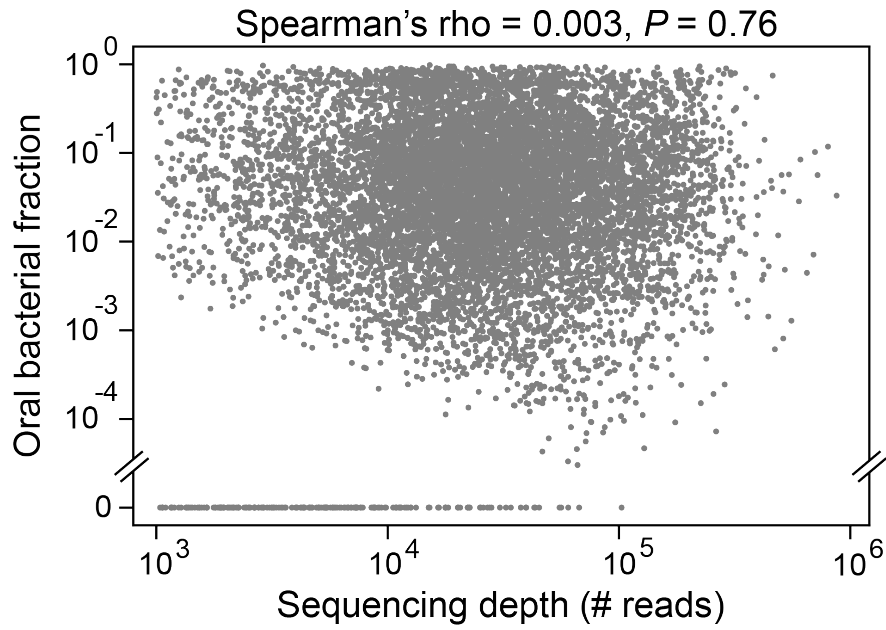
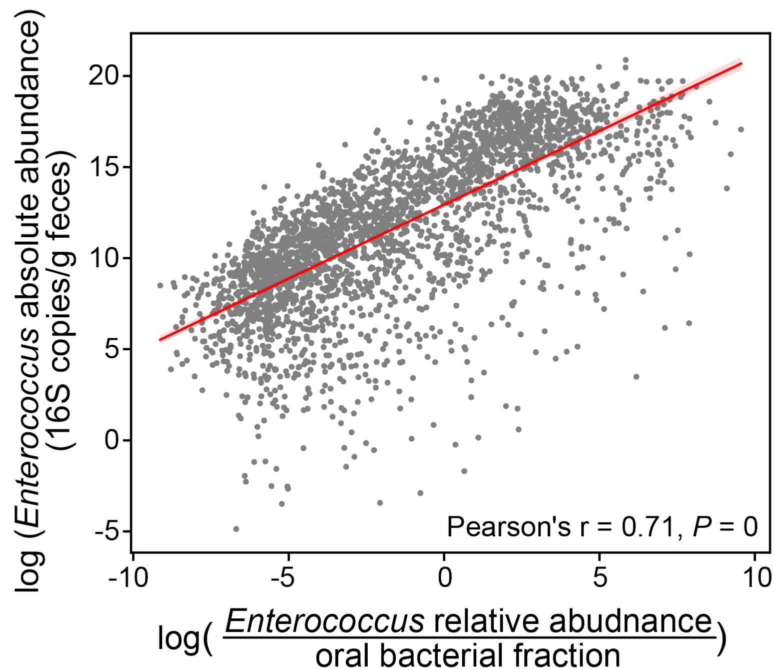


Figure S1: No association between oral bacterial fraction and sequencing depth across fecal samples of MSKCC allo-HCT recipients. Each dot represents a fecal sample (n = 10,433). The statistical test used was Spearman's correlation.

13



14 **Figure S2: Estimating *Enterococcus* absolute abundance through the *Enterococcus*-to-oral
15 bacteria relative abundance ratio.** Each data point represents a fecal sample from MSKCC
16 allo-HCT recipients. We excluded samples with zero relative abundance of *Enterococcus* or zero
17 oral bacterial fraction from both the plot and the Pearson's correlation analysis, resulting in a
18 subset of 2,765 samples. The red line represents the best linear fit, with the shading in the same
19 color indicating its ±95% confidence interval. The absolute abundance of *Enterococcus* on the *y*-
20 axis was calculated by multiplying its relative abundance by the total bacterial load. The
21 statistical test used was Pearson's correlation.

