Human gut microbes' transmission, persistence, and contribution to lactose tolerance

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
John Liam Fitzstevens
aus Rochester, New York, Vereinigte Staaten

Tübingen
2023

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	08.12.2023
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatterin:	Prof. Dr. Ruth E. Ley
2. Berichterstatterin:	Prof. Dr. Nadine Ziemert
3. Berichterstatter:	Prof. Dr. Andrew H. Moeller

Table of Contents

Acknowledgements

I'd like to thank Ruth Ley for having mentored me the past 5.5 years. I remember sitting in her office as a young student, losing track of time as we talked about how microbial transmission may function between people, how infants acquire microbes and what their fate is - things that were only abstract ideas at the time. I couldn't have asked for a more efficacious advisor in making the transition from ideation to hypothesis-testing, and story-telling. The most important thing I've learned here is that big things don't happen without big teams, and in this thesis, there are three keystone postdocs and one group leader: Victor Schmidt, Taichi Suzuki, Xiaoying Liu, and Nicholas Youngblut. From Victor I learned the difference that being friendly makes, and how to work in the lab; from Taichi how to craft a maximally-compelling narrative; from Xiaoying, half the things I know about *Bifidobacterium*; and from Nick, how to become a bioinformatician. I'd like to thank Nadine Ziemert and Detlef Weigel who, alongside Ruth and Nick, served on my thesis advisory committee.

The return to the office from COVID reaffirmed how important human-contact is in work; my different iterations of office-clusters have been the foundation of that. First with Hagay Enav, Andrea Bórbon, and Claudi Frick, followed by Jess Sutter, and lastly with Kelsey Huus, Stacey Heaver and Meghna Chatterjee. Albane Ruaud, Jacobo de la Cuesta, and Tanja Schön, Guillermo Luque, thank you for helping me analyze data and stay up on state-of-the-art bioinformatics. I'd like to thank our teams at the Centre de Recherches Médicales de Lambaréné, 108 Military Hospital in Hanoi, and Institute for Tropical Medicine Tübingen, in particular, Mirabeau Mbong, Nguyen Thu Ha, and Anne Pfleiderer for being so reliable and committed to excellence.

Clear thinking is done by a clear mind; I have many people to thank for catalyzing my mental and emotional hygiene. Firstly, my parents, Chris and Mark Fitzstevens, and my sister, Maia Fitzstevens. Next, the people with whom I spent the most time over the past 5 years, my WG: Mai Lan Doan, Tom Schellenberg, Jasmin Schels, and Felix Räsch. Next, Schwimmen für alle Kinder, in particular Dagmar Müller and Gerd Müller, who gave me the chance to give back to my local community and have become my family away from home. Thank you Klara Röbl, Franziska Ecker and Manus Schmedding for the Gute Laune Brigade memories in Gabon, and Pauline von Boehmer, for teaching me Schwäbisch. Healthy bodies catalyze clear minds: thank

you Mikel Zhobro, Jia-Jie Zhu, Ghadeer Elmkaiel, Felix, and Albane for the workouts, Nils Rohbohm and Nico Spatscheck for the cycling, and Marco Podobnik for attacking all 20 alpine-hiking stages of the Via Alpina together.

I'd like to thank Alex Baugh, my biology professor at Swarthmore College who instilled in me the importance of forethought, and who mentored me through decision-making after my bachelor's degree, including speaking highly of the Max Planck Society. Thank you Elizabeth Brownell, who took a chance on me as a 20-year old summer intern, and first opened my mind to the concept of the human gut microbiome, which is what led me down this path.

Summary

Biology and medicine are focused heavily on human genetics' impact on the body, which is reasonable as the first human genome was completely sequenced only last year. There are, however, 1,000 times more genes in the bacteria, viruses and archaea that live in the gut - collectively referred to as its microbiota - than in the human genome itself. They augment the genes of their human hosts, and their presence in the gastrointestinal tract, where nutrients are broken down and absorbed, suggest that consideration of the human genome alone may afford an incomplete understanding of how the body functions. Biomedical data have a reputation for being noisy; much of what is understood today as randomness might be explainable in the future if we transition to a more integrative model of human biology - one that considers the genomes of both the host and microbiota.

Here, I study how gut microbes have been involved in human adaptation to different environments. Since this topic is broad, I divided it into three aims. Aim one examines whether the microbiome accounts for the fact that genetics does not always predict phenotype in lactose tolerance. Aim two studies whether gut microbes are acquired from humans' mothers. Aim three determines whether gut microbes acquired during infancy persist into, and throughout adulthood. Together, these aims address how humans acquire their gut microbes, for how long they persist in the gut, and what the metabolic consequences for the host are.

In aim one, I investigate the role of gut microbiota in lactose tolerance. This is because the lactose tolerance phenotype is more prevalent than the genotype, lactase-persistence. To address this discrepancy, I traveled with my colleague Victor Schmidt to Gabon and Vietnam, where lactase persistence is rare. We genotyped over 800 women, including in Germany, where lactase persistence is common, and measured their lactose tolerance phenotype. Approximately 20% of lactase-persisters were lactose tolerant by clinical standards. I then tested the hypothesis that this phenotype was driven by their microbiota by surveying their stool metagenomes, but saw no differences between tolerant and intolerant lactase non-persisters. To gain a better understanding of how the stool microbiota metabolize lactose, my colleague Xiaoying Liu cultured the stools with lactose in vitro, enabling measurement of not only H_2, but also whether the lactose is metabolized at all. Her results showed that some stools produced gas from lactose, others metabolized little to no lactose, and the rest metabolized lactose into soluble products.

These results showed that lactose tolerance could come from either no metabolism of lactose, or metabolism into soluble products. The genus *Bifidobacterium* was associated with the latter. Our results for the first time link microbial metabolism and taxa to non-genetic lactose tolerance.

Aim one's results show that consideration of the human genome alone does, in fact, afford an incomplete understanding of how the body functions. This impacts how we understand our ancestors' adaptation to novel environments during their migration out of Africa. Dairying preceded the evolution of genetic lactose tolerance over 10,000 years ago: we propose that *Bifidobacterium* initially facilitated dairying by making it tolerable, because without any dairying, there would have been no positive selection for lactase persistence. The microbe and host-genetic adaptation yield the same outcome, in that they have enabled humans to exploit ruminants as converters of grass to energy in the form of milk. We conclude from this that *Bifidobacterium* has been buffering selection for the lactase-persistence genotype for at least 15,000 years.

In my second aim, I examined mother-to-child transmission of microbial strains in Germany, Gabon and Vietnam. I hypothesized that mothers share more gut microbial strains with their own infants than those of other mothers. To test this, I collected stool samples from not only the aforementioned women in Gabon, Vietnam, and Germany, but also from their children. I assembled genomes from both sets of stool metagenomes to determine the extent to which gut microbes from different mothers and infants are related. We used this to determine that strains of microbes, particularly of the genus *Prevotella*, are indeed shared vertically (between mothers and their children). We observed this vertical transmission in all three sampled countries, but also to a lesser extent, horizontally (between unrelated mothers and infants) in Gabon and Vietnam. This suggests that gut microbial strains persist in human populations in host lineage-specific fashion.

My third aim was to determine the temporal stability of microbes in the human gut. My hypothesis was that the evidence for vertical lineage-specificity of gut microbes shown in aim two is predicated on strains acquired during infancy persisting throughout hosts' lives. To test this, I collected stools from a 13 year-old human, and sequenced them alongside stools frozen since his infancy. Indeed, we found the persistence of two microbial strains: one belonging to the species *Prevotella copri* - the same taxon for which we found abundant evidence of vertical transmission - and the other to *Alistipes_A ihumii*. They had both accrued less than than 40 single nucleotide polymorphisms since the participant's infancy, which is within the expected

range that microbes have been shown to evolve at in the human gut. They shared several phenotypes with the species we'd previously shown to codiversify with humans, namely a high sensitivity to oxygen. We also identified three strains that persisted in the gut of his mother despite the same passage of 13 years, suggesting that gut microbes can persist not only from infancy to adolescence, but also throughout adulthood.

Taken together, these three aims' results indicate that gut microbiota exist in an intimate metabolic association with their hosts. They are acquired vertically, but also horizontally in communities with reduced dispersal limitation. They can confer a phenotype previously thought to be conferred only by the host's own genome: lactose tolerance. And we could demonstrate that they persist in our guts, from the period of infant-acquisition through adolescence but also throughout adulthood - for at least 13 years, thereby enabling their eventual transmission to a next generation of infant hosts. This is of great consequence to evolutionary biology. It means that evolution acts not solely on the human genome, but on the genomes of humans in concert with those of their microbial gut symbionts. These aims demonstrate how an integrative understanding of host-microbe interplay yields a more refined understanding of the human body, blurring the line between evolutionary biology and modern medicine.

Zusammenfassung

Biologie und Medizin konzentrieren sich stark auf die Auswirkungen der menschlichen Genetik auf den Körper, vorangetrieben durch die erste vollständige Sequenzierung des menschlichen Genoms im letzten Jahr. Jedoch gibt es 1.000-mal mehr Gene in den Bakterien, Viren und Archaea, die im Darm leben und kollektiv als Mikrobiota bezeichnet werden, als im menschlichen Genom selbst. Diese mikrobiellen Gene erweitern die Gene ihrer menschlichen Wirte und ermöglichen somit zusätzliche metabolische Prozesse. Ihre Anwesenheit im Magen-Darm-Trakt, wo Nährstoffe abgebaut und aufgenommen werden, legt nahe, dass allein die Betrachtung des menschlichen Genoms kein vollständiges Verständnis dafür bietet, wie der Körper funktioniert. Eine große Hürde für biomedizinische Daten ist ihre Ungenauigkeit, aufgrund der eingeschränkten Möglichkeiten alle relevanten physiologischen Variablen zu erfassen und zu bewerten. Vieles von dem, was heute als Zufälligkeit verstanden wird, könnte in der Zukunft erklärbar sein, wenn wir zu einem integrativen Modell der menschlichen Biologie übergehen - einem Modell, das die Genome sowohl des Menschen als auch ihrer Mikrobiota berücksichtigt.

In meiner Forschung untersuche ich, wie Darmmikroben in die Anpassung des Menschen an unterschiedliche Umgebungen involviert waren. Da dieses Thema breit angelegt ist, habe ich es in drei Teile unterteilt. Teil eins untersuchte, ob das Mikrobiom dafür verantwortlich ist, dass die Gene des Menschen nicht immer den Phänotyp bei Laktosetoleranz vorhersagen vermag. Teil zwei untersuchte, ob eine Übertragung von Darmmikroben von den Müttern auf die Kinder in Menschen stattfindet. Teil drei fokussierte sich auf die Frage, ob Darmmikroben, die während der Kindheit erworben wurden, im Erwachsenenalter bestehen bleiben. Gemeinsam adressieren diese Fragestellungen, wie Menschen ihre Darmmikroben erwerben, wie lange sie im Darm verbleiben, und welche Stoffwechselkonsequenzen sich daraus für den Wirt ergeben.

Im Rahmen von Ziel Nummer eins untersuche ich die Rolle der Darmmikrobiota bei der Laktosetoleranz. Dies liegt daran, dass der Phänotyp der Laktosetoleranz häufiger vorkommt als der Genotyp der Laktasepersistenz. Um dies zu untersuchen, reiste ich zusammen mit meinem Kollegen Victor Schmidt in zwei Länder, Gabun und Vietnam, in welchen Laktasepersistenz selten vorkommt. Zusammen mit Daten aus Deutschland, wo Laktasepersistenz weit verbreitet ist, sammelten wir Daten von 800 Frauen. Neben der Bestimmung ihres

Laktosetoleranz-Phänotyps wurden sie genotypisiert. Etwa nur 20% der laktasepersistenten Personen waren nach klinischen Maßstäben laktosetolerant. Anschließend testete ich die Hypothese, dass dieser Phänotyp der phänotypische Laktosetolenranz trotz genetischer Laktoseintoleranz durch das Mikrobiota angetrieben wird. Hierfür untersuchte ich ihre Stuhlmetagenome, stellte aber keine Unterschiede zwischen phänotypisch toleranten und nicht toleranten genotypischer laktoseintoleranten Personen fest. Um ein besseres Verständnis dafür zu bekommen, wie das Mikrobiota Laktose metabolisieren kann, führte meine Kollegin Xiaoying Liu *in vitro* Experimente durch. Sie kultivierte Stuhlproben mit Laktose und ermöglichte H2-Messungen sowie direkte Messungen der Laktose-Metabolisierung. Ihre Ergebnisse zeigten, dass einige Mikrobiota aus Stuhlproben Gas aus Laktose produzierten, andere wenig bis keine Laktose verstoffwechselten, und der Rest Laktose in lösliche Produkte metabolisierte. Schlussfolgerlich könnte Laktosetoleranz entweder durch keinen Laktosestoffwechsel oder durch einen Stoffwechsel in lösliche Produkte entstehen könnte. Mit Letzterem wurde die Gattung Bifidobacterium in Verbindung gebracht. Somit verknüpfen unsere Ergebnisse erstmals den mikrobiellen Stoffwechsel und Taxa mit der nicht-genetischen Laktosetoleranz.

Die Ergebnisse von Ziel eins zeigen, dass die Betrachtung des menschlichen Genoms allein kein vollständiges Verständnis dafür bietet, wie der Körper funktioniert. Dies hat Auswirkungen auf unser Verständnis der Anpassung unserer Vorfahren an neue Umgebungen während ihrer Migration aus Afrika. Die Milchproduktion erfolgte vor mehr als 10.000 Jahren, bevor sich die genetische Laktosepersistenz entwickelte: Wir schlagen vor, dass Bifidobacterium die Milchproduktion anfangs ermöglicht hat, indem es sie verträglich machte, denn ohne Milchproduktion hätte es keine positive Selektion für Laktasepersistenz gegeben. Die Anpassung von Mikroben und Wirtsgenetik führt zu demselben Ergebnis, nämlich dass sie es den Menschen ermöglicht haben, Wiederkäuer als Umwandler von Gras in Energie in Form von Milch zu nutzen. Daraus folgern wir, dass Bifidobacterium die Selektion für das Laktasepersistenz-Genotyp seit mindestens 15.000 Jahren gepuffert hat.

In meinem zweiten Teil habe ich die Übertragung von mikrobiellen Stämmen von Müttern auf ihre Kinder in Deutschland, Gabun und Vietnam untersucht. Ich habe die Hypothese aufgestellt, dass Mütter mehr Darmmikrobenstämme mit ihren eigenen Kindern teilen als mit Kindern anderer Mütter. Um dies zu überprüfen, habe ich Stuhlproben nicht nur von den zuvor erwähnten Frauen in Gabun, Vietnam und Deutschland, sondern auch von ihren Kindern

gesammelt. Ich habe Genome aus beiden Sets von Stuhlmetagenomen zusammengestellt, um den Grad der Verwandtschaft zwischen Darmmikroben verschiedener Mütter und Kinder zu bestimmen. Wir haben festgestellt, dass Mikrobenstämme, insbesondere der Gattung *Prevotella*, tatsächlich vertikal geteilt werden (zwischen Müttern und ihren Kindern). Wir haben diese vertikale Übertragung in allen drei untersuchten Ländern beobachtet, aber auch in geringerem Maße horizontal (zwischen nicht verwandten Müttern und Kindern) in Gabun und Vietnam. Diese Beobachtungen legen nahe, dass Darmmikrobenstämme in menschlichen Populationen in einer spezifischen Weise persistieren, die von der Abstammungslinie des Wirts abhängt.

Mein drittes Ziel war es, die Stabilität von Mikroben im menschlichen Darm während des menschlichen Lebens zu bestimmen. Meine Hypothese war, dass die vertikale Abstammung von Darmmikroben, die in Teil 2 gezeigt wurde, auf dem Bestehenbleiben der mikrobiellen Stämme, die während der Kindheit erworben wurden, im Verlauf des Lebens beruht. Um dies zu testen, analysierte ich Stuhlproben von einem 13-jährigen Menschen zusammen mit Stuhlproben, die seit seiner Kindheit eingefroren waren. Tatsächlich fand ich zwei mikrobiellen Stämme, die beständig seit früher Kindheit an in der Versuchsperson vorhanden waren: Der erste Stamm gehörte zu *Prevotella copri* - dasselben Taxon, das bereits im vorherigen Abschnitt mit vertikaler Übertragung im Menschen in Zusammenhang gebracht wurde. Der zweite Stamm gehörte zu *Alistipes_A ihumii*. Beide hatten seit der Kindheit des Teilnehmers weniger als 40 Einzelnukleotidpolymorphismen akkumuliert, was im erwarteten Bereich der mikrobiellen Entwicklung von Genen im menschlichen Darm liegt. Diese zwei Stämme teilen einige Phänotypen mit Mikroben, die gemeinsam mit Menschen diversifiziert sind. Hierzu gehört eine hohe Empfindlichkeit gegenüber Sauerstoff. Des weiteren identifizierten wir auch drei Stämme, die im Darm seiner Mutter trotz des gleichen Zeitraums von 13 Jahren persistierten. Die Ergebnisse beider Versuchspersonen weisen darauf hin, dass Darmmikroben nicht nur von der Kindheit bis zur Adoleszenz, sondern auch im Erwachsenenalter persistieren können.

Zusammenfassend zeigen die Ergebnisse dieser drei Ziele, dass Darmmikrobiota in enger metabolischer Verbindung zu ihren Wirten stehen. Sie werden vertikal, aber auch horizontal in Gemeinschaften mit eingeschränkter Ausbreitung übertragen. Sie können einen Phänotyp verleihen, der zuvor nur durch das eigene Genom des Wirts vermittelt wurde: Laktosetoleranz. Und wir konnten zeigen, dass sie in unseren Därmen bestehen bleiben, von der Phase des Erwerbs in der Kindheit über die Adoleszenz bis hin zum Erwachsenenalter - für mindestens 13

Jahre, was ihre spätere Übertragung auf eine nächste Generation von Kind-Wirten ermöglicht. Dies ist von großer Bedeutung für die Evolutionsbiologie. Dies bedeutet, dass die Evolution nicht nur auf das menschliche Genom wirkt, sondern auf die Genome der Menschen im Zusammenspiel mit denen ihrer mikrobiellen Darm-Symbionten. Diese Ziele zeigen, wie ein integriertes Verständnis der Wechselwirkung zwischen Wirt und Mikrobe zu einem präziseren Verständnis des menschlichen Körpers führt und die Grenze zwischen Evolutionsbiologie und moderner Medizin verwischt.

Publications

Published

TA Suzuki*, **JL Fitzstevens***, VT Schmidt, H Enav, K Huus, M Mbong, A Grießhammer, A Pfeiderer, BR Adegbite, JF Zinsou, M Esen, TP Velavan, AA Adegnika, LH Song, TD Spector, AL Muehlbauer, N Marchi, H Kang, L Maier, R Blekhman, L Ségurel, G Ko, ND Youngblut, P Kremsner, RE Ley, (2022) **Codiversification of gut microbiota with humans**. Science 2022.10.12.462973.
Scientific ideas/data generation/analysis & interpretation/writing: 30/80/20/5%

Advanced manuscripts; awaiting submission to target journals

JL Fitzstevens*, X Liu*, A Ruaud, VT Schmidt, M Mbong, J Rauch, T Suzuki, TH Nguyen, A Arzamasov, D Rodionov, BR Adegbite, JF Zinsou, M Esen, TP Velavan, AA Adegnika, L Huu Song, P Kremsner, A Osterman, ND Youngblut, AV Tykaht, RE Ley, ***Bifidobacterium*** **confer lactose tolerance in genetically-intolerant humans**.
Scientific ideas/data generation/analysis & interpretation/writing: 70/60/80/95%

JL Fitzstevens, M Angenent, LT Angenent, RE Ley, **Microbiota persist in human infant and adult guts for up to 13 years**.
Scientific ideas/data generation/analysis & interpretation/writing: 90/100/100/100%

* co-first author

Introduction

Homo sapiens accomplished an incredible feat during their 300,000 years of existence: populating planet earth (Beyer et al. 2021). Explanations for how this was possible are dominated by narratives about how the forces of evolution - mutation, gene flow, genetic drift, and natural selection - have acted on the human genome (Fan et al. 2016; Marciniak and Perry 2017; Mathieson 2020). Such a focused effort on the human genome is important, as it is indeed the primary unit upon which natural selection acts, and was not fully sequenced until 2022 (Moran and Sloan 2015; Nurk et al. 2022). The discovery of billions of bacteria, viruses and archaea inhabiting the human gastrointestinal tract - the gut microbiota - suggest that humans are part microbial: the 20,000 genes of their own genomes are complemented by those of their gut microbes, which exceed a staggering 22 million (Willyard 2018; Tierney et al. 2019). This genetic supplementation yields the profound implication that microbes may participate in host adaptation to new environments (Suzuki and Ley 2020). It suggests that humans of the same genotype could have different fitness because of their gut microbiota, and that natural selection can, at times, act not on one in isolation, but rather on the two together (Moran and Sloan 2015). This thesis focuses on a single question: How do gut microbes assist in human adaptation to new environments? This question is vast - too vast for one thesis to address comprehensively - so it is broken into three aims.

Aim 1 uses the microbiome to account for human phenotype-genotype disparity. Candidate human genetic adaptations are well documented (Fan et al. 2016). The microbiota may have played a role in how human genotypes arose (Suzuki and Ley 2020), and several of them have already been positively or negatively associated with microbial taxa: *AMY1* (a high-starch diet) with *Ruminococcus*, *PLD1* (energy harvest) with *Akkermansia*, and *LCT* (dairying) with *Bifidobacterium (Everard et al. 2013; Goodrich et al. 2017; Poole et al. 2019; Suzuki and Ley 2020)*. More humans consume dairy than have the genetic adaptation to it (Blekhman et al. 2015; Goodrich, Davenport, Beaumont, et al. 2016; Goodrich, Davenport, Waters, et al. 2016), and those without the genetic adaptation have more *Bifidobacterium (Bonder et al. 2016)*. We tested whether *Bifidobacterium* confers lactose tolerance in the absence of the host-genetic adaptation - both in Gabon and Vietnam, where the host-genetic adaptation is rare, and in Germany, where it is prevalent.

Aim 2 examines whether or not gut microbes are transmitted between consecutive host generations. If gut microbes perform consequential tasks for the host, but are not transmitted to the next generation (vertical transmission), those benefits may be lost. To determine whether vertical transmission exists, we surveyed the stool metagenomes of mothers and their children from Gabon, Vietnam and Germany, and identified the extent to which microbial strains were shared between them. Previous studies had identified strain-sharing events between mothers and their infants, but their sampling was heavily-biased towards western populations (Asnicar et al. 2017; Yassour et al. 2018; Korpela et al. 2018). As a result of this, the impact of westernization on how gut microbes pass from one generation of human host to the next remains largely unknown. These previous studies determined strain-relatedness by mapping metagenome reads to developer-curated reference databases (Nayfach et al. 2016; Shi et al. 2022; Beghini et al. 2021). This dilutes the differences between a given dataset's strains with the differences they have with strains in the western-biased reference databases. We overcame this by mapping metagenome reads to genomes that we ourselves assembled from Gabon, Vietnam and Germany (Olm et al. 2021), reducing database-bias in the first ever study to comparatively examine microbial transmission in both westernized and non-westernized human populations.

It is not known for how long a gut microbe can survive in the human gut. This is because longitudinal datasets cost time, and no one has ever tested whether microbial strains persist in the human gut for longer than 5 years (Bäckhed et al. 2015; Yassour et al. 2016; Chu et al. 2017; Vatanen et al. 2018). These same studies also suffered from the aforementioned pitfall of reference-based alignment bioinformatics (Bäckhed et al. 2015; Yassour et al. 2016; Chu et al. 2017; Vatanen et al. 2018). Here, we had the unprecedented opportunity to collect stool samples from an adolescent, from whom we also had frozen stool samples from their first 2.5 years of life. We surveyed the stool metagenomes at each time point, spanning his first 2.5 and 12th - into his 13th - year of life, assembled microbial genomes, and determined strain relatedness. We also did the same for his mother, enabling us to test whether strains persist in the human gut not only during the transition from infancy to adolescence, but also throughout adulthood.

Taken together, these three aims advance our understanding of how gut microbes have, and continue to contribute to humans' adaptation to diverse environments on earth. In aim one, we studied one such adaptation - that of adults to dairy - and whether *Bifidobacterium* contributes to it. Aims two and three studied how such contributions are possible; whether gut

microbes are ubiquitous amongst humans or inherited vertically from their mothers, and whether such microbes persist after infant-acquisition into, and throughout, adulthood. We conclude that infants acquire gut microbes from their mothers, which persist in their guts not only into adolescence but also into adulthood, throughout which they are capable of performing tasks once thought to be only conferrable by the host's own genome. Host-genetic explanations for humans' numerous adaptations to life on earth are incomplete without considering potential interactions with, and contributions from, their gut microbes.

Chapter one: *Bifidobacterium* confer lactose tolerance in genetically-intolerant humans

(Advanced manuscript; awaiting submission to target journal)

J. Liam Fitzstevens[1†], Xiaoying Liu[1†], Albane Ruaud[1], Victor T. Schmidt[1], Mirabeau Mbong Ngwese[1], Julia Rauch[1], Taichi Suzuki[1], Nguyen Thu Ha[2,3], Aleksandr Arzamasov[4], Dmitry Rodionov[4], Bayode R. Adegbite[5,6], Jeannot F. Zinsou[5,6], Meral Esen[5,7,8], Thirumalaisamy P. Velavan[2,5], Ayola A. Adegnika[5,6,7,9], Le Huu Song[2,3], Peter Kremsner[5,6,7,8], Andrei Osterman[4], Nicholas D. Youngblut[1], Alexander V. Tyakht[1], and Ruth E. Ley[1,7*]

[*] Corresponding author. Email: rley@tuebingen.mpg.de

[†] These authors contributed equally to this work.

[1] Department of Microbiome Science, Max Planck Institute for Biology, Tübingen, Germany.

[2] Vietnamese-German Center for Medical Research, Hanoi, Vietnam.

[3] 108 Military Central Hospital, Hanoi, Vietnam.

[4] Sanford Burnham Prebys Medical Discovery Institute, La Jolla, California, USA.

[5] Institute for Tropical Medicine, University of Tübingen, Tübingen, Germany.

[6] Centre de Recherches Médicales de Lambaréné, Lambaréné, Gabon.

[7] German Center for Infection Research, Tübingen, Germany.

[8] Cluster of Excellence EXC 2124 Controlling Microbes to Fight Infections, University of Tübingen, Tübingen, Germany.

[9] Fondation pour la Recherche Scientifique, Cotonou, Bénin.

Abstract

The lactase-persistence (LP) genotype allows digestion of the milk sugar lactose in adults and confers lactose tolerance. Genetically lactase non-persistent (LNP) individuals can also be lactose tolerant, but responsible microbiota remain elusive. Here, we assessed lactose tolerance as H_2-production in breath after lactose dose, LP/LNP genotype, and gut microbiome metagenomic diversity in 483 adults from Gabon (100% LNP), Vietnam (99% LNP), and Germany (23% LNP). In all three populations, ~ 20% of LNP were lactose tolerant though microbiomes differed. In-vitro lactose addition to stool showed low H_2 production stemmed either from minimal breakdown of lactose, or breakdown producing metabolites of the Bifid shunt pathway - lactate and acetate - and the growth of *Bifidobacterium*. Our results indicate that *Bifidobacterium* can confer lactose tolerance across populations, including where the LP genotype is rare, and may have facilitated functional take-over by the human genome when dairying first began 12,000 years ago.

Introduction

Lactose is human milk's primary carbohydrate and infants' main energy source (Asp and Dahlqvist 1974; Ségurel and Bon 2017). Lactase, a beta-galactosidase, hydrolyzes it in the small intestine into glucose and galactose, which are absorbed into the blood (Asp and Dahlqvist 1974; Ségurel and Bon 2017). Lactase is encoded by the LCT gene, whose transcription typically downregulates after weaning (Ingram et al. 2009; Anguita-Ruiz, Aguilera, and Gil 2020). In such genetically lactase non-persistent (LNP) individuals - who comprise the majority of humans - dietary lactose passes unhydrolyzed from the small to large intestine, where it is fermented by gut microbiota. This produces gasses (particularly hydrogen), bloating and abdominal pain: the lactose intolerant phenotype (Misselwitz et al. 2013). Lactase persistence and lactose tolerance are most commonly diagnosed by performing an overnight fast, consuming lactose and measuring whether there are substantial increases in blood glucose (indicative of lactase persistence) and breath hydrogen (indicative of lactose intolerance) (Itan et al. 2010). However, sometimes LNP individuals do not make hydrogen, which poses an enigma.

One of the strongest and most recent examples of natural selection on the human genome is the LCT transcriptional enhancer *MCM6*, which for approximately one third of humans extends lactase expression from infancy to adulthood, known as the lactase persistent (LP) genotype (Ingram et al. 2009; Ségurel and Bon 2017; Anguita-Ruiz, Aguilera, and Gil 2020). This positive selection occurred between 7,000 - 10,000 years ago in independent dairying societies, each with different *MCM6* single nucleotide polymorphisms (Gallego Romero et al. 2012; Ségurel and Bon 2017; Jones et al. 2015). LP confers lactose tolerance, but the modern prevalence of the phenotype exceeds that of the genotype (Itan et al. 2010; Ranciaro et al. 2014; Hollfelder et al. 2021). The same dissonance is reported in our ancestors: dairying evolved as a cultural practice 5,000 years before the emergence of LP (Ségurel and Bon 2017). This is often explained by cultural evolution with Mongolians as an example, where less than five percent of the population is LP but fermented dairy consumption is prevalent (Jeong et al. 2018; Curry 2018). Fermentation, however, only reduces the concentration of lactose in dairy by 20 - 50 % (Alm 1982). It has been suggested that the gut microbiota may confer lactose tolerance to genetically intolerant individuals (Ranciaro et al. 2014; Jeong et al. 2018; Hollfelder et al. 2021; Segurel et al. 2020).

A microbe with the metabolic potential to confer such tolerance is the genus *Bifidobacterium*, whose species can uptake lactose, hydrolyze it with intracellular beta-galactosidase and ferment resulting glucose with the "Bifid shunt" pathway (Parche et al. 2006; Catanzaro, Sciuto, and Marotta 2021). Microbiome genome-wide association studies (GWAS) performed in the UK and USA suggest this may be the case, as LNPs harbor more *Bifidobacterium* than do LPs (Blekhman et al. 2015; Goodrich, Davenport, Beaumont, et al. 2016; Goodrich, Davenport, Waters, et al. 2016); an association dependent on host milk-consumption (Bonder et al. 2016). This led to the hypothesis that *Bifidobacterium* can outcompete LNP hosts for lactose (Goodrich, Davenport, Waters, et al. 2016; Suzuki and Ley 2020), nevertheless being in symbiosis because its metabolism produces short-chain fatty acids (SCFAs) as energy for the host. Here, we tested not only whether this competition exists, but also what the consequence for the host is; namely whether it could confer tolerance to the milk sugar. Consuming dietary lactose infused with live *Bifidobacterium* spp. reduces breath hydrogen (T. Jiang, Mustapha, and Savaiano 1996), but whether this capability innately exists within LNPs guts remains unstudied.

To investigate the role of the gut microbiome in lactose tolerance, we conducted a human intervention study in which we surveyed stool microbiomes, assessed tolerance by measuring hydrogen and glucose from lactose, and genotyped LP/LNP. We enrolled 483 participants in three countries: Gabon (n = 152), Vietnam (n = 190) and Germany (n = 141). Approximately 20% of LNP participants in each country were lactose tolerant, which we termed "microbially-acquired lactose tolerance" (MALT). There were however no differences in stool metagenomes of MALT and intolerant LNP individuals. To characterize the microbial processes underlying this phenotype, we cultivated a subset of LNP stool samples with lactose (n = 149), sequenced their metagenomes and measured the metabolites produced. Stools responded in one of four different ways: little hydrolysis of lactose without fermentation ("Inactive"), little hydrolysis with weak fermentation ("Weak"), hydrolysis and fermentation producing high hydrogen ("Gassy"), or hydrolysis and fermentation, producing low hydrogen ("Tolerant"). The tolerant group was enriched in *Bifidobacterium* both in-vivo and -vitro, and the metabolic byproducts of canonical Bifid shunt pathway, lactate and acetate. We conclude that MALT is a widespread trait, comprised of LNPs with metabolically-active and inactive microbiota.

Bifidobacterium drives metabolically-active MALT, and likely enabled humans to practice dairying before the evolution of LP.

Methods

Recruitment

Participants qualified for the study in Gabon (n = 152), Vietnam (n = 190) and Germany (n = 141) with the following inclusion criteria: (1) between 18-40 y old, (2) born in the country-of-testing, (3) no severe allergy to dairy, (4) not currently pregnant, (5) have no known underlying medical conditions that may reasonably put themselves or other participants at risk, and (6) willingness to follow the low-starch requirements for their pre-fast meal, fast overnight and not consume anything prior to arrival at the study location (Supp Table [Metadata_and_phenotyping]). Our field staff screened participants upon arrival at the clinic and explained the above criteria, and all protocols, risks, and study motivations in local languages before informed consent was provided. The following local ethical bodies authorized our protocols for use in this study: Gabon: Comité National d'Ethique, protocol number: N0025/2017/PR/SG/CNE; Vietnam: Scientific Ethics Review Committee and 108 Military Central Hospital, protocol number: 108MCH/RES/VGCARE-03-16012018; Germany: Ethik-Kommission an der Medizinischen Fakultät der Eberhard-Karls-Universität und am Universitätsklinikum Tübingen, protocol number: 529/2018BO1.

Lactose tolerance phenotyping, genotyping, and dietary surveys

We assessed lactose tolerance using breath hydrogen tests (BHTs) and blood glucose tests (BGTs) according to standard clinical protocols (Rezaie et al. 2017). Briefly, we measured baseline breath gas and blood glucose levels, followed by a 25 g dose of lactose monohydrate (Campro Scientific, Berlin, Germany) dissolved in 250 ml of water. We subsequently measured breath gas and blood glucose levels at 30, 60, 120, and 180 min after the dose. We used AlveoSampler bags and syringes (Quintron, Milwaukee, USA) to collect breath exhalant and measured its hydrogen, methane and carbon dioxide with a daily-calibrated BreathTracker SC (Quintron, Milwaukee, USA). We reported gas values after correction for CO_2, as per manufacturer's instructions. We collected BGT data with Accu-Chek Guide blood glucose

monitors as per manufacturer's instructions (Roche, Basil, Switzerland). We operationalized lactose tolerance as an increase in breath hydrogen below 30 ppm and lactase persistence as an increase in blood glucose equal to or greater than 20 mg / dl. We used a chi-squared test to determine whether the HBT or BGT was a better predictor of LP.

To minimize fermentation in the large intestine from non-lactose sources on the day of testing, participants were restricted to a set of approved low-starch ingredients for dinner the evening prior (e.g., white rice, white pasta, boiled chicken, eggs). They were also required to conduct an overnight fast for at least 12 h, and were not permitted to consume any food or drink besides water during the fast and study duration (from 8:00 AM until 12:00). Smoking and exercise were not permitted the morning of their visit to the clinic.

We assessed lactase persistence genotypes by collecting a saliva sample from each participant with Saliva DNA Collection and Preservation Devices (Norgen, Thorold, Canada). We extracted salivary genomic DNA using PowerSoil DNA extraction kits (Qiagen, Hilden, Germany) and sequenced host DNA on an Infinium Global Screening Array (Illumina, SanDiego, USA) at the University Hospital of Bonn, Life & Brain Research Centre. We assessed the genotype of each participant at three lactase persistence SNPs (rs41525747, rs4988235, and rs41380347) using Plinkv1.9 (Purcell et al. 2007). Participants collected stool from themselves using a Fe-Col collection kit (Alpha Labs, Eastleigh, UK), which we froze within 8 h of submission at -20 °C. We transported stool samples to Germany on dry ice and stored them at -80 °C.

Participants conducted a health and dietary survey where the primary objectives were to assess the relative degree of dairy intake across the different populations studied and whether or not participants had symptoms during testing. Since each location had a unique suite of dairy products available (e.g. condensed milk in Gabon, yogurt-coffee in Vietnam, or pizza in Germany), we were unable to standardize questionnaires across regions. We therefore applied a lactose score to each product in each of our regional surveys, based on available nutritional information for the product. We then quantified the total lactose intake for each participant based on the sum across all dietary products with lactose. Finally, we built a distribution of lactose scores across all regions, and applied an arbitrary yet directly comparable 'lactose score' to each participant. We note that these scores can only be used for comparisons within this study and do not reflect a generalizable assessment of lactose intake in g / day.

Associating blood glucose and breath hydrogen with LP genotype

Given that we measured two types of lactose tolerance phenotype data - both breath hydrogen and blood glucose tests - we wanted to determine which of the two was a better predictor of the LP genotype. To do this, we performed a χ^2-test of genotype versus TRUE/FALSE phenotype (glucose rise > 20 mg / dl or H_2 rise > 30 ppm) with 1 degree of freedom on the 481 samples for whom we had genotype, H_2 rise, and glucose rise data.

Generalized linear models

For generalized linear models, all features and interactions of interest were included in the model and sequentially removed until all were significant. When possible, contrasts were adjusted for the Country variable. For instance in Model 1, given the lower H_2 rise for German individuals as seen by eye (Figure [In_vivo_phenotyping]), contrasts were set to first evaluate the effect of "Germany" versus "Gabon" and "Vietnam" and then of "Gabon" versus "Vietnam". The p-values were adjusted using the Benjamini-Hochberg method with the number of all tested variables and number of contrasts of categorical variables (e.g., "Country" was counted twice since two tests were performed). Note that for Model 1, which was a two-part model; (i) sequential removing of variables was conducted independently for each part, (ii) p-value correction was done for both parts together, (iii) predicted values were computed such that if the first part of the model predicted H_2 rise > 1 ppm with a probability > 0.5, then the predicted value of the second part was used, otherwise the predicted value was 0.

Microbiome sequencing

We extracted genomic DNA from frozen stool with PowerSoil DNA extraction kits (Qiagen, Hilden, Germany). We prepared shotgun metagenomic libraries as per (Karasov et al. 2018) with slight modifications. Briefly, 1 ng of purified gDNA was used in a Nextera (Illumina, San Diego, USA) Tn5 tagmentation reaction to fragment and ligate adaptors in a single reaction, followed by a 14-cycle PCR to add sample-specific barcodes. We purified libraries with Mag-Bind TotalPure NGS beads (Omega Biotech, Norcross, USA), which we pooled and quantified as above. We size-selected libraries to 400-700 base pairs with a BluePippen (Sage Science, Beverly, USA). We concentrated and purified them further as needed using DNA Clean

& Concentrator-5 (Zymo Research, Irvine, USA) and sequenced them on a HiSeq 3000 (Illumina, San Diago, USA) with 150 paired end cycles.

Data processing and profiling

We performed metagenome quality-control with a pipeline described by Youngblut et al. (Youngblut et al. 2020). Briefly, we used Skewer 0.2.2 (H. Jiang et al. 2014) and the bbtools "bbduk" command to trim adapters and the bbtools "bbmap" command to filter reads that mapped to the human genome (GRCh37/hg19). We assessed read quality with Fastqc 0.11.7 and multiQC 1.5a. We subsampled metagenomes to 5 million reads and profiled their taxonomies with KrakEN 2 and Bracken 2.2, and genes and pathways with HUMAnN 3.0.0.alpha.3 (Beghini et al. 2021), each using custom databases curated by Struo2 (Youngblut and Ley 2021) with GTDB release 95 (Parks et al. 2022).

Before performing machine learning on the 483 in vivo metagenomes, we filtered their profiling output to taxa of prevalence ≥ 6.5 % (2719 genera) and pathways of prevalence ≥ 5 % (p = 125 pathways). We used the same cutoffs before running machine learning on the 299 in vitro metagenomes, which yielded 943 genera and 121 pathways at t = 0 h (149 metagenomes) and 1033 genera and 144 pathways at t = 4.5 h (150 metagenomes).

Extracellular β-galactosidase activity in LNPs

We aliquoted 100 - 200 mg of stool in 1.5 ml Eppendorf tubes with 100 mM MES buffer at pH 6 (6 µl buffer / mg fecal sample) and vortexed them at 2000 rpm for 5 min until thoroughly homogenized. We centrifuged the tubes at 15,000 rpm for 30 min and stored the supernatants at -20 °C before performing the β-galactosidase activity assay (Promega Corp., Madison, WI). Briefly, we added 50 µl supernatant and 50 µl 2 X Assay buffer in 96-well plates, incubated them at 37 ºC for 30 min and measured the absorbance at 420 nm every 2 min, vortexing the plates at 120 rpm for 40 s before each measurement. We generated standards by correlating the commercial β-galactosidase concentration (U / µl) with the slope of the absorbance. One unit β-galactosidase hydrolyzes 1.0 µmols of o-nitrophenyl-b-D-galactopyranoside (ONPG) to o-nitrophenol and galactose per min at pH 6 and 37 ºC. We recorded extracellular β-galactosidase activity in U / g feces.

Amending LNP stools with lactose

We aliquoted 0.5 g fecal samples with 4 mL 30 mM lactose buffered with 100 mM MES at pH 6 in Hungate anaerobic culture tubes in a glove box. We performed in vitro cultivation at 37 ºC for 4.5 h with continuous shaking at 600 rpm. We measured headspace H_2, CH_4 and CO_2 concentrations at 1.5 h, 3 h and 4.5 h with a gas chromatograph (SRI8610C; SRI Instruments, Torrence, USA) with a packed column (3' x 1/8" S.S. HayeSep D Packed Teflon; Restek, Bellefonte, USA) at 50 ºC with a thermal conductivity detector (TCD) at 120 ºC and a flame ionization (FID) detector. We used N_2 carrier gas with an input pressure of 207000 Pa. We calculated the headspace gas concentrations with the ideal gas equation, with units mmol of gas in the headspace per liter of culture.

We measured non-gaseous metabolites by collecting liquid samples before and after cultivation. We centrifuged them at 12,000 g for 30 m, the supernatants from which we filtered through a 0.22 μm syringe filter and preserved at -20 °C for HPLC analysis. We measured the supernatants' concentrations of lactose and the metabolites glucose, galactose, acetate, lactate, propionate, butyrate, succinate, glycerol and ethanol in supernatants with HPLC (Shimadzu LC40, Kyoto, Japan) with an Aminex HPX-87H column (Bio-Rad, California, USA). We performed separations at a flow rate of 0.6 mL / minute at 60 °C with a refractive index detector, a UV detector and 5 mM H_2SO_4 as the mobile phase. We calculated carbon consumption by subtracting the total mass of lactose, glucose and galactose at the end of incubation from the total mass of lactose (mg) at the beginning.

Comparing in vitro and in vivo H_2 production

To test whether samples that produced low hydrogen in vitro similarly produced low hydrogen in vivo, we collapsed the Inactive, Tolerant and Weak groups into one "low" in vitro H_2 production group and compared its in vivo H_2 production to that of the sole "high" group: Gassy. We compared their means with a Wilcoxon test.

Metagenome analysis

Overview

We performed all analyses in R (R Core Team 2018) using the tidyverse R package suite (Wickham et al. 2019). We used the phyloseq R package to rarefy in vivo metagenomes to their minimum even depth (324,778) before calculating observed and Shannon diversity metrics with its "estimate_richness" function (McMurdie and Holmes 2013). We calculated Bray Curtis distance matrices on the non-rarefied in vivo metagenomes with phyloseq's "distance" function; ordinating and plotting them with the "ordinate" and "plot_ordination" functions. We calculated differential abundance tests for in vivo metagenones by country with the DESeq2 R package (Love, Huber, and Anders 2014); for the in vitro metagenomes, we used a Wilcoxon signed-rank test to determine differences microbial growth on lactose, followed by a Benjamini-Hochberg p-value correction. We performed Wilcoxon, ANOVAs and Tukey HSD posthoc tests with the ggpubr R package "stat.test" and "stat_compare_means" functions (Kassambara 2020).

Machine learning models

We performed model selection on 120 cross-validation (CV) sets with 80-20 % train-test dataset split. We fitted the model with highest average predictive accuracy on the test sets to all data. For models with relative abundances of taxa, we performed feature selection on the train set before fitting the predictive model. Our feature selection was adapted from the gRRF algorithm (Deng and Runger 2013) to perform faster with the ranger R package (Wright, Wager, and Probst 2020) and to take the taxonomic structure into account when species and genus relative abundances were included (Ruaud et al. 2022). For feature selection, we tuned the gamma parameter for feature selection, as well as the k parameter when taking the taxonomy into account. For the predictions of random forest models, the number of trees and maximal depth were tuned; random forests were computed with the ranger R package (Wright, Wager, and Probst 2020). We compared models using one-sided Wilcoxon signed-rank tests with accuracies of CV sets matched between the two models. Models were interpreted with endoR (Ruaud et al. 2022).

For the in vitro models, we first filtered out all taxa that did not significantly increase in relative abundance over the course of the experiment in one of the groups (one-sided Wilcoxon signed-rank test and adjusted p-value < 0.05).

Genome assembly from metagenomes

We assembled metagenomes as outlined by Suzuki and colleagues (Suzuki et al. 2022), from all three sets of our metagenomes: the in vivo set, and those from the beginning and end of cultivation with lactose in vitro. Briefly, we used metaSPAdes v3.15.4 to assemble contigs. We differentially coverage-binned them with MetaBat v2.15.0, MaxBin v2.7.7 and VAMB v3.0.2, and selected the highest-quality, non-redundant MAGs with Das-Tool v1.1.4. Lastly, we dereplicated the MAGs with CheckM2 v1.0.1, and dRep v2.0.0 set to 99.99 % average nucleotide identity (ANI). We taxonomically-classified these non-redundant MAGs with GTDB-Tk database release 95 (Chaumeil et al. 2019), of which 209 were of the genus *Bifidobacterium*.

Culturing and sequencing of *Bifidobacterium* isolates

We isolated *Bifidobacterium* species from frozen fecal samples stored in -80 °C. We obtained feces from 10 adult individuals each from Vietnam (Son La, mean age = 25.7 years old, mean weight = 52.7 kg, mean infant age = 7.7 months old) and Gabon (Babongo, mean age = 30.8 years old, mean weight = 44.2 kg, mean infant age = 8.5 months). The isolation protocol was based on a previous study (Fouhy et al. 2015) with modifications. We grew all cultures under anaerobic conditions using BD GasPak EZ (East Rutherford, NJ, US).

For the isolation process, we serially diluted one gram of frozen fecal sample (10^{-4}–10^{-7}) in maximum recovery diluent (MRD) (Oxoid Ltd, Basingstoke, Hampshire, UK). We plated the dilutions in triplicate on Man Rogosa Sharpe (MRS, Oxoid, Basingstoke, UK) agar supplemented with 0.05% (wt/vol) L-cysteine hydrochloride (Merck, Germany), 5% (w/v) lactose (Win Lab, Gemini House, Middlesex, Hab 7ET, UK), 100 µg/ml Mupirocin (Fluka), and 50 units of Nystatin. We then incubated the plates at 37 °C for 72 hours, and picked approximately 50 colonies per individual from plates with the highest dilution. We cultured individual colonies in liquid MRS media in 2 ml 96-well deep plates. We extracted DNA using PureLink Pro 96 Genomic DNA Kit (Invitrogen, Carlsbad, CA, USA) and included negative

control samples in every plate. We quantified DNA concentrations using the PicoGreen protocol and the Gen5 microplate reader.

To further screen for Bifidobacterium spp., we developed a qPCR screening approach using the KiCqStart SYBR Green qPCR ReadyMix (Sigma-Aldrich, St. Louis, MO, USA) and previously described *Bifidobacterium*-specific primers (F-bifido: 5'-CGCGTCYGGTGTGAAAG-3', and R-bifido: 5'-CCCCACATCCAGCATCCA-3') (Delroisse et al. 2008). We retained isolates showing melting temperatures between 86.5-88.0°C, Cq values < 22.0, and DNA concentration > 0.25 ng/ul. This screening approach excluded around 30 % of the total isolates that were non-*Bifidobacterium*, resulting in 735 *Bifidobacterium*-candidate isolates for subsequent sequencing. We performed prepared libraries and sequenced isoaltes in the same way as for the aforementioned stool metagenomes.

Isolate-genome assembly and bioinformatic processing

We assembled post-quality control reads de novo on a per-sample basis. We subsampled reads to 1,000,000 reads per sample using seqtk v1.2. We normalized reads using the bbnorm command of bbtools v37.78 (target = 50, k = 31, minkmers = 15, and prefilter = t) and assembled contigs with SPAdes v3.11.1. The assembly parameters included setting off coverage cutoff, applying the "careful" parameter, and setting the minimal scaffold length to 500. Subsequently, we refined assemblies using Pilon v1.22 with a chunk size of 10,000,000, and assessed their quality with CheckM v1.0.16. We performed anannotation with Prokka v1.13, and assigned taxonomies with CheckM and sourmash v2.0.0a4 (scaled = 10,000 and k = 31). Out of the 735 candidate-*Bifidobacterium* isolates, 531 isolates classified as belonging to the Bifidobacteriaceae family. Furthermore, 503 isolates exhibited CheckM-estimated completeness > 95 % and contamination < 10 %. We dereplicated those isolate genomes to ANI 99.999 % and taxonomically-classified them with the same software versions as described for the MAGs, which yielded 59 non-redundant *Bifidobacterium* isolate genomes.

