

Sugar-rich foods exacerbate antibiotic-induced microbiome disruption

<https://doi.org/10.1038/s41586-026-11077-3>

Received: 4 March 2025

Accepted: 25 August 2026

Published online: 30 September 2026

Open access

 Check for updates

Anqi Dai^{1,21}, Annamaria Ballweg^{1,21}, William Jogia^{2,3}, Madhumitha Rangesa¹, Peter A. Adintori¹, Mirae Baichoo¹, Tyler Funnell⁴, Teng Fei⁵, Nicholas R. Waters¹, Abrielle Swartz¹, Sukanya Sahu¹, Brianna Gipson¹, Sandeep S. Raj¹, Eiko Hayase⁶, Zoe Pierce¹, Kenton Wu¹, Natalie Smith⁴, Justin Neuberger¹, Adam Warren¹, Marina Burgos da Silva⁷, Oriana Miltiadous⁸, Corrado Zuanelli Brambilla⁹, Marissa Lubin Buchan^{1,10}, Tatanisha K. Peets¹⁰, Ana Gradissimo¹, Luigi A. Amoretti¹, Caichen Duan^{2,3,11}, Chenzhen Zhang^{2,3}, Fanny Matheis^{2,3}, Alexis P. Sullivan^{2,3}, John B. Slingerland⁴, Annelie G. Clurman¹, Daniel G. Brereton¹, Paul A. Giardina¹, Antonio L. C. Gomes⁴, Abigail J. Johnson^{12,13}, Dan Knights^{14,15}, Robert R. Jenq⁴, Kate A. Markey^{16,17}, Miguel-Angel Perales^{1,18}, Sergio A. Giralt^{1,18}, Marcel R. M. van den Brink^{20,22}✉, Jonas Schluter^{2,3,19,22}✉ & Jonathan U. Peled^{1,18,22}✉

Diet shapes the composition of the gut microbiota¹; however, the specific effects of distinct food groups on microbiome dynamics are unclear, particularly in circumstances of extreme perturbation. Here we evaluated the relationship between diet and intestinal microbiome dynamics by precisely tracking 9,419 meals consumed by 173 patients who were hospitalized for haematopoietic cell transplantation and analysing subsequent microbiome changes. Bayesian inference applied to data from 158 patients with paired longitudinal microbiome samples revealed that the intake of sweets and sugars during antibiotic exposure predicted exacerbated microbial dysbiosis, manifesting as lowered α -diversity and greater expansion of the pathobiont *Enterococcus*. Experiments in mice also showed that sucrose supplementation increased and prolonged antibiotic-induced *Enterococcus* expansion. These data suggest that avoiding a diet rich in simple sugars during antibiotic treatment may mitigate microbiota disruption. Further studies in independent cohorts will offer opportunities to generalize these findings and evaluate microbiota-sparing interventions.

The intestinal microbiota modulates host metabolism and immunity, and perturbations of microbiome composition are linked to various disease states². Although a precise definition of a homeostatic human microbiome composition remains unclear^{3,4}, a number of disease states share dysbiotic patterns in microbial community compositions, often characterized by loss of diversity and expansion of facultative anaerobes⁵. Understanding the factors that influence microbiome disruption and dysbiosis is key to deciphering the interactions between host and commensal organisms and to designing therapeutic strategies that target the microbiome.

Host and environmental factors shape intestinal microbiome compositions; in particular, a major effect of diet has been observed both

in mice and humans^{1,6–10}. However, most of these human analyses were either in small dietary-intervention trials of individuals with chronic conditions or of healthy volunteers with unperturbed, stable microbiomes. Severe perturbations, such as those that occur during acute illness or intensive medical treatment are less well understood^{4,11}. Controlled studies in mice and humans have indicated that the recovery of mouse microbiome composition after antibiotics can be impaired or facilitated by withholding or augmenting dietary fibre, respectively^{10,12,13}. Moreover, while some studies suggest that dietary perturbations exert effects on microbial composition within hours^{14–17}, most human studies have correlated faecal microbiome compositions with long-term habitual diet^{7,18–20} or variations on months-long timescales^{21,22}

¹Adult Bone Marrow Transplantation Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ²Institute for Systems Genetics, NYU Langone Health Grossman School of Medicine, New York, NY, USA. ³Department of Microbiology, NYU Langone Health Grossman School of Medicine, New York, NY, USA. ⁴Department of Hematology and Hematopoietic Cell Transplantation, City of Hope National Medical Center, Los Angeles, CA, USA. ⁵Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ⁶Departments of Genomic Medicine and Stem Cell Transplantation Cellular Therapy, Division of Cancer Medicine, University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁷Human Oncology & Pathogenesis Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ⁸Department of Pediatrics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ⁹Hematology Unit, Department of Cellular Therapy, Hematology and Laboratory Medicine, Azienda Ospedaliero-Universitaria Senese, Siena, Italy. ¹⁰Food and Nutrition Services, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ¹¹Department of Biomedical Engineering, NYU Tandon School of Engineering, New York, NY, USA. ¹²Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA. ¹³Riddet Institute, Massey University, Palmerston North, New Zealand. ¹⁴Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, USA. ¹⁵Biotechnology Institute, University of Minnesota, Saint Paul, MN, USA. ¹⁶Translational Science and Therapeutics Division, Fred Hutchinson Cancer Center, Seattle, WA, USA. ¹⁷Division of Medicine, University of Washington, Seattle, WA, USA. ¹⁸Weill Cornell Medical College, New York, NY, USA. ¹⁹Laura and Isaac Perlmutter Cancer Center, NYU Langone Health, New York, NY, USA. ²⁰Present address: City of Hope, Duarte, CA, USA. ²¹These authors contributed equally: Anqi Dai, Annamaria Ballweg. ²²These authors jointly supervised this work: Marcel R.M. van den Brink, Jonas Schluter, Jonathan U. Peled. ✉e-mail: mvandenbrink@coh.org; jonas.schluter@nyulangone.org; peledj@mskcc.org