22 **Table S1: Oral bacterial ASV, taxonomy, and total abundance in mouse feces. a**, Sequence
23 and taxonomy of all 53 oral bacterial ASVs found in mouse feces, with specific oral ASVs for
24 each mouse indicated in the table. **b**, Total relative and absolute abundances of oral and gut
25 bacteria in fecal samples. We determined the oral or gut bacterial load (absolute abundance) by
26 multiplying the total bacterial load by the respective oral or gut bacterial fraction (relative
27 abundance). Bacterial load units for fecal and oral samples are 16S copies per gram of feces and
28 per swab, respectively. Samples with fewer than 1,000 reads were excluded from analysis and
29 are labelled in the "Exclusion" column. DSS: Dextran Sulfate Sodium. **c**, FEAST-estimated oral
30 bacterial fractions in post-treatment fecal samples. FEAST was run ten times for each fecal
31 sample.
32
33 See separate Excel table

34 **Table S2: Pairwise Adonis test of compositional similarity between pre- and post-treatment**
35 **fecal and oral microbiome samples.** Adonis test is a two-sided permutation-based multivariate
36 analysis that assesses the differences in community composition between groups of microbiome
37 samples. In our analysis, microbiome samples were divided into four groups: pre-treatment oral
38 samples (group label: PreOral), pre-treatment fecal samples (group label: PreFecal), post-
39 treatment oral samples (group label: PostOral), and post-treatment fecal samples (group label:
40 PostFecal). The table below displays the R^2 values and FDR-corrected *P* values for all six pairs
41 of comparisons and for each mouse group. Bray-Curtis distance metric was used to measure the
42 similarity between microbiome samples. Notably, post-antibiotic-treatment fecal samples
43 exhibited greater similarity to pre-treatment oral samples ($R^2 = 0.15$, $P = 0.055$) than to pre-
44 treatment fecal samples ($R^2 = 0.20$, $P = 0.032$).
45

Mouse group	**Sample group 1**	**Sample group 2**	**R^2**	***P* value**
No treatment	PreOral (n = 4)	PostOral (n = 9)	0.21	0.058
	PostOral (n = 9)	PostFecal (n = 10)	0.76	0.001
	PreFecal (n = 4)	PostOral (n = 9)	0.71	0.003
	PreOral (n = 4)	PostFecal (n = 10)	0.92	0.003
	PreOral (n = 4)	PreFecal (n = 4)	0.95	0.044
	PreFecal (n = 4)	PostFecal (n = 10)	0.09	0.305
Antibiotic treatment	PreOral (n = 8)	PostOral (n = 8)	0.14	0.066
	PostOral (n = 8)	PostFecal (n = 12)	0.12	0.066
	PreFecal (n = 7)	PostOral (n = 8)	0.42	0.002
	PreOral (n = 8)	PostFecal (n = 12)	0.15	0.055
	PreOral (n = 8)	PreFecal (n = 7)	0.33	0.004
	PreFecal (n = 7)	PostFecal (n = 12)	0.20	0.032
Dextran sulfate sodium treatment	PreOral (n = 3)	PostOral (n = 5)	0.41	0.086
	PostOral (n = 5)	PostFecal (n = 9)	0.10	0.273
	PreFecal (n = 5)	PostOral (n = 5)	0.26	0.055
	PreOral (n = 3)	PostFecal (n = 9)	0.56	0.029
	PreOral (n = 3)	PreFecal (n = 5)	0.86	0.033

	PreFecal (n = 5)	PostFecal (n = 9)	0.32	0.029

46

47 **Table S3: Sequence and taxonomy of 178 oral bacterial ASVs identified from healthy**
48 **human individuals involved in the HMP dataset.**
49
50 See separate Excel table

Table S4: Oral bacterial ASV, taxonomy, and total abundance in human feces of MSKCC allo-HCT recipients. a, Sequence and taxonomy of 127 oral bacterial ASVs. **b**, Total relative and absolute abundances of oral and gut bacteria in fecal samples. We determined the oral or gut bacterial load (absolute abundance) by multiplying the total bacterial load by the respective oral or gut bacterial fraction (relative abundance). Bacterial load unit is 16S copies per gram of feces.

