Genomic and epidemiological evidence of bacterial transmission from probiotic capsule to blood in ICU patients

Idan Yelin [1,10], Kelly B. Flett[2,3,9,10], Christina Merakou[3,4,10], Preeti Mehrotra[2,5], Jason Stam[6], Erik Snesrud[6], Mary Hinkle[6], Emil Lesho[6], Patrick McGann[6], Alexander J. McAdam[2,3,7], Thomas J. Sandora [2,3,11]*, Roy Kishony [1,8,11]* and Gregory P. Priebe [2,3,4,11]*

Probiotics are routinely administered to hospitalized patients for many potential indications[1] but have been associated with adverse effects that may outweigh their potential benefits[2–7]. It is particularly alarming that probiotic strains can cause bacteremia[8,9], yet direct evidence for an ancestral link between blood isolates and administered probiotics is lacking. Here we report a markedly higher risk of *Lactobacillus* bacteremia for intensive care unit (ICU) patients treated with probiotics compared to those not treated, and provide genomics data that support the idea of direct clonal transmission of probiotics to the bloodstream. Whole-genome-based phylogeny showed that *Lactobacilli* isolated from treated patients' blood were phylogenetically inseparable from *Lactobacilli* isolated from the associated probiotic product. Indeed, the minute genetic diversity among the blood isolates mostly mirrored pre-existing genetic heterogeneity found in the probiotic product. Some blood isolates also contained de novo mutations, including a non-synonymous SNP conferring antibiotic resistance in one patient. Our findings support that probiotic strains can directly cause bacteremia and adaptively evolve within ICU patients.

Probiotics are increasingly administered to hospitalized patients[1]. These supplementary products have shown benefit in acute infectious diarrhea, antibiotic-associated diarrhea, and ulcerative colitis[2,10,11]. In the intensive care unit (ICU), additional indications are being explored, including prevention of ventilator-associated pneumonia, pancreatitis, and sepsis[12–14]. However, studies on the efficacy and adverse effects of probiotic treatment in ICU patients show conflicting results, and their use remains controversial[4–7,15]. Adverse outcomes, including bacteremia, have been reported and may preclude their use in specific populations such as those with a compromised immune system or disorders of the gastrointestinal tract[8,9,16,17].

Bacteremia that appears during the course of probiotic treatment can involve *Lactobacillus* species similar to those in probiotics, yet as these species are also common in the human gastrointestinal microbiome, pinpointing the source of these infections has been challenging[18]. Studies that use pulsed-field gel electrophoresis have previously revealed strain-level similarity between blood and probiotic isolates[19], but higher genomic resolution is required to establish direct clonal ancestry and the possibility of direct transmission of probiotic bacteria to the blood.

In the context of bacterial pathogens, whole-genome methods have been powerful in identifying transmission links and within-host adaptation[20–23]. Constructing a single nucleotide polymorphism (SNP)-level phylogeny of isolates from patients can unravel ancestral links between lineages and likely paths of transmission. Whole-genome comparison of isolates can also reveal adaptive mutations important for the survival of the pathogen within the host[20,22]. Yet, despite their established power, the use of these whole-genome approaches for tracing of probiotic strains has so far been limited. Here, we apply whole-genome analysis and phenotyping to blood isolates and probiotic strains administered to ICU patients.

Analysis of cases of *Lactobacillus* bacteremia in ICU patients at Boston Children's Hospital showed that patients receiving *Lactobacillus rhamnosus* strain GG (LGG) probiotics had a markedly higher risk of developing *Lactobacillus* bacteremia than those who received no probiotics. Over a period of 5.5 years, a total of 22,174 patients were treated in an ICU, and 522 of these patients received LGG-containing probiotic—typically through a feeding tube—as part of their treatment. Analysis of recorded *Lactobacillus* bacteremia among these patients showed a significantly greater risk for patients who received the LGG-containing probiotic than for those who did not; 6 of the 522 patients had *Lactobacillus* bacteremia (1.1%, patients R1–R6; Supplementary Table 1a) compared to only 2 of the 21,652 patients who did not receive the LGG probiotic (0.009%, patients N1–N2; $P = 4.8 \times 10^{-9}$, Fisher's exact test; Supplementary Table 1a). Furthermore, all six of the ICU blood isolates from patients receiving the LGG probiotic were identified by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry as *Lactobacillus rhamnosus*, while the two isolates from patients not receiving the LGG probiotic were identified as other *Lactobacillus* species (Supplementary Table 1a). The ICU patients receiving probiotics containing LGG are therefore at

[1]Department of Biology, Technion–Israel Institute of Technology, Haifa, Israel. [2]Division of Infectious Diseases, Department of Pediatrics, Boston Children's Hospital, Boston, MA, USA. [3]Harvard Medical School, Boston, MA, USA. [4]Division of Critical Care Medicine, Department of Anesthesiology, Critical Care and Pain Medicine, Boston Children's Hospital, Boston, MA, USA. [5]Division of Infectious Diseases, Beth Israel Deaconess Medical Center, Boston, MA, USA. [6]Walter Reed Army Institute of Research, Silver Spring, MD, USA. [7]Department of Laboratory Medicine, Boston Children's Hospital, Boston, MA, USA. [8]Department of Computer Science, Technion–Israel Institute of Technology, Haifa, Israel. [9]Present address: Novant Health Eastover Pediatrics, Charlotte, NC, USA. [10]These authors contributed equally: I. Yelin, K. B. Flett, C. Merakou. [11]These authors jointly supervised this work: T. J. Sandora, R. Kishony, G. P. Priebe. *e-mail: thomas.sandora@childrens.harvard.edu; rkishony@technion.ac.il; gregory.priebe@childrens.harvard.edu

markedly higher risk of developing *Lactobacillus rhamnosus* bacteremia (6 out of 522 compared to 0 out of 21,652; $P = 1.8 \times 10^{-10}$, Fisher's exact test). Further, the LGG probiotic bacteremia rate of 1.1% that we observed is also much higher than the annual rate of LGG probiotic bacteremia (0.00007%) reported in the general population[19]. Yet, *L. rhamnosus* bacteremia can occasionally also appear in patients not receiving these probiotics: during the study period, there were an additional ten cases of *Lactobacillus* bacteremia among approximately 93,000 non-ICU patients (Patients N3–N12, Supplementary Table 1b), and four of these ten isolates were identified by MALDI-TOF as *Lactobacillus rhamnosus* (Patients N5, N9–N11, Supplementary Table 1b). None of the ten non-ICU patients were receiving a probiotic at the time of the bacteremia. Taken together, these results suggest that ICU patients who receive probiotics containing LGG are at much higher risk of developing *Lactobacillus* bacteremia than patients who do not, but it is difficult to prove a direct causal relationship as occasional cases of *L. rhamnosus* bacteremia also appear in patients who do not receive these probiotics.

To achieve better ancestral resolution, we next used whole-genome sequencing to determine strain-level similarity among the blood and probiotic isolates. We performed whole-genome sequencing of all ten *L. rhamnosus* blood isolates (six from patients receiving probiotic and four from patients not receiving probiotic), as well as 16 isolates from each of three probiotic capsules of different lots (probiotic batches 1–3; Supplementary Table 2, Methods and Fig. 1a). To quantify strain-level relatedness among these isolates, we started by measuring their distance to all available *L. rhamnosus* genomes (GenBank, October 2017; Supplementary Table 3). Illumina reads of each isolate were aligned to each of these genomes, and the fraction of aligned reads, affected both by gene content similarity and SNP density, was quantified as a measure of similarity. We found that all six blood isolates and all probiotic product isolates shared the same closest reference genome—an LGG genome (GenBank chromosome ID: FM179322)—suggesting high relatedness between these two sets of isolates (Fig. 1b). By contrast, all four *L. rhamnosus* blood isolates from patients not receiving probiotics were more similar to other strains, indicating that they were not derived from the probiotic product (Fig. 1b).

To further increase genomic resolution, we next compared the genomes of the blood and LGG probiotic isolates by alignment to the reference genome. Analysis of gene content of the isolates identified only a single deletion; one of the probiotic isolates of batch 2 had a large deletion of a region which included 82 genes of the reference genome FM179322 (genes 384–465, Fig. 2). Strains were also almost identical at the single-nucleotide level; analysis of SNP-level variations identified a total of only 23 SNPs among all isolates (Methods; for the list of SNPs, see Supplementary Table 4). Indeed, the greatest SNP distance between any isolate and the last common ancestor was not more than six SNPs. Two SNPs were shared by all isolates, separating them from the reference genome, and indicating that the blood and probiotic isolates share a more recent last common ancestor than the LGG clone deposited in GenBank (Fig. 1c). Moreover, the blood and probiotic isolates were phylogenetically inseparable; there was no mutation that strictly separated these two groups (Fig. 1c and Supplementary Table 4).

