria sample prepared on the Illumina HiSeq 2500 platform in Shanghai Personal Biotechnology company (Shanghai, China). Quality control was carried out with FastQC, AdapterRemoval (v2.1.7) was used to remove the joint contamination³¹ and SOAPec (v2.0) software was used to carry out quality correction based on Kmer frequency³².

SNP-Calling and construction of the phylogenetic tree of mutants

To estimate the evolutionary distance between isolates across the 3 models and phylogenetic tree construction, we aligned all short reads to the reference genome of *Lactobacillus plantarum* HNU082 for SNPs identification. Reads were aligned using Bowtie2³³ (Alignment parameters: bowtie2 -p -x -no-mixed -very-sensitive -n-ceil 0,0.01 -1 -2 | SAMtools sort -O bam -@ 24 -o - > *.bam). Candidate SNPs were identified and filtered with SAMtools³⁴. In particular, genomic positions were considered to be candidate SNP positions if at least one pair of isolates were discordant on the called base and both members of the pair had: FQ scores (produced by Bftools) less than 60, at least 7 reads that aligned to each of the forward strands and reverse strand and a major allele frequency of at least 90%. If the median coverage across samples at a candidate position was less than 10 reads or if 33% or more of the isolates failed to meet the filters described above, this position was discarded⁸. We generated a neighbor-joining tree from the concatenated list of variable positions from conserved genomic regions present in all isolates from all samples by MEGA-X and iTOL software³⁵. When computing the distance between each pair of isolates, we only used variable positions that had unambiguous nucleotide calls from both isolates.

Calculating the distance to the most recent common ancestor (dMRCA) and the mutation type

To calculate dMRCA for each model at each time point, we counted the number of positions at which the called allele was different from the ancestral allele for each isolate, assessing only SNP positions that were polymorphic among isolates from the particular time point, and averaged the results. SNPs were categorized into 6 types, based on the chemical nature of the single nucleotide changes. we computed the mutation spectrum for each model and then computed the mean and standard deviation of each of the 6 types. The frequency of the mutation type of G-C to A-T was the most abundant in all three models, but the frequency of the mutation type of A-T to G-C was significantly higher in the zebrafish model than that in the other two models.

Identification of the mobile elements of Lp082 during ingestion

To identify mobile elements of the strain Lp082, we performed the pan-genome analysis of all strains (N=109) isolated in this study, including the original reference strain and all isolates from different hosts and time points. In general, the whole genomic sequencing reads of each isolate was assembled by SPAdes with the parameters as below: `spades.py -t 24 -k 21,33,55,77,89 –careful –only-assembler`³⁶. To delineate the core genome of all strains, we mapped the assembled contigs (including only those with the length of >500 bp) of each strain against the reference strain of Lp082 using BLASTn³⁷ with identity levels $\geq 90\%$ and an e-value $< 1e-5$. Then, we obtained a set of strain-specific contigs by excluding the contigs that mapped to the core-genome. Further, to remove the potential redundancy in the strain-specific sequences, we pairwise compared these sequences using BLAT³⁸ with an identity level $\geq 90\%$ and a match length $\geq 85\%$, which resulted in a set of non-redundant contigs specific to strains that can be also termed as the accessory genome. We reason that all the contigs in the accessory genome of this strain can be inferred as the “mobile elements” inserted into the reference genome (Lp082) during microbial colonization.

To investigate the presence/absence of each potential mobile element in a given isolate, we calculated the read coverage and depth of each isolate against the accessory genome: coverage >80% and depth >60x were considered to be present, while coverage <20% or depth <60x were regarded as absent³⁹. Accordingly, we calculated the coverage and depth of each mobile element in each isolate (**Table S5**).

Identification of the genes with parallel evolution among the models

We counted a gene as under parallel evolution if, in at least one host, the gene had multiple independent SNPs and more than 1 SNP per 2,000 bp (to account for the fact that long genes are more likely to be mutated multiple times by chance). Cases in which two

SNPs in the same gene that always occurred together in the same isolates were not included as parallel evolution⁸. based on these principles, a total of 5 genes related to carbohydrate utilization and transposase were identified under parallel evolution.

Evolutionary dynamics of the strain Lp082 in the gut

While too few probiotic isolates ($n=7$) in the zebrafish model over the 28-days experiment, we only characterized the evolutionary dynamics of the probiotic strain within the human and mouse model. To increase the reproducibility of SNPs studied, we only focused on SNPs that should be identified from at least 4 isolates. For each of the 22 SNPs that met this criterion, we calculated the frequency of reads at each SNP position that agreed with the mutation (derived) allele. To fill the time points where no strain was isolated, we generated a continuous relative abundance of SNPs over time by continuous bezier interpolation⁴³. To confirm the parent and child relationship of these SNPs, the phylogenetic trees were constructed for all strains isolated in the human and mouse model. To visualize parent and child lineages separately, we subtracted the relative abundance of a parent sublineage by the sum of relative abundances of its child sublineages. When the combined relative abundance of child sublineages

exceeded that of their parent sublineage, we set the frequency of the parent sublineage to 0. At last, the muller plot was employed for dynamic visualization⁴⁰. We did not build the model of the evolutionary dynamics of this probiotic in the zebrafish model, as too few isolates (only seven) were found.

Metagenomic DNA extraction and shotgun metagenomic sequencing and data quality control

The QIAamp® DNA Stool Mini Kit (Qiagen, Hilden, Germany) was used for DNA extraction from the fecal samples. The quality of the extracted DNA was assessed by 0.8% agarose gel electrophoresis, and the OD 260/280 was measured by spectrophotometry. All of the DNA samples were subjected to shotgun metagenomic sequencing by using an Illumina HiSeq 2500 instrument in the Novogene Company (Beijing, China). Libraries were prepared with the paired-end reads (2 x 150 bp). The raw reads were trimmed using Sickle(<https://github.com/najoshi/sickle>) and subsequently aligned to the human genome to remove the host DNA fragments.