See separate Excel table

58 **Table S5: Antibiotics associated with intestinal domination by oral bacterial ASVs in**
59 **MSKCC allo-HCT recipients.** A total of 291 patients with at least 10 samples between day -10
60 and 40 relative to transplantation were included. The hazard ratios quantify the relative risk of
61 intestinal domination by any single oral ASV that exceeds 30% in relative abundance compared
62 to no domination. Oral and intravenous vancomycin were separated because vancomycin does
63 not reach adequate levels in human gut when given intravenously. Distinct routes of
64 administration were combined for all other antibiotics. *P* values were corrected for multiple
65 comparisons using FDR. Significant associations ($P < 0.05$) are highlighted in red. CI:
66 confidence interval. The statistical test used for data analysis was Cox proportional hazard with
67 time-dependent covariates.
68

Antibiotic	**Hazard ratio (95% CI)**	***P* value**
macrolide derivatives	4.5e-8 (>0)	0.996
metronidazole	0.26 (0.04-1.93)	0.568
fluoroquinolones	0.45 (0.26-0.77)	0.022
sulfamethoxazole/trimethoprim	0.52 (0.06-4.27)	0.720
aztreonam	0.63 (0.15-2.68)	0.720
cephalosporins	0.93 (0.50-1.72)	0.880
intravenous vancomycin	0.94 (0.59-1.49)	0.880
carbapenems	1.36 (0.72-2.57)	0.580
linezolid	1.57 (0.69-3.58)	0.580
oral vancomycin	1.77 (1.18-2.64)	0.022
piperacillin/tazobactam	2.24 (1.37-3.65)	0.015
aminoglycosides	2.95 (0.39-22.18)	0.580

69

70 **Table S6. Regression slope between relative abundance of selected bacterial genera and**
71 **total bacterial load in feces of MSKCC allo-HCT recipients.** Bacterial genera that dominated
72 at least 100 fecal samples with a relative abundance exceeding 30% were selected for linear
73 regression analysis in log-log space. In the table, we presented the mean relative abundance of
74 each genus across HMP samples collected from the gastrointestinal tract, oral cavity, nasal
75 cavity, skin, and urogenital tract. The five genera with the highest mean relative abundance in the
76 gastrointestinal tract, compared to the other four body sites, are highlighted in red. Notably, they
77 all exhibited positive slopes.
78
79 See separate Excel table

80 **Table S7: High-quality genomes of *Streptococcus* species assembled from shotgun**
81 **metagenomics data.** Metagenome-assembled genomes (MAGs) were constructed from 22 fecal
82 samples from MSKCC allo-HCT recipients. The iRep value is an index of replication and serves
83 as an indicator of bacterial replication rate[1]. iRep values that did not meet all genome and
84 mapping quality requirements are marked as “n/a”. Assembled MAGs containing *Streptococcus*
85 ASV_8 are highlighted in red.
86
87 See separate Excel table

88 **Table S8: Oral bacterial fraction in feces is associated with survival of MSKCC allo-HCT**
89 **recipients**. All-cause mortality and GVHD-related mortality were analyzed by a Cox
90 proportional hazard model and a Fine-Gray subdistribution hazard model, respectively. Both
91 models were adjusted for *Enterococcus* absolute abundance, age, underlying disease, graft
92 source, and intensity of conditioning regime. *P* values were FDR-corrected for multiple
93 comparisons. Abbreviations: CI (confidence interval), AML (Acute Myeloid Leukemia), ALL
94 (Acute Lymphocytic Leukemia), MDS (Myelodysplastic syndrome), MPN (Myeloproliferative
95 neoplasm). The statistical test used for data analysis was Cox proportional hazard with time-
96 dependent covariates.
97

Endpoint event	Covariate	Hazard ratio	95% CI	*P* value
All-cause mortality	Oral bacterial fraction	4.02	2.11-7.68	6.3e-5
	log(*Enterococcus* absolute abundance)	1.19	1.13-1.25	3.5e-11
	Age	1.03	1.02-1.04	1.0e-12
	Disease			
	AML/ALL/MDS/MPN	Reference category		
	Others	1.30	1.02-1.65	0.053
	Graft source			
	T-cell depletion	Reference category		
	Unmodified	1.21	0.92-1.58	0.193
	Cord	1.33	0.89-1.99	0.193
	Conditioning intensity			
	Ablative	Reference category		
	Non-ablative	0.46	0.30-0.72	0.001
	Reduced intensity	0.90	0.67-1.21	0.491
GVHD-related mortality	Oral bacterial fraction	4.23	1.69-10.6	0.006
	log(*Enterococcus* absolute abundance)	1.22	1.13-1.32	1.5e-6
	Age	1.03	1.01-1.04	2.3e-4