Much of the genetic diversity among blood isolates mirrored pre-existing genetic diversity within the probiotic capsules. We identified 11 genomic positions that were polymorphic across blood isolates (Fig. 2). Three of these mutations, all of which were non-synonymous, were recurring mutations, observed in more than one blood isolate (H294Q in CamS, H248Y in GlvA, and Q1827R in SpcB; Fig. 2 and Supplementary Table 4). These repeatedly occurring blood isolate mutations were all shown to pre-exist in the probiotic product (these same loci were diverse within each of the three batches of the probiotic product, Fig. 2). Furthermore, the *camS* and
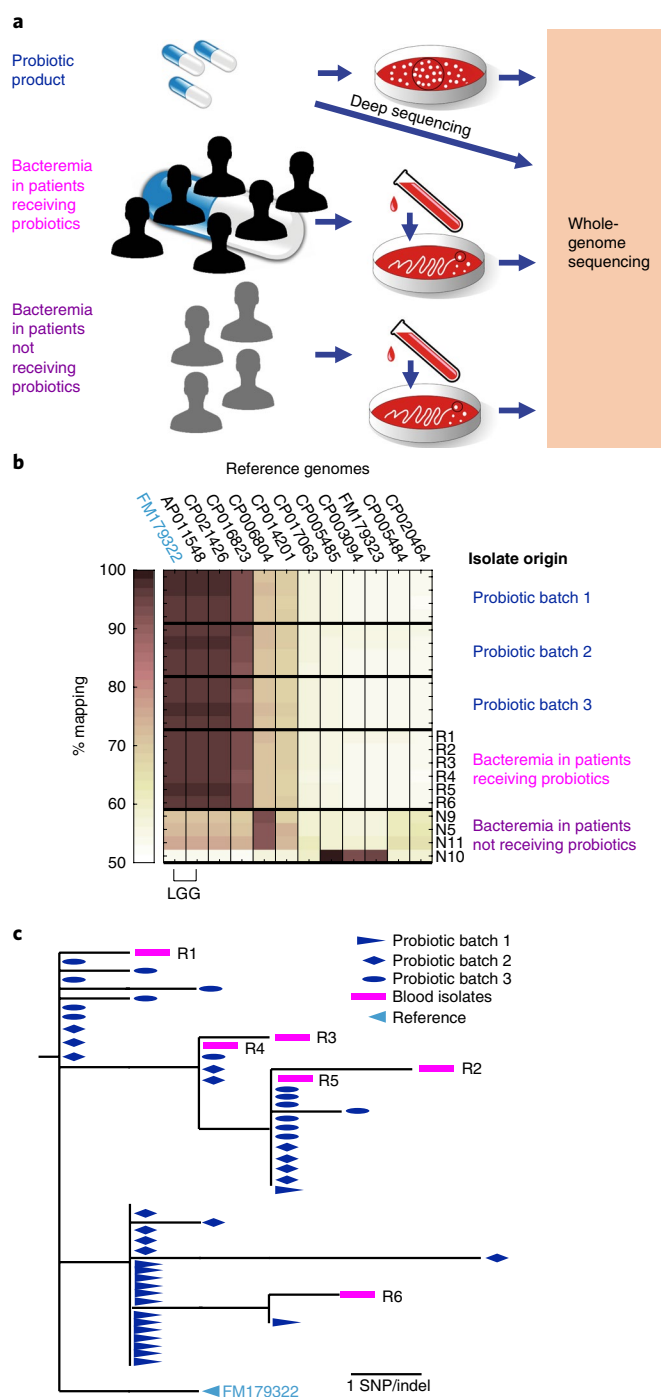

Fig. 1 | Genomic evidence for *L. rhamnosus* transmission from probiotic capsule to the blood of patients. a, Schematic for whole-genome sequencing of *L. rhamnosus* probiotic isolates, blood isolates from ICU patients ($n = 6$) receiving probiotics, and blood isolates from non-ICU patients ($n = 4$) who were not receiving probiotics. Black circles represent sequencing of multiple individual colonies for each probiotic batch but a single colony for each blood isolate. **b**, Similarity between *L. rhamnosus* isolates and available reference genomes are shown as the fraction of reads aligned to each reference. Isolates are identified by their source: four representative isolates from each of three probiotic product batches, the six blood isolates from patients receiving probiotics, and the four blood *L. rhamnosus* isolates from patients not receiving probiotics. **c**, Phylogenetic analysis of all 54 sequenced LGG isolates: 16 isolates from each of three separate probiotic batches (blue), and the six blood isolates from Patients R1 to R6 (magenta).
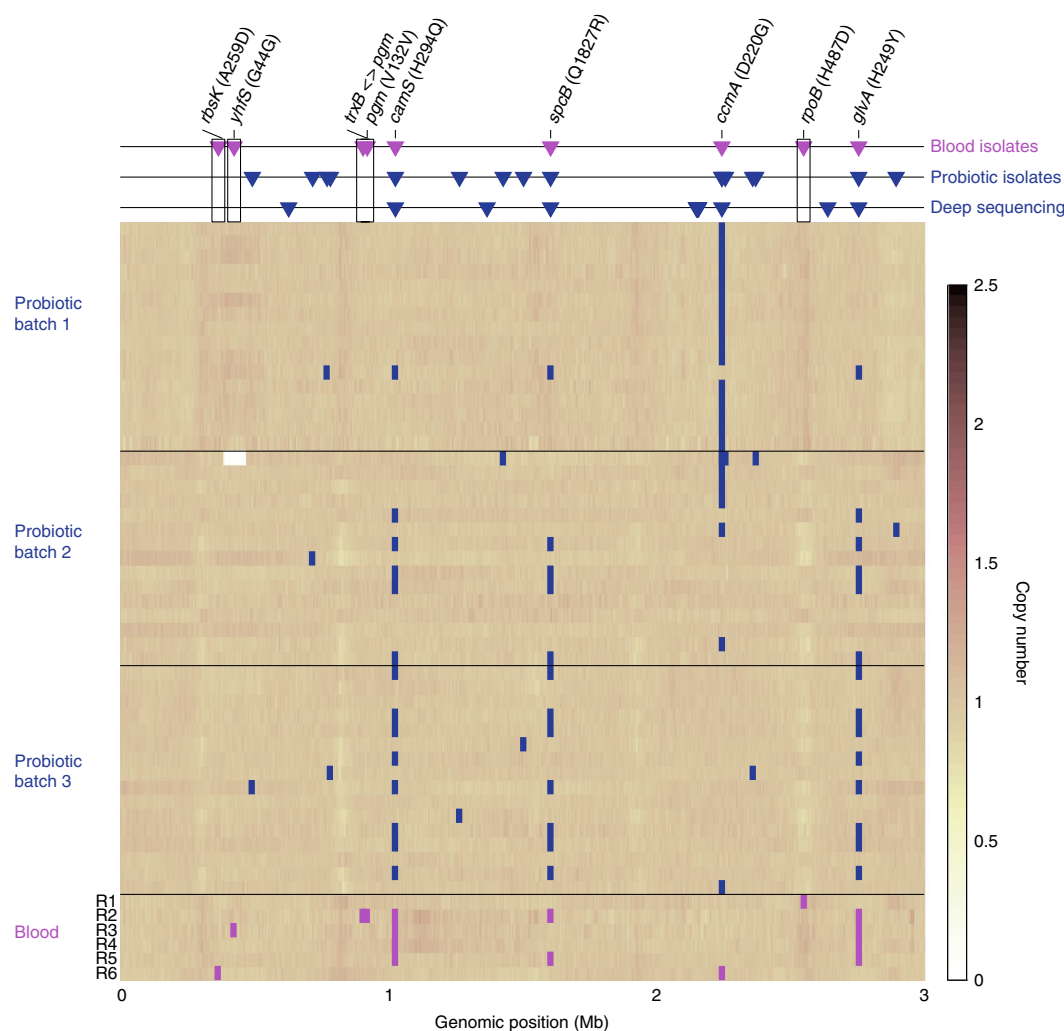
Fig. 2 | Coverage of the LGG reference genome for the probiotic and blood *Lactobacillus rhamnosus* isolates. For each isolate (row in matrix) SNPs are marked as squares (magenta for blood isolates, blue for probiotic product isolates). Triangles (top panel) indicate all mutations identified in blood isolates (magenta triangles) and probiotic product (blue triangles) compared to the LGG reference genome FM179322. For the probiotic product, these are either high-quality SNPs in whole-genome sequencing (middle row) or diversity identified by deep sequencing of the product (bottom row, see Methods). Annotation is included for all SNPs identified in blood isolates. SNPs identified only in blood isolates are indicated with black open rectangles. Mb, megabase pairs.

glvA SNPs always appeared together in both the blood and probiotic isolates. One other polymorphic locus, identified in a single blood isolate, was also found to pre-exist in the probiotic product (a D220G mutation in the ABC transporter CcmA). Overall, correspondence between blood isolate mutations and pre-existing diversity within the probiotic product further supports the likelihood of transmission of bacteria from probiotic to blood.

