Uniform bacterial genetic diversity along the gut

Michael Wasney[1]*, Leah Briscoe[2]*, Richard Wolff[3], Hans Ghezzi[4], Carolina Tropini[5,6,7], Nandita Garud[1,2,3]

[1]University of California, Los Angeles, Human Genetics, Los Angeles, CA, [2]University of California, Los Angeles, Interdepartmental Program in Bioinformatics, Los Angeles, CA, [3]University of California, Los Angeles, Ecology and Evolutionary Biology, Los Angeles, CA, [4] University of British Columbia, Department of Bioinformatics, Vancouver, Canada, [5]University of British Columbia, Department of Microbiology and Immunology, Vancouver, Canada, [6]University of British Columbia, School of Biomedical Engineering, Vancouver, Canada, [7]Canadian Institute for Advanced Research, Humans and the Microbiome Program, Toronto, Canada

*These authors contributed equally to this work.

Abstract

Environmental gradients throughout the digestive tract shape spatial variation in the composition and abundance of bacterial species along the gut. However, much less is known about how genetic diversity *within* bacterial species is distributed along the gut. Understanding this distribution is important because bacterial genetic variants confer traits that influence both microbiome function and host physiology, including local inflammation and nutrient metabolism. Thus, to understand how the microbiome functions at a mechanistic level, it is essential to understand how genetic diversity is organized along the gut. In this study, we profiled genetic diversity of approximately 30 common gut commensal bacteria in five regions along the gut lumen in germ-free mice colonized with the same healthy human stool sample. Although species composition varied considerably along the gut, genetic diversity within species was substantially more uniform. Driving this uniformity were similar strain frequencies along the gut, implying that multiple, genetically divergent strains of the same species can coexist within a host without spatially segregating. Additionally, the approximately 60 unique evolutionary adaptations arising within mice tended to sweep throughout the gut, showing little gut region specificity. We then analyzed metagenomic samples collected along the guts of conventional

mice and healthy humans and found similar dynamics with their natural microbiomes, suggesting that uniform bacterial genetic diversity may be common to multiple host species. Together, our findings demonstrate that uniform spatial distribution of genetic diversity along the gastrointestinal tract is a robust feature of mammalian gut ecosystems.

Introduction

The spatial organization of microbiota along the gut plays a critical role in microbiome function[1]. Previous research in humans and mice has established that different regions of the gut, such as the small and large intestines, harbor distinct assemblages of bacterial species[2,3], and that disruption of this organization is associated with disease[4]. Spatial organization in the microbiome is thought to arise from distinct niches imposed by environmental and physiological gradients along the gut[3,5,6]. However, it remains unknown whether such gradients drive similar spatial organization of within-species bacterial genetic diversity, such as by determining the occupancy of genetically distinct strains or individual genetic variants along the gut. Understanding spatial organization at this level is crucial, as bacterial genetic diversity drives traits that support proper microbiome function and are associated with multiple host-relevant phenotypes[7], including ability to digest food[8], local inflammation in specific regions of the gut[9], antibiotic resistance[10], metabolic capabilities[8,11–14], and pathogen resistance[15].

There are several ecological and evolutionary mechanisms that could result in either i) spatially segregated or ii) well-mixed genetic variation along the gut (**Figure 1A**). Spatial segregation could arise because environmental gradients select for different strains and mutations along the gut, co-colonizing strains occupy distinct physical niches to reduce competition, or slow migration rates limiting the spread of genetic variation. Contrastingly, spatial uniformity could arise because adaptive variants are globally adaptive along the gut, co-colonizing strains occupy distinct non-physical (e.g., metabolic) niches, or rapid migration rates.

Previous studies have sought to examine the spatial distribution of strains and individual nucleotide variants. One study looking at a small number of commensal bacterial species found that different strains can coexist along the human gut[16], but did not determine whether strain frequencies of these co-colonizing strains were variable in different gut regions or whether

57 coexistence was a general phenomenon. Other recent work found that single nucleotide variants
58 (SNVs) can arise locally along the gut for a few bacterial species[16,17], though these studies did
59 not demonstrate that spatially specific evolutionary changes were adaptive or common across
60 bacterial species or host. In another case, a uropathogenic strain of *Enterococcus gallinarum*
61 evolved adaptive mutations specific to the mucosa, which caused inflammation[9]. However, it is
62 unclear how commonly adaptive mutations spatially segregate along the longitudinal axis of the
63 gastrointestinal (GI) tract (i.e., from small to large intestine), particularly among commensal
64 microbes not facing any strong selective perturbations such as antibiotics. Here we will test the
65 hypothesis that strains and their adaptations may also display spatial segregation, similar to the
66 spatial heterogeneity observed at the species level.

67 To date, characterizing the spatial structure of genetic diversity along the mammalian gut has
68 been challenging because of the invasive nature of directly sampling the GI tract, the limited
69 bacterial biomass obtained from such attempts, and because of the difficulties in culturing a
70 broad diversity of taxa found natively in the human gut microbiome. However, these challenges
71 can be circumvented using a combination of experimental and statistical approaches. First, the
72 use of a gnotobiotic mouse model for studying the gut microbiome enables the study of human
73 gut-specific species of bacteria *in vivo* in a controlled manner as a single inoculum can be
74 administered to multiple mice, allowing us to measure variability across replicates[18–21]. Second,
75 the use of mice overcomes the hurdle of invasive sampling in humans as the full gut contents can
76 be extracted after sacrifice and then processed for sequencing. Third, the use of shotgun
77 metagenomic sequencing enables the characterization of genetic diversity across a broad panel of
78 bacterial taxa, including those undetectable using culture-based methods[22]. Finally, recent
79 statistical innovations in phasing genotypes of individual strains from shotgun metagenomic data
80 enables the ability to quantify strain frequencies and their evolutionary modifications[23–25].

81 In this study, we capitalize on these advances to assess the distribution of strains and
82 evolutionary modifications along the gut. Specifically, we analyzed the frequencies of bacterial
83 strains and the evolutionary modifications they carry within luminal samples obtained from
84 along the GI tracts of germ-free mice eight weeks after they were inoculated with the
85 microbiome of a single healthy human (**Figure 1B**). We previously analyzed this same
86 humanized mouse dataset with 16S rRNA sequencing of luminal contents and found significant

87 shifts in abundance of taxonomic family members along the gut[3]. By using metagenomic
88 sequencing to analyze the same samples at a deeper level, we were able to estimate strain
89 frequencies and identify evolutionary modifications. We report that the frequency of strains and
90 evolutionary modifications found in most bacterial species was unexpectedly uniform across the
91 gut. Confirming that this uniformity extends to other natural microbiomes, we applied the same
92 analyses to (i) a cohort of conventional mice with their own naturally acquired, host-adapted
93 microbiomes and (ii) a cohort of healthy humans whose luminal contents from multiple gut
94 regions were sampled using ingestible capsule devices swallowed over several days. In all
95 systems, we observed spatial uniformity of genetic diversity along the GI tract, suggesting that
96 this pattern may be a common feature of multiple mammalian gut microbiomes. These results
97 further imply that species abundance differences along the gut may play a greater role than
98 genetic differences in responding to environmental gradients along the gut, and that this could be
99 a robust feature of gut microbial ecology across both mouse and human hosts.

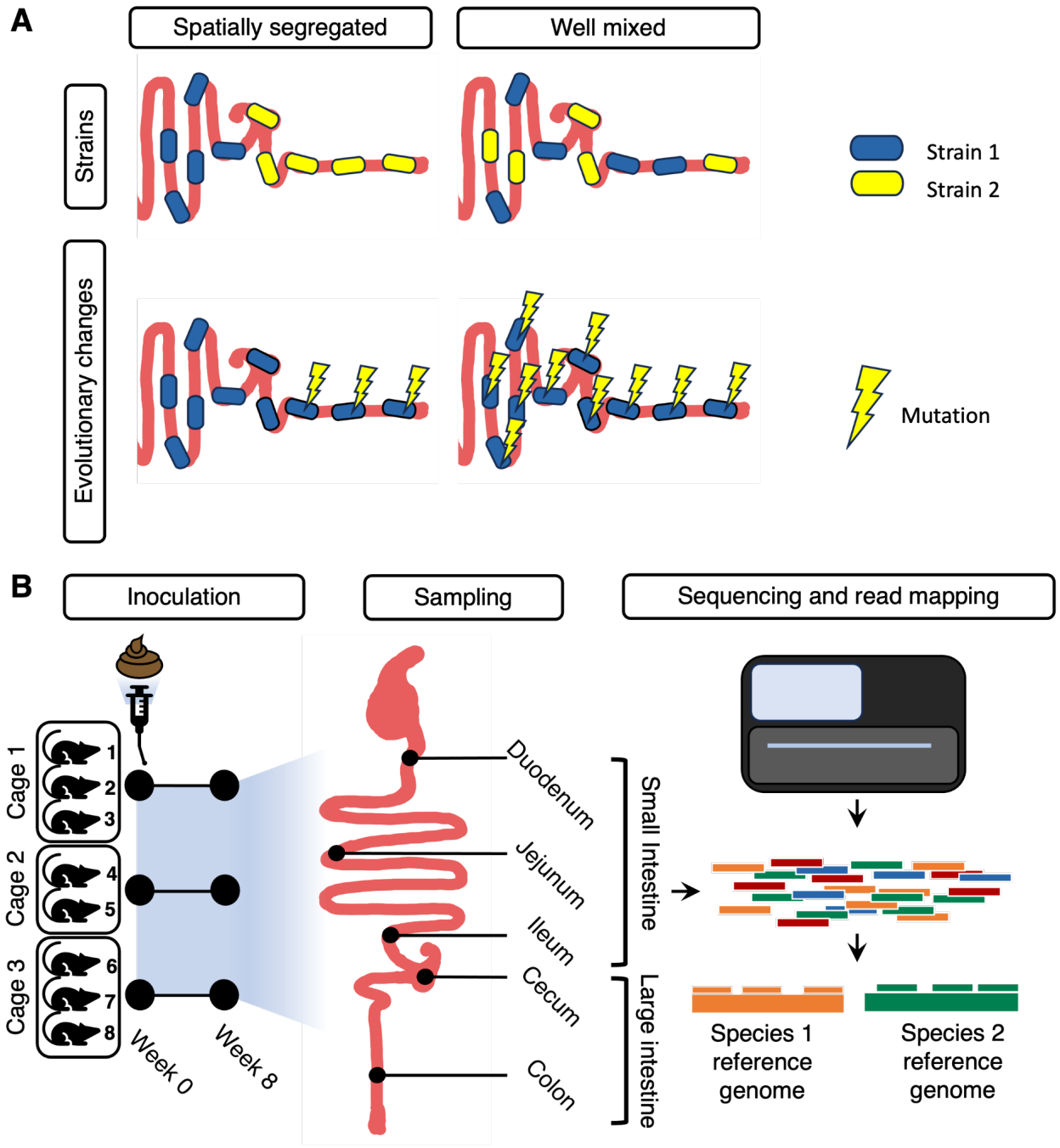


Figure 1. Hypotheses and experimental design. (A) Possible organization of genetic diversity along the gut. Strains and their evolutionary changes may be either spatially segregated, or evenly mixed. **(B)** Experimental design. Eight germ-free Swiss Webster mice were inoculated with a single healthy human-derived stool sample[3]. Subsequently, the microbiota was allowed to equilibrate over eight weeks before shotgun sequencing was performed on samples collected from five regions along the gut (small intestine: duodenum, jejunum, and ileum; large intestine:

107 cecum and colon). The duodenum, jejunum, and ileum are considered regions of the small
108 intestine, and cecum and colon are considered regions of the large intestine. To quantify species
109 abundances and SNV frequencies, metagenomic reads were mapped against a reference database
110 of genomes from almost 6,000 bacterial species.

111 Results

112 To investigate the effect of gut region on microbial strain diversity, we performed shotgun
113 sequencing on microbiomes from eight mice that were orally gavaged germ-free Swiss Webster
114 mice with a single stool sample (hereafter referred to as the inoculum) from a healthy human
115 donor in a previous study[3] (**Figure 1B**). Luminal contents were collected from five intestinal
116 regions in each mouse: the duodenum, jejunum, ileum, cecum, and colon. DNA was extracted
117 and metagenomic sequencing was performed on each of the five mouse intestinal samples as
118 well as the original human inoculum (**Methods**).

119 To infer species abundances and SNV frequencies, we mapped shotgun reads to a database of
120 reference genomes[26] for the most prevalent and abundant bacterial species present in these mice
121 (**Figure 1B**) (see **Methods** for more details on the bioinformatic pipeline used). Summaries of
122 species and genetic-level diversity have been previously quantified from stool with similar
123 approaches (e.g., Nayfach *et al.*, 2016)[26], but here we investigate this diversity along the gut.

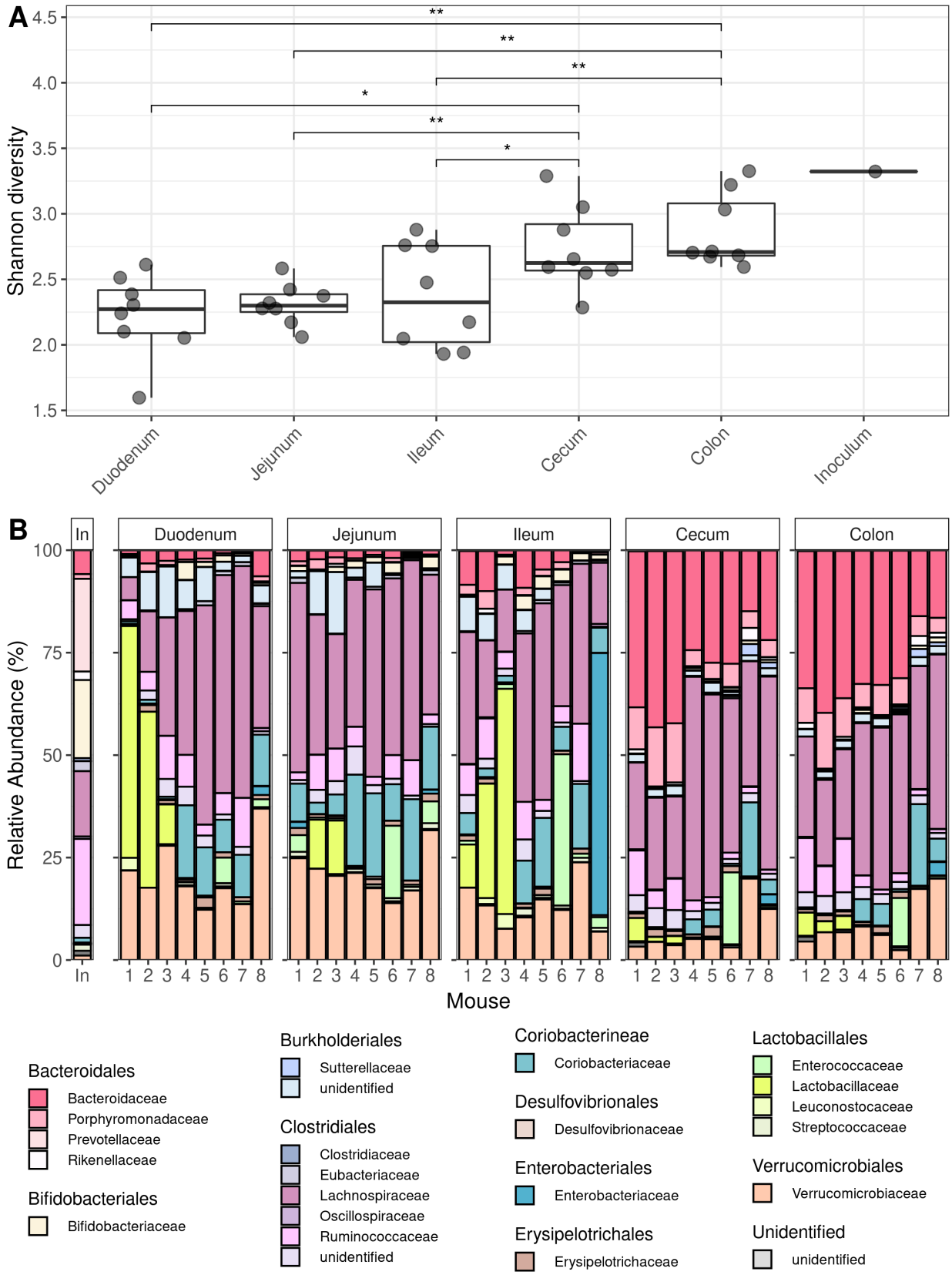
124 Bacterial species and functional diversity differ along the gut

125 First, we examined whether bacterial species diversity, measured using alpha diversity (the
126 Shannon index), varied along the tract of the gut. We hypothesized the large intestine would
127 have higher alpha diversity compared to other regions of the gut, as previous studies in humans[27]
128 and mice[19,20] have found that species diversity is highest in the colon due to lower flow rates and
129 the relatively high abundance of host- and food-derived complex polysaccharides[5,6,28]. As
130 expected, gut region was a major driver of alpha diversity, with the large intestine having
131 significantly higher diversity than all regions in the small intestine (average fold change per
132 mouse = 1.22; Wilcoxon signed-rank test, p-value = 0.0156) (**Figure 2A**).