	Disease			
	AML/ALL/MDS/MPN	Reference category		
	Others	1.31	0.90-1.91	0.319
	Graft source			
	T-cell depletion	Reference category		
	Unmodified	1.07	0.67-1.70	0.777
	Cord	1.43	0.76-2.70	0.420
	Conditioning intensity			
	Ablative	Reference category		
	Non-ablative	0.90	0.47-1.72	0.777
	Reduced intensity	1.14	0.71-1.84	0.777

98

99 **Table S9: Microbiome datasets used in this study.**

100

101 See separate Excel table

205 **Supplementary Note 1: Discussion on the discrepancy between our approach and FEAST**
206 **for estimating oral bacterial fraction**
207 As shown in Extended Data Fig. 2, our approach and FEAST show quantitative
208 agreement in estimating oral bacterial fractions when the fraction, as estimated by our approach,
209 exceeds 0.0001. However, for fractions smaller than 0.0001, our approach yields much smaller
210 estimated values than FEAST. Specifically, in five samples (F_Day_7_DSS_E,
211 F_Day_3_Abx_D, F_Day_3_Control_E, F_Day_7_Control_E, F_Day_3_DSS_E), substantial
212 differences arise between FEAST and our approach, with the ratio of FEAST to our approach
213 exceeding 10 (i.e., different magnitudes).
214 This discrepancy can be largely attributed to two reasons. In the case of samples
215 F_Day_7_DSS_E and F_Day_3_Abx_D, where our approach estimated a zero oral bacterial
216 fraction, the values estimated by FEAST show considerable variability across different runs. In
217 the case of sample F_Day_7_DSS_E, FEAST provided a range between 4.4e-31 and 1.2e-3 for
218 the minimum and maximum oral bacterial fractions in different runs. Similarly, for sample
219 F_Day_3_Abx_D, FEAST estimates span from 2.5e-255 to 6.7e-3. Even for samples where the
220 average FEAST estimation agrees with the oral bacterial fraction estimated by our approach,
221 notable variations exist from run to run. For example, in the case of sample
222 F_Day_7_Control_C, the FEAST-estimated oral bacterial fractions vary between 7.6e-19 and
223 7.7e-4. Another example, F_Day_3_Control_A, exhibits variations ranging from 8.5e-7 to 9.1e-
224 4. Overall, the large variability underscores the inherent difficulty in accurately quantifying oral
225 bacterial transmission using FEAST when such transmission events are infrequent. For a detailed
226 table of all FEAST estimates in each individual run, refer to Table S1c.
227 Moreover, FEAST calculates the fraction of oral bacteria in fecal samples on an
228 individual mouse basis. In contrast, our approach defines oral ASVs using predefined thresholds
229 for mean relative abundance and prevalence across both oral and fecal samples collected from all
230 mice (refer to Methods in the main text, Section: Identification of oral ASVs in fecal samples).
231 Consequently, oral sequences identified by our approach must exhibit a sufficiently high relative
232 abundance in the oral cavity and a correspondingly low relative abundance in the gut across the
233 majority of mice. This stringent criterion ensures a more conservative approach, enhancing
234 confidence that oral ASVs detected in fecal samples genuinely originate from bacteria colonizing
235 the oral cavity rather than the gut. By incorporating both oral and fecal samples from all mice,
236 our methodology aims for a robust quantification of transmitted oral bacteria in the gut. Due to
237 this conservative nature, it is understandable that our approach consistently generated lower
238 estimated oral bacterial fractions for the five samples compared to those estimated by FEAST.
239 For the two reasons mentioned above, we believe it may not be suitable to compare our
240 approach against FEAST when fecal samples contain extremely low fraction (<0.0001 in our
241 method) of oral bacteria.