In addition to the six blood isolate mutations that were found to pre-exist in the probiotic product, we identified five blood isolate mutations that did not appear in the isolates from the probiotic product, suggesting de novo evolution within the patient (Fig. 2, Supplementary Table 5). These 5 mutations were not found in any of the 16 genomes isolated from each of the 3 capsules. To further test for their possible existence in the product, we deep-sequenced capsules from five different batches, obtained both from the hospital and from a commercial pharmacy, and identified diverse loci (batches 2–6, Methods; Fig. 2, Extended Data Fig. 1, Supplementary Table 6). No pre-existing genotypic diversity was found at the loci of the five blood-isolate-specific mutations (Fig. 2, Supplementary Table 7). One of these five mutations was in an intergenic promoter mutation, two were non-synonymous coding mutations (H487D

in the RNA polymerase RpoB (Fig. 3a, Extended Data Fig. 2) and A259D near the active site of the RbsK ribokinase (Extended Data Fig. 3)), and two were synonymous mutations (at G44 of the YhfS transferase and at V132 of phosphoglucomutase). These mutations, existing in the blood but not identified in the probiotic product, could represent de novo mutations selected within the patient.

The blood-isolate-specific mutation in the *rpoB* RNA polymerase gene (H487D) appeared in an isolate from patient R1, who had received *L. rhamnosus GG* and the rifampin derivative rifaximin concurrently during the 3 months prior to bacteremia. This mutation, which changes a specific residue in the cleft of the RpoB DNA-binding site, is known to provide resistance to rifampin (Fig. 3a, Extended Data Fig. 2)[24–26]. Antibiotic susceptibility measurements showed that this blood isolate was indeed resistant to rifampin, whereas all other blood isolates as well as a probiotic isolate containing no SNPs were sensitive to this antibiotic (Fig. 3b). By contrast, susceptibilities to other antibiotics were nearly identical among the blood and probiotic isolates (Supplementary Table 8), suggesting that the R1 isolate adapted specifically to resist rifampin. Interestingly, although rifampin resistance mutations at other *rpoB* positions typically confer decreased fitness, mutations at the H487
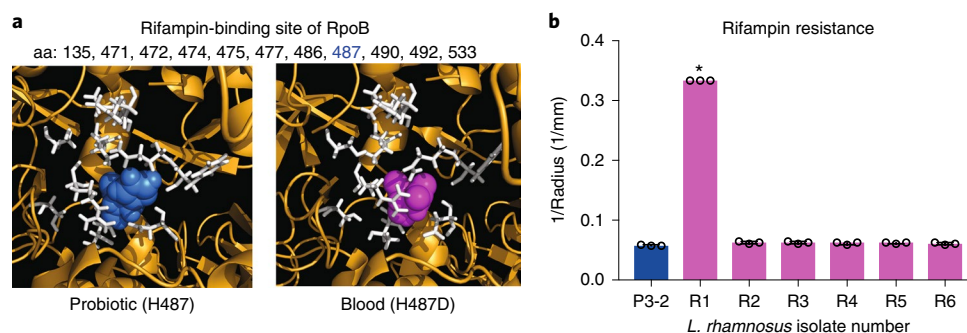
Fig. 3 | The *L. rhamnosus* blood-isolate-specific *rpoB* SNP occurs at the rifampin-binding site and confers rifampin resistance. a, Predicted structure of *LGG* RNA polymerase β-subunit RpoB showing the rifampin-binding site (white) with H487 of the probiotic (blue, left) and D487 of the blood isolate from Patient R1 (magenta, right). **b**, Rifampin susceptibility testing of blood isolates of each patient (R1–R6) compared to a probiotic isolate with no SNPs (P3-2). Bars depict the medians of three independent experiments, and error bars show the interquartile ranges. *$P = 0.0021$ for R1 compared to P3-2 by Kruskal-Wallis test followed by Dunn's multiple comparisons test. The blood isolate from Patient R1 was resistant based on zone cutoffs for *Staphylococcus aureus* (Supplementary Table 8).

position can retain fitness similar to that of the wild type[27]. Indeed, the R1 isolate carrying the *rpoB* mutation showed no significant fitness cost compared with the probiotic strain containing no SNPs (Supplementary Table 9). The specificity of the *rpoB* mutation to the patient receiving rifampin, together with its associated resistance and growth phenotypes, further suggest that the probiotic strains may acquire adaptive mutations that increase their fitness in the host environment.

We further considered other adaptive phenotypes. Survival in serum or human whole blood was similar among the probiotic and blood isolates (Supplementary Fig. 1), whereas the level of biofilm formation (Extended Data Fig. 4), which may lead to increased adhesion to a central venous line (CVL) and/or enhanced survival in the gastrointestinal tract, was significantly higher in the LGG blood and probiotic isolates compared to the non-LGG blood isolates (from Patients N5, N9–N11) and to the probiotic isolate P2-1 containing an 82-gene deletion, which includes the *spaCBA* pilus genes critical for biofilm[28] (Supplementary Table 10). These results suggest that biofilm is not required for bacteremia and that the LGG probiotic products may contain mutants with markedly different biofilm phenotypes.

Our patient population was critically ill, but the patients who developed LGG bacteremia while receiving probiotics did not have the typical risk factors for *Lactobacillus* bacteremia such as severe immune compromise or bowel disintegrity. Furthermore, in a case-control study to compare potential risk factors for bacteremia in these 6 cases with 16 matched control ICU patients who received probiotics but did not have bacteremia (Methods), we found no significant differences in device utilization, vasopressor support, recent surgery, diarrhea, parenteral nutrition, or antibiotic exposure (Supplementary Table 11). Although the low number of patients in this case-control study may limit statistical power, our inclusion of a control group with case-control methodology represents a significant improvement over prior descriptive studies in understanding specific risk factors within the ICU. The lack of strong differences between the patients who had bacteremia and the control group that did not have bacteremia suggests that the ICU patients at risk for transmission of probiotics from product to blood may not be easily identifiable.

The exact mechanism of transmission from probiotic to blood is unclear. Nearly all of these patients had a central line, and direct contamination of the central line with a probiotic strain or with stool containing the probiotic strain could lead to the observed probiotic bacteremia. Alternatively, the probiotic bacteria could have translocated across the bowel wall. The antibiotic-resistance-related

adaptation we observed could appear either prior to or immediately following the transmission of the bacteria to the blood. Our results suggest that these adaptive mutations are absent in the probiotic capsule and therefore evolved within the host environment. Nevertheless, given possible genomic variations among batches of the probiotic product, we cannot exclude the possibility that some of these presumably blood-specific mutations were present in the specific capsules given to each patient. In any case, whether they appear through rare mutations that already exist in the product or de novo during treatment, these emerging antibiotic-resistant probiotic bacteria could potentially undermine treatment efficacy. It would be interesting in future studies to explore the importance of other de novo mutations with additional in vitro phenotyping, or in animal models.

In summary, our epidemiological analysis uncovers a statistically and clinically significant risk for bacteremia associated with the use of probiotic *Lactobacilli* in the ICU, and genome-level analysis identified six independent cases of transmission of probiotics from capsule to blood in ICU patients treated with probiotics. Our results also provide evidence of within-host evolution of the probiotic, including acquisition of antibiotic resistance. Probiotics have shown significant benefits for acute infectious diarrhea, antibiotic-associated diarrhea, and ulcerative colitis[2,10,11]. However, our findings highlight that as ICU patients have increased risk for probiotic-associated bacteremia, these potential benefits must be weighed against this risk when considering the continued use of probiotics in the ICU.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-019-0626-9.

References
1. Sarah, H. Y., Jernigan, J. A. & McDonald, L. C. Prevalence of probiotic use among inpatients: a descriptive study of 145 US hospitals. *Am. J. Infect. Control* **44**, 548–553 (2016).
2. Szajewska, H. What are the indications for using probiotics in children? *Arch. Dis. Child.* **101**, 398–403 (2016).
3. Barraud, D., Bollaert, P.-E. & Gibot, S. Impact of the administration of probiotics on mortality in critically ill adult patients: a meta-analysis of randomized controlled trials. *Chest* **143**, 646–655 (2013).

4. Barraud, D. et al. Probiotics in the critically ill patient: a double blind, randomized, placebo-controlled trial. *Intensive Care Med.* **36**, 1540–1547 (2010).

5. Honeycutt, T. C. B. et al. Probiotic administration and the incidence of nosocomial infection in pediatric intensive care: a randomized placebo-controlled trial. *Pediatr. Crit. Care Med.* **8**, 452–458 (2007). quiz 464.

6. Suez, J. et al. Post-antibiotic gut mucosal microbiome reconstitution is impaired by probiotics and improved by autologous FMT. *Cell* **174**, 1406–1423.e16 (2018).

7. Zmora, N. et al. Personalized gut mucosal colonization resistance to empiric probiotics is associated with unique host and microbiome features. *Cell* **174**, 1388–1405.e21 (2018).