Identification of metagenomic species, microbial functional genes, and metabolic pathway annotation

Bracken⁴¹ was applied for metagenomic species identification and abundance estimation (Table S6). For metagenomic functional features and metabolic pathway annotation (Table S7), HUMAnN2 was performed by using the UniRef90 database⁴². More information was listed in “code availability”. Accordingly, we got the relative abundances of intestinal microbial taxonomic compositions, gene families and metabolic pathways, respectively. To perform the differential abundance analysis, we introduced a compositionality-aware statistical method as Morton et al demonstrated⁴³. In the microbiota analysis, we performed a widely-used centered log ratio of microbial relative abundance to verify our analysis results. The R package “composition” was applied to implement the centered log-ratio (CLR) transformation for raw relative abundance profiles.

Co-evolution analysis based on shotgun metagenomic data of gut microbiota

Based on the correlation of the strain Lp082 with the indigenous microbes, we employed the MIDAS (Metagenomic Intra-Species Diversity Analysis System) to perform intestinal microbiota mutations annotation⁴⁴. Briefly, a reference genome database including 33 species with the abundance more than 0.1% and the species closely related to Lp082 was constructed. Then we mapped the shotgun metagenomic sequencing reads from each microbiome to the reference bacterial genomic database and quantified nucleotide variation along the entire genome (i.e. the read depth and observed alleles at each position). The data from multiple samples was used for SNP calling of individual intestinal species. The samples at baseline in each host were set as the reference for the calculation of bacterial mutations occurred within hosts at other time points. The SNPs profiles for these intestinal microbes among the human and mouse models both in probiotic and placebo groups (**Table S8**).

Experimental verification of SNPs and the phenotypic changes associated with SNPs

For SNPs verification, we retrieved the upstream and downstream 100-bp sequences of

each SNP in the final set and designed primers (**Table S3**) for PCR amplification. The Sanger sequencing results of PCR products were used to verify the nucleotide status of SNPs acquired from the *in silico* analysis. After verification, 22 SNPs out of 71 putative SNPs were finally confirmed.

Further, we examined the potential phenotypic improvements of this strain caused by the selective single-nucleotide mutations *in vitro*, including the ability of rhamnose utilization, trehalose utilization (**Fig. S2C**) and acid tolerance. For the polysaccharides utilization experiments, the rhamnose/trehalose was set as the only source of carbohydrate in broth and the bacterial growth conditions including growth rate and the number of colonies were calculated. For the acid tolerance experiment, the mutational strains and the ancient Lp082 strain were inoculated into the broth with pH values of 2, 2.5, 3.0, and 7.0 respectively, and the strain survival rates were calculated.

Statistical analysis

All statistical analyses were performed using R software (version 3.5.1). PCoA analysis was performed in R using the *ade4* package. CLR transformation was performed by the “*zcomposition*” package. The heatmap was constructed using the “*pheatmap*” package, and the evolutionary dynamics were built using the “*ggmuller*” package. The differential abundances of various profiles were tested with the Wilcoxon rank-sum test and were considered significantly different at $p < 0.05$. For boxplot construction, the package “*ggpubr*” was used. The edges of the network were calculated by the “*SpiecEasi*” co-occurrence correlations²⁴ and visualized in Cytoscape (Version 3.7.1). For more details, please refer to the document “Coevolution project analysis code and figure reproducibility”.

Data and code availability

The sequence data reported in this paper have been deposited in the NCBI database (genome resequencing and metagenomic sequencing data: PRJNA594992, PRJNA597371, PRJNA590026, PRJNA597969 and PRJNA588621). All the project analysis code had been deposited in Github:<https://github.com/zhjch321123/Host-dependent-co-evolution-of-supplemented-probiotic>

Declarations

Acknowledgments

The project was supported by the National Natural Science Foundation of China (No. 31701577). CA is supported by NIH T32 OD017863. AZ is supported by NIH K08 DK102902, R21 MH117780, R01 HL148801. This work was supported in part by NIDDK grant P30 DK120515, P30 DK063491, and UL1 TR001442 and NCCIH grant DP1-AT010885. This work was also supported by IBM Research through the AI Horizons Network, UC San Diego AI for Healthy Living program in partnership with the UC San Diego Center for Microbiome Innovation.

Author Contributions

The research topic was developed by J.Z. and R.K. The experiment was performed by S.J., D.H. and C.M. Data collection was performed by D.H. and S.J. Data analysis was performed by S.H., J.Z., S.J., C.A., P.B., Y.V. and Q.Z. The manuscript was written and revised by S.H., J.Z., C.A., M.E., A.Z., V.C., P.B., Y.V., Y.L. and R.K. All authors read and approved the final manuscript for submission.

Competing interest statement

AZ is a founder and equity-holder in Tortuga Biosciences.