242 **Supplementary Note 2: Theoretical relationship between oral bacterial fraction and total**
243 **bacterial load in fecal samples**
244 By definition, the relative and absolute abundance of oral and gut bacteria in a fecal
245 sample are related through the following equation:

246 $$F_{total} = \frac{F_{oral}}{f_{oral}} = \frac{F_{gut}}{1 - f_{oral}} \quad \text{Eq. S1}$$

247 Here, f_{oral} represents the relative abundance of oral bacteria, F_{oral} and F_{gut} represent the
248 absolute abundance of oral and gut bacteria respectively, and $F_{total}(= F_{oral} + F_{gut})$ is the total
249 bacterial load.
250
251 Pure *Marker* hypothesis: When an increase in f_{oral} is solely driven by gut bacterial depletion,
252 F_{oral} remains constant (let the constant be K_1). In this scenario, Eq. S1 can be rewritten as

253 $$F_{total} = \frac{K_1}{f_{oral}} \quad \text{Eq. S2}$$

254 or on the log-log scale (base b)

255 $$log_b F_{total} = log_b K_1 - log_b f_{oral} \quad \text{Eq. S3}$$

256 Eq. S3 indicates that the derivative of log-transformed total bacterial load with respect to log-
257 transformed oral bacterial fraction is -1:

258 $$\frac{d\, log_b F_{total}}{d\, log_b f_{oral}} = -1 \quad (pure\ Marker\ hypothesis) \quad \text{Eq. S4}$$

259
260 Pure *Expansion* hypothesis: When an increase in f_{oral} is solely driven by absolute expansion of
261 oral bacterial population, F_{gut} remains constant (let the constant be K_2). In this scenario, Eq. S1
262 can be rewritten as

263 $$F_{total} = \frac{K_2}{1 - f_{oral}} \quad \text{Eq. S5}$$

264 or on the log-log scale (base b)

265 $$log_b F_{total} = log_b K_2 - log_b(1 - f_{oral}) \quad \text{Eq. S6}$$

266 From Eq. S6, we found that the derivative of log-transformed total bacterial load with respect to
267 log-transformed oral bacterial fraction is positive:

268 $$\frac{d\, log_b F_{total}}{d\, log_b f_{oral}} = \frac{f_{oral}}{1 - f_{oral}} > 0 \quad (pure\ Expansion\ hypothesis) \quad \text{Eq. S7}$$

269

270 **Supplementary Note 3: Validation of oral bacterial ASVs identified from healthy**
271 **individuals in patients with inflammatory bowel disease**
272 Since human microbiota composition is body site-specific, we hypothesized that the
273 reference set of 178 oral ASVs identified from a large cohort of healthy individuals recruited by
274 the Human Microbiome Project (HMP) can be applied to other healthy individuals and patients
275 for the detection of oral bacteria in their gut. To test this hypothesis, we analyzed a publicly
276 available dataset from patients with inflammatory bowel disease (IBD) and their healthy
277 controls[2]. This dataset includes paired fecal and saliva samples from a total of 43 healthy
278 controls (HC), 16 patients with Crohn's disease (CD), and 42 patients with ulcerative colitis
279 (UC).
280 By varying three cutoff parameters used to identify oral ASVs from HMP, we showed
281 that the estimated oral fraction in the fecal samples of the IBD cohort participants is largely
282 robust against variations in these parameter values (Extended Data Fig. 4a). Among the three
283 cutoffs, the cutoff for mean relative abundance (θ_a, see Methods in the main text for details) has
284 the strongest impact. Across all healthy individuals, the mean oral bacterial fraction in feces was
285 found to be 1.2%, which closely aligns with a previously reported estimate of 2%[3]. This fraction
286 exhibited a nearly three-fold increase to 4.2% and 4.3% in CD and UC patients respectively,
287 supporting the notion that IBD is associated with an enrichment of oral bacteria in the gut[4]. Most
288 importantly, more than 84% (mean values: 92.2% for HC, 89.8% for CD, and 84.1% for UC) of
289 the oral ASVs detected in the fecal samples from the reference set were also found in their
290 corresponding saliva samples (Extended Data Fig. 4b). In addition, these oral ASVs accounted
291 for over 87% (mean values: 93.4% for HC, 94.6% for CD, and 87.8% for UC) of the estimated
292 total relative abundance of oral bacteria in the feces (Extended Data Fig. 4c). In summary, we
293 validated that the reference set of oral ASVs identified from the HMP dataset can be used to infer
294 oral ASVs in the fecal samples from other non-HMP healthy individuals and patients even in the
295 absence of paired oral samples.