8. Kunz, A. N., Noel, J. M. & Fairchok, M. P. Two cases of *Lactobacillus* bacteremia during probiotic treatment of short gut syndrome. *J. Pediatr. Gastroenterol. Nutr.* **38**, 457–458 (2004).

9. Salminen, M. K. et al. *Lactobacillus* bacteremia, clinical significance, and patient outcome, with special focus on probiotic *L. rhamnosus* GG. *Clin. Infect. Dis.* **38**, 62–69 (2004).

10. Thomas, D. W. & Greer, F. R. Clinical report—probiotics and prebiotics in pediatrics. *Pediatrics* **6**, 1217–1231 (2010)

11. Ghouri, Y. A. et al. Systematic review of randomized controlled trials of probiotics, prebiotics, and synbiotics in inflammatory bowel disease. *Clin. Exp. Gastroenterol.* **7**, 473–487 (2014).

12. Theodorakopoulou, M., Perros, E., Giamarellos-Bourboulis, E. J. & Dimopoulos, G. Controversies in the management of the critically ill: the role of probiotics. *Int. J. Antimicrob. Agents* **42** (Suppl.), S41–S44 (2013).

13. Zhang, G.-Q., Hu, H.-J., Liu, C.-Y., Shakya, S. & Li, Z.-Y. Probiotics for preventing Late-onset sepsis in preterm neonates: a PRISMA-Compliant systematic review and meta-analysis of randomized controlled trials. *Medicine* **95**, e2581 (2016).

14. Siempos, I. I., Ntaidou, T. K. & Falagas, M. E. Impact of the administration of probiotics on the incidence of ventilator-associated pneumonia: a meta-analysis of randomized controlled trials. *Crit. Care Med.* **38**, 954–962 (2010).

15. Gu, W.-J., Wei, C.-Y. & Yin, R.-X. Lack of efficacy of probiotics in preventing ventilator-associated pneumonia probiotics for ventilator-associated pneumonia: a systematic review and meta-analysis of randomized controlled trials. *Chest* **142**, 859–868 (2012).

16. Oudhuis, G. J., Bergmans, D. C. J. J. & Verbon, A. Probiotics for prevention of nosocomial infections: efficacy and adverse effects. *Curr. Opin. Crit. Care* **17**, 487–492 (2011).

17. Lolis, N. et al. *Saccharomyces boulardii* fungaemia in an intensive care unit patient treated with caspofungin. *Crit. Care* **12**, 414 (2008).

18. Lebeer, S., Vanderleyden, J. & De Keersmaecker, S. C. J. Genes and molecules of lactobacilli supporting probiotic action. *Microbiol. Mol. Biol. Rev.* **72**, 728–764 (2008).

19. Salminen, M. K. et al. *Lactobacillus* bacteremia during a rapid increase in probiotic use of *Lactobacillus rhamnosus* GG in Finland. *Clin. Infect. Dis.* **35**, 1155–1160 (2002).

20. Chung, H. et al. Global and local selection acting on the pathogen *Stenotrophomonas maltophilia* in the human lung. *Nat. Commun.* **8**, 14078 (2017).

21. Lieberman, T. D. et al. Genomic diversity in autopsy samples reveals within-host dissemination of HIV-associated *Mycobacterium tuberculosis*. *Nat. Med.* **22**, 1470–1474 (2016).

22. Lieberman, T. D. et al. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat. Genet.* **43**, 1275–1280 (2011).

23. Smith, E. E. et al. Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc. Natl Acad. Sci. USA.* **103**, 8487–8492 (2006).

24. Watanabe, Y., Cui, L., Katayama, Y., Kozue, K. & Hiramatsu, K. Impact of *rpoB* mutations on reduced vancomycin susceptibility in *Staphylococcus aureus*. *J. Clin. Microbiol.* **49**, 2680–2684 (2011).

25. Hua, X. et al. Global effect of *rpoB* mutation on protein expression in *Enterococcus faecium*. *Jundishapur J. Microbiol.* **9**, e37322 (2016).

26. Campbell, E. A. et al. Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *EMBO J.* **24**, 674–682 (2005).

27. Enne, V. I., Delsol, A. A., Roe, J. M. & Bennett, P. M. Rifampicin resistance and its fitness cost in *Enterococcus faecium*. *J. Antimicrob. Chemother.* **53**, 203–207 (2004).

28. Lebeer, S. et al. Functional analysis of *Lactobacillus rhamnosus* GG pili in relation to adhesion and immunomodulatory interactions with intestinal epithelial cells. *Appl. Environ. Microbiol.* **78**, 185–193 (2012).

Methods

Patient inclusion criteria and clinical data. Eighteen cases of *Lactobacillus* bacteremia were identified through usual surveillance activities of the Infection Prevention and Control program at Boston Children's Hospital from January 2009 to June 2014 (6 patients receiving probiotics, 12 patients not receiving probiotics; Supplementary Table 1). Based on pharmacy records from January 2009 to June 2014, there were 15,736 probiotic doses administered to 645 ICU patients, including 5,859 (37%) in a medical ICU; 4,080 (26%) in an intermediate care program (ICP); 3,560 (23%) in a medical–surgical ICU; 2,114 (13%) in a cardiac ICU; and 123 (0.8%) in a neonatal ICU, all at a single center (Boston Children's Hospital). The majority (522 of 645, 81%) of the ICU patients who received probiotics, including all 6 patients who developed bacteremia, received a probiotic containing LGG. ICU patients who were prescribed probiotics received a median of 8 (IQR 3–23) doses, with a range of 1–347 doses. Twenty-four percent of doses were given by mouth, 62% by gastrostomy or jejunostomy tube, and 14% by nasogastric or nasojejunal tube. The average numbers of doses per month did not change significantly over time, with 217 doses per month in 2009, 216 doses per month in 2010, 244 doses per month in 2011, 268 doses per month in 2012, 239 doses per month in 2013, and 249 doses per month in 2014. Probiotics were administered to only 3% of ICU patients (645 ICU patients among a total of 22,174 patients admitted to these ICUs during the study period). Probiotics were most commonly prescribed because patients had been receiving them prior to ICU admission. There were no ICU-specific guidelines for probiotic administration. Nearly all patients had a CVL at the time of the bacteremia (only Patients R2, N1, and N5 did not), so nearly all of these bacteremias met CDC criteria for central line-associated bloodstream infection (CLABSI). We did not examine the details of probiotic doses administered to non-ICU patients. The study was approved by the Boston Children's Hospital Institutional Review Board.

Our study was not designed to assess the clinical impact of bacteremia, although we do note that these episodes of bacteremia manifested initially as clinically active infection and that nearly all of the patients were treated with intravenous antibiotics directed at *Lactobacillus*. CVLs were removed during treatment from two of the five patients receiving probiotics who had a CVL and from nine of the ten patients not receiving probiotics who had a CVL. Notably, two of the six cases of LGG bacteremia (patients R2 and R4) and one of the four cases of non-LGG *L. rhamnosus* bacteremia (patient N5) were considered by their treating physicians as potential contaminants or transient bacteremias and were not specifically treated with a long course of antibiotics. However, as patient R4 had a CVL, the bacteremia was classified as a CLABSI. None of the patients had endocarditis, and none died within 7 days of bacteremia.

Isolation of bacteria from probiotic capsules and blood. Blood isolates were frozen at the time of isolation by the clinical microbiology laboratory and were later streaked on MRS-agar plates, and the *Lactobacillus* species were identified by MALDI-TOF. To isolate DNA for whole-genome sequencing, single colonies were picked (16 colonies per probiotic product batch and a single colony for blood samples) to inoculate MRS broth (BD Biosciences, 288130), and overnight cultures were frozen. To isolate individual bacteria from probiotic capsules, we used two complementary techniques. For probiotic product batch 1, a capsule was streaked on a CDC Anaerobe Blood Agar plate, bacterial lawn was scraped off the plate, frozen (−80 °C, glycerol) and then streaked to single colonies on MRS-agar plates (BD Biosciences, 288210) incubated at 37 °C. For product batches 2 and 3, capsules were resuspended and thoroughly vortexed in PBS, streaked on MRS-agar plates incubated at 37 °C until colonies showed, and then re-streaked to purity.

Whole-genome sequencing of individual isolates and deep sequencing. For single isolates, DNA was extracted from frozen overnight bacterial cultures derived from single colonies picked as described above (Macherey-Nagel, NucleoSpin 96 Tissue). For capsule deep sequencing (probiotic product batches 2–6, Supplementary Table 2), DNA was extracted from 100 µl (>10[8] cells) of resuspended capsule (same kit as above). Batch 1 was not available for deep sequencing. Nextera sequencing libraries were prepared[29] and sequenced in an Illumina HiSeq 2500 machine in rapid mode to produce 125base paired-end reads. DNA extraction and library preparation for single-isolate deep sequencing used as control (see "Genomic analysis of capsule deep sequencing" below) was carried out in a similar manner to other single isolates, and sequencing was done in the same Illumina run as the capsule DNA deep sequencing.