Phenotyping *Bifidobacterium* isolate genomes and MAGs

We phenotyped the 209 *Bifidobacterium* MAGs together with the 59 *Bifidobacterium* isolate genomes with Traitar v3.0.1, and inferred their phylogeny with PhyloPhlAn v3.0.3, which we annotated with iTOL v5 (Asnicar et al. 2020; Letunic and Bork 2021). We further predicted

whether or not each *Bifidobacterium* genome could metabolize particular carbohydrates, ranging from monosaccharides, disaccharides and polysaccharides using a modified binary metabolic phenotyping-pipeline originally described by Iablokov and colleagues (Iablokov et al. 2021). To avoid false-positives and -negatives for carbohydrate phenotypes, we filtered the aforementioned genomes to satisfy "high-quality" MIMAG standards (CheckM2-estimated completeness >= 90 % and contamination <= 5 %), qualifying 139 of the original 209 for carbohydrate phenotyping. (Bowers et al. 2017; Chklovski et al. 2022). We used a linear mixed model to model the total number of carbohydrates metabolizable by each genome, with GTDB species-designation as a random effect, and completeness, contamination, genome type (MAG or isolate genome), and breath hydrogen production from lactose as fixed effects. We repeated the same analysis with binary logistic regression - whether or not each carbohydrate had a "1" or '0" phenotype for each genome - with the same aforementioned explanatory variables.

We had three high-quality *Bifidobacterium* MAGs from participants whose baseline hydrogen measurements were above 30 ppm - likely due to failure to comply with the overnight fasting protocol - which excluded them from the preceding breath hydrogen-based analyses. That has, however, no bearing on our determination of whether carbohydrate metabolic potential of *Bifidobacterium* differs by country, so we added those three to the set of 139, and repeated the same generalized linear- and binary logistic regression-models without breath-hydrogen as an explanatory variable.

Strain-sharing analysis

We generated sub-species representative genomes (SSRGs) by dereplicating our isolate and metagenome-assembled genomes at 98 % average nucleotide identity with CheckM2 v1.0.1 and dRep v2.0.0. We identified single nucleotide polymorphisms by mapping our metagenome reads to those SSRGs with inStrain v1.0.0, thereby generating inStrain profiles; one per metagenome per SSRG. We did this for not only the adult participants of our study, but also their children. We pairwise-compared all profiles with `inStrain compare` to produce popANI metrics (Olm et al. 2021). As per inStrain documentation, we defined a shared strain as any pairwise comparison where at least 50% of the SSRG is compared, and the popANI is at least 99.999%.

Results

Recruitment & metagenomic libraries

We recruited 483 participants for this study; 152 in Gabon (Mean Age 26.4 y ± 0.51 SE), 190 in Vietnam (Mean Age 27.5 y ± 0.38 SE), and 141 in Germany (Mean Age 29.4 y ± 0.44 SE) (Figure 1 A; Table S1). All participants were female and born in the country of sampling. We generated an average 3.7 million reads for each in vivo metagenome (Gabon Mean ± SE: 3,618,585 ± 60,421; Vietnam Mean ± SE: 3,811,863 ± 61,970; Germany Mean ± SE: 3,666,806 ± 78,023) (Table S1).

Microbiome profiling

Gabonese microbiomes had greater species richness and Shannon diversity than those in Vietnam and Germany (Figure 1 B). We detected more species in Vietnam than in Germany, but Shannon diversity was greater in Germany (Figure 1 B), indicating that though there were fewer species detected in the latter, their abundances were more even in each metagenome. Principal component analysis of Bray-Curtis dissimilarity produced clusters from each country, with Gabon and Vietnam more diffuse than Germany (Figure 1 C). The genera most differentially abundant in Gabon, as compared to Vietnam and Germany, were Treponema_D, Succinivibrio, Anaerovibrio, CAG-568, RUG 572 and UBA1436. The genus most abundant in Vietnam as compared to Gabon and Germany was Actinotalea; those most abundant in Germany were CAG-314, CAG-267, 51-20 and CAG-495 (Figure S1). Bacteroidaceae was the most abundant phylum in each country (Figure 1D). The taxonomic bar plot for the sole Vietnamese LP is given in plot S2.

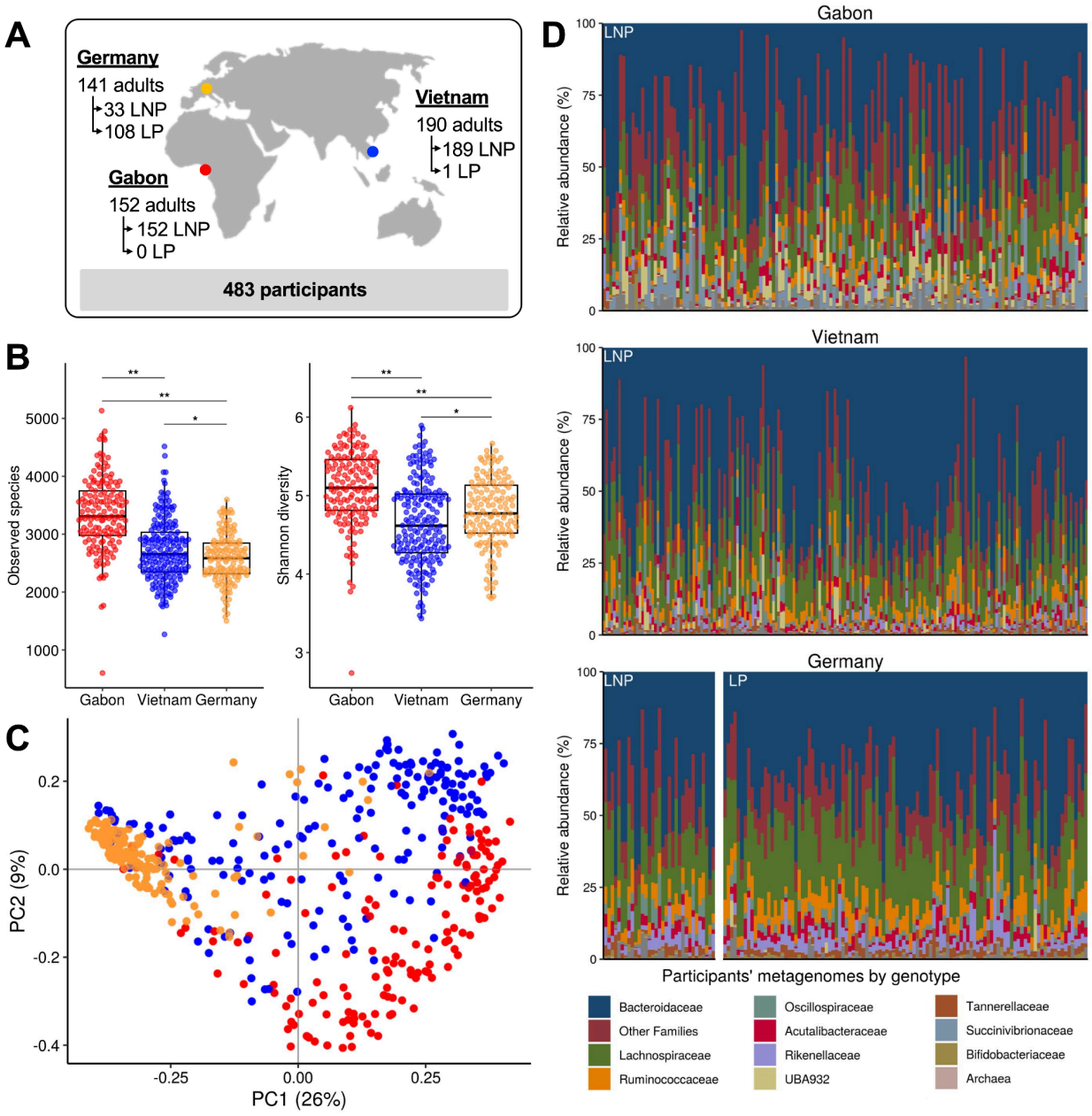

Figure 1. Study recruitment, genotyping, and gut microbiome survey. (A) Summary of recruitment and lactase persistence genotyping for participants in Gabon, Vietnam and Germany. (B) Alpha diversity of participants' stool metagenomes, rarefied to even sequencing depth. P-values are based on Wilcoxon rank sum tests, corrected for multiple comparison with the Benjamini-Hochberg procedure; * = $p < 0.05$; ** = $p < 5x10^{-10}$. (C) Principle coordinate analysis of Bray Curtis dissimilarities between stool metagenomes with colors corresponding to the sampling locations in panel A. The first two principal components are plotted with the percent variance explained by each. (D) Microbial taxa relative abundances from stool metagenomes,

with individual participants from each country plotted by LP/LNP genotype, in order of increasing age on the x-axis. Each color represents either a single bacterial family, all archaea, or the sum of all other bacterial families represented in each metagenome.

Lactase persistence-genotyping and lactose tolerance-phenotyping

As expected, the LNP genotype was very prevalent in Gabon (100 %) and Vietnam (99.5 %) and less prevalent in Germany (23.4 %; Figure 1 A; Table S1). Surprisingly, 21 % of LNP participants made little breath H_2 from lactose (< 30 ppm) such that they did not qualify for a lactose-intolerant diagnosis (20.4, 22.2 and 18.2 % in Gabon, Vietnam and Germany, respectively) (Figure 2 B; Table S1). The % of LNP participants that were lactose tolerant by both BHT and BGT ranged from 0 - 2 by country (Table S1). This indicates that LNP participants making little H_2 from lactose was not the outcome of small-intestinal lactase (i.e., false-positives for the LNP genotype), but rather due to the hypothesized microbially-conferred tolerance. As expected given this result, glucose rise better discriminated LP from LNP participants than did H_2 rise (Glucose rise < 20 mg/dl: $\chi^2_{glucose} = 284.44$; H_2 rise < 30 ppm: $\chi^2_{H2} = 209.01$). Only in Germany was the BGT less sensitive than BHT: 99.1 % of LP participants were tolerant by BHT but only 79.6 % were tolerant by BGT (Glucose rise < 20: $\chi^2_{glucose} = 47.01$; H_2 rise < 30: $\chi^2_{H2} = 98.91$). Participants that reported gastrointestinal symptoms during the 4 hours of testing had hydrogen rises greater in magnitude than those that did not (Wilcoxon signed-rank test, p-value = 0.045).

Modeling glucose production from lactose

To understand the association of covariates with glucose rise, we ran a linear model on the genotype, age, BMI, and covariate-interactions (n = 140) of German participants, since that was the only country wherein both LP and LNP were prevalent. As expected, genotype was significantly associated with glucose rise. From all other covariates and their interactions, only age showed a trend for negative association glucose rise (p-value = 0.06, adjusted p-value = 0.15) (Figure S2). We repeated the same model as above, except for only LP participants. From the 108 samples with complete metadata, one was from Vietnam so it was removed (model fitted on 107 Germans). No covariate was significantly associated with glucose rise.

Since substantial LNPs produced glucose from lactose (Figure 2 B), we used the same model above on LNPs-only to associate covariates with their glucose rises (n = 337 samples), first comparing {Gabon, Germany} vs Vietnam; then Gabon vs Germany. Gabonese and Germans had significantly higher glucose rises than Vietnamese (adjusted p-values = 0.0195) and Gabonese than Germans (adjusted p-value = 0.0061). Though age was not significantly associated with glucose rises (negative association, adjusted p-value = 0.0598), there was a significant interaction effect between age and country Germany (adjusted p-value = 0.0108), such that German LNP glucose rises decreased with age.

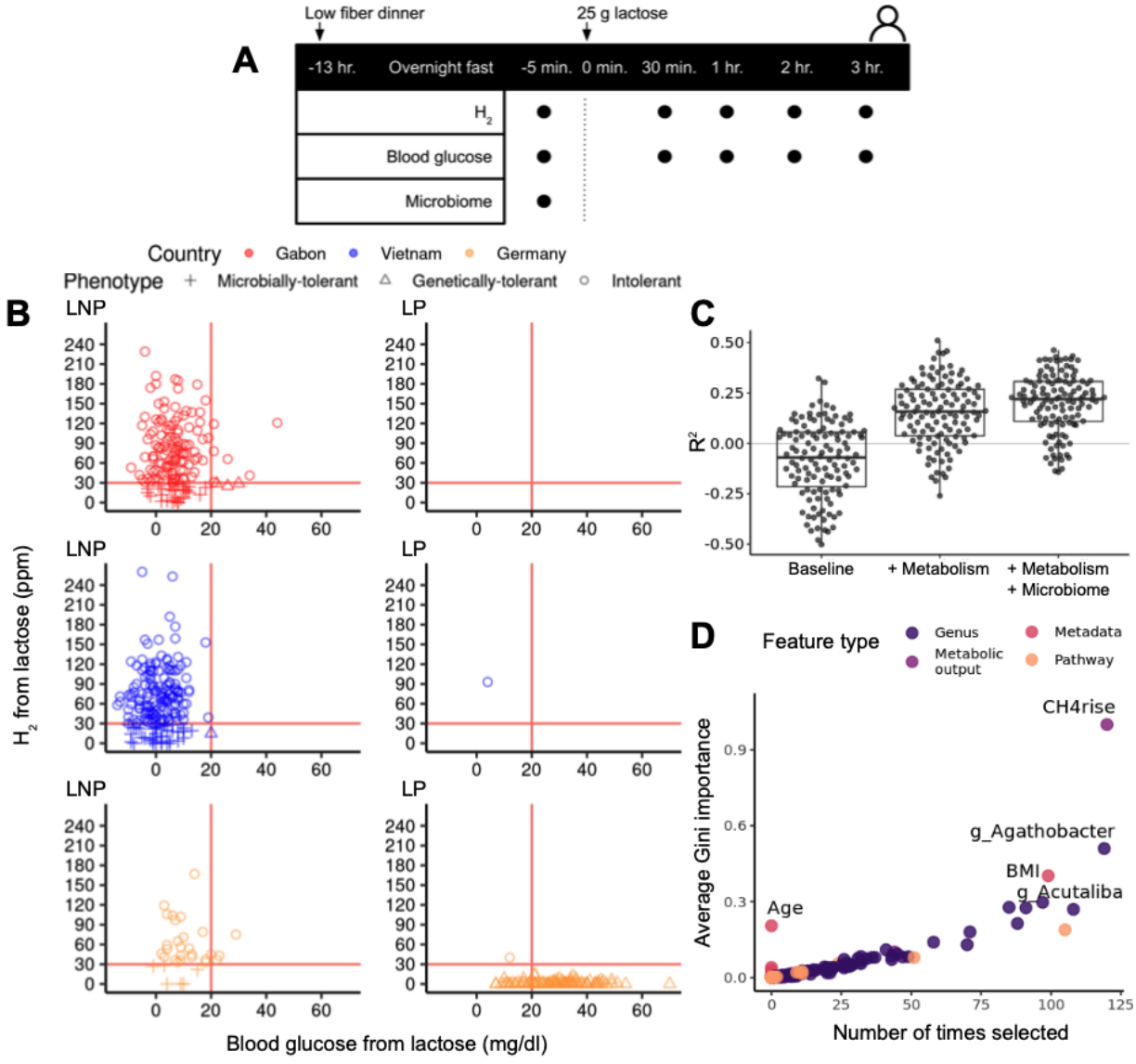

Figure 2. Lactose tolerance phenotyping. (A) Participants ingested 25 g lactose monohydrate after an overnight fast to assess lactose malabsorption with hydrogen breath testing and blood

glucose monitoring for 3 h. Microbiome samples were collected before ingestion of the lactose. (B) The breath hydrogen and blood glucose made by each participant from lactose, separated by their lactase persistence genotype. Participants with a hydrogen rise above 30 ppm (denoted by a red line) were lactose intolerant; those below 30 ppm were lactose tolerant. The clinical cutoff for lactose tolerance as determined by glucose is given by the line at 20 mg / dl; those above the line are genetically-tolerant, and those below it are genetically-intolerant. (C) The variance in hydrogen rise explained by 1) participant metadata alone, 2) participant metadata with the lactose tolerance tests' metabolic results, and 3) both with the addition of the microbiome, which comprised both genus and pathway abundances. (D) Feature importance of generalized linear models in explaining variance in breath hydrogen, colored by type of data.

Modeling in vivo H_2-production from lactose

LP and LNP, without microbiome

We fit a generalized linear model to correlate H_2 rise with glucose rise, age, CH_4 rise, country, and BMI, and tested for interaction effects of glucose rise with BMI and age in samples with the necessary metadata (n = 445). Due to the data distribution, we used a two-part generalized model: 1) probability of H_2 rise > 1 ppm across all samples, and 2) for samples with H_2 rise > 1 ppm (n = 353 samples): regression of H_2 rise with a Gamma distribution. The model explained 35 % of the observed variance in H_2 rise across all data (R^2 = 0.35; p-values and coefficients given in Table S2).

As expected, glucose is negatively associated with the probability of H_2 rise > 1 ppm. H_2 rise significantly increased with lower glucose rise (adjusted p-values = $1.5x10^{-2}$ and $1.6x10^{-5}$ for each part of the model). There was a trend for higher CH_4 production with increasing H_2 production, consistent with CH_4 being produced from H_2, although the correlation was not significant after p-value correction (adjusted p-values = $4.6x10^{-2}$ and $9.4x10^{-2}$ for each part of the model). Age was positively associated with H_2 rise (adjusted p-values = $1.5x10^{-2}$ and $1.4x10^{-2}$ for each part of the model), such that older individuals were more likely to produce more H_2 after lactose ingestion. Age was positively associated with the probability of H_2 rise > 1 ppm in people with glucose rises <= 15 mg/dL (15/108 LP and 313/337 LNP); the closer to 15 mg/dL, the smaller the effect of age. For those with glucose rises > 15 mg/dL, age was negatively associated

with the probability of H2 rise > 1 ppm (93/108 = 86 % of the AA/AG individuals and 24/337 = 7 % of the LNP individuals).

Country was the covariate with the highest coefficient (Table S2), indicating a strong influence of the country - and its confounding factors - on lactose fermentation in the gut. Germans produced less H_2 from lactose than do Gabonese and Vietnamese (adjusted p-values = $1.29x10^{-7}$ and $1.13x10^{-3}$ for each part of the model); when an H_2 rise occurred (i.e., H_2 rise > 1 ppm; 2nd part of the model), it was lower in Gabonese than in Vietnamese participants (adjusted p-value = $1.7x10^{-2}$ for the 2nd part of the model). Finally, BMI was negatively associated with H_2 rise (adjusted p-value = $2.8x10^{-3}$).

LNP only, without microbiome

We fit a generalized linear model to associated metadata with LNP participants' H_2 rises (Gamma distribution, link = log, n = 337 samples, pseudocount of 1 since Gamma function does not accept zeros, features: age, BMI, beta-galactosidase, 'Country: Germany vs others', 'Country: Gabon vs Vietnam'). As was the case when modeling H_2 rise with both LNP and LP participants, age was positively associated with H_2 production (Table S3), and BMI and 'Country: Germany vs others' were both negatively associated with H_2 rise (Table S3).

LNP only, with microbiome

To determine the contribution of the microbiome to in vivo gas production, we subset LNP participants with a glucose rise < 20 mg/dL (n = 323 samples). Our rationale was to measure the percentage of explained variance in H_2 rise due to microbial features by comparing models (regression random forest models with 0.8 train-test split and 120 CV sets) with and without these features. If the explained variance increases by including microbial features, the increase would correspond to the conditional dependence of microbiome data and H_2 rise given all features in the baseline. Our baseline, or null model, consisted of H_2 rise predicted using host information: age, BMI, and country. We tested sets of features alone or combined: (i) metabolic output: CH_4 rise and extracellular β-galactosidase activity, (ii) taxonomic profiles: relative abundances of genera, and (iii) metabolic profiles: relative abundances of MetaCyc pathways. The R^2 of the baseline was -0.085±0.18, which is a poor accuracy similar to random guessing.

First, we verified that metabolic outputs associated with lactose degradation (extracellular β-galactosidase activity) and H_2 production (CH_4 rise) were good predictors of H_2 rise following lactose ingestion. As expected, the model greatly improved by including these variables (R^2 ~ 0.15 ± 0.16, average variance explained by metabolic outputs = 0.24 ± 0.075; Wilcoxon signed-rank test, adjusted p-value = 5.0×10^{-21}). Next, we assessed the variance explained by the gut microbiome. Individually, relative abundances of genera and pathways significantly increased the accuracy of the baseline model (Wilcoxon signed-rank test, respective adjusted p-values = 2.1×10^{-19} and 4.1×10^{-16}). In particular, taxonomic profiles explained 0.13 ± 0.095 of variance in H_2 rise (R^2 ~ 0.037 ± 0.15) and metabolic pathways explained 0.092 ± 0.090 (R^2 ~ -0.029 ± 0.17). Combining taxonomic profiles to metabolic outputs significantly increased the accuracy of models compared to metabolic outputs alone (average increase = 0.040 ± 0.085; Wilcoxon signed-rank test, respective p-value = 2.8×10^{-6}). However, combining pathways and metabolic outputs did not increase the accuracy compared to (average decrease = -0.017 ± 0.075; the decrease occurred because of the noise added by the metabolic pathways).

We hypothesized that relative abundances of the *Bifidobacterium* genus and species would be responsible for the explained variance by taxonomic profiles. Although they significantly increased model accuracy compared to baseline (Wilcoxon signed-rank test, p-value = 1.38×10^{-4}), relative abundances of *Bifidobacterium* genus and species did not account for the majority of the explained variance due to taxonomic profiles (average increase = 0.030 ± 0.085 whereas taxonomic profiles explained 0.13 ± 0.095 of the H_2 rise variance) and did not improve the model with metabolic outputs when combined to it (average decrease = -0.033 ± 0.066).

To investigate which variables were responsible for the explained variance in H_2 rise, we looked at which metabolic outputs, genera, and pathways were most consistently selected across CV sets in models with the three types of features (R^2 ~ 0.20 ± 0.15) and at their average importance across CV sets (Figure S3 A). On average across CV sets, 23 ± 8.5 features were selected to make predictions. CH_4 rise and the genus *Agathobacter* (family *Lachnospiraceae*) were almost always selected and had the highest average feature importances (respectively selected in 120 and 119 of 120 CV sets; Figure S3 A). BMI, the genus *Acutalibacter* (family *Ruminococcaceae*), and the superpathway of L-serine and glycine biosynthesis I were next in terms of number of times selected and average relative abundance (each selected in 99, 108, and 105 CV sets, respectively; Figure S3 A). Samples with H_2 rise close to the mean were the best

predicted on average across CV sets, such that predicted values varied around 62 ± 15, min-max = 26-97, although the observed values had an average of 67 ± 45, min-max = 0-260 (Figure S3 B-C). This means that despite features being associated with H_2 rise, low H_2 rises were not well predicted. Contrary to the in vivo model output that excluded the microbiome, the country of origin of individuals was not important (Figure S3 A).

CH_4 rise was a better predictor in the model than methanogens' relative abundances. While removing CH_4 rise from the predictors led to a drop in accuracy (model with metadata, betagal, and relative abundances of genera: R^2 = 0.051 ± 0.15), methanogens' relative abundances were not even used for predictions (e.g., *Methanobrevibacter* A, the most abundant and prevalent human gut methanogen, was selected in only 32 of the 120 CV sets). This suggests that relative abundances are not an accurate proxy of metabolic activity (Figure S4).

In summary, we could explain up to 0.20 ± 0.15 of variance in *in vivo* H_2 production from lactose using metabolic outputs, taxonomic and metabolic profiles, and host covariates. CH_4 rise and the genus *Agathobacter* were the most important predictors. Nonetheless, low H_2 rise samples were not well predicted. Hence, to investigate low H_2 production from lactose digestion by gut microbiota, we set up controlled experiments to only look at the effect of microorganisms.

Modeling in vitro H_2-production from lactose

Experimental design

To understand the microbial differences between LNPs that produce high- and low-H_2 from lactose, we randomly selected stool samples from 75 of them that met, and 75 of them that did not meet HBT criteria for lactose tolerance (H_2 rise < 30 ppm and H_2 baseline < 30 ppm). We incubated these 150 samples with lactose for 4.5 h, measuring their metabolites and sequencing their mirobiomes at the beginning and end of the experiment.

Stools responded to lactose in four different ways: (i) no H_2 production due to no metabolic activity (i.e., "Inactives"), (ii) low H_2 production due to low carbon consumption ("Weak"), and of the high carbon consumers, those with (iii) high- or (iv) low-H_2 production ("Gassy" and "Tolerant", respectively, divided using k-means clustering) (Figure 3 B; Table S4; Figure S5). The relative abundances of several taxa increased during cultivation with lactose, the greatest magnitude of which was the genus *Bifidobacterium* in Tolerant, followed by *Collinsella*

in Gassy, and *Escherichia* in the Weak and Inactive groups (Figure 3 C). Countries were significantly associated with groups (χ^2-test, $\chi^2 = 14$, df = 6, p-value = 0.029), such that: (i) the majority of German samples belonged to the Tolerant and Gassy groups (8 and 6 of the 16 German samples, respectively), (ii) most Vietnamese samples had a low activity as they belonged to the Inactive and Weak groups (23 and 22 of the 69 Vietnamese samples, respectively), and (iii) Gaboneses samples were substantially represented in all groups (Table S4). The Tolerant and Gassy groups had more extracellular beta galactosidase than the Inactive and Weak groups (Figure S6). Host dietary lactose scores did not differ by in vitro classification (Figure S7).

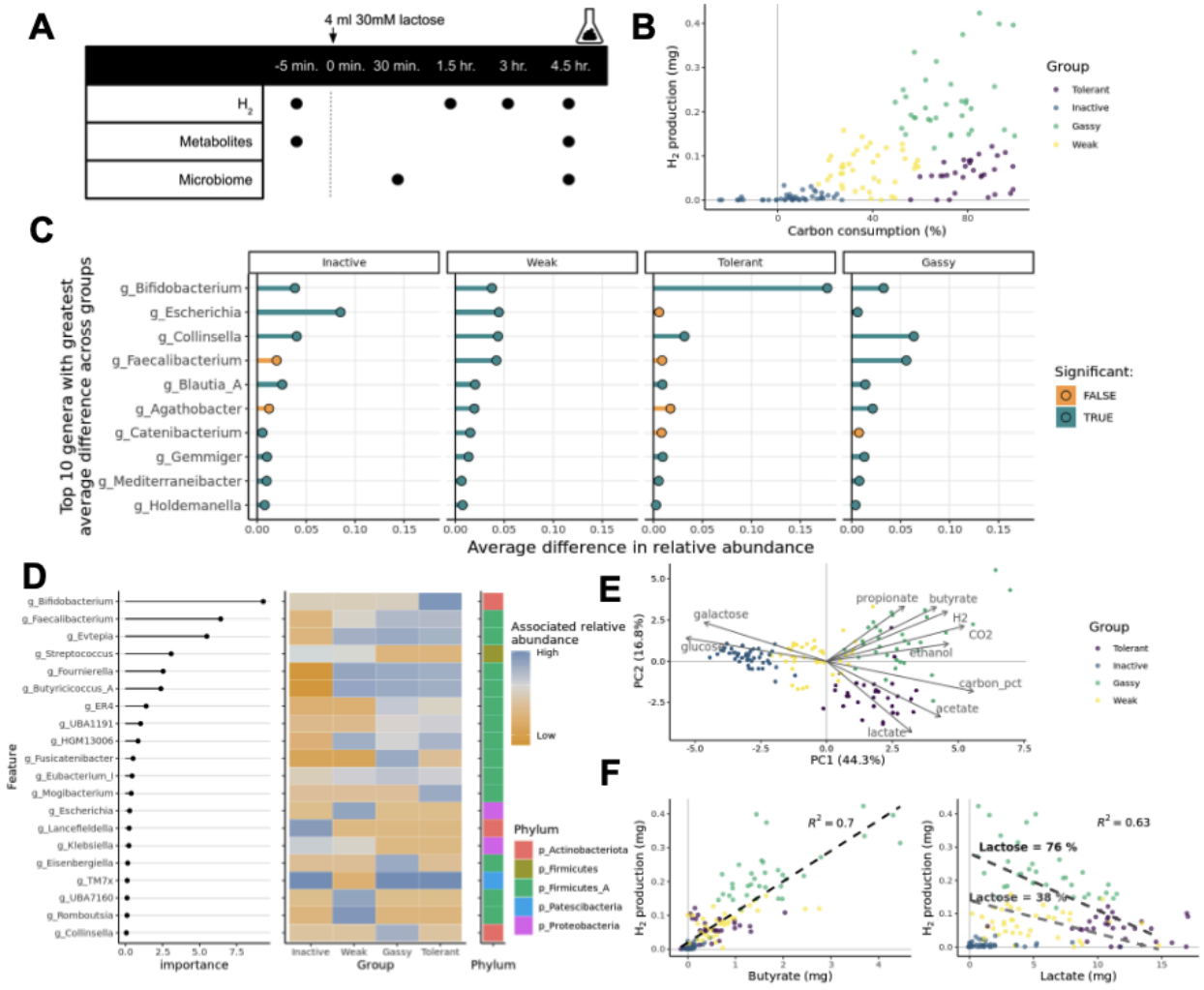

Figure 3. LNP stool microbiotas' response to lactose. (A) Stool samples were cultivated with lactose for 4.5 hours, during which microbial taxa and metabolites, including hydrogen, were monitored. (B) Sample groupings based on the percent of lactose's carbon consumed, and the

hydrogen produced from it. (C) Changes in relative abundance during cultivation of the ten most differentially-abundant taxa, by the four lactose metabolic groupings from (B). P-values < 0.05 after Benjemini-Hochberg correction are marked in blue. (D) Features of greatest importance in a random forest model which classified metagenome membership to each of the four metabolic groups from (B), as interpreted by endoR. Associations of genera with each group being that of high or low relative abundance are given in blue and yellow, respectively. (E) Principal component (PC) analysis on all metabolites measured at the end of the experiment. Metabolites most correlated with each PC are shown in grey; the closer the arrows, the more correlated the variables are on the PC. (F, left) Correlation of butyrate with hydrogen production. (F, right) Correlation of lactate and hydrogen production at high and low % lactose consumption.

Metabolites produced from lactose

Principal component analysis (PCA) associated metabolites with the aforementioned metabolic-response groups (Figure 3 E). PC1 separates high- from low-carbon consumers (Figure 3 E). PC2 separates samples with high lactate and acetate production from those with high hydrogen, butyrate, and propionate production. Glucose and galactose were associated with the Inactive group, indicating that those monosaccharides had been cleaved from lactose, but not fermented. Of the high-carbon consumers, those producing high hydrogen were associated with the metabolites propionate and butyrate, and those producing low hydrogen with lactate and acetate. H_2 and butyrate were the most positively correlated variables on the 2 PCs (Figure 3 E and F; Spearman's test, S = 75582, df = 148, p-value = 2.4×10^{-40}, $R^2 = 0.70$). H_2 and lactate production were negatively correlated and dependent on carbon consumption (ANOVA on ranks, p-values for carbon consumption < 2×10^{-16}, lactate = 4.8×10^{-2}, interaction effect of carbon consumption and lactate = 6.3×10^{-11}; the coefficient of determination of the model, $R^2 = 0.63$; Figure 3 F). The black line in Figure 3 F corresponds to a linear regression of H_2 versus carbon consumption, lactate, and their interaction effect at a fixed carbon consumption of 76 % (the Gassy- and Tolerant-groups' mean carbon consumption); the grey line is similar but at a fixed carbon consumption of 38 % (the Weak group's mean carbon consumption).

On average, 44 ± 33 % of the lactose was consumed (Figure S8 A). Glucose and galactose were produced in equal amounts, but glucose was preferably consumed over galactose (Figure S8 B). Remarkably, 12 samples had negative theoretical carbon consumption, meaning

that glucose and galactose concentrations were higher at the end of the experiment than what would be expected from the lactose cleavage. These samples belonged to 5 Vietnamese and 7 Gabonese individuals and did not show any trend for similar microbiome profiles at the beginning of the experiment (Figure S8 C-D).

Predicting metabolic group with random forest

We fit random forest classifiers to predict metabolic groups using relative abundances of microbial taxa and pathways at the end of the experiment as features. We limited our analysis to genera and pathways that significantly increased in at least one of the groups (one-sided Wilcoxon signed-rank test and adjusted p-value < 0.05; Table S5). The overall accuracy across CV sets was 0.51 ± 0.082 (a random classifier of four groups having an expected accuracy of 0.25), with the balanced accuracy of each group being: Inactives = 0.74 ± 0.092; Weak = 0.53 ± 0.11; Tolerant = 0.80 ± 0.095; Gassy = 0.67 ± 0.097. Given the good accuracy, the model was fit on all samples for interpretation: 114 features were selected, of which 29 were metabolic pathways. Models using taxa and pathway abundances from the start of the experiment performed less well, which was expected as the effects of lactose on the metagenomes were not as exacerbated (Table S5). The classifier was interpreted using endoR (Ruaud et al. 2022).

Linking the in-vitro and vivo datasets

To determine whether samples that produced low hydrogen in vitro correspondingly made less hydrogen in vivo, we grouped the ‘Inactives’, ‘Weak’, and ‘Tolerant’ groups into one ‘low’ group. Their mean in vivo hydrogen production was indeed lower than that of samples that produced high hydrogen in vitro (i.e. ‘Gassy’), as tested with a Wilcoxon test (Figure S9). We next tested whether levels of *Bifidobacterium* differed between the four in vitro clusters using an ANOVA on log10-transformed relative abundances, and found that indeed, in vivo *Bifidobacterium* was enriched in the Tolerant group only (Figure 5C). There were no statistically-significant differences in in vivo symptom intensity by in vitro-classified group, though we do note that the only participants to report the most severe symptoms did belong to the Gassy group (Figure S10).

Phenotyping Bifidobacterium genomes' carbohydrate metabolism

Species-identity and CheckM2-estimated completeness explained variance in the 139 phenotyped *Bifidobacterium* genomes' total number of metabolizable carbohydrates; country and amount of breath hydrogen from lactose did not. Species-identity and completeness similarly explained variance in whether or not each individual genome could metabolize each carbohydrate, but country and breath hydrogen did not. Despite these differences by species, all but two could metabolize lactose and glucose. The prevalence of genomes with the potential to metabolize different carbohydrates varied by sampling location (Figure 4; carbohydrate metadata are provided in Table S6). This is explainable by the aforementioned fact that carbohydrate metabolism differs by species, and the relative abundances of those species differed by country.

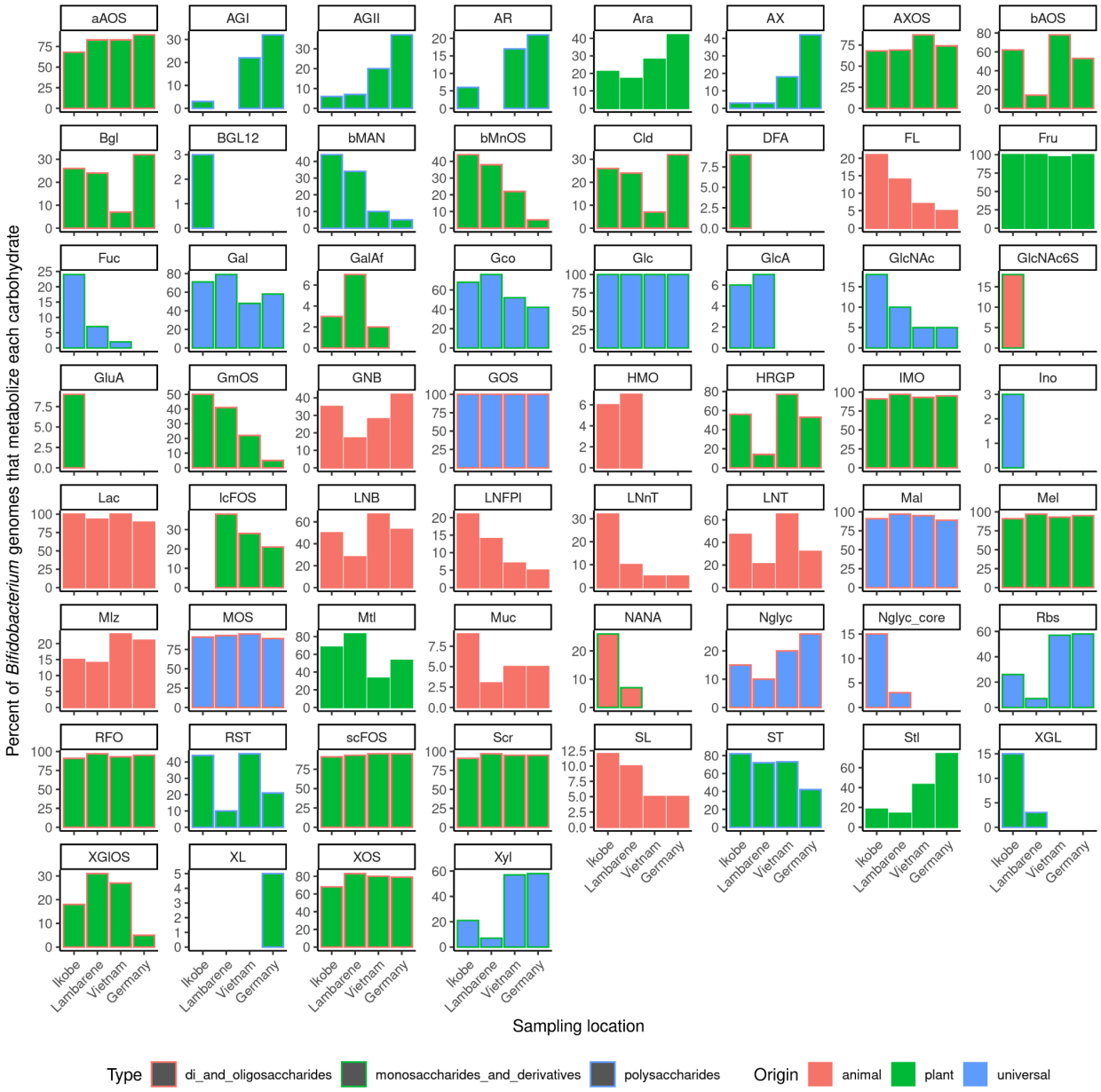

Figure 4. Biogeography of *Bifidobacterium* carbohydrate repertoires. Percentages of *Bifidobacterium* genomes by location with the ability to metabolize each phenotyped-carbohydrate. Whether or not the carbohydrate is of animal, plant, or universal origin is given by the bar colors red, green and blue, respectively. Monosaccharides and their derivatives, di- and oligo-saccharides, and polysaccharides are distinguished by green, red, and blue bar borders, respectively. Descriptions of each carbohydrate are provided in Table S6.

Bifidobacterium genome phenotyping

Bifidobacterium strain isolation yielded 59 non-redundant genomes. We assembled 7456 MAGs, of which 5,950 were over 50% complete, had less than 5% contamination and were non-redundant (i.e., were dereplicated at 99.99% ANI). 150 of the 5,950 genomes were *Bifidobacterium* MAGs. We inferred the phylogeny of all 209 *Bifidobacterium* genomes (Figure 5 A). We inferred their phenotypes: all had beta-galactosidase, and all but 1 (*Bifidobacterium* angulatum; Germany) ferment lactose (Figure 5 A).

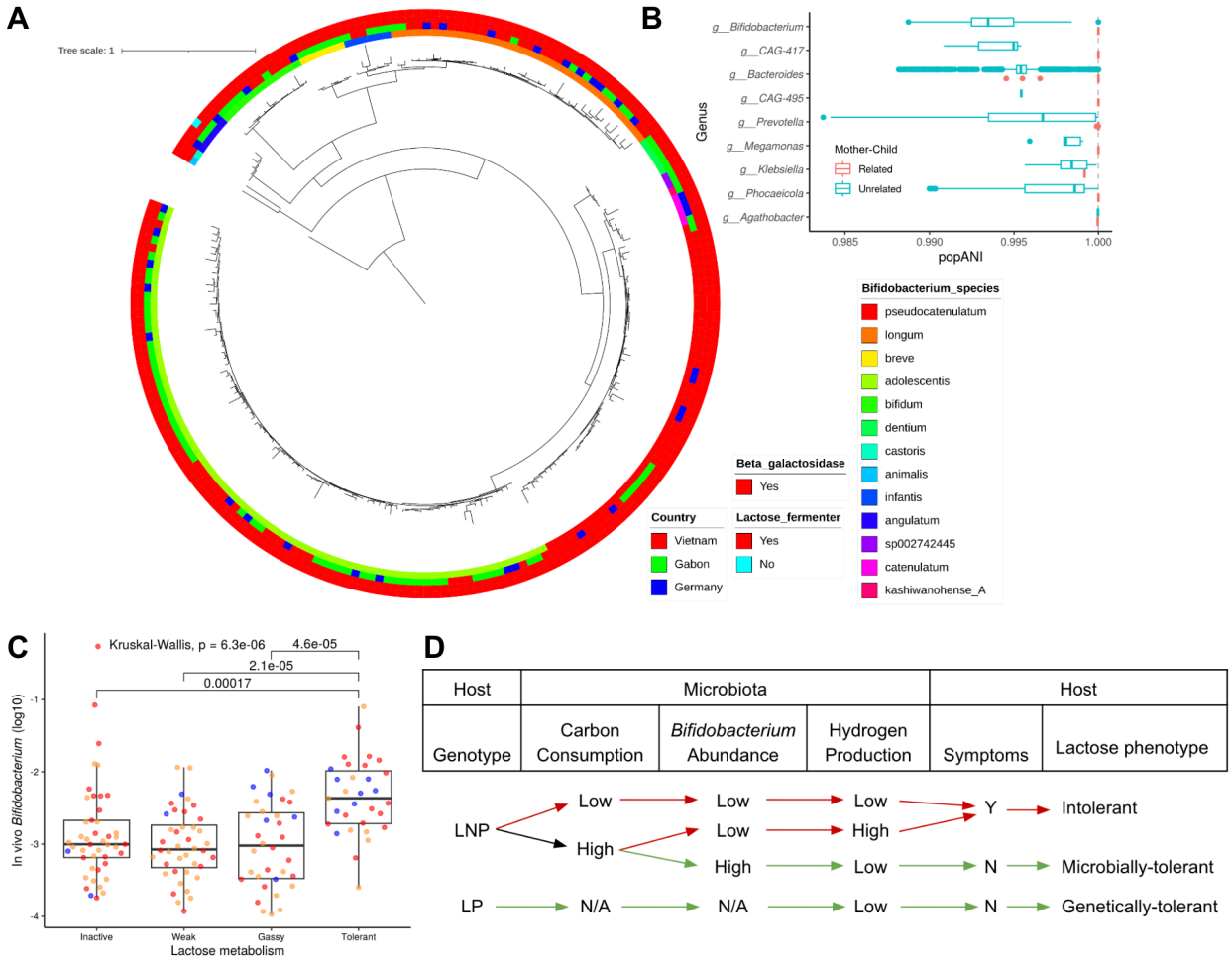

Figure 5. *Bifidobacterium* genomes, strain-sharing, and summary. (A) Phylogeny of 209 *Bifidobacterium* genomes, annotated by their country of origin, species membership, and whether or not they have beta-galactosidase and the ability to ferment glucose. (B) The relatedness of strains for which comparisons between both related- and unrelated- mother-child pairs could be performed. A popANI greater to or equal than 99.999 % indicates a strain-sharing

event, and is denoted by the grey line. Genera are ordered by decreasing difference in the median popANIs of their strains compared between related-, and unrelated- mother-child pairs. (C) In vivo *Bifidobacterium* abundance of the different participants, separated by their lactose metabolism group membership in vitro. P values calculated with a Kruskal-Wallis test and Dunn's post hoc test, of which only those < 0.05 are shown. (D) Summary of how a Bifidobacterium-mediated interaction between host genotype and microbiota determines whether or not LNPs are LT. Dependence on the microbiome in determining LNP host phenotype is shown with colored arrows: green and red depict tolerance and intolerance, respectively.

Strain sharing

We generated 3,212 sub-species representative genomes (SSRGs) by dereplicating the 5,590 non-redundant MAGs together with the 59 isolate genomes. We collected stool from 351 of the total 483 total participants' infants, yielding a total of 351 mother-infant pairs. One of the mothers had 2 infants, yielding a final mother-infant dataset of 351 lactose-phenotyped mothers and 353 infants. We used these 704 metagenomes and 3,212 SSRGs to perform 18,755 strain comparisons, 11,615 of which compared at least 50 % of the SSRG. 1 of these was a comparison of *Bifidobacterium* strains in related mother-infant pairs, and it was a strain-sharing event (Figure 5 B). 82 other strain-sharing events were identified, though they were not of strains belonging to *Bifidobacterium* (Figure 5 B).

Discussion

We measured the gas and glucose that 483 study participants in Gabon, Vietnam and Germany made from lactose, genotyped LP/LNP, and surveyed their gut microbes to investigate the contribution of the microbiome to lactose tolerance. As expected, LP individuals - the majority of which were in Germany, where there is a history of dairying - made high glucose and low hydrogen from lactose; they were lactose tolerant (LT). Genotype similarly predicted phenotype for the majority of LNP participants, who made low glucose and high hydrogen from lactose, qualifying them for a lactose intolerant diagnosis. However, one fifth of LNPs were LT. This was the case in each of the three countries, despite their differing LNP prevalences (25 - 100 %), and locations on different continents. This confirms previous reports of a LP-LT genotype-phenotype disconnect (Itan et al. 2010; Hollfelder et al. 2021), whose authors could only speculate resulted from false negatives for LP (missed LP-conferring alleles) and/or a contribution of gut microbiota (Ranciaro et al. 2014; Jeong et al. 2018). We are the first to test both possibilities, as we novelly-combined both lactose phenotyping and genotyping with a survey of the microbiome, which we hypothesized accounts for this discrepancy.

The 20% of LNP individuals with LT did not have substantial glucose rises, indicating that their low-hydrogen production from lactose was not the result of missed LP-conferring alleles in their genomes. Their microbiomes were not however different from the 80% of LNP individuals that were intolerant, so we cultivated LNP stools with lactose in vitro. We observed 4 different responses of LNP gut microbiota to lactose: one of little hydrolysis with no fermentation (“Inactive”), little hydrolysis with weak fermentation (“Weak”), hydrolysis with fermentation yielding high-hydrogen (“Gassy”) and hydrolysis with fermentation yielding low-hydrogen (“Tolerant”). This identifies two gut-microbial mechanisms by which LNP individuals can be LT - one of metabolic activity and one of metabolic inactivity. Having ruled out the possibility of false negatives for LP, we conclude that the LP-LT disconnect is in fact a contribution of gut microbiota, which we term “microbially-acquired lactose tolerance” (MALT).

Bifidobacterium drives metabolically-active MALT because it was the singular microbial genus enriched in stools classified as Tolerant in vitro, compared to those of the Weak, Inactive and Gassy groups. Being able to extrapolate what we observed in vitro to the in vivo context is dependent on two criteria, both of which we met: 1) Stools that made less hydrogen from lactose in vitro were from hosts that made less hydrogen from lactose in vivo (Inactive, Weak and

Tolerant, compared to Gassy), and 2) Individuals with Tolerant-classified stools had the same metagenomic-enrichment for *Bifidobacterium* in vivo that they did in vitro. This explains why we did not observe stool metagenome differences in vivo-classified tolerant and intolerant LNPs: the latter were comprised of vitro-classified Gassy stools only (i.e., those without *Bifidobacterium* enrichment), but the former comprised the Inactive-, Weak- and Tolerant-classified stools grouped together, of which only the Tolerant-classified subset were *Bifidobacterium*-enriched. Host-microbiome GWAS of the past overcame microbially-inactive MALT's dilution of the *Bifidobacterium* signal in tolerant LNPs by having comparatively larger sample sizes. Those GWAS were an important step towards generating the hypothesis that microbes may contribute to lactose tolerance, but by nature of being non-interventional, insufficient to test it (Goodrich, Davenport, Waters, et al. 2016).