References

- 1 Hill, C. *et al.* Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepato***11**, 506-514, doi:10.1038/nrgastro.2014.66 (2014).
- 2 Mallon, C. A., Elsas, J. D. V. & Salles, J. F. Microbial invasions: the process, patterns, and mechanisms. *Trends Microbiol***23**, 719-729, doi:10.1016/j.tim.2015.07.013 (2015).
- 3 Crook, N. *et al.* Adaptive Strategies of the Candidate Probiotic E. coli Nissle in the Mammalian Gut. *Cell Host Microbe***25**, 499-512 e498, doi:10.1016/j.chom.2019.02.005 (2019).
- 4 Pacheco, A. R. *et al.* Fucose sensing regulates bacterial intestinal colonization. *Nature***492**, 113-117, doi:10.1038/nature11623 (2012).
- 5 Leatham, M. P. *et al.* Precolonized human commensal Escherichia coli strains serve as a barrier to E. coli O157:H7 growth in the streptomycin-treated mouse intestine. *Infect Immun***77**, 2876-2886, doi:10.1128/IAI.00059-09 (2009).
- 6 Ferreira, A., Crook, N., Gasparini, A. J. & Dantas, G. Multiscale Evolutionary Dynamics of Host-Associated Microbiomes. *Cell***172**, 1216-1227, doi:10.1016/j.cell.2018.02.015 (2018).
- 7 Yelin, I. *et al.* Genomic and epidemiological evidence of bacterial transmission from probiotic capsule to blood in ICU patients. *Nat Med***25**, 1728-1732, doi:10.1038/s41591-019-0626-9 (2019).
- 8 Zhao, S. *et al.* Adaptive Evolution within Gut Microbiomes of Healthy People. *Cell Host Microbe***25**, 656-667 e658, doi:10.1016/j.chom.2019.03.007 (2019).
- 9 Walter, J., Maldonado-Gomez, M. X. & Martinez, I. To engraft or not to engraft: an ecological framework for gut microbiome modulation with live microbes. *Curr Opin Biotechnol***49**, 129-139, doi:10.1016/j.copbio.2017.08.008 (2018).
- 10 Maldonado-Gomez, M. X. *et al.* Stable Engraftment of Bifidobacterium longum AH1206 in the Human Gut Depends on Individualized Features of the Resident Microbiome. *Cell Host Microbe***20**, 515-526, doi:10.1016/j.chom.2016.09.001 (2016).

- 11 Storelli, G. *et al.* Lactobacillus plantarum promotes Drosophila systemic growth by modulating hormonal signals through TOR-dependent nutrient sensing. *Cell Metab***14**, 403-414, doi:10.1016/j.cmet.2011.07.012 (2011).
- 12 Schwarzer, M. *et al.* Lactobacillus plantarum strain maintains growth of infant mice during chronic undernutrition. *Science***351**, 854-857, doi:10.1126/science.aad8588 (2016).
- 13 Kostic, A. D., Howitt, M. R. & Garrett, W. S. Exploring host-microbiota interactions in animal models and humans. *Genes Dev***27**, 701-718, doi:10.1101/gad.212522.112 (2013).
- 14 Derrien, M. & van Hylckama Vlieg, J. E. Fate, activity, and impact of ingested bacteria within the human gut microbiota. *Trends Microbio***23**, 354-366, doi:10.1016/j.tim.2015.03.002 (2015).
- 15 McNulty, N. P. *et al.* The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. *Sci Transl Med***3**, 106ra106, doi:10.1126/scitranslmed.3002701 (2011).
- 16 Suez, J. *et al.* Post-Antibiotic Gut Mucosal Microbiome Reconstitution Is Impaired by Probiotics and Improved by Autologous FMT. *Cell***174**, 1406-1423 e1416, doi:10.1016/j.cell.2018.08.047 (2018).
- 17 Garud, N. R., Good, B. H., Hallatschek, O. & Pollard, K. S. Evolutionary dynamics of bacteria in the gut microbiome within and across hosts. *PLoS Bio***17**, e3000102, doi:10.1371/journal.pbio.3000102 (2019).
- 18 Kundu, P., Blacher, E., Elinav, E. & Pettersson, S. Our Gut Microbiome: The Evolving Inner Self. *Cell***171**, 1481-1493, doi:10.1016/j.cell.2017.11.024 (2017).
- 19 Sanchez, A. & Gore, J. feedback between population and evolutionary dynamics determines the fate of social microbial populations. *PLoS Bio***11**, e1001547, doi:10.1371/journal.pbio.1001547 (2013).
- 20 Martino, M. E. *et al.* Bacterial Adaptation to the Host's Diet Is a Key Evolutionary Force Shaping Drosophila-Lactobacillus Symbiosis. *Cell Host Microbe***24**, 109-119 e106, doi:10.1016/j.chom.2018.06.001 (2018).
- 21 Zhang, J. *et al.* Metagenomic approach reveals microbial diversity and predictive microbial metabolic pathways in Yucha, a traditional Li fermented food. *Sci Rep***6**, 32524, doi:10.1038/srep32524 (2016).
- 22 Hempel, S. *et al.* Probiotics for the prevention and treatment of antibiotic-associated diarrhea: a systematic review and meta-analysis. *JAMA***307**, 1959-1969, doi:10.1001/jama.2012.3507 (2012).
- 23 Shao, Y. *et al.* Lactobacillus plantarum HNU082-derived improvements in the intestinal microbiome prevent the development of hyperlipidaemia. *Food Funct***8**, 4508-4516, doi:10.1039/c7fo00902j (2017).