Genomic data analysis of isolates. Illumina reads were filtered to remove reads contaminated by the Nextera adapter or low-quality bases (>2 bases with a Phred Score of <20), yielding an average of 1.13 M reads per sample (s.d. = 2.9 × 10[5]). These reads were aligned to indicated reference genomes using Bowtie 1.2.1.1, allowing a maximum of 3 mismatches per read. The fraction of aligned reads was used to determine the distance between isolates and reference genomes. Alignment to the closest reference genome (Genbank FM179322) was analysed further. Base calling was done using SAMtools and BCFtools 0.1.19. A genome position was identified as a SNP if more than a single allele was identified across isolates using a quality threshold of FQ < −80. Phylogeny was based on the identified SNPs and

was determined by the PHYLIP dnapars algorithm, which carries out unrooted parsimony.

Gene content analysis. For each isolate, a 'raw copy number' for each gene was calculated as the median base coverage across the gene divided by the median coverage across the genome of the isolate. To remove gene-specific biases, this raw copy number was further normalized by the median raw value of the gene across all isolates, yielding the gene copy number used to identify deleted genes. For Fig. 2, a similar analysis of genomic coverage was performed, in which for each 6 kb region the mean read coverage was divided by the median coverage across the genome and normalized by the median of this region-specific value across isolates.

Genomic analysis of capsule deep sequencing. Reads were filtered and aligned to the reference genome as described for single-isolate analysis above (GenBank: FM179322). This resulted in coverage of >97% of the reference genome. Per batch, median coverage of these positions was 372–1,268× (Supplementary Table 2). As a control, a single-isolate colony (batch 2, isolate 15) was also sequenced at high depth (median coverage 897×). To identify variable loci in the probiotic batches, for each probiotic batch and genomic position, we performed a Fisher's exact test to compare the number of reads calling the reference with the alternative base in the probiotics versus the single-isolate control. To control for multiple comparisons, a P value of $(0.05)/(\text{genome length}) = 1.66 \times 10^{-8}$ was used for calling within-batch diversity.

Bacterial strains and culture conditions for in vitro assays. *L. rhamnosus* bacteria were grown at 37 °C with 5% CO_2 for 48 h on Trypticase Soy Agar II with 5% sheep blood (BD Biosciences). Liquid culture was performed using MRS broth (Sigma-Aldrich) supplemented with 0.001% Tween 80 (MRST) (Sigma-Aldrich) at 37 °C with 5% CO_2 for 24 h statically. *Pseudomonas aeruginosa* PAO1 *galU* mutant[30] were grown at 37 °C overnight on Trypticase Soy Agar II (BD Biosciences). Liquid culture was performed using LB broth Miller (Fisher BioReagents) at 37 °C, at 200 rpm overnight. Modified TSB (mTSB) medium consisted of 15 g l⁻¹ TSB (BD Biosciences) and 20 g l⁻¹of Bacto-proteose peptone no. 3 (BD Biosciences) was used for the biofilm assay.

Biofilm assay. The assay for biofilm formation was based on a previous report, with minor modifications[31]. In brief, 3×10^7 colony-forming units (c.f.u.) were added to 200 µl of mTSB in three replicates in flat bottom polystyrene 96-well plates (Costar) at 37 °C with 5% CO_2 for 72 h. Bacteria were removed by inverting the plate. The plate was then washed with water, and attached bacteria were stained for 30 min with 200 µl 0.1% (wt/vol) crystal violet in an isopropanol–methanol–PBS solution (volume ratio of 1:1:18). Plates were washed with water, left to dry for 15 min, and then 150 µl of 33% glacial acetic acid was added to each well. Biofilm was measured at 570 nm (Versa max, Molecular devices).

Antibiotic resistance testing. Disk diffusion susceptibility testing was performed by the Boston Children's Hospital Infectious Diseases Diagnostic Laboratory using the standard methods for *Staphylococcus aureus* (specific disk diffusion methods for *Lactobacillus* are not available)[32].

Competition assay. The single culture and competition assays were based on a previous report, with modifications[33]. In brief, bacteria of the probiotic strain containing no SNPs (P3-2) and of the blood isolate R1, from overnight MRST liquid cultures were adjusted to an OD₆₀₀ of 0.05, either in single culture or mixed in a 1:1 ratio. During growth in MRST broth at 37 °C with 5% CO_2 for 24 h CFUs were determined every 2 h by serial dilutions on MRST-agar plates for the single cultures and on both MRST-agar and MRST-agar with 1 µg ml⁻¹ rifampicin (Research Products International) for the competition cultures.

BLASTP of RpoB protein. The protein accession numbers of the RpoB protein from the bacteria *LGG* (CAR88393.1), *S. aureus* M1112 (EWR31828.1), *Stahylococcus epidermidis* RP62A (AAW53580.1), *Enterococcus faecium* 343-3 (AAO00728.1), *E. faecium* 38-15 (AAO00731.1), *E. faecium* 40-4 (AAO00730.1), *Escherichia coli* K-12 substr. MG1655 (NP_418414.1), *Bacillus velezensis* CC09 (ANB47365.1) were used in COBALT for amino acid alignment from NCBI (https://www.st-va.ncbi.nlm.nih.gov/tools/cobalt).

Whole blood killing assay. Bacteria were grown on trypticase soy agar (TSA) with 5% sheep blood (BD Biosciences) and incubated for 48 h at 37 °C with 5% CO_2. Overnight cultures in MRST medium were washed once in PBS (Boston Bio-products) and adjusted to give 10⁶ CFU per 50 µl. Next, 50 µl of each strain were added to 450 µl of heparinized blood from a healthy donor. Inoculum c.f.u. values were determined by serial dilutions on TSA with 5% sheep blood. After 1 h and 3 h of rotation at 37 °C, serial dilutions were plated to determine the number of surviving c.f.u. In parallel, static tubes were held at 37 °C as a non-phagocytosis control for all time points (0 h, 1 h, and 3 h).

Serum sensitivity assays. Bacteria were grown on TSA with 5% sheep blood (BD Biosciences) and incubated 48 h at 37 °C with 5% CO_2. Static overnight cultures in

MRST medium at 37 °C with 5% CO_2 were washed in PBS once and then diluted in PBS plus 1 mM $CaCl_2$ and 1 mM $MgCl_2$, and 100 µl aliquots were placed in a sterile 96-well plate to give a final inoculum of approximately 5×10^6 c.f.u. per well. Pooled male, type AB human serum (Sigma-Aldrich) was diluted in PBS plus 1 mM $CaCl_2$ and 1 mM $MgCl_2$ to give twice the desired final concentration. Final serum concentrations that were used were 50% and 25%. Human serum (50%) that was heat-inactivated by incubation at 56 °C for 30 min, and 0% serum served as controls. Equal volumes (100 µl) of sera and bacterial suspensions were mixed and incubated at 37 °C for 1 h with gentle shaking. An aliquot from each well was serially diluted and then plated on TSA with 5% sheep blood after incubation for 48 h at 37 °C with 5% CO_2 for enumeration. A serum-sensitive, rough lipopolysaccharide (LPS) strain of *P. aeruginosa* (PAO1 *galU*[30]), grown on TSA at 37 °C overnight and then in liquid culture in LB at 37 °C overnight, was used as a positive control.

Case-control study methods. As all cases of *Lactobacillus* bacteremia in patients receiving probiotics occurred in an ICU, cases were matched with up to 3 control patients who had received probiotics in an ICU within 90 d of the case and had similar or longer ICU exposure prior to censoring. Controls were selected randomly using incidence density sampling. Censoring occurred on the date of bacteremia diagnosis or, for controls, on the date of discharge, death, or transfer from the ICU. Five cases had 3 controls identified, whereas 1 case with a particularly long ICU stay, had only 1 possible control identified. As a result, 6 cases were compared with 16 controls in our analysis.

Patient data for the 6 cases and 16 controls were collected retrospectively by chart review onto a standardized case report form. Temperature, white blood cell (WBC) count, C-reactive protein (CRP), and 30 d mortality were collected from the date of bacteremia or censoring. Clinical variables previously associated with either *Lactobacillus* bacteremia or with central line-associated bloodstream infections in general were collected for the 7 days prior to bacteremia, death or discharge[34–37].

Immunodeficiency was defined as active cancer diagnosis, solid organ or stem cell transplant, primary immunodeficiency, receipt of immunosuppressant within 6 weeks prior to censoring, or neutropenia (absolute neutrophil count (ANC) or total WBC count of <500 occurring for at least two days and within three calendar days before or after the date of censoring). Medical device data included endovascular prosthetic material, CVL, tracheostomy tube, gastrostomy tube, and urinary catheter. Gastrointestinal breakdown included documentation of mucositis, diarrhea, or skin breakdown around the gastrostomy or jejunostomy insertion site. Diarrhea was identified by documentation in the physician or nursing notes or by stool output of >20 ml kg^{-1} in a 24 h period as per the CDC definition of mucosal barrier injury[38]. Antibiotic data included information on oral and intravenous antibiotics regardless of indication.