133 Next, we hypothesized that the relative abundance of individual bacterial community members
134 differ along the gut, as factors such as oxygen availability, antimicrobial peptides, and carbon-
135 source favor different dominant taxa in the mammalian small and large intestines,
136 respectively[1,2,29,30]. To quantify shifts in community composition along the gut, we examined the
137 relative abundance of bacterial families, expecting to see saccharolytic, obligate anaerobic taxa
138 dominate in the large intestine and rapidly dividing facultative anaerobes dominate in the small
139 intestine[2,6,26,28,29]. Supporting our prediction, several families in the order Bacteroidales—an
140 order of anaerobic bacteria comprising many species which are involved in breaking down
141 complex sugars in the colon[31]—were enriched in the large intestine (**Figure 2B, Supplementary**
142 **Figure 1**), including the families Bacteroidaceae (median log2 fold change per mouse = 0.338;
143 Wilcoxon signed-rank test, p-value = 0.0143), Porphyromonadaceae (median log2 fold change
144 per mouse = 8.58×10^{-2}; Wilcoxon signed-rank test, p-value = 0.0143), and Rikenellaceae
145 (median log2 fold change per mouse = 6.23×10^{-3}; Wilcoxon signed-rank test, p-value =
146 0.0209). Meanwhile, the family Lactobacillaceae, which is known to colonize the human small
147 intestine with the aid of unique mucus-binding proteins (MUBs) which adhere to the small
148 intestinal mucosal layer[29], was enriched in the small intestine relative to the large intestine
149 (median log2 fold change per mouse = -0.121; Wilcoxon signed-rank test, p-value = 0.0225)
150 (**Supplementary Figure 1**). These enrichments in the small intestine confirm that despite human
151 stool more closely resembling the community composition of the human large intestine[30],
152 inoculation with a human fecal sample is sufficient to support colonization by small intestine-
153 associated taxa. Together, these results are in line with previous work showing that the
154 composition of microbial taxa differs along the gut[3,28].

155 Lastly, to investigate community-wide functional diversity along the gut, using HUMAnN3[32] we
156 mapped metagenomic reads to gene families and MetaCyc pathways. Across all samples,
157 HUMAnN3 identified 808,204 gene families corresponding to 470 MetaCyc pathways, which
158 were subsequently aggregated at the community level by summing read counts across taxa.
159 These pathway profiles were then analyzed using MaAsLin3[33] to test for differences between the
160 small intestine and large intestine. The model incorporated gut location, cage, and read count as
161 fixed effects and used subject-level stratification for paired comparisons—revealing 80 pathways
162 significantly associated with gut location ($q < 0.05$). Among these, 33 pathways showed highly
163 significant associations ($q < 0.001$), and their coefficients are shown in **Supplementary Figure 2**

164 **(Supplementary Table 2)**. The same families we found to be enriched in the large intestine were
165 also the source of many of the pathways most significantly associated with the large intestine,
166 namely Bacteroidaceae (25 pathways) and Rikenellaceae (14 pathways). Meanwhile, a family
167 enriched in the small intestine, Lactobacillaceae, was the source of 7 pathways significantly
168 associated with the small intestine.


A
Shannon diversity
4.5
4.0
3.5
3.0
2.5
2.0
1.5
**
**
**
*
**
*
Duodenum
Jejunum
Ileum
Cecum
Colon
Inoculum
B
Relative Abundance (%)
100
75
50
25
0
In
Duodenum
Jejunum
Ileum
Cecum
Colon
1 2 3 4 5 6 7 8
Mouse
Bacteroidales
Bacteroidaceae
Porphyromonadaceae
Prevotellaceae
Rikenellaceae
Bifidobacteriales
Bifidobacteriaceae
Burkholderiales
Sutterellaceae
unidentified
Clostridiales
Clostridiaceae
Eubacteriaceae
Lachnospiraceae
Oscillospiraceae
Ruminococcaceae
unidentified
Coriobacterineae
Coriobacteriaceae
Desulfovibrionales
Desulfovibrionaceae
Enterobacteriales
Enterobacteriaceae
Erysipelotrichales
Erysipelotrichaceae
Lactobacillales
Enterococcaceae
Lactobacillaceae
Leuconostocaceae
Streptococcaceae
Verrucomicrobiales
Verrucomicrobiaceae
Unidentified
unidentified

169

170 **Figure 2. Increase in taxonomic diversity and change in community membership along the**
171 **length of the gut. (A)** Alpha diversity (Shannon Index) estimates for bacteria in different regions
172 of the gut. Two-sided Wilcoxon rank sum tests were performed between all possible pairs of
173 small intestinal versus large intestinal samples (* indicates $p \leq 0.05$; ** indicates $p \leq 0.01$). **(B)**
174 Relative abundance of bacterial families in the five gut regions and inoculum. Bacterial families
175 are grouped by order in the legend.

176 **Bacterial genetic diversity along the gut**

177 Having confirmed that taxonomic diversity varies spatially along the gut, we next investigated
178 whether the same was true of within-species genetic diversity. First, we quantified the
179 distribution of nucleotide diversity (π) for the 30 most abundant species detected in the inoculum
180 and mice (**Methods**) (**Supplementary Figure 3**, **Supplementary Table 3**). Across these
181 species, nucleotide diversity measured in the inoculum was highly variable, ranging from
182 2.2×10^{-4} in *Odoribacter splanchnicus* to 1.2×10^{-2} in *Bacteroides vulgatus,* spanning two
183 orders of magnitude (**Figure 3A**). High values of π ($> 10^{-3}$) are generally inconsistent with the
184 colonization and subsequent diversification of a single strain during a host's lifetime, and instead
185 support the presence of two or more strains that diverged long before colonization[23]. Meanwhile,
186 low values of π are more consistent with colonization by a single strain[23]. Thus, the large range
187 of π values observed across species indicates that in some instances, there are multiple co-
188 colonizing strains of the same species, and in other instances there is only one strain.

189 Average nucleotide diversity values in the mice were positively correlated with nucleotide
190 diversity in the inoculum (Spearman $\rho = 0.845$, p-value = 7.55×10^{-7}) (**Figure 3A,**
191 **Supplementary Figure 4**). However, despite this overall similarity, a range of nucleotide
192 diversity values was observed within individual mice, with some mice having levels comparable
193 to those seen in the inoculum and others having substantially lower values, suggesting variation
194 in lineage structure across mice. Specifically, for the species with the highest values of
195 nucleotide diversity in the inoculum ($\pi > 1 \times 10^{-3}$), values across mice were often bimodally
196 distributed. For example, π values for *Bacteroides vulgatus* were either all clustered around 10^{-2}
197 in some mice or around 10^{-3} and lower for other mice (**Figure 3A**). This variation suggests that
198 in some mice, multiple strains colonized, and in other mice, a single strain colonized even when

multiple were present in the inoculum. However, π never spanned an order of magnitude within a single mouse, potentially implying more similar strain colonization dynamics along the gut within mice than across mice.

To understand if nucleotide diversity differed between regions of the gut, we measured F_{ST}, or fixation index, between the small and large intestine for each species. F_{ST} is a distance metric where high values indicate large genetic differentiation between two samples[39,40]. We hypothesized that F_{ST} would be elevated between gut regions within hosts, either due to evolutionary changes accruing or variable strain frequencies. Contrary to this expectation, we observed that F_{ST} between the small and large intestine of the same host was on average low for the 30 species examined (0.119), especially relative to the average F_{ST} between mice at the same gut region (0.225) (**Figure 3B**). This difference was more pronounced in species with high nucleotide diversity in the inoculum ($\pi > 10^{-3}$) (Wilcoxon rank sum test, p-value = 2.3×10^{-7}) compared to species with low diversity ($\pi \leq 10^{-3}$) (Wilcoxon rank sum test, p-value=0.043).

Species with high nucleotide diversity in the inoculum may have larger differences in F_{ST} within versus between mice than species with low nucleotide diversity in the inoculum because the former exhibits more variable strain colonization dynamics. Specifically, when multiple strains are present in the inoculum, there is a potential for different combinations of strains to colonize different mice, leading to higher F_{ST} and differences in π between mice. Within mice, F_{ST} tends to be low and π tends to be uniform for these species, suggesting that similar combinations of strains are found all along the gut tract of the same host. Meanwhile, when there is only one strain present in the inoculum, the same strain should colonize all mice and gut regions. This is consistent with the observed low values of F_{ST} and more uniform values of π both within and between mice for these latter species. In the next section, we further investigate whether similar combinations of strains are found along the gut tract when nucleotide diversity is high.

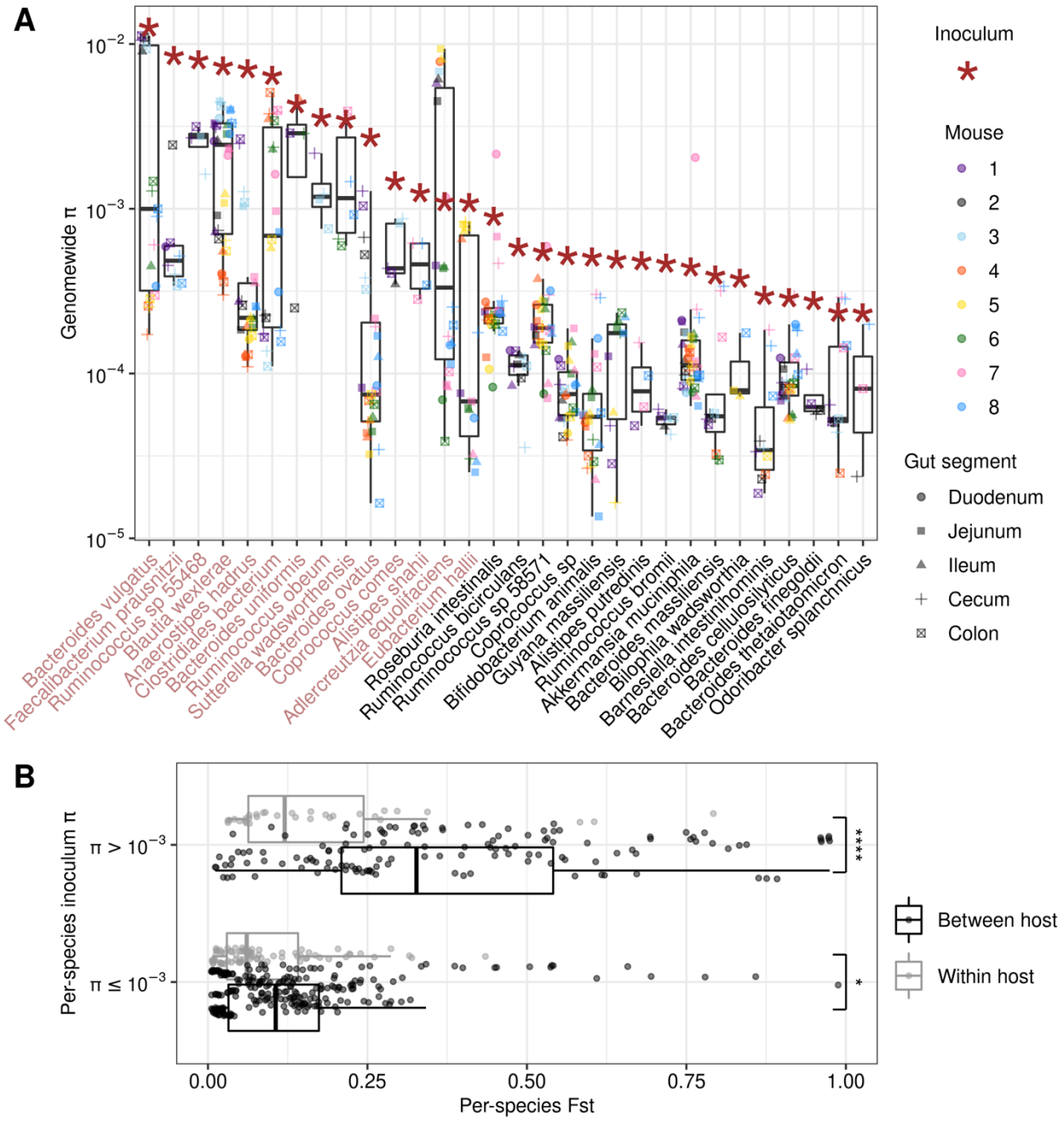


Figure 3: Genetic diversity measured in 30 gut commensal species. (A) Nucleotide diversity (π) estimates for the 30 most abundant and prevalent species (**Methods**). Each dot represents nucleotide diversity (π) in a specific gut region of an individual mouse. Dot color corresponds to the mouse, while dot shape indicates the gut region. Red asterisks denote nucleotide diversity measured in the inoculum for that species. Species names in light red correspond to those species with $\pi > 1 \times 10^{-3}$ in the inoculum. **(B)** Fixation index (F_{ST}) for between all pairs of mouse

230 samples for the 30 species analyzed in panel **(A)**. Each data point represents a single species for a
231 given pair of samples, with points colored according to whether the sample comparison is within
232 host (grey) or between hosts (black). Data points are further separated based on nucleotide
233 diversity of the species in the inoculum. Two separate Wilcoxon rank sum tests were performed
234 comparing the distributions of F_{ST} within versus between hosts for species with $\pi > 10^{-3}$ and
235 species with $\pi \leq 10^{-3}$, respectively (* indicates $p \leq 0.05$; **** indicates $p \leq 0.0001$).

236 **Strain frequencies are uniform along the gut within hosts**

237 We next asked whether the identities and frequencies of strains colonizing different parts of the
238 gut of a given mouse are the same or different when multiple strains are present in the inoculum.
239 Variation in occupancy or frequency along the gut would suggest that physical niche partitioning
240 is important in maintaining multiple conspecific strains in the same host that might otherwise
241 compete in the same physical niche, or that distinct environmental pressures along the gut select
242 for different strains.

243 To understand whether frequencies of co-colonizing strains of the same species vary along the
244 gut, we inferred strain frequencies for the 14 species with levels of nucleotide diversity $> 10^{-3}$
245 in the inoculum, as these samples had sufficiently high enough diversity to potentially harbor at
246 least two strains (**Figure 3**). To infer strain frequencies, we used an algorithm previously applied
247 to metagenomic time series data[24,25] (**Methods**). For the species *Alistipes shahii*, *Bacteroides*
248 *ovatus*, *Coprococcus comes*, and *Faecalibacterium prausnitzii*, only one strain was inferred to be
249 present within the mouse samples. For the rest of the species, we inferred that two strains were
250 present (**Figure 3**; **Supplementary Figure 5**). We next asked whether the frequencies of strains
251 for species with two strains were uniform or variable along the gut.