- 24 Kurtz, Z. D. *et al.* Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput Biol***11**, e1004226, doi:10.1371/journal.pcbi.1004226 (2015).
- 25 Schroder, G. & Lanka, E. TraG-like proteins of type IV secretion systems: functional dissection of the multiple activities of TraG (RP4) and TrwB (R388). *J Bacteriol***185**, 4371-4381, doi:10.1128/jb.185.15.4371-4381.2003 (2003).
- 26 Gibson, G. R. *et al.* Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nat Rev Gastroenterol Hepatol***14**, 491-502, doi:10.1038/nrgastro.2017.75 (2017).
- 27 Succi, M. *et al.* Pre-cultivation with Selected Prebiotics Enhances the Survival and the Stress Response of *Lactobacillus rhamnosus* Strains in Simulated Gastrointestinal Transit. *Front Microbiol***8**, 1067, doi:10.3389/fmicb.2017.01067 (2017).
- 28 Tillisch, K. *et al.* Consumption of fermented milk product with probiotic modulates brain activity. *Gastroenterology***144**, 1394-1401, 1401 e1391-1394, doi:10.1053/j.gastro.2013.02.043 (2013).
- 29 Sanders, M. E. *et al.* Safety assessment of probiotics for human use. *Gut Microbes***1**, 164-185, doi:10.4161/gmic.1.3.12127 (2010).
- 30 Kwok, L. Y. *et al.* The impact of oral consumption of *Lactobacillus plantarum* P-8 on faecal bacteria revealed by pyrosequencing. *Benef Microbes***6**, 405-413, doi:10.3920/BM2014.0063 (2015).
- 31 Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes***9**, 88, doi:10.1186/s13104-016-1900-2 (2016).
- 32 Luo, R. *et al.* SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience***1**, 18, doi:10.1186/2047-217X-1-18 (2012).
- 33 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods***9**, 357-359, doi:10.1038/nmeth.1923 (2012).
- 34 Yang, J., Ding, X., Sun, X., Tsang, S. Y. & Xue, H. SAMSVM: A tool for misalignment filtration of SAM-format sequences with support vector machine. *J Bioinform Comput Biol***13**, 1550025, doi:10.1142/S0219720015500250 (2015).
- 35 Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res***44**, W242-245, doi:10.1093/nar/gkw290 (2016).
- 36 Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol***19**, 455-477, doi:10.1089/cmb.2012.0021 (2012).

- 37 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol***215**, 403-410, doi:10.1016/S0022-2836(05)80360-2 (1990).
- 38 Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome Res***12**, 656-664, doi:10.1101/gr.229202 (2002).
- 39 Song, Y. *et al.* Genomic Variations in Probiotic *Lactobacillus plantarum* P-8 in the Human and Rat Gut. *Front Microbio***9**, 893, doi:10.3389/fmicb.2018.00893 (2018).
- 40 Moskalensky, A. E. *et al.* Method for the simulation of blood platelet shape and its evolution during activation. *PLoS Comput Biol***14**, e1005899, doi:10.1371/journal.pcbi.1005899 (2018).
- 41 Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol***20**, 257, doi:10.1186/s13059-019-1891-0 (2019).
- 42 Franzosa, E. A. *et al.* Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods***15**, 962-968, doi:10.1038/s41592-018-0176-y (2018).
- 43 Morton, J. T. *et al.* Establishing microbial composition measurement standards with reference frames. *Nat Commun***10**, 2719, doi:10.1038/s41467-019-10656-5 (2019).
- 44 Nayfach, S., Rodriguez-Mueller, B., Garud, N. & Pollard, K. S. An integrated metagenomics pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome Res***26**, 1612-1625, doi:10.1101/gr.201863.115 (2016).

Table

Table 1. The detailed information of the 22 SNPs

SNP ID	Location	Aa	Alt	Gene ID	MT	AAC	Gene name	Biological process	Protein annotation
SNP01	63020	C	A	Gene0056	N	Q86H	HP	NA	Transposase
SNP10	695080	T	C	Gene0658	N	S194P	hcaR	Transcription regulation	Transcriptional activator for 3-phenylpropionic acid catabolism
SNP22	1283071	A	G	Gene1205	N	L186S	HP	NA	Hypothetical protein
SNP23	1338981	A	C	Gene1257	N	D484E	ram2	Protein farnesylation	Bacterial alpha-L-rhamnosidase 6 hairpin glycosidase
SNP27	1589637	G	T	Gene1470	N	L827M	trePP	Carbohydrate metabolic process	Trehalose 6-phosphate phosphorylase
SNP28	1590767	T	G	Gene1470	N	D450A	trePP	Carbohydrate metabolic process	Trehalose 6-phosphate phosphorylase
SNP32	1861048	T	C	Gene1717	S	T164T	yagU	Response to acidic Ph	Inner membrane protein response to acidic
SNP41	2799363	A	G	Gene2601	S	E328E	spa	Virulence	Immunoglobulin G-binding protein
SNP43	2854610	A	C	Gene2651	N	D202A	gor	Glutathione metabolic process	Glutathione reductase
SNP47	3007289	T	G	Gene2804	S	A96A	int4	DNA integration	Transposase
SNP46	3007122	T	G	Gene2804	N	S41A	int4	DNA integration	Transposase
SNP50	3212834	T	G	Gene3003	N	F99L	ybhF	Transport	ABC transporter ATP-binding protein
SNP51	3213000	G	T	Gene3003	N	V155L	ybhF	Transport	ABC transporter ATP-binding protein
SNP55	3310964	C	T	Gene3110	N	G130S	tnp2PF3	Transposition	Transposase
SNP56	3311024	T	C	Gene3110	N	T110A	tnp2PF3	Transposition	Transposase
SNP57	3311155	A	C	Gene3110	N	V66G	tnp2PF3	Transposition	Transposase
SNP58	3311174	T	C	Gene3110	N	N60D	tnp2PF3	Transposition	Transposase
SNP60	3311545	T	C	Gene3111	N	K35E	tnp2PF3	Transposition	Transposase
SNP59	3311468	G	A	Gene3111	S	S60S	tnp2PF3	Transposition	Transposase
SNP61	1976363	A	G	NA	NA	NA	NA	NA	NA
SNP62	1976367	G	T	NA	NA	NA	NA	NA	NA
SNP71	3271378	G	T	NA	NA	NA	NA	NA	NA