In addition to *Bifidobacterium*'s DNA being overly-abundant in Tolerant-classified stools, so too were the metabolic end-products of its canonical Bifid shunt pathway, lactate and acetate. The Bifid shunt ferments only hexoses, meaning dietary lactose that reaches the colon needs to have first been hydrolyzed. We confirmed that both the Tolerant and Gassy groups - the two that were metabolically-active - had more beta-galactosidase than their Inactive and Weak counterparts. In the Tolerant group, *Bifidobacterium* fermented the resulting glucose via the Bifid shunt into lactate and acetate while making relatively little hydrogen. Glucose fermentation in the Gassy group - enriched in the genera *Collinsella*, *Faecalibacterium* and *Agathobacter* - produced high hydrogen and butyrate. As the only group of the four to make high hydrogen, Gassy represents the lactose metabolism of the 80%-majority of LNP individuals that made high hydrogen in vivo, as confirmed by the fact that Gassy-classified stools' hosts made more hydrogen during in vivo phenotyping than did the hosts of the Inactives, Weak and Tolerant groups.

The 20% of LNP individuals with MALT comprise not only stool metagenomes classified as Tolerant in vitro, but also those of the Inactive and Weak groups, as all three share the same decisive outcome in host phenotyping: low hydrogen production from lactose. Abundance of E. coli was enriched in the Inactive and Weak groups, whose lac operon, which is repressed in the presence of any glucose, activates in lactose's presence and contains both a permease to facilitate lactose's transport into the cell and beta-galactosidase to hydrolyze it (Maloy and Hughes 2013). This hydrolysis was represented in the metabolic profiles of the

Inactive samples, wherein glucose and galactose were the defining metabolites, and in Weak samples, wherein that same hydrolysis was followed only by weak fermentation. Unmetabolized lactose has been hypothesized to increase osmotic load in the colon, thereby causing diarrhea (Xue et al. 2020), but there were no differences in in vivo-reported symptoms between the in vitro-classified groups (though we do note that the only severe symptoms were reported by Gassy-classified participants). A weakness of this study is that we did not follow-up with participants about their symptoms after the 3 hours of in vivo phenotyping. As a result, we cannot conclude whether the Inactive and Weak groups, both of which underlie MALT as much as the Tolerant group, are true- or false-positives for lactose tolerance. That outstanding question should be the focus of further study.

MALT affords complete lactose tolerance in that it reduces gas production to levels matching those of LP individuals, but it yields less energy for the host, since absorbing gut microbe-produced SCFAs into the blood yields less energy than converting lactose to glucose in the small intestine (Moffett et al. 2020). Producing less gas does however mean that *Bifidobacterium* made dairy a more tolerable energy-source for ancestral humans. We conclude that this, together with fermentation, helps account for the 5,000-year gap between the evolution of dairying and the LP genotype. We presume that conferral of tolerability was dependent on the abundance of *Bifidobacterium* available to metabolize lactose relative to that of other members of the gut microbiota, because we showed here that other microbes produce gas from it. We propose that in periods of extreme nutritional scarcity, before LP evolved, there was positive selection on humans with greater abundances of *Bifidobacterium*, as only they were able to 1) extract energy from lactose in the form of SCFAs, and 2) sufficiently tolerate dairy to be able to consume it at all, thereby enabling uptake of its other nutritional components, including its proteins, fats, vitamins, and minerals. This may explain why there is a signature of the genus having codiversified with humans (Suzuki et al. 2022). The counter-argument that LNP human ancestors without MALT-conferring *Bifidobacterium* could still reap dairy's other nutritional rewards does not hold, as the diarrhea ensuing from lactose intolerance is an experience some have described as being similar to dysentery (Guerrant et al. 1992). This reduced gut-transit time reduces energy-extraction from the rest of the LNP diet, thereby further increasing the fitness disparity between LNPs with and without MALT-conferring *Bifidobacterium*.

Taken together, we conclude that *Bifidobacterium* facilitated the evolution of LP. This is because without dairying, which *Bifidobacterium* enabled, there would not have been positive selection for the non-synonymous mutations in MCM6 that first conferred LP 10,000 years ago. Instead, the MCM6 mutations would have emerged in several individuals, with genetic drift as its only chance of survival in the population, which is much less likely. This also means that *Bifidobacterium* buffered position selection on LP after it arose, because it afforded an intermediary level of fitness between those with LP and those without it. This is how we explain why LP has not gone to fixation in Germany such that 100% of our participants had it; instead, only 77% did. The prevalence of MALT in German LNPs was high, at 20%, which suggests that the signature of ancestral buffering still exists, and that it may also exist in other countries.

This does not, however, explain why MALT's prevalence in Gabon and Vietnam was also 20%, because they do not have a thousands of years-old history of dairying, meaning there can exist no legacy of ancestral buffering of *Bifidobacterium* on LP. Study participants in Gabon and Vietnam did however report consuming some dairy in their diets. A difference between them, and Germans dairying prior to LP 15,000 years ago, is that only the latter were faced with nutrient scarcity, and the risk of starvation. This means that the evolutionary context in which MALT has emerged in Gabon and Vietnam is different. Positive selection is not acting on both the host and *Bifidobacterium* together, as it was for ancestral Germans threatened with food scarcity, but rather on the microbiota alone: microbes that extract maximal ATP from 1 mol of lactose (i.e., *Bifidobacterium*) have more fitness in dairying LNP guts than microbial competitors with less extraction efficiency (e.g., *E. coli*).

Our carbohydrate metabolic potential analysis revealed a second explanation for why MALT is prevalent in all three countries: *Bifidobacterium* metabolism is not restricted to lactose, but can rather metabolize a diverse array of dietary carbohydrates. This means that infants, whose levels of *Bifidobacterium* are high during breastfeeding, have their populations of that genus sustained after weaning because those microbes metabolize carbohydrates in their solid food diets, even though those differ by country. This explains a facet of how *Bifidobacterium* strains persist in guts throughout hosts' entire lifespans, an important prerequisite for the recent discovery that members of this genus are transmitted between host generations (Suzuki et al. 2022). Our strain-sharing analysis here could only perform 1 comparison of *Bifidobacterium* strains within a single related-mother infant pair, which is not robust enough to study its

transmission dynamics. We conclude that this failure to detect shared-strains is because the *Bifidobacterium* species abundant in children are generally less abundant in adults, and vice versa, leading to a failure to be able to call SNPs on our assembled genomes with the metagenomes of both mothers and their infants.

Mongolia offers an intriguing fourth example of this host genotype-microbiome interplay. Only 5% of the population is LP, but they have been dairying for the past 5,000 years, and as much as 30% of their calories are from dairy (Curry 2018; Wilkin et al. 2020). Some explain this with the fact that fermentation is prevalent there: they recruit microbes to pre-metabolize lactose outside the body, enabling consumption of that lactose-reduced dairy without issue (Curry 2018). The same explanation has been given for the discrepancies between LP and lactose tolerance observed in other populations. Indeed, adding lactose-metabolizing microbes - including *Bifidobacterium* - to dairy before human consumption has been shown to reduce breath hydrogen from lactose (T. Jiang, Mustapha, and Savaiano 1996; Aguilera et al. 2021; Masoumi et al. 2021), as has giving LNPs *Bifidobacterium* probiotics, reviewed recently by Mysore Saisrasad and colleagues (Mysore Saiprasad, Moreno, and Savaiano 2023). Others have however proposed that microbes intrinsic to the human gut could also be increasing dairy's tolerability (Segurel et al. 2020; Suzuki and Ley 2020), and we are the first to test that hypothesis. We implicated *Bifidobacterium* in this microbially-acquired lactose tolerance phenotype, and indeed, the relative abundance of *Bifidobacterium* in Mongolian adults is as high as 80% (Liu et al. 2017).

A recent non-interventional study attempted to test our hypothesis and arrived at the opposite conclusion, correlating *Bifidobacterium* with lactose intolerance in LNPs (Brandao Gois et al. 2022). They genotyped for LNP, but rather than administer lactose and measure gas as we did, participants self-reported both gastrointestinal symptoms and dairy consumption for 7 days. Their association of *Bifidobacterium* with increased lactose intolerance symptoms is explained by a confounder, which they even measured themselves: dairy consumption. LNPs that consume more dairy are already known to have more *Bifidobacterium* (Bonder et al. 2016). Extrapolating from our results, LNPs consuming large amounts of dairy will provide gut *Bifidobacterium* with more lactose than they alone can metabolize, making that substrate available to the other members of the microbiota - those whose metabolism of lactose is associated with lactose

intolerance symptoms. Rather than contradict our conclusion, their outcome provides more evidence to support it.

We posit that LP went close to fixation in populations wherein the genotype emerged during times of extreme nutrient scarcity; when the threat of starvation was so high, that LP - the ability to maximize energy extraction from lactose - was needed to survive. We postulate that this was the case 15,000 years ago in northern Europe, where LP is prevalent, but not in Mongolia 5,000 years ago and still today, where MALT and fermentation prior to consumption are sufficient to make dairy tolerable (Curry 2018). In other words, whether LP moves close to fixation or recruitment of *Bifidobacterium* are sufficient, depends on when dairying starts relative to whether or not a society has surpassed its developmental time-point wherein food is scarce.

Despite the aforementioned study weakness that we only monitored symptoms within the 3-hour phenotyping window, there was nonetheless a difference in hydrogen production, such that those with symptoms made more hydrogen. While the majority of participants submitted samples before testing began, a small minority of them submitted them later in the day - after the lactose had reached their colons - and were not recorded as such. The resulting metagenomes were presumably mildly-enriched for taxa that responded to the dose. Nonetheless, this confounder did not prevent the observed enrichment of metabolically-active MALT LNP metagenomes for *Bifidobacterium*, both in vitro and in vivo.

In conclusion, we identified a new form of lactose tolerance, conferred not by human hosts' own genomes but rather by their gut microbiotas' *Bifidobacterium*. It affords complete lactose tolerance, as both LP and MALT are associated with low-hydrogen production. Though LP results in more glucose from dairy, MALT enabled LNP-consumption of it throughout human history, which we propose facilitated dairying before the evolution of LP, and has been buffering selection for that genotype ever since.

References

Aguilera, Gabriela, Constanza Cárcamo, Sandra Soto-Alarcón, and Martin Gotteland. 2021. "Improvement in Lactose Tolerance in Hypolactasic Subjects Consuming Ice Creams with High or Low Concentrations of Bifidobacterium Bifidum 900791." *Foods (Basel, Switzerland)* 10 (10). https://doi.org/10.3390/foods10102468.

Alm, L. 1982. "Effect of Fermentation on Lactose, Glucose, and Galactose Content in Milk and Suitability of Fermented Milk Products for Lactose Intolerant Individuals." *Journal of Dairy Science* 65 (3): 346–52.

Anguita-Ruiz, Augusto, Concepción M. Aguilera, and Ángel Gil. 2020. "Genetics of Lactose Intolerance: An Updated Review and Online Interactive World Maps of Phenotype and Genotype Frequencies." *Nutrients* 12 (9). https://doi.org/10.3390/nu12092689.

Asnicar, Francesco, Andrew Maltez Thomas, Francesco Beghini, Claudia Mengoni, Serena Manara, Paolo Manghi, Qiyun Zhu, et al. 2020. "Precise Phylogenetic Analysis of Microbial Isolates and Genomes from Metagenomes Using PhyloPhlAn 3.0." *Nature Communications* 11 (1): 2500.

Asp, N-G, and A. Dahlqvist. 1974. "Intestinal β-Galactosidases in Adult Low Lactase Activity1 and in Congenital Lactase Deficiency." *Enzyme* 18: 84–102.

Beghini, Francesco, Lauren J. McIver, Aitor Blanco-Míguez, Leonard Dubois, Francesco Asnicar, Sagun Maharjan, Ana Mailyan, et al. 2021. "Integrating Taxonomic, Functional, and Strain-Level Profiling of Diverse Microbial Communities with bioBakery 3." *eLife* 10 (May). https://doi.org/10.7554/eLife.65088.

Blekhman, Ran, Julia K. Goodrich, Katherine Huang, Qi Sun, Robert Bukowski, Jordana T. Bell, Timothy D. Spector, et al. 2015. "Host Genetic Variation Impacts Microbiome Composition across Human Body Sites." *Genome Biology* 16 (September): 191.

Bonder, Marc Jan, Alexander Kurilshikov, Ettje F. Tigchelaar, Zlatan Mujagic, Floris Imhann, Arnau Vich Vila, Patrick Deelen, et al. 2016. "The Effect of Host Genetics on the Gut Microbiome." *Nature Genetics* 48 (11): 1407–12.

Bowers, Robert M., Nikos C. Kyrpides, Ramunas Stepanauskas, Miranda Harmon-Smith, Devin Doud, T. B. K. Reddy, Frederik Schulz, et al. 2017. "Minimum Information about a Single Amplified Genome (MISAG) and a Metagenome-Assembled Genome (MIMAG) of Bacteria and Archaea." *Nature Biotechnology* 35 (8): 725–31.

Brandao Gois, M. F., Trishla Sinha, Johanne E. Spreckels, Arnau Vich Vila, Laura A. Bolte, Rinse K. Weersma, Cisca Wijmenga, Jingyuan Fu, Alexandra Zhernakova, and Alexander Kurilshikov. 2022. “Role of the Gut Microbiome in Mediating Lactose Intolerance Symptoms.” *Gut*.

Catanzaro, Roberto, Morena Sciuto, and Francesco Marotta. 2021. “Lactose Intolerance: An Update on Its Pathogenesis, Diagnosis, and Treatment.” *Nutrition Research* 89 (May): 23–34.

Chaumeil, Pierre-Alain, Aaron J. Mussig, Philip Hugenholtz, and Donovan H. Parks. 2019. “GTDB-Tk: A Toolkit to Classify Genomes with the Genome Taxonomy Database.” *Bioinformatics* 36 (6): 1925–27.

Chklovski, Alex, Donovan H. Parks, Ben J. Woodcroft, and Gene W. Tyson. 2022. “CheckM2: A Rapid, Scalable and Accurate Tool for Assessing Microbial Genome Quality Using Machine Learning.” *bioRxiv*. https://doi.org/10.1101/2022.07.11.499243.

Curry, Andrew. 2018. “Early Mongolians Ate Dairy, but Lacked the Gene to Digest It.” *Science* 362 (6415): 626–27.

Delroisse, Jean-Marc, Anne-Lise Boulvin, Isabelle Parmentier, Robin Dubois Dauphin, Micheline Vandenbol, and Daniel Portetelle. 2008. “Quantification of Bifidobacterium Spp. and Lactobacillus Spp. in Rat Fecal Samples by Real-Time PCR.” *Microbiological Research* 163 (6): 663–70.

Deng, Houtao, and George Runger. 2013. “Gene Selection with Guided Regularized Random Forest.” *Pattern Recognition* 46 (12): 3483–89.

Fouhy, Fiona, Jennifer Deane, Mary C. Rea, Órla O’Sullivan, R. Paul Ross, Grace O’Callaghan, Barry J. Plant, and Catherine Stanton. 2015. “The Effects of Freezing on Faecal Microbiota as Determined Using MiSeq Sequencing and Culture-Based Investigations.” *PloS One* 10 (3): e0119355.

Gallego Romero, Irene, Chandana Basu Mallick, Anke Liebert, Federica Crivellaro, Gyaneshwer Chaubey, Yuval Itan, Mait Metspalu, et al. 2012. “Herders of Indian and European Cattle Share Their Predominant Allele for Lactase Persistence.” *Molecular Biology and Evolution* 29 (1): 249–60.

Goodrich, Julia K., Emily R. Davenport, Michelle Beaumont, Matthew A. Jackson, Rob Knight, Carole Ober, Tim D. Spector, Jordana T. Bell, Andrew G. Clark, and Ruth E. Ley. 2016.

"Genetic Determinants of the Gut Microbiome in UK Twins." *Cell Host & Microbe* 19 (5): 731–43.

Goodrich, Julia K., Emily R. Davenport, Jillian L. Waters, Andrew G. Clark, and Ruth E. Ley. 2016. "Cross-Species Comparisons of Host Genetic Associations with the Microbiome." *Science* 352 (6285): 532–35.

Guerrant, R. L., J. B. Schorling, J. F. McAuliffe, and M. A. de Souza. 1992. "Diarrhea as a Cause and an Effect of Malnutrition: Diarrhea Prevents Catch-up Growth and Malnutrition Increases Diarrhea Frequency and Duration." *The American Journal of Tropical Medicine and Hygiene* 47 (1 Pt 2): 28–35.

Hollfelder, Nina, Hiba Babiker, Lena Granehäll, Carina M. Schlebusch, and Mattias Jakobsson. 2021. "The Genetic Variation of Lactase Persistence Alleles in Sudan and South Sudan." *Genome Biology and Evolution* 13 (5). https://doi.org/10.1093/gbe/evab065.

Iablokov, Stanislav N., Pavel S. Novichkov, Andrei L. Osterman, and Dmitry A. Rodionov. 2021. "Binary Metabolic Phenotypes and Phenotype Diversity Metrics for the Functional Characterization of Microbial Communities." *Frontiers in Microbiology* 12 (May): 653314.

Ingram, Catherine J. E., Charlotte A. Mulcare, Yuval Itan, Mark G. Thomas, and Dallas M. Swallow. 2009. "Lactose Digestion and the Evolutionary Genetics of Lactase Persistence." *Human Genetics* 124 (6): 579–91.

Itan, Yuval, Bryony L. Jones, Catherine J. E. Ingram, Dallas M. Swallow, and Mark G. Thomas. 2010. "A Worldwide Correlation of Lactase Persistence Phenotype and Genotypes." *BMC Evolutionary Biology* 10 (February): 36.

Jeong, Choongwon, Shevan Wilkin, Tsend Amgalantugs, Abigail S. Bouwman, William Timothy Treal Taylor, Richard W. Hagan, Sabri Bromage, et al. 2018. "Bronze Age Population Dynamics and the Rise of Dairy Pastoralism on the Eastern Eurasian Steppe." *Proceedings of the National Academy of Sciences of the United States of America* 115 (48): E11248–55.

Jiang, Hongshan, Rong Lei, Shou-Wei Ding, and Shuifang Zhu. 2014. "Skewer: A Fast and Accurate Adapter Trimmer for next-Generation Sequencing Paired-End Reads." *BMC Bioinformatics* 15 (June): 182.

Jiang, T., A. Mustapha, and D. A. Savaiano. 1996. "Improvement of Lactose Digestion in Humans by Ingestion of Unfermented Milk Containing Bifidobacterium Longum." *Journal of Dairy Science* 79 (5): 750–57.

Jones, Bryony Leigh, Tamiru Oljira, Anke Liebert, Pawel Zmarz, Nicolas Montalva, Ayele Tarekeyn, Rosemary Ekong, et al. 2015. “Diversity of Lactase Persistence in African Milk Drinkers.” *Human Genetics* 134 (8): 917–25.

Karasov, Talia L., Juliana Almario, Claudia Friedemann, Wei Ding, Michael Giolai, Darren Heavens, Sonja Kersten, et al. 2018. “Arabidopsis Thaliana and Pseudomonas Pathogens Exhibit Stable Associations over Evolutionary Timescales.” *Cell Host & Microbe*. https://doi.org/10.1016/j.chom.2018.06.011.

Kassambara, Alboukadel. 2020. “ggpubr:‘ggplot2’ Based Publication Ready Plots.” *R Package Version 0. 4. 0* 438.

Letunic, Ivica, and Peer Bork. 2021. “Interactive Tree Of Life (iTOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation.” *Nucleic Acids Research* 49 (W1): W293–96.

Liu, Wenjun, Jiachao Zhang, Chunyan Wu, Shunfeng Cai, Weiqiang Huang, Jing Chen, Xiaoxia Xi, et al. 2017. “Corrigendum: Unique Features of Ethnic Mongolian Gut Microbiome Revealed by Metagenomic Analysis.” *Scientific Reports* 7 (January): 39576.

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550.

Maloy, Stanley, and Kelly Hughes. 2013. *Brenner’s Encyclopedia of Genetics*. Academic Press.

Masoumi, Seyed Jalil, Davood Mehrabani, Mehdi Saberifiroozi, Mohammad Reza Fattahi, Fariba Moradi, and Masoud Najafi. 2021. “The Effect of Yogurt Fortified with Lactobacillus Acidophilus and Bifidobacterium Sp. Probiotic in Patients with Lactose Intolerance.” *Food Science & Nutrition* 9 (3): 1704–11.

McMurdie, Paul J., and Susan Holmes. 2013. “Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data.” *PloS One* 8 (4): e61217.

Misselwitz, Benjamin, Daniel Pohl, Heiko Frühauf, Michael Fried, Stephan R. Vavricka, and Mark Fox. 2013. “Lactose Malabsorption and Intolerance: Pathogenesis, Diagnosis and Treatment.” *United European Gastroenterology Journal* 1 (3): 151–59.

Moffett, John R., Narayanan Puthillathu, Ranjini Vengilote, Diane M. Jaworski, and Aryan M. Namboodiri. 2020. “Acetate Revisited: A Key Biomolecule at the Nexus of Metabolism, Epigenetics and Oncogenesis-Part 1: Acetyl-CoA, Acetogenesis and Acyl-CoA Short-Chain Synthetases.” *Frontiers in Physiology* 11 (November): 580167.

Mysore Saiprasad, Sindusha, Olivia Grace Moreno, and Dennis A. Savaiano. 2023. “A Narrative

Review of Human Clinical Trials to Improve Lactose Digestion and Tolerance by Feeding Bifidobacteria or Galacto-Oligosacharides." *Nutrients* 15 (16). https://doi.org/10.3390/nu15163559.

Olm, Matthew R., Alexander Crits-Christoph, Keith Bouma-Gregson, Brian A. Firek, Michael J. Morowitz, and Jillian F. Banfield. 2021. "inStrain Profiles Population Microdiversity from Metagenomic Data and Sensitively Detects Shared Microbial Strains." *Nature Biotechnology*, January, 1–10.

Parche, Stephan, Manfred Beleut, Enea Rezzonico, Doris Jacobs, Fabrizio Arigoni, Fritz Titgemeyer, and Ivana Jankovic. 2006. "Lactose-over-Glucose Preference in Bifidobacterium Longum NCC2705: glcP, Encoding a Glucose Transporter, Is Subject to Lactose Repression." *Journal of Bacteriology* 188 (4): 1260–65.

Parks, Donovan H., Maria Chuvochina, Christian Rinke, Aaron J. Mussig, Pierre-Alain Chaumeil, and Philip Hugenholtz. 2022. "GTDB: An Ongoing Census of Bacterial and Archaeal Diversity through a Phylogenetically Consistent, Rank Normalized and Complete Genome-Based Taxonomy." *Nucleic Acids Research* 50 (D1): D785–94.

Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel A. R. Ferreira, David Bender, Julian Maller, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81 (3): 559–75.

Ranciaro, Alessia, Michael C. Campbell, Jibril B. Hirbo, Wen-Ya Ko, Alain Froment, Paolo Anagnostou, Maritha J. Kotze, et al. 2014. "Genetic Origins of Lactase Persistence and the Spread of Pastoralism in Africa." *American Journal of Human Genetics* 94 (4): 496–510.

R Core Team. 2018. "R: A Language and Environment for Statistical Computing." Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rezaie, Ali, Michelle Buresi, Anthony Lembo, Henry Lin, Richard McCallum, Satish Rao, Max Schmulson, Miguel Valdovinos, Salam Zakko, and Mark Pimentel. 2017. "Hydrogen and Methane-Based Breath Testing in Gastrointestinal Disorders: The North American Consensus." *The American Journal of Gastroenterology* 112 (5): 775–84.

Ruaud, Albane, Niklas Pfister, Ruth E. Ley, and Nicholas D. Youngblut. 2022. "Interpreting Tree Ensemble Machine Learning Models with endoR." *PLoS Computational Biology* 18 (12): e1010714.

Ségurel, Laure, and Céline Bon. 2017. “On the Evolution of Lactase Persistence in Humans.” *Annual Review of Genomics and Human Genetics* 18 (August): 297–319.

Segurel, Laure, Perle Guarino-Vignon, Nina Marchi, Sophie Lafosse, Romain Laurent, Céline Bon, Alexandre Fabre, Tatyana Hegay, and Evelyne Heyer. 2020. “Why and When Was Lactase Persistence Selected for? Insights from Central Asian Herders and Ancient DNA.” *PLoS Biology* 18 (6): e3000742.

Suzuki, Taichi A., J. Liam Fitzstevens, Victor T. Schmidt, Hagay Enav, Kelsey E. Huus, Mirabeau Mbong Ngwese, Anne Grießhammer, et al. 2022. “Codiversification of Gut Microbiota with Humans.” *Science* 377 (6612): 1328–32.

Suzuki, Taichi A., and Ruth E. Ley. 2020. “The Role of the Microbiota in Human Genetic Adaptation.” *Science* 370 (6521): eaaz6827.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the Tidyverse.” *Journal of Open Source Software* 4 (43): 1686.

Wilkin, Shevan, Alicia Ventresca Miller, William T. T. Taylor, Bryan K. Miller, Richard W. Hagan, Madeleine Bleasdale, Ashley Scott, et al. 2020. “Dairy Pastoralism Sustained Eastern Eurasian Steppe Populations for 5,000 Years.” *Nature Ecology & Evolution* 4 (3): 346–55.

Wright, M. N., S. Wager, and P. Probst. 2020. “Ranger: A Fast Implementation of Random Forests.” *R Package Version 0.12*. https://xscode.com/imbs-hl/ranger.

Xue, Hong, Min Zhang, Jinxin Ma, Ting Chen, Fengyun Wang, and Xudong Tang. 2020. “Lactose-Induced Chronic Diarrhea Results From Abnormal Luminal Microbial Fermentation and Disorder of Ion Transport in the Colon.” *Frontiers in Physiology* 11 (July): 877.

Youngblut, Nicholas D., Jacobo de la Cuesta-Zuluaga, Georg H. Reischer, Silke Dauser, Nathalie Schuster, Chris Walzer, Gabrielle Stalder, Andreas H. Farnleitner, and Ruth E. Ley. 2020. “Large-Scale Metagenome Assembly Reveals Novel Animal-Associated Microbial Genomes, Biosynthetic Gene Clusters, and Other Genetic Diversity.” *mSystems* 5 (6). https://doi.org/10.1128/mSystems.01045-20.

Youngblut, Nicholas D., and Ruth E. Ley. 2021. “Struo2: Efficient Metagenome Profiling Database Construction for Ever-Expanding Microbial Genome Datasets.” *PeerJ* 9

(September): e12198.

Supplemental Figures

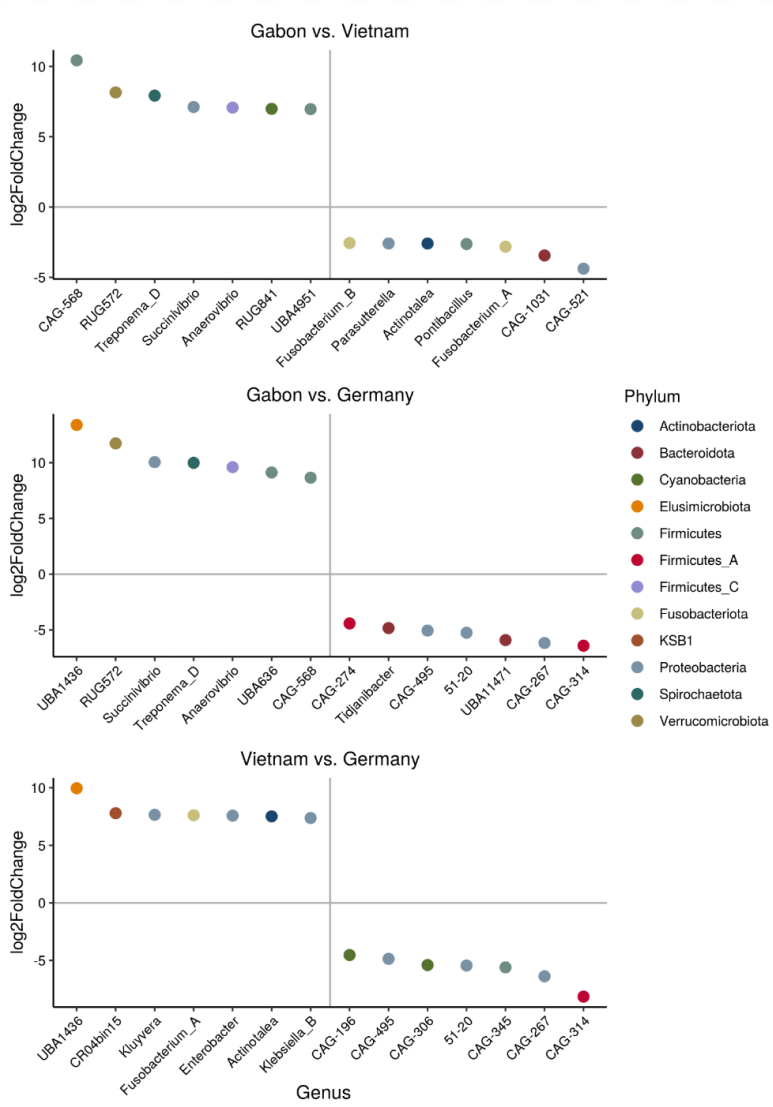

Figure S1. Differential abundance of genera between each country's in vivo metagenomes. Values calculated with DESeq2; genera are colored by phyla.

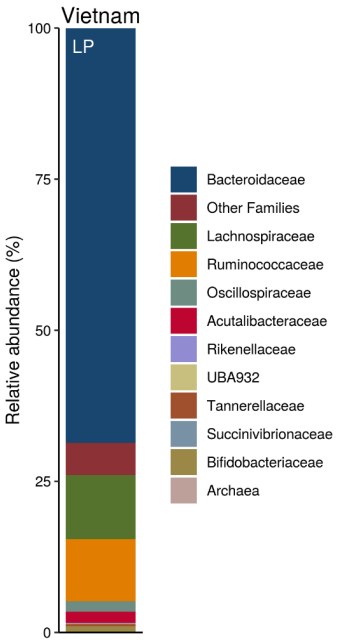

Figure S2. Relative abundances of microbial taxa sequenced from the single LP participant from Vietnam. Each color represents either a single bacterial family, all archaea, or the sum of all other bacterial families represented in each metagenome; colors correspond to the same taxa as given in Figure 1 D.

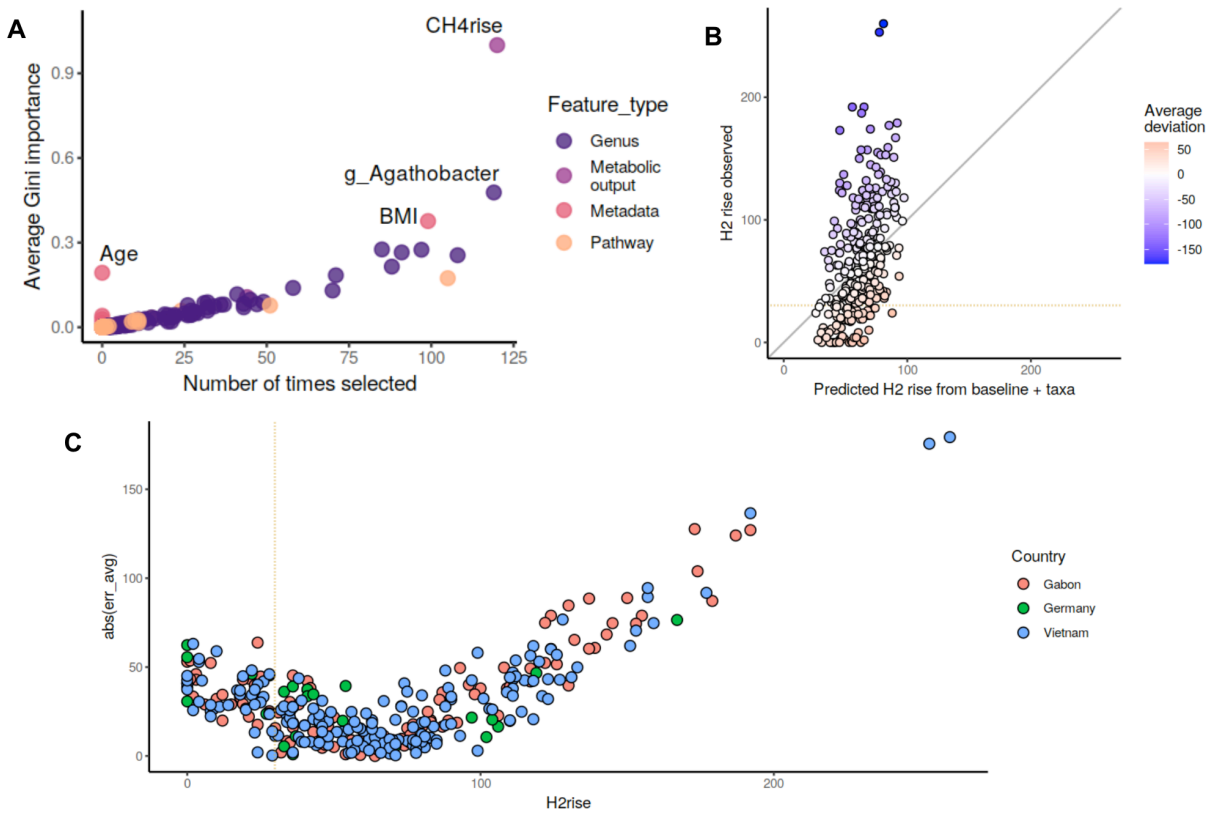

Figure S3. Accuracy of the random forest regression model predicting in vivo H_2 rise with host covariates and microbial relative abundances. (A) Number of times features were selected and average Gini importance across the 120 cross validation sets with 80-20 % train-test dataset split. (B) Comparison of the average predicted value with the observed one. The color indicates the difference on average between the predicted and observed values. The grey line indicates perfect prediction (y = x); the yellow dotted line (H_2 rise = 30 ppm) indicates the threshold delimiting lactose tolerant from intolerant individuals. (C) Absolute deviation difference on average between the predicted and observed values (y-axis) according to the observed value (x-axis). Points are colored by sampling country; accuracy of predictions did not depend on the country; the yellow dotted line (H_2 rise = 30 ppm) indicates the threshold differentiating lactose-tolerant from -intolerant individuals.

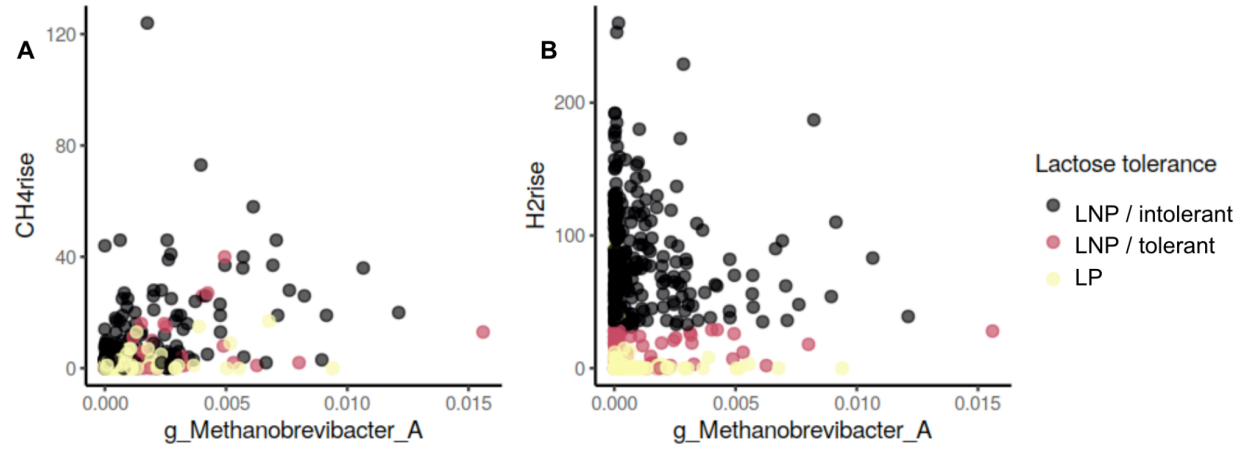

Figure S4. Relative abundances of *Methanobrevibacter* A, the most abundant and prevalent human gut methanogen, plotted against production of breath (A) CH_4 or (B) H_2 from lactose in vivo.

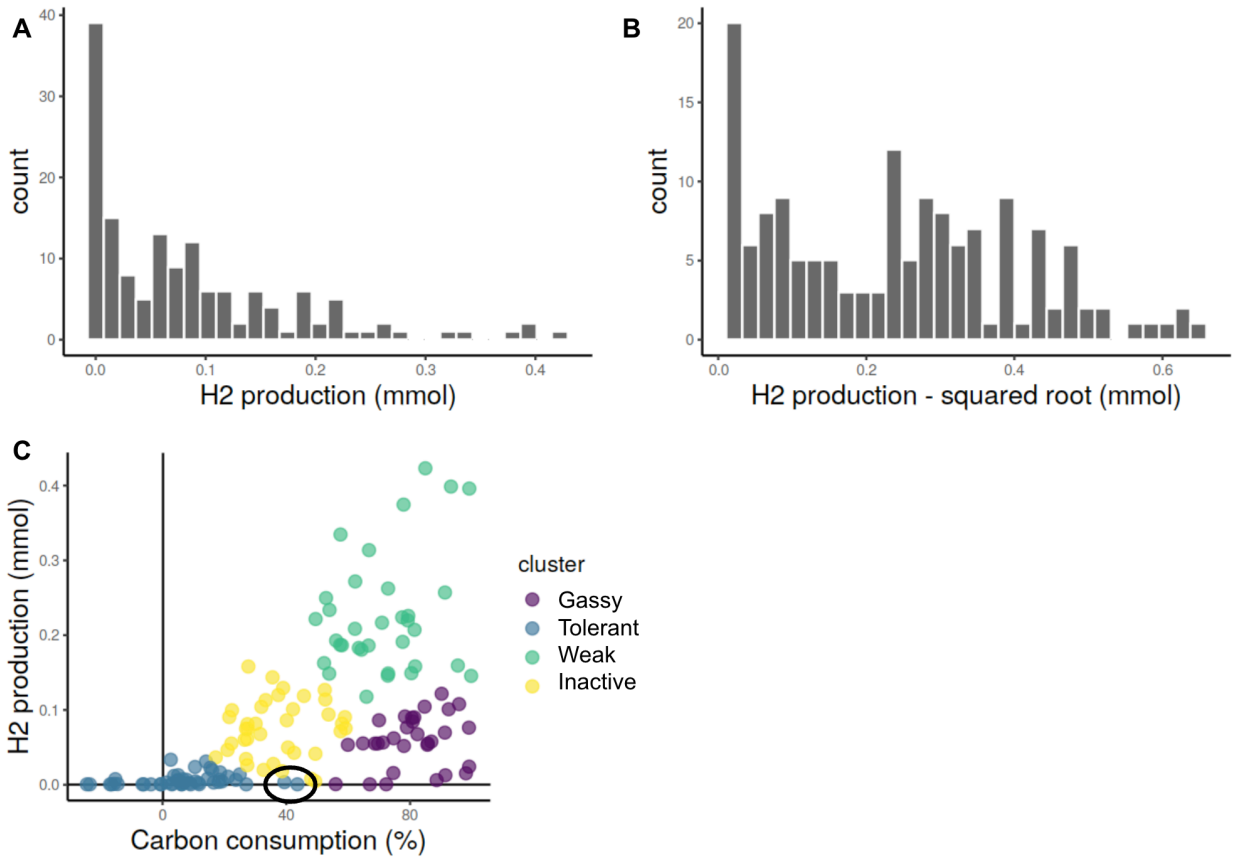

Figure S5. Clustering of vitro samples based on H_2 production and carbon consumption. (A) Distribution of H_2 production in vitro samples. (B) Distribution of H_2 production in in vitro samples after square root transformation. (C) K-means clustering of samples based on H_2

production (square root transformed) and carbon consumption. The two circled samples from cluster 2, which group we then named “inactives”, were manually reclassified as cluster 4; this cluster was then named LC for low carbon consumption.

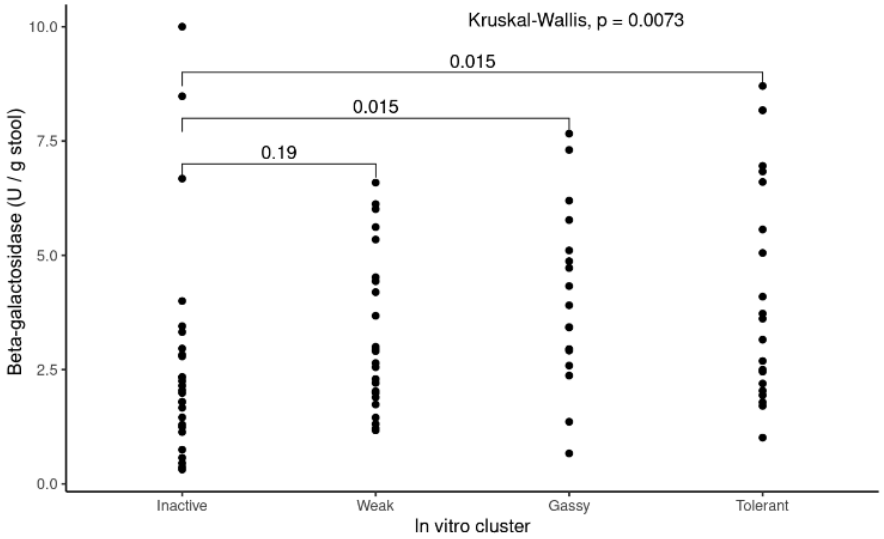

Figure S6. Extracellular beta-galactosidase assessed by in-vitro cluster.

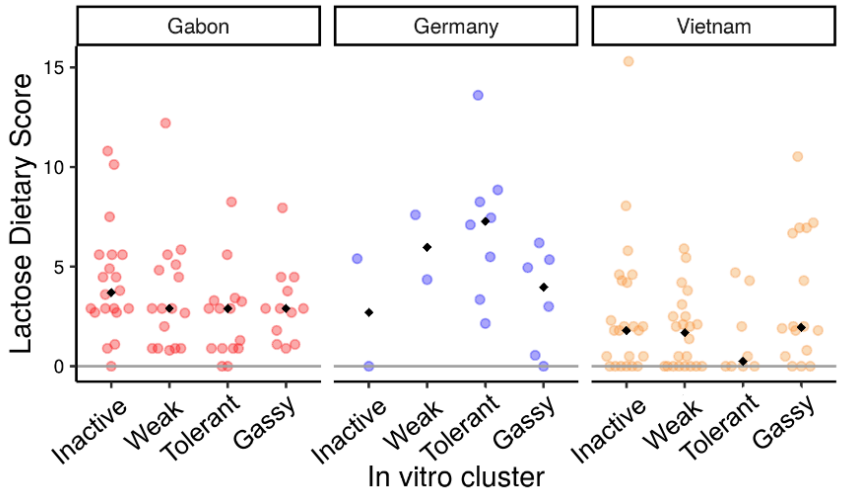

Figure S7. Dietary lactose score by in vitro cluster group by country.

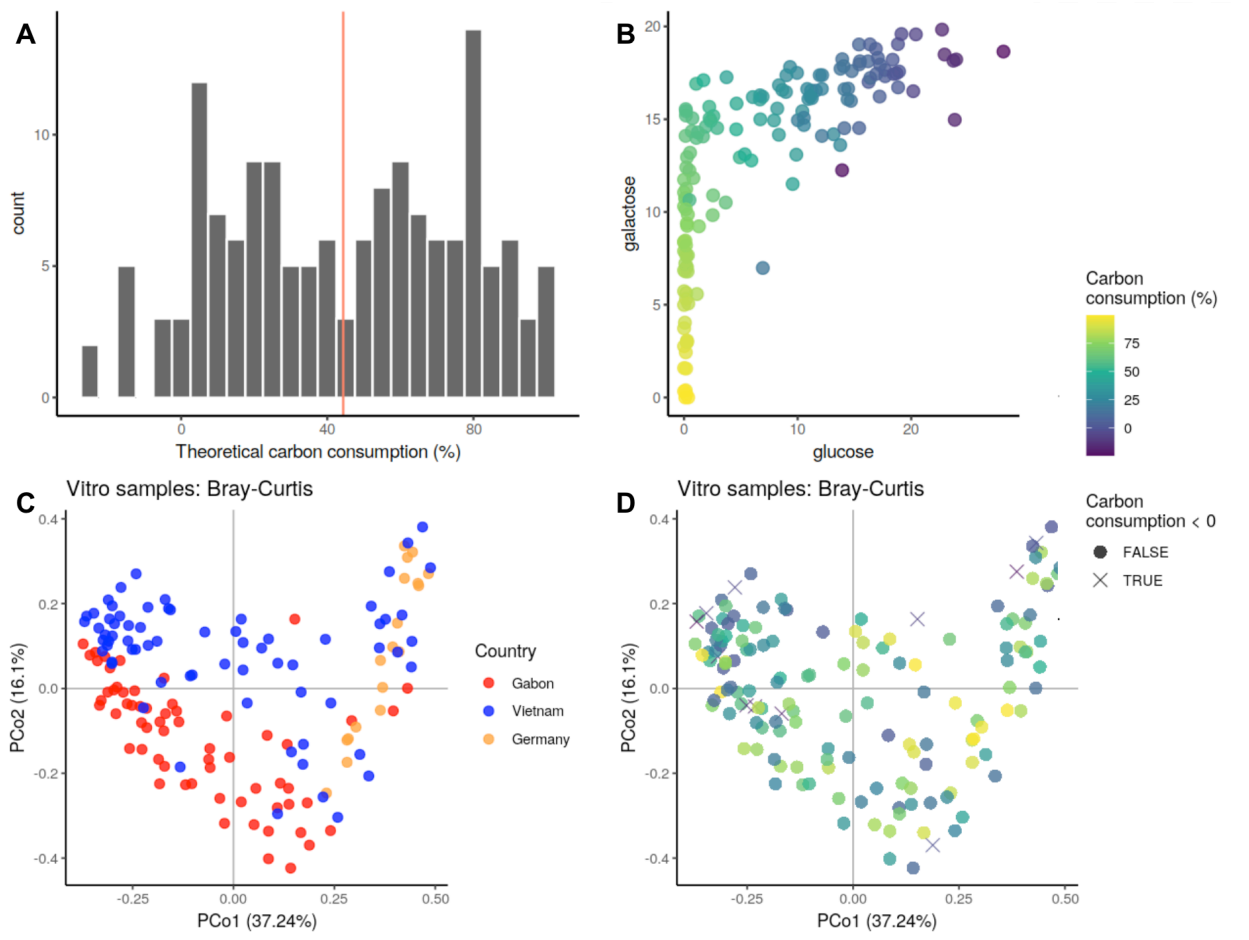

Figure S8. Gut microbiotas' consumption of lactose during 4.5 h of anaerobic cultivation with 36 mg of lactose. (A) Theoretical lactose consumed. (B) Glucose and galactose produced during the experiment (difference of final concentrations with starting ones). Points are colored by carbon consumption. (C-D) Principal coordinates of Bray-Curtis distances calculated with relative abundances of genera in samples at the beginning of the experiments. (C) Samples are colored by country. (D) Samples are colored by theoretical carbon consumption and represented by a cross if the theoretical carbon consumption was negative.

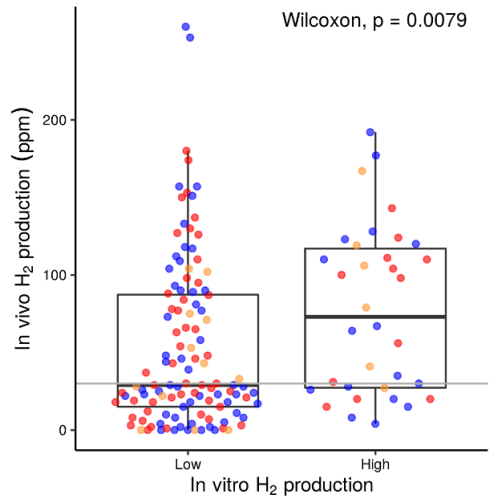

Figure S9. In vivo hydrogen production of participants whose stools were classified as low- and high-hydrogen producers in vitro, respectively. The "Low" hydrogen producers in vitro represent the Inactive, Weak and Tolerant groups, combined, and the "High" group are the Gassy samples.

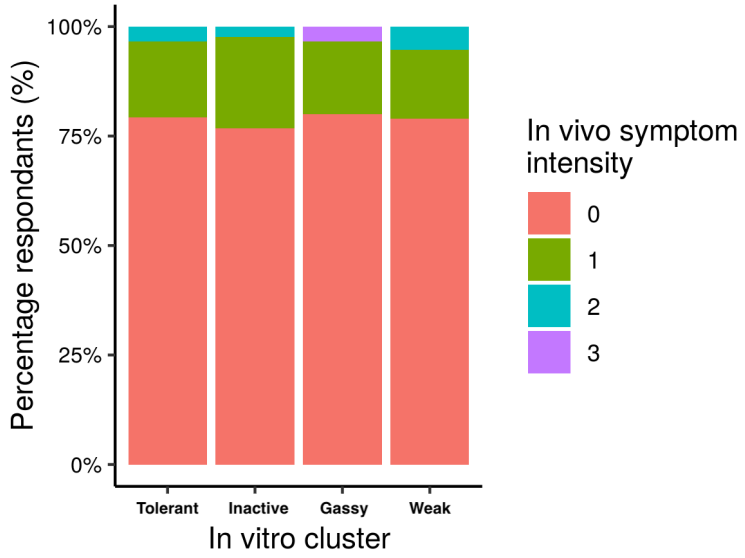

Figure S10. In vivo symptoms by in vitro classification. Study respondents' self-reported symptom discomfort at the end of the 3 h lactose tolerance phenotyping; "0", "1", "2", and "3" corresponded to no, mild, moderate, and severe gastrointestinal distress. They are grouped by their in vitro stool classifications.