252 Application of the strain phasing algorithm revealed that frequencies of co-colonizing strains
253 were relatively uniform across gut regions within hosts for most bacterial species and mice,
254 contrary to the expectation that their frequencies may differ due to niche partitioning or
255 environmental selection. For example, for the species *Blautia wexlerae*, the two strains were able
256 to coexist at roughly 40:60% frequency along all five regions of the gut in mouse 8, and for the
257 species *Adlercreutzia equilofaciens*, the two strains were able to coexist at roughly 20:80%

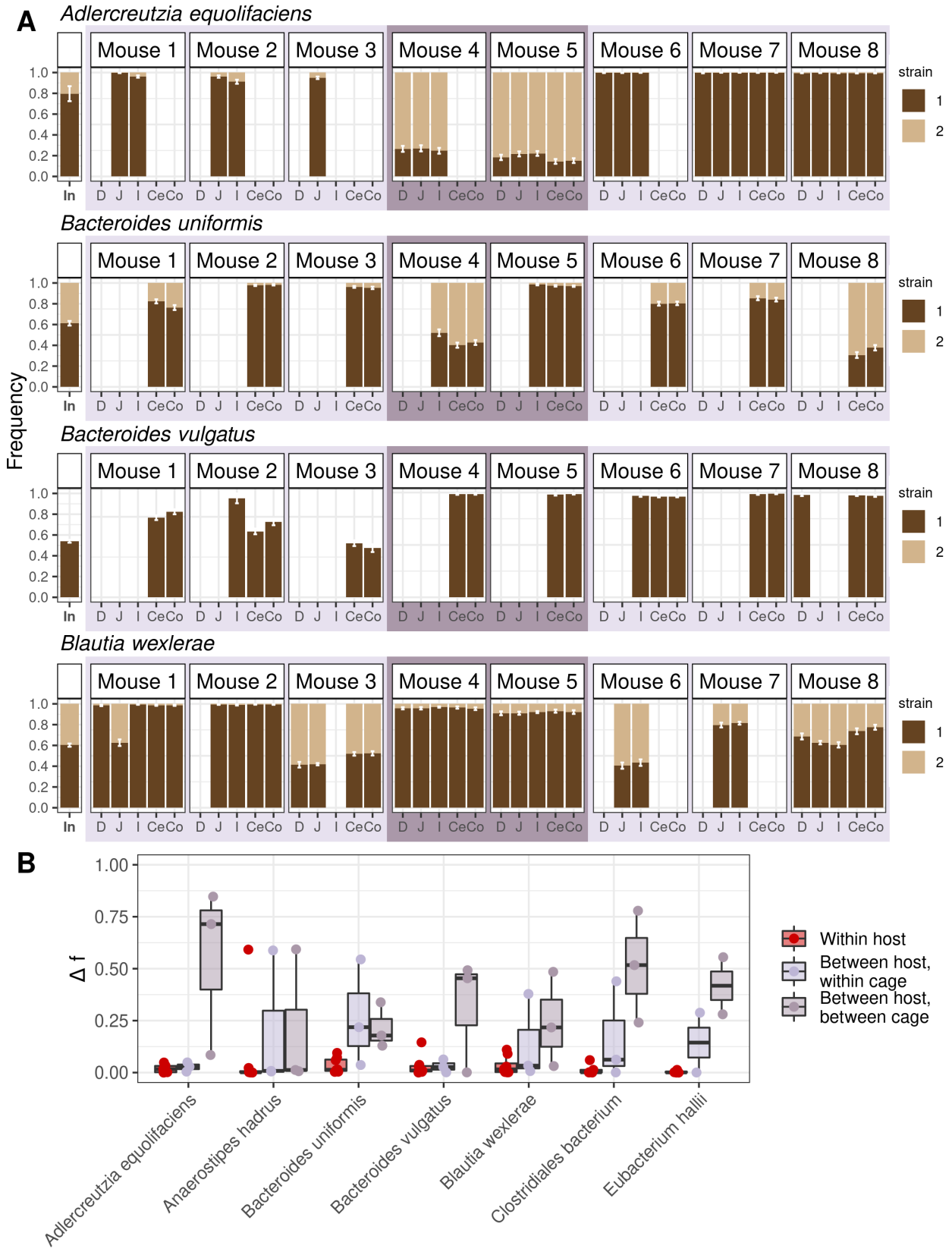
258 frequency in all parts of the gut in mice 4 and 5 (**Figure 4A**). To understand whether co-
259 colonizing strains generally displayed uniform frequencies along the gut, we calculated change
260 in major strain frequency (Δf) between mouse samples for 7 species that were present in
261 multiple mice, cages, and gut regions (**methods**). Across the seven species, median change in
262 frequency between gut regions within a host was only 9.36×10^{-3}, much smaller than the
263 median change in frequency we observed across mice within the same cage (3.49×10^{-2}) or
264 between mice in different cage (3.09×10^{-1}) (**Figure 4B**).

265 To confirm that strain frequencies were generally more homogeneous within versus between
266 hosts for all species in our dataset, we performed a series of ANOVA tests to estimate the
267 proportion of variance in major strain frequencies attributable to gut region ("gut region"),
268 mouse identity ("mouse"), and cage identity ("cage") (**Supplementary Figure 6**). Consistent
269 with the observations of uniformity in genetic variation along the gut, gut region explained the
270 lowest variance for all seven bacterial species (average variance explained = 2.70%), suggesting
271 that gut region is less important in explaining strain frequencies than mouse identity or cage
272 identity. By contrast, mouse and cage explained substantially greater variation (average
273 combined variance explained = 90.0%). These observations support the idea that strain
274 frequencies along the gut were generally uniform across species and reveal that multiple strains
275 were able to coexist in the same region of the gut. Furthermore, they imply some ecological
276 mechanism other than spatial segregation must be responsible for permitting strain coexistence
277 within the same host.

278 Social transmission of bacteria, through behaviors like coprophagy[14,41], may also facilitate
279 between-host spread of strains, resulting in similar strain frequencies between mice in the same
280 cage. Supporting this notion, strain frequencies were more similar among cohoused mice
281 compared to mice housed in different cages, indicating that migration of strains between mice
282 plays a role in homogenizing frequencies of strains. For example, for the species *A.*
283 *equilofaciens,* which was present in multiple mice and cages, the within-cage variance of strain
284 frequencies measured in the jejunum ranged from 1.65×10^{-6} in cage 3 to 1.29×10^{-3} in
285 cage 2, whereas the between-cage variance was at least two orders of magnitude greater,
286 1.18×10^{-1} (**Supplementary Figure 7, Supplementary Table 4**). These results align with
287 previous studies that have found coprophagic behavior in mice can facilitate the spread of strains

288 and adaptive variants in *Bacteroides thetaiotaomicron*[42] and *Escherichia coli*[14,41] between co-
289 housed mice, enforcing similar strain frequencies between mice[14]. However, not all species have
290 similar within-cage strain frequencies (e.g., *B. uniformis*, **Figure 4**), suggesting barriers to
291 between-host migration in these cases. This species-specific migration effect is reflected in the
292 fact that "cage" explains most of the variance in strain frequency for four species, but very little
293 variance in the remaining three species (**Supplementary Figure 6**).

294 Together, these results indicate that co-colonizing strains do not need to occupy physically
295 distinct niches to colonize the same host. To the contrary, strain frequencies tended to be
296 relatively uniform along the gut. This uniformity may be driven by rapid migration rates between
297 gut regions, which also likely results in greater similarity in strain frequencies among mice of the
298 same cage versus different cages due to social transmission.


A
Adlercreutzia equolifaciens
Bacteroides uniformis
Bacteroides vulgatus
Blautia wexlerae
Mouse 1
Mouse 2
Mouse 3
Mouse 4
Mouse 5
Mouse 6
Mouse 7
Mouse 8
Frequency
strain
1
2
In
D J I CeCo
B
Δ f
Within host
Between host, within cage
Between host, between cage
Adlercreutzia equolifaciens
Anaerostipes hadrus
Bacteroides uniformis
Bacteroides vulgatus
Blautia wexlerae
Clostridiales bacterium
Eubacterium hallii

299

300 **Figure 4. Strain frequencies are similar along the gut but may differ between mice housed**
301 **in different cages. (A)** Strain frequencies for four bacterial species detected in the guts of eight
302 mice inoculated with the same human stool sample. A total of 14 bacterial species had high
303 levels of inoculum nucleotide diversity ($\pi \geq 1 \times 10^{-3}$). See **Supplementary Figure 5** for the
304 strain frequencies of the ten species not shown here). "Strain 1" and "strain 2" represent the same
305 strain across all samples for a given species. Error bars represent bootstrapped 95% confidence
306 intervals. Mice 1-3 were co-housed in cage 1, mice 4 and 5 in cage 2, and mice 6-8 in cage 3,
307 with the housing scheme indicated by the alternating light and dark purple boxes. Absence of a
308 strain frequency bar in a sample indicates insufficient coverage or species absence (**Methods**).
309 **(B)** Change in strain frequency Δf was calculated for seven species present in multiple mice,
310 cages, and gut regions. (**Methods**). Each dot represents the change in major strain frequency
311 between a pair of samples from within the same host (red), between hosts in the same cage (light
312 purple), and between hosts in the different cages (dark purple).

313 **Evolutionary changes are not localized and spread throughout the gut**

314 Even if bacterial strain frequencies are relatively homogenous along the gut, it is possible that
315 individual genetic variants within species are locally restricted along the gut, potentially due to
316 local adaptation. To assess the possibility of local adaptation, we asked if any SNVs have
317 extreme frequency differences between gut regions, or if they too are homogeneously distributed.

318 SNV frequency changes between gut regions could arise within a host as a consequence of (i)
319 evolutionary changes (e.g., adaptation), (ii) fluctuations in frequencies of genetically distinct
320 strains, or (iii) sampling error. Here, we were interested in distinguishing SNV frequency
321 changes arising as a consequence of evolution from those arising from strain fluctuations or
322 sampling error. To distinguish SNV frequency changes arising due to evolution as opposed to
323 strain fluctuations, we inferred the genotype of one of the strains present in the sample, or 'quasi-
324 phased' the genotype[23] and then tracked SNV frequency changes on the background of that
325 genotype between pairs of samples. Doing so allowed us to detect new mutations modifying a
326 resident lineage as opposed to alleles rising to high frequency due to strain fluctuations. To
327 distinguish SNV frequency changes due to evolution (e.g., adaptation) versus sampling error, we
328 considered only large allele frequency changes ($f \leq 0.2$ in one sample to $f \geq 0.8$ in another

329 sample) between samples with a median read coverage of 20 or more and at sites with a
330 minimum read coverage of 20 or more in both samples, which are unlikely to arise due to
331 sampling error or drift[23,43]. Previously, we showed that the genotype of the dominant lineage in
332 samples with a single, high-frequency strain can be confidently quasi-phased[23], and that
333 evolutionary changes with large allele frequency changes can subsequently be detected
334 (**Methods**).

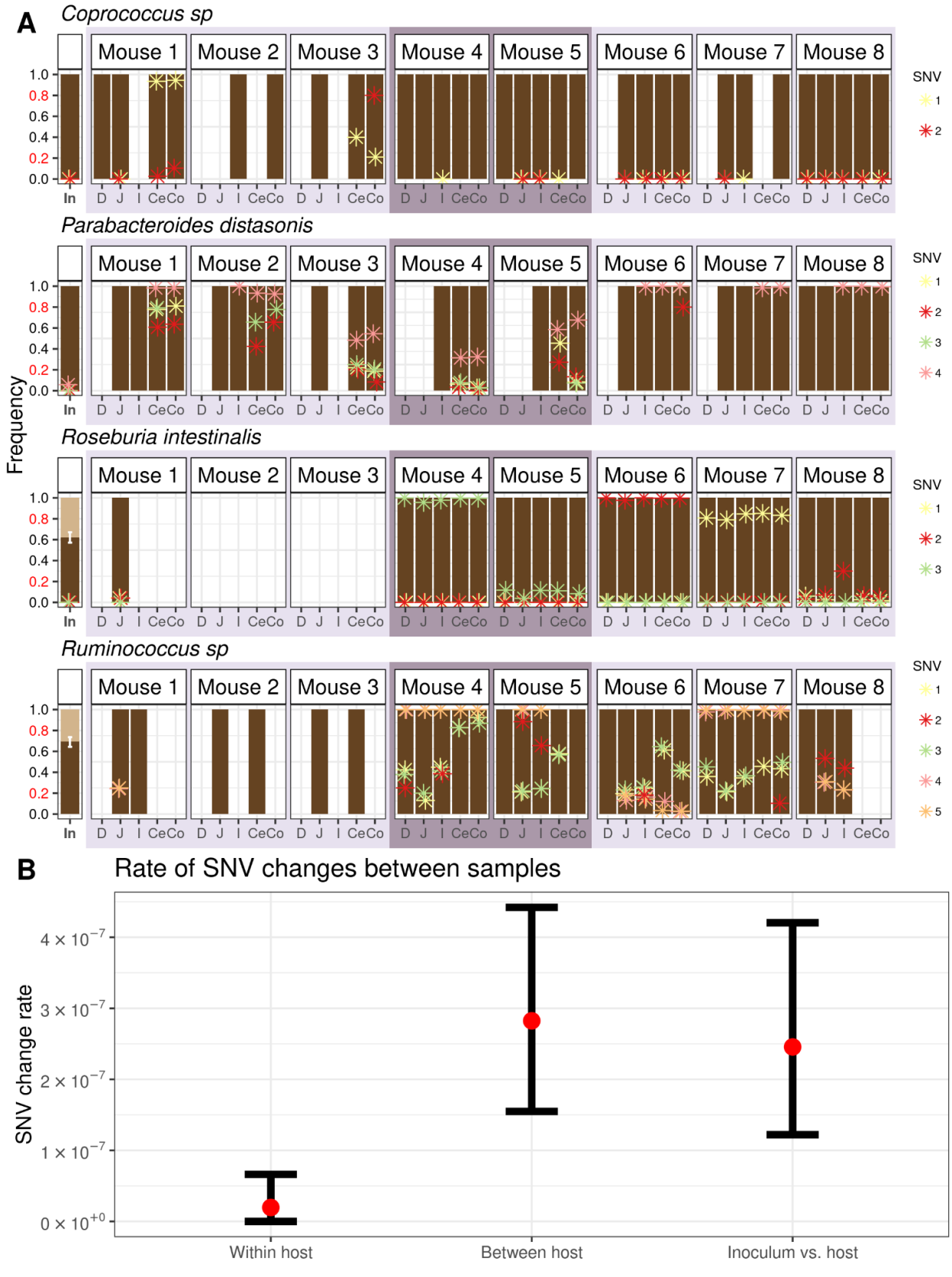
335 Using this method, we identified 366 quasi-phaseable genomes from 62 bacterial species present
336 in the intestinal contents of the eight mice and the inoculum (**Supplementary Figure 3**). Next,
337 we detected SNV frequency changes between pairs of quasi-phaseable genomes. These pairs
338 were between the inoculum and any mouse (10 bacterial species), different gut regions *within*
339 mice (spanning 33 bacterial species), and gut regions *between* mice (28 bacterial species). To
340 avoid comparing distantly diverged lineages that were not descended from the same inoculum
341 strain, we only considered pairs with ≤20 SNV frequency changes. We observed a total of 65
342 unique SNV frequency changes across all possible pairs of quasi-phaseable genomes in our
343 dataset (**Figure 5A**, **Supplementary Figure 3, Supplementary Figure 8**, **Supplementary**
344 **Table 5**).

345 By having an inoculum to compare to, we can ascertain if a lack of SNV differences along the
346 gut is due to the no evolutionary changes accruing, or the global spread of an evolutionary
347 change throughout the gut. We observed 11 SNV frequency changes rising from low frequency
348 in the inoculum to high frequency in mice in the 10 species with at least one inoculum-mouse
349 sample quasi-phaseable genome pair (**Figure 5A**, **Supplementary Figure 8**, **Supplementary**
350 **Table 5**). The rapid timescales over which these SNVs rose to high frequency are consistent with
351 these SNV frequency changes potentially representing adaptations occurring during the host
352 colonization process. Only one of these SNVs, detected in *Coprococcus* sp., showed spatial
353 differentiation in a single host (SNV frequency of 0 in jejunum versus 0.935 in cecum of mouse
354 1) (**Figure 5A**, **Supplementary Table 5**), indicating that advantageous SNVs tend to spread
355 along the gut. Since not all SNVs observed within mice had sufficient coverage to be analyzed in
356 the inoculum, we next examined SNV frequency changes arising between gut regions.
357 Supporting the observation that SNVs spread along the gut, we observed only 4 unique *within-*
358 *host* SNV frequency changes in 3 out of the 33 species, all between the small intestine and large

359 intestine, including the SNV in *Coprococcus* sp. (**Supplementary Table 5**). By contrast, we
360 observed 57 unique *between-host* SNV frequency changes in 15 of the 28 species (**Figure 5A**,
361 **Supplementary Figure 8**, **Supplementary Table 5**). The rate of SNV frequency changes
362 observed between hosts was significantly higher than that within hosts (bootstrapped 95%
363 confidence interval for between-host SNV changes is 1.54×10^{-7} differences/bp -
364 4.42×10^{-7} differences/bp; bootstrapped 95% confidence interval for within-host SNV changes
365 is 0 differences/bp - 6.62×10^{-8} differences/bp) (**Figure 5B**). The low rate of SNV frequency
366 changes within mice relative to between-host and inoculum-mouse SNV frequency changes rates
367 suggests that potentially advantageous SNVs overwhelmingly transmit along the gut within
368 hosts.