Note: "Aa" represented the original single nucleotide; "Alt" represented the mutated single nucleotide;

"MT" represented Mutation Type, Synonymous mutations or Non-synonymous mutations;

"AAC" represented amino acid changes;

Figures

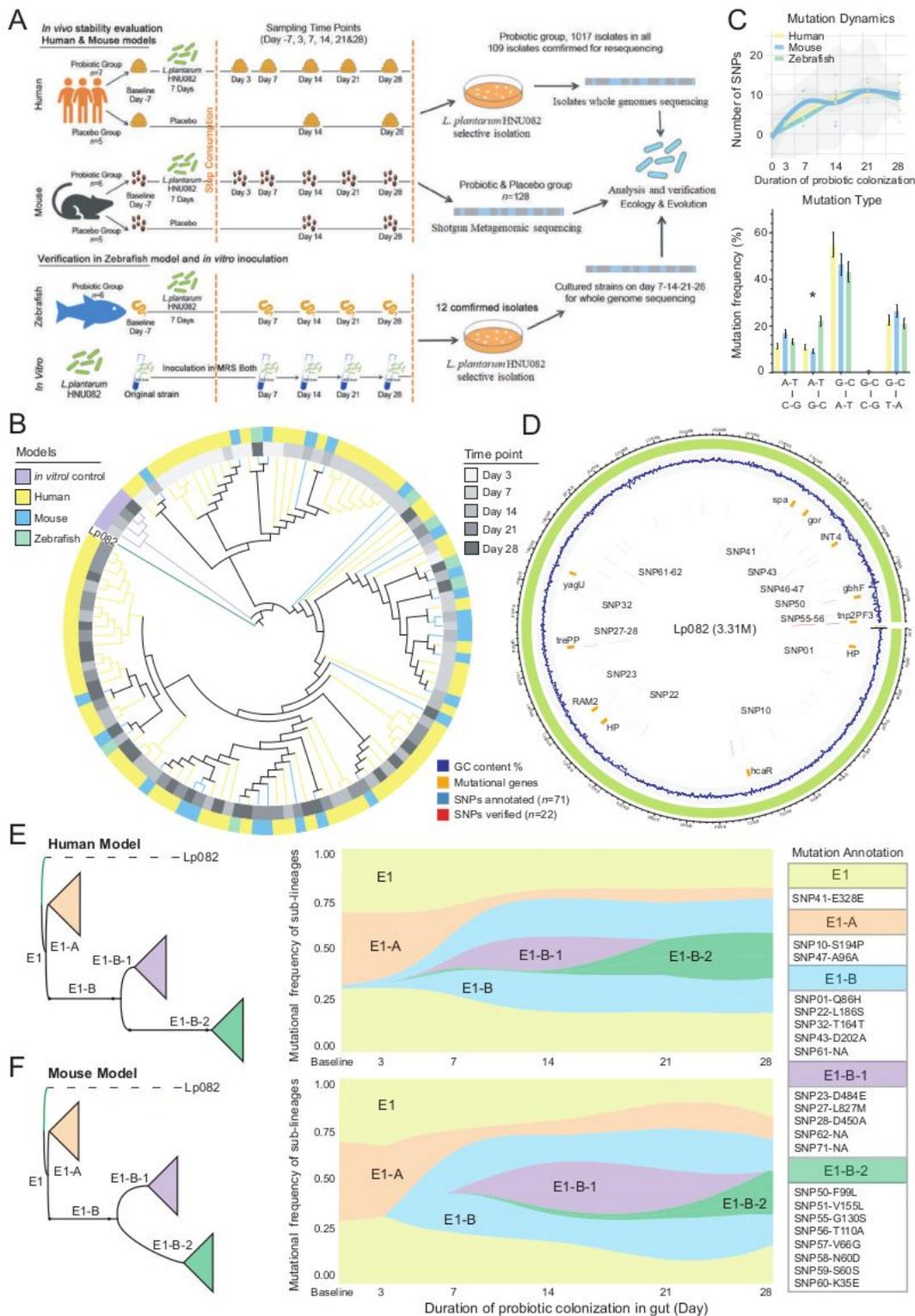


Figure 1

The in vivo adaptive evolutionary dynamics of probiotics over a four-weeks sampling period in the gut of humans, mice and zebrafish. (A) The experiment design. We used Lp082 as a model probiotic strain to explore the effects of host-derived selection pressure (the mouse and human models were applied, the zebrafish model was used for further verification) on the genetic stability of the ingested probiotics and the impact of probiotic mutations on the indigenous gut microbiome of different hosts (the mouse and