For the case-control study, odds ratios for continuous and categorical variables were generated by exact conditional logistic regression using SAS 9.4.

Statistical analysis. Fig. 3b: Kruskal–Wallis test ($P = 0.0297$, Kruskal–Wallis statistic = 13.99) followed by Dunn's multiple comparisons test to P3-2 were performed. $\alpha = 0.05$. Statisical results relating to Fig. 3b: See Supplementary Table 12a.

Extended Data Fig. 4: 3 independent experiments were performed on different days. In each experiment, each bacterial isolate had 3 technical replicates. $P < 0.0001$ by ANOVA with Tukey's multiple comparisons test for the pairwise comparison of any of the isolates making biofilm (defined as $OD_{570} > 1$) compared to either P2-1, N5, N9, N10, N11, or medium control. There were no statistically significant differences among the isolates making biofilm or among the isolates not making biofilm. $F = 38.93$. d.f. = 42. For statistical results relating to Extended Data Fig. 4: see Supplementary Table 12b.

Supplementary Fig. 1a: 2 independent experiments on different days were performed. Error bars show the interquartile range of 3 technical replicates for all apart from 50% heat-inactivated (h.i.) PAO1 *galU* control, which had 2 technical replicates. *$P = 0.0448$ for PAO1 *galU* 50% serum versus 50% h.i. by Kruskal–Wallis test ($P = 0.0297$, Kruskal–Wallis statistic = 26.88) followed by Dunn's multiple comparisons test.

Supplementary Fig. 1b: 2 independent experiments were performed on different days. Error bars show s.d. of 3 technical replicates. The ratios t_{1h}/t_{0h} and t_{3h}/t_{0h}, ($t =$ time), were used for statistical analysis. For t_{1h}/t_{0h}, $P = 0.1893$ by one-way ANOVA followed by Dunnett's multiple comparisons test. $F = 1.677$. There were no statistically significant differences upon multiple pairwise comparisons against P3-2. $P = 0.7058$ for P3-2 versus P1-1, $P = 0.9998$ for P3-2 versus R1, $P = 0.5002$ for P3-2 versus R2, $P = 0.9020$ for P3-2 versus R3, $P = 0.9547$ for P3-2 versus R4, $P = 0.8192$ for P3-2 versus R5, $P = 0.2698$ for P3-2 versus R6. d.f. = 15. For t_{3h}/t_{0h}: $P = 0.1901$ by one-way ANOVA followed by Dunnett's multiple comparisons test. $F = 1.658$. There were no statistically significant differences upon multiple pairwise comparisons against P3-2. $P = 0.8893$ for P3-2 versus P1-1, $P = 0.9998$ P3-2 versus R1, $P = 0.8571$ for P3-2 versus R2, $P = 0.9998$ for P3-2 versus R3, $P = 0.8957$ for P3-2 versus R4, $P = 0.2034$ for P3-2 versus R5, $P = 0.9353$ P3-2 versus R6. d.f. = 16.

Extended Data Table 9: 3 independent experiments on different days were performed. In each experiment, 3 independent bacterial cultures of each strain

were used. Values shown are the median with 25% and 75% percentiles. For single culture experiments, doubling time: $P > 0.9999$, Mann–Whitney $U = 4$, number of divisions: $P > 0.9999$, Mann–Whitney $U = 4$ by unpaired two-tailed Mann–Whitney test. For the competition culture experiment, doubling time: $P = 0.1$, Mann–Whitney $U = 0$, no. of divisions: $P = 0.1$, Mann–Whitney $U = 0$ by unpaired two-tailed Mann–Whitney test.

Please refer to the Nature Research Reporting Summary for additional details.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data and code availability

Sequence data are available in the NCBI SRA repository under BioProjectID PRJNA562050 with accession numbers SRX6757122–SRX6757178. BioSample accession numbers are: SAMN12632778–SAMN12632834. Figure 3, Extended Data Figure 4, Supplementary Table 9, and Supplementary Figure 1 have associated raw data. All other data are available from the corresponding authors upon reasonable request.

References

29. Cai, X.-C. et al. Rifampicin-Resistance mutations in the *rpoB* Gene in *Bacillus velezensis* CC09 have pleiotropic effects. *Front. Microbiol.* **8**, 178 (2017).
30. Wi, Y. M. et al. Rifampicin resistance in *Staphylococcus epidermidis*: molecular characterisation and fitness cost of *rpoB* mutations. *Int. J. Antimicrob. Agents* **51**, 670–677 (2018).
31. Xu, M., Zhou, Y. N., Goldstein, B. P. & Jin, D. J. Cross-resistance of *Escherichia coli* RNA polymerases conferring rifampin resistance to different antibiotics. *J. Bacteriol.* **187**, 2783–2792 (2005).
32. Baym, M. et al. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS ONE* **10**, e0128036 (2015).
33. Priebe, G. P. et al. The *galU* Gene of *Pseudomonas aeruginosa* is required for corneal infection and efficient systemic spread following pneumonia but not for infection confined to the lung. *Infect. Immun.* **72**, 4224–4232 (2004).
34. Lebeer, S., Verhoeven, T. L. A., Perea Vélez, M., Vanderleyden, J. & De Keersmaecker, S. C. J. Impact of environmental and genetic factors on biofilm formation by the probiotic strain *Lactobacillus rhamnosus* GG. *Appl. Environ. Microbiol.* **73**, 6768–6775 (2007).
35. CLSI. *Performance Standards for Antimicrobial Susceptibility Testing.* 27th ed. Supplement M100. Wayne, P. A. Clinical and Laboratory Standards Institute (2017).
36. Mariam, D. H., Mengistu, Y., Hoffner, S. E. & Andersson, D. I. Effect of *rpoB* mutations conferring rifampin resistance on fitness of mycobacterium tuberculosis. *Antimicrob. Agents Chemother.* **48**, 1289–1294 (2004).
37. Wylie, M. C. et al. Risk factors for central line–associated bloodstream infection in pediatric intensive care units. *Infect. Control Hosp. Epidemiol.* **31**, 1049–1056 (2010).
38. De Groote, M. A., Frank, D. N., Dowell, E., Glode, M. P. & Pace, N. R. *Lactobacillus rhamnosus* GG bacteremia associated with probiotic use in a child with short gut syndrome. *Pediatr. Infect. Dis. J.* **24**, 278–280 (2005).
39. Ledoux, D., Labombardi, V. J. & Karter, D. *Lactobacillus acidophilus* bacteraemia after use of a probiotic in a patient with AIDS and hodgkin's disease. *Int. J. STD AIDS* **17**, 280–282 (2006).
40. Gouriet, F., Million, M., Henri, M., Fournier, P.-E. & Raoult, D. *Lactobacillus rhamnosus* bacteremia: an emerging clinical entity. *Eur. J. Clin. Microbiol. Infect. Dis.* **31**, 2469–2480 (2012).
41. See, I. et al. Mucosal barrier injury laboratory-confirmed bloodstream infection: results from a field test of a new national healthcare safety network definition. *Infect. Control Hosp. Epidemiol.* **34**, 769–776 (2013).

Author contributions

K.B.F., G.P.P., and T.J.S. conceived the study. R.K., I.Y, J.S., E.S., M.H., E.L., and P. McGann performed whole-genome sequencing and data analysis. C.M. performed

Extended Data Fig. 1 | Deep sequencing identifies loci of diversity across probiotic product batches. Five probiotic batches (batches P2-P6, see Supplementary Table 2) were sequenced at high depth together with a single colony. In each batch, for each position in the reference genome, a two-sided Fisher's exact test was carried out to determine differences in diversity between the batch-derived sequences and the colony-derived ones, and the respective P values were plotted. Significant loci ($P < 1.66 \times 10^{-8}$) are marked with labels A–O (for details see Supplementary Table 6). A single locus of increased diversity in the colony in comparison to only one of the probiotic batches (P3) was also observed (green).