369 To understand whether SNVs spreading along the gut showed additional signals of being
370 adaptive beyond extreme allele frequency changes on short time scales, we asked if the same
371 SNV changes arise in multiple mice, as parallelism is a hallmark of adaptation[44–48]. Adaptive
372 variation can arise in multiple mice via independent recurrent *de novo* mutation or the parallel
373 rise in frequency of standing genetic variation (SGV) already present in the inoculum[48]. Of the
374 21 unique SNVs that were observed to undergo an extreme allele frequency change between any
375 pair of samples and had sufficient coverage in the inoculum (minimum of 20 reads), 15 had non-
376 zero frequency in the inoculum, indicating that SGV is an important source of genetic variation
377 in these mice (**Figure 5A**, **Supplementary Figure 8**, **Supplementary Table 5**). Of these, five
378 nonsynonymous and one synonymous SNVs were observed to undergo parallel sweeps from low
379 frequency in the inoculum to high frequency in multiple independent cages. One of these parallel
380 SNVs was in *P. distasonis* in a gene encoding a TonB-dependent transporter, which rose from
381 low frequency in the inoculum ($f = 0.057$) to high frequency in multiple mice in all three
382 cages (ranging from an average $f = 0.319$ in mouse 4 to an average $f = 1.00$ in mouse 6),
383 consistent with parallelism (**Figure 5A**, **Supplementary Table 5**). TonB-dependent transporters
384 are ubiquitous outer membrane-associated proteins involved in nutrient transport[49] with
385 structural variation having been found to be associated with niche partitioning by different
386 *Bacteroides* taxa[50]. Moreover, TonB-dependent transporters have been found to be a common
387 target of adaptation in the human gut microbiome[47,51,52]. Notably, all the SNVs observed to
388 undergo parallelism were not found to be differentiated along the gut within hosts. This fact

389 would seem to suggest that even when SNVs are adaptive, they are either broadly adaptive to the
390 entire gut, or alternatively have high migration rates that outweighs local selection pressures,
391 perhaps due to mixing and peristalsis along the gut[53].


A
Coprococcus sp
Parabacteroides distasonis
Roseburia intestinalis
Ruminococcus sp
Mouse 1
Mouse 2
Mouse 3
Mouse 4
Mouse 5
Mouse 6
Mouse 7
Mouse 8
Frequency
SNV
In
D J I CeCo
B
Rate of SNV changes between samples
SNV change rate
Within host
Between host
Inoculum vs. host

392

393 **Figure 5: Evolutionary changes are uniformly found along the gut. (A)** SNVs exhibiting
394 extreme allele frequency changes (allele frequency $f \leq 0.2$ in one sample to $f \geq 0.8$ in another
395 sample) between any pair of samples sharing the same quasi-phaseable strain genome (< 20 SNV
396 frequency changes) are shown for *Coprococcus* sp., *P. distasonis*, *R. intestinalis*, and
397 *Ruminococcus* sp. Asterisks represent the frequency of a given SNV, with each SNV within a
398 species having its own unique color. Samples lack asterisks for a particular SNV when those loci
399 do not meet the minimum coverage requirement of 20 reads to infer frequency. When two strains
400 of the same species are present, strain abundances of co-colonizing strains are represented by
401 two different shades of brown (see **Figure 4** for legend). The remaining 11 species in which
402 SNV frequency differences were detected are shown in **Supplementary Figure 8**. **(B)**
403 Comparison of the rate of SNV frequency changes aggregated across species within mice,
404 between mice, and between inoculum and mice. Colored in red are the observed SNV frequency
405 change rates aggregated across species. Bars represent bootstrapped 95% Cis (**Methods**).

406 **Genetic homogeneity is generalizable to the gut of conventional mice and healthy humans**

407 Having established that uniform strain frequencies and spread of evolutionary adaptations along
408 the gut is a common feature in humanized mice, we next investigated whether these same
409 patterns also arise in hosts with established native microbiota, specifically conventional mice and
410 healthy humans.

411 We first analyzed a cohort of nine conventional mice that were cohoused until weaning at three
412 weeks of age before being housed into three cages of three mice each. At six weeks of age, we
413 sampled and sequenced luminal contents from the same five gut regions that were previously
414 examined in humanized mouse guts (**Methods**). Large intestinal samples harbored higher alpha
415 diversity and a different family-level composition than small intestinal samples (**Methods**)
416 (**Supplementary Figure 9**), confirming that the expected spatial organization does exist at
417 higher taxonomic levels along the guts of the conventional mice.

418 Further replicating our findings from humanized mice, we found that in conventional mice, co-
419 colonizing strains of the same species exhibited largely uniform frequencies across gut regions.
420 This pattern contrasts with the substantial variation in strain frequencies between mice, even if
421 co-housed (**Figure 6A**, **Supplementary Figure 10**). Among the nine bacterial species with co-

422 colonizing strains that were present in multiple mice, cages, and gut regions, median change in
423 frequency Δf between gut segments within hosts was only 2.21×10^{-2}, an order of magnitude
424 smaller than the median frequency changes observed between mice in the same cage
425 (2.16×10^{-1}) and in different cages (1.42×10^{-1}) (**Figure 6B**). Corroborating these results, gut
426 region explained the least variation (average variance explained = 0.75%), whereas mouse
427 identity explained the most variation (average variance explained = 78.39%) (**Supplementary**
428 **Figure 11**). Cage explained a moderate amount of variation (average variance explained =
429 19.97%). These results indicate that uniform frequencies of strains along the gut is not unique to
430 humanized mice and may be a common feature of mouse microbiota.

431 Additionally, we did not observe evolutionary changes accruing between gut regions within
432 hosts even while changes arose between hosts at a rate of 1.03×10^{-6} changes/bp (**Figure 6C**).
433 These results indicate that any evolutionary changes arising within hosts sweep globally
434 throughout the gut.

435 Next, to determine whether spatially uniform patterns of genetic diversity are a feature of human
436 guts as well, we analyzed previously published metagenomic data[54] collected using ingestible
437 sampling capsules designed to sample luminal contents from four successive regions of the
438 gastrointestinal tract: pyloric sphincter to the upper small intestine, upper to mid-small intestine,
439 mid- to lower small intestine and the lower small intestine into the ascending colon. In this
440 dataset, 15 healthy adult subjects ingested four capsule types twice daily for as long as a week.

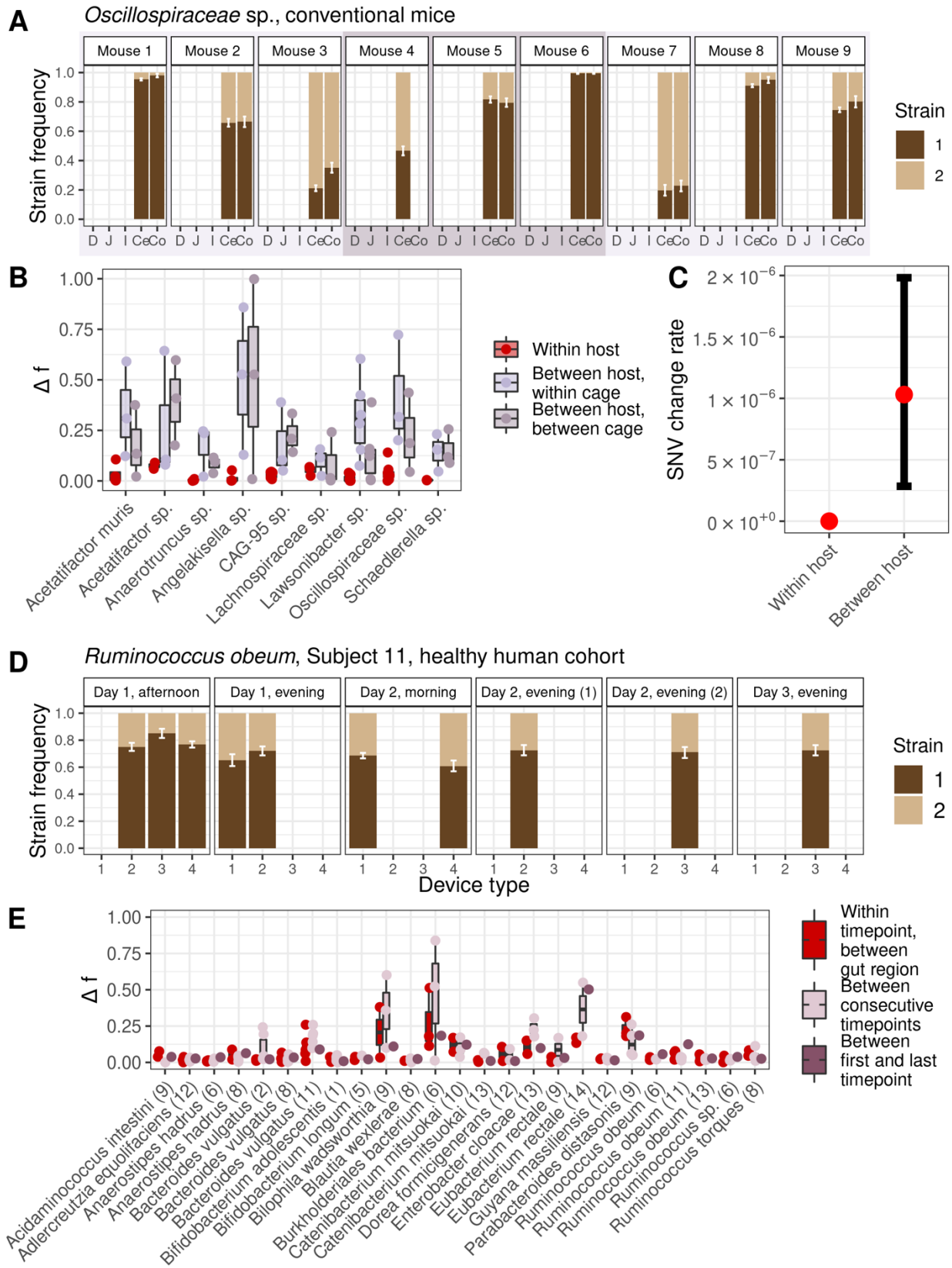
441 We hypothesized that strain frequencies in humans might vary more along the gut than in mice,
442 given the greater length and compartmentalization of the human intestine, as well as the potential
443 for host-specific niche adaptation over long residence times[25,27]. Contrary to this expectation,
444 strain frequencies were generally homogeneous not only along the gut but also across timepoints.
445 For example, the frequency of the major strain of *Ruminococcus obeum* in Subject 11 fluctuated
446 around a mean of 72% with a variance of 0.4% across gut locations and timepoints (**Figure 6D**).
447 More generally, across the 25 species x host pairs with co-colonizing strains in multiple
448 timepoints and device types, the median change in major strain frequency Δf between gut
449 regions within the same timepoint was only 3.79×10^{-2}. This was the same order of magnitude
450 as the median change observed between consecutive timepoints (2.92×10^{-2}), or even between

451 the first and last observed timepoint for a species x host pair (3.55×10^{-2}) (**Figure 6E**). These
452 results indicate that strain frequency tends to be uniform along the human gut as well as across
453 hourly and daily timescales.

454 While the majority of species x host pairs never showed a temporal or spatial strain fluctuation in
455 strain frequency exceeding 25%, there were several outlier species x host that harbored strains
456 undergoing large frequency temporal shifts, and to a lesser degree, spatial shifts. However, these
457 fluctuations tended to be transient. For example, the frequency of the major strain of
458 *Burkholderiales bacterium* in Subject 6 fluctuated between a minimum of ~10% and maximum
459 of ~95% within and across timepoints, but displayed no consistency in the direction of change
460 across either space or time (**Supplementary Figure 12**). We observed this in other species x
461 host pairs as well, illustrating that strain fluctuations that do occur along the gut and over time
462 are short-lived, consistent with previous work showing that stably co-colonizing strains can
463 fluctuate in frequency around a carrying capacity over longer timescales[25].

464 We next asked whether extreme SNV frequency changes arising as a consequence of within-host
465 evolution were spatially segregated. To do so, we identified SNV frequency changes between
466 distinct regions of the gut for each of the 56 host x species pairs (representing 35 species
467 detected in 11 subjects) with enough data in at least two capsules for a given time point. Out of
468 the 56 pairs, only 7 of these (12.5% of species x host pairs) had detectable SNV frequency
469 changes between gut regions within a single timepoint (**Supplementary Table 6**), indicating that
470 spatial segregation of SNVs is a feature in only a minority of samples.

471 Finally, we investigated whether the SNV changes arising along the gut are observable over
472 multiple time points, or whether they are transient. Persistence would suggest that environmental
473 gradients maintain spatial segregation of genetic variation along the gut over time. Of the seven
474 host x species pairs with spatial differences, four had sufficient temporal and spatial data (at least
475 two timepoints with at least two device types per timepoint). In three of four host x species pairs,
476 spatial differences did not persist over multiple time points, with allele frequencies often
477 reverting to their original state (**Supplementary Figure 13**, **Supplementary Table 6**). These
478 results imply that sustained extreme allele frequency changes along the gut are rare.


A
Oscillospiraceae sp., conventional mice
Mouse 1
Mouse 2
Mouse 3
Mouse 4
Mouse 5
Mouse 6
Mouse 7
Mouse 8
Mouse 9
Strain frequency
D J I CeCo
Strain
1
2
B
Δ f
Within host
Between host, within cage
Between host, between cage
Acetatifactor muris
Acetatifactor sp.
Anaerotruncus sp.
Angelakisella sp.
CAG-95 sp.
Lachnospiraceae sp.
Lawsonibacter sp.
Oscillospiraceae sp.
Schaedlerella sp.
C
SNV change rate
Within host
Between host
D
Ruminococcus obeum, Subject 11, healthy human cohort
Day 1, afternoon
Day 1, evening
Day 2, morning
Day 2, evening (1)
Day 2, evening (2)
Day 3, evening
Strain frequency
Device type
Strain
1
2
E
Δ f
Within timepoint, between gut region
Between consecutive timepoints
Between first and last timepoint
Acidaminococcus intestini (9)
Adlercreutzia equolifaciens (12)
Anaerostipes hadrus (6)
Anaerostipes hadrus (8)
Bacteroides vulgatus (2)
Bacteroides vulgatus (8)
Bacteroides vulgatus (11)
Bifidobacterium adolescentis (1)
Bifidobacterium longum (5)
Bilophila wadsworthia (9)
Blautia wexlerae (8)
Burkholderiales bacterium (6)
Catenibacterium mitsuokai (10)
Catenibacterium mitsuokai (13)
Dorea formicigenerans (12)
Enterobacter cloacae (13)
Eubacterium rectale (9)
Eubacterium rectale (14)
Guyana massiliensis (12)
Parabacteroides distasonis (9)
Ruminococcus obeum (6)
Ruminococcus obeum (11)
Ruminococcus obeum (13)
Ruminococcus sp. (6)
Ruminococcus torques (8)

479

Figure 6. Uniform genetic variation along the gut in conventional mice and healthy humans. (A) Strain frequencies for two co-colonizing strains of an *Oscillospiriceae* species long the guts of nine conventional mice housed across three cages. Error bars represent bootstrapped 95% confidence intervals. Mice 1-3 were co-housed in cage 1, mice 4-6 in cage 2, and mice 7-9 in cage 3, as indicated by the light and dark purple boxes. Absence of a strain frequency bar indicates insufficient coverage or species absence in a given sample. For additional strain frequency plots corresponding to species analyzed in **(B)**, see **Supplementary Figure 10**. **(B)** Change in strain frequency Δf was calculated for seven species present in multiple mice, cages, and gut regions. (**Methods**). Each dot represents the change in major strain frequency between a pair of samples from within the same host (red), between hosts in the same cage (light purple), and between hosts in the different cages (dark purple). **(C)** Comparison of the rate of SNV frequency changes within and between mice. Colored in red are the observed SNV frequency change rates aggregated across species. Bars represent bootstrapped 95% CIs. **(D)** Strain frequencies for two co-colonizing strains of *Ruminococcus obeum* were inferred from capsule devices that sampled luminal contents along the gut of a healthy human subject (subject 11) across six timepoints. Device type 1 targeted the pyloric sphincter to the upper small intestine, device type 2 targeted the upper to mid-small intestine, device type 3 targeted the mid- to lower small intestine, and device type 4 targeted the lower small intestine into the ascending colon[54]. Error bars represent bootstrapped 95% confidence intervals. Timepoints represent the time at which capsules were swallowed, with days being relative to the first timepoint plotted and times of day being coarsened into "morning" (before 12 pm PST), "afternoon" (after 12pm and before 8 pm PST), and "evening" (after 8 pm PST) bins. For additional strain frequency plots corresponding to species analyzed in **(E)**, see **Supplementary Figure 12**. **(E)** Change in strain frequency Δf was calculated between samples belonging to 25 species x host pairs with co-colonizing strains present in multiple timepoints and device types (**Methods**). Subject is indicated by the parenthetical. Each dot represents the change in major strain frequency between samples from different device types swallowed during the same timepoint, (red), collected during consecutive timepoints (light purple), and collected during the first and last timepoint (dark purple). Strain frequency change over the course of the entire sampling period rarely exceeded changes occurring between consecutive timepoints, suggesting that hourly or daily shifts in strain frequency are transient.