human model). First, we sequenced the complete genome of this model probiotic strain. We next isolated the probiotics from the feces of hosts at different time points to identify genetic mutations using whole-genome resequencing. Simultaneously, the original strain was continuously inoculated in vitro and sequenced to assess genetic mutation in the absence of host selective pressure. Next, we employed the metagenomic sequencing method to characterize the impact of probiotics ingestion on resident gut microbiota in humans and mice as compared to the placebo groups. (B) Phylogenetic tree constructed based on the SNPs of all Lp082 isolates. The different grey bars and color bars represent the strains isolated from different time points and different hosts, and the Lp082 was set as the root strain. The isolates were dominated by 22 SNPs during the probiotic colonization, especially in the first two time points of the human model. (C) The number of SNPs detected at every sampling time point (including the day 3, 7, 14, 21 and 28) in human, mouse and zebrafish models (top panel). No significant difference in the number of SNPs at the end of the experiment was found among the three models. The mutation type of G-C to A-T was the most frequently detected in all three models, but the frequency of the mutation type of A-T to G-C was significantly higher in the zebrafish model than that in the other two models (bottom panel). (D) All confirmed SNPs and their gene locations are marked on the reference genome of Lp082. (E-F, left panel) The simplified phylogenetic tree based on Lp082 isolates from the human (E) and mouse (F) model in all sampling time points. The tree is rooted in the ancient probiotic strain consumed and evolved into 3 groups (labeled in different colors) based on 21 SNPs in 5 branches. The evolutionary relationships of the 5 branches are visualized as E1, E1-A/B, and E1-B-1/2. (E-F, right panel) The evolutionary dynamics of the 5 branches (in different colors) in the human and mouse model were constructed based on the mutation frequency of the 21 SNPs. A major difference in the evolutionary dynamics of branch E1-B-1 was observed between the human and mouse model, which consisted of the temporary evolutionary divergence in the second week of probiotic colonization.

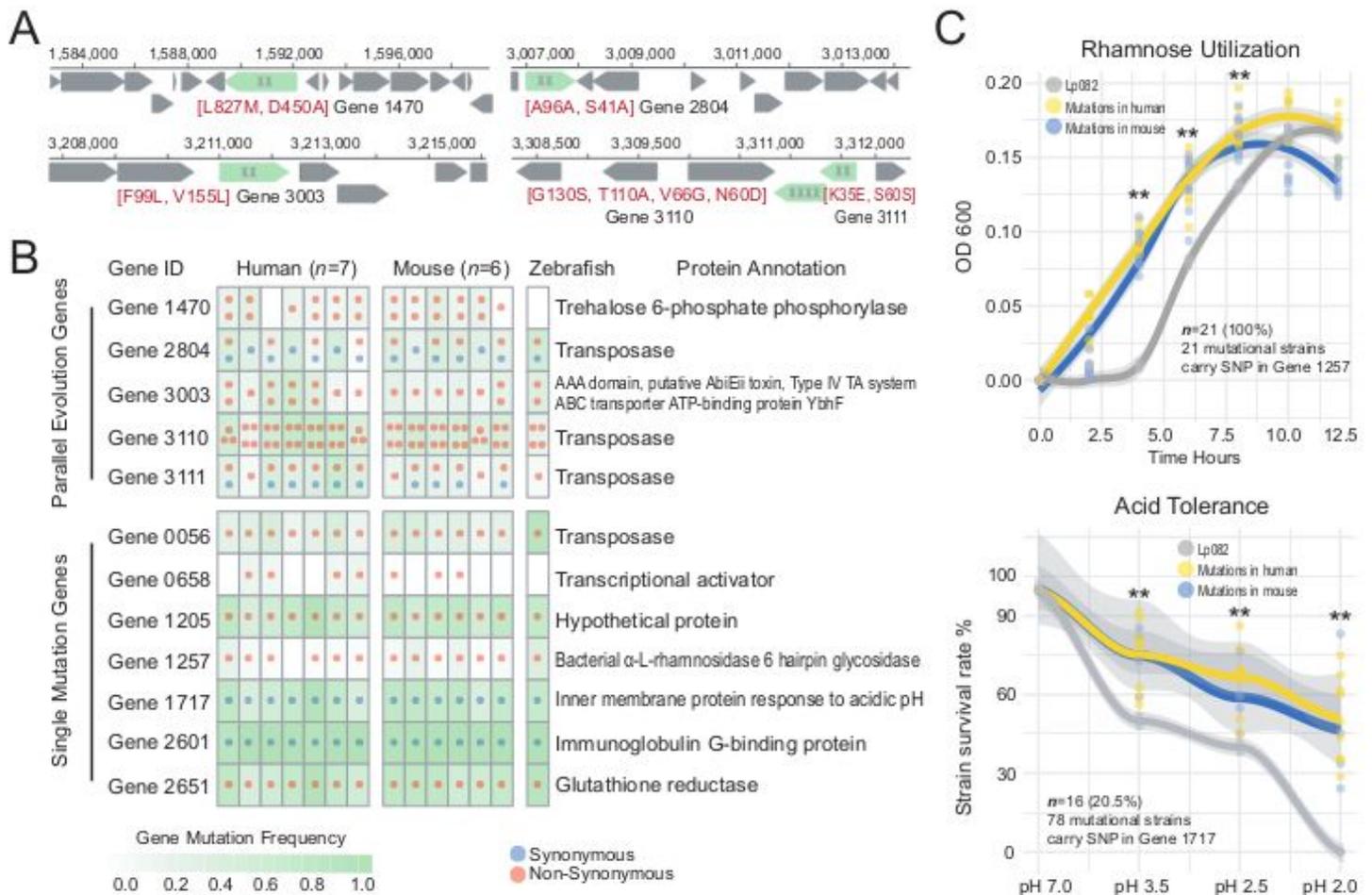


Figure 2

Genes that underwent in vivo evolution and the enhanced competitive fitness of the probiotic in mammalian hosts. (A) The locations of the five parallel evolutionary genes in the Lp082 chromosome. The genes highlighted by green color represent the parallel evolutionary genes and the “x” symbols represent the number of mutations detected in the gene. Annotations in red color represent the mutations that lead to functional changes in the amino acid sequence (ie. L(Leu) to M(Met)). (B) Five probiotic genes underwent parallel evolution in different host species, while the other seven genes only have a single mutation. Each dot in the grid chart represents an independent mutation event and is also colored by types of mutation (synonymous or non-synonymous). The color within each cell represents the mutation frequency of a gene in all probiotic isolates from a host subject. (C) Either non-synonymous or synonymous SNPs can confer the growth benefits of probiotics in the host gut. The phenotypic verification experiments of probiotic isolates related to the rhamnose utilization (top panel) and acid tolerance (bottom panel). A non-synonymous SNP identified in Gene 1257 (annotated as bacterial alpha-L-rhamnosidase 6 hairpin glycosidase) from 21 mutational isolates (100%). This SNP conferred growth benefits to all isolates by utilizing the rhamnose more rapidly than the original strain. Remarkably, a synonymous SNP in Gene 1717 (annotated as inner membrane protein response to acidic) was found in

78 isolates. Among these, 16 mutational isolates (20.5%) improved the performance of this strain in acid tolerance regardless of a synonymous SNP.