296 **Supplementary Note 4: Discussion on associations between biofilm-forming capacity of**
297 ***Streptococcus*, *Actinomyces* and *Abiotrophia* and their fecal relative abundance in MSKCC**
298 **allo-HCT recipients**

299 All three bacterial genera have the ability to form biofilm in the oral cavity[5,6]. However, it
300 remains uncertain if they can form biofilms in the lower gastrointestinal tract. *Streptococcus*
301 *thermophilus*, the most dominant *Streptococcus* species in the MSKCC allo-HCT cohort (see
302 Table S7), is a poor biofilm producer due to its inability to firmly attach to surfaces[7].
303 Consistently, exopolysaccharides derived from *S. thermophilus* do not interact with mucin[8],
304 indicating that this species may be unable to form biofilm in colonic mucus.

305 Furthermore, we used PICRUSt2[9] to predict KEGG[10] pathway abundances from 16S
306 amplicon sequencing data (10,433 samples). The biofilm-forming potential of intestinal
307 microbiota was quantified by the sum of relative abundance of three KEGG pathways containing
308 the keyword "biofilm" (ko05111, ko02025, ko02026). We observed a positive, but statistically
309 insignificant, Spearman correlation between this biofilm index and the relative abundance of oral
310 *Streptococcus* (Spearman's rho = 0.004, P = 0.661). For oral *Actinomyces* and *Abiotrophia*, the
311 correlations were statistically significant but negative (Spearman's rho = -0.075 and -0.056, P =
312 9.1e-14 and 2.7e-8, respectively). Using a subset of 3,108 samples with 16S qPCR data, we
313 found that the biofilm-forming capacity is positively associated with the total bacterial load
314 (Spearman's rho = 0.073, P = 6.3e-5). Here, all P values have been adjusted for multiple
315 comparisons.

316 Taken together, although we cannot entirely rule out the possibility, both literature studies
317 and our findings do not provide evidence for a link between intestinal dominations by the three
318 oral genera and their biofilm-forming capacity.

319 **Supplementary Note 5: Discussion on associations between antibiotic exposure and oral**
320 **bacterial domination in the fecal samples from MSKCC allo-HCT recipients**

321 During allo-HCT, most patients received multiple antibiotics for prophylactic and
322 treatment purposes for different periods. Due to different spectra of antibiotics, their
323 administration led to diverse dynamics of oral bacteria translocated to the intestine, as observed
324 in fecal samples by 16S rRNA amplicon sequencing (Fig. 4a). To isolate the individual effect of
325 each antibiotic, we used a time-dependent Cox proportional hazards model (see Methods in the
326 main text for details). The Cox model revealed that piperacillin-tazobactam, a combination of β-
327 lactam and β-lactamase inhibitor with a broad spectrum of antibacterial activity, had the most
328 significant positive influence on intestinal domination by any single oral ASV with a relative
329 abundance that exceeds 30% (Table S5; Hazard ratio = 2.24; 95% confidence interval, 1.37-3.65;
330 $P = 0.015$). This association aligns with a previous finding showing that piperacillin-tazobactam
331 caused the most pronounced depletion of anaerobic gut commensals in the same patient cohort[11].

332 To validate the anaerobe-depleting effect of piperacillin-tazobactam, we reanalyzed a
333 previous study[12] that investigated the dynamics of gut microbiota during and after total gut
334 decontamination (using oral piperacillin-tazobactam) and selective gut decontamination (using
335 oral polymyxin-neomycin) in children undergoing allo-HCT at the Leiden University Medical
336 Center in the Netherlands. Indeed, the children who received piperacillin-tazobactam exhibited
337 significantly higher total fraction of oral bacteria in their feces compared to those who received
338 oral polymyxin-neomycin across all transplantation stages (Extended Data Fig. 5).

339 Our analysis also revealed a significant negative association between fluoroquinolone
340 antibiotics with oral bacterial domination (Table S5; Hazard ratio = 0.45; 95% confidence
341 interval, 0.26-0.77; $P = 0.022$). Fluoroquinolones are commonly used as prophylactic agents to
342 reduce the incidence of gram-negative bacterial infections, including *Pseudomonas* and
343 *Enterobacteriaceae*, in patients with neutropenia[13]. One possible explanation for this negative
344 association is that fluoroquinolones are generally more effective against aerobic and facultative
345 anaerobic bacteria, such as those found in the oral flora, than against anaerobic bacteria[14].
346 Another potential explanation is that all allo-HCT recipients were initially prescribed
347 fluoroquinolone prophylaxis but may be switched to other agents, such as piperacillin-
348 tazobactam, when fevers or bloodstream infections occur. As a result, fecal samples exposed only
349 to fluoroquinolones may be enriched for samples collected before acute febrile episodes or from
350 patients with more benign treatment courses who never got a fever and thus remained on
351 fluoroquinolone prophylaxis.