511 ## Discussion

512 In this study, we analyzed the frequencies of bacterial strains and evolutionary changes along the
513 guts of eight mice inoculated with the same human-stool-derived inoculum. In sharp contrast to
514 the species-level heterogeneity reported here and in previous studies[3,6,28,29,55–63], we found
515 substantial *genetic* homogeneity along the gut (**Figure 4, Figure 5**, **Supplementary Figure 5**,
516 **Supplementary Figure 8)**. Spatial uniformity was a feature of not only the bacterial species
517 analyzed in humanized mice, but also those in conventional mice and healthy humans, indicating
518 that uniformity is a robust feature of mammalian gut ecosystems. Together, these results indicate
519 that spatial segregation of strains is not a requisite for their coexistence and that environmental
520 gradients do not necessarily generate large strain fluctuations or allele frequency differences
521 along the gut lumen.

522 Uniformity likely arises due to rapid migration rates of bacteria between gut regions and, in the
523 case of the mice studies, due to rapid migration between mice (**Figure 4-6**, **Supplementary**
524 **Figure 5-7**, **Supplementary Figure 11**). This latter conclusion is supported by the observation
525 that strain frequencies are more similar between co-housed mice than between mice housed in
526 different cages, consistent with recent work demonstrating extensive social transmission of
527 bacteria among co-housed mice[14,41]. Additionally, this conclusion is supported by the
528 observation that evolutionary changes at the SNV level are usually found at similar frequencies
529 along the gut, suggesting migration of variants arising from a single origin as opposed to the
530 same evolutionary change arising in different locations along the gut independently of each
531 other. The possibility of rapid migration facilitating uniformity is consistent with recent work
532 showing that there is significant mixing of luminal contents along the gut[53].

533 Uniform strain frequencies are not only found along the length of the gut but also over time,
534 reflecting stable coexistence of strains. Stable coexistence is a common feature of healthy human
535 gut microbiomes[25,27] and in experimental evolution studies[64,65]. Our own analysis of the healthy
536 human cohort sampled over multiple time points show that while strains tended to fluctuate in
537 frequency more across time than they did across space (i.e., between capsule) within any one
538 timepoint, these temporal strain frequency fluctuations were short-lived and tended to revert to
539 the overall trend at subsequent sampling points (**Figure 6E**, **Supplementary Figure 12**). Stable

540 coexistence of conspecific strains is thought to be facilitated at least in part by metabolic niche
541 partitioning[66]. Our findings support the role of metabolic niche partitioning in stably maintaining
542 strain diversity, as spatial segregation does not seem required to reduce competition among
543 strains.

544 Future work employing higher-resolution sampling strategy across both time and space could
545 reveal bacterial genetic heterogeneity that was not observed in this study. For example, sampling
546 the gut prior to eight weeks post-inoculation might have captured distinct spatial distribution of
547 alleles before they migrated to all gut regions. Additionally, in this study we focused on extreme
548 allele frequency changes, as these are unlikely to be confounded with sampling error in shotgun
549 data and instead are likely to have arisen only due to adaptation. However, more subtle allele
550 frequency differences may manifest along the gut, especially among neutral genetic variants,
551 consistent with recent human microbiome studies analyzing metagenome-assembled genomes
552 reporting locally restricted, potentially neutral evolutionary changes in a limited number of
553 species[16,17].

554 In addition, while we focused on luminal contents from five gut regions and observed uniform
555 genetic variation, other physiological structures found in the mammalian gut could exhibit
556 greater spatial heterogeneity. For example, different crypts in the intestinal epithelium could
557 potentially be occupied by distinct strains, similar to that of skin pores[67]. Crypts are known to
558 have communities distinct from the lumen[55,68], and may perform ecological filtering at the strain
559 and nucleotide-level as well. Similarly, the mucosa displays different biogeography from the
560 lumen[56], and mucosal surfaces are known to generate population structure in culture[69], promote
561 site-specific binding in certain bacterial species that have co-evolved with their hosts[70], and drive
562 adaptive differentiation relative to populations in the gut lumen[9]. However, the technical
563 challenges in obtaining sufficient bacterial biomass for metagenomics sequencing from crypts or
564 mucosa is significant[71]. The continued improvement in host-depletion and fine-scale sampling
565 strategies (e.g., Wu-Woods *et al.*, 2023)[17] will enable investigation of bacterial population
566 structure at unprecedented spatial resolution.

567 Our findings in healthy mammalian hosts raise the question of whether spatial genetic structure
568 may emerge under disease conditions or altered environments. Local inflammation or

569 environmental changes could generate distinct niches that foster local bacterial adaptation or
570 filter for specific, disease-adapted strains. In future work, it will be interesting to investigate if
571 spatial structure at the genetic level arises in diseased hosts. Intestinal disorders such as
572 inflammatory bowel disease (IBD) are known to generate local inflammation phenotypes and
573 disrupt gut microbiome composition[72,73] in specific gut regions[4,74,75]. For example, in a recent
574 mouse study, adaptive SNVs in a species of pathogenic bacteria were found to appear locally in
575 different colonic tissues, enabling translocation of a pathogenic bacteria from the colonic lumen
576 to the mucosa, and subsequently to the liver[9]. Thus, in pathogenic bacteria and disease scenarios,
577 we might observe specific strains or local adaptations associated with these acute conditions
578 along the gut, supporting our original hypothesis that environmental heterogeneity might induce
579 genetic diversity along the gut.

580 In conclusion, our findings demonstrate that bacterial strains and their evolutionary changes are
581 uniformly distributed along the gut lumen of across different mammalian hosts, in contrast to the
582 spatial heterogeneity observed at higher bacterial taxonomic levels. These findings indicate that
583 spatial segregation of strains is not a requisite for their co-existence, and that environmental
584 gradients do not result in local adaptation among alleles experiencing the largest allele frequency
585 changes. Future studies aimed at understanding the dynamics of migration facilitating this
586 uniformity and the metabolic strategies employed by strains to co-exist will be crucial for
587 elucidating the dynamics of strain-level diversity and its functional consequences within the gut
588 ecosystem.

589
Methods

590
Humanized mouse experimental design

591 For the humanized mouse study, we performed shotgun sequencing on the luminal contents of
592 mice we previously analyzed with 16S rRNA sequencing[3]. All animal experiments were
593 conducted in accordance with the ethical guidelines of the University of British Columbia's
594 (UBC) animal care procedures, following protocol numbers A19-0078 and A19-0122 approved
595 by the Animal Care Committee. Swiss Webster male and female mice at nine weeks of age were
596 used and were provided with an autoclaved standard diet (Purina LabDiet 5K67). Briefly, eight

597 germ-free Swiss Webster mice were orally gavaged with the same human stool sample (the
598 inoculum) from a healthy adult human previously labeled TL1[3]. This set up is akin to previous
599 studies conducting controlled mouse studies with human inoculums[18–21]. The eight mice were
600 housed in three different cages in the following configurations: cage 1 had mice 1-3 (sex male),
601 cage 2 had mice 4 and 5 (sex male), and cage 3 had mice 6-8 (sex female). Given that human
602 microbiota can rapidly colonize the mouse in a matter of days to weeks, we allowed for a total of
603 eight weeks to colonize and equilibrate in the mice[18–21], allowing for stabilization of community
604 structure and recovery following oral gavage. This period included six weeks on a standard
605 rodent diet (LabDiet 5k67), followed by two weeks in which mice in cages 2 and 3 were exposed
606 to a fiber-rich diet of 30% guar gum (TestDiet 5BSE) while those in cage 1 remained on the
607 same chow. With this experimental set up, we were able to ascertain replicability and
608 generalizability of the genetic homogeneity observed along the gut in three independent cages
609 with varying diets and sexes, as well as contrast within-host, across-host, and across-cage
610 dynamics (**Figure 1**).

611 Eight weeks after inoculation, mice were sacrificed using carbon dioxide with secondary cervical
612 dislocation, and luminal contents were collected from five intestinal regions: the duodenum,
613 jejunum, ileum, cecum, and colon. The intestines were carefully sectioned using a razor blade
614 into the corresponding regions and the contents were squeezed into 1.5 mL-microcentrifuge
615 tubes. DNA extraction was performed on luminal content samples and the original inoculum
616 (TL1) from 96-well plates using the DNeasy PowerSoil Pro Kit (Catalog number: 47016). For
617 more information on the approach used, see the methods section of Ng *et al.*, 2023[3].

618 **Conventional mice experimental design**

619 Conventional mice microbiomes were obtained from nine male Swiss Webster mice from the
620 same litter and with their own microbiomes at six weeks of age. All animal experiments were
621 conducted in accordance with the ethical guidelines of the University of British Columbia's
622 (UBC) animal care procedures, following protocol numbers A19-0078 and A19-0122 approved
623 by the Animal Care Committee. Mice were cohoused until three weeks of age, at which point
624 they were separated into three cages (three mice per cage). They were provided with an
625 autoclaved standard diet (Purina LabDiet 5K67) for the duration of the six weeks. At six weeks,

626 mice were sacrificed and luminal contents were collected from the same five intestinal regions as
627 with the humanized mice (duodenum, jejunum, ileum, cecum, and colon), and DNA extraction
628 was performed using an identical approach.

629 **Metagenomic Sequencing**

630 Metagenomic sequencing was performed on luminal content samples from humanized and
631 conventional mouse cohorts and the inoculum using a NovaSeq with 150 bp paired end reads.
632 Humanized mouse samples were sequenced to a depth ranging from 23M to 136M reads (median
633 = 33M reads), and the inoculum was sequenced to a depth of 119M reads. Samples from mice 7,
634 8 and the inoculum were sequenced in a separate batch than from samples from mice 1-6 at the
635 highest end of the range of sequencing depths. Conventional mouse samples were sequenced to a
636 depth ranging from 39M to 75M reads (median = 45M reads) (**Table S1**).

637 **Estimation of species abundances, SNVs, and CNVs**

638 For the humanized mice and data from Shalon *et al.*, 2023[54], we used the Metagenomic Intra-
639 species Diversity Analysis System 1 (MIDAS 1; version 1.2, downloaded November 24, 2021)[26]
640 to calculate relative species abundances, gene copy number variants (CNVs), and single
641 nucleotide variants (SNVs). MIDAS 1 calculated these values by mapping reads against a
642 database of genomes representing 5,952 bacterial species[26,76] found in the human gut
643 microbiome.

644 The conventional mice microbiomes harbored bacteria native to mice and thus was not well-
645 suited to the MIDAS 1 database used to process the humanized mouse and human microbiomes,
646 as the reference genomes in the database corresponded only to human-associated bacterial
647 species resulting in fewer reads being mapped. We therefore used MIDAS 3 (version 3.0.1;
648 downloaded August 5, 2024) to construct a mouse-specific database using isolate and
649 metagenome assembled genome (MAG) reference genomes from the Mouse Gastrointestinal
650 Bacteria Catalogue (MGBC)[77] using a standard approach[78]. MIDAS 3 was subsequently used to
651 map reads to this database and calculate relative species abundances, CNVs, and SNVs.

652

653 While the outputs from both pipelines were generated with permissive parameters, as described
654 below, we subsequently applied several stringent post-processing filters to rule out metagenomic
655 artifacts and to accurately estimate species, CNVs, and SNVs. The post processing pipeline we
656 use is also briefly described below and reflects the pipeline developed in Garud *et al.*, 2019[23].

657 **Estimating species relative abundance**

658 For humanized mouse and Shalon *et al.*, 2023 data, MIDAS 1 was used to estimate species
659 presence and abundance by mapping reads to 15 single-copy marker genes from 5,952
660 representative species[26,76]. Reads were mapped with default MIDAS 1 settings. Species with a
661 median single copy marker gene coverage ≥3 reads were marked as present within a sample.
662 These read coverages were then used by MIDAS 1 to compute species relative abundance. For
663 conventional mouse data, MIDAS 3 was used estimated species presence and abundance in a
664 similar manner, this time to 15 single-copy marker genes from 1,093 representative species[77] in
665 the custom-built mouse-specific database. Species were marked as present if they had a median
666 marker coverage ≥ 2 reads, per the recommendation of MIDAS 3 developers[78,79].

667 **Quantifying pathway relative abundance**

668 To quantify changes in pathway abundances along the gut, we used HUMAnN 3 (Beghini et al.,
669 2021) version 3.9. HUMAnN3 was downloaded on August 11, 2025 and run with default
670 settings using the CHOCOPhlAn pangenome database version v201901_v31 and the UniRef90-
671 DIAMOND 2019-01 protein database obtained through the HUMAnN3 download utility both on
672 August 11, 2025. UniRef gene families were mapped to KEGG Ortholog (KO) groups using the
673 HUMAnN3-provided mapping file map_ko_uniref90.txt.gz (downloaded December 5, 2025).
674 Gene family and pathway read counts were summed across all taxa to generate community-wide
675 functional profiles. Pathway abundance was automatically normalized in the MaAsLin3 analysis
676 described below via the "total sum scaling" setting.
677
678 To identify pathways differing between the upper and lower gut, we analyzed the raw pathway
679 count matrix using the MaAsLin3[33] R package version 1.2.0. MaAsLin3 was downloaded on
680 December 7, 2025 and run with the default parameters on the raw pathway count matrix
681 (minimum feature abundance of 0.0001, minimum prevalence of 0.1, and default FDR

682 correction). We specified the model formula: ~ coarse_position + cage + read_counts +
683 strata(subject_id) with normalization = TSS, transformation = LOG. Gut location, read count and
684 cage were modeled as fixed-effect covariates and stratification by subject ID. P-values were
685 adjusted for multiple comparisons using the Benjamini–Hochberg false discovery rate (FDR)
686 procedure, as implemented in MaAsLin3, and pathways with $q < 0.05$ were considered
687 significant.
688
689 **Quantifying SNV frequencies**

690 We estimated SNV frequencies for each species in each sample. Downstream, these frequencies
691 were used to estimate nucleotide diversity, identify quasi-phaseable samples, and quantify SNV
692 frequency differences between pairs of quasi-phaseable samples. To estimate SNV frequencies in
693 humanized mouse and Shalon et al., 2023 samples, reads were mapped to representative
694 reference genomes using default MIDAS 1 mapping thresholds and arguments[23]. SNV
695 frequencies were estimated in conventional mouse data using MIDAS 3, with mapping
696 thresholds and arguments set to match MIDAS 1 defaults.