Supplemental Tables

Table S1. Participant metadata, lactose-genotyping and -phenotyping overview.

	Country			All
	Gabon	Vietnam	Germany	
N	152	190	141	483
% Female	100	100	100	100
% Born In-Country	100	100	100	100
Age Mean ± SE	26.4 ± 0.51	27.5 ± 0.38	29.4 ± 0.44	27.7 ± 0.26
BMI Mean ± SE	24.8 ± 0.509	21.6 ± 0.200	23.8 ± 0.350	23.1 ± 0.198
Metagenome reads Mean ± SE	3,618,585 ± 60,421	3,811,863 ± 61,970	3,666,806 ± 78,023	3,708,692 ± 38,516
LNP (% of country total)	100.0	99.5	23.4	77.4
Mean H2 rise ± SE	71.9 ± 3.88	65.3 ± 3.23	54.5 ± 6.30	67.0 ± 2.35
% LT by HBT	20.4	22.2	18.2	21.1
% LT by HBT only	18.4	21.7	18.2	20.1
% LT by HBT & BGT	2.0	0.5	0.0	1.1
Mean BGT rise ± SE	6.95 ± 0.629	1.05 ± 0.456	10.3 ± 1.20	4.28 ± 0.399
% LT by BGT	5.9	0.5	12.1	3.7
% LT by BGT only	3.9	0.0	12.1	2.7
Mean CH4 rise ± SE	11.0 ± 1.31	3.67 ± 0.363	2.58 ± 0.756	6.56 ± 0.597
LP (% of country total)	0.0	0.5	76.6	22.6
Mean H2 rise ± SE	-	93.0 ± N/A	1.86 ± 0.452	2.70 ± 0.950
% LT by HBT	-	0.0	99.1	0.9
% LT by HBT only	-	0.0	19.4	19.3

% LT by HBT & BGT	-	0.0	79.6	78.9
Mean BGT rise ± SE	-	4.0 ± N/A	29.0 ± 1.11	28.7 ± 1.12
% LT by BGT	-	0.0	79.6	78.9
% LT by BGT only	-	0.0	0.0	0.0
Mean CH4 rise ± SE	N/A	3.00 ± N/A	1.17 ± 0.275	1.18 ± 0.273

Abbreviations	
BMI	body mass index
LNP	lactase non-persistant
LP	lactase persistant
LT	lactose tolerant
HBT	hydrogen breath test (lactose tolerant: rise < 30 ppm)
BGT	blood glucose test (lactose tolerant: rise > 20 mg / dL)

Table S2. Metadata p values in in vivo hydrogen modeling for both LNPs and LPs.

Model	Feature (a)	Coefficient	P-value	Adjusted p-value (b)	Significant
Binomial: Y = log(p(H2 rise > 1 ppm) / p(H2 rise <= 1 ppm))	Glucoserise	0.19	7.41E-03	1.67E-02	TRUE
	Age	0.15	6.95E-03	1.67E-02	TRUE
	Country: Gemany vs Gabon & Vietnam	-2.12	8.05E-09	1.45E-07	TRUE
	Country: Gabon vs Vietnam	-0.08	8.71E-0	1.00E+00	FALSE

			1		
	CH4rise	0.13	2.89E-02	5.20E-02	FALSE
	Glucoserise:Age	-0.01	1.13E-03	4.06E-03	TRUE
Gamma: Y = log(H2 rise) for individuals with H2 rise > 1 ppm	Glucoserise	-0.02	1.96E-06	1.76E-05	TRUE
	Age	0.02	5.16E-03	1.55E-02	TRUE
	BMI	-0.03	7.07E-04	3.18E-03	TRUE
	Country: Gemany vs Gabon & Vietnam	-0.33	2.12E-04	1.27E-03	TRUE
	Country: Gabon vs Vietnam	-0.19	9.68E-03	1.94E-02	TRUE
	CH4rise	0.01	6.45E-02	1.05E-01	FALSE
(a)	The full model comprised the following features: glucose rise, age, BMI, glucose rise * age, glucose rise*BMI, country, CH4 rise, lactose score. Features were removed sequentially untill all were significant.				
(b)	p-values were corrected using the Benjamini-Hochberg method with n = 18 (i.e., taking into account that all tested features and interactions).				

Table S3. Metadata p values in in vivo hydrogen modeling for LNPs only.

Feature	Coefficient	p-value	Adjusted p-value	Significant

Age	0.02	0.0063	0.0377	TRUE
BMI	-0.03	0.001	0.0121	TRUE
Betagal	0.04	0.0287	0.0688	FALSE
Country: Germany vs others	-0.22	0.0114	0.0455	TRUE
Country: Gabon vs Vietnam	-0.15	0.0282	0.0688	FALSE

Table S4. Sample groupings by their carbon consumption and hydrogen production in vitro.

Group	Percent carbon consumed (%, mean ± sd)	H2 produced (mg, mean ± sd)	Country			
			Gabon	Vietnam	Germany	Total
Inactive	5.20 ± 12.81	0.01 ± 0.01	20	23	2	45
Weak	37.53 ± 11.81	0.07 ± 0.04	16	22	2	40
Gasssy	71.49 ± 14.25	0.23 ± 0.08	15	8	8	34
Tolerant	80.57 ± 11.47	0.06 ± 0.03	12	16	6	31

Table S5. Predicting metabolic group with random forest with genera and pathway relative abundances as features.

		Balanced accuracy					
In vitro experiment	Overall accuracy	Inactives	LCLH	HCLH	HCHH	# features selected	# selected features that were pathways
Beginning	0.4961±0.0914	0.6709±0.0896	0.5734±0.0785	0.7018±0.1200	0.6294±0.1296	146	1

End	0.5172±0.0818	0.7438±0.0920	0.5322±0.1122	0.8027±0.0953	0.6659±0.09707	114	29

Table S6. Carbohydrate phenotype descriptions and metadata.

Phenotype	Description	Type_group	Detailed_group	Origin	Comment
Glc	D-glucose utilization	monosaccharides_and_derivatives	monosaccharides	universal	D-Glucose (Glc) is an aldohexose. Glc is a building block of many dietary di-, oligo-, and polysaccharides, including sucrose, lactose, raffinose, starch, glycogen, and cellulose. Glc is absorbed and metabolized by the host. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Glc utilizers
Gal	D-galactose utilization	monosaccharides_and_derivatives	monosaccharides	universal	D-Galactose (Gal) is an aldohexose. Gal is a building block of many dietary oligo- and polysaccharides of plant origin (e.g., raffinose, galactan, galactomannan), lactose, human milk oligosaccharides, glycosaminoglycans, N- and O-linked glycans of plant/animal origin (including the human host). Gal is absorbed and metabolized by the host and fermented by gut microbiota. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Gal utilizers
Rbs	D-ribose utilization	monosaccharides_and_derivatives	monosaccharides	universal	D-Ribose (Rbs) is an aldopentose. Rbs is a building block of ribonucleotides, universally essential metabolites such as ATP, NAD, etc. Rbs and Rbs-5P are typically produced from glucose by the pentose phosphate pathway in most cell types (including the host); however, Rbs is rarely present in nature in free form. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Rbs utilizers

Fuc	L-fucose utilization	monosaccharides_and_derivatives	deoxymonosaccharides	universal	L-Fucose (Fuc) is a deoxyhexose. Fuc is a building block of human milk oligosaccharides, N- and O-linked glycans of plant/animal origin (including the human host). Fuc is also found in side chains of some plant polysaccharides, for example, xyloglucans. Fuc can be salvaged by the host and used for protein fucosylation. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Fuc utilizers
FL	2'FL and 3FL utilization	di_and_oligosaccharides	oligosaccharides	animal	2'- and 3-fucosyllactoses (FLs) are trisaccharides with structures Fuc(a1-2)Gal(b1-4)Glc and Fuc(a1-3)[Gal(b1-4)]Glc, respectively. 2'FL is the most abundant oligosaccharide in the milk of mothers with secretor status, whereas 3FL is abundant in the milk of non-secretors. FLs are not metabolized by the host. Strains harboring (i) a complete set of genes encoding transporters and catabolic pathways OR (ii) extracellular α-fucosidases are deemed FL utilizers
LNFPI	LDFT and LNFP I utilization	di_and_oligosaccharides	oligosaccharides	animal	2'- and 3-fucosyllactoses (FLs) are trisaccharides with structures Fuc(a1-2)Gal(b1-4)Glc and Fuc(a1-3)[Gal(b1-4)]Glc, respectively. 2'FL is the most abundant oligosaccharide in the milk of mothers with secretor status, whereas 3FL is abundant in the milk of non-secretors. The host does not metabolize FLs. Strains harboring (i) a complete set of genes encoding transporters and catabolic pathways OR (ii) extracellular α-fucosidases (GH29 and GH95) are deemed FL utilizers
GlcNAc	N-acetylglucosamine utilization	monosaccharides_and_derivatives	aminosugars	universal	N-acetylglucosamine (GlcNAc) is an aminohexose. GlcNAc is a building block of various polysaccharides (murein, chitin), human milk oligosaccharides,

					glycosaminoglycans, N- and O-linked glycans of plant/animal origin (including the human host). Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed GlcNAc utilizers
LNB	Lacto-N-biose utilization	di_and_oligosaccharides	disaccharides	animal	Lacto-N-biose (LNB) is a disaccharide with the structure Gal(b1-3)GlcNAc. LNB is a structural component of type I chain human milk oligosaccharides and glycolipids; it is released by extracellular lacto-N-biosidases (GH20 and 136). The host does not metabolize LNB. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed LNB utilizers
LNT	LNT utilization	di_and_oligosaccharides	oligosaccharides	animal	Lacto-N-tetraose (LNT) is a tetrasaccharide with the structure Gal(b1-3)GlcNAc(b1-3)Gal(b1-4)Glc. LNT is a major type I chain oligosaccharide found in human milk. The host does not metabolize LNT. Strains harboring (i) a complete set of genes encoding transporters and catabolic pathways OR (ii) extracellular lacto-N-biosidases (GH20, GH136) are deemed LNT utilizers
LNnT	LNnT utilization	di_and_oligosaccharides	oligosaccharides	animal	Lacto-N-tetraose (LNnT) is a tetrasaccharide with the structure Gal(b1-4)GlcNAc(b1-3)Gal(b1-4)Glc. LNnT is a major type II chain oligosaccharide found in human milk. The host does not metabolize LNnT. Strains harboring (i) a complete set of genes encoding transporters and catabolic pathways OR (ii) extracellular β-galactosidases (GH2) and β-N-acetylglucosaminidases (GH20) are deemed LNnT utilizers

HMO	HMO utilization via H1 cluster ABC transporters	di_and_oligosaccharides	oligosaccharides	animal	Human milk oligosaccharides (HMOs) are oligosaccharides with DP 3-20 found in human milk. The host does not metabolize HMOs. Strains harboring a specific gene cluster (H1) encoding multiple predicted HMO transporters and glycoside hydrolases are deemed intracellular utilizers of multiple HMO species (instead of individual oligosaccharides)
Nglyc	N-glycan utilization	di_and_oligosaccharides	oligosaccharides	universal	N-glycans (Nglyc) are oligosaccharides covalently attached to protein at asparagine (Asn) residues by an N-glycosidic bond. N-glycans decorate many prokaryotic and eukaryotic (plant, fungal, animal) proteins; all eukaryotic N-glycans begin with GlcNAcb1–Asn. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Nglyc utilizers
Nglyc_core	N-glycan core utilization	di_and_oligosaccharides	disaccharides	universal	N-glycan core (Nglyc_core) is a glycopeptide with the structure Fuc(a1-6)GlcNAcb1-Asn-peptide. Nglyc_core is released during the cleavage of the N,N′-diacetylchitobiose core of N-glycans (predominantly complex and hybrid) by bacterial endo-β-N-acetylglucosaminidases (e.g., GH18). Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Nglyc_core utilizers
GNB	Galacto-N-biose utilization	di_and_oligosaccharides	disaccharides	animal	Galacto-N-biose (GNB) is a disaccharide with the structure Gal(b1-3)GalNAc. GNB is a structural component of Core I/II mucin O-glycans; it can be released by endo-α-N-acetylgalactosaminidases (GH101). The host does not metabolize GNB. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed GNB utilizers

Muc	Mucin O-glycan degradation	di_and_oligosaccharides	oligosaccharides	animal	Mucin O-glycans (Muc) are oligosaccharides attached through GalNAc to protein at serine (Ser) or threonine (Thr) residues by an O-glycosidic bond. Muc are the primary constituents of mucins, glycoproteins that are expressed on various mucosal sites of the body, especially the intestinal tract. Strains harboring a complete set of genes encoding extracellular glycoside hydrolases that step-by-step degrade O-glycan chains are deemed Muc degraders
GlcNAc6S	N-acetylglucosamine-6-sulfate utilization	monosaccharides_and_derivatives	aminosugars	animal	N-acetylglucosamine-6-sulfate utilization (GlcNAc6S) is a sulfated derivative of GlcNAc. GlcNAc6S can be found in terminal or branched positions of mucin O-glycans; it is released by GH20 sulfoglycosidases secreted, for example, by Bifidobacterium bifidum. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed GlcNAc6S utilizers
NANA	N-acetylneuraminic acid utilization	monosaccharides_and_derivatives	sialic_acids	animal	N-acetylneuraminic acid (NANA) is the predominant sialic acid (9-carbon backbone 2-keto acid sugar) in mammalian cells (including the host). NANA is a building block of human milk oligosaccharides, complex N-glycans, O-glycans, and lipid-associated glycoconjugates (gangliosides). NANA can be recycled or degraded by the host. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed NANA utilizers
SL	Sialyllactose utilization	di_and_oligosaccharides	oligosaccharides	animal	3'- and 6'-sialyllactose (SLs) are trisaccharides with structures NeuNAc(a2-3)Gal(b1-4)Glc and NeuNAc(a2-6)Gal(b1-4)Glc, respectively. SLs are among the most abundant sialylated human milk oligosaccharides and are also found in the milk of other mammals. The host does not metabolize SLs. Strains harboring (i)

					a complete set of genes encoding catabolic pathways OR (ii) extracellular α-sialidases are deemed SL utilizers
Mal	Maltose utilization	di_and_oligosaccharides	disaccharides	universal	Maltose (Mal) is a disaccharide with the structure Glc(a1-4)a-Glc. Mal is produced from starch and glycogen by extracellular α-amylases (including those secreted by the host). Maltose can be found in starch-derived products like maltodextrin and corn syrup. The host metabolizes Mal. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Mal utilizers
MOS	Maltooligosaccharide utilization	di_and_oligosaccharides	oligosaccharides	universal	Maltooligosaccharides (MOS) are oligomers (DP 2-8) composed of 1,4-linked α-D-Glcp residues. Example: maltotriose with the structure Glc(a1-4)Glc(a1-4)a-Glc. MOS are produced from starch, pullulan, and glycogen by extracellular α-amylases and pullulanases (including those secreted by the host). The host metabolizes MOS. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed MOS utilizers
ST	Starch degradation	polysaccharides	homopolysaccharides	plant	Starch (ST) is a polysaccharide most green plants produce for energy storage. ST consists of α-D-Glcp residues that form 1,4-linked linear (amylose) and 1,6-branched chains (amylopectin). The host can metabolize ST. Strains harboring genes encoding extracelluar α-amylases (GH13_28 or GH13_32) AND a complete set of genes encoding transporters and catabolic pathways for the released maltose and maltooligosaccharides are deemed ST degraders
RST	Resistant starch degradation	polysaccharides	homopolysaccharides	plant	Resistant starch (RST) is starch that reaches the large intestine without being fully digested due to native supramolecular structure and morphology OR chemical modifications. The

					host does not metabolize RST. Strains harboring genes encoding extracellular RST-binding α-amylases (GH13_28) AND a complete set of genes encoding transporters and catabolic pathways for the released maltose and maltooligosaccharides are deemed RST degraders
IMO	Panose and isomaltooligosaccharide utilization	di_and_oligosaccharides	disaccharides;oligosaccharides	plant	Panose and isomaltooligosaccharides (IMOs) are oligosaccharides with structures Glc(a1-6)Glc(a1-4)Glc, Glc(a1-6)Glc(isomaltose), Glc(a1-6)Glc(a1-6)Glc (isomaltotriose), respectively. IMOs are found naturally in some foods and manufactured commercially from starch, which is enzymatically converted into a mixture of isomaltooligosaccharides. The host does not metabolize IMOs. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed IMO utilizers
Mlz	Melezitose utilization	di_and_oligosaccharides	oligosaccharides	animal	Melezitoze (Mlz) is a trisaccharide with the structure Glc(a1-3)Fruf(b2-1a)Glc. Mlz is the primary trisaccharide in honeydew, especially in the more common honeydew of aphids that live on spruces, constituting up to 70% of the sugar fraction. The host does not metabolize Mlz. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Mlz utilizers
Cld	Cellobiose and cellodextrin utilization	di_and_oligosaccharides	disaccharides;oligosaccharides	plant	Cellobiose and cellodextrin (Cld) are oligomers (DP 2-8) composed of 1,4-linked β-D-Glcp residues. Cld are produced from cellulose by extracellular endo-β-1,4-glucanases/cellulases from various GH families. The host does not metabolize Cld. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Cld utilizers

Bgl	Beta-glucose oligosaccharide utilization	di_and_oligosaccharides	disaccharides;oligosaccharides	plant	Beta-glucose oligosaccharides (Bgl) are oligomers (DP 2-8) composed of 1,2/3/6-linked β-D-Glcp residues. Bgl are produced from various beta-glucans of plant, microbial, and fungal origin by extracellular endo-β-1,4-glucanases from various GH families. The host does not metabolize Bgl. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Bgl utilizers
BGL12	1,2-beta-oligoglucan utilization	polysaccharides	homopolysaccharides	plant	1,2-β-oligoglucan (BGL12) is an oligomer (DP 17–24) composed of 1,2-linked β-D-Glcp residues. BGL12 in cyclic form is secreted by some bacteria, e.g., Rhizobium phaseoli and Brucella. The host does not metabolize BGL12. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed BGL12 utilizers
mBGL	Mixed beta-glucose-containing glucan utilization	polysaccharides	heteropolysaccharides	plant	Mixed β-glucose-containing glucan (mBGL) is a polysaccharide composed of b-D-Glcp and potentially b-D-Galp residues. The existence of this glycan (potentially of plant or bacterial origin) is predicted based on the presence of a gene cluster (termed bgl cluster in Barratt et al., 2022) that would encode the catabolic machinery, namely multiple glycoside hydrolases: endo-β-glucanase (GH30_3), β-glucosidase (GH3), and a GH2-family enzyme. Strains harboring the bgl gene cluster are deemed mBGL utilizers
Mel	Melibiose utilization	di_and_oligosaccharides	disaccharides	plant	Melibiose (Mel) is a disaccharide with the structure Gal(a1-6)Glc. Mel is abundant in soybeans and other legumes and seeds. The host does not metabolize Mel. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Mel utilizers

RFO	Raffinose family oligosaccharide utilization	di_and_oligosaccharides	oligosaccharides	plant	Raffinose family oligosaccharide (RFOs) is a group of oligomers including raffinose, stachyose, and verbascose with structures Gal(a1-6)Glc(a1-2b)Fruf, Gal(a1-6)Gal(a1-6)Glc(a1-2b)Fruf, and Gal(a1-6)Gal(a1-6)Gal(a1-6)Glc(a1-2b)Fruf, respectively. RFOs are abundant in soybeans and other legumes and seeds. The host does not metabolize RFOs. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed RFO utilizers
Lac	Lactose utilization	di_and_oligosaccharides	disaccharides	animal	Lactose (Lac) is a disaccharide with the structure Gal(b1-4)Glc. Lac is the most abundant carbohydrate in the milk of humans and other mammals. The host utilizes Lac. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Lac utilizers
GOS	Galactooligosaccharide utilization	di_and_oligosaccharides	disaccharides;oligosaccharides	universal	Galactooligosaccharides (GOS) are oligomers (DP 2-8) composed of 1,3/4/6-linked β-D-Galp. GOS are produced from type I galactans by extracellular endo-β-galactanases (GH53) and from type II arabinogalactans by extracellular exo-β-1,3-galactanases (GH43_24) and exo-β-galactobiohydrolases (GH30_5). GOS are also found in small quantities in mammalian milk and can be synthesized enzymatically (this subtype of GOS usually harbors a D-Glcp residue at the reducing end). The host does not metabolize GOS. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed GOS utilizers
AGI	Type I galactan and arabinogalactan	polysaccharides	homopolysaccharides;heteropolysaccharides	plant	Type I galactans and arabinogalactans (AGI) are members of a group of plant polysaccharides called hemicelluloses. For example, potato galactan has a linear backbone of 1,4-linked β-D-Galp residues,

	degradation				some of which are substituted at C(O)3 with α-L-Araf residues. The host does not metabolize AGI. Strains harboring genes encoding extracellular endo-β-1,3/4-galactanases (GH53) AND a complete set of genes encoding transporters and catabolic pathways for the released galactooligosaccharides are deemed AGI degraders
AGII	Type II arabinogalactan degradation	polysaccharides	heteropolysaccharides	plant	Type II arabinogalactans (AGII) are plant polysaccharides modifying arabinogalactan proteins widely distributed in plant species. For example, larch AGII has a branched structure comprising a 1,3-linked β-D-Galp chains backbone and 1,6-linked β-D-Galp side chains. Side chains in larch AGII are modified by α-L-Araf and α-L-Arap residues. The host does not metabolize AGII. Strains harboring genes encoding specific extracellular glycoside hydrolases (including exo-β-1,3-galactanases (GH43_24) and exo-β-galactobiohydrolases (GH30_5)) AND a complete set of genes encoding transporters and catabolic pathways for the released galactooligosaccharides are deemed AGII degraders
GalAf	Galactosylarabinose utilization	di_and_oligosaccharides	disaccharides	plant	Galactosylarabinose (GalAf) is a disaccharide with the structure Gal(a1-3)Araf. GalAf is a structural component of gum arabic arabinogalactan protein, from which it can be released by 3-O-alpha-D-galactosyl-alpha-L-arabinofuranosidases (GH39). The host does not metabolize GalAf. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed GalAf utilizers
GA	Gum arabic	polysaccharides	heteropolysacchari	plant	Gum arabic (GA) is a subtype of polysaccharides modifying arabinogalactan

	degradation		des		proteins characterized by heavy modification of side chains by α-L-Araf, α-L-Arap, α-L-Rhap, α-L-Fucp, GlcA, and 4-O-methyl-GlcA residues. The host does not metabolize GA. Strains harboring genes encoding a specific set of extracellular glycoside hydrolases (including key α-1,3/4-L-arabinofuranosidase (GH43_22_34)) that remove side-chain decorations AND a complete set of genes encoding transporters and catabolic pathways for the released galactooligosaccharides are deemed GA degraders
Fru	D-fructose utilization	monosaccharides_and_derivatives	monosaccharides	plant	D-Fructose (Fru) is a ketohexose. Scr is a building block of many plant di-, oligo-, and polysaccharides, including sucrose, raffinose, fructooligosaccharides, inulin, and levan. Scr enriched in fruits, vegetables, honey, and common sweeteners. The host inefficiently absorbs and metabolizes Fru. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Fru utilizers
Scr	Sucrose utilization	di_and_oligosaccharides	disaccharides	plant	Sucrose (Scr) is a disaccharide with the structure Fruf(b2-1a)Glc. Scr is abundant in fruits, vegetables, and nuts. The host metabolizes Scr. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Scr utilizers
scFOS	Short-chain fructooligosaccharide utilization	di_and_oligosaccharides	oligosaccharides	plant	Short-chain fructooligosaccharides (scFOS) are a group of oligosaccharides with low DP (3-5), including 1-kestose and nystose with structures Fruf(b2-1)Fruf(b2-1a)Glc and Fruf(b2-1)Fruf(b2-1)Fruf(b2-1a)Glc, respectively. scFOS are produced from inulin by enzymatic hydrolysis or from sucrose by enzymatic synthesis (transglycosylation). The host does not metabolize scFOS. Strains harboring a complete set of genes encoding

					transporters and catabolic pathways are deemed scFOS utilizers
lcFOS	Long-chain fructooligosaccharide utilization	di_and_oligosaccharides	oligosaccharides	plant	Long-chain fructooligosaccharides (lcFOS) are oligomers (DP 6-20) composed of 2,1- or 2,6-linked β-D-Fruf residues. 2,1-linked lcFOS constitute a low-DP fraction of inulin from chicory root (although most inulin chains have larger DP). In addition, lcFOS can be produced from inulin (2,1-) and levan (2,6-) by extracellular inulinases and levanases (GH32 and GH68). The host does not metabolize lcFOS. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed lcFOS utilizers
DFA	Difructose dianhydride utilization (production)	di_and_oligosaccharides	disaccharides	plant	Difructose anhydrides (DFAs) are cyclic disaccharides produced by the condensation of two D-Fruf molecules. Example: DFA I with the structure α-D-Fruf-1,2′:2,1′-β-D-Fruf. DFAs can be generated by thermal, acidic, or enzymatic treatment of inulin-, sucrose-, or fructose-rich substrates. The host does not metabolize DFAs. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed DFA utilizers
bMnOS	Beta-mannose oligosaccharide utilization	di_and_oligosaccharides	disaccharides;oligosaccharides	plant	Beta-mannose oligosaccharides (bMnOS) are oligomers (DP 2-8) composed of 1,4-linked β-Manp residues. Example: mannotriose Man(b1-4)-Man(b1-4)b-Man. bMnOS are produced from β-mannan and glucomannan by extracellular endo-β-1,4-mannanases (GH5 and GH26). The host does not metabolize bMnOS. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed bMnOS utilizers
GmOS	Glucomannan oligosacch	di_and_oligosaccharides	disaccharides;oligosacchari	plant	Glucomannan oligosaccharides (GmOS) are oligomers (DP 2-8) are composed of 1,4-linked β-Manp and β-Glcp residues.

	aride utilization		des		Example: mannosyl-glucose with the structure Man(b1-4)b-Glc. GmOS are produced from glucomannan by extracellular endo-β-1,4-mannanases (GH5_8 and GH26). The host does not metabolize GmOS. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed GmOS utilizers
bMAN	Beta-mannan family polysaccharide degradation	polysaccharides	homopolysaccharides;heteropolysaccharides	plant	Beta-mannan family polysaccharides (bMANs) are members of a group of plant polysaccharides called hemicelluloses. bMANs have a linear backbone of 1,4-linked β-D-Manp residues that can alternate with β-D-Glcp residues (glucomannan). Some of the β-D-Manp residues can be substituted at C(O)6 with side chains starting with α-D-Galp residues (galactomannan). The host does not metabolize bMANs. Strains harboring genes encoding extracellular endo-β-1,4-mannanases (GH5_8 and GH26) AND a complete set of genes encoding transporters and catabolic pathways for the released beta-mannose oligosaccharides are deemed (bMAN) degraders
Xyl	D-xylose utilization	monosaccharides_and_derivatives	monosaccharides	universal	D-xylose (Xyl) is an aldopentose. Xyl is a building block of various plant polysaccharides (arabinoxylan, xylan, xyloglucan, etc.). Xyl can be absorbed by the host and incorporated into proteoglycans, where it functions as a linker between the core protein and glycosaminoglycan chains. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Xyl utilizers
XOS	Xylooligosaccharide utilization	di_and_oligosaccharides	disaccharides;oligosaccharides	plant	Xylooligosaccharides (XOS) are oligomers (DP 2-8) composed of 1,4-linked β-D-Xylp residues. Example: xylotriose with the structure Xyl(b1-4)Xyl(b1-4)Xyl. XOS are produced from xylans and arabinoxylans by

					extracellular glycoside hydrolases (e.g., GH10-family endo-β-1,4-xylanases). The host does not metabolize XOS. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed XOS utilizers
AXOS	Arabinoxylooligosaccharide utilization	di_and_oligosaccharides	oligosaccharides	plant	Arabinoxylooligosaccharides (AXOS) are oligomers (DP 2-8) composed of 1,4-linked β-D-Xylp residues, some of which are mono- or disubstituted at C(O)2 or C(O)3 with α-L-Araf residues. Example: A3XXX with the structure Araf(a1-3)Xyl(b1-4)Xyl(b1-4)b-Xyl. AXOS are produced from arabinoxylans by extracellular glycoside hydrolases (e.g., GH10-family endo-β-1,4-xylanases). The host does not metabolize AXOS. Strains harboring (i) a complete set of genes encoding transporters and catabolic pathways OR (ii) genes encoding specific extracellular α-L-arabinofuranosidases (GH43) AND a complete set of genes encoding transporters and catabolic pathways for the released α-L-Araf residues are deemed AXs degraders
XL	Xylan degradation	polysaccharides	homopolysaccharides	plant	Xylans (XLs) are members of a group of plant polysaccharides called hemicelluloses. XLs have a linear backbone of 1,4-linked β-D-Xylp residues. The host does not metabolize XLs. Strains harboring genes encoding endo-β-1,4-xylanases (GH10) AND a complete set of genes encoding transporters and catabolic pathways for the released xylooligosaccharides are deemed XL degraders
AX	Arabinoxylan degradation	polysaccharides	heteropolysaccharides	plant	Arabinoxylans (AXs) are members of a group of plant polysaccharides called hemicelluloses. AXs have a linear backbone of 1,4-linked β-D-Xylp residues, some of which are mono-/disubstituted at C(O)2 or C(O)3 with α-L-Araf residues. AXs are not metabolized

					by the host. Strains harboring genes encoding (i) endo-β-1,4-xylanases (GH10) AND a complete set of genes encoding transporters and catabolic pathways for the released arabinoxylooligosaccharides, OR (ii) genes encoding specific extracellular α-L-arabinofuranosidases (GH43) AND a complete set of genes encoding transporters and catabolic pathways for the released α-L-Araf residues are deemed AXs degraders
Ara	L-arabinose utilization	monosaccharides_and_derivatives	monosaccharides	plant	L-arabinose (Ara) is an aldopentose. Ara is a building block of various plant polysaccharides (arabinoxylan, arabinogalactan, and arabinan, etc.) and glycoproteins. The host does not metabolize Ara. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Ara utilizers
aAOS	Alpha-arabinooligosaccharide utilization	di_and_oligosaccharides	disaccharides;oligosaccharides	plant	Alpha-arabinooligosaccharides (aAOS) are oligomers (DP 2-8) composed of 1,5-linked α-L-Araf residues, some of which can be substituted at C(O)2 or C(O)3 with α-L-Araf residues. Example: arabinotriose with the structure Araf(a1-5)Araf(a1-5)Araf. aAOS are produced from arabinans by extracellular glycoside hydrolases (e.g., endo-α-1,5-L-arabinanases). The host does not metabolize aAOS. Strains harboring (i) a complete set of genes encoding transporters and catabolic pathways (ii) OR genes encoding specific extracellular α-L-arabinofuranosidases (GH43_22) AND a complete set of genes encoding transporters and catabolic pathways for the released α-L-Araf residues are deemed aAOS utilizers
bAOS	Beta-arabinooligosaccharide utilization	di_and_oligosaccharides	disaccharides;oligosaccharides	plant	Beta-arabinooligosaccharides (bAOS) are oligomers (DP 2-3) composed of 1,2-linked β-L-Araf residues. Example: β-arabinobiose with the structure Araf(b1-2)Araf. bAOS are

					produced from plant hydroxyproline‐rich glycoproteins by extracellular β-L-arabinobiosidases (GH121). The host does not metabolize bAOS. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed bAOS utilizers
AR	Arabinan degradation	polysaccharides	homopolysaccharides	plant	Arabinans (AR) are plant polysaccharides often associated with pectins, where they form extensive side chains. AR have a linear backbone of 1,5-linked α-L-Araf residues, some of which are substituted at C(O)2 or C(O)3 with α-L-Araf residues. The host does not metabolize AR. Strains harboring genes encoding specific extracellular α-L-arabinofuranosidases (GH43_22) AND a complete set of genes encoding transporters and catabolic pathways for the released α-L-Araf residues are deemed AR degraders
HRGP	Hydroxyproline-rich glycoprotein oligoarabinofuranoside degradation	di_and_oligosaccharides	oligosaccharides	plant	Hydroxyproline-rich glycoprotein oligoarabinofuranosides (HRGP) are oligosaccharides attached to hydroxyproline residues (Hyp) in Hyp-rich proteins of plant origin by an O-glycosidic bond. Example: Ara3‐Hyp and Ara4‐Hyp with structures Araf(b1-2)Araf(b1-2)ArafbHyp, and Araf(a1-3)Araf(b1-2)Araf(b1-2)ArafbHyp, respectively. The host does not metabolize HRGP. Strains harboring genes encoding specific extracellular glycoside hydrolases (e.g., β-L-arabinobiosidases) AND a complete set of genes encoding transporters and catabolic pathways for the released bAOS are deemed HRGP degraders
XGlOS	Xyloglucan oligosaccharide utilization	di_and_oligosaccharides	disaccharides;oligosaccharides	plant	Xyloglucan oligosaccharides (XGlOS) are oligomers (DP2-8) composed of 1,4-linked β-D-Glcp residues, some of which are substituted at C(O)6 with side chains starting with α-D-Xylp residues. XGlOS are produced

					from xyloglucans by extracellular glycoside hydrolases (e.g., GH5-family endo-β-1,4-glucanases). The host does not metabolize XGlOS. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed XGlOS utilizers
XGL	Xyloglucan degradation	polysaccharides	heteropolysaccharides	plant	Xyloglucans (XGLs) are members of a group of plant polysaccharides called hemicelluloses. XGLs have a linear backbone of 1,4-linked β-D-Glcp residues, some of which are substituted at C(O)6 with side chains starting with α-D-Xylp residues. The host does not metabolize XGL. Strains harboring genes encoding endo-β-1,4-glucanases (GH5) AND a complete set of genes encoding transporters and catabolic pathways for the released xyloglucan oligosaccharides are deemed XGL degraders
GlcA	D-glucuronate utilization	monosaccharides_and_derivatives	uronic_acids	universal	Glucuronate (GlcA) is a hexuronic acid. GlcA is a building block of glycosaminoglycans (e.g., chondroitin and dermatan sulfates), microbial (gellan and xanthan gums) and plant (gum arabic, rhamnogalacturonan I/II) polysaccharides. The host does not metabolize GlcA. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed GlcA utilizers
GluA	Glucuronide utilization	di_and_oligosaccharides	oligosaccharides	plant	Glucuronides (GluA) are oligosaccharides containing glucuronic acid released during the degradation of gum arabic and rhamnogalacturonan I. The host does not metabolize GluA. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed GluA utilizers
Gco	D-gluconate utilization	monosaccharides_and_derivati	sugar_acids	universal	D-Gluconate (Gco) is a hexonate sugar acid. Gco can be found in fruits, rice, meat, dairy products, wine, honey, and vinegar. Gco is

		ves			also used in various foods as coagulants for tofu, agents to adjust the pH of ham, sausage, and noodles, and as acidifiers and calcium fortifiers for beverages. The host does not metabolize Gco. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Gco utilizers
Glt	D-galacto nate utilization	monosacc harides_an d_derivati ves	sugar_aci ds	animal	D-Galactonate (Glt) is a hexonate sugar acid. Glt is an intermediate in D-galactose metabolism in certain bacteria, such as Stenotrophomonas maltophilia. Humans also produce small amounts of D-galactonate, and its levels increase in galactosemic patients. Various reports have demonstrated the presence of D-galactonate in mammalian tissues and body secretions. The host does not metabolize Glt. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Glt utilizers
Ino	Inositol utilization	monosacc harides_an d_derivati ves	sugar_alc ohols	universal	Inositol (Ino) or myo-inositol, is a carbocyclic sugar alcohol. Ino is synthesized by the host from glucose. Ino is abundant in the brain and other mammalian tissues mediating cell signaling and osmoregulation. Ino is also enriched in various foods such as cantaloupe and citrus fruits, bran, and seeds in the form of phytic acid, which is not directly bioavailable to the host. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Ino utilizers
Mtl	D-mannito l utilization	monosacc harides_an d_derivati ves	sugar_alc ohols	plant	Mannitol (Mtl) is a hexose sugar alcohol. Mtl is the most abundant polyol in nature, produced by bacteria, yeasts, fungi, algae, lichens, and many plants. Mtl is used as a sweetener in candies and chewing gum. The host does not metabolize Mtl. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Mtl utilizers

Stl	D-sorbitol utilization	monosaccharides_and_derivatives	sugar_alcohols	plant	Sorbitol (Stl) or glucitol is a hexose sugar alcohol. Stl is used as a sweetener ingredient in manufactured products (including diet drinks and ice cream), mints, cough syrups, and sugar-free chewing gum. Stl is present in fruits, particularly in their dried forms, such as prunes and dried pears. Stl is inefficiently (slowly) metabolized by the host. Strains harboring a complete set of genes encoding transporters and catabolic pathways are deemed Stl utilizers

Chapter two: Codiversification of gut microbiota with humans

(Manuscript published in Science, September 16 2022)

Taichi A. Suzuki[1†], J. Liam Fitzstevens[1†], Victor T. Schmidt[1], Hagay Enav[1], Kelsey E. Huus[1], Mirabeau Mbong Ngwese[1], Anne Grießhammer[2], Anne Pfleiderer[3], Bayode R. Adegbite[3,4], Jeannot F. Zinsou[3,4], Meral Esen[3,5,6], Thirumalaisamy P. Velavan[3,7], Ayola A. Adegnika[3,4,5,8], Le Huu Song[7,9], Timothy D. Spector[10], Amanda L. Muehlbauer[11], Nina Marchi[12], Hyena Kang[13], Lisa Maier[2,6], Ran Blekhman[14], Laure Ségurel[12,15], GwangPyo Ko[13], Nicholas D. Youngblut[1], Peter Kremsner[3,4,5,6], Ruth E. Ley[1,6*]

[*] Corresponding author. Email: rley@tuebingen.mpg.de

[†] These authors contributed equally to this work.

[1] Department of Microbiome Science, Max Planck Institute for Biology, Tübingen, Germany.

[2] Interfaculty Institute of Microbiology and Infection Medicine, University of Tübingen, Tübingen, Germany.

[3] Institute for Tropical Medicine, University of Tübingen, Tübingen, Germany.

[4] Centre de Recherches Médicales de Lambaréné, Lambaréné, Gabon.

[5] German Center for Infection Research, Tübingen, Germany.

[6] Cluster of Excellence EXC 2124 Controlling Microbes to Fight Infections, University of Tübingen, Tübingen, Germany.

[7] Vietnamese German Center for Medical Research, Hanoi, Vietnam.

[8] Fondation pour la Recherche Scientifique, Cotonou, Bénin.

[9] 108 Military Central Hospital, Hanoi, Vietnam.

[10] Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.

[11] Department of Ecology, Evolution, and Behavior, University of Minnesota, Minneapolis, MN, USA.

[12] Eco-anthropologie, Muséum National d'Histoire Naturelle, CNRS, Université de Paris, Paris, France.

[13] Department of Environmental Health Sciences, Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea.

[14] Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN, USA.

[15] Laboratoire de Biométrie et Biologie Evolutive, CNRS, Université Lyon 1, Villeurbanne, France.

Abstract

The gut microbiomes of human populations worldwide have many core microbial species in common. However, within a species, some strains can show remarkable population specificity. The question is whether such specificity arises from a shared evolutionary history (codiversification) between humans and their microbes. To test for codiversification of host and microbiota, we analyzed paired gut metagenomes and human genomes for 1225 individuals in Europe, Asia, and Africa, including mothers and their children. Between and within countries, a parallel evolutionary history was evident for humans and their gut microbes. Moreover, species displaying the strongest codiversification independently evolved traits characteristic of host dependency, including reduced genomes and oxygen and temperature sensitivity. These findings all point to the importance of understanding the potential role of population-specific microbial strains in microbiome-mediated disease phenotypes.

Main text

Across populations, humans share many of the same bacterial and archaeal species in their gut microbiomes (1–3). Within these cosmopolitan species, different strains can dominate in different populations (4–7). Strain variation can arise in several ways, from the uptake of new strains to their in situ evolution (8). When strains and their hosts evolve in parallel, they codiversify, and as a result, their phylogenies are congruent. Codiversification provides opportunities to develop intimate host-microbial relationships across multiple generations (9).

Previous work showed that a small subset of gut bacterial lineages speciated with hominid ancestors (10), but whether such patterns of codiversification extended within host species, and specifically within humans, remained to be demonstrated. There are reasons not to expect to see codiversification with humans: Our diets have changed with time, our populations have expanded across the world, and modern lifestyles may have blurred any signals (11). The identification of species that codiversified with humans has important implications for understanding how humans evolved with their microbiomes and how strains within species may interact with specific host populations (12).

Several human gut microbes are thought to have followed patterns of human migration out of Africa. A notable example is the stomach-dwelling bacterium *Helicobacter pylori*, the causative agent of gastritis and stomach cancer. Cultured isolates of H. pylori show spatial patterns of strain diversity consistent with human migration patterns (13). A few prevalent gut microbial species, including *Prevotella copri*, are also thought to have tracked human migration, given how patterns of metagenome-derived strain variation mapped onto continents (4–6). Strain distributions that map onto human migration patterns are suggestive of codiversification, as geographic origins tend to reflect human genetic origins, especially at the continental level (14, 15). But on finer geographic and population scales, such as within countries, the assumption that geography can stand in for genotype weakens (15). A host phylogeny is required to directly test for codiversification by comparison to microbial phylogenies. Such comparative phylogenetic analyses would also allow microbial taxa to be ranked by the degree of cophylogeny they display.

Given the paucity in the public domain of matched human genotype and gut metagenome datasets required for testing for codiversification, especially for undersampled regions (16), we generated new paired datasets from individuals that we sampled in Gabon, Vietnam, and

Germany. We also leveraged existing datasets for subjects from Cameroon, the Republic of Korea (South Korea), and the UK by generating fecal metagenomes and/or host genotype data (17–20) (Fig. 1A and table S1). In addition, we collected fecal metagenomes from children whose mothers were study participants in Gabon, Vietnam, and Germany (Fig. 1A and table S2). Altogether, our combined dataset of 839 adults and 386 children allowed us to assess codiversification between humans and gut microbial species shared across and within populations.

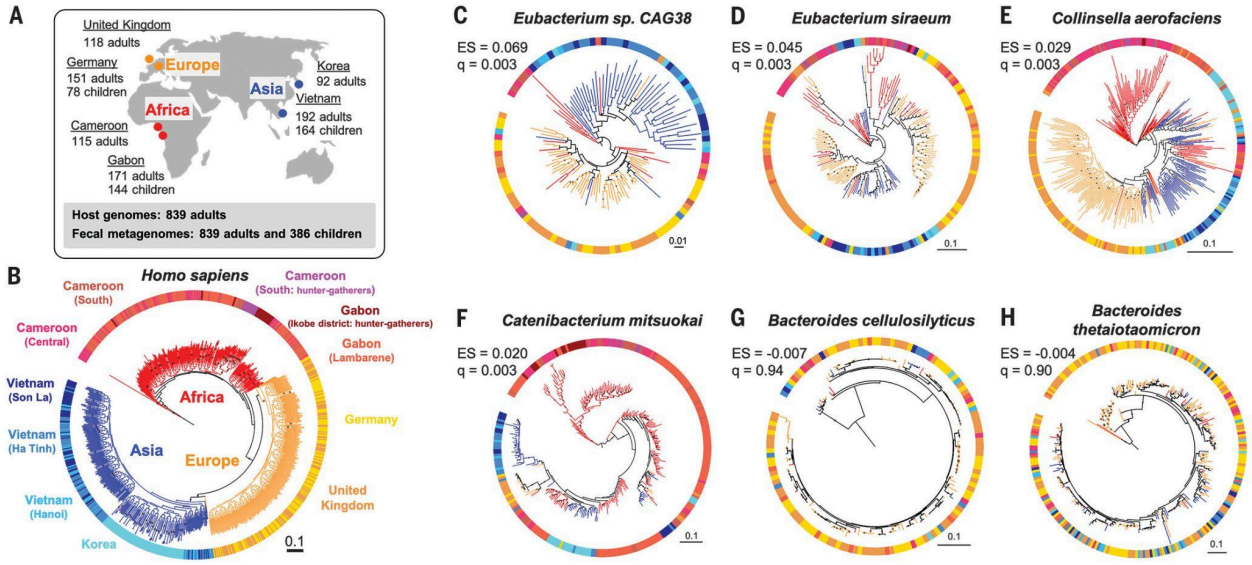

Fig. 1. The human phylogeny and selected bacterial phylogenies. (A) Sampling locations and sizes. (B) A maximum likelihood phylogeny of human subjects based on 20,506 single-nucleotide polymorphisms. Tree branch colors indicate continental origins. Outer strip colors indicate finer geographic locations, and labels refer to sampling locations. (C to H) Maximum likelihood phylogenies for six bacterial species based on species-specific marker genes. PACo effect size (ES) and *q* values (q) are shown. Bootstrap values >50% are plotted on branches, and all phylogenies are rooted at the midpoint. Colors of branches and outer strips correspond to sampling locations shown in (B). The scale bars show substitutions per site for all phylogenies.

Using 20,506 single-nucleotide polymorphisms, we created a maximum likelihood phylogeny to represent the genetic relatedness of the human subjects. As expected, humans clustered into three robust major groups matching their geographic origins (21), where individuals from Asia and Europe formed sister clades nested within individuals from Africa (Fig. 1B). We selected bacterial and archaeal species present in the guts of ≥100 adults with ≥10 individuals per major human group and ≥1 individual per country (see methods in the supplementary materials and table S3). We then created phylogenies for the resulting 59 species using two methods: (i) species-specific marker genes with StrainPhlan3 (7) and (ii) metagenome-assembled genomes (MAGs) using PhyloPhlAn (22). MAG-based trees were obtained for 33 of 59 taxa in adults.

Among the 59 taxa assessed for codiversification, 36 taxa have phylogenies that are more similar to the human host phylogeny than to a permuted host phylogeny [$q < 0.05$, PACo positive effect size (ES); see methods]. *Eubacterium* species showed the largest ES ($q < 0.003$; Fig. 1, C and D, and table S4). Similar results were obtained using two other methods, Parafit (23) and Phytools (24) (tables S4 and S5). Seven species that showed significant codiversification across all three tests included *Collinsella aerofaciens* (Fig. 1E), *Catenibacterium mitsuokai* (Fig. 1F), *Eubacterium rectale*, and *P. copri* (table S4). In contrast, *Bacteroides*, *Alistipes*, and *Parabacteroides* species generally showed the least evidence of cophylogeny (Fig. 1, G and H, and table S4). Results were robust to sample size (tables S6 and S7 and fig. S1) and to bootstrap support for 36 to 50% of taxa (tables S4 and S8). Overall, species within the *Firmicutes* phylum showed more evidence of cophylogeny (Wilcoxon rank sum test, $P = 7.6 \times 10{-}6$) than others, and *Bacteroidetes* species showed least (Wilcoxon rank sum test, $P=1.5\times10{-}6$).

We also searched for a codiversification signal within each country. For a subset of taxa, we observed codiversification in multiple countries independently (table S9). Within-country tests included fewer individuals, so the codiversification signal tended to be weaker. Overall, 20 of 59 taxa showed positive ES with uncorrected $P < 0.05$ in at least one country, but only one taxon remained significant after false discovery rate (FDR) correction: P. copri within Gabon ($q = 0.042$; table S9). Notably, three species (*P. copri*, *Coprococcus eutactus*, and *E. rectale*) had uncorrected $P < 0.05$ in three countries independently (table S9). These within-country results indicate that codiversification is robust to hosts living in a shared environment and suggest that codiversification is not driven by continental-scale processes alone.

We sampled the gut metagenomes of children (average age = 7.4 months) of the genotyped participants in Gabon, Vietnam, and Germany. Using the mothers' genotypes allowed us an unprecedented opportunity to test for codiversification of the child gut microbiome. Among the 20 most prevalent child taxa tested (table S10), nine showed evidence of codiversification ($q < 0.05$) (table S11). All four *Bifidobacterium* species tested showed significant PACo ES ($q < 0.05$) (Fig. 2, A to D, and table S11). According to MAG-based phylogenies, *Bifidobacterium longum* showed the strongest evidence of codiversification in children (table S12). The signals of codiversification for several taxa also extended within countries in children in Gabon and Germany, but not in Vietnam (table S11). These results show that microbes common to the gut in early childhood have also codiversified with humans.