a

Probiotic (H487)   Blood (H487D)   Overlap
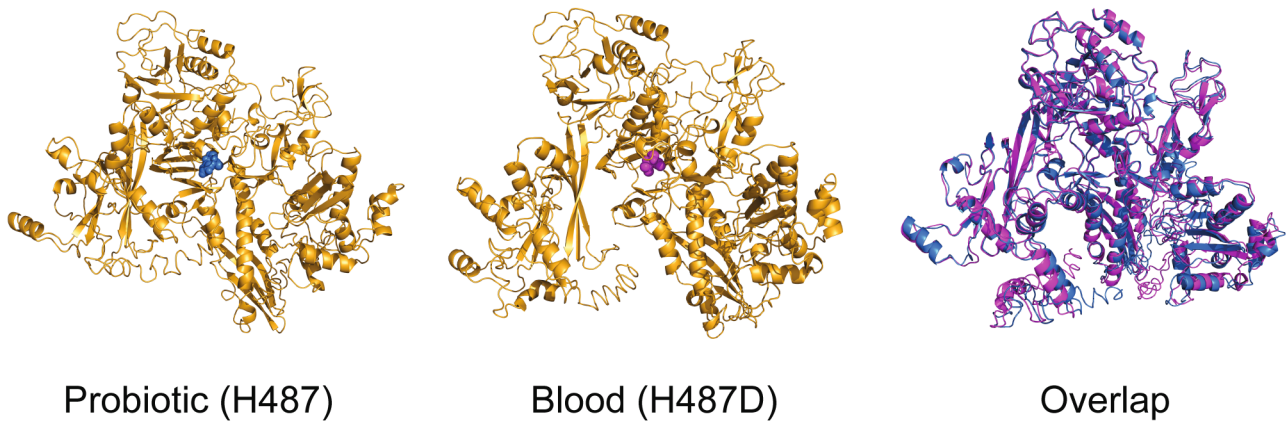
b

aa: 471, 474, 490, 651, 921, 929, 1039

DNA-binding site of RpoB
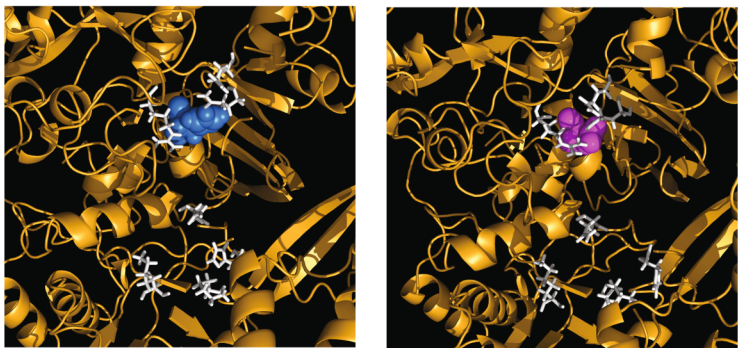
c

Rifampin Cluster I RpoB region

|			
Lactobacillus rhamnosus GG	464	KEFFGSSQLSQFMDQTNPLGELTHKRRLSALGPGGLTRDRAGYEV	508
Lactobacillus rhamnosus GG rif[R]	464	KEFFGSSQLSQFMDQTNPLGELTDKRRLSALGPGGLTRDRAGYEV	508
Staphylococcus aureus M1112	458	KEFFGSSQLSQFMDQANPLAELTHKRRLSALGPGGLTRERAQMEV	502
Staphylococcus aureus M1112 rif[R]	458	KEFFGSSQLSQFMDQTNPLGELTDKRRLSALGPGGLTRDRAGYEV	502
Staphylococcus epidermidis RP62A	458	KEFFGSSQLSQFMDQANPLAELTHKRRLSALGPGGLTRERAQMEV	502
Staphylococcus epidermidis RP62A rif[R]	458	KEFFGSSQLSQFMDQTNPLGELTYKRRLSALGPGGLTRDRAGYEV	502
Enterococcus faecium 343-3	466	KEFFGSSQLSQFMDQTNPLGELTHKRRLSALGPGGLTRDRAGYEV	510
Enterococcus faecium 343-3 rif[R]	466	KEFFGSSQLSQFMDQTNPLGELTYKRRLSALGPGGLTRDRAGYEV	510
Enterococcus faecium 40-4	466	KEFFGSSQLSQFMDQTNPLGELTHKRRLSALGPGGLTRDRAGYEV	510
Enterococcus faecium 40-4 rif[R]	466	KEFFGSSQLSQFMDQTNPLGELTYKRRLSALGPGGLTRDRAGYEV	510
Enterococcus faecium 38-15	466	KEFFGSSQLSQFMDQTNPLGELTHKRRLSALGPGGLTRDRAGYEV	510
Enterococcus faecium 38-15 rif[R]	466	KEFFGSSQLSQFMDQTNPLGELTQKRRLSALGPGGLTRDRAGYEV	510
Escherichia coli str. K12 substr. MG1655	503	KEFFGSSQLSQFMDQNNPLSEI THKRRI SALGPGGLTRERAG FEV	547
Escherichia coli str. K12 substr. MG1655 rif[R]	503	KEFFGSSQLSQFMDQNNPLSEI TYKRRI SALGPGGLTRERAG FEV	547
Bacillus velezensis CC09	459	KEFFGSSQLSQFMDQNNPLSEI THKRRI SALGPGGLTRERAG FEV	503
Bacillus velezensis CC09 rif[R]	459	KEFFGSSQLSQFMDQNNPLSEI TYKRRI SALGPGGLTRERAG FEV	503
Bacillus velezensis CC09 rif[R]	459	KEFFGSSQLSQFMDQNNPLSEI TDKRRI SALGPGGLTRERAG FEV	503
Bacillus velezensis CC09 rif[R]	459	KEFFGSSQLSQFMDQNNPLSEI TRKRRI SALGPGGLTRERAG FEV	503
Bacillus velezensis CC09 rif[R]	459	KEFFGSSQLSQFMDQNNPLSEI TCKRRI SALGPGGLTRERAG FEV	503
		* * * * * * * * * * * * * * * * *** * * * *** ***** * *** ** **	

Extended Data Fig. 2 | see figure caption on next page.

Extended Data Fig. 2 | The blood-isolate-specific *rpoB* SNP does not perturb the RpoB predicted structure but occurs near the DNA-binding site and is associated with rifampin resistance in other bacterial species. (**a**) Predicted structures of *L. rhamnosus* GG RNA polymerase β-subunit RpoB with histidine at position 487 seen in the probiotic (blue, left), aspartic acid at position 487 seen in the blood isolate from Patient R1 (magenta, middle), and overlap (right). (**b**) Predicted DNA-binding site amino acids are shown in white, with the histidine (blue) of the probiotic (left) and the aspartic acid (magenta) of blood isolate from Patient R1 (right) shown compared to the DNA-binding positions. (**c**) Amino acid (aa) sequence alignment of the rifampin cluster I of the RpoB protein from *L*GG and other genera. Numbering begins and ends at the first and last aa of the cluster; asterisks depict evolutionarily conserved aa residues; red asterisk shows the conservation across species of the histidine. In magenta, aa substitution H487D of the *L. rhamnosus* GG rifampin-resistant isolate (Patient R1) found in this study, H481D of *S. aureus* M1112 rifampin-resistant isolate[24], and H482D of *B. velezensis* rifampin-resistant isolate[39]; in orange, substitution H481Y of *S. epidermidis* RP62A rifampin-resistant isolate[40], H489Y of *E. faecium* 343-3 rifampin-resistant isolate[27], H489Y of *E. faecium* 40-4 rifampin-resistant isolate[27], H526Y of *E. coli* K-12 substr. MG1655 rifampin-resistant isolate[41], and H482Y of *B. velezensis* rifampin-resistant isolate[39]; in lavender, substitution H489Q of *E. faecium* 38–15 rifampin-resistant isolate[27]; in brown, substitution H482R of *B. velezensis* rifampin-resistant isolate[39]; in turquoise, substitution H482C of *B. velezensis* rifampin-resistant isolate[39].
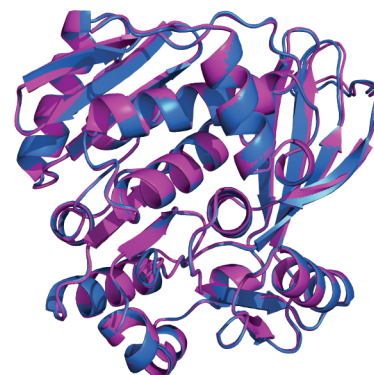
a
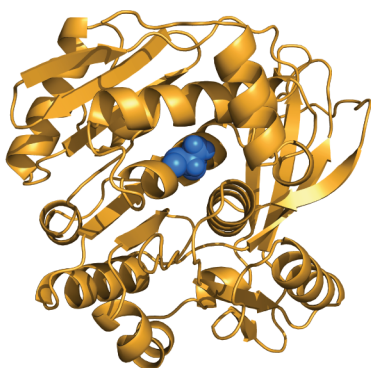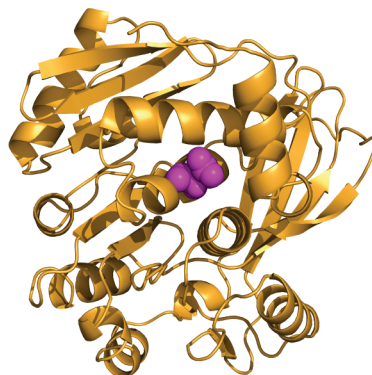
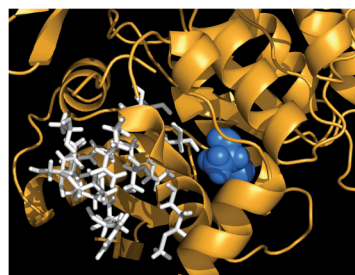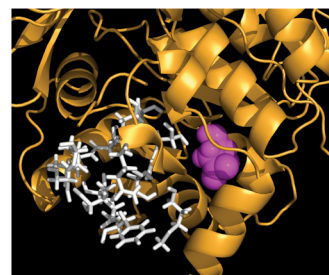Probiotic (A259)   Blood (A259D)   Overlap

b

aa: **186, 222, 223, 224, 225, 227, 241, 242, 244, 246,252, 253, 256, 279, 282, 286**