352 **Supplementary Note 6: Slow growth of *Streptococcus* ASV_8 in fecal samples from**
353 **MSKCC allo-HCT recipients**
354 Among all identified oral ASVs*, Streptococcus* ASV_8 had the highest mean relative
355 abundance among all fecal samples of the MSKCC allo-HCT recipients. We previously
356 published shotgun metagenomics data[15] for 395 fecal samples from these patients. Among the
357 395 samples, 19 contains at least 10% *Streptococcus* ASV_8, as indicated by the paired 16S
358 rRNA sequencing. Using a published bioinformatic pipeline (see Methods in the main text for
359 details), we obtained 22 high-quality metagenome-assembled genomes (MAGs) of *Streptococcus*
360 *spp.* from these 19 metagenomic samples. Among the 22 MAGs, four contains ASV_8, and all
361 were annotated as *S. thermophilus*. We then calculated the replicate rate of these MAGs through
362 iRep[1]. The iRep index can accurately estimate the ratio between the coverage at the origin and
363 terminus of replication, which is proportional to replication rate. We found an averaged iRep
364 index of 1.35 (Table S7), indicating that, on average, only 35% cells are replicating. This iRep-
365 based estimation suggests that bacteria containing *Streptococcus* ASV_8 grew slowly in the
366 intestine of MSKCC allo-HCT recipients.

367 1. Brown, C.T., Olm, M.R., Thomas, B.C. & Banfield, J.F. Measurement of bacterial
368 replication rates in microbial communities. *Nat. Biotechnol.* **34**, 1256-1263 (2016).

369 2. Imai, J., *et al.* A potential pathogenic association between periodontal disease and
370 Crohn's disease. *JCI Insight* **6**, e148543 (2021).

371 3. Schmidt, T.S., *et al.* Extensive transmission of microbes along the gastrointestinal tract.
372 *Elife* **8**, e42693 (2019).

373 4. Read, E., Curtis, M.A. & Neves, J.F. The role of oral bacteria in inflammatory bowel
374 disease. *Nat. Rev. Gastroenterol. Hepatol.* **18**, 731-742 (2021).

375 5. Larsen, T. & Fiehn, N.E. Dental biofilm infections–an update. *APMIS* **125**, 376-384
376 (2017).

377 6. Abdullah, N., *et al.* Intraoral appliances for in situ oral biofilm growth: a systematic
378 review. *J. Oral Microbiol.* **11**, 1647757 (2019).

379 7. Couvigny, B., *et al.* *Streptococcus thermophilus* biofilm formation: a remnant trait of
380 ancestral commensal life? *PLoS One* **10**, e0128099 (2015).

381 8. Arias, S.L. & Brito, I.L. Biophysical determinants of biofilm formation in the gut. *Curr.*
382 *Opin. Biomed. Eng.* **18**, 100275 (2021).

383 9. Douglas, G.M., *et al.* PICRUSt2 for prediction of metagenome functions. *Nat.*
384 *Biotechnol.* **38**, 685-688 (2020).

385 10. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic*
386 *Acids Res.* **28**, 27-30 (2000).

387 11. Morjaria, S., *et al.* Antibiotic-induced shifts in fecal microbiota density and composition
388 during hematopoietic stem cell transplantation. *Infect. Immun.* **87**, e00206-00219 (2019).

389 12. Bekker, V., *et al.* Dynamics of the gut microbiota in children receiving selective or total
390 gut decontamination treatment during hematopoietic stem cell transplantation. *Biol.*
391 *Blood Marrow Transplant.* **25**, 1164-1171 (2019).

392 13. Cometta, A., Calandra, T., Bille, J. & Glauser, M.P. *Escherichia coli* resistant to
393 fluoroquinolones in patients with cancer and neutropenia. *N. Engl. J. Med.* **330**, 1240-
394 1241 (1994).

395 14. Appelbaum, P.C. Quinolone activity against anaerobes. *Drugs* **58**, 60-64 (1999).

396 15. Yan, J., *et al.* A compilation of fecal microbiome shotgun metagenomics from
397 hematopoietic cell transplantation patients. *Sci. Data* **9**, 1-11 (2022).

398 16. Stein-Thoeringer, C., *et al.* Lactose drives *Enterococcus* expansion to promote graft-
399 versus-host disease. *Science* **366**, 1143-1149 (2019).