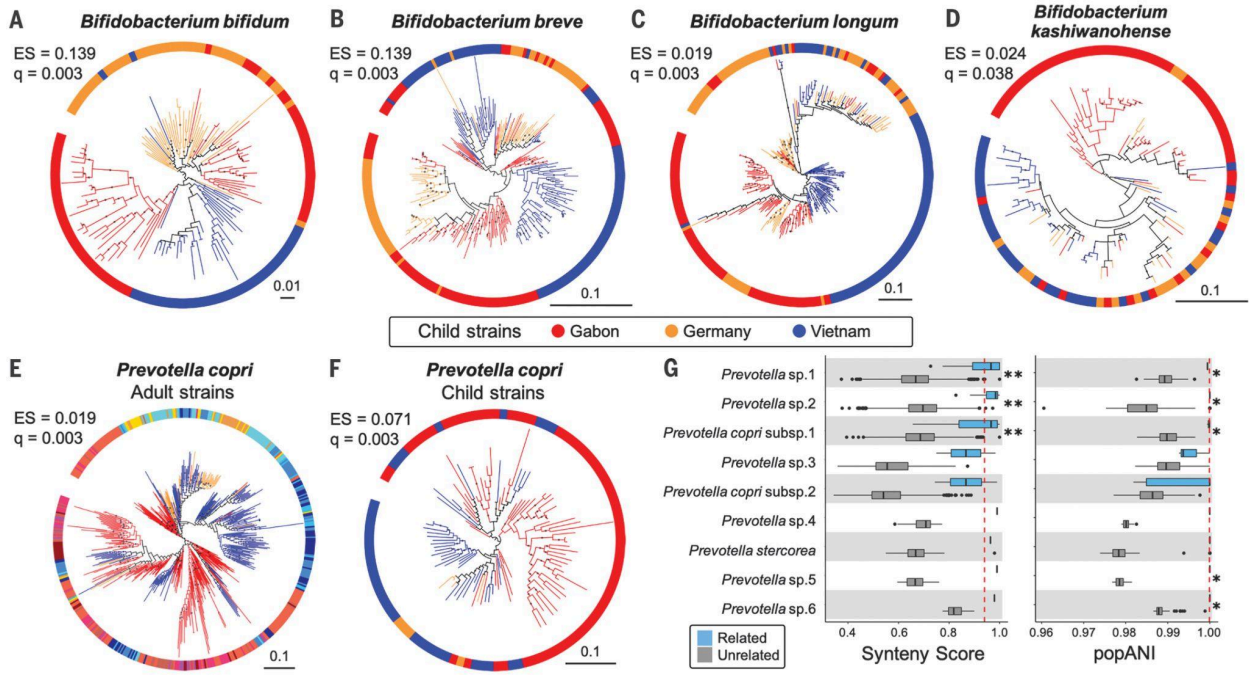

Fig. 2. Bacterial phylogenies derived from children's microbiomes and strain sharing with their mothers. (**A** to **D**) Four species of Bifidobacterium show evidence of cophylogeny based on mothers' genotypes. (**E**) Phylogeny of P. copri strains in adults and (**F**) in children. The colors in (E) correspond to those in Fig. 1B. Bootstrap values ≥50% are shown as black dots on branches, and the phylogenies are rooted at the midpoint. The scales show substitutions per site. (**G**) Prevotella strain sharing between mothers and their own children ("related," blue boxplots) compared with sharing between women and unrelated children ("unrelated," gray boxplots).

(Left) Strain comparisons using SynTracker (39). (Right) Strain comparisons using inStrain (40). Dashed red lines indicate the thresholds for strain sharing events [0.96 for synteny; 0.99999 for popANI (population-level average nucleotide identity)]. *$P < 0.05$ and **$P < 5 \times 10{-}5$ using Wilcoxon-Mann-Whitney test. Codiversification test results for all 20 common child taxa are reported in table S11.

There is little overlap in species composition between adult and child microbiomes; nevertheless, of the overlapping 12 species detected, *P. copri* (Fig. 2, E and F) and *Blautia wexlerae* showed evidence of codiversification in both adults ($q < 0.01$; table S4) and children ($q < 0.01$; table S11). In addition, we observed that mothers and their children share the same strains of *P. copri* (Fig. 2G and fig. S2). For mother-child pairs, strain sharing is often interpreted as vertical transmission, but acquisition of strains from a shared environment cannot be excluded (8). Indeed, our data also support strain sharing between community members: Within sampling locations in Gabon and Vietnam, we observed instances of the same strains in the microbiomes of mothers and unrelated children (fig. S2 and tables S13 and S14). Strain sharing is known among families and socially engaging individuals in the human species (25) and other social animal species (26, 27). Although vertical transmission from parents to offspring over long time periods can result in patterns of codiversification, strain transmission between related individuals in the same communities may also contribute to these patterns (28).

Modern humans emerged in Africa before colonizing the rest of the world (29). Microbial species that migrated with their human hosts may also show signatures of out-of-Africa patterns, and, indeed, some bacterial species exhibit such patterns (4–6). To test for an African origin for the species tested here for cophylogeny, we quantified the number and direction of strain transfer events by applying stochastic character mapping. Consistent with out-of-Africa migration events, when the 10 most highly and 10 least highly ranked taxa (by PACo ES) were compared, the top 10 had significantly greater proportions of transfer events from Africa to the rest of the regions compared with the bottom 10 (Wilcoxon rank sum test, $P = 0.029$) (fig. S3). Because this analysis does not require host genotype data, we added data from 1219 public fecal metagenomes derived from other human populations and from wild primates (tables S15 and S16). Trends were the same with the expanded dataset (Wilcoxon rank sum test, $P = 0.089$) (Fig. 3, A to D, and fig. S4). Additionally, the top taxa also showed significantly more transfer events

from Asia to other regions (Fig. 3, A, E, and F) and fewer transfer events from America to other regions (Fig. 3A). Our results underscore the fact that each species has its own story. Caveats include inaccurate assumptions of host genetic origin based on sampling locations, or a complex history of strain transmission events among different human populations. As expected from codiversification patterns observed in some bacterial families with hominids (10), for the top taxa, primate strains tend to be basal in relation to all human strains (Fig. 3, D and F). In contrast, taxa with least evidence of cophylogeny had primate strains nested within human strains (fig. S5).

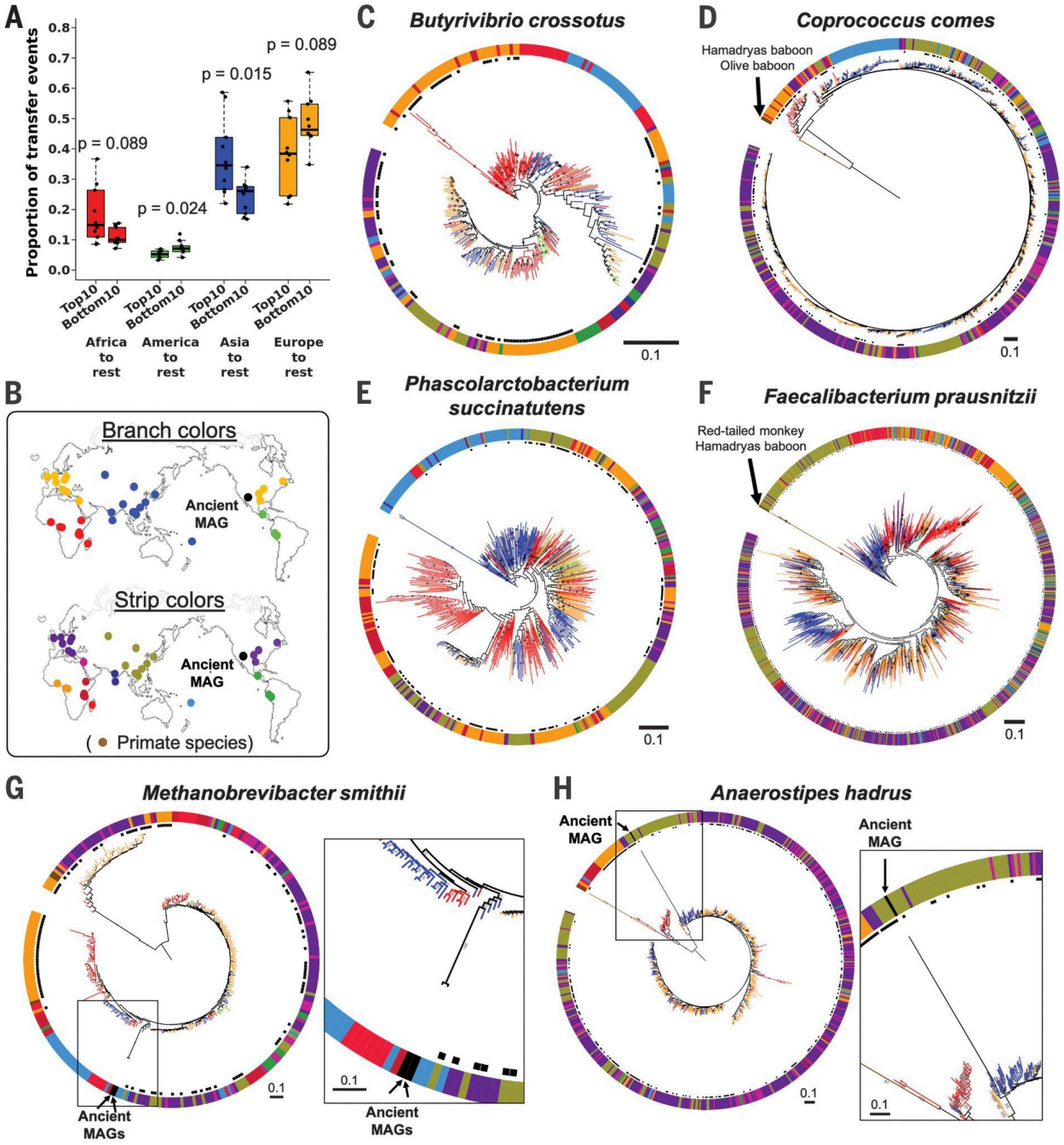

Fig. 3. Strain transfer events and microbial phylogenies including data from public metagenomes. (**A**) Results from stochastic character mapping on microbial phylogenies including six countries from this study and public metagenomes. The boxplots compare the occurrence of transfer events between sampling regions between the top 10 and bottom 10 taxa identified by PACo ES. P values are based on the Wilcoxon rank sum test. (**B**) Sampling locations and color keys correspond to the panels that follow. The colors of the branches and

outer color strip indicate the estimated host genetic structure based on sampling locations (21). Black dots next to the color strip indicate samples from the original six countries. Example phylogenies: (**C**) *Butyrivibrio crossotus*, where African strains are basal; (**D**) Coprococcus comes, where primate strains are basal, followed by African strains; (**E**) *Phascolarctobacterium succinatutens*, where Asian strains are basal; (**F**) *Faecalibacterium prausnitzii*, where strains from primates are basal, followed by strains from Asia. (**G** and **H**) Examples of microbial phylogenies with ancient MAGs recovered from paleofeces of Native Americans. Bootstrap values ≥50% are shown on branches. All trees were rooted at the midpoint. The scale bars show substitutions per site.

Our inference of strain transfer events is based on present-day strains, yet ancient DNA analysis can provide a snapshot directly from the past. We added high-quality ancient MAGs recovered from 1000- to 2000-year-old paleofeces of Native North American tribes (30) to the phylogenies of five gut microbial species (Fig. 3, G and H, and fig. S6). Consistent with the known migration history of the Americas (31), the ancient MAGs were most closely related to strains from modern East Asians, with high bootstrap values for species with significant codiversification (*Methanobrevibacter smithii* and *Anaerostipes hadrus*; Fig. 3, G and H); this was not the case for taxa that did not show significant codiversification (fig. S6).

We hypothesized that species that codiversified with their hosts are better adapted to the host environment than those that did not. Therefore, we predicted codiversified species (high PACo ES) to be enriched in features characteristic of host adaptation, including genome reduction, enrichment in AT content and essential functions, and depletion of non-essential functions (32). To test for these traits genomically, we collected publicly available genome sequences for the 59 species (table S17). As expected, the degree of codiversification was inversely correlated to genome size (Fig. 4A) and positively correlated with genomic AT content (fig. S7A). The relationship with genome size, but not with AT content, remained significant after correcting for phylogenetic relatedness (see methods), indicating that genome size reduction arose independently in codiversified taxa.

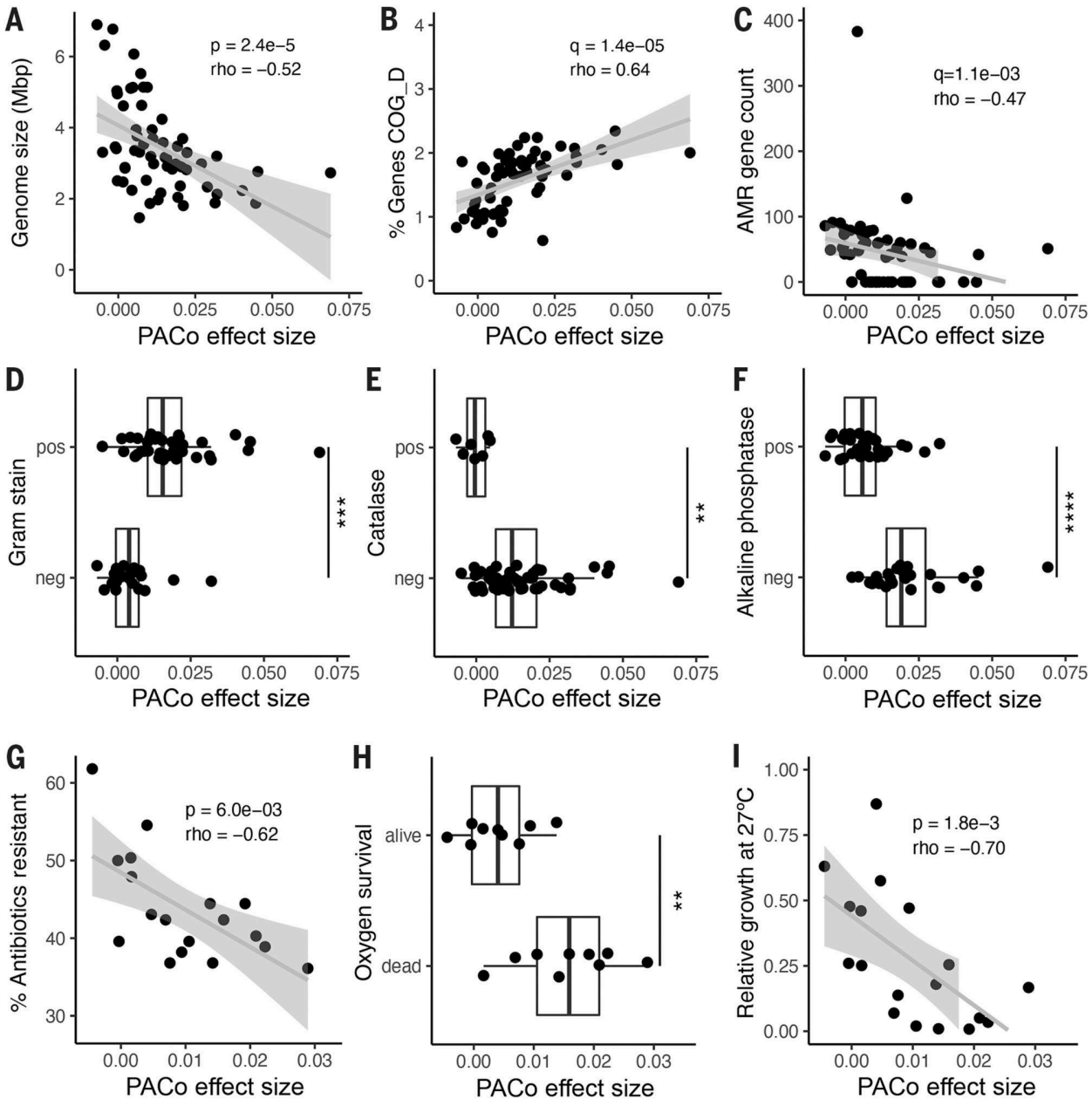

Fig. 4. Genomic and functional features correlated with codiversification. PACo effect size correlated with: (**A**) Median genome size per species. (**B**) Percentage of total genes in the genome annotated to COG D for cell cycle and replication. (**C**) Number of antimicrobial resistance (AMR) markers annotated per genome. (**D**) Predicted Gram stain per species. (**E**) Predicted catalase activity per species. (**F**) Predicted alkaline phosphatase activity. (**G**) Percentage of antibiotics to which the species was resistant in vitro in a panel of 144 common antimicrobials. (**H**) Survival in vitro after 48 hours of O2 exposure for a subset of culturable

species. (**I**) Relative growth of each species in vitro at 27°C compared with 37°C. (A) to (F): n = 59 species; (G) to (I): n = 18 species. Statistical significance was determined by Spearman's correlation [(A) to (C), (G), and (I)] or by Wilcoxon rank sum test [(D) to (F) and (H)]. Exploratory analyses were corrected using FDR across all gene categories and predicted traits [*q* value; (B) to (F)]. pos, positive for the given trait; neg, negative for the given trait. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. Exact *P* values are reported in table S18.

To further explore the genomic signatures of codiversification, we tested for differences in 67 genomic features, including 23 functional categories [clusters of orthologous groups (COGs)], in addition to pseudogenes, antibiotic resistance markers, plasmid markers, and 41 traits predicted from genomic content (see methods and table S18). Overall, ES correlated with 24 of 67 genomic features (FDR-adjusted $P < 0.05$), and five retained significance after correction for phylogenetic relatedness (table S18). A random forest model including these genomic characteristics accurately predicted ES (PACo $q < 0.01$), with a mean area under the curve of 0.83 ± 0.22 (SD) across a fivefold cross-validation.

PACo ES was significantly correlated with the proportion of the genome dedicated to essential functions such as replication, transcription, and translation (Fig. 4B, fig. S7B, and table S18). In contrast, greater ES was associated with fewer pseudogenes (fig. S7C) and antibiotic resistance markers Fig. 4C, and a smaller proportion of the genome dedicated for nonessential functions, such as secretion and cell wall biogenesis (fig. S7D and table S18). Gram-positive species were enriched for higher ES overall (Fig. 4D). A number of predicted traits related to environmental survival, including oxygen sensitivity (Fig. 4E), inorganic phosphate scavenging (Fig. 4F), and the use of diverse energy sources (fig. S7, E to I, and table S18), were reduced on average in species with high ES.

To directly test the functions predicted from genome-based observations, we assessed in vitro phenotypes of a representative set of 18 culturable species (see methods and table S19). Consistent with the reduction in antibiotic resistance markers, codiversified species exhibited significantly reduced antibiotic resistance in a previously published drug screen of 144 antimicrobial compounds (Fig. 4G) (33). Consistent with the predicted loss of catalase activity, species with higher ES were significantly more likely to die upon exposure to atmospheric oxygen (Fig. 4H). Increased temperature sensitivity has been associated with coevolved insect

symbionts and is expected in gut microbes that enjoy a temperature-stable niche (34). Accordingly, we observed that ES was significantly correlated with poor relative growth at below-host temperature (27°C) (Fig. 4I). All in vitro phenotypic associations retained significance after correction for phylogenetic relatedness, even in cases where the corresponding genomic prediction did not (table S18).

Taken together, the features that associate with cophylogeny are highly reminiscent of those commonly seen in host-associated microbes (35–37) and coevolved insect symbionts (32). Patterns of host-microbial codiversification alone do not necessarily imply interactions or adaptations between hosts and microbes (9, 28). However, together with the observed functional attributes, such as smaller genomes and oxygen and temperature sensitivity, codiversified species likely evolved host dependency.

By expanding metagenome collections into poorly characterized populations, and pairing metagenome and host genomic data obtained from the same individuals, we have identified common members of the human gut microbiota that have independently codiversified with human populations. These codiversified species have repeatedly and independently acquired traits that suggest limited survival capabilities outside of the host (2, 28, 35). Loss of an environmental reservoir can facilitate dependence on resources produced by other gut microbes and/or the host and lead to reduced genome size (35–37). The selection pressure on efficient host-to-host transmission could result in strain sharing between related individuals, or those living in proximity, such as we observed in our populations. Many of the traits characteristic of codiversified species likely adapted to the niche of the animal gut (not necessarily human), and whether humans reciprocally adapted to these microbial species or strains remains to be investigated. The list of codiversified species provides a starting point to investigate host-microbial coevolution in humans (12).

The list of human health conditions linked to the microbiome ranges from mulnutrition to allergies and cardiovascular disease. The incidence of these diseases is population specific, and the diversity of microbiomes is also population specific. Several of the species that codiversified with humans, such as *P. copri* (4), *E. rectale* (5), and *B. longum* (38), are known to vary in their functional capacity according to population. An awareness of differences in gut microbial strains between populations has already led to the notion that probiotics for treating malnutrition should be locally sourced (38). The microbiome is a therapeutic target for personalized medicine, and

our results underscore the importance of a population-specific approach to microbiome-based therapies.

References and notes

1. V. K. Gupta, S. Paul, C. Dutta, Front. Microbiol. 8, 1162 (2017).
2. P. I. Costea et al., Mol. Syst. Biol. 13, 960 (2017).
3. E. Pasolli et al., Cell 176, 649–662.e20 (2019).
4. A. Tett et al., Cell Host Microbe 26, 666–679.e7 (2019). 5. N. Karcher et al., Genome Biol. 21, 138 (2020).
6. B. D. Merrill et al., bioRxiv 2022.03.30.486478 [Preprint] (2022). https://doi.org/10.1101/2022.03.30.48647.
7. D. T. Truong, A. Tett, E. Pasolli, C. Huttenhower, N. Segata, Genome Res. 27, 626–638 (2017).
8. H. Enav, F. Bäckhed, R. E. Ley, Cell Host Microbe 30, 627–638 (2022).
9. N. A. Moran, D. B. Sloan, PLOS Biol. 13, e1002311 (2015).
10. A. H. Moeller et al., Science 353, 380–382 (2016).
11. A. H. Nishida, H. Ochman, Nat. Commun. 12, 5632 (2021).
12. T. A. Suzuki, R. E. Ley, Science 370, eaaz6827 (2020).
13. D. Falush et al., Proc. Natl. Acad. Sci. U.S.A. 98, 15056–15061 (2001).
14. J. Novembre et al., Nature 456, 98–101 (2008).
15. N. A. Rosenberg et al., Science 298, 2381–2385 (2002). 16. R. J. Abdill, E. M. Adamowicz, R. Blekhman, PLOS Biol. 20, e3001536 (2022).
17. A. Lokmer et al., PLOS ONE 14, e0211139 (2019).
18. G. Even et al., Front. Cell. Infect. Microbiol. 11, 533528 (2021). 19. M. Y. Lim et al., Gut 66, 1031–1038 (2017).
20. H. Xie et al., Cell Syst. 3, 572–584.e3 (2016).
21. P. Duda, Jan Zrzavý, Sci. Rep. 6, 29890 (2016).
22. F. Asnicar et al., Nat. Commun. 11, 2500 (2020).
23. P. Legendre, Y. Desdevises, E. Bazin, Syst. Biol. 51, 217–234 (2002). 24. L. J. Revell, Methods Ecol. Evol. 3, 217–223 (2012).

25. I. L. Brito et al., Nat. Microbiol. 4, 964–971 (2019).
26. J. Tung et al., eLife 4, e05224 (2015).
27. A. H. A. H. Moeller et al., Sci. Adv. 2, e1500997–e1500997 (2016).
28. M. Groussin, F. Mazel, E. J. Alm, Cell Host Microbe 28, 12–22 (2020).
29. R. Nielsen et al., Nature 541, 302–310 (2017).
30. M. C. Wibowo et al., Nature 594, 234–239 (2021). 31. T. Goebel, M. R. Waters, D. H. O'Rourke, Science 319, 1497–1502 (2008).
32. J. P. McCutcheon, N. A. Moran, Nat. Rev. Microbiol. 10, 13–26 (2011).
33. L. Maier et al., Nature 555, 623–628 (2018).
34. K. E. Huus, R. E. Ley, mSystems 6, e0070721 (2021).
35. H. P. Browne et al., Genome Biol. 22, 204 (2021).
36. S. Nayfach, Z. J. Shi, R. Seshadri, K. S. Pollard, N. C. Kyrpides, Nature 568, 505–510 (2019).
37. S. A. Frese et al., PLOS Genet. 7, e1001314 (2011).
38. M. J. Barratt et al., Sci. Transl. Med. 14, eabk1107 (2022).
39. H. Enav, R. E. Ley, bioRxiv 2021.10.06.463341 [Preprint] (2021). https://doi.org/10.1101/2021.10.06.463341.
40. M. R. Olm et al., Nat. Biotechnol. 39, 727–736 (2021).
41. T. A. Suzuki, L. Fitzstevens, K. Huus, N. Youngblut, leylabmpi/ codiversification: Zenodo release, version 1.0.1, Zenodo (2022); https://doi.org/10.5281/zenodo.6947454.
42. T. A. Suzuki, J. L. Fitzstevens, N. D. Youngblut, R. E. Ley, Phylogenies related to "Codiversification of gut microbiota with humans," Dryad (2022); https://doi.org/ 10.5061/dryad.qrfj6q5k2.

Acknowledgements

We thank T. H. Nguyen, E. Cosgrove, A. Clark, A. Kostic, M. Taylor, Native American tribe officers (M. Bremer, J. Aguilar, J. Charlie, R. Williams, B. Lewis, S. Anton, A. Garcia-Lewis, and B. Bernstein), S. Dauser and members of the Department of Microbiome Science, and four anonymous reviewers. Funding: This work was supported by the Max Planck Society. T.D.S. was funded by the Wellcome Trust, Medical Research Council, European Union, Chronic

Disease Research Foundation, Zoe Global Ltd., the National Institute for Health Research–funded BioResource, and the Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. L.S. was supported by an Agence Nationale de la Recherche grant (MICROREGAL, ANR-15- CE02-0003). R.B. was supported by NIH grant R35-GM128716. Author contributions: Writing – original draft, Conceptualization, Methodology, and Visualization: T.A.S., J.L.F., H.E., K.E.H., N.D.Y., and R.E.L. Software, Validation, and Formal analysis: T.A.S., J.L.F., H.E., K.E.H., and N.D.Y. Investigation: T.A.S., J.L.F., V.T.S., K.E.H., M.M.N., A.G., A.P., B.R.A., J.F.Z., A.L.M., N.M., H.K., L.M., R.B., L.S., G.K., N.D.Y., and R.E.L. Resources: J.L.F., V.T.S., A.P., B.R.A., J.F.Z., M.E., T.P.V., A.A.A., L.H.S., T.D.S., A.L.M., N.M., H.K., L.M., R.B., L.S., G.K., N.D.Y., P.K., and R.E.L. Data curation: T.A.S., J.L.F., V.T.S., K.E.H., M.M.N., A.L.M., N.M., H.K., R.B., L.S., and N.D.Y. Project administration: T.A.S., J.L.F., V.T.S., L.M., R.B., L.S., G.K., and R.E.L. Supervision: L.M., R.B., L.S., G.K., N.D.Y., P.K., and R.E.L. Funding acquisition: T.D.S., L.M., R.B., L.S., G.K., and R.E.L. Competing interests: G.K. is the founder and a board member of KoBioLabs, Inc. T.D.S is a cofounder of ZOE Ltd., a personalized nutrition company. Data and materials availability: All data used in this study are free to access. The raw sequence data and MAGs are available from the European Nucleotide Archive under the study accession numbers PRJEB40256, PRJEB9584, PRJEB32731, PRJEB27005, PRJEB30834, and PRJEB46788. All sample metadata used in this study are provided in the supplementary materials. The code used in data analysis is outlined on Zenodo (41), and phylogenies and alignments are available in Dryad (42).

Supplementary materials

science.org/doi/10.1126/science.abm7759

Tables S1 to S19
References (43–80)
MDAR Reproducibility Checklist

View/request a protocol for this paper from Bio-protocol.

Submitted 12 October 2021; accepted 10 August 2022
10.1126/science.abm7759

Material and Methods

Participant recruitment

Five hundred and fourteen adult women were recruited for the study from three countries: Gabon (n = 171, average age = 27 years old), Vietnam (n = 192, average age = 28 years old), and Germany (n = 151, average age = 29 years old) (Table S1). Inclusion criteria included: women between the ages of 18 and 40; no severe allergy to dairy products; not currently self-reported as pregnant; no known underlying medical conditions; local residency. Participants were screened for the above criteria by field staff and/or physicians upon arrival at the participating clinics, and all protocols, risks, and study motivations were carefully explained to participants in local languages. 386 children of women who participated in the study were included for microbiome analysis: Gabon (n = 144, average age = 8.0 months), Vietnam (n = 164, average age = 6.7 months), and Germany (n = 78, average age = 7.7 months) (see Table S2 for detail). The average ages of the children did not differ significantly between populations (Kruskal-Wallis p = 0.43), and aside from 3 children in Gabon, all were breastfed. All participants provided informed consent for participation in the study for themselves, and as applicable, for their child. All human research was approved by local ethical committees. The study protocols for this study are as follows: Gabon, issuing authority: Comité National d'Ethique, Protocol number: N0025/2017/PR/SG/CNE; Vietnam, issuing authority: Scientific Ethics Review Committee and 108 Military Central Hospital, Protocol number: 108MCH/RES/VGCARE-03-16012018; Germany, issuing authority: Ethik-Kommission an der Medizinischen Fakultät der Eberhard-Karls-Universität und am Universitätsklinikum Tübingen. 529/2018BO1.

Gut metagenome data generation

Participants in Gabon, Vietnam, and Germany used Fe-Col collection kits (Alpha Labs, Eastleigh, UK) to collect stool from themselves and/or their child. Samples were frozen on dry ice within 8 hours of collection and transported on dry ice for storage at -80°C. Sample collection in Korea is described in (*19*). Genomic DNA was extracted from frozen stool using PowerSoil DNA extraction kits (Qiagen, Hilden, Germany). Metagenomic library preparation was performed as per reference (*43*) with slight modifications. Briefly, 1ng of purified gDNA

was used in a Nextera (Illumina, San Diego, USA) Tn5 tagmentation reaction to fragment and ligate adaptors in a single reaction, followed by a 14 cycle PCR to add sample specific barcodes. Libraries were purified using Mag-Bind TotalPure NGS beads (Omega Biotech, Norcross, USA), pooled, and quantified. Size selection (400-700bp) was performed on a BluePippin (Sage Science, Beverly, USA). Libraries were concentrated and further purified as needed using DNA Clean & Concentrator-5 (Zymo Research, Irvine, USA). Sequencing was conducted on a HiSeq 3000 System (Illumina, San Diago, USA) with 150 paired-end sequencing. We used a quality control pipeline described in (*43*). Briefly, adapter trimming and quality control filtering was conducted using Skewer0.2.2 and bbtools "bbduk" command. Reads mapping to the human genome (GRCh37/hg19) were filtered using bbtools "bbmap" command. The read quality was assessed using Fastqc 0.11.7 and multiQC 1.5a. The same pipeline was applied to all publicly available metagenomes that were used in this study, including the previously published metagenomes from Cameroon (PRJEB27005 and PRJEB30834)(*17*, *18*) the UK (PRJEB13747)(*20*), and the 1326 public metagenomes from various studies (see below).

Metagenome-assembled genome (MAG) generation and analysis

Paired-end reads were subsampled to ≤ 20 million per sample with seqtk 1.3. We used metaSPAdes 3.12.0 for per-metagenome *de novo* assemblies. Contig binning with differential coverage was conducted with three binners: MetaBAT 2.15, MaxBin 2.2.7, and VAMB 3.0.2. We ran each binner 3 times with different parameter sets (MetaBAT: [-prob_threshold 0.6, prob_threshold 0.7, -prob_threshold 0.8], MaxBin: [--maxP 92 --maxEdges 150, --maxP 94 -maxEdges 325, --maxP 97 --maxEdges 500], VAMB: [-l 24 -n 384 384, -l 32 -n 512 512, -l 40 -n 768 768]) for a total of 9 binning methods applied to each metagenome assembly. Only contigs ≥1.5 kbp were used for binning. Per-metagenome differential coverage binning with coverage calculated from all metagenomes would require an excessive number of read mapping jobs (*i.e.,* all pairwise combinations of metagenomes), so we utilized a subsampling approach. Specifically, for each metagenome assembly, we mapped reads to the assembled contigs from the target metagenome and reads from 39 other randomly selected metagenomes. Bowtie 2.4.1 was used for mapping reads to contigs. To reduce biases in coverage estimated resulting from varying sampling depths among metagenomes, we subsampled to ≤5 million paired-end reads prior to mapping in order to reduce biases from samples with high sequencing depths (*e.g.,* >20

million). We used DAS-Tool (1.1.2 for the current study and Lokmer et al. (2019) (*17*); 1.1.3 for the Even et al. (2021) (*18*)) to select the highest quality, non-redundant contig bins (MAGs) from all 9 binning methods, with quality based on CheckM 1.1.3-estimated completeness and contamination.

The raw MAGs were filtered by CheckM-estimated completeness (≥ 50%) and contamination (< 5%) and then dereplicated at 99.9 and 95% ANI with dRep (3.2.0 for the current study; 3.2.2. For Lokmer et al. (2019) (*17*) and Even et al. (2021) (*18*)). Prodigal 2.6.3 was used for gene prediction. Multi-locus genome phylogenies were inferred via PhyloPhlAn 3.0.2, with DIAMOND 2.0.9 used for phylogenetic marker homology searches, MAFFT 7.480 used for the multiple sequence alignment, FastTree 2.1.10 used for initial tree inference, and RAxML 8.2.12 used for final tree inference (starting from the phylogeny inferred from FastTree). Outgroups for each clade-level phylogeny were automatically selected from sister clades (*e.g.,* sister genus). Specifically, the GTDB Release 95 genome metadata was used to identify genomes in sister clades, and the outgroups were selected with preference given to genomes with the highest CheckM-estimated completeness and contamination.

Mother-child strain sharing analysis

We used CheckM and dRep to dereplicate 6,234 filtered MAGs assembled from mother-child metagenomes at 95% ANI and to identify 796 species representative genomes (SRGs). Species were assigned to both GTDB and NCBI using gtdb_to_taxdump v0.1.6(*44*). We mapped metagenome reads from 386 mother-child pairs to the SRGs with inStrain 1.5.3's "profile" function (--min_cov 5 --min_freq 0.05 --min_genome_coverage 0.1 --min_read_ani 0.95 -database_mode --skip_plot_generation) and used the "compare" function to calculate popANI and percent_genome_compared (--min_cov 5 --min_freq 0.05 --ani_threshold 0.99999 -database_mode --skip_plot_generation). Pairwise comparisons with popANI ≥99.999% and percent_genome_compared ≥50% were identified as strain sharing events. To test whether there are significant strain sharing events within related mother-child pairs compared to within unrelated mother-child pairs more than expected by chance, we used the hypergeometric distribution R function "phyper", with correction for multiple comparisons using the R function "p.adjust" (--*method* = "BH"; Table S12).

We used SynTracker (*39*) to quantify the relatedness of strains in related and unrelated mother-child pairs using the 6,234 filtered MAGs, with each SRG as a reference genome, with default parameters (--perc_identity = 97, --qcov_hsp_perc=70, --maxSep=15, --maxGap = 15). Average pairwise synteny scores were calculated by randomly selecting 30 homologous 5kbp regions/pairwise. Significance tests for the synteny scores and popANI of strains between related and unrelated mother-child pairs were performed using the R function "wilcox.test" (--*alternative* = less). Multiple testing correction was performed using the "p.adjust" function (-*method* = "BH"). Note that unlike the marker-based phylogenies described below (*i.e.,* studying the dominant strain per individual), results from inStrain and SynTracker represent multiple strains per individual.

Data generation of human genotype data

We generated human genotype data from the participants in four countries (Gabon, Vietnam, and Germany, and Cameroon) and used published genotype data for the participants in Korea (*19*) and the United Kingdom (*20*). In Gabon, Vietnam, and Germany, we collected saliva samples from each participant using Saliva DNA Collection and Preservation Devices (Norgen, Thorold, Canada). Genomic DNA from saliva was extracted using PowerSoil DNA extraction kits (Qiagen, Hilden, Germany). DNA was analyzed using the Infinium Global Screening Array (Illumina, SanDiego, USA) and genotyped at the University Hospital of Bonn, Life & Brain Research Centre. For Cameroon participants, host genotype data was generated for individuals that were collected in 2013 and 2017 with published gut metagenomes (*17*, *18*). For each participant, approximately 2 ml of saliva was collected in a 15 ml falcon tube, to which we added the same volume of a homemade buffer (5mM TRIS, 5mM EDTA, 5mM sucrose, 10mM NaCl and 1% SDS, pH=8). DNA was then extracted following the procedure from (*45*). For samples collected in 2013, the extracted DNA was then genotyped on an Illumina Omni2.5 genotyping array at the University of Chicago Genomics Core, in Chicago, USA. For samples collected in 2017, after quantification using a PicoGreen assay and qPCR assays to check for amplification and clustering, extracted DNA was genotyped on an Illumina Multi-Ethnic Global array at the University of Minnesota Genomics Center.

Using the GenomeStudio software, we excluded individuals with a call rate below 0.95, as well as markers that failed genotyping, present on non-autosomal chromosomes, with a call rate below 95% or a cluster separation below 0.3. For the previously published genotype data from the UK (*20*) and Korea (*19*), we excluded close relatives (*i.e.*, one individual per twin) and individuals that did not have paired fecal metagenomes. Finally, genome assembly versions were converted to GRCh38 using UCSC Genome Browser LiftOver and all datasets were merged in PLINK v1.9. After merging, further quality control was conducted where variants that are nonbiallelic, minor allele frequency of < 0.05, missing call rates of > 0.9, and deviated from HardyWeinberg equilibrium ($P< 10^{-5}$) were removed. This resulted in 20,506 SNPs that overlapped across all six datasets.

Human phylogenetic inference from SNP data

We created 100 maximum likelihood trees in SNPhylo (version 20180901)(*46*) including 839 individuals using 20,506 SNPs, which were further filtered to remove uninformative SNPs based on the following parameters (ld_threshold = 0.1, maf_threshold = 0.05, and missing_rate = 0.1). This resulted in around 9,000 SNPs to create the host phylogeny (each bootstrapped tree uses a slightly different number of SNPs). We picked the best tree out of the 100 trees based on the maximum likelihood score and plotted the bootstrap values (referred as "HostBestTree") using "plotBS" in phangorn R package. To account for branches that have low bootstrap support, a majority-rule consensus tree (*47*) was also created where the best tree with branches with bootstrap values < 50% were collapsed (referred as "HostConsTree") using the function "as.polytomy" in ggtree R package. Although we acknowledge that the bootstrap value 50% is low, the threshold was chosen to allow comparisons with the results obtained from the best maximum likelihood phylogenies with minimal changes in the phylogeny size of host and microbes (see below). The main results are reported using HostBestTree. The host tree was rooted by midpoint. Final trees were annotated in iTOL v6 (*48*).

Microbial phylogenetic inference from StrainPhlAn data

We created microbial phylogenies from adult and child metagenomes using StrainPhlAn v3.0 (*7*). First, we picked the top 100 prevalent taxa in metagenomic samples of adults (n=839) and children (n=386) using MetaPhlAn3 (v3.0.1) using the following parameters (--tax_lev a -min_cu_len 2000). The taxonomy is based on NCBI. Next, we used StrainPhlan3 to generate microbial phylogenies using the following parameters: samples2markers.py (--breadth_threshold 80) and strainphlan.py (--phylophlan_mode accurate --marker_in_n_samples 10 -sample_with_n_markers 10). To allow for between-population comparisons, we selected taxa in adults that represent a total of ≥ 100 individuals, ≥ 10 individuals each from major human grouping; Africa (Cameroon and Gabon), Asia (Korea and Vietnam), and Europe (Germany and UK) and ≥ 1 individual per country, which resulted in 59 taxa (Tables S3&S4). For child samples, we picked the top 20 taxa that had the largest trees due to the smaller sample size (Tables S10&S11). The best tree out of 100 trees based on the maximum likelihood score was selected (referred as "BacBestTree"), and a majority rule consensus tree was created (referred to as "BacConsTree"). We chose StrainPhlan3 for the main method to create strain-level phylogenies because it is suited for creating phylogenetic trees of the most dominant strain per individual. All trees generated by StrainPhlAn3 are rooted by midpoint. All MAG-based trees are rooted by an outgroup. Final trees were annotated in iTOL v6 (*48*).

Public metagenomes and phylogenetic trees

We downloaded and curated public metagenomes accessible at the time of the study (Tables S15&S16). We retained stool samples from subjects with no associated disease status (healthy or control group), no current use of antibiotics, and with an age ranging from 12-65 years old (age mean ± s.d. = 36.7 ± 13.6 years old). All samples that indicated non-local individuals (*e.g.,* travelers) were also excluded. For visualization purposes, the collection locality of the individual was used as a proxy for host genetic groups based on Duda and Zrzavy (2016)(*21*). The assignment of genetic structure and color used in Fig. 3 can be found in Table S15. When samples exceeded 100 individuals per population per study, we randomly selected 100 individuals to reduce the sampling bias. Samples were quality filtered using the same QC pipeline described above. When multiple samples were available per individual, we combined

the reads. Using StrainPhlAn3 with the same parameters described above, we were able to extract marker genes from fecal metagenomes derived from 1219 human individuals representing 19 countries in 21 studies (*2–4*, *20*, *49–65*), and 107 primate individuals representing 27 species of primates from 12 countries in 2 studies (*66*, *67*). The number of sequence reads and detailed sample information for all public metagenomes used in this study are shown in Table S16. Phylogenies for the 10 top-ranking and bottom-ranking taxa were created using StrainPhlAn3 with the same parameters described above. High-quality ancient MAGs from paleofeces described in (*30*) were shared by Alex Kostic with support from the Peabody Museum. Six ancient MAGs that were identified at the species level were added to the microbial phylogenies as reference genomes in StrainPhlAn3 (Fig. S6). All trees were annotated in ITOL v6 (*48*).

Cophylogeny statistics

We mainly report the cophylogeny test results based on principle coordinate analysis and procrustes distance (referred as "PACo") from the function "PACo" in paco R package because it is reported to have greater statistical power and lower type I error rates compared to other methods (*68*). We also used two additional methods to quantify the degree of codiversification between the host tree and the microbial tree: (i) function "parafit" in the ape R package, which is based on principle coordinate analyses and euclidean distance (referred as "Parafit")(*23*), and (ii) the function "cospeciation" in phytools R package, which is based on Robinson-Foulds distance (referred as "Phytools")(*24*). All tests were conducted with 999 permutations. A recent study investigating large host-microbe phylogenies noted that PACo and Parafit tests resulted in significant host-restricted associations ($p < 0.001$) even when random host trees were used (*11*). This observation was true in our study where all taxa resulted in p values < 0.001 when one-toone host-microbe pairs were considered. Thus, for PACo and Parafit, we collapsed highly similar microbial strains as a same strain to create one-to-many host-microbe pairs using the function "cutreeDynamic" in dynamicTreeCut R package with the following parameters: cutHeight = 0.99 and minClusterSize = 2. To represent the degree of host-microbial cophylogeny, we calculated cophylogeny effect size (ES) for each taxon by representing the percent change between the observed phylogenetic distance (obs_dist: distance between microbial tree and host tree) and null phylogenetic distance (null_dist: average permuted

distance between microbial tree and host tree) as the following: (null_dist - obs_dist) / (null_dist). Thus, positive ES indicates evidence of cophylogeny (obs_dist < null_dist) and negative ES indicates no evidence of cophylogeny (obs_dist > null_dist). Unlike Nishida and Ochman (2021) (*11*), none of the taxa with negative ES resulted in p value < 0.05 using our approach by collapsing highly related strains as mentioned above. To account for the uncertainty of phylogenetic topologies, we applied the three codiversification tests on majority rule consensus trees, where branches with bootstrap values < 50% were collapsed (HostConsTree vs BacConsTree), in addition to best trees based on the maximum likelihood score (HostBestTree vs BacBestTree). We acknowledge that collapsing branches with bootstrap values < 50% is low, but collapsing branches with higher bootstrap values will significantly reduce the number of branches in some microbial phylogenies. Thus, to compare the results between BestTree and ConTree in all adult and child taxa with minimal differences in the number of tips, we decided to use majority-rule consensus phylogenies. Overall, the number of strains represented in a microbial tree did not bias the three codiversification test results for both between- and within-country analyses (Table S6&S7). ES and q values from the three codiversification tests were compared using Spearman's rho correlation (Table S5). To test for host-microbial cophylogeny using child taxa, we used the host phylogeny based on the mother's genotype and microbial phylogeny based on the child's metagenome and tested for codiversification using PACo.

Transfer events between Africa, Europe and Asia

To quantify the occurrence and directions of host-switch/transfer events, we used stochastic character mapping based on MCMC (*69*) implemented as SIMMAP in "phytools" R package (*24*). We applied the character mapping on (i) marker-based and MAG-based trees of adults in this study (six countries), (ii) marker-based trees of adults after adding public metagenomes (23 countries), (iii) marker-based trees of adults only using public metagenomes (19 countries), and (iv) marker-based trees of children in this study (three countries). Briefly, we treated "host regions" (*i.e.,* "Africa", "America", "Asia", and "Europe") as microbial traits, inferred ancestral character states on microbial phylogeny using the equal rates model, repeated this 100 times, and calculated the average number of character changes and direction of host transfer events. Because larger trees will give a greater number of character changes on average,

we divided the number of transfer events for each category by the total number of transfer events of that microbial tree. Occurrence of transfer events from Africa to Asia and from Africa to Europe were combined as "Africa to rest". The same aggregation was applied to the two other categories; "Asia to rest" and "Europe to rest". Correlations between occurrence of transfer events and PACo ES were based on Spearman rho correlation. Pairwise comparisons between categories of transfer events were based on Wilcoxon rank sum test. FDR-corrected *p* values are indicated as *q* values.

Genome feature analyses and trait prediction

Selection of genomes.

All genomes that were available from NCBI for each abundant/prevalent species in adults (n = 59) on November 3, 2021 were downloaded via NCBI_genome_download. Genomes were quality filtered (minimum completeness >80 and contamination <5 via CheckM (*70*)) and the top 10 genomes per species were retained for further analysis. If fewer than 10 genomes were available for a species, all genomes passing minimum quality filtering were used. A list of the genome accessions considered per species can be found in Table S17 (n = 495 genomes total).

Trait profiling.

Genome size and GC content were calculated as the median value per species across the top filtered genomes. COG categories were annotated per genome via Prodigal (*71*) and EggNOG-mapper (*72*), and genes mapped to each COG were calculated as the percentage of the total genes annotated per genome. Pseudogenes per genome were annotated using dfast (*73*) and calculated as a percentage of the genes annotated per genome. Antibiotics markers per genome were annotated using ABRicate (https://github.com/tseemann/abricate) and the CARD database (*74*). Phenotypic traits were predicted per genome using Traitar. Trait confidence from Traitar (*75*) is expressed as an integer from 0 to 5; we converted this to a categorical prediction, with [0,2.5] considered "negative" and [2.5,5] considered "positive" for the trait. Prior to statistical testing, traits with near-constant values across all species genomes were filtered out using the mlr package (*76*) at 10% threshold (10% of features must differ from the mode value). Resultant *p* values from all statistical tests involving the COGs, pseudogenes, antibiotics markers and trait

predictions were aggregated and corrected using False Discovery Rate with a significance threshold of 0.05. FDR-corrected *p* values are indicated as *q* values.

Phylogenetic Generalized Least Squares (PGLS) correction.

A phylogenetic tree spanning all high-quality genomes used for genomic profiling was constructed in Anvio (*77*), using all available bacterial marker genes (71 reference genes from the Bacteria_71 HMM source, which were shared across all 495 annotated genomes) and according to the default parameters outlined in the phylogenomics workflow(https://merenlab.org/2017/06/07/phylogenomics/). The tree was rooted to *M. smithii*. PGLS was then applied to traits of interest (genome size, GC content, all genomic traits passing FDR correction) using a lambda model implemented in the phylolm package in R (*78*). PGLS was performed by randomly selecting one genome per species to both prune the tree and select the corresponding trait information. This random selection of tree tips and subsequent PGLS calculation with the resulting tree was repeated for 100 iterations. The mean *p* value across all iterations was used.

Machine learning.