697 To obtain accurate estimates of SNV frequencies, we imposed coverage requirements. MIDAS 1
698 and 3 report the read coverage D for each site in the reference genome for each sample. We
699 excluded any sample with a mean $D < 5X$ across all protein coding sites, as was done in Garud *et*
700 *al.*, 2019. We further excluded from analysis species non-zero coverage at ≥ 40% of reference
701 sites within a sample. Additionally, to minimize the possibility of any mapping errors, we
702 implemented two controls. First, as described below and as was done in Garud *et al.*, 2019[23], we
703 exclude genes with copy number greater than 3 in at least one sample in our study, as they were
704 potential candidates for read stealing. Second, we exclude individual sites with a coverage value
705 $< 0.3\overline{D}$ or $> 3\overline{D}$ where $\overline{D}$ is mean genome-wide coverage, as these sites had abnormal coverages
706 outside the range of expectations for single copy genes.

707 Finally, SNV frequency changes were identified only if median coverage values between two
708 samples from the same mouse were within a factor of three of each other, as we expect that
709 coverage should be relatively constant within a mouse. Below we describe additional coverage
710 filters that were imposed for various calculations, including SNV frequency changes.

711 **Annotating SNVs as synonymous versus nonsynonymous**

712 MIDAS 1 annotated SNVs as synonymous and nonsynonymous based on the reading frames of
713 genes as annotated in the PATRIC database[80], and MIDAS 3 annotated SNVs as synonymous
714 and nonsynonymous based on annotations from Prokka[81].

715 **Quantifying gene copy number**

716 We estimate copy number variation (CNV) to exclude any genes from further analysis that may
717 be potential candidates for read stealing or donating. To estimate CNVs in humanized mice and
718 Shalon et al., 2023 data, reads were mapped to pangenomes using default MIDAS 1 mapping
719 thresholds and arguments[23]. CNVs were estimated in conventional mouse data using MIDAS 3
720 with mapping thresholds and arguments set to match MIDAS 1 defaults.

721 To minimize read mapping errors, we identified genes with abnormally high gene copy number
722 values, which may be potentially indicative of erroneous read recruitment[82]. MIDAS 1 and 3
723 estimates CNVs by mapping reads to a *pangenome* consisting of several sequenced isolates
724 belonging to the same species. Average gene coverage was calculated as the number of reads
725 mapping to that gene normalized by gene length. Both versions of MIDAS compute copy
726 number, c, as the ratio between a gene's average coverage and the median coverage of a set of 15
727 single-copy marker genes belong to that species. Genes with a copy number $c > 3$ in at least one
728 sample in our study were excluded during post-processing, as they were potential candidates for
729 stealing. To exclude any additional genes that may suffer from erroneous read recruitment, we
730 further filtered out any genes with a copy number $c > 3$ in the human samples analyzed in
731 Garud *et al.*, 2019[23]. Finally, any genes sharing ≥ 95% average nucleotide identity with any other
732 gene across species boundaries as identified in a set of "blacklisted" genes in Garud *et al.*, 2019
733 (section A.iv)[23] were excluded to further prevent erroneous read mapping.

734 **Calculating alpha and beta diversity of gut samples**

735 Alpha diversity was computed for humanized and conventional mouse data using the Shannon
736 diversity index as implemented by the Vegan package in R[83] (**Figure 2A**, **Supplementary**
737 **Figure 9A**). To assess if Shannon diversity between gut regions was significantly different, we
738 used a paired Wilcoxon signed-rank test comparing pairs of samples from one region versus the

739 other from the same mouse. To calculate fold change in alpha diversity between the small and
740 large intestines, alpha diversity was averaged across small intestinal regions (duodenum,
741 jejunum, and ileum) and large intestinal regions (cecum and colon), and the per mouse fold
742 change was calculated as the ratio of the averaged large intestinal and small intestinal alpha
743 diversity values. The fold change values presented in the main text represent the average of these
744 per-mouse estimates.

745 To produce the stacked bar plots in **Figure 2B** and **Supplementary Figure 9B**, species
746 abundances within taxonomic families were summed. Species with less than 0.1% abundance
747 were excluded. To test for differential abundance of specific families between the small and
748 large intestines of humanized mice, relative abundances for each family were averaged across the
749 small intestine (duodenum, jejunum, and ileum) and large intestine (cecum and colon) in each
750 host, the wilcox.test() in R was used to apply a paired Wilcoxon signed-rank test to compare
751 average large and small intestinal relative abundances (paired by mouse) for each species,
752 generating per-species estimates for the median log2 fold change in mean relative abundance
753 between large and small intestinal samples in the same mouse, as well as p-values and 95%
754 confidence intervals to assess significance (**Supplementary Figure 1**).

755 **Calculating nucleotide diversity**

756 For every humanized mouse and sample and the inoculum, we estimated π, a population-level
757 metric of nucleotide diversity that represents the probability of two randomly sampled genomes
758 from the population have different alleles at a base pair. To do so, we used the formula for
759 nucleotide diversity applied in Schloissnig *et al.*, 2013[84], which accounts for total read counts:

760 $$\pi(S,G) = \frac{1}{|G|}\sum_{i=1}^{|G|}\sum_{B_1\in\{ATGC\}}\sum_{B_2\in\{ATGC\}\setminus B_1}\frac{x_{i,B_1}}{D_i}\frac{x_{i,B_2}}{D_i - 1}$$

761 Here, S is the sample; G is the genome of the focal species, with $|G|$ being its size; i is the locus
762 in the genome; $B_1$and B_2 are the reference and alternative alleles, respectively; x_{i,B_j} is the
763 number of reads with allele B_j at locus i; and D_i is the total read depth at locus i. This formula is
764 an extension of a previously proposed π estimator on next generation sequencing data devised by

765 Begun *et al.*, 2007[85]. Sites with read coverage less than four were excluded from π calculations,
766 such that $|G|$ is equivalent to the total number of sites with adequate coverage rather than the
767 total size of the genome.

768 Samples from mice 7 and 8 were sequenced to a substantially higher depth compared to samples
769 from mice 1 through 6 (**Supplementary Table 1**), resulting in elevated levels of coverage per bp
770 for bacterial species in mice 7 and 8. As a consequence, there was an increased opportunity to
771 detect higher levels of nucleotide diversity in mice 7 and 8 merely due to systematic differences
772 in coverage. To be able to compare nucleotide diversity levels between mice, we applied a
773 random down sampling per nucleotide site such that the depth of mice 7 and 8 matched the
774 median coverage of mice 1 through 6.

775 Distributions of π in **Figure 3A** were plotted for the 30 species which met the following criteria:
776 minimum coverage of four reads in the inoculum and at least three mouse samples representing
777 at least two different mouse hosts. Species were only considered in samples in which at least
778 500,000 nucleotides met the minimum coverage threshold.

779 **Calculating F_{ST}**

780 π estimates were used to compute genome-wide F_{ST} as a quotient of the mean genome-wide π
781 subtracted from 1:

782 $$F_{ST}(S_1, S_2, G) = 1 - \frac{\pi_{within}}{\pi_{between}} = 1 - \frac{\left(\pi(S_1, G) + \pi(S_2, G)\right)/2}{\pi(S_1, S_2, G)}$$

783 Where π_{within} is given by the formula for $\pi(S, G)$ provided above, and pairwise $\pi_{between}$ is
784 given by the following formula also derived in Schloissnig *et al.*, 2013*:*

785 $$\pi(S_1, S_2, G) = \frac{1}{|G|} \sum_{i=1}^{|G|} \sum_{B_1 \in \{ATGC\}} \sum_{B_2 \in \{ATGC\} \setminus B_1} \frac{x_{i,B_1,S_1}}{D_{i,S_1}} \frac{x_{i,B_2,S_2}}{D_{i,S_2}}$$

786 Where k in S_k designates sample 1 or 2, respectively. We used ileum and colon to represent the
787 small and large intestine, respectively because those regions had the highest mean read coverage

788 across all samples in each gut region. In order to accurately estimate allele frequencies, each site
789 was only included in the F_{ST} calculation if a minimum of 4 reads was observed at that site for
790 both samples being compared.

791 **Inferring strain frequencies**

792 Previous work has shown that hosts tend to be colonized by multiple genetically distinct strains
793 (typically 1-4) of the same species[23,86]. When levels of recombination between strains within a
794 host is sufficiently low, these strains may co-exist as genetically distinct subpopulations[24].
795 Typically, two randomly drawn strains from different hosts harbor on order 10^3 - 10^4 SNV
796 changes[23,47,84,87], and two strains residing within a host are expected to be similarly diverged[23].
797 We leverage this expected divergence between strains to distinguish between the strain
798 genotypes and infer their overall frequencies.

799 Inferring strain genotypes and frequencies from short read sequencing data is a difficult problem
800 because linkage between co-segregating SNVs residing on the same strain backbone are
801 destroyed by the fact that they are unlikely to reside on the same short read. Even perfectly
802 linked SNVs may be difficult to assign to the correct strain given the potential for sampling
803 noise. However, recent work has shown that strain inference is feasible when there are many
804 samples collected from the same hosts, as alleles belonging to the same strain will display highly
805 correlated allele frequencies across samples[24,25,65,88–90].

806 We leverage recent work by Roodgar *et al.*, 2021[24] and extended by Wolff *et al.,* 2023[25] to infer
807 strain identity and frequency in samples from all three datasets analyzed in this study. The papers
808 cited above inferred strain genotypes and frequencies from *temporal* data. Here, we use *spatial*
809 data both within mice and across mice, and a mixture of *spatial* and *temporal* data within human
810 hosts sampled in Shalon *et al.*, 2023 to infer strain genotypes and frequencies within samples.
811 Because humanized mice were inoculated with the same human sample and conventional mice
812 shared homogenized microbiomes due to co-housing, we inferred strain frequency jointly across
813 all mouse samples belonging to the same dataset, reasoning that mice from the same dataset all
814 shared the same subset of strains. For Shalon *et al.*, 2023 data, we inferred strain frequency

815 jointly across all samples belonging to the same host, likewise reasoning that these samples
816 carried the same subset of strains.

817 Briefly, the approach clusters SNVs segregating on the same strain background by detecting
818 correlations between SNVs with highly similar allele frequencies across samples. The degree of
819 correlation between a pair of SNVs i and j is measured by distance metric d between allele
820 frequency trajectories $\hat{f}_i$ and $\hat{f}_j$:

821 $$d(\hat{f}_i, \hat{f}_j) = \frac{1}{S}\sum_{s=1}^{S} \frac{2(D_{is} + D_{js})(\hat{f}_{is} - \hat{f}_{js})^2}{(\hat{f}_{is} + \hat{f}_{js})(1 - \hat{f}_{is} + 1 - \hat{f}_{js})},$$

822 where S is the number of samples, $\hat{f}_{is}$ and $\hat{f}_{js}$ are the frequency of the alleles in a sample s, and
823 D_{is} and D_{js} represent the read depth of the alleles in sample s. SNVs that are in perfect linkage
824 with one another are expected to have a $d = 0$.

825 To ensure that high-quality and informative sites were included in our stain inference procedure,
826 we only considered loci that had a read coverage of $D \geq 10$ and were polymorphic in at least
827 20% of samples in which the species was detected. Filtering out non-polymorphic loci prevented
828 the algorithm from generating clusters at allele frequencies of 1 or 0 that reflect low frequency
829 variation within a strain population as opposed to actual genetically distinct strains.

830 Detecting correlated clusters of SNVs belonging to the same strain background can be
831 confounded without knowing if the alternative or reference allele belongs to one strain versus
832 another. To address this issue, as done in Roodgar *et al.*, 2021[24], we set d between a pair of loci
833 to the minimum of $d(\hat{f}_i, \hat{f}_j)$ and $d(\hat{f}_i, 1 - \hat{f}_j)$ to ensure that polarization differences (i.e., calling
834 an allele frequency 0.2 instead of 0.8, based on the reference allele used in the reference genome)
835 between loci did not prevent them from being clustered together.

836 To identify SNVs linked on the same strain background, we applied a greedy, network-based
837 algorithm to extract large cluster SNVs with highly correlated allele frequencies across samples.
838 The algorithm begins by forming a network in which nodes represent SNVs, with pairs SNVs
839 connected by an edge if they have a distance $d < 3.5$, which represents a maximum distance
840 threshold below which pairs of SNVs are likely to be linked on the same strain background, as

841 demonstrated in Roodgar *et al.*, 2021[24]. Next, the algorithm identifies a focal SNV as the SNV
842 with the maximum number of connections (i.e., edges) and extracts all polymorphic SNVs
843 connected to it. This cluster of SNVs was designated as a strain if it included at least 10^3 SNVs,
844 as this represents the lower bound for the typical number of alleles that are expected to segregate
845 between genetically distinct strains[23,47,84,87]. Note that SNVs connected to the focal SNV were
846 only extracted if they were also connected to at least 25% of other SNVs within a cluster. SNVs
847 included in the initial cluster were removed from the overall set of variable sites, and the
848 clustering process was repeated until no additional clusters of at least 10^3 SNVs were identified.
849 Strain frequencies were then inferred as the mean SNV allele frequencies in each cluster
850 (**Supplementary Figure 14**). However, despite requiring strain clusters to have a minimum of
851 10^3 SNVs, not all SNVs met our coverage requirement of $D \geq 10$ across all samples.
852 Therefore, we only report strain frequencies for samples in which every cluster has at least 100
853 high coverage SNVs ($D \geq 10$) to ensure that the inferred strain frequencies were well-
854 supported.

855 95% confidence intervals for strain frequencies in each sample were generated using a stationary
856 bootstrapping approach[91]. Specifically, 1,000 bootstrap samples each consisting of 100 SNVs
857 were drawn with replacement from each strain cluster within each sample, and subsequently the
858 mean of the subsampled SNV frequencies in each bootstrap sample was computed as strain
859 frequency. The upper and lower bounds of the 95% confidence interval was calculated as the 2.5
860 and 97.5 percentile of the bootstrapped strain frequency distribution.

861 In addition to the filters used by Wolff *et al.*, 2023, we required minor alleles to have a read
862 support of at least four reads to prevent rare mutations and sequencing errors from erroneously
863 giving rise to spurious low-frequency clusters inferred from conventional and humanize mouse
864 data. For strain frequency profiles in Shalon *et al.*, 2023, we filtered out spurious low-frequency
865 clusters and lowly abundant strains using a different approach by requiring all strains rise above
866 a frequency of 5% in at least one sample to be included in downstream analyses.

867 We visually inspected all strain frequency profiles and either applied manual modifications to the
868 algorithm or excluded from downstream analysis all profiles which violated one of the following
869 expectations. In a scenario when only one strain is present, we expect *zero* clusters to be

detected. In a scenario in which two strains are present, we expect only *one* cluster, representing the alleles segregating between the two strains. The relative frequencies of the two strains can be inferred as $\bar{f}$ and $1 - \bar{f}$, where $\bar{f}$ is the mean frequency of the clustered SNVs. In a scenario in which three strains are present, we expect *three* clusters the frequencies of which roughly sum to 1 in each sample, representing the SNV segregating between each of the three pairs of strains. This second scenario did not occur in our humanized or conventional mouse data. In the Shalon *et al.*, 2023 data, the strain frequency inference pipeline did at times yield a maximum of three strain clusters. However, manual inspection revealed that these clusters tended to sum to an overall frequency significantly greater or less than 1 in some samples, suggesting that the strain frequency pipeline was failing to accurately capture all strains or spuriously inferring too many. As a result, we excluded species x host pairs with three strain clusters and limited our analysis to species x host pairs that had two strains. At times, the strain inference algorithm outputs two clusters, which is inconsistent with a scenario of either 1, 2 or 3 strains colonizing and instead may be arising from high variance in allele frequency. This happened once in the humanized mouse data with *B. uniformis*, the strain inference algorithm inferred *two* clusters (**Supplementary Figure 15A**). It was visually apparent in this case that the two clusters likely represented the same group of SNVs segregating between two strains' genetic backgrounds because of their highly correlated trajectories, but were separated into two distinct clusters due to there being a large variance in individual SNV frequency. In this singular instance, we reclustered by using the mean SNV frequency of all SNVs in both clusters as a focal SNV (**Supplementary Figure 15B**) and extracting all other SNVs connected to it by an edge in the network (i.e., with $d < 3.5$). This method produced a single cluster. The algorithm was able to infer strain frequencies for all other species analyzed in this study from the humanized and conventional mouse datasets without the need for *ad hoc* modifications to the pipeline. For several species x host pairs in the Shalon *et al.*, 2023 data, the strain inference algorithm yielded two clusters. These species x host pairs were excluded from downstream analysis of strain frequency in favor of species x host pairs with more unambiguous strain phasing results, all of which had two strains (*i.e.*, a single cluster).