ADP-binding site of ribokinase
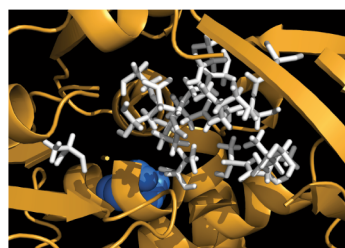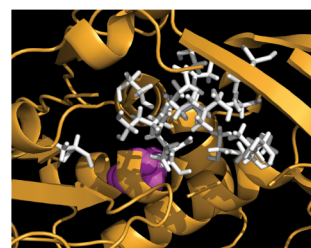

Probiotic (A259)   Blood (A259D)

aa: **10, 12, 14, 39, 40, 41, 44, 108, 110, 139, 141, 250, 251, 254, 291**

**Adenosine-binding site of ribokinase**

Probiotic (A259)   Blood (A259D)

Extended Data Fig. 3 | The blood-isolate-specific ribokinase SNP does not perturb the predicted structure of ribokinase but occurs near the active site. (**a**) Predicted structures of probiotic ribokinase with A259 (blue, left), blood isolate from Patient R1 with ribokinase A259D SNP (magenta, middle) and overlap (right). (**b**) The predicted binding site amino acids of ribokinase for adenosine are shown in white, with the alanine 259 (blue) of the probiotic (left) and the aspartic acid (magenta) of blood isolate 1 (right) shown compared to the adenosine-binding positions.
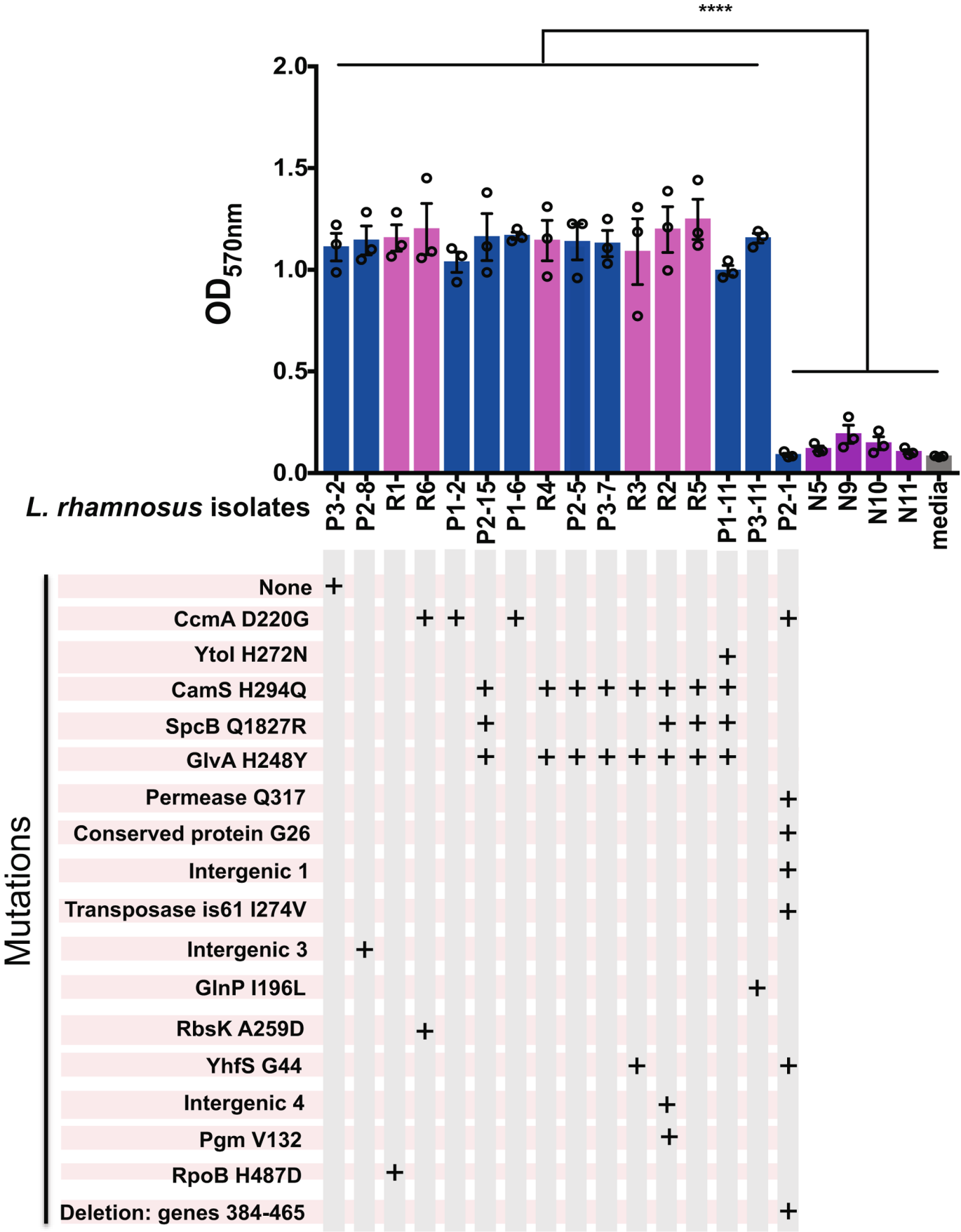
Extended Data Fig. 4 | see figure caption on next page.

Extended Data Fig. 4 | Biofilm formation of probiotic and blood *L. rhamnosus* isolates. Blood isolates from patients receiving (R1-R6) and those not receiving probiotics (N5, N9, N10, N11), as well as selected probiotic isolates, were tested for biofilm formation. Isolates are grouped by similar mutations, as depicted in the grid below the isolate labels. Isogenic probiotic isolates from different probiotic capsules were used as controls, if available, as were controls for mutations found in blood isolates, when available. In Px-y, x is probiotic batch number, y is probiotic isolate number. Bars represent means of three independent experiments performed on different days, with three technical replicates per isolate in each experiment. Error bars depict the s.e.m. ****$P < 0.0001$ by ANOVA followed by Tukey's multiple comparisons test for the pairwise comparison of any of the isolates making biofilm (defined as $OD_{570} > 1$) compared to either P2-1, N5, N9, N10, N11, or medium control. There were no statistically significant differences among the isolates making biofilm or among the isolates not making biofilm.

nature research

Corresponding author(s):   Thomas J. Sandora, Roy Kishony, Gregory P. Priebe

Last updated by author(s): Sep 20, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Illumina HiSeq 2500 software (HCS version 2) was used for DNA sequencing.
Data analysis	Illumina sequencing data analysis tools: Bowtie 1.2.1.1 for short read alignment, SAMtools and BCFtools 0.1.19 for SNP calling. Custom code was used to automate these tools and visualize output. COBALT was accessed on the NCBI website: https://www.st-va.ncbi.nlm.nih.gov/tools/cobalt

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Sequence data are available under BioProjectID PRJNA562050. BioSample accession numbers are: SAMN12632778-SAMN12632834. Figure 3, Extended Data Figure 4, Supplementary Table 9, and Supplementary Figure 1 have associated raw data. All other data are available from the corresponding authors upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All patients that developed Lactobacillus bacteremia while receiving probiotics during our study period were included. For the case-control analysis, a 3:1 matching scheme was chosen given the small sample size and potential variability in control subjects. Sample size calculations were not performed for the overall study. Given the infrequency of our outcome, all patients that developed Lactobacillus bacteremia while receiving probiotics during our study period (n=6) were included. The study period was chosen based on the availability and reliability of infection prevention and microbiology data during this time period.
Data exclusions	There were no data exclusions
Replication	All in vitro experiments were performed 2-3 times, and all findings were replicated.
Randomization	As a retrospective study, randomization was not performed.
Blinding	As a retrospective case-control study, blinding was not performed. Data analysis was done according to methods pre-specified before data collection to prevent bias during analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☒ ☐	Palaeontology
☒ ☐	Animals and other organisms
☐ ☒	Human research participants
☒ ☐	Clinical data

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Six patients with Lactobacillus bacteremia and sixteen matched controls, all from the intensive care unit, had their charts reviewed for clinical factors that could be associated with the development of bacteremia. As a retrospective study, consent was waived by our IRB (protocol #P00018055). The covariate-relevant population characteristics of the participants in the case-control study are shown in Supplementary Table 11. For the whole blood killing assays, blood was obtained from human volunteers with informed consent on a separate IRB-approved protocol (protocol #P00025848).
Recruitment	See above.
Ethics oversight	Boston Children's Hospital IRB

Note that full information on the approval of the study protocol must also be provided in the manuscript.