We conducted machine learning with the goal of predicting cophylogeny (PACo *q* value < 0.01) with all variables used for trait profiling (see above). Binary random forest classifiers were trained across 5 cross validation (CV) folds, with Boruta (*79*) used for variable selection within each fold, and area under the receiver operating characteristics curve (AUC) calculated for each fold. To control for evolutionary relatedness among genomes (observations), we blocked the CV folds by genus (*i.e.,* genomes in the train or test splits did not belong to the same genus). The following R packages were used: randomForest (*80*), Boruta (*79*), and mlr (*76*).

In vitro phenotyping

Selection of taxa.

Taxa were selected from an established collection of human gut bacterial isolates (*33*) according to their ability to grow in vitro in the same rich media (mGAM) within 24h, allowing for parallel growth phenotyping. Strain information is listed in Table S19. We confirmed that this subset of 18 species did not differ significantly from the full list in terms of mean PACo ES

(p=0.45, Wilcoxen Rank Sum test); variance of the PACo ES (p=0.29, Levene's test for Homogeneity of Variance) or distribution of bacterial phyla (p=0.11, Chisquared test). We acknowledge however that the in vitro taxa tend towards an underrepresentation of Firmicutes and a lower mean PACo ES (Table S19), a bias which is driven by the fastidious growth requirements and/or lack of isolated type strains for many of the highly cophylogenetic Firmicutes.

Oxygen sensitivity.

Each strain was inoculated from a 100 µL single-use glycerol stock aliquot into 10 mL of anaerobic mGAM and grown anaerobically at 37ºC for 24 hours. Strains were then subcultured 1/10000 into 10 mL fresh anaerobic mGAM for 16h overnight. Subcultures were pelleted under atmospheric conditions (8000xg 2 min), washed in aerobic PBS, and normalized to an O.D.600 of 1.0 in a final volume of 1 mL within a 2 mL tube. Tubes were then incubated aerobically at 37ºC without shaking for 48h. Viability of the endpoint culture was assessed by plating 100 µL onto solid medium in an anaerobic hood and incubating at 37ºC anaerobically for 48h. Supplemented BHI agar (BHI-s) was routinely used for all strains with the exception of NT5011 (*R. intestinalis*), which failed to grow on BHIs agar during time-zero controls and was therefore assessed on solidified mGAM agar. A total absence of growth was scored as "dead" and at least one viable colony present was scored as "alive" (limit of detection: 10 CFU/mL). The oxygen survival assay was performed at least three independent times per species and the consensus survival of all three replicates taken. For three species which showed inconsistent viability after 48h due to CFU counts approaching the limit of detection, consensus survival was scored as the most frequent outcome out of five independent replicates. These three species could also be removed entirely from analysis without affecting the outcome (n=15, p=0.002).

Temperature sensitivity.

50 µL of each strain was inoculated from a 100 µL single-use glycerol stock aliquot into 5 mL of anaerobic mGAM and grown anaerobically at 37ºC for 24 hours. Strains were then subcultured 1/10000 into 5 mL fresh anaerobic mGAM for 16h overnight. Subcultures were normalized to an O.D. of 0.02 in mGAM under anaerobic conditions and further diluted 1:2 into a 100 µL volume per well within a 96-well plate. Half of one plate (48 technical replicates) was

assessed per biological replicate for each strain. Plates were sealed with a BreathEasy gas-permeable membrane and O.D.600 was read over 24 hours using an automated stacker and plate reader placed within an incubator set to 37ºC or 27ºC. Three independent biological replicates were performed per strain per temperature. One replicate of strain NT5009 at 37ºC and one replicate of strain NT5075 at 37ºC were subsequently excluded due to contamination of the working stock and contamination of the plate, respectively.

An area-under-the-curve (AUC) was calculated for each growth curve, which takes into account the lag phase, growth rate and yield of the culture, as has been previously described (*33*). Briefly, each growth curve was zeroed to the starting O.D., the AUC was calculated, and each AUC for each species replicate was subsequently centered and scaled to the respective 37ºC control. The average normalized AUC of all three biological replicates per species was used to yield the value "Normalized growth at 27ºC" for subsequent analysis.

Antibiotics resistance.

Antibiotics resistance data was obtained from a screen of 144 antimicrobial compounds described in Maier et al (2018)(*33*) for the same n=18 species isolates used for oxygen and temperature sensitivity (Table S19). Resistance to each drug was determined by minimum inhibitory concentration as previously described (*33*). Overall antibiotics resistance per species was calculated as the percentage of drugs to which the species was resistant out of the total 144 drugs assayed. PGLS was performed as described above for all phenotypic assays, except that only one species representative was analyzed per species in phenotypic assays; therefore, the random selection of each genome per species applied only to the tree pruning. Iteration was still repeated for 100 iterations. The mean *p* value of all iterations was used.

Supplemental Results

Tests for sample size bias.

To test whether sample size (the number of tips in the microbial phylogeny) affected the cophylogeny test results, we ran Spearman correlations between cophylogeny results (effect size (ES) or *q* value) and sample size for three different cophylogeny tests (PACo, Parafit, and Phytools) on two different tree types (BestTree and ConsTree) and two datasets (adults and children) (Table S6, Fig. S1A&B). None of the correlations were significant, suggesting that sample size is not biasing the test results using all individuals. Similarly, when we tested for correlations between sample size and cophylogeny test results within countries, none of the correlations were significant after correcting for FDR (*q* value > 0.05) (Table S7). Based on uncorrected *p* values of Spearman correlations, there was a trend that sample size correlated with PACo ES in adults (rho = 0.152, p = 0.008 Fig. S1C and Table S7). This trend was mostly driven by microbial species that had within-country sample sizes of <25 (Fig. S1C). None of the taxa with evidence of within-country cophylogeny in adults (q < 0.05) have sample size <25 (Fig. S1D, Table S9). Thus, the effect of sample size on PACo results should be minimal.

Cophylogeny tests.

To compare PACo results with other cophylogeny tests, we also applied Parafit (*23*) and Phytools (*24*) (Table S5). Overall, ES and *q* values from the three codiversification tests tended to yield similar results based on best maximum likelihood phylogenies using Spearman's rho correlation, especially for PACo and Parafit (Table S5). Although PACo is reported to have lower type I error rates and greater statistical power compared to Parafit (*68*), the ES from the two tests are highly correlated in our dataset (rho = 0.806, p = 1.4e-14) because the two tests use similar algorithms (Table S5). In contrast, Phytools is based on Robinson-Foulds distances and shows positive but non-significant correlations with either PACo or Parafit (rho = 0.031, p = 0.82 or rho = 0.100, p = 0.449, respectively), and the least number of significant taxa (*q* value < 0.05) (Table S5). The correlations were more variable when majority-rule consensus phylogenies were used (Table S5, see below). Regardless of the differences among tests, among the 36 taxa that showed positive ES in PACo, 16 taxa showed positive ES in all three tests with *q* value < 0.1. Among them, seven taxa showed *q* value < 0.05 across all tests, including *Collinsela aerofaciens* (Fig. 1H), *Catenibacterium mitsuokai* (Fig. 1G), *Prevotella copri* (Fig.

2E), *Doera longicatena*, *Coprococcus eutactus*, *Roseburia faecis*, and *Eubacterium rectale*. Despite the significant variation in quantifying cophylogeny among tests, the top and bottom taxa are generally consistent irrespective of the test used (Table S4).

Bootstrap support.

To account for branches with low bootstrap support on both host and microbial phylogenies, we created majority-rule consensus phylogenies by collapsing the branches with bootstrap support ≤ 50 and re-ran the three cophylogeny tests (Table S4). For example, among the 36 taxa that showed positive ES and q value < 0.05 using PACo based on best maximum likelihood phylogenies, 13 taxa (36%) showed positive ES and q value < 0.05 using consensus phylogenies. Parafit and Phytools also resulted in similar results, where 17/35 (49%) and 4/8 (50%) taxa remained robust to bootstrap support, respectively. Among the six tests conducted (three cophylogeny tests on two types of trees), *Collinsela aerofaciens* (Fig. 1E), *Catenibacterium mitsuokai* (Fig. 1F), *Prevotella copri* (Fig. 2E), and *Doera longicatena* showed positive ES and q value < 0.05 across five tests. Despite the fact that there are no significant correlations among the three cophylogeny tests when they are applied to majority-rule consensus phylogenies (ES or q values, Table S5), *Prevotella copri* was the only taxa that showed positive ES and q value < 0.05 across six tests. Furthermore, despite the little overlap in species between adult and child microbiomes (12 taxa), we found *P. copri* (Fig. 2E&F) and *B. wexlerae* showed evidence of copylogeny in both adults (q value < 0.01) and children (q value < 0.01) using the best maximum likelihood phylogenies (Tables S4&S11). Using majority-rule consensus phylogenies, *P. copri* and *B. longum* showed evidence of cophylogeny in both adults (q value < 0.05) and children (q value < 0.01). Although collapsing the branches with low bootstrap support did affect some of the cophylogeny test results, results from some taxa were robust to this change in the tree structure.

MAG-based phylogenies.

We created MAG-based phylogenies to compare the cophylogeny test results with marker-based phylogenies created by StrainPhlan. We were able to create MAG-based phylogenies for 33 out of the 59 taxa. Only 10 taxa met the prevalence criteria in marker-based phylogenies (≥100 adults with ≥10 individuals per major human group and ≥1 individual per

country). Among the 10 taxa, *Roseburia faecis* was the only taxa that showed evidence of cophylogeny based on PACo (uncorrected p value < 0.05) which is consistent with the marker-based results (Table S8). When we do not consider the prevalence criteria, a total of five taxa showed evidence of cophylogeny using MAG-based best maximum likelihood phylogenies based on PACo (uncorrected p value < 0.05): three taxa (*Roseburia faecis*, *Catenibacterium mitsuokai,* and *Bifidobacterium longum*) showed consistent results with marker-based phylogeny results, and two *Parabacteroides* species (*P. merdae* and *P. distasonis*) showed inconsistent results from marker-based phylogeny results (Table S8). The inconsistent results of *Parabacteroides species* are likely due to small sample size (<54 strains) and underrepresentation of non-western samples (>72% are from Germany and UK samples) in the MAGbased phylogeny. In children, MAG-based phylogenies were created for 10 out of the 20 taxa.

The only significant PACo result using MAG-based phylogeny was *Bifidobacterium longum* (q = 0.01, based on consensus phylogeny), which also showed significant cophylogeny using marker-based phylogeny (q = 0.003 and q = 0.007, based on best maximum likelihood tree and consensus tree, respectively). Although the MAG-based analyses in adults and children were underpowered, some of the top taxa were robust to different tree building methods.

Mother-child strain sharing.

We used two independent approaches to search for shared strains in mother-child pairs: inStrain (*40*), a tool based on population-level nucleotide strain diversity (popANI), and SynTracker (*39*), based on the synteny of long-stretches of DNA. We identified bacterial species with strains more similar between mothers and their own children compared to unrelated children. Of the 63 taxa (adult and child taxa combined) assessed for codiversification, 20 were tested for strain sharing events (strain sharing is defined using the following cutoffs: 0.96 for synteny, 0.99999 for popANI) between related and unrelated motherchild pairs. Of these, notably, strains of *P. copri* and other *Prevotella* spp. were shared between mothers and their children using two independent methods (Fig. 2G, Fig. S2). *Eubacterium rectale* (PACo, ES = 0.011, q = 0.003) and *Bacteroides vulgatus* (PACo, ES = 0.009, q = 0.018) also showed evidence of mother-child strain sharing (Fig. S2).

Notably, we also observed strain sharing between mothers and children who were not their own, for individuals sampled in the same location (Fig. S2). For taxa where we find no evidence of codiversification but evidence of strain sharing, such as *Bacteroides uniformis* (Fig. S2), this pattern may be due to changes in lifestyle across time (*e.g.,* greater sanitation or isolation reducing group-level transmission). These observations suggest a model where statistically significant patterns of cophylogeny can be generated with or without strict vertical transmission (*e.g.,* shared microbial source) as long as strain sharing between more genetically related individuals exceeds that of less related individuals.

Supplemental Figures

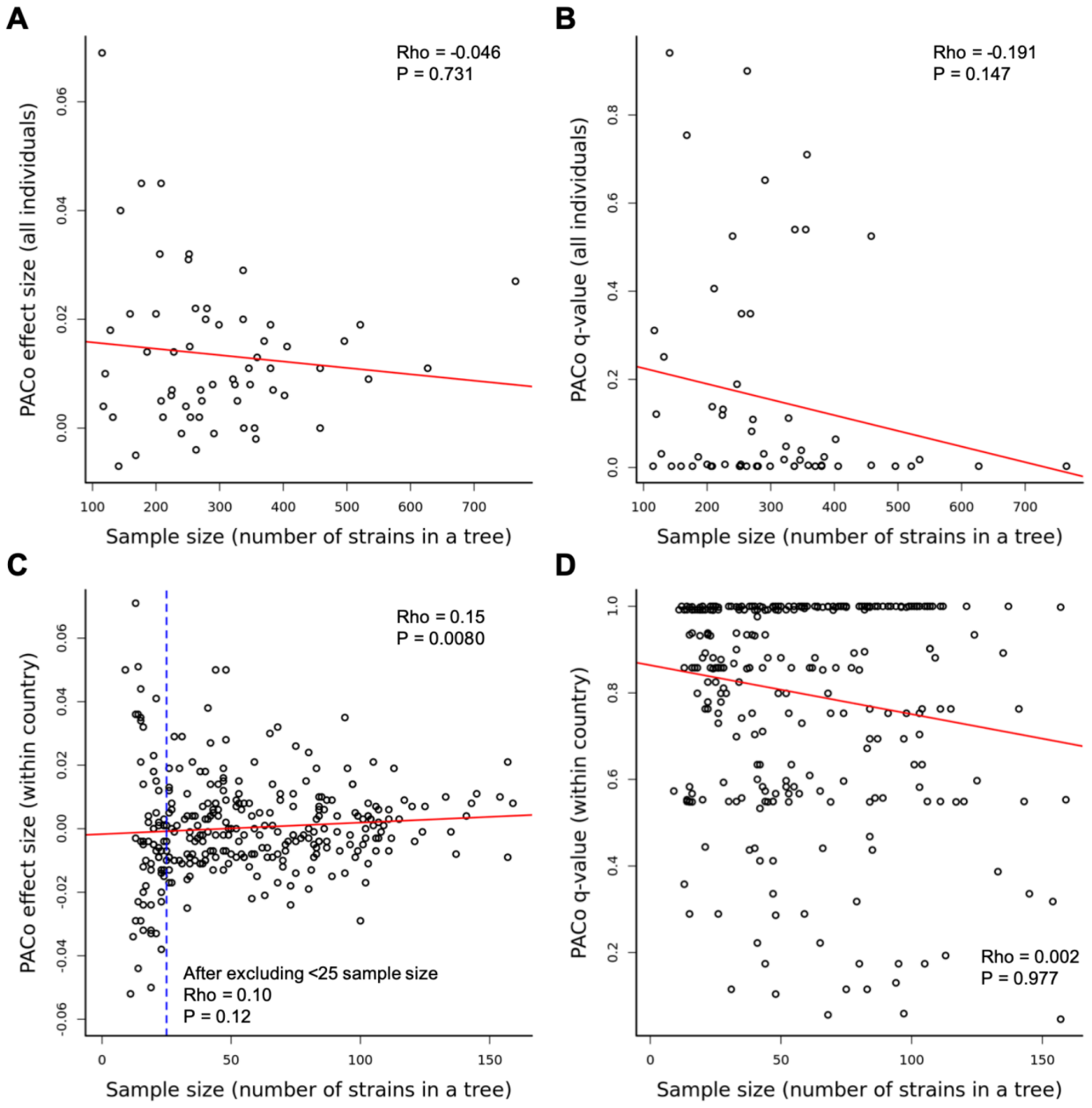

Figure S1. Correlations between sample size and PACo results. Correlations between sample size (number of strains in a given microbial phylogeny) and (A) PACo effect size (ES) and (B) PACo *q* values using all individuals. Each dot represents one of the 59 microbial taxa tested. Correlations between sample size and (C) PACo ES and (D) PACo *q* value from within-country analyses. Each dot is the 59 microbial taxa tested in six countries. In panel C, after excluding within-country analyses involving less than 25 sample size (blue dashed line), sample size does

not significantly correlate with PACo ES. Spearman correlation rho values and *p* values are shown.

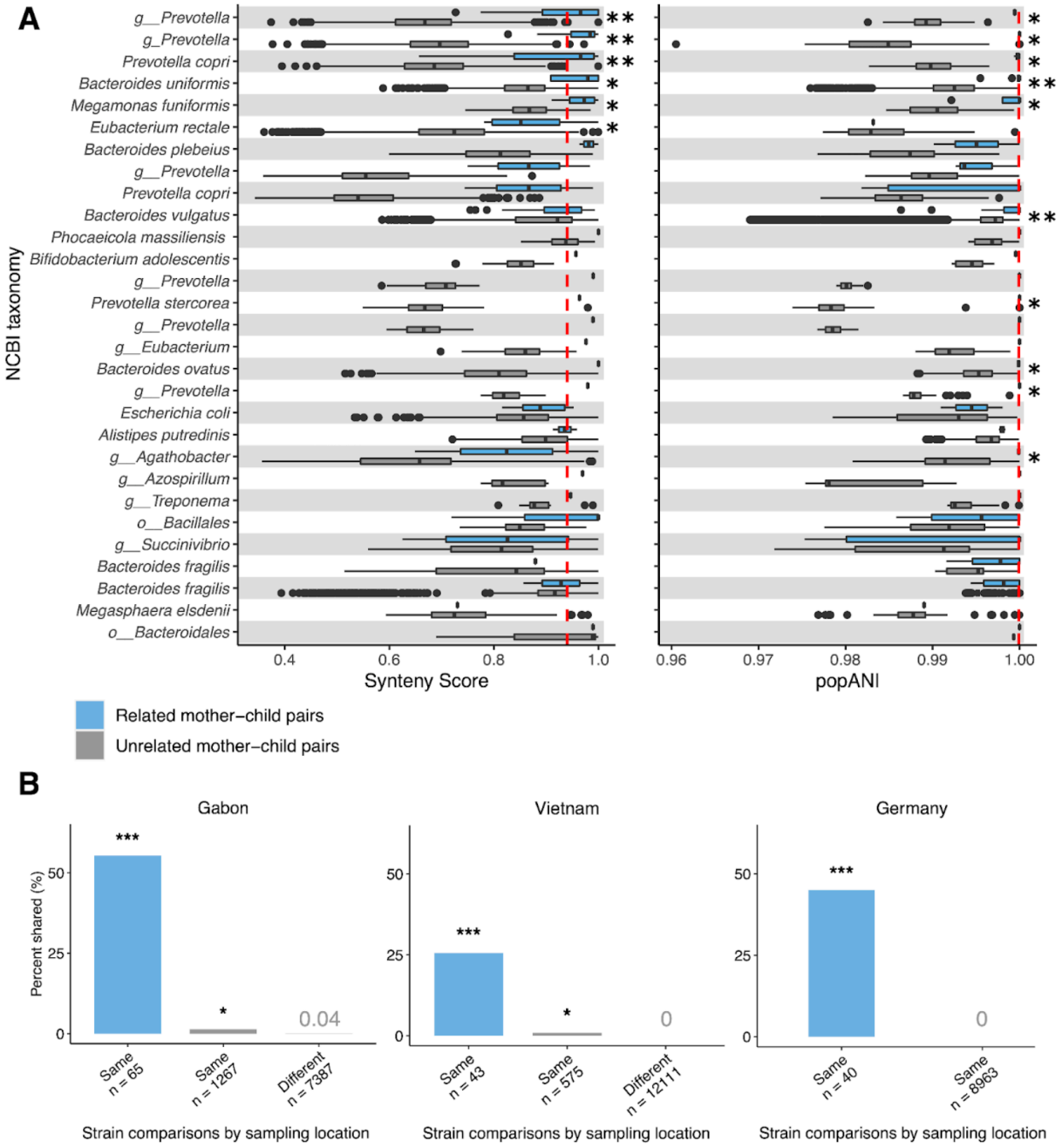

Figure S2. Strain sharing assessed within related and unrelated mother-child pairs in Gabon, Vietnam, and Germany. (A) The relatedness of strains per species within related and unrelated mother-child pairs, using SynTracker (left) and inStrain (right). Species listed are those identified in both related and unrelated mother-child pairs. Dashed red lines indicate the pairwise comparison thresholds for strain sharing events (0.96 for synteny; 0.99999 for

popANI). (B) Strain sharing within related and unrelated mother-child pairs by country using inStrain. The percentage of total strain comparisons ("n") identified as strain sharing events between related mother-child pairs, unrelated mother-child pairs sampled in the same location, and unrelated mother-child pairs sampled in different locations. Stars correspond to *q* values (corrected Wilcoxon-Mann-Whitney test for Fig. S2A and corrected hypergeometric test for Fig. S2B; * = $p < 0.05$; ** = $p < 5x10^{-5}$; *** = $p < 5x10^{-25}$; Table S13). Comparisons were calculated with samples collected from 6 locations in Gabon (average inter-location distance = 114km), 4 locations in Vietnam (average inter-location distance = 238km) and 1 location in Germany (Table S14).

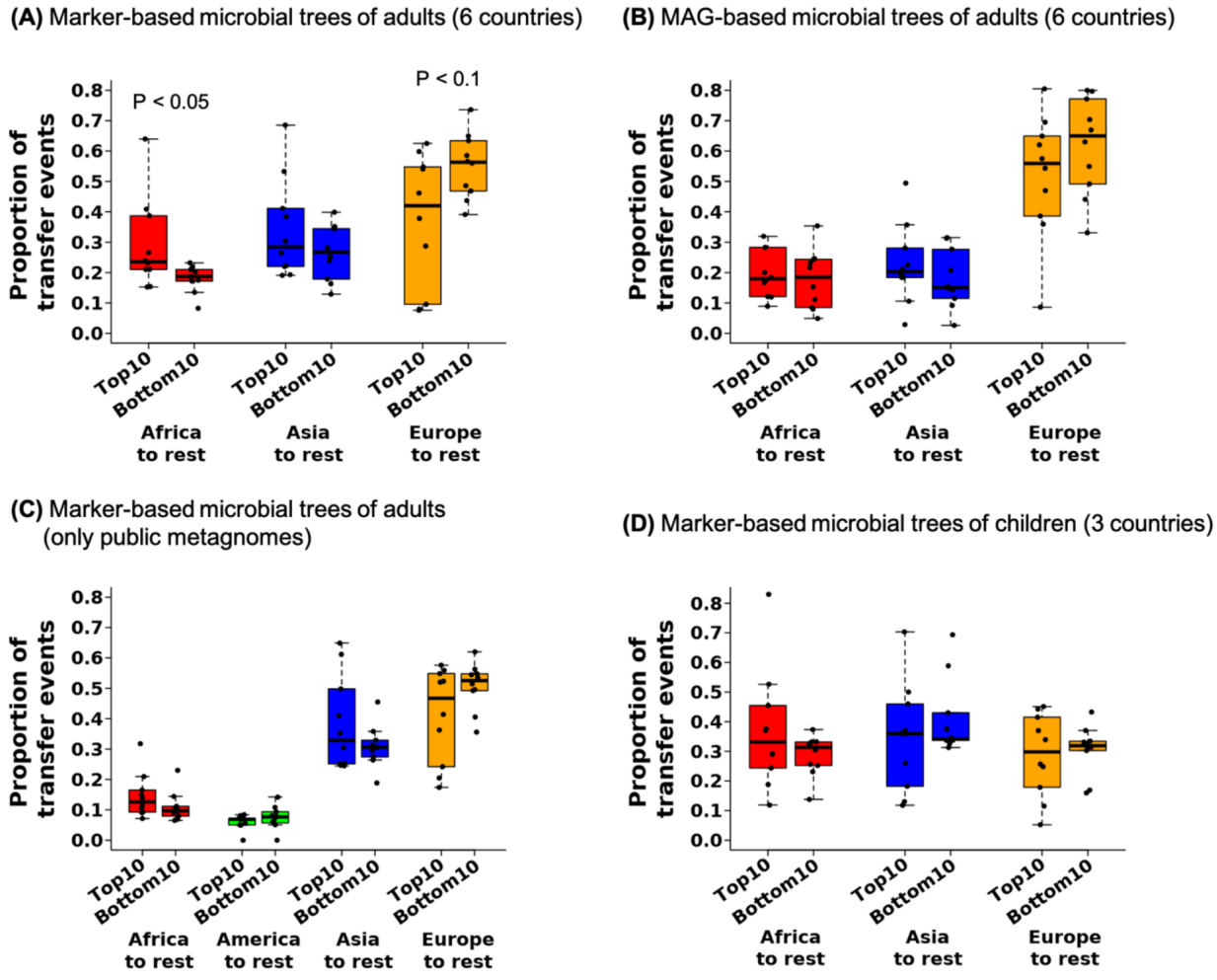

Figure S3. Boxplots comparing occurrence of transfer events between sampling regions for the top 10 and bottom 10 taxa identified by PACo effect sizes using different tree-building

approaches and populations (Table S4). Results of stochastic character mapping on microbial trees based on (A) marker genes including adult metagenomes from six countries, (B) MAGs including adult metagenomes from six countries, (C) marker genes including adult human metagenomes from only the public dataset, and (D) marker genes including child metagenomes from three countries. *p* values are based on Wilcoxon rank sum tests.

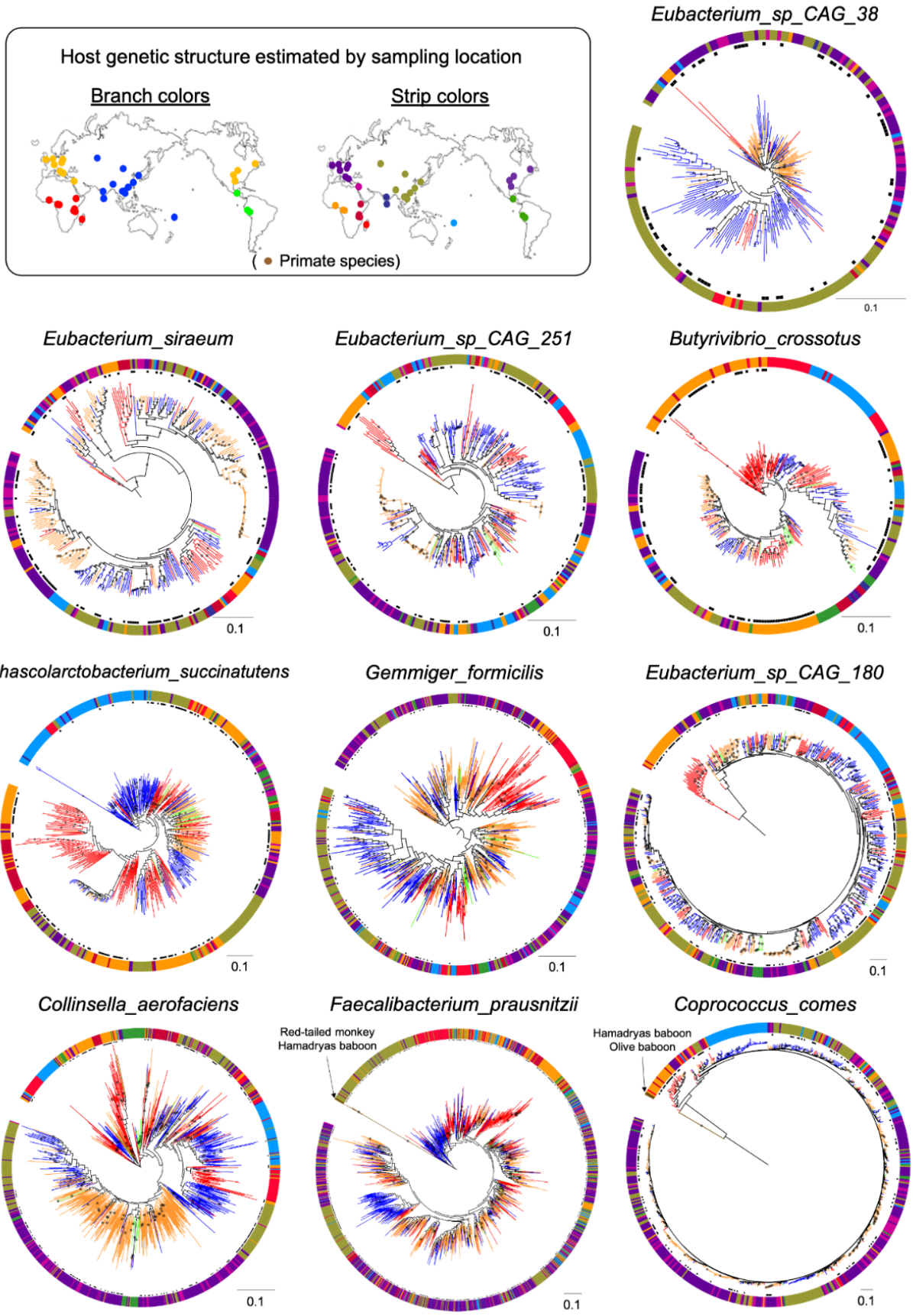
Host genetic structure estimated by sampling location
Branch colors
Strip colors
(Primate species)
Eubacterium_sp_CAG_38
Eubacterium_siraeum
Eubacterium_sp_CAG_251
Butyrivibrio_crossotus
Phascolarctobacterium_succinatutens
Gemmiger_formicilis
Eubacterium_sp_CAG_180
Collinsella_aerofaciens
Faecalibacterium_prausnitzii
Coprococcus_comes
Red-tailed monkey
Hamadryas baboon
Hamadryas baboon
Olive baboon
0.1

Figure S4. Microbial phylogenies of the top 10 taxa (based on PACo) with addition of public metagenomes. Colors of the branches and outer colorstrip indicate the host genetic structure estimated from sampling locations(*21*). Arrows indicate strains from Primate hosts. The black dots next to the outer colorstrip are samples from the initial analyses with six populations with paired fecal metagenomes and human genomes (*i.e.*, Gabon, Cameroon, Korea, Vietnam, Germany, and the UK). Bootstrap values ≥50% are shown. All trees were rooted at the midpoint. The scales show substitutions per site. Note that phylogenies for *Butyrivibrio crossotus, Coprococcus comes*, *Phascolarctobacterium succinatutens*, and *Faecalibacterium prausnitzii*, were included in Figure 3, and duplicated here to facilitate comparisons.

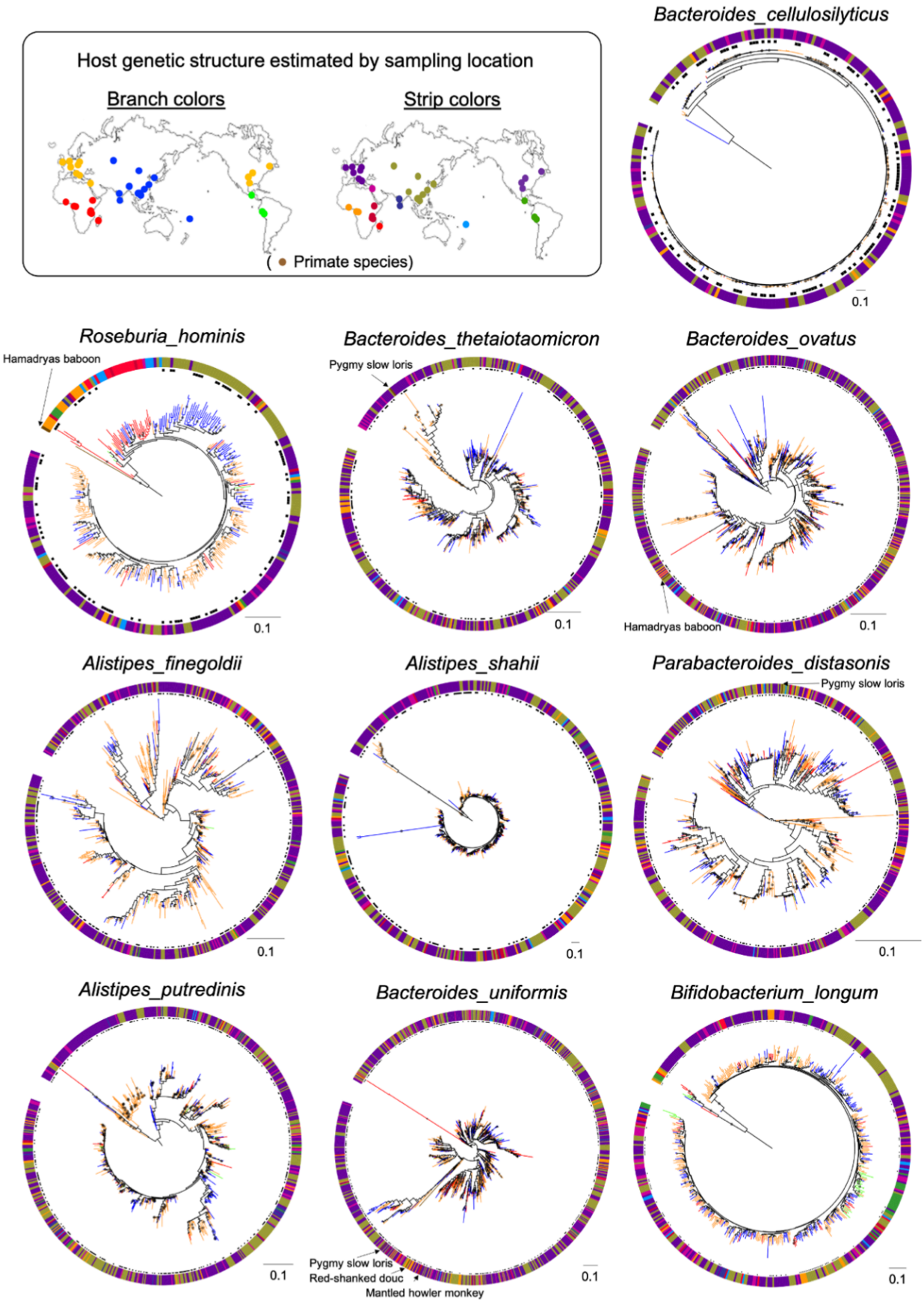
Host genetic structure estimated by sampling location
Branch colors
Strip colors
(Primate species)
Bacteroides_cellulosilyticus
0.1
Roseburia_hominis
Hamadryas baboon
0.1
Bacteroides_thetaiotaomicron
Pygmy slow loris
0.1
Bacteroides_ovatus
Hamadryas baboon
0.1
Alistipes_finegoldii
0.1
Alistipes_shahii
0.1
Parabacteroides_distasonis
Pygmy slow loris
0.1
Alistipes_putredinis
0.1
Bacteroides_uniformis
Pygmy slow loris
Red-shanked douc
Mantled howler monkey
0.1
Bifidobacterium_longum
0.1

Figure S5. Microbial phylogenies of the bottom 10 taxa (based on PACo) with addition of public metagenomes. Colors of the branches and outer colorstrip indicate the host genetic structure estimated from sampling locations(*21*). Arrows indicate strains from Primate hosts. The black dots next to the outer colorstrip are samples from the initial analyses with six populations with paired fecal metagenomes and human genomes (*i.e.*, Gabon, Cameroon, Korea, Vietnam, Germany, and the UK). Bootstrap values ≥50% are shown on branches. All trees were rooted at the midpoint. The scales show substitutions per site.

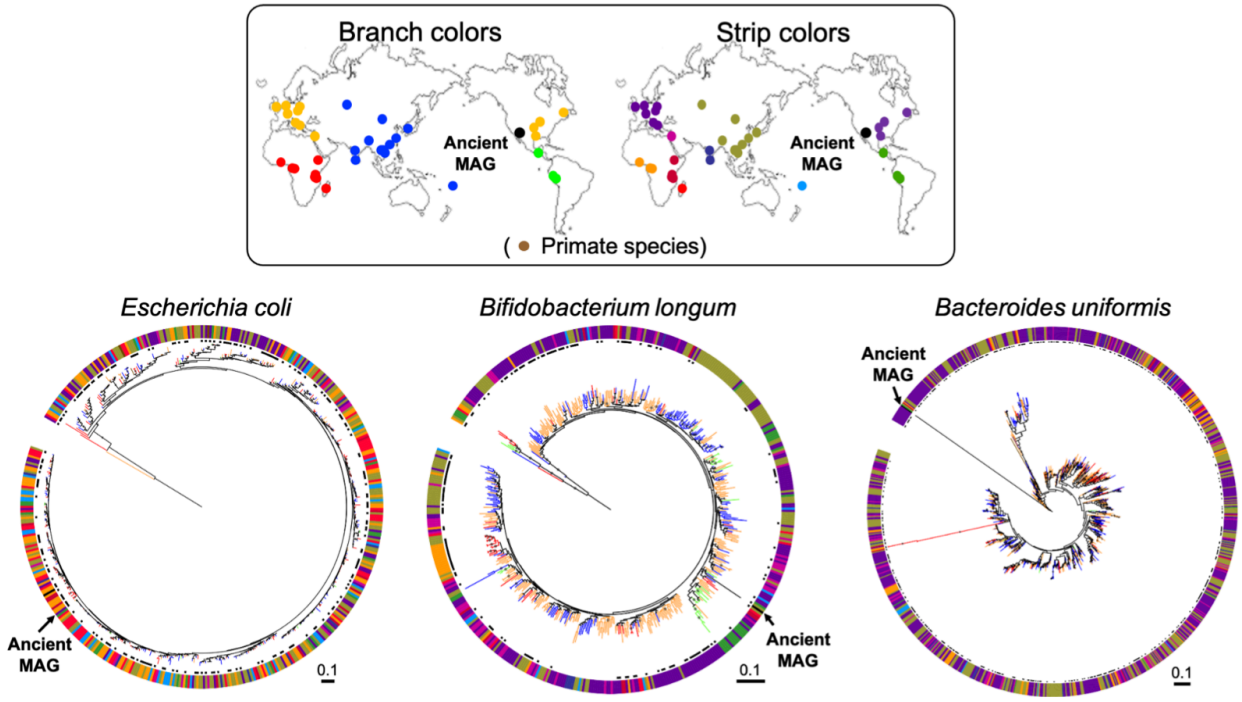

Figure S6. Microbial phylogenies that represent ancient MAGs from Native Americans. Of the five species that we recovered ancient MAGs for, these three species have nonsignficant PACo ES (two species that have significant PACo ES are shown in Fig. 3G&H). The colors of the branches and outer colorstrip indicate the host genetic structure estimated from sampling locations(*21*). Samples from the initial analyses are shown as black dots next to the outlier color strip. Arrows indicate strains from paleofeces of Native Americans. Bootstrap values ≥50% are shown on branches. All trees were rooted at the midpoint. The scales show substitutions per site.

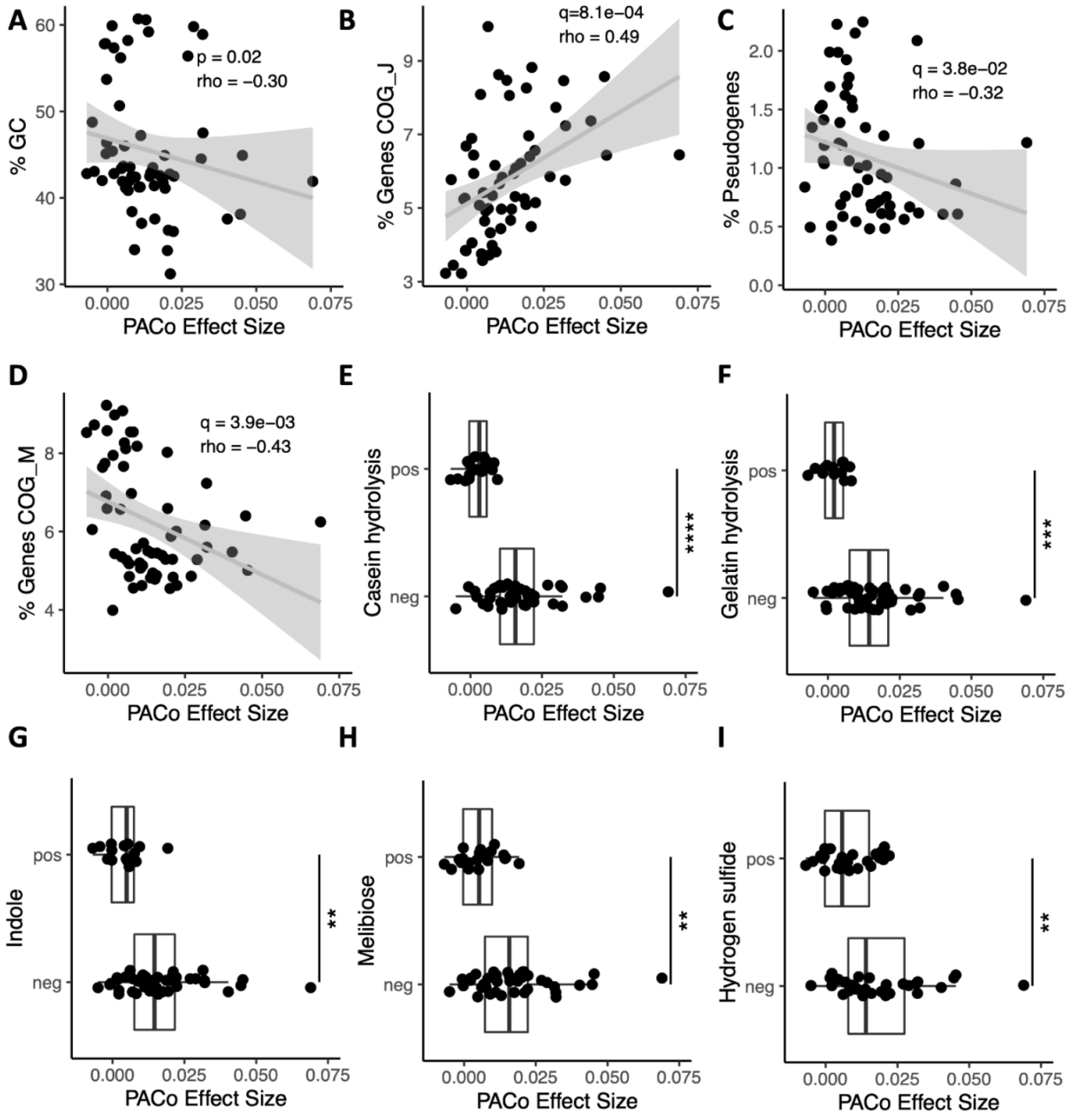

Figure S7. Selected genomic features that correlate with cophylogeny effect size. (A) Median %GC content per species. (B) Percentage of total genes in the genome annotated to COG J for translation. (C) Percentage of pseudogenes per total genes in the genome. (D) Percentage of total genes in the genome annotated to COG M for cell wall/membrane/envelope biogenesis. (EI) Predicted activity for (E) casein hydrolysis, (F) gelatin hydrolysis, (G) indole metabolism, (H) melibiose metabolism and (I) hydrogen sulfide metabolism. N=59 species for all panels. Statistical significance was determined by Spearman's correlation (A-D) or by Wilcoxon rank

sum test (E-I). Exploratory analyses were corrected using False Discovery Rate across all gene categories and predicted traits (q value; B-I). COG, cluster of orthologous genes; pos, positive for the given trait; neg, negative for the given trait. **p value <0.01, *** p value <0.001, ****p value <0.0001. Exact p values are reported in Table S18.

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S19

MDAR Reproducibility Checklist

References and Notes

1. V. K. Gupta, S. Paul, C. Dutta, Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity. *Front. Microbiol.* **8**, 1162 (2017). doi:10.3389/fmicb.2017.01162 Medline

2. P. I. Costea, L. P. Coelho, S. Sunagawa, R. Munch, J. Huerta-Cepas, K. Forslund, F. Hildebrand, A. Kushugulova, G. Zeller, P. Bork, Subspecies in the global human gut microbiome. *Mol. Syst. Biol.* **13**, 960 (2017). doi:10.15252/msb.20177589 Medline

3. E. Pasolli, F. Asnicar, S. Manara, M. Zolfo, N. Karcher, F. Armanini, F. Beghini, P. Manghi,
   A. Tett, P. Ghensi, M. C. Collado, B. L. Rice, C. DuLong, X. C. Morgan, C. D. Golden, C. Quince, C. Huttenhower, N. Segata, Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell* **176**, 649–662.e20 (2019). doi:10.1016/j.cell.2019.01.001 Medline

4. A. Tett, K. D. Huang, F. Asnicar, H. Fehlner-Peach, E. Pasolli, N. Karcher, F. Armanini, P.
   Manghi, K. Bonham, M. Zolfo, F. De Filippis, C. Magnabosco, R. Bonneau, J. Lusingu, J. Amuasi, K. Reinhard, T. Rattei, F. Boulund, L. Engstrand, A. Zink, M. C. Collado, D.

R. Littman, D. Eibach, D. Ercolini, O. Rota-Stabelli, C. Huttenhower, F. Maixner, N. Segata, The *Prevotella copri* complex comprises four distinct clades underrepresented in Westernized populations. *Cell Host Microbe* **26**, 666–679.e7 (2019). doi:10.1016/j.chom.2019.08.018 Medline

5. N. Karcher, E. Pasolli, F. Asnicar, K. D. Huang, A. Tett, S. Manara, F. Armanini, D. Bain, S. H. Duncan, P. Louis, M. Zolfo, P. Manghi, M. Valles-Colomer, R. Raffaetà, O. RotaStabelli, M. C. Collado, G. Zeller, D. Falush, F. Maixner, A. W. Walker, C. Huttenhower, N. Segata, Analysis of 1321 *Eubacterium rectale* genomes from metagenomes uncovers complex phylogeographic population structure and subspecies functional adaptations. *Genome Biol.* **21**, 138 (2020). doi:10.1186/s13059-020-02042-y Medline

6. B. D. Merrill, M. M. Carter, M. R. Olm, D. Dahan, S. Tripathi, S. P. Spencer, B. Yu, S. Jain, N. Neff, A. R. Jha, E. D. Sonnenburg, J. L. Sonnenburg, Ultra-deep sequencing of Hadza hunter-gatherers recovers vanishing microbes. bioRxiv 2022.03.30.486478 [Preprint] (2022). https://doi.org/10.1101/2022.03.30.48647.