When applying this algorithm to the humanized mouse data, it in some instances inferred only single strain for a bacterial species across all samples despite the inoculum displaying high

900 nucleotide diversity, suggestive of multiple strains being present in the inoculum. This
901 phenomenon may occur when applying the algorithm to a collection of samples in which a strain
902 is present in only one sample (*e.g.,* the inoculum) but absent in all others (*e.g.*, the mouse
903 samples). As a result, the algorithm would be unable to detect correlations between the SNVs
904 private to a single sample and therefore fail to detect that strain. We identified when this might
905 be happening in the humanized mouse data by looking for species in which only a single strain
906 was inferred (i.e., the algorithm produced no clusters), but which had at least 1,000 SNVs in the
907 inoculum that were polymorphic and had a read support of four reads for the minor allele. When
908 this scenario occurred, we assumed two strains were present in the inoculum and inferred the
909 strain frequency $f_{strain\ 1,inoculum}$ of the first strain as the mean of the allele frequencies of these
910 polymorphic SNVs in the inoculum, and the strain frequency $f_{strain\ 2,inoculum}$ of the second
911 strain as $1 - f_{strain\ 1,inoculum}$. We generated 95% confidence intervals by applying the same
912 stationary bootstrapping approach to the polymorphic SNVs identified in the inoculum.

913 **Calculating change in major strain frequency Δf between samples**

914 The change in major strain frequency Δf was calculated as the magnitude of the difference in
915 frequencies of one of the strains between two samples. Because all samples with co-colonizing
916 strain analyzed in this study harbored two strains, Δf was the same regardless of whether strain 1
917 or 2 was used to calculate the frequency change. For the humanized and conventional mouse
918 data, we calculated Δf between samples for co-colonizing species present in at least at least two
919 mice, at least two cages, and at least two gut regions. When visualizing the distribution of Δf
920 values, we categorized pairwise comparisons as being "within host", "between host, within
921 cage", and "between host, between cage". For each species, we selected a single "within host"
922 pair per mouse, a single "between host, within cage" pair per cage, and a single "between host,
923 between cage" pair for the three possible cage-wise comparisons. For "within host" comparison
924 within each host, we selected the pair of samples that maximized the distance between the two
925 sampling sites, breaking ties by selecting the pair of samples with the sample that is furthest
926 down the gut (i.e., closest to the colon). For between host comparisons (both within and between
927 cage), we selected the pair of samples that represented the same gut region or two gut regions
928 proximal to one another along the gut, breaking ties by selecting pairs of samples with the
929 sample that is furthest down the gut (i.e., closest to the colon). For "between host, between cage"

comparisons, we ensured that no mouse was involved in more than one cross-cage comparison to eliminate redundancy. For Shalon et al., 2023 data, we calculated Δf between samples belonging to co-colonized species x host pairs if the species were present in at least two timepoints and at least two device types. When visualizing the distribution of Δf values, we categorized pairwise comparisons as being "within timepoint, between gut region", "between consecutive timepoints", and "between host first and last timepoint". For each species x host pair, we selected a single "within timepoint" pair per timepoint, a single "between consecutive timepoints" pair for each consecutive pair of timepoints, and a single "between host first and last timepoint" pair. For "within timepoint" comparisons within each timepoint, we selected the sample pair that maximized the distance between device types (i.e., biasing towards device type 1 versus device type 4 comparisons, if available), breaking ties by selecting the sample pair with the maximum device type number (i.e., biasing towards pairs with device type 4). For between timepoint comparisons (both consecutive and first and last timepoints), the sample pair that minimized the distance between device types and was furthest along the gut was selected.

Performing ANOVA of major strain frequencies

To quantify the variance in major strain frequency across samples, we applied an ANOVA approach. In the humanized mouse dataset, the major strain was designated as the strain at frequency $f > 0.5$ in the inoculum. In conventional mouse and Shalon *et al.*, 2023 data, the major strain was designated as the strain at frequency $f > 0.5$ in a majority of samples. The centered log ratio (CLR) transformation was applied to major strain frequency in all samples in which that species was detected, a commonly used approach to remove compositional artifacts when applying non-compositional approaches such as ANOVA[92,93]. For each species, we built an ANOVA model using the Stats package in R. In the humanized and conventional mouse data, the model was specified with CLR-transformed major strain frequency as the dependent variable and cage, mouse, and gut region as the independent categorical variables, and was applied jointly to all samples belonging to the same study. Variance explained was calculated as the sum of squares between groups corresponding to each variable divided by the sum of squares total. We applied several prevalence requirements to ensure that the amount of variance explained by each dependent variable was properly quantified. In humanized and conventional mouse data, we only performed ANOVA on species detected across at least two cages (with the species detectable in

960 at least two mice in those cages), at least two mice (with the species detectable in at least two gut
961 regions in those mice), and two gut regions (with the species detectable in these same gut regions
962 in at least two mice).

963 **Identifying quasi-phaseable lineages**

964 SNV frequency differences between samples can arise due to evolution, or due to shifts in the
965 frequency of co-colonizing strains, which are likely to harbor thousands of differing nucleotides
966 between them. To distinguish between SNV frequency differences that arise due to evolution
967 versus those arising from strain frequency shifts, we identified quasi-phaseable lineages using an
968 approach developed in Garud *et al.*, 2019[23]. By inferring quasi-phaseable lineages, we then were
969 able to confidently infer evolutionary changes.

970 Briefly, the approach works by identifying samples in which a particular species has a single
971 dominant strain at $f \geq 0.8$, hereafter referred to as a quasi-phaseable sample. In such samples,
972 the allele corresponding to the dominant strain can be inferred with statistical confidence.
973 Specifically, the alleles with frequency $f \geq 0.8$ at genomic loci with coverage $D \geq 20$ are
974 assigned to the quasi-phaseable sample. Only samples with a median coverage of 20X for the
975 species of interest are processed through the quasi-phaseable pipeline. Employing such stringent
976 coverage and allele frequency thresholds ensures that the incorrect allele is not inferred, as it is
977 statistically improbable that sampling error would give rise to the incorrect allele[23].

978 We made an important alteration to the quasi-phaseable pipeline in its application to the
979 humanized and conventional mouse data relative to its application in Garud *et al.*, 2019[23].
980 Normally, the quasi-phaseable pipeline designates a sample as having a quasi-phaseable strain if
981 the number of intermediate frequency SNVs (i.e., SNVs with frequency $0.2 < f < 0.8$) which a
982 species harbors within the sample are 10% or less than the number of SNV differences that are
983 typically observed for that species between independent hosts. This heuristic arises from the
984 assumption that independent hosts should harbor genetically distinct strains that have diverged
985 over long evolutionary timescales, and is born out by empirical evidence[23]. The 10% threshold is
986 calibrated independently for each species, as genetically distinct strains of one species are not
987 necessarily expected to show the same level of nucleotide differentiation as genetically distinct

988 strains of a different species. In previous applications of the quasi-phaseable pipeline, the
989 number of SNV differences between independent hosts has been assessed from the dataset used
990 in those studies. For example, Garud *et al.*, 2019[23] studied evolutionary dynamics within 693
991 unique individuals sampled as a part of the human microbiome project[94,95], the TwinsUK
992 registry[96], a cohort of Chinese subjects[97], and four additional young twins[98]. Because most hosts
993 included in this compiled dataset were unrelated (and in some cases, on different continents), the
994 assumption that their resident strains were distantly diverged was a reasonable one. By contrast,
995 the humanized and conventional mice are expected to have low between-host strain diversity,
996 either because they were inoculated with the stool from a single human host or because their
997 microbiomes were homogenized early on in development through co-housing. As a result, we
998 could not accurately quantify the number of differences that are expected between two unrelated
999 hosts, making it impossible to detect quasi-phaseable strains using this dataset alone. We took
1000 two different approaches to overcome this issue in humanized and conventional mice,
1001 respectively. In humanized mice, we calculated the expected number of between-host SNV
1002 differences for each focal species based on the aforementioned data used in Garud *et al.*, 2019[23],
1003 which we had previously processed through MIDAS 1. Because this resource was not available
1004 for the conventional mouse data processed through MIDAS 3, we designated samples as quasi-
1005 phaseable if intermediate frequency sites (i.e., SNVs with frequency $0.2 < f < 0.8$) occurred at
1006 a rate of 1×10^{-4} per bp or less, consistent with rates observed in quasi-phaseable samples from
1007 humanized mice. Results published in this study (**Figure 3A**) and elsewhere[23] suggest that this is
1008 a highly conservative threshold for the maximum amount of intermediate frequency diversity we
1009 would expect to see if only a single strain were present within a host.

1010 **Inferring SNV frequency changes**

1011 To detect SNV frequency changes arising as a result of evolution, we follow the approach in
1012 Garud *et al.*, 2019. A SNV frequency change is defined as an allele with low frequency in one
1013 quasi-phaseable sample ($f \leq 0.2$) and high frequency in another quasi-phaseable sample ($f \geq$
1014 0.8). SNV differences were only considered at loci that had a coverage $D \geq 20$ in both samples
1015 being compared. An extreme allele frequency change of this magnitude and this minimum depth
1016 is statistically unlikely to arise between quasi-phaseable strains due to sampling error or drift
1017 given large population sizes and short timescales considered in this study[23,43]. Pairs of samples

1018 harboring ≤ 20 SNV differences were considered to have the same inoculum strain evolving over
1019 short timescales (e.g., the eight-week timescale considered in this study), as larger amounts of
1020 SNV differences, such as $\mathcal{O}(10^3)$ or greater are inconsistent with within-host rates of
1021 diversification[47,82,84,87]. Consequently, pairs of quasi-phaseable strains with > 20 SNV differences
1022 were excluded from this evolutionary analysis. For more information, see the Supplementary
1023 Information for Garud *et al.*, 2019[23].

1024 **Bootstrapping confidence intervals for SNV difference rates**

1025 To measure whether there were significant differences in the number of SNV frequency changes
1026 occurring between samples within hosts, between hosts, and between the inoculum and mouse
1027 hosts in the humanized mouse data, we employed a bootstrap approach to compute confidence
1028 intervals for the rate of SNV difference between quasi-phaseable pairs of samples with the same
1029 dominant strain. In 1,000 bootstrap iterations, we sampled 100 quasi-phaseable sample pairs (not
1030 necessarily of the same species) with replacement from each comparison category: within host,
1031 between host, and between inoculum and host. For the within-cage comparison category, we
1032 only considered quasi-phaseable pairs from different hosts. We computed the actual rate of SNV
1033 frequency changes for a particular comparison category as the total number of SNVs displaying
1034 extreme allele frequency changes across all sampled quasi-phaseable pairs divided by the total
1035 number of high coverage loci across all sampled quasi-phaseable pairs. We calculated 95%
1036 confidence intervals using the cumulative distribution function (CDF) of the overall SNV change
1037 difference rate in each comparison category generated from the 1,000 bootstrap iterations, taking
1038 the 2.5 and 97.5 percentile as the upper and lower bound, respectively (**Figure 5B**). We applied
1039 an identical approach for assessing whether SNV frequency change rates were significantly
1040 different between samples from the same host or different hosts in the conventional mouse data
1041 (**Figure 6C**).

1042 Data Availability Statement

1043 Raw metagenomics sequencing reads are available on NCBI under BioProject [project accession
1044 number will be made available upon publication]. All necessary metadata, as well as the source

code for the sequencing pipeline, downstream analyses, and figure generation, are available at GitHub (https://github.com/garudlab/Wasney-Briscoe/).

Acknowledgments

The authors thank Kristin Harper, PhD, of Harper Health & Science Communications, LLC, for providing editorial support in accordance with Good Publication Practice guidelines, and Alison Feder, PhD, for her helpful edits. The authors also thank the Garud and the Tropini labs for their valuable inputs. In particular, the authors thank Peter Laurin for his statistical advice and feedback on the manuscript. Finally, the authors thank Dr. Sidhartha Sinha and members of the Sinha Lab at Stanford University for helpful discussions. This work was funded by NIGMS NIH award R35GM151023 (to NRG), NSF CAREER award (no. 2240098) (to NRG), a Paul Allen Distinguished Investigator Award (to CT and NRG), a Canadian Institutes of Health Research Team Grant: Canadian Microbiome Initiative 2 (to CT), Crohn's and Colitis Canada, Canadian Institute for Advanced Research (to CT), and the Michael Smith Foundation for Health Research Scholar Award (to CT).

References

1. McCallum, G. & Tropini, C. The gut microbiota and its biogeography. *Nat. Rev. Microbiol.* **22**, 105–118 (2024).
2. Gu, S. *et al.* Bacterial Community Mapping of the Mouse Gastrointestinal Tract. *PLOS ONE* **8**, e74957 (2013).
3. Ng, K. M. *et al.* Single-strain behavior predicts responses to environmental pH and osmolality in the gut microbiota. *mBio* **14**, e00753-23 (2023).
4. D'Adamo, G. L. *et al.* Bacterial clade-specific analysis identifies distinct epithelial responses in inflammatory bowel disease. *Cell Rep. Med.* **4**, (2023).
5. Morowitz, M. J., Carlisle, E. & Alverdy, J. C. Contributions of Intestinal Bacteria to Nutrition and Metabolism in the Critically Ill. *Surg. Clin. North Am.* **91**, 771–785 (2011).
6. Miller, B. M., Liou, M. J., Lee, J.-Y. & Bäumler, A. J. The longitudinal and cross-sectional heterogeneity of the intestinal microbiota. *Curr. Opin. Microbiol.* **63**, 221–230 (2021).
7. Dapa, T. *et al.* Within-host evolution of the gut microbiome. *Curr. Opin. Microbiol.* **71**, 102258 (2023).
8. Dapa, T., Ramiro, R. S., Pedro, M. F., Gordo, I. & Xavier, K. B. Diet leaves a genetic signature in a keystone member of the gut microbiota. *Cell Host Microbe* **30**, 183-199.e10 (2022).
9. Yang, Y. *et al.* Within-host evolution of a gut pathobiont facilitates liver translocation. *Nature* **607**, 563–570 (2022).
10. Collins, J. *et al.* Dietary trehalose enhances virulence of epidemic Clostridium difficile. *Nature* **553**, 291–294 (2018).