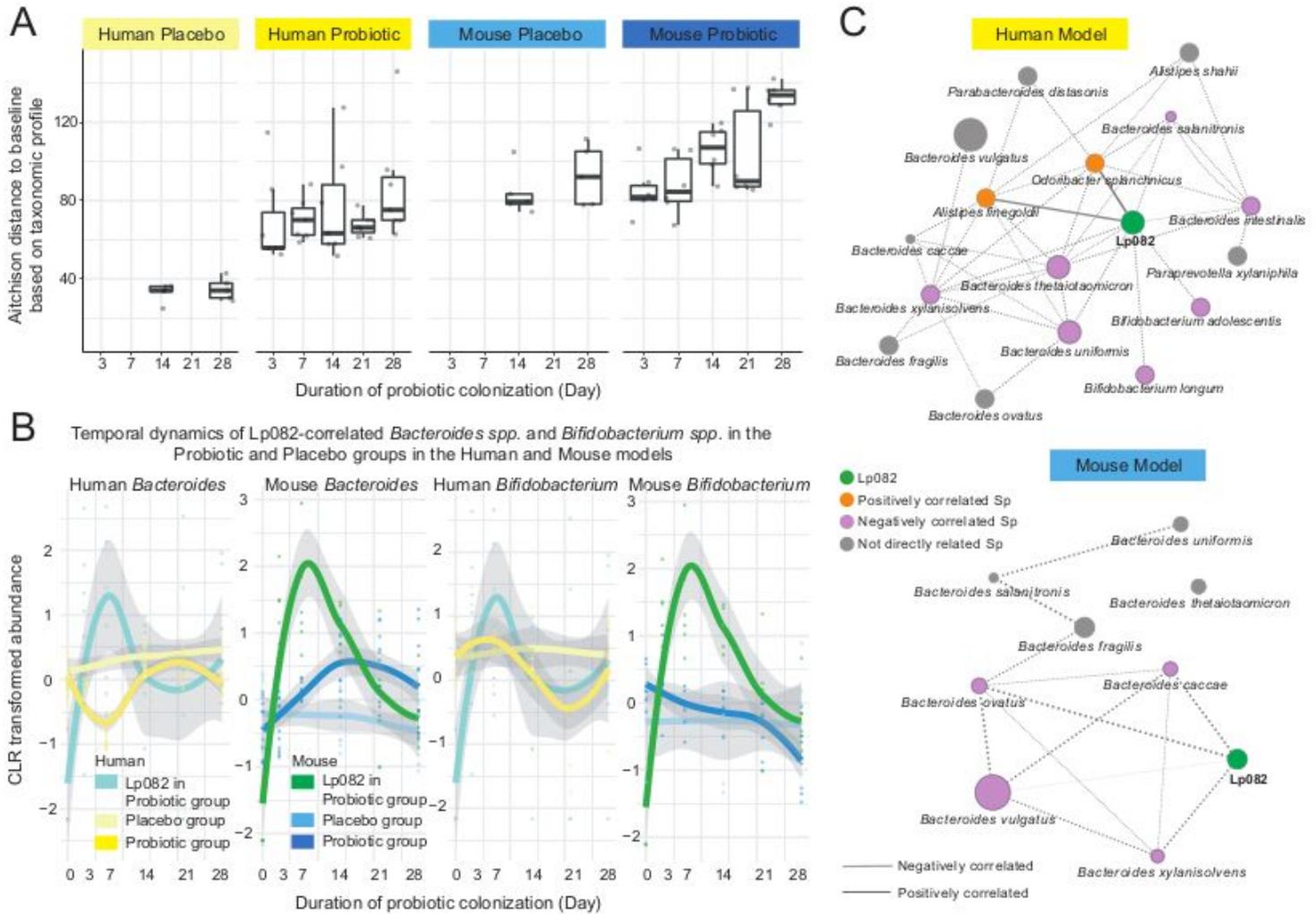


Figure 3

The microbiome response in the gut to the probiotic engraftment in humans and mice. (A) The Aitchison distance between the samples in baseline and other time points both in the human and mouse model. The colored points represented the samples at different time points. We found that interindividual heterogeneity in the human gut microbiota is greatest among those in the gut microbiota of all host models. (B) The temporal dynamic of 19 intestinal species belonged to the genera *Bacteroides* and *Bifidobacterium*, which were commonly identified in both human and mouse models. The dynamics of the same species residing in both humans and mice were highly divergent, suggesting niche-specific adaptation strategies of resident gut microbes responding to probiotic invasion. (C) The ecological relationships between the Lp082 and resident intestinal microbes are visualized by the co-occurrence networks in the human and mouse model. The nodes in different colors respectively represent Lp082 (green), the probiotic positively correlated species (orange), the probiotic negatively correlated species (purple), and the species indirectly correlated with the probiotic (grey). The correlation strength between nodes (species) was calculated by the SpiecEasi based on the CLR-transformed microbial relative

abundance. The thickness of an edge represents the correlation strength between two nodes. A dashed edge indicates a negative correlation while a solid edge indicates a positive correlation between microbial species.