7. D. T. Truong, A. Tett, E. Pasolli, C. Huttenhower, N. Segata, Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res.* **27**, 626–638 (2017). doi:10.1101/gr.216242.116 Medline

8. H. Enav, F. Bäckhed, R. E. Ley, The developing infant gut microbiome: A strain-level view. *Cell Host Microbe* **30**, 627–638 (2022). doi:10.1016/j.chom.2022.04.009 Medline

9. N. A. Moran, D. B. Sloan, The hologenome concept: Helpful or hollow? *PLOS Biol.* **13**, e1002311 (2015). doi:10.1371/journal.pbio.1002311 Medline

10. A. H. Moeller, A. Caro-Quintero, D. Mjungu, A. V. Georgiev, E. V. Lonsdorf, M. N. Muller, A. E. Pusey, M. Peeters, B. H. Hahn, H. Ochman, Cospeciation of gut microbiota with hominids. *Science* **353**, 380–382 (2016). doi:10.1126/science.aaf3951 Medline

11 A. H. Nishida, H. Ochman, Captivity and the co-diversification of great ape microbiomes. *Nat. Commun.* **12**, 5632 (2021). doi:10.1038/s41467-021-25732-y Medline

12. T. A. Suzuki, R. E. Ley, The role of the microbiota in human genetic adaptation. *Science*

370, eaaz6827 (2020). doi:10.1126/science.aaz6827 Medline

13. D. Falush, C. Kraft, N. S. Taylor, P. Correa, J. G. Fox, M. Achtman, S. Suerbaum, Recombination and mutation during long-term gastric colonization by *Helicobacter pylori*: Estimates of clock rates, recombination size, and minimal age. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 15056–15061 (2001). doi:10.1073/pnas.251396098 Medline

14. J. Novembre, T. Johnson, K. Bryc, Z. Kutalik, A. R. Boyko, A. Auton, A. Indap, K. S. King, S. Bergmann, M. R. Nelson, M. Stephens, C. D. Bustamante, Genes mirror geography within Europe. *Nature* **456**, 98–101 (2008). doi:10.1038/nature07331 Medline

15. N. A. Rosenberg, J. K. Pritchard, J. L. Weber, H. M. Cann, K. K. Kidd, L. A. Zhivotovsky,
M. W. Feldman, Genetic structure of human populations. *Science* **298**, 2381–2385 (2002). doi:10.1126/science.1078311 Medline

16. R. J. Abdill, E. M. Adamowicz, R. Blekhman, Public human microbiome data are dominated by highly developed countries. *PLOS Biol.* **20**, e3001536 (2022). doi:10.1371/journal.pbio.3001536 Medline

17. A. Lokmer, A. Cian, A. Froment, N. Gantois, E. Viscogliosi, M. Chabé, L. Ségurel, Use of shotgun metagenomics for the identification of protozoa in the gut microbiota of healthy individuals from worldwide populations with various industrialization levels. *PLOS ONE*
14, e0211139 (2019). doi:10.1371/journal.pone.0211139 Medline

18. G. Even, A. Lokmer, J. Rodrigues, C. Audebert, E. Viscogliosi, L. Ségurel, M. Chabé, Changes in the human gut microbiota associated with colonization by *Blastocystis* sp. and *Entamoeba* spp. in non-industrialized populations. *Front. Cell. Infect. Microbiol.* **11**, 533528 (2021). doi:10.3389/fcimb.2021.533528 Medline

19. M. Y. Lim, H. J. You, H. S. Yoon, B. Kwon, J. Y. Lee, S. Lee, Y.-M. Song, K. Lee, J. Sung, G. Ko, The effect of heritability and host genetics on the gut microbiota and metabolic syndrome. *Gut* **66**, 1031–1038 (2017). doi:10.1136/gutjnl-2015-311326 Medline

20. H. Xie, R. Guo, H. Zhong, Q. Feng, Z. Lan, B. Qin, K. J. Ward, M. A. Jackson, Y. Xia, X. Chen, B. Chen, H. Xia, C. Xu, F. Li, X. Xu, J. Y. Al-Aama, H. Yang, J. Wang, K. Kristiansen, J. Wang, C. J. Steves, J. T. Bell, J. Li, T. D. Spector, H. Jia, Shotgun metagenomics of 250 adult twins reveals genetic and environmental impacts on the gut microbiome. *Cell Syst.* **3**, 572–584.e3 (2016). doi:10.1016/j.cels.2016.10.004 Medline

21. P. Duda, Jan Zrzavý, Human population history revealed by a supertree approach. *Sci. Rep.*
6, 29890 (2016). doi:10.1038/srep29890 Medline

22. F. Asnicar, A. M. Thomas, F. Beghini, C. Mengoni, S. Manara, P. Manghi, Q. Zhu, M. Bolzan, F. Cumbo, U. May, J. G. Sanders, M. Zolfo, E. Kopylova, E. Pasolli, R. Knight, S. Mirarab, C. Huttenhower, N. Segata, Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. *Nat. Commun.* **11**, 2500 (2020). doi:10.1038/s41467-020-16366-7 Medline

23 P. Legendre, Y. Desdevises, E. Bazin, A statistical test for host-parasite coevolution. *Syst. Biol.* **51**, 217–234 (2002). doi:10.1080/10635150252899734 Medline

24. L. J. Revell, phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012). doi:10.1111/j.2041-210X.2011.00169.x

25. I. L. Brito, T. Gurry, S. Zhao, K. Huang, S. K. Young, T. P. Shea, W. Naisilisili, A. P. Jenkins, S. D. Jupiter, D. Gevers, E. J. Alm, Transmission of human-associated microbiota along family and social networks. *Nat. Microbiol.* **4**, 964–971 (2019). doi:10.1038/s41564-019-0409-6 Medline

26. J. Tung, L. B. Barreiro, M. B. Burns, J.-C. Grenier, J. Lynch, L. E. Grieneisen, J. Altmann, S. C. Alberts, R. Blekhman, E. A. Archie, Social networks predict gut microbiome composition in wild baboons. *eLife* **4**, e05224 (2015). doi:10.7554/eLife.05224 Medline

27. A. H. A. H. Moeller, S. Foerster, M. L. Wilson, A. E. Pusey, B. H. Hahn, H. Ochman, Social behavior shapes the chimpanzee pan-microbiome. *Sci. Adv.* **2**, e1500997–e1500997
(2016). doi:10.1126/sciadv.1500997 Medline

28. M. Groussin, F. Mazel, E. J. Alm, Co-evolution and co-speciation of host-gut bacteria systems. *Cell Host Microbe* **28**, 12–22 (2020). doi:10.1016/j.chom.2020.06.013 Medline

29. R. Nielsen, J. M. Akey, M. Jakobsson, J. K. Pritchard, S. Tishkoff, E. Willerslev, Tracing the peopling of the world through genomics. *Nature* **541**, 302–310 (2017). doi:10.1038/nature21347 Medline

30. M. C. Wibowo, Z. Yang, M. Borry, A. Hübner, K. D. Huang, B. T. Tierney, S. Zimmerman, F. Barajas-Olmos, C. Contreras-Cubas, H. García-Ortiz, A. Martínez-Hernández, J. M.
Luber, P. Kirstahler, T. Blohm, F. E. Smiley, R. Arnold, S. A. Ballal, S. J. Pamp, J. Russ, F. Maixner, O. Rota-Stabelli, N. Segata, K. Reinhard, L. Orozco, C. Warinner, M. Snow, S. LeBlanc, A. D. Kostic, Reconstruction of ancient microbial genomes from the human gut. *Nature* **594**, 234–239 (2021). doi:10.1038/s41586-021-03532-0 Medline

31. T. Goebel, M. R. Waters, D. H. O'Rourke, The late Pleistocene dispersal of modern humans in the Americas. *Science* **319**, 1497–1502 (2008). doi:10.1126/science.1153569 Medline

32. J. P. McCutcheon, N. A. Moran, Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* **10**, 13–26 (2011). doi:10.1038/nrmicro2670 Medline

33. L. Maier, M. Pruteanu, M. Kuhn, G. Zeller, A. Telzerow, E. E. Anderson, A. R. Brochado, K. C. Fernandez, H. Dose, H. Mori, K. R. Patil, P. Bork, A. Typas, Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* **555**, 623–628 (2018). doi:10.1038/nature25979 Medline

34. K. E. Huus, R. E. Ley, Blowing hot and cold: body temperature and the microbiome. *mSystems* **6**, e0070721 (2021). doi:10.1128/mSystems.00707-21 Medline

35. H. P. Browne, A. Almeida, N. Kumar, K. Vervier, A. T. Adoum, E. Viciani, N. J. R. Dawson, S. C. Forster, C. Cormie, D. Goulding, T. D. Lawley, Host adaptation in gut Firmicutes is associated with sporulation loss and altered transmission cycle. *Genome Biol.* **22**, 204 (2021). doi:10.1186/s13059-021-02428-6 Medline

36 S. Nayfach, Z. J. Shi, R. Seshadri, K. S. Pollard, N. C. Kyrpides, New insights from uncultivated genomes of the global human gut microbiome. *Nature* **568**, 505–510 (2019).
doi:10.1038/s41586-019-1058-x Medline

37. S. A. Frese, A. K. Benson, G. W. Tannock, D. M. Loach, J. Kim, M. Zhang, P. L. Oh, N. C. K. Heng, P. B. Patil, N. Juge, D. A. Mackenzie, B. M. Pearson, A. Lapidus, E. Dalin, H. Tice, E. Goltsman, M. Land, L. Hauser, N. Ivanova, N. C. Kyrpides, J. Walter, The evolution of host specialization in the vertebrate gut symbiont *Lactobacillus reuteri*. *PLOS Genet.* **7**, e1001314 (2011). doi:10.1371/journal.pgen.1001314 Medline

38. M. J. Barratt, S. Nuzhat, K. Ahsan, S. A. Frese, A. A. Arzamasov, S. A. Sarker, M. M. Islam, P. Palit, M. R. Islam, M. C. Hibberd, S. Nakshatri, C. A. Cowardin, J. L. Guruge, A. E. Byrne, S. Venkatesh, V. Sundaresan, B. Henrick, R. M. Duar, R. D. Mitchell, G. Casaburi, J. Prambs, R. Flannery, M. Mahfuz, D. A. Rodionov, A. L. Osterman, D. Kyle, T. Ahmed, J. I. Gordon, *Bifidobacterium infantis* treatment promotes weight gain in Bangladeshi infants with severe acute malnutrition. *Sci. Transl. Med.* **14**, eabk1107 (2022). doi:10.1126/scitranslmed.abk1107 Medline

39. H. Enav, R. E. Ley, SynTracker: a synteny based tool for tracking microbial strains. bioRxiv 2021.10.06.463341 [Preprint] (2021). https://doi.org/10.1101/2021.10.06.463341.

40. M. R. Olm, A. Crits-Christoph, K. Bouma-Gregson, B. A. Firek, M. J. Morowitz, J. F. Banfield, inStrain profiles population microdiversity from metagenomic data and sensitively detects shared microbial strains. *Nat. Biotechnol.* **39**, 727–736 (2021). doi:10.1038/s41587-020-00797-0 Medline

41. T. A. Suzuki, L. Fitzstevens, K. Huus, N. Youngblut, leylabmpi/codiversification: Zenodo release, version 1.0.1, Zenodo (2022); https://doi.org/10.5281/zenodo.6947454.

42. T. A. Suzuki, J. L. Fitzstevens, N. D. Youngblut, R. E. Ley, Phylogenies related to "Codiversification of gut microbiota with humans," Dryad (2022); https://doi.org/10.5061/dryad.qrfj6q5k2.

43. N. D. Youngblut, J. de la Cuesta-Zuluaga, G. H. Reischer, S. Dauser, N. Schuster, C. Walzer, G. Stalder, A. H. Farnleitner, R. E. Ley, Large-scale metagenome assembly reveals novel animal-associated microbial genomes, biosynthetic gene clusters, and other genetic diversity. *mSystems* **5**, e01045-20 (2020). doi:10.1128/mSystems.01045-20 Medline

44. N. Youngblut, W. Shen, nick-youngblut/gtdb_to_taxdump: Zenodo release, version 0.1.1, Zenodo (2020); https://doi.org/10.5281/zenodo.3696964.

45. D. Quinque, R. Kittler, M. Kayser, M. Stoneking, I. Nasidze, Evaluation of saliva as a source of human DNA for population and association studies. *Anal. Biochem.* **353**, 272–277
(2006). doi:10.1016/j.ab.2006.03.021 Medline

46. T.-H. Lee, H. Guo, X. Wang, C. Kim, A. H. Paterson, SNPhylo: A pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics* **15**, 162 (2014). doi:10.1186/1471-2164-15-162 Medline

47. M. T. Holder, J. Sukumaran, P. O. Lewis, A justification for reporting the majority-rule consensus tree in Bayesian phylogenetics. *Syst. Biol.* **57**, 814–821 (2008). doi:10.1080/10635150802422308 Medline

48 I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021). doi:10.1093/nar/gkab301 Medline

49. S. Rampelli, S. L. Schnorr, C. Consolandi, S. Turroni, M. Severgnini, C. Peano, P. Brigidi, A. N. Crittenden, A. G. Henry, M. Candela, Metagenome sequencing of the Hadza hunter-gatherer gut microbiota. *Curr. Biol.* **25**, 1682–1693 (2015). doi:10.1016/j.cub.2015.04.055 Medline

50. S. A. Smits, J. Leach, E. D. Sonnenburg, C. G. Gonzalez, J. S. Lichtman, G. Reid, R. Knight, A. Manjurano, J. Changalucha, J. E. Elias, M. G. Dominguez-Bello, J. L. Sonnenburg, Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania.
Science **357**, 802–806 (2017). doi:10.1126/science.aan4834 Medline

51. N. Qin, F. Yang, A. Li, E. Prifti, Y. Chen, L. Shao, J. Guo, E. Le Chatelier, J. Yao, L. Wu, J.

Zhou, S. Ni, L. Liu, N. Pons, J. M. Batto, S. P. Kennedy, P. Leonard, C. Yuan, W. Ding, Y. Chen, X. Hu, B. Zheng, G. Qian, W. Xu, S. D. Ehrlich, S. Zheng, L. Li, Alterations of the human gut microbiome in liver cirrhosis. *Nature* **513**, 59–64 (2014). doi:10.1038/nature13568 Medline

52. W. Liu, J. Zhang, C. Wu, S. Cai, W. Huang, J. Chen, X. Xi, Z. Liang, Q. Hou, B. Zhou, N. Qin, H. Zhang, Unique Features of Ethnic Mongolian Gut Microbiome revealed by metagenomic analysis. *Sci. Rep.* **6**, 34826 (2016). doi:10.1038/srep34826 Medline53. D. B. Dhakan, A. Maji, A. K. Sharma, R. Saxena, J. Pulikkan, T. Grace, A. Gomez, J. Scaria, K. R. Amato, V. K. Sharma, The unique composition of Indian gut microbiome, gene catalogue, and associated fecal metabolome deciphered using multi-omics approaches. *Gigascience* **8**, giz004 (2019). doi:10.1093/gigascience/giz004 Medline

54. Z. Jie, H. Xia, S.-L. Zhong, Q. Feng, S. Li, S. Liang, H. Zhong, Z. Liu, Y. Gao, H. Zhao, D. Zhang, Z. Su, Z. Fang, Z. Lan, J. Li, L. Xiao, J. Li, R. Li, X. Li, F. Li, H. Ren, Y. Huang, Y. Peng, G. Li, B. Wen, B. Dong, J.-Y. Chen, Q.-S. Geng, Z.-W. Zhang, H. Yang, J.

Wang, J. Wang, X. Zhang, L. Madsen, S. Brix, G. Ning, X. Xu, X. Liu, Y. Hou, H. Jia, K. He, K. Kristiansen, The gut microbiome in atherosclerotic cardiovascular disease. *Nat. Commun.* **8**, 845 (2017). doi:10.1038/s41467-017-00900-1 Medline

55. L. A. David, A. Weil, E. T. Ryan, S. B. Calderwood, J. B. Harris, F. Chowdhury, Y. Begum, F. Qadri, R. C. LaRocque, P. J. Turnbaugh, Gut microbial succession follows acute secretory diarrhea in humans. *mBio* **6**, e00381-15 (2015). doi:10.1128/mBio.00381-15 Medline

56. D. Zeevi, T. Korem, N. Zmora, D. Israeli, D. Rothschild, A. Weinberger, O. Ben-Yacov, D. Lador, T. Avnit-Sagi, M. Lotan-Pompan, J. Suez, J. A. Mahdi, E. Matot, G. Malka, N. Kosower, M. Rein, G. Zilberman-Schapira, L. Dohnalová, M. Pevsner-Fischer, R. Bikovsky, Z. Halpern, E. Elinav, E. Segal, Personalized nutrition by prediction of glycemic responses. *Cell* **163**, 1079–1094 (2015). doi:10.1016/j.cell.2015.11.001 Medline

57. M. Schirmer, R. D'Amore, U. Z. Ijaz, N. Hall, C. Quince, Illumina error profiles: Resolving fine-scale variation in metagenomic sequencing data. *BMC Bioinformatics* **17**, 125
(2016). doi:10.1186/s12859-016-0976-y Medline

58 G. Zeller, J. Tap, A. Y. Voigt, S. Sunagawa, J. R. Kultima, P. I. Costea, A. Amiot, J. Böhm, F. Brunetti, N. Habermann, R. Hercog, M. Koch, A. Luciani, D. R. Mende, M. A. Schneider, P. Schrotz-King, C. Tournigand, J. Tran Van Nhieu, T. Yamada, J.
Zimmermann, V. Benes, M. Kloor, C. M. Ulrich, M. von Knebel Doeberitz, I. Sobhani, P. Bork, Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol. Syst. Biol.* **10**, 766 (2014). doi:10.15252/msb.20145645 Medline

59. P. Ferretti, E. Pasolli, A. Tett, F. Asnicar, V. Gorfer, S. Fedi, F. Armanini, D. T. Truong, S. Manara, M. Zolfo, F. Beghini, R. Bertorelli, V. De Sanctis, I. Bariletti, R. Canto, R. Clementi, M. Cologna, T. Crifò, G. Cusumano, S. Gottardi, C. Innamorati, C. Masè, D. Postai, D. Savoi, S. Duranti, G. A. Lugli, L. Mancabelli, F. Turroni, C. Ferrario, C. Milani, M. Mangifesta, R. Anzalone, A. Viappiani, M. Yassour, H. Vlamakis, R. Xavier, C. M. Collado, O. Koren, S. Tateo, M. Soffiati, A. Pedrotti, M. Ventura, C. Huttenhower, P. Bork, N. Segata, Mother-to-infant microbial transmission from different body sites shapes the developing infant gut microbiome. *Cell Host Microbe* **24**, 133–145.e5 (2018). doi:10.1016/j.chom.2018.06.005 Medline

60. S. Louis, R.-M. Tappu, A. Damms-Machado, D. H. Huson, S. C. Bischoff, Characterization of the gut microbial community of obese patients following a weight-loss intervention using whole metagenome shotgun sequencing. *PLOS ONE* **11**, e0149564 (2016). doi:10.1371/journal.pone.0149564 Medline

61. Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012). doi:10.1038/nature11234 Medline

62. E. C. Pehrsson, P. Tsukayama, S. Patel, M. Mejía-Bautista, G. Sosa-Soto, K. M. Navarrete, M. Calderon, L. Cabrera, W. Hoyos-Arango, M. T. Bertoli, D. E. Berg, R. H. Gilman, G.
Dantas, Interconnected microbiomes and resistomes in low-income human habitats.

Nature **533**, 212–216 (2016). doi:10.1038/nature17672 Medline

63. F. Raymond, A. A. Ouameur, M. Déraspe, N. Iqbal, H. Gingras, B. Dridi, P. Leprohon, P.-L. Plante, R. Giroux, È. Bérubé, J. Frenette, D. K. Boudreau, J.-L. Simard, I. Chabot, M.-C. Domingo, S. Trottier, M. Boissinot, A. Huletsky, P. H. Roy, M. Ouellette, M. G. Bergeron, J. Corbeil, The initial state of the human gut microbiome determines its reshaping by antibiotics. *ISME J.* **10**, 707–720 (2016). doi:10.1038/ismej.2015.148 Medline

64. A. J. Obregon-Tito, R. Y. Tito, J. Metcalf, K. Sankaranarayanan, J. C. Clemente, L. K. Ursell, Z. Zech Xu, W. Van Treuren, R. Knight, P. M. Gaffney, P. Spicer, P. Lawson, L. Marin-Reyes, O. Trujillo-Villarroel, M. Foster, E. Guija-Poma, L. Troncoso-Corzo, C. Warinner, A. T. Ozga, C. M. Lewis, Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat. Commun.* **6**, 6505 (2015). doi:10.1038/ncomms7505 Medline

65. I. L. Brito, S. Yilmaz, K. Huang, L. Xu, S. D. Jupiter, A. P. Jenkins, W. Naisilisili, M. Tamminen, C. S. Smillie, J. R. Wortman, B. W. Birren, R. J. Xavier, P. C. Blainey, A. K. Singh, D. Gevers, E. J. Alm, Mobile genes in the human microbiome are structured from global to individual scales. *Nature* **535**, 435–439 (2016). doi:10.1038/nature18927 Medline

66 K. R. Amato, J. G. Sanders, S. J. Song, M. Nute, J. L. Metcalf, L. R. Thompson, J. T. Morton, A. Amir, V. J. McKenzie, G. Humphrey, G. Gogul, J. Gaffney, A. L. Baden, G. A. O. Britton, F. P. Cuozzo, A. Di Fiore, N. J. Dominy, T. L. Goldberg, A. Gomez, M. M. Kowalewski, R. J. Lewis, A. Link, M. L. Sauther, S. Tecot, B. A. White, K. E. Nelson, R. M. Stumpf, R. Knight, S. R. Leigh, Evolutionary trends in host physiology outweigh dietary niche in structuring primate gut microbiomes. *ISME J.* **13**, 576–587 (2019). doi:10.1038/s41396-018-0175-0 Medline

67. N. D. Youngblut, G. H. Reischer, W. Walters, N. Schuster, C. Walzer, G. Stalder, R. E. Ley, A. H. Farnleitner, Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. *Nat. Commun.* **10**, 2200 (2019). doi:10.1038/s41467-019-10191-3 Medline

68. J. A. Balbuena, R. Míguez-Lozano, I. Blasco-Costa, PACo: A novel Procrustes application to cophylogenetic analysis. *PLOS ONE* **8**, e61048 (2013). doi:10.1371/journal.pone.0061048 Medline

69. J. P. Huelsenbeck, R. Nielsen, J. P. Bollback, Stochastic mapping of morphological characters. *Syst. Biol.* **52**, 131–158 (2003). doi:10.1080/10635150390192780 Medline

70. D. H. Parks, M. Imelfort, C. T. Skennerton, P. Hugenholtz, G. W. Tyson, CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015). doi:10.1101/gr.186072.114 Medline

71. D. Hyatt, G.-L. Chen, P. F. Locascio, M. L. Land, F. W. Larimer, L. J. Hauser, Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010). doi:10.1186/1471-2105-11-119 Medline

72. J. Huerta-Cepas, K. Forslund, L. P. Coelho, D. Szklarczyk, L. J. Jensen, C. von Mering, P. Bork, Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol. Biol. Evol.* **34**, 2115–2122 (2017). doi:10.1093/molbev/msx148 Medline

73. Y. Tanizawa, T. Fujisawa, Y. Nakamura, DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* **34**, 1037–1039 (2018). doi:10.1093/bioinformatics/btx713 Medline

74. B. Jia, A. R. Raphenya, B. Alcock, N. Waglechner, P. Guo, K. K. Tsang, B. A. Lago, B. M. Dave, S. Pereira, A. N. Sharma, S. Doshi, M. Courtot, R. Lo, L. E. Williams, J. G. Frye, T. Elsayegh, D. Sardar, E. L. Westman, A. C. Pawlowski, T. A. Johnson, F. S. L. Brinkman, G. D. Wright, A. G. McArthur, CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573 (2017). doi:10.1093/nar/gkw1004 Medline

75. A. Weimann, K. Mooren, J. Frank, P. B. Pope, A. Bremges, A. C. McHardy, From genomes to phenotypes: Traitar, the microbial trait analyzer. *mSystems* **1**, e00101-16 (2016).
doi:10.1128/mSystems.00101-16 Medline

76. B. Bischl, M. Lang, L. Kotthoff, J. Schiffner, J. Richter, E. Studerus, G. Casalicchio, Z. M. Jones, mlr: Machine learning in R. *J. Mach. Learn. Res.* **17**, 5938–5942 (2016).

77. A. M. Eren, E. Kiefl, A. Shaiber, I. Veseli, S. E. Miller, M. S. Schechter, I. Fink, J. N. Pan, M. Yousef, E. C. Fogarty, F. Trigodet, A. R. Watson, Ö. C. Esen, R. M. Moore, Q. Clayssen, M. D. Lee, V. Kivenson, E. D. Graham, B. D. Merrill, A. Karkman, D.

Blankenberg, J. M. Eppley, A. Sjödin, J. J. Scott, X. Vázquez-Campos, L. J. McKay, E. A. McDaniel, S. L. R. Stevens, R. E. Anderson, J. Fuessel, A. Fernandez-Guerra, L. Maignien, T. O. Delmont, A. D. Willis, Community-led, integrated, reproducible multiomics with anvi'o. *Nat. Microbiol.* **6**, 3–6 (2021). doi:10.1038/s41564-020-00834-3 Medline

78. L. s. T. Ho, C. Ané, A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol.* **63**, 397–408 (2014). doi:10.1093/sysbio/syu005 Medline

79. M. B. Kursa, W. R. Rudnicki, Feature selection with the Boruta package. *J. Stat. Softw.* **36**, 1–13 (2010). doi:10.18637/jss.v036.i11

80. W. Liaw, Classification and regression by randomForest. *R News* **2**, 18–22 (2002).

Other Supplementary Material not included in manuscript:

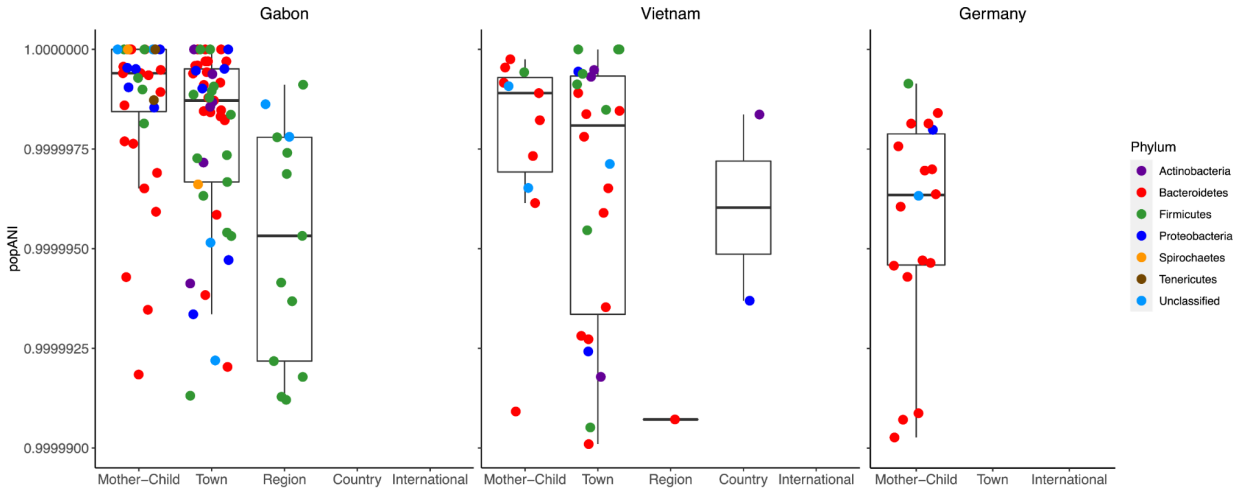

Figure not in manuscript. Pairwise comparisons of mothers' gut microbial strains with those of their children, and those of unrelated children, all of which had the minimum popANI "same strain" threshold of 0.99999. popANI of 1 indicates genome clonality, with lower scores indicating more single nucleotide polymorphisms (SNPs) differentiating the strains' genomes. The pairwise comparison scores are divided into several categories: those of mothers with their own children, and mothers with unrelated children having been sampled in the same town, region, country, or internationally. The phylum of each compared strain is indicated by its color.

Results/Discussion not included in manuscript:

Though there is insufficient statistical power to test this, these preliminary suggests that there is a gradient of strain-relatedness; the most-related being those shared between related mothers and infants, followed by those of unrelated mothers and infants living in the same town, with an increasing number of SNPs differentiating strains as the geographic distance between those unrelated mothers and infants increases.

Chapter three: Microbes persist in human infant and adult guts for up to 13 years

(Advanced manuscript; awaiting submission to target journal)

Liam Fitzstevens[1], Miles Angenent[2], Largus T. Angenent[3,4,5,6,7], and Ruth E. Ley[1,6,#]

correspondence: rley@tuebingen.mpg.de

[1] Department of Microbiome Science, Max Planck Institute for Biology, Tübingen, Germany.

[2] Geschwister-Scholl-Schule, Tübingen, Germany.

[3] Environmental Biotechnology Group, Department of Geosciences, University of Tübingen, Tübingen, Germany.

[4] AG Angenent, Max Planck Institute for Biology, Tübingen, Germany.

[5] Department of Biological and Chemical Engineering, Aarhus University, Aarhus, Denmark.

[6] Novo Nordisk Foundation CO_2 Research Center, Aarhus University, Aarhus, Denmark.

[7] Cluster of Excellence EXC 2124 Controlling Microbes to Fight Infections, University of Tübingen, Tübingen, Germany.

Abstract

It is unknown whether microbial strains acquired by human infants persist in their guts beyond age 5. This is due to the logistical challenge of conducting such extensive longitudinal studies. In this case study, we sampled the stool of a 13-year old human, from whom we also had frozen stool from their first 2.5 years of life. To determine whether gut microbial strains persist over such long periods of time, we sequenced their stool metagenomes from both infancy and adolescence - including several from the parents, assembled genomes from metagenomes (MAGs), and called single-nucleotide polymorphisms (SNPs) by mapping metagenome reads to those MAGs. To account for evolution over those elapsed 13 years, we defined a persistent strain as instances where 99.999 % of the genome is identical (e.g., a microbe with a 4 kilobase pair genome has a "same strain"-threshold of 40 SNPs). We detected 2 strains that spanned infancy into adolescence: one belonging to the species *Prevotella copri*, and the second to *Alistipes_A ihumii*. The *Prevotella copri* strain was first detected on day 4 of life and again on day 294, after which it remained detectable throughout the majority of the adolescent samples. The *Alistipes_A ihumii* strain was first detected later in adolescence, and was also identified as being present in most of the adolescent stool samples. We also found that strains persisted within and between the participant's parents, belonging to the species *Phocaeicola mediterraneensis*, *Alistipes_A ihumii*, *Bacteroides uniformis*, and *Alistpies putredinis*, as well as a novel bacterium we name *Persistoides*, for which there was no phylum-level classification. While a recent cross-sectional study suggested that certain strains persist in humans over very long periods of time, this is the first direct evidence that individual strains do, in fact, persist from infancy into adolescence.

Introduction

Humans have two sets of genes: those encoded by their own genome, and those encoded by their microbial symbionts. The human genes amount to approximately 22,000, are inherited vertically from parents, and are stable throughout life (Willyard 2018). The microbial genes in the gut alone outnumber those of the human by a staggering 22 million, and start to be acquired from the environment after birth (Tierney et al. 2019; Kennedy et al. 2023). They are distributed amongst the genomes of thousands of bacterial, archaeal, and viral strains, which colonize infant guts via transmission from diverse sources, including the hospitals in which they are born and their parents (Lou et al. 2021; Olm et al. 2021; Enav, Bäckhed, and Ley 2022; Suzuki et al. 2022). This vertical transmission of gut microbes between host generations has contributed to a subset of them having recently been shown to codiversify with humans, wherein patterns of phylogenetic divergence between symbiont microbes reflect those of their human hosts (Suzuki et al. 2022).

Beyond vertical transmission, codiversification of gut microbes with humans is contingent on a second biological phenomenon: the persistence of these strains in the gut throughout their human hosts' lifetimes. This is because in a recent study of a westernized population wherein codiversification was detected, not a single strain was horizontally-transmitted; none were shared between unrelated mothers and infants (Suzuki et al. 2022). Without strain-persistence in the gut, these strictly vertically-acquired microbes would go extinct in their host prior to their transmission to the next host generation, eliminating their signature of host-codiversification that we know to exist. In light of this prediction, strains have not been shown to exist for longer than 5 years in infant guts, or even those in adult guts, due to the logistical challenges of conducting such lengthy longitudinal studies (Faith et al. 2013; Bäckhed et al. 2015; Yassour et al. 2016; Chu et al. 2017; Vatanen et al. 2018).

Here, we conducted a case-study to determine whether microbial strains acquired during a human's infancy were still present in their gut during adolescence. The first sample collection effort for this project was conducted 15 years ago, during the first 2.5 years of the participant's life (Koenig et al. 2011). We collected new stool samples from the same participant at ages 12 and 13, and sequenced their metagenomes, together with those from their infancy. We assembled genomes from those metagenomes, which we then dereplicated, filtered, taxonomically-profiled,

and mapped metagenome-reads to in order to detect gut microbial strains (Olm et al. 2021). This, as well as any alignment-based approach for strain-resolved metagenomics, is made challenging by the fact that microbes abundant during infancy differ from those abundant during adulthood. Our hypothesis is that many of the same strains are present in both infancy and adulthood, but at very low abundances, as that would maintain codiversification in populations wherein transmission is strictly vertical. This very-low abundance means that strains will be at the limit of detection; we deeply-sequenced the metagenomes to an average of 10 million paired-end reads to capture as much strain-sharing between time points as possible.

Methods

Sampling

We froze stool samples collected throughout the infant's first 2.5 years of life, along with those of his mother, at -80 °C for 13 years. This original sampling was approved by Washington University in St. Louis' Internal Review Board (protocol 09-0039). Permission to collect new stool samples from the same participants during the then-infant's 12th and 13th years of life, along with those of the mother and father, was granted by the University of Tübingen Medical Faculty's ethical committee (protocol 672/2018B02), and were also stored at -80 °C.

Gut metagenomes

We extracted DNA from stool samples with DNeasy Powersoil HTP 96 kits (Qiagen, Venlo, Netherlands). We prepared libraries as previously described by Karasov and colleagues (Karasov et al. 2018). Briefly, we size-selected fragments to 400-700 base pairs with BluePippin, and used an Illumina HiSeq 3000 to sequence pooled libraries with paired-end 150 base pair reads. We performed quality control on these metagenomes as described by Youngblut and colleagues (Youngblut et al. 2020). We taxonomically profiled that output using Kraken2 and Bracken v2, with a custom database curated by Struo2, using GTDB release 207 (Lu et al. 2017; Youngblut and Ley 2021; Parks et al. 2022). We performed data analysis in R with the tidyverse package suite (Wickham et al. 2019). We rarefied metagenomes to the lowest common read

count before calculating Shannon diversity and unweighted Bray Curtis dissimilarity with the package Phyloseq v3.17.

Metagenome-assembled genome (MAG) generation and analysis

We assembled MAGs as described by Suzuki and colleagues (Suzuki et al. 2022). Briefly, we de novo assembled contigs with metaSPAdes v3.15.5, and binned them with differential coverage using MetaBat v2.15.0, MaxBin v2.2.7 and VAMB v.4.1.1. We used Das-Tool v1.1.6 to select the non-redundant bins of highest-quality. We dereplicated these MAGs to 98% average nucleotide identity (ANI) with CheckM2 v1.0.1 and dRep v2.0.0, thereby generating sub-species representative genomes (SSRGs). We filtered the SSRGs to a minimum CheckM2-estimated completeness of 50% and maximum contamination of 5%, and taxonomically classified them using GTDB-Tk database release 207 (Olm et al. 2017; Chaumeil et al. 2019; Chklovski et al. 2022).

Strain sharing analysis and phenotyping

We used inStrain v1.3.0 to map reads from all metagenomes to the SSRGs, which generated inStrain `profiles` that catalog the SNPs in each metagenome, and pairwise-compared those `profiles` with inStrain `compare` to compute popANI (Olm et al. 2021). We then used inStrain to hierarchically-cluster the output of these pairwise comparisons to generate strain clusters, each of which represented a unique strain. We used Traitar v3.0.1 to phenotype strain genomes.

Results

Composition of gut microbial phyla was variable during the first year of life (Figure 1, A, B, and C). The first and second weeks were taxonomically dominated by *Firmicutes* and *Proteobacteria*, respectively, and weeks 3-10 by *Firmicutes_A*, each reaching peak relative abundances of 75 % (Figure 1, C). This was followed by a bloom of *Actinobacteria*, which at its peak on day 139 also reached a relative abundance of 75 % (Figure 1, C). *Bacteriodota* started increasing in relative abundance by month 5, and reached 100 % relative abundance during month 7 (Figure 1, C). By month 10, the microbiome consisted mostly of *Bacteriodota* and

Firmicutes_A, which remained the case through the final infant sampling at age 2.3 years (Figure 1, C). The same two phyla remained most abundant at age 12, which remained the case at age 13 (Figure 1, C). Shannon diversity, which accounts for both the number of microbial species detected and their evenness in the population, increased with infant age (Figure 1, A). It did not reach its peak of 5 by the final infancy timepoints, but rather as an adolescent (Figure 1, A). On a principal component of analysis plot of unweighted Bray Curtis distances, metagenomes formed four clusters, each comprising metagenomes of the host with progressively increasing age, such that the fourth cluster was formed by the oldest infant sample and those from adolescence, consistent with metagenomes from both being dominated by *Bacteriodota* and *Firmicutes_A* (Figure 1, B).

We assembled 1,659 metagenome-assembled genomes (MAGs). We dereplicated them at 99.9 % ANI to create a final set of 892 non-redundant MAGs, and dereplicated those further to 98 % ANI, thereby generating 269 sub-species representative genomes (SSRGs). We mapped each of our 81 metagenomes to these SSRGs to identify SNPs, and compared those SNPs between samples, which amounted to a total of 6,903 pairwise comparisons of 414 strains. 3,652 of the comparisons were identified as being the same strain (i.e., popANI >= 99.999 %). As expected, the vast majority of these strain-sharing events were between time points during either infancy (Figure S1), or during adolescence (Figure S2) (3492/3652; 96 %), but the rest were detected in both (Figure 1, D) (160/3652; 4 %).

We identified two strains present in at least one of the child's infant and adolescent stool samples, one belonging to *Prevotella copri* and the second to *Alistipes_A ihumii* (Figure 1, D). We inferred microbial phenotypes from these strains' genomes with Traitar; both shared many phenotypes, including that they are gram-negative, obligate anaerobes that can hydrolyze and grow on lactose, and ferment glucose (Figure 1, E). Neither of them were phenotyped as being motile, spore formers, or having catalase. We detected the same strains of *Phocaeicola mediterraneensis, Alistepes putredinis*, and *Alistipes_A ihumii* in the mother, despite 13 years having elapsed between sampling (Figure 1, E). We also detected a strain of a novel bacterium with no phylum classification in GTDB in the mother's stool on the infant's third day of life, which we detected in the infant's stool 13 years later (Figure 1, E), to which we assign a genus name *Persistoides*. We observed the same within-family persistence for a strain of *Bacteroides*

uniformis, detected in several of the infant samples and again in the father 12 years later (Figure 1, E).

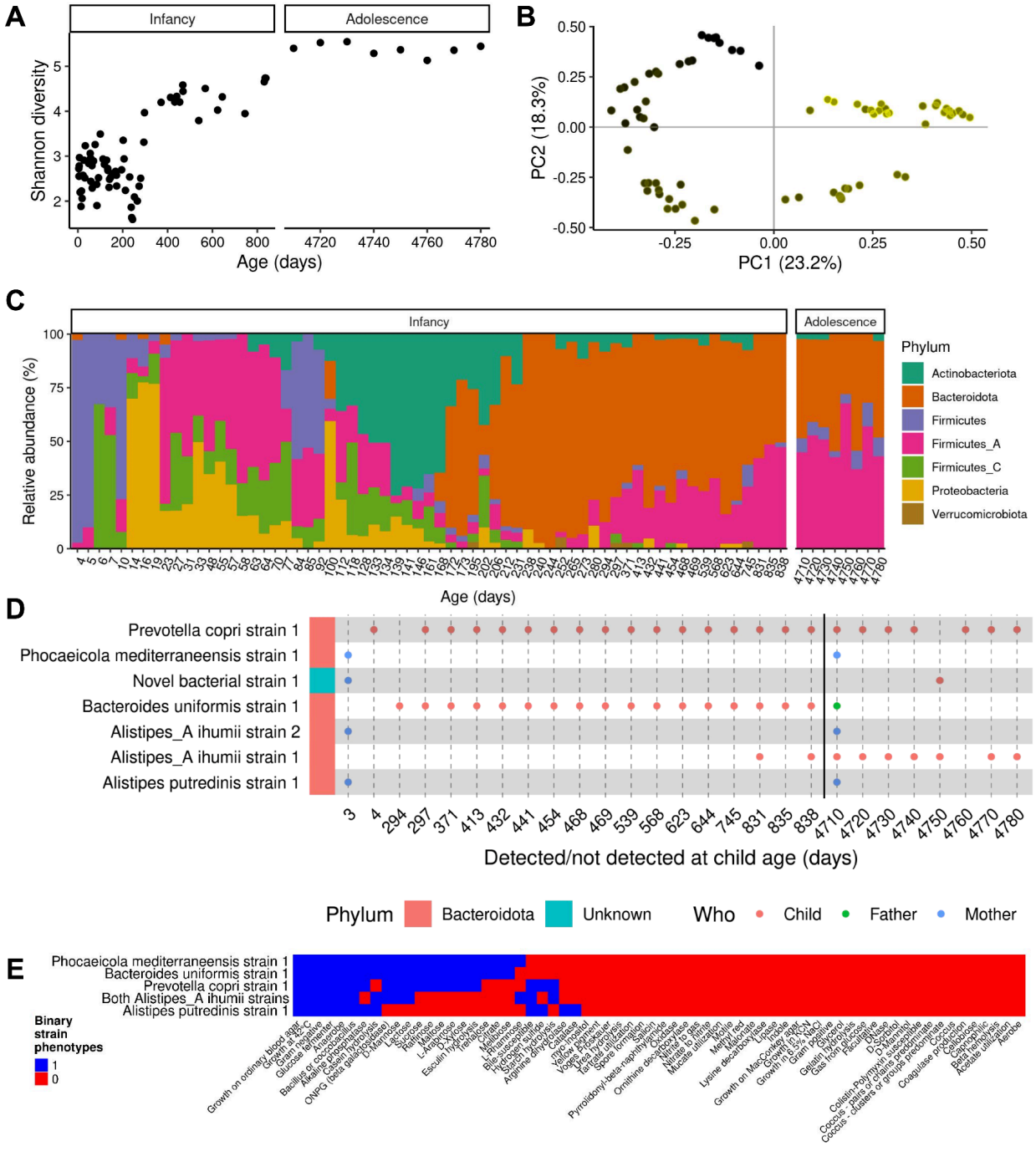

Figure 1. Microbial diversity, taxonomic composition, persistent strains and their predicted phenotypes in a human gut over 13 years. (A) Metagenomic Shannon diversity over time. (B)

Principal component analysis of unweighted Bray Curtis distances, and yellow and black circles depicting the youngest to oldest metagenomes from infancy to adolescence, respectively. (C) Phylum-level composition of the gut microbiome over time. (D) Microbial strains acquired during infancy detected in at least one adolescent time point. (E) Predicted phenotypes of the six strains shared between infancy and adolescence, for which there exists a GTDB taxonomic assignment.

Discussion

We sampled the same individual's stool during infancy and adolescence to determine whether microbes persisted in their gut, despite over 13 years having elapsed. This was indeed the case for two microbial strains. The first persistent strain belonged to the species *Prevotella copri*, which we detected as early as on the host's 4th day of life. The next persistent strain was of the species *Bacteroides uniformis*, and was not first detected until day 294. They were both predicted based on their genomes to grow on lactose; they each encoded beta-galactosidase, and the ability to ferment the resulting glucose. The participant breastfed until day 371 of life, implicating breastmilk and its main carbohydrate, lactose, in providing substrate with which these persistent strains could originally survive after colonizing the host's gut environment. Both strains were predicted from their genomes to be neither oxygen-tolerant nor spore-forming, suggesting that they had, and continue to have, a reliance on their human host for survival.

We also collected stool samples from the then-infant's mother that spanned the same 13 years, and identified three strains to have persisted in her gut; one of the species *Phocaeicola mediterraneensis*, the second of *Alistepes_A ihumii*, and the third of *Alistipes putredinis*. Finally, we identified two examples of strain persistence not within the same individual, but rather between individuals of the same family: one *Bacteroides uniformis* strain detected in the infant's gut and that of their father's 12 years later (from whom we only had stool samples collected during his son's adolescence), and the same between the mother and child, for a bacterial strain whose phylum has no known taxonomic classification. The fundamental weakness of this analysis is that we are operating at the limits of detection for these strains, meaning that our false-negative rate for their detection is very high. This is because when calling SNPs for a given SSRG, we need at least 50 % coverage from metagenome reads, and this is very hard to achieve

for microbes that live in very-low abundance in their host. This means that we are, with certainty, severely under-reporting the true temporal-stability of microbial strains, both in their duration of gut persistence, and the number of strains for which that is the case. Future studies can overcome this by enriching shotgun metagenomic libraries for DNA of specific taxa. Despite this methodological shortcoming of our study, it is the first analysis to evaluate strain persistence in the human gut for longer than five years, in either an infant or an adult (Faith et al. 2013; Bäckhed et al. 2015; Yassour et al. 2016; Chu et al. 2017; Vatanen et al. 2018).

Strains that we identified in this study as having persisted in the humans guts for over a decade belonged both to taxa that we recently found to have signal of codiversification with humans (e.g., *Prevotella copri*), and those that we did not (e.g., *Alistipes* and *Bacteroides uniformis*) (Suzuki et al. 2022). This is consistent with our understanding that gut microbial taxa not undergoing codiversification with humans are frequently horizontally transmitted; colonized by so many hosts, that their high prevalence makes them shared across populations, and by extension, are also shared within hosts over time. Strains that were persistent in this study, but for which there was no evidence of codiversification in our previous work (e.g., *Alistipes* and *Bacteroides uniformis*), had genomes phenotyped as being oxygen-intolerant, which suggests that they have other genomic strategies that confer their high inter-individual transmission. Codiversifying strains are on the other hand reliant on vertical transmission, a transmission mode recently shown to be prevalent in *Prevotella copri* (Suzuki et al. 2022), but also on their survival in their host's gut at least until they can be passed to that host's offspring. Here, we could provide evidence to suggest that this is the case for *Prevotella copri*: not only is it vertically-transmitted, but it can persist in the gut until at least adolescence - when the host becomes capable of sexual reproduction - which suggests that this persistence is the second half of the mechanism by which it is codiversifying with humans.

References

Bäckhed, Fredrik, Josefine Roswall, Yangqing Peng, Qiang Feng, Huijue Jia, Petia Kovatcheva-Datchary, Yin Li, et al. 2015. "Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life." *Cell Host & Microbe* 17 (5): 690–703.

Chaumeil, Pierre-Alain, Aaron J. Mussig, Philip Hugenholtz, and Donovan H. Parks. 2019. "GTDB-Tk: A Toolkit to Classify Genomes with the Genome Taxonomy Database." *Bioinformatics* 36 (6): 1925–27.

Chklovski, Alex, Donovan H. Parks, Ben J. Woodcroft, and Gene W. Tyson. 2022. "CheckM2: A Rapid, Scalable and Accurate Tool for Assessing Microbial Genome Quality Using Machine Learning." *bioRxiv*. https://doi.org/10.1101/2022.07.11.499243.

Chu, Derrick M., Jun Ma, Amanda L. Prince, Kathleen M. Antony, Maxim D. Seferovic, and Kjersti M. Aagaard. 2017. "Maturation of the Infant Microbiome Community Structure and Function across Multiple Body Sites and in Relation to Mode of Delivery." *Nature Medicine* 23 (3): 314–26.

Enav, Hagay, Fredrik Bäckhed, and Ruth E. Ley. 2022. "The Developing Infant Gut Microbiome: A Strain-Level View." *Cell Host & Microbe* 30 (5): 627–38.

Faith, Jeremiah J., Janaki L. Guruge, Mark Charbonneau, Sathish Subramanian, Henning Seedorf, Andrew L. Goodman, Jose C. Clemente, et al. 2013. "The Long-Term Stability of the Human Gut Microbiota." *Science* 341 (6141): 1237439.

Karasov, Talia L., Juliana Almario, Claudia Friedemann, Wei Ding, Michael Giolai, Darren Heavens, Sonja Kersten, et al. 2018. "Arabidopsis Thaliana and Pseudomonas Pathogens Exhibit Stable Associations over Evolutionary Timescales." *Cell Host & Microbe* 24 (1): 168–79.e4.

Kennedy, Katherine M., Marcus C. de Goffau, Maria Elisa Perez-Muñoz, Marie-Claire Arrieta, Fredrik Bäckhed, Peer Bork, Thorsten Braun, et al. 2023. "Questioning the Fetal Microbiome Illustrates Pitfalls of Low-Biomass Microbial Studies." *Nature* 613 (7945): 639–49.

Koenig, Jeremy E., Aymé Spor, Nicholas Scalfone, Ashwana D. Fricker, Jesse Stombaugh, Rob Knight, Largus T. Angenent, and Ruth E. Ley. 2011. "Succession of Microbial Consortia in the Developing Infant Gut Microbiome." *Proceedings of the National Academy of Sciences of the United States of America* 108 Suppl 1 (March): 4578–85.

Lou, Yue Clare, Matthew R. Olm, Spencer Diamond, Alexander Crits-Christoph, Brian A. Firek, Robyn Baker, Michael J. Morowitz, and Jillian F. Banfield. 2021. “Infant Gut Strain Persistence Is Associated with Maternal Origin, Phylogeny, and Traits Including Surface Adhesion and Iron Acquisition.” *Cell Reports Medicine* 2 (9): 100393.

Lu, Jennifer, Florian P. Breitwieser, Peter Thielen, and Steven L. Salzberg. 2017. “Bracken: Estimating Species Abundance in Metagenomics Data.” *PeerJ Computer Science* 3 (January): e104.

Olm, Matthew R., Christopher T. Brown, Brandon Brooks, and Jillian F. Banfield. 2017. “dRep: A Tool for Fast and Accurate Genomic Comparisons That Enables Improved Genome Recovery from Metagenomes through de-Replication.” *The ISME Journal* 11 (12): 2864–68.

Olm, Matthew R., Alexander Crits-Christoph, Keith Bouma-Gregson, Brian A. Firek, Michael J. Morowitz, and Jillian F. Banfield. 2021. “inStrain Profiles Population Microdiversity from Metagenomic Data and Sensitively Detects Shared Microbial Strains.” *Nature Biotechnology*, January, 1–10.

Parks, Donovan H., Maria Chuvochina, Christian Rinke, Aaron J. Mussig, Pierre-Alain Chaumeil, and Philip Hugenholtz. 2022. “GTDB: An Ongoing Census of Bacterial and Archaeal Diversity through a Phylogenetically Consistent, Rank Normalized and Complete Genome-Based Taxonomy.” *Nucleic Acids Research* 50 (D1): D785–94.

Suzuki, Taichi A., J. Liam Fitzstevens, Victor T. Schmidt, Hagay Enav, Kelsey E. Huus, Mirabeau Mbong Ngwese, Anne Grießhammer, et al. 2022. “Codiversification of Gut Microbiota with Humans.” *Science* 377 (6612): 1328–32.