1079 11. Paepe, M. D. *et al.* Trade-Off between Bile Resistance and Nutritional Competence Drives
1080 Escherichia coli Diversification in the Mouse Gut. *PLOS Genet.* **7**, e1002107 (2011).
1081 12. Conway, T. & Cohen, P. S. Commensal and Pathogenic Escherichia coli Metabolism in the Gut.
1082 in *Metabolism and Bacterial Pathogenesis* 343–362 (John Wiley & Sons, Ltd, 2015).
1083 doi:10.1128/9781555818883.ch16.
1084 13. Barroso-Batista, J. *et al.* Specific Eco-evolutionary Contexts in the Mouse Gut Reveal
1085 Escherichia coli Metabolic Versatility. *Curr. Biol. CB* **30**, 1049-1062.e7 (2020).
1086 14. Vasquez, K. S. *et al.* Quantifying rapid bacterial evolution and transmission within the mouse
1087 intestine. *Cell Host Microbe* **29**, 1454-1468.e4 (2021).
1088 15. Kim, S. G. *et al.* Microbiota-derived lantibiotic restores resistance against vancomycin-resistant
1089 Enterococcus. *Nature* **572**, 665–669 (2019).
1090 16. Dubinsky, V., Dotan, I. & Gophna, U. Strains Colonizing Different Intestinal Sites within an
1091 Individual Are Derived from a Single Founder Population. *mBio* **0**, e03456-22 (2023).
1092 17. Wu-Woods, N. J. *et al.* Microbial-enrichment method enables high-throughput metagenomic
1093 characterization from host-rich samples. *Nat. Methods* **20**, 1672–1682 (2023).
1094 18. Johansson, M. E. V. *et al.* Normalization of Host Intestinal Mucus Layers Requires Long-Term
1095 Microbial Colonization. *Cell Host Microbe* **18**, 582–592 (2015).
1096 19. Seedorf, H. *et al.* Bacteria from diverse habitats colonize and compete in the mouse gut. *Cell* **159**,
1097 253–266 (2014).
1098 20. Turnbaugh, P. J. *et al.* The Effect of Diet on the Human Gut Microbiome: A Metagenomic
1099 Analysis in Humanized Gnotobiotic Mice. *Sci. Transl. Med.* **1**, 6ra14 (2009).
1100 21. Sonnenburg, E. D. *et al.* Diet-induced extinction in the gut microbiota compounds over
1101 generations. *Nature* **529**, 212–215 (2016).
1102 22. Tikhonov, M., Leach, R. W. & Wingreen, N. S. Interpreting 16S metagenomic data without
1103 clustering to achieve sub-OTU resolution. *ISME J.* **9**, 68–80 (2015).
1104 23. Garud, N. R., Good, B. H., Hallatschek, O. & Pollard, K. S. Evolutionary dynamics of bacteria in
1105 the gut microbiome within and across hosts. *PLOS Biol.* **17**, e3000102 (2019).
1106 24. Roodgar, M. *et al.* Longitudinal linked-read sequencing reveals ecological and evolutionary
1107 responses of a human gut microbiome during antibiotic treatment. *Genome Res.* **31**, 1433–1446 (2021).
1108 25. Wolff, R., Shoemaker, W. & Garud, N. Ecological Stability Emerges at the Level of Strains in the
1109 Human Gut Microbiome. *mBio* **14**, e02502-22 (2023).
1110 26. Nayfach, S., Rodriguez-Mueller, B., Garud, N. & Pollard, K. S. An integrated metagenomics
1111 pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome*
1112 *Res.* **26**, 1612–1625 (2016).
1113 27. Faith, J. J. *et al.* The Long-Term Stability of the Human Gut Microbiota. *Science* **341**, 1237439
1114 (2013).
1115 28. Tropini, C., Earle, K. A., Huang, K. C. & Sonnenburg, J. L. The Gut Microbiome: Connecting
1116 Spatial Organization to Function. *Cell Host Microbe* **21**, 433–442 (2017).
1117 29. Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota.
1118 *Nat. Rev. Microbiol.* **14**, 20–32 (2016).
1119 30. Vaga, S. *et al.* Compositional and functional differences of the mucosal microbiota along the
1120 intestine of healthy individuals. *Sci. Rep.* **10**, 14977 (2020).
1121 31. Zafar, H. & Saier, M. H. Gut Bacteroides species in health and disease. *Gut Microbes* **13**,
1122 1848158.

1123 32. Beghini, F. *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial
1124 communities with bioBakery 3. *eLife* **10**, e65088 (2021).
1125 33. Nickols, W. A. *et al.* MaAsLin 3: Refining and extending generalized multivariable linear models
1126 for meta-omic association discovery. *BioRxiv Prepr. Serv. Biol.* 2024.12.13.628459 (2024)
1127 doi:10.1101/2024.12.13.628459.
1128 34. Li, X. *et al.* Uncovering de novo polyamine biosynthesis in the gut microbiome and its alteration
1129 in inflammatory bowel disease. *Gut Microbes* **17**, 2464225 (2025).
1130 35. Kennedy, M. S. *et al.* Dynamic genetic adaptation of *Bacteroides thetaiotaomicron* during murine
1131 gut colonization. *Cell Rep.* **42**, 113009 (2023).
1132 36. Candeliere, F. *et al.* β-Glucuronidase Pattern Predicted From Gut Metagenomes Indicates
1133 Potentially Diversified Pharmacomicrobiomics. *Front. Microbiol.* **13**, (2022).
1134 37. Walter, J. Ecological Role of Lactobacilli in the Gastrointestinal Tract: Implications for
1135 Fundamental and Biomedical Research. *Appl. Environ. Microbiol.* **74**, 4985–4996 (2008).
1136 38. Mager, L. F. *et al.* Microbiome-derived inosine modulates response to checkpoint inhibitor
1137 immunotherapy. *Science* **369**, 1481–1489 (2020).
1138 39. Wright, S. The Genetical Structure of Populations. *Ann. Eugen.* **15**, 323–354 (1949).
1139 40. Wright, S. The genetical structure of populations. *Ann. Eugen.* **15**, 323–354 (1951).
1140 41. Frazão, N. & Gordo, I. Shared Evolutionary Path in Social Microbiomes. *Mol. Biol. Evol.* **40**,
1141 msad153 (2023).
1142 42. Vasquez, K. S. *et al.* High-resolution lineage tracking of within-host evolution and strain
1143 transmission in a human gut symbiont across ecological scales. 2024.02.17.580834 Preprint at
1144 https://doi.org/10.1101/2024.02.17.580834 (2024).
1145 43. Ghosh, O. M. & Good, B. H. Emergent evolutionary forces in spatial models of luminal growth
1146 and their application to the human gut microbiota. *Proc. Natl. Acad. Sci.* **119**, e2114931119 (2022).
1147 44. Wichman, H. A., Badgett, M. R., Scott, L. A., Boulianne, C. M. & Bull, J. J. Different
1148 Trajectories of Parallel Evolution During Viral Adaptation. *Science* **285**, 422–424 (1999).
1149 45. Woods, R., Schneider, D., Winkworth, C. L., Riley, M. A. & Lenski, R. E. Tests of parallel
1150 molecular evolution in a long-term experiment with Escherichia coli. *Proc. Natl. Acad. Sci.* **103**, 9107–
1151 9112 (2006).
1152 46. Lieberman, T. D. *et al.* Parallel bacterial evolution within multiple patients identifies candidate
1153 pathogenicity genes. *Nat. Genet.* **43**, 1275–1280 (2011).
1154 47. Zhao, S. *et al.* Adaptive Evolution within Gut Microbiomes of Healthy People. *Cell Host Microbe*
1155 **25**, 656-667.e8 (2019).
1156 48. Wong, D. P. G. H. & Good, B. H. Quantifying the adaptive landscape of commensal gut bacteria
1157 using high-resolution lineage tracking. *Nat. Commun.* **15**, 1605 (2024).
1158 49. Gómez-Santos, N., Glatter, T., Koebnik, R., Świątek-Połatyńska, M. A. & Søgaard-Andersen, L.
1159 A TonB-dependent transporter is required for secretion of protease PopC across the bacterial outer
1160 membrane. *Nat. Commun.* **10**, 1–18 (2019).
1161 50. Moreno-Gámez, S. *et al.* Niche partitioning by resource size in the gut microbiome.
1162 2025.11.13.688124 Preprint at https://doi.org/10.1101/2025.11.13.688124 (2025).
1163 51. Mills, R. H. *et al.* Multi-omics analyses of the ulcerative colitis gut microbiome link Bacteroides
1164 vulgatus proteases with disease severity. *Nat. Microbiol.* **7**, 262–276 (2022).
1165 52. Chen, D. W. & Garud, N. R. Rapid evolution and strain turnover in the infant gut microbiome.
1166 *Genome Res.* **32**, 1124–1136 (2022).

1167 53. Cremer, J., Arnoldini, M. & Hwa, T. Effect of water flow and chemical environment on
1168 microbiota growth and composition in the human colon. *Proc. Natl. Acad. Sci.* **114**, 6438–6443 (2017).
1169 54. Shalon, D. *et al.* Profiling the human intestinal environment under physiological conditions.
1170 *Nature* **617**, 581–591 (2023).
1171 55. Pédron, T. *et al.* A Crypt-Specific Core Microbiota Resides in the Mouse Colon. *mBio* **3**,
1172 10.1128/mbio.00116-12 (2012).
1173 56. Li, H. *et al.* The outer mucus layer hosts a distinct intestinal microbial niche. *Nat. Commun.* **6**,
1174 8292 (2015).
1175 57. Yasuda, K. *et al.* Biogeography of the Intestinal Mucosal and Lumenal Microbiome in the Rhesus
1176 Macaque. *Cell Host Microbe* **17**, 385–391 (2015).
1177 58. Desai, M. S. *et al.* A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier
1178 and Enhances Pathogen Susceptibility. *Cell* **167**, 1339-1353.e21 (2016).
1179 59. Chassaing, B. & Gewirtz, A. T. Identification of Inner Mucus-Associated Bacteria by Laser
1180 Capture Microdissection. *Cell. Mol. Gastroenterol. Hepatol.* **7**, 157–160 (2019).
1181 60. Mas-Lloret, J. *et al.* Gut microbiome diversity detected by high-coverage 16S and shotgun
1182 sequencing of paired stool and colon sample. *Sci. Data* **7**, 92 (2020).
1183 61. Duncan, K., Carey-Ewend, K. & Vaishnava, S. Spatial analysis of gut microbiome reveals a
1184 distinct ecological niche associated with the mucus layer. *Gut Microbes* **13**, 1874815 (2021).
1185 62. Mondragón-Palomino, O. *et al.* Three-dimensional imaging for the quantification of spatial
1186 patterns in microbiota of the intestinal mucosa. *Proc. Natl. Acad. Sci.* **119**, e2118483119 (2022).
1187 63. Guo, Y. *et al.* Oral pathobiont Klebsiella chaperon usher pili provide site-specific adaptation for
1188 the inflamed gut mucosa. *Gut Microbes* **16**, 2333463 (2024).
1189 64. Lenski, R. E. Experimental evolution and the dynamics of adaptation and genome evolution in
1190 microbial populations. *ISME J.* **11**, 2181–2194 (2017).
1191 65. Goyal, A., Bittleston, L. S., Leventhal, G. E., Lu, L. & Cordero, O. X. Interactions between
1192 strains govern the eco-evolutionary dynamics of microbial communities. *eLife* **11**, e74987 (2022).
1193 66. Feng, L. *et al.* Identifying determinants of bacterial fitness in a model of human gut microbial
1194 succession. *Proc. Natl. Acad. Sci.* **117**, 2622–2633 (2020).
1195 67. Conwill, A. *et al.* Anatomy promotes neutral coexistence of strains in the human skin
1196 microbiome. *Cell Host Microbe* **30**, 171-182.e7 (2022).
1197 68. Lee, S. M. *et al.* Bacterial colonization factors control specificity and stability of the gut
1198 microbiota. *Nature* **501**, 426–429 (2013).
1199 69. Jin, X. *et al.* Culturing of a complex gut microbial community in mucin-hydrogel carriers reveals
1200 strain- and gene-associated spatial organization. *Nat. Commun.* **14**, 3510 (2023).
1201 70. Gutiérrez-García, K. *et al.* A conserved genetic basis for commensal-host specificity through live
1202 imaging of colonization dynamics. 2024.04.19.590229 Preprint at
1203 https://doi.org/10.1101/2024.04.19.590229 (2024).
1204 71. Tang, Q. *et al.* Current Sampling Methods for Gut Microbiota: A Call for More Precise Devices.
1205 *Front. Cell. Infect. Microbiol.* **10**, (2020).
1206 72. Lloyd-Price, J. *et al.* Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases.
1207 *Nature* **569**, 655–662 (2019).
1208 73. Nishihara, Y. *et al.* Mucosa-associated gut microbiota reflects clinical course of ulcerative colitis.
1209 *Sci. Rep.* **11**, 13743 (2021).

1210 74. Atreya, R. & Siegmund, B. Location is important: differentiation between ileal and colonic
1211 Crohn's disease. *Nat. Rev. Gastroenterol. Hepatol.* **18**, 544–558 (2021).
1212 75. Kordahi, M. C. *et al.* Genomic and functional characterization of a mucosal symbiont involved in
1213 early-stage colorectal cancer. *Cell Host Microbe* **29**, 1589-1598.e6 (2021).
1214 76. Wu, D., Jospin, G. & Eisen, J. A. Systematic identification of gene families for use as 'markers'
1215 for phylogenetic and phylogeny-driven ecological studies of bacteria and archaea and their major
1216 subgroups. *PloS One* **8**, e77033 (2013).
1217 77. Beresford-Jones, B. S. *et al.* The Mouse Gastrointestinal Bacteria Catalogue enables translation
1218 between the mouse and human gut microbiotas via functional mapping. *Cell Host Microbe* **30**, 124-
1219 138.e8 (2022).
1220 78. Smith, B. J. *et al.* Accurate estimation of intraspecific microbial gene content variation in
1221 metagenomic data with MIDAS v3 and StrainPGC. 2024.04.10.588779 Preprint at
1222 https://doi.org/10.1101/2024.04.10.588779 (2024).
1223 79. Zhao, C., Dimitrov, B., Goldman, M., Nayfach, S. & Pollard, K. S. MIDAS2: Metagenomic
1224 Intra-species Diversity Analysis System. *Bioinformatics* **39**, btac713 (2023).
1225 80. Wattam, A. R. *et al.* Improvements to PATRIC, the all-bacterial Bioinformatics Database and
1226 Analysis Resource Center. *Nucleic Acids Res.* **45**, D535–D542 (2017).
1227 81. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069
1228 (2014).
1229 82. Garud, N. R., Good, B. H., Hallatschek, O. & Pollard, K. S. Evolutionary dynamics of bacteria in
1230 the gut microbiome within and across hosts. *PLOS Biol.* **17**, e3000102 (2019).
1231 83. OKSANEN, J. Vegan : community ecology package. *Httpveganr-Forg.-Proj.*
1232 https://cir.nii.ac.jp/crid/1570291225091856896 (2010).
1233 84. Schloissnig, S. *et al.* Genomic variation landscape of the human gut microbiome. *Nature* **493**, 45–
1234 50 (2013).
1235 85. Begun, D. J. *et al.* Population Genomics: Whole-Genome Analysis of Polymorphism and
1236 Divergence in Drosophila simulans. *PLOS Biol.* **5**, e310 (2007).
1237 86. Truong, D. T., Tett, A., Pasolli, E., Huttenhower, C. & Segata, N. Microbial strain-level
1238 population structure and genetic diversity from metagenomes. *Genome Res.* **27**, 626–638 (2017).
1239 87. Van Rossum, T., Ferretti, P., Maistrenko, O. M. & Bork, P. Diversity within species: interpreting
1240 strains in microbiomes. *Nat. Rev. Microbiol.* **18**, 491–506 (2020).
1241 88. Quince, C. *et al.* STRONG: metagenomics strain resolution on assembly graphs. *Genome Biol.*
1242 **22**, 214 (2021).
1243 89. Kang, J. T. L. *et al.* Long-term ecological and evolutionary dynamics in the gut microbiomes of
1244 carbapenemase-producing Enterobacteriaceae colonized subjects. *Nat. Microbiol.* **7**, 1516–1524 (2022).
1245 90. Zheng, W. *et al.* High-throughput, single-microbe genomics with strain resolution, applied to a
1246 human gut microbiome. *Science* **376**, eabm1483 (2022).
1247 91. Politis, D. N. & Romano, J. P. The Stationary Bootstrap. *J. Am. Stat. Assoc.* **89**, 1303–1313
1248 (1994).
1249 92. Tools for compositional data with a total - Vera Pawlowsky-Glahn, Juan José Egozcue, David
1250 Lovell, 2015. https://journals.sagepub.com/doi/abs/10.1177/1471082X14535526.
1251 93. Greenacre, M. *Compositional Data Analysis in Practice*. (Chapman and Hall/CRC, New York,
1252 2018). doi:10.1201/9780429455537.

1253 94. Human Microbiome Project Consortium. A framework for human microbiome research. *Nature*
1254 **486**, 215–221 (2012).
1255 95. Lloyd-Price, J. *et al.* Strains, functions and dynamics in the expanded Human Microbiome
1256 Project. *Nature* **550**, 61–66 (2017).
1257 96. Xie, H. *et al.* Shotgun Metagenomics of 250 Adult Twins Reveals Genetic and Environmental
1258 Impacts on the Gut Microbiome. *Cell Syst.* **3**, 572-584.e3 (2016).
1259 97. Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature*
1260 **490**, 55–60 (2012).
1261 98. Korpela, K. *et al.* Selective maternal seeding and environment shape the human gut microbiome.
1262 *Genome Res.* **28**, 561–568 (2018).
1263