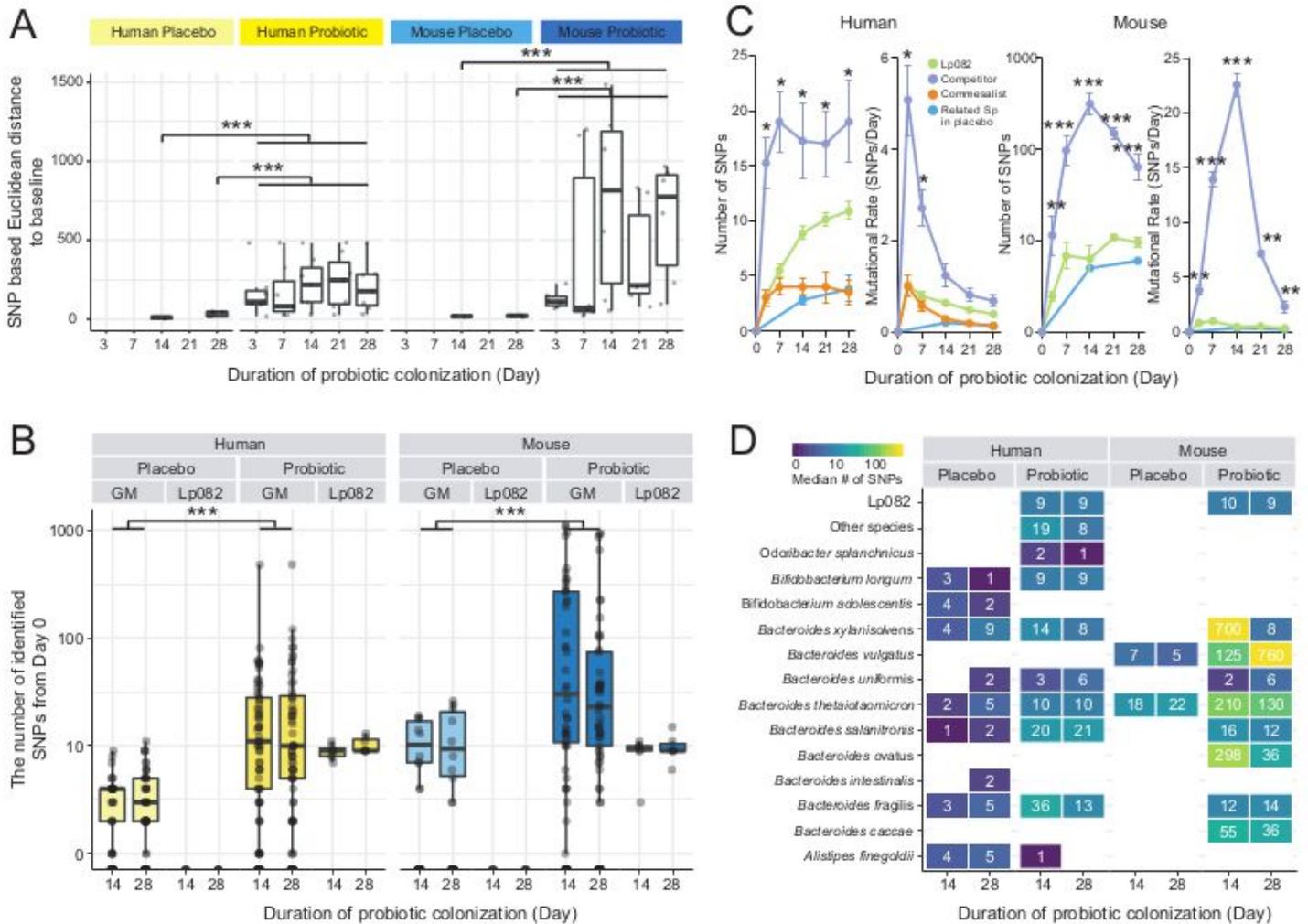


Figure 4

The rapid co-evolution of the ingested probiotic and resident gut microbiota of humans and mice within 28 days. (A) The Euclidean distance based on the number of SNPs identified from Day 0 to other time points during the probiotic colonization in human and mouse models. It strongly indicated an intensive evolutionary response in resident microbiota due to probiotic intake. (B) The distribution of the mutations identified in the probiotic Lp082 and resident gut microbiota from the probiotic and placebo group in both human and mouse models. Each dot in the boxplot represents the number of SNPs that occurred on a microbial strain in the gut of a host subject as compared to that on Day 0. The “GM” is the abbreviation of “gut microbiota”. (C) The ecological relationship with Lp082 determined the number of SNPs that occurred on a resident microbial strain. Overall, the number of SNPs of probiotic “competitors” (orange, such as *Bacteroides* spp. and *Bifidobacterium* spp.) was significantly greater than that of “non-competitors” of this probiotic (orchid) or this probiotics (green) in the human model. In the mouse model, the adaptive mutations occurring in those probiotic competitors were one to two orders of magnitude

more than those identified in the probiotics over the 28-days sampling period. The least mutations were identified in the placebo group (blue line). (D) The heat map indicates the median number of SNPs identified in each microbial species at Day 14 and Day 28 in each host group compared to Day 0. Asterisks: statistical significance (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$)

Supplementary Files

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

- [SupFigures.pdf](#)
- [Allsuppltables.xlsx](#)
- [SupFigures.pdf](#)