Tierney, Braden T., Zhen Yang, Jacob M. Luber, Marc Beaudin, Marsha C. Wibowo, Christina Baek, Eleanor Mehlenbacher, Chirag J. Patel, and Aleksandar D. Kostic. 2019. “The Landscape of Genetic Content in the Gut and Oral Human Microbiome.” *Cell Host & Microbe* 26 (2): 283–95.e8.

Vatanen, Tommi, Eric A. Franzosa, Randall Schwager, Surya Tripathi, Timothy D. Arthur, Kendra Vehik, Åke Lernmark, et al. 2018. “The Human Gut Microbiome in Early-Onset Type 1 Diabetes from the TEDDY Study.” *Nature* 562 (7728): 589–94.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the Tidyverse.” *Journal of Open*

Source Software 4 (43): 1686.

Willyard, Cassandra. 2018. “New Human Gene Tally Reignites Debate.” *Nature* 558 (7710): 354–55.

Yassour, Moran, Tommi Vatanen, Heli Siljander, Anu-Maaria Hämäläinen, Taina Härkönen, Samppa J. Ryhänen, Eric A. Franzosa, et al. 2016. “Natural History of the Infant Gut Microbiome and Impact of Antibiotic Treatment on Bacterial Strain Diversity and Stability.” *Science Translational Medicine* 8 (343): 343ra81.

Youngblut, Nicholas D., Jacobo de la Cuesta-Zuluaga, Georg H. Reischer, Silke Dauser, Nathalie Schuster, Chris Walzer, Gabrielle Stalder, Andreas H. Farnleitner, and Ruth E. Ley. 2020. “Large-Scale Metagenome Assembly Reveals Novel Animal-Associated Microbial Genomes, Biosynthetic Gene Clusters, and Other Genetic Diversity.” *mSystems* 5 (6). https://doi.org/10.1128/mSystems.01045-20.

Youngblut, Nicholas D., and Ruth E. Ley. 2021. “Struo2: Efficient Metagenome Profiling Database Construction for Ever-Expanding Microbial Genome Datasets.” *PeerJ* 9 (September): e12198.

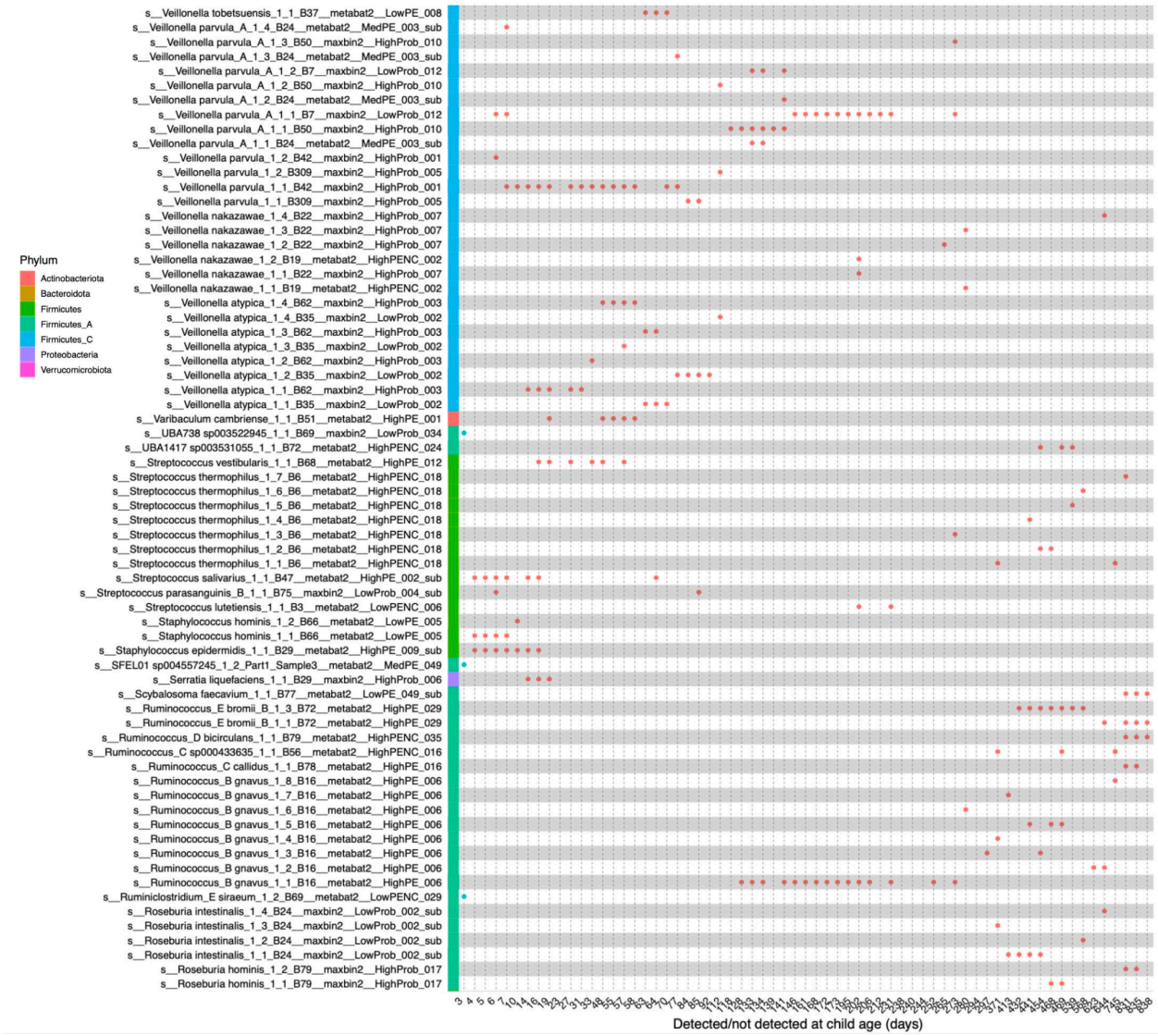

Figure S1. Microbial strains detected only during the child's infant sampling period (first panel of four). Each individual strain is given as a row, with the time points we identified them in indicated by circles, colored by the participant in whose gut microbiome we detected it (red = infant; blue = mother). The phylum of each strain is color-coded in the first column, and the name of the species to which the strain belongs is given in its name.

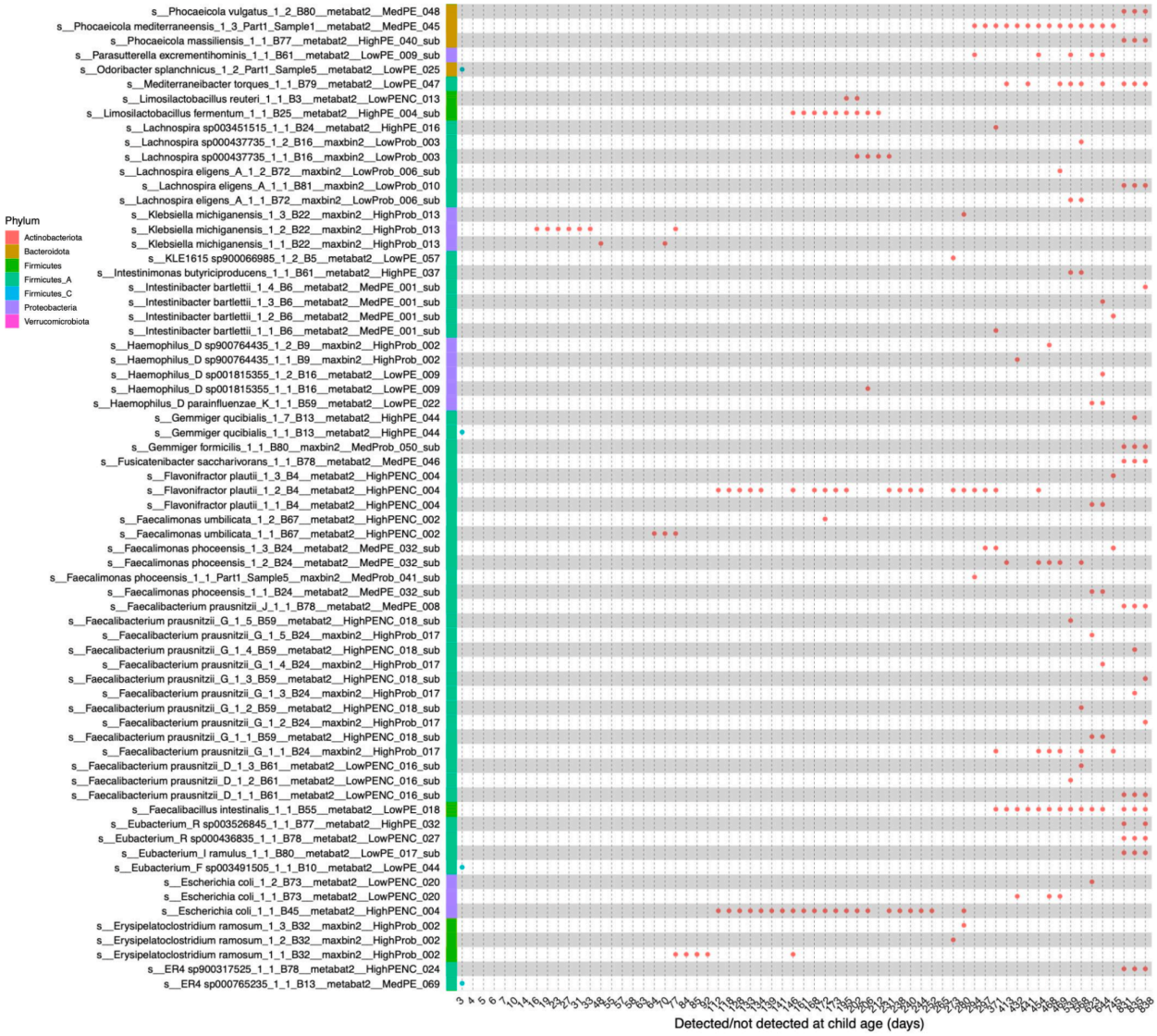

Figure S1. Microbial strains detected only during the child's infant sampling period (second panel of four). See first panel for figure legend.

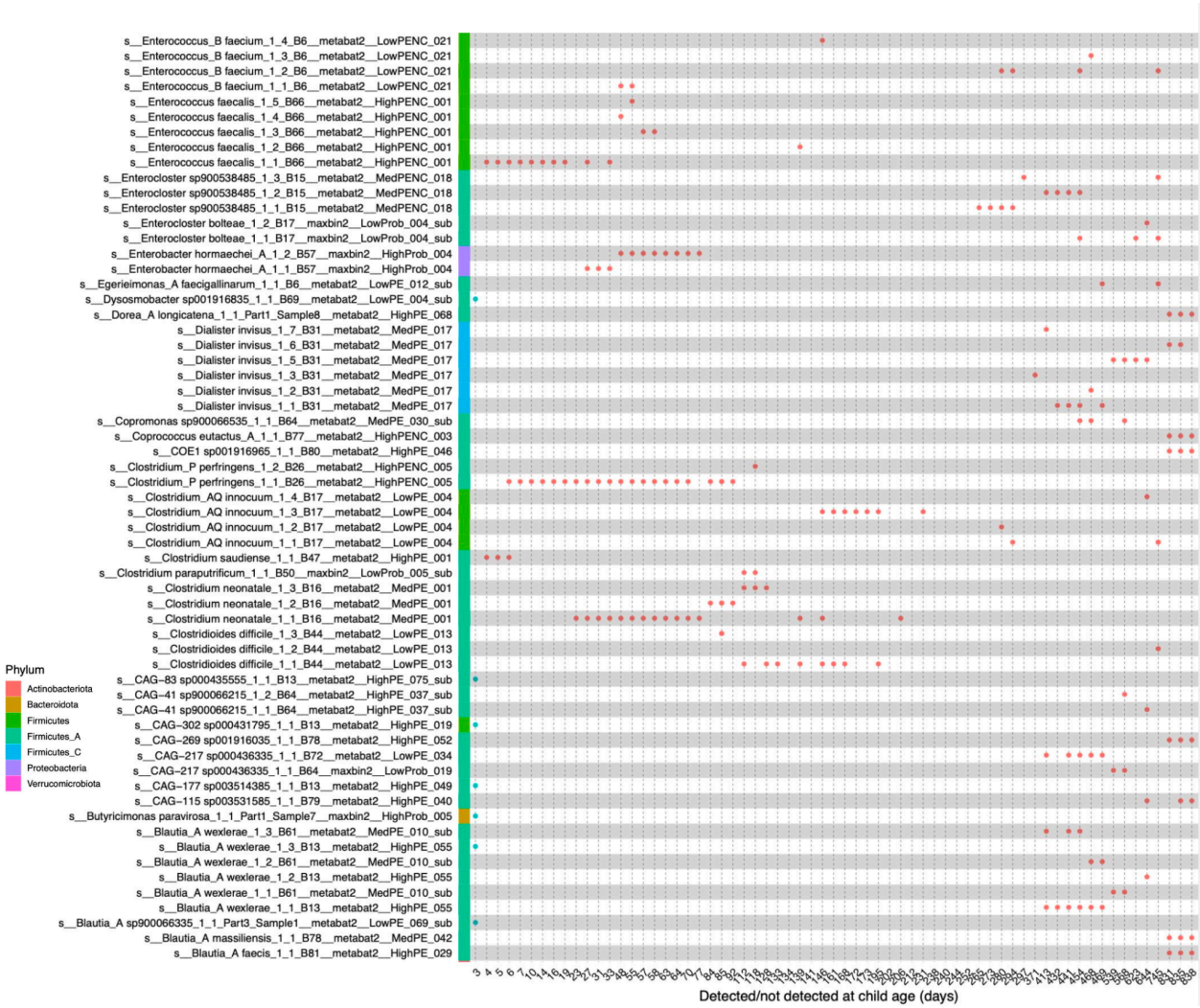

Figure S1. Microbial strains detected only during the child's infant sampling period (third panel of four). See first panel for figure legend.

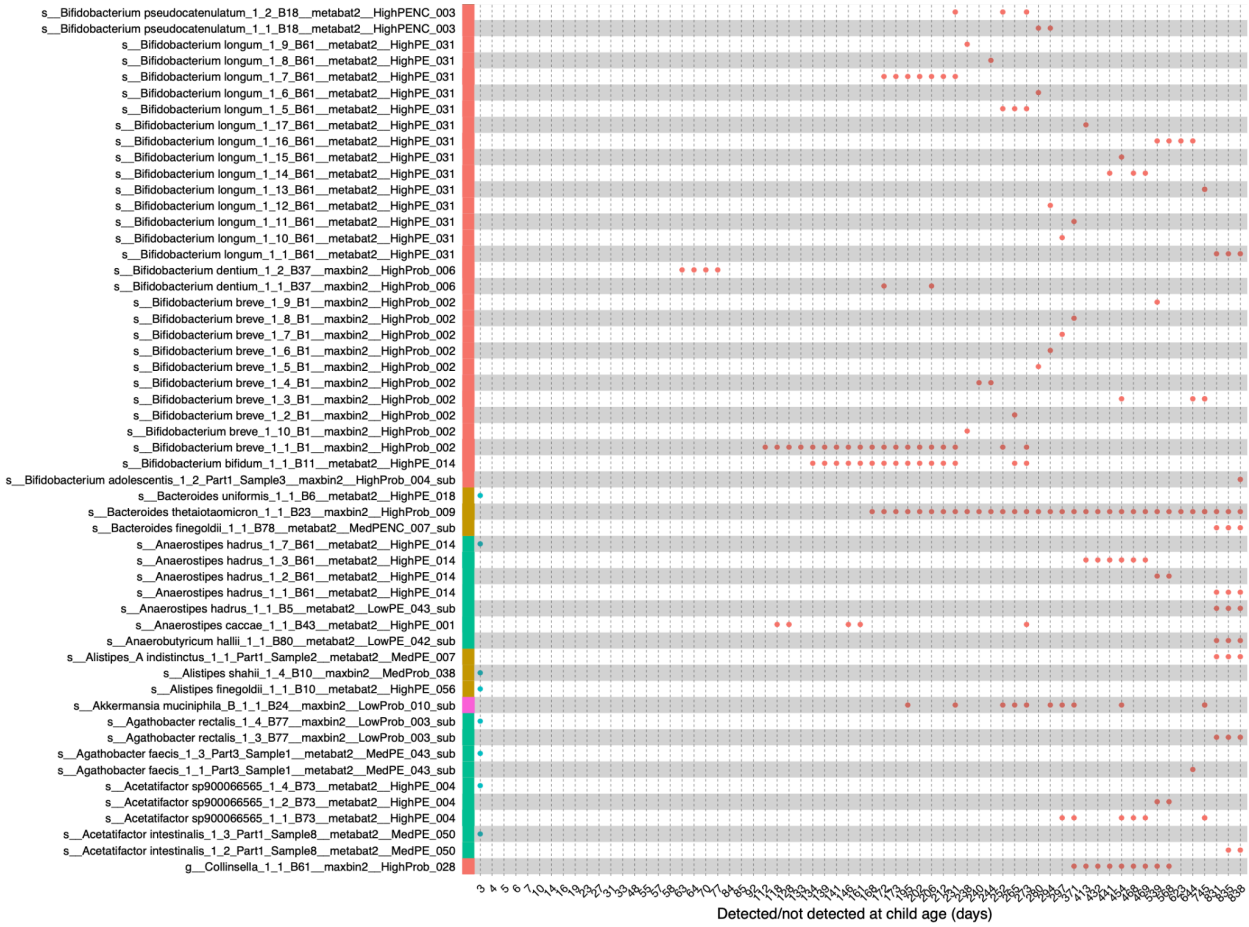

Figure S1. Microbial strains detected only during the child's infant sampling period (fourth panel of four). See first panel for figure legend.

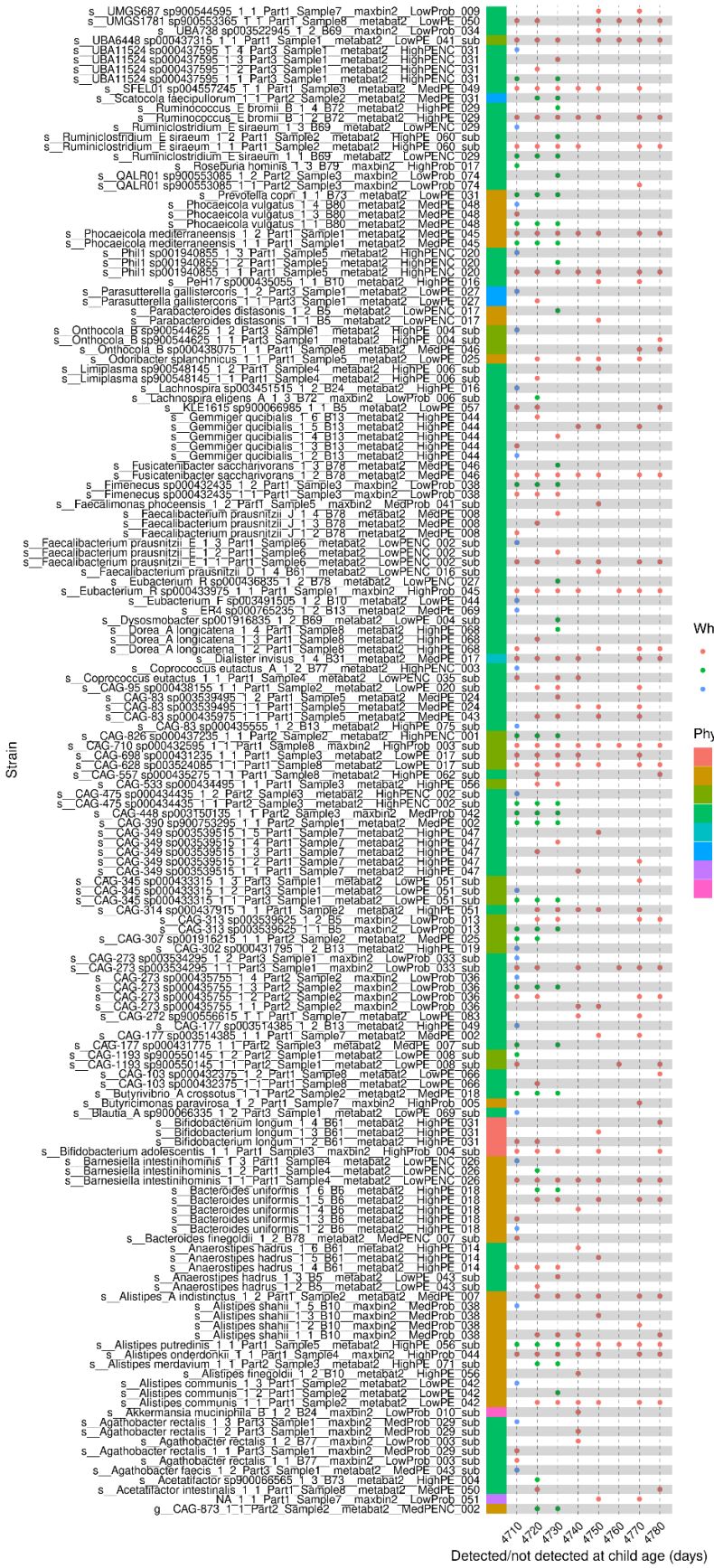
Strain
s__UMGS687 sp900544595_1_1_Part1_Sample7__maxbin2__LowProb_009
s__UMGS1781 sp900553365_1_1_Part1_Sample8__metabat2__LowPE_050
s__UBA738 sp003522945_1_2_B69__maxbin2__LowProb_034
s__UBA6448 sp000437315_1_1_Part1_Sample1__metabat2__LowPE_041_sub
s__UBA11524 sp000437595_1_4_Part3_Sample1__metabat2__HighPENC_031
s__UBA11524 sp000437595_1_3_Part3_Sample1__metabat2__HighPENC_031
s__UBA11524 sp000437595_1_2_Part3_Sample1__metabat2__HighPENC_031
s__UBA11524 sp000437595_1_1_Part3_Sample1__metabat2__HighPENC_031
s__SFEL01 sp004557245_1_1_Part1_Sample3__metabat2__MedPE_049
s__Scatocola faecipullorum_1_1_Part2_Sample2__metabat2__MedPE_031
s__Ruminococcus_E bromii_B_1_4_B72__metabat2__HighPE_029
s__Ruminococcus_E bromii_B_1_2_B72__metabat2__HighPE_029
s__Ruminiclostridium_E siraeum_1_3_B69__metabat2__LowPENC_029
s__Ruminiclostridium_E siraeum_1_2_Part1_Sample2__metabat2__HighPE_060_sub
s__Ruminiclostridium_E siraeum_1_1_Part1_Sample2__metabat2__HighPE_060_sub
s__Ruminiclostridium_E siraeum_1_1_B69__metabat2__LowPENC_029
s__Roseburia hominis_1_3_B79__maxbin2__HighProb_017
s__QALR01 sp900553085_1_2_Part2_Sample3__maxbin2__LowProb_074
s__QALR01 sp900553085_1_1_Part2_Sample3__maxbin2__LowProb_074
s__Prevotella copri_1_1_B73__metabat2__LowPE_031
s__Phocaeicola vulgatus_1_4_B80__metabat2__MedPE_048
s__Phocaeicola vulgatus_1_3_B80__metabat2__MedPE_048
s__Phocaeicola vulgatus_1_1_B80__metabat2__MedPE_048
s__Phocaeicola mediterraneensis_1_2_Part1_Sample1__metabat2__MedPE_045
s__Phocaeicola mediterraneensis_1_1_Part1_Sample1__metabat2__MedPE_045
s__Phil1 sp001940855_1_3_Part1_Sample5__metabat2__HighPENC_020
s__Phil1 sp001940855_1_2_Part1_Sample5__metabat2__HighPENC_020
s__Phil1 sp001940855_1_1_Part1_Sample5__metabat2__HighPENC_020
s__PeH17 sp000435055_1_1_B10__metabat2__HighPE_016
s__Parasutterella gallistercoris_1_2_Part3_Sample1__metabat2__LowPE_027
s__Parasutterella gallistercoris_1_1_Part3_Sample1__metabat2__LowPE_027
s__Parabacteroides distasonis_1_2_B5__metabat2__LowPENC_017
s__Parabacteroides distasonis_1_1_B5__metabat2__LowPENC_017
s__Onthocola_B sp900544625_1_2_Part3_Sample1__metabat2__HighPE_004_sub
s__Onthocola_B sp900544625_1_1_Part3_Sample1__metabat2__HighPE_004_sub
s__Onthocola_B sp000435075_1_1_Part1_Sample8__metabat2__MedPE_046
s__Odoribacter splanchnicus_1_1_Part1_Sample5__metabat2__LowPE_025
s__Limiplasma sp900548145_1_2_Part1_Sample4__metabat2__HighPE_006_sub
s__Limiplasma sp900548145_1_1_Part1_Sample4__metabat2__HighPE_006_sub
s__Lachnospira eligens_A_1_3_B72__maxbin2__LowProb_006_sub
s__Lachnospira sp003451515_1_2_B24__metabat2__HighPE_016
s__KLE1615 sp900066985_1_1_B5__metabat2__LowPE_057
s__Gemmiger qucibialis_1_6_B13__metabat2__HighPE_044
s__Gemmiger qucibialis_1_5_B13__metabat2__HighPE_044
s__Gemmiger qucibialis_1_4_B13__metabat2__HighPE_044
s__Gemmiger qucibialis_1_3_B13__metabat2__HighPE_044
s__Gemmiger qucibialis_1_2_B13__metabat2__HighPE_044
s__Fusicatenibacter saccharivorans_1_3_B78__metabat2__MedPE_046
s__Fusicatenibacter saccharivorans_1_2_B78__metabat2__MedPE_046
s__Fimenecus sp000432435_1_2_Part1_Sample3__maxbin2__LowProb_038
s__Fimenecus sp000432435_1_1_Part1_Sample3__maxbin2__LowProb_038
s__Faecalimonas phoceensis_1_2_Part1_Sample5__maxbin2__MedProb_041_sub
s__Faecalibacterium prausnitzii_J_1_4_B78__metabat2__MedPE_008
s__Faecalibacterium prausnitzii_J_1_3_B78__metabat2__MedPE_008
s__Faecalibacterium prausnitzii_J_1_2_B78__metabat2__MedPE_008
s__Faecalibacterium prausnitzii_E_1_3_Part1_Sample6__metabat2__LowPENC_002_sub
s__Faecalibacterium prausnitzii_E_1_2_Part1_Sample6__metabat2__LowPENC_002_sub
s__Faecalibacterium prausnitzii_E_1_1_Part1_Sample6__metabat2__LowPENC_002_sub
s__Faecalibacterium prausnitzii_D_1_4_B61__metabat2__LowPENC_016_sub
s__Eubacterium_R sp000436835_1_2_B78__metabat2__LowPENC_027
s__Eubacterium_R sp000433975_1_1_Part1_Sample1__maxbin2__HighProb_045
s__Eubacterium_F sp003491505_1_2_B10__metabat2__LowPE_044
s__ER4 sp000765235_1_2_B13__metabat2__MedPE_069
s__Dysosmobacter sp001916835_1_2_B69__metabat2__LowPE_004_sub
s__Dorea_A longicatena_1_4_Part1_Sample8__metabat2__HighPE_068
s__Dorea_A longicatena_1_3_Part1_Sample8__metabat2__HighPE_068
s__Dorea_A longicatena_1_2_Part1_Sample8__metabat2__HighPE_068
s__Dialister invisus_1_4_B31__metabat2__MedPE_017
s__Coprococcus eutactus_A_1_2_B77__metabat2__HighPENC_003
s__Coprococcus eutactus_1_1_Part1_Sample4__metabat2__LowPENC_035_sub
s__CAG-95 sp000438155_1_1_Part1_Sample2__metabat2__LowPE_020_sub
s__CAG-83 sp003539495_1_2_Part1_Sample5__metabat2__MedPE_024
s__CAG-83 sp003539495_1_1_Part1_Sample5__metabat2__MedPE_024
s__CAG-83 sp000435975_1_1_Part1_Sample5__metabat2__MedPE_043
s__CAG-83 sp000435555_1_2_B13__metabat2__HighPE_075_sub
s__CAG-826 sp000437235_1_1_Part2_Sample2__metabat2__HighPENC_001
s__CAG-710 sp000432595_1_1_Part1_Sample8__maxbin2__HighProb_003_sub
s__CAG-698 sp000431235_1_1_Part1_Sample3__metabat2__LowPE_017_sub
s__CAG-628 sp003524085_1_1_Part1_Sample8__metabat2__LowPE_017_sub
s__CAG-557 sp000435275_1_1_Part1_Sample8__metabat2__HighPE_062_sub
s__CAG-533 sp000434495_1_1_Part1_Sample3__metabat2__HighPE_056
s__CAG-475 sp000434435_1_2_Part2_Sample3__metabat2__HighPENC_002_sub
s__CAG-475 sp000434435_1_1_Part2_Sample3__metabat2__HighPENC_002_sub
s__CAG-448 sp003150135_1_1_Part2_Sample3__maxbin2__MedProb_042
s__CAG-390 sp900753295_1_1_Part2_Sample1__metabat2__MedPE_002
s__CAG-349 sp003539515_1_5_Part1_Sample7__metabat2__HighPE_047
s__CAG-349 sp003539515_1_4_Part1_Sample7__metabat2__HighPE_047
s__CAG-349 sp003539515_1_3_Part1_Sample7__metabat2__HighPE_047
s__CAG-349 sp003539515_1_2_Part1_Sample7__metabat2__HighPE_047
s__CAG-349 sp003539515_1_1_Part1_Sample7__metabat2__HighPE_047
s__CAG-345 sp000433315_1_3_Part3_Sample1__metabat2__LowPE_051_sub
s__CAG-345 sp000433315_1_2_Part3_Sample1__metabat2__LowPE_051_sub
s__CAG-345 sp000433315_1_1_Part3_Sample1__metabat2__LowPE_051_sub
s__CAG-314 sp000437915_1_1_Part1_Sample2__metabat2__HighPE_051
s__CAG-313 sp003539625_1_2_B5__maxbin2__LowProb_013
s__CAG-313 sp003539625_1_1_B5__maxbin2__LowProb_013
s__CAG-307 sp001916215_1_1_Part2_Sample2__metabat2__MedPE_025
s__CAG-302 sp000431795_1_2_B13__metabat2__HighPE_019
s__CAG-273 sp003534295_1_2_Part3_Sample1__maxbin2__LowProb_033_sub
s__CAG-273 sp003534295_1_1_Part3_Sample1__maxbin2__LowProb_033_sub
s__CAG-273 sp000435755_1_4_Part2_Sample2__maxbin2__LowProb_036
s__CAG-273 sp000435755_1_3_Part2_Sample2__maxbin2__LowProb_036
s__CAG-273 sp000435755_1_2_Part2_Sample2__maxbin2__LowProb_036
s__CAG-273 sp000435755_1_1_Part2_Sample2__maxbin2__LowProb_036
s__CAG-272 sp900556615_1_1_Part1_Sample7__metabat2__LowPE_083
s__CAG-177 sp003514385_1_2_B13__metabat2__HighPE_049
s__CAG-177 sp003514385_1_1_Part1_Sample7__metabat2__MedPE_002
s__CAG-177 sp000431775_1_1_Part2_Sample3__metabat2__MedPE_007_sub
s__CAG-1193 sp900550145_1_2_Part2_Sample1__metabat2__LowPE_008_sub
s__CAG-1193 sp900550145_1_1_Part2_Sample1__metabat2__LowPE_008_sub
s__CAG-103 sp000432375_1_2_Part1_Sample8__metabat2__LowPE_066
s__CAG-103 sp000432375_1_1_Part1_Sample8__metabat2__LowPE_066
s__Butyrivibrio_A crossotus_1_1_Part2_Sample2__metabat2__MedPE_018
s__Butyricimonas paravirosa_1_2_Part1_Sample7__maxbin2__HighProb_005
s__Blautia_A sp900066335_1_2_Part3_Sample1__metabat2__LowPE_069_sub
s__Bifidobacterium longum_1_4_B61__metabat2__HighPE_031
s__Bifidobacterium longum_1_3_B61__metabat2__HighPE_031
s__Bifidobacterium longum_1_2_B61__metabat2__HighPE_031
s__Bifidobacterium adolescentis_1_1_Part1_Sample3__maxbin2__HighProb_004_sub
s__Barnesiella intestinihominis_1_3_Part1_Sample4__metabat2__LowPENC_026
s__Barnesiella intestinihominis_1_2_Part1_Sample4__metabat2__LowPENC_026
s__Barnesiella intestinihominis_1_1_Part1_Sample4__metabat2__LowPENC_026
s__Bacteroides uniformis_1_6_B6__metabat2__HighPE_018
s__Bacteroides uniformis_1_5_B6__metabat2__HighPE_018
s__Bacteroides uniformis_1_4_B6__metabat2__HighPE_018
s__Bacteroides uniformis_1_3_B6__metabat2__HighPE_018
s__Bacteroides uniformis_1_2_B6__metabat2__HighPE_018
s__Bacteroides finegoldii_1_2_B78__metabat2__MedPENC_007_sub
s__Anaerostipes hadrus_1_6_B61__metabat2__HighPE_014
s__Anaerostipes hadrus_1_5_B61__metabat2__HighPE_014
s__Anaerostipes hadrus_1_4_B61__metabat2__HighPE_014
s__Anaerostipes hadrus_1_3_B5__metabat2__LowPE_043_sub
s__Anaerostipes hadrus_1_2_B5__metabat2__LowPE_043_sub
s__Alistipes_A indistinctus_1_2_Part1_Sample2__metabat2__MedPE_007
s__Alistipes shahii_1_5_B10__maxbin2__MedProb_038
s__Alistipes shahii_1_3_B10__maxbin2__MedProb_038
s__Alistipes shahii_1_2_B10__maxbin2__MedProb_038
s__Alistipes shahii_1_1_B10__maxbin2__MedProb_038
s__Alistipes putredinis_1_1_Part1_Sample5__metabat2__HighPE_056_sub
s__Alistipes onderdonkii_1_1_Part1_Sample4__maxbin2__HighProb_044
s__Alistipes merdavium_1_1_Part2_Sample3__metabat2__HighPE_071_sub
s__Alistipes finegoldii_1_2_B10__metabat2__HighPE_056
s__Alistipes communis_1_3_Part1_Sample2__metabat2__LowPE_042
s__Alistipes communis_1_2_Part1_Sample2__metabat2__LowPE_042
s__Alistipes communis_1_1_Part1_Sample2__metabat2__LowPE_042
s__Akkermansia muciniphila_B_1_2_B24__maxbin2__LowProb_010_sub
s__Agathobacter rectalis_1_3_Part3_Sample1__maxbin2__MedProb_029_sub
s__Agathobacter rectalis_1_2_Part3_Sample1__maxbin2__MedProb_029_sub
s__Agathobacter rectalis_1_2_B77__maxbin2__LowProb_003_sub
s__Agathobacter rectalis_1_1_Part3_Sample1__maxbin2__MedProb_029_sub
s__Agathobacter rectalis_1_1_B77__maxbin2__LowProb_003_sub
s__Agathobacter faecis_1_2_Part3_Sample1__metabat2__MedPE_043_sub
s__Acetatifactor sp900066565_1_3_B73__metabat2__HighPE_004
s__Acetatifactor intestinalis_1_1_Part1_Sample8__metabat2__MedPE_050
NA_1_1_Part1_Sample7__maxbin2__LowProb_051
g__CAG-873_1_1_Part2_Sample2__metabat2__MedPENC_002
4710
4720
4730
4740
4750
4760
4770
4780
Detected/not detected at child age (days)
Who
Child
Father
Mother
Phylum
Actinobacteriota
Bacteroidota
Firmicutes
Firmicutes_A
Firmicutes_C
Proteobacteria
Unknown
Verrucomicrobiota

Figure S2. Microbial strains detected only during the child's adolescent sampling period. Each individual strain is given as a row, with the time points we identified them in indicated by circles, colored by the participant in whose gut microbiome we detected it. The phylum of each strain is color-coded in the first column, and the name of the species to which the strain belongs is given in its name.

Discussion

Taken together, these three projects describe a very intimate relationship between humans and their gut microbial symbionts. We found that gut microbes can confer a tolerance to lactose which has, until now, been thought to be conferred only by the host's own genome. This microbially-derived phenotype is 100% tolerant from the perspective of hydrogen production - typically the sole focus of medical doctors' diagnoses of patients - because LNPs enriched for *Bifidobacterium* can make no hydrogen from lactose just as LPs do. The energetic consequence for the hosts are however different, because the energetic yield from glucose absorbed in the small intestine (genetic tolerance) is greater than the ATP the host can make after gut-uptake of the *Bifidobacterium*-derived acetate (microbially-acquired tolerance). In other words, hydrogen and symptoms meet tolerance criteria, but the host-energy derived from lactose is only intermediary in microbially-acquired tolerance. This means that early in the cultural-adoption of dairying, there was likely not only positive selection for LP, but also LNPs enriched with *Bifidobacterium*. Selection was likely so strong that LNPs without sufficient *Bifidobacterium* for microbial-tolerance died during starvation events. Studies of how humans have adapted to the cultural practice of dairying have thus far centered on their own genetic adaptation; this study reveals that human-only frameworks do not accurately reflect the biology of selection.

We are the first to study vertical gut microbiome transmission in both western and non-western populations under the same project. This is important because the ways in which metagenomic libraries are made, sequenced, and their resulting metagenomes bioinformatically-processed and analyzed has a large impact on outcome, which renders comparing the results between studies of little utility when we are concerned about where SNPs are located at resolutions of 99.999% ANI. We controlled for these confounders in our study, giving us the unprecedented opportunity to directly compare dynamics of human gut microbe transmission in western and non-western populations. Consistent with other studies, mothers are a significant source of microbes for their infant. However, unlike in Gabon and Vietnam, where they were prevalent, there was not a single strain-sharing event between unrelated mothers and infants in Germany. This was the case despite having conducted all of our sampling in the same German town (as compared to the multiple towns of the Gabonese and Vietnamese strain-sharing datasets), meaning we over-sampled German mother-infant pairs living closeby, thereby grossly

inflating the probability of detecting unrelated strain-sharing in that country. We nonetheless detected zero.

This suggests that the codiversification patterns we observed in Gabon and Vietnam were driven by both horizontal- and vertical-transmission, whereas they were driven solely by vertical transmission in Germany. This is in alignment with basic ecological theory, which posits that environments with lower dispersal limitation (i.e., non-western ones) exhibit fewer island dynamics (Koskella, Hall, and Metcalf 2017). There is, however, a technical limitation to this claim of such magnitude, that it may be rendered invalid: we were operating at the limits of strain-detection. This is because despite advances in the sensitivity of strain-resolved metagenomics, the state-of-the-art has been constrained for years by the fact that taxa abundant in adults are present at very low abundance in infants, and vice versa. This is in part the outcome of a trend for preference of next-generation-sequencing- over culture-based approaches in this sub-field (Fitzstevens et al. 2017). We have, however, reached the point where short-read metagenomics cannot garner sufficient coverage on very-low-abundance representative genomes to enable robust analysis. Deeper sequencing should in theory solve this problem, but the detection of species in our dataset already saturated at 3 million paired-end reads; we sequenced beyond that, and nonetheless found no unrelated mother-infant strain sharing in Germany. Long-read sequencing may suffer from the same problem; reverting to isolating strains in vitro is the best way to guarantee we don't have false negatives for unrelated strain-sharing in Germany. Enriching metagenome libraries with hybridization is however a molecular technique that should be developed further, as it could work in both short- and long-read systems, where with efficient probe-binding, it could be more easily deployed at scale in high-throughput workflows.

Operating at the limit of strain-detection was the fundamental weakness of not only the strain-sharing project in Gabon, Vietnam and Germany, but also the longitudinal dataset, because just as the gut ecosystems of mothers and their infants are so different, so too were those of the longitudinally-studied participant's infant and adolescent guts. inStrain's false positive rate is very low, but we cannot conclude that representative genomes with insufficient coverage in metagenomes did not persist in the longitudinal dataset. In this sense, this was truly a discovery dataset; one with which we could ask the question whether any strains persisted at all. My aforementioned outlook for the field of strain-resolved metagenomics - to rely again on in vitro culture directly from stool and to develop reliable hybridization protocols from stool

metagenomes - will enable future researchers of both transmission- and persistence-questions to conduct studies with not only a low false-positive rate, as we enjoyed here, but also a low false-negative rate. This will further strengthen our understanding of how gut microbes persist both within individuals, and also within lineages of human hosts over time, which we identified here as being the foundation of their ability to confer phenotypes once thought to be conferrable only by host genomes.

References

Asnicar, Francesco, Serena Manara, Moreno Zolfo, Duy Tin Truong, Matthias Scholz, Federica Armanini, Pamela Ferretti, et al. 2017. "Studying Vertical Microbiome Transmission from Mothers to Infants by Strain-Level Metagenomic Profiling." *mSystems* 2 (1). https://doi.org/10.1128/mSystems.00164-16.

Bäckhed, Fredrik, Josefine Roswall, Yangqing Peng, Qiang Feng, Huijue Jia, Petia Kovatcheva-Datchary, Yin Li, et al. 2015. "Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life." *Cell Host & Microbe* 17 (5): 690–703.

Beghini, Francesco, Lauren J. McIver, Aitor Blanco-Míguez, Leonard Dubois, Francesco Asnicar, Sagun Maharjan, Ana Mailyan, et al. 2021. "Integrating Taxonomic, Functional, and Strain-Level Profiling of Diverse Microbial Communities with bioBakery 3." *eLife* 10 (May). https://doi.org/10.7554/eLife.65088.

Beyer, Robert M., Mario Krapp, Anders Eriksson, and Andrea Manica. 2021. "Climatic Windows for Human Migration out of Africa in the Past 300,000 Years." *Nature Communications* 12 (1): 4889.

Blekhman, Ran, Julia K. Goodrich, Katherine Huang, Qi Sun, Robert Bukowski, Jordana T. Bell, Timothy D. Spector, et al. 2015. "Host Genetic Variation Impacts Microbiome Composition across Human Body Sites." *Genome Biology* 16 (September): 191.

Bonder, Marc Jan, Alexander Kurilshikov, Ettje F. Tigchelaar, Zlatan Mujagic, Floris Imhann, Arnau Vich Vila, Patrick Deelen, et al. 2016. "The Effect of Host Genetics on the Gut Microbiome." *Nature Genetics* 48 (11): 1407–12.

Chu, Derrick M., Jun Ma, Amanda L. Prince, Kathleen M. Antony, Maxim D. Seferovic, and Kjersti M. Aagaard. 2017. "Maturation of the Infant Microbiome Community Structure and Function across Multiple Body Sites and in Relation to Mode of Delivery." *Nature Medicine* 23 (3): 314–26.

Everard, Amandine, Clara Belzer, Lucie Geurts, Janneke P. Ouwerkerk, Céline Druart, Laure B. Bindels, Yves Guiot, et al. 2013. "Cross-Talk between Akkermansia Muciniphila and Intestinal Epithelium Controls Diet-Induced Obesity." *Proceedings of the National Academy of Sciences of the United States of America* 110 (22): 9066–71.

Fan, Shaohua, Matthew E. B. Hansen, Yancy Lo, and Sarah A. Tishkoff. 2016. "Going Global by Adapting Local: A Review of Recent Human Adaptation." *Science* 354 (6308): 54–59.

Fitzstevens, John L., Kelsey C. Smith, James I. Hagadorn, Melissa J. Caimano, Adam P. Matson, and Elizabeth A. Brownell. 2017. “Systematic Review of the Human Milk Microbiota.” *Nutrition in Clinical Practice: Official Publication of the American Society for Parenteral and Enteral Nutrition* 32 (3): 354–64.

Goodrich, Julia K., Emily R. Davenport, Michelle Beaumont, Matthew A. Jackson, Rob Knight, Carole Ober, Tim D. Spector, Jordana T. Bell, Andrew G. Clark, and Ruth E. Ley. 2016. “Genetic Determinants of the Gut Microbiome in UK Twins.” *Cell Host & Microbe* 19 (5): 731–43.

Goodrich, Julia K., Emily R. Davenport, Andrew G. Clark, and Ruth E. Ley. 2017. “The Relationship Between the Human Genome and Microbiome Comes into View.” *Annual Review of Genetics* 51 (November): 413–33.

Goodrich, Julia K., Emily R. Davenport, Jillian L. Waters, Andrew G. Clark, and Ruth E. Ley. 2016. “Cross-Species Comparisons of Host Genetic Associations with the Microbiome.” *Science* 352 (6285): 532–35.

Korpela, Katri, Paul Costea, Luis Pedro Coelho, Stefanie Kandels-Lewis, Gonneke Willemsen, Dorret I. Boomsma, Nicola Segata, and Peer Bork. 2018. “Selective Maternal Seeding and Environment Shape the Human Gut Microbiome.” *Genome Research* 28 (4): 561–68.

Koskella, Britt, Lindsay J. Hall, and C. Jessica E. Metcalf. 2017. “The Microbiome beyond the Horizon of Ecological and Evolutionary Theory.” *Nature Ecology & Evolution* 1 (11): 1606–15.

Marciniak, Stephanie, and George H. Perry. 2017. “Harnessing Ancient Genomes to Study the History of Human Adaptation.” *Nature Reviews. Genetics* 18 (11): 659–74.

Mathieson, Iain. 2020. “Human Adaptation over the Past 40,000 Years.” *Current Opinion in Genetics & Development* 62 (June): 97–104.

Moran, Nancy A., and Daniel B. Sloan. 2015. “The Hologenome Concept: Helpful or Hollow?” *PLoS Biology* 13 (12): 1–10.

Nayfach, Stephen, Beltran Rodriguez-Mueller, Nandita Garud, and Katherine S. Pollard. 2016. “An Integrated Metagenomics Pipeline for Strain Profiling Reveals Novel Patterns of Bacterial Transmission and Biogeography.” *Genome Research* 26 (11): 1612–25.

Nurk, Sergey, Sergey Koren, Arang Rhie, Mikko Rautiainen, Andrey V. Bzikadze, Alla Mikheenko, Mitchell R. Vollger, et al. 2022. “The Complete Sequence of a Human

Genome.” *Science* 376 (6588): 44–53.

Olm, Matthew R., Alexander Crits-Christoph, Keith Bouma-Gregson, Brian A. Firek, Michael J. Morowitz, and Jillian F. Banfield. 2021. “inStrain Profiles Population Microdiversity from Metagenomic Data and Sensitively Detects Shared Microbial Strains.” *Nature Biotechnology*, January, 1–10.

Poole, Angela C., Julia K. Goodrich, Nicholas D. Youngblut, Guillermo G. Luque, Albane Ruaud, Jessica L. Sutter, Jillian L. Waters, et al. 2019. “Human Salivary Amylase Gene Copy Number Impacts Oral and Gut Microbiomes.” *Cell Host & Microbe* 25 (4): 553–64.e7.

Shi, Zhou Jason, Boris Dimitrov, Chunyu Zhao, Stephen Nayfach, and Katherine S. Pollard. 2022. “Fast and Accurate Metagenotyping of the Human Gut Microbiome with GT-Pro.” *Nature Biotechnology* 40 (4): 507–16.

Suzuki, Taichi A., and Ruth E. Ley. 2020. “The Role of the Microbiota in Human Genetic Adaptation.” *Science* 370 (6521): eaaz6827.

Tierney, Braden T., Zhen Yang, Jacob M. Luber, Marc Beaudin, Marsha C. Wibowo, Christina Baek, Eleanor Mehlenbacher, Chirag J. Patel, and Aleksandar D. Kostic. 2019. “The Landscape of Genetic Content in the Gut and Oral Human Microbiome.” *Cell Host & Microbe* 26 (2): 283–95.e8.

Vatanen, Tommi, Eric A. Franzosa, Randall Schwager, Surya Tripathi, Timothy D. Arthur, Kendra Vehik, Åke Lernmark, et al. 2018. “The Human Gut Microbiome in Early-Onset Type 1 Diabetes from the TEDDY Study.” *Nature* 562 (7728): 589–94.

Willyard, Cassandra. 2018. “New Human Gene Tally Reignites Debate.” *Nature* 558 (7710): 354–55.

Yassour, Moran, Eeva Jason, Larson J. Hogstrom, Timothy D. Arthur, Surya Tripathi, Heli Siljander, Jenni Selvenius, et al. 2018. “Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life.” *Cell Host & Microbe* 24 (1): 146–54.e4.

Yassour, Moran, Tommi Vatanen, Heli Siljander, Anu-Maaria Hämäläinen, Taina Härkönen, Samppa J. Ryhänen, Eric A. Franzosa, et al. 2016. “Natural History of the Infant Gut Microbiome and Impact of Antibiotic Treatment on Bacterial Strain Diversity and Stability.” *Science Translational Medicine* 8 (343): 343ra81.