and relied on recall-based surveys, the imprecision and limitations of which have been well described^{23,24}. Taken together, while diet is assumed to be a major determinant of microbiome composition, an understanding of precise diet–microbiota interactions in humans under real-world disease and treatment conditions is lacking.

Patients with blood cancers who undergo allogeneic haematopoietic cell transplantation (allo-HCT) are typically hospitalized for several weeks, during which they receive chemotherapy and, frequently, antibiotics. During this time, they exhibit substantial changes in nutritional intake²⁵, as well as severe microbiome perturbation^{4,26}, including loss of α -diversity and expansion of facultative anaerobe organisms, most often *Enterococcus*²⁷. These shifts are associated with adverse clinical outcomes, including mortality^{26,27}. These microbial–ecological disruptions are driven largely by antibiotics^{4,28}, as well as by intestinal inflammation induced by chemotherapy, irradiation and graft-versus-host disease (GVHD)^{29,30}. Here we hypothesize that nutrition contributes to these microbiome dynamics, and we reasoned that microbiome perturbations offer an opportunity to reveal these relationships by tracking microbiome composition alongside precise dietary records.

The weeks during which patients undergoing allo-HCT are hospitalized facilitate frequent collection of both stool specimens and high-precision dietary intake data; considering these data in light of such therapy-induced perturbations enables causal-inference approaches^{31,32}. We therefore aimed to quantify associations between dietary intake and microbiome composition to reveal mechanistic insights into diet-inducible microbiome modulation. For this, we collected daily dietary intake data, paired with longitudinally collected faecal samples. We profiled faecal microbiome compositions and revealed associations between key microbiome features and dietary components. Intake of sweets during antibiotic exposure predicted lowered faecal microbiome α -diversity, as well as increased *Enterococcus* abundance, suggesting a model whereby simple dietary sugars favour the expansion of pathobionts that persisted through antibiotic treatment. We validated this prediction in a mouse model in which dietary sugar supplementation exacerbated and prolonged antibiotic-induced *Enterococcus* blooms. This study advances our understanding of microbiome dynamics outside of the steady state and identifies potential interventional approaches; for example, short-term sugar avoidance is a strategy that could be readily evaluated in HCT and across patient cohorts treated with various antibiotics.

Allo-HCT begins with an intensive course of chemotherapy conditioning that can cause painful mouth and throat lesions (mucositis), nausea and poor appetite (Fig. 1a). Consequently, patients often have a poor diet and lose weight during a period of increased caloric needs²⁵. These symptoms typically persist until the graft is established in the bone marrow (engraftment) and produces neutrophils—observed in this cohort after a median of 12 days (range, 8–37 days). The severe degree of nutritional perturbation is underscored by the observation that a subset of patients in the cohort had such a low dietary intake that they required nutrition support, either through intravenous infusion (total parenteral nutrition (TPN); $n = 23$ patients for a median of 11 days; range, 2–63 days) or a gastric tube (enteral nutrition, EN; $n = 5$ patients for a median of 76 days; range, 5–156 days).

High-frequency diet–microbiome data

We recorded the precise intake of food items consumed during 9,419 meals (Fig. 1b) consumed by 173 allo-HCT recipients with blood cancers (Supplementary Table 1; range, 8–128 meals per patient) and also collected faecal samples during this period (Fig. 1c). We mapped meal records to a hierarchical food nomenclature (Food and Nutrient Database for Dietary Studies, FNDDS) that classifies foods into nested categories, resulting in nine distinct food groups. A food taxonomy constructed from these categories (Fig. 1d) facilitated analysis of specific

food types at various hierarchical aggregation levels, as well as the application of metrics from ecology to summarize nutritional intake data⁸. Thus, the high-frequency, granular dietary intake data in this study takes the form of, for example, ‘1 day before the faecal sample was collected on day –5 relative to HCT, the patient consumed 142 g of baked, broiled or roasted chicken breast’, which is distinct from the resolution offered by typical food-frequency instruments (for example, ‘chicken two times per week’).

To visualize the high-dimensional dietary dataset, we applied the TaxUMAP algorithm³ such that each point represents a patient’s total food consumption on a single day; gradients across dietary TaxUMAP space indicated that patients consumed more calories early during transplantation in the form of more complex meals (Fig. 1e–h), an observation that was confirmed in plots of daily caloric intake (Fig. 1i; $P < 0.01$) and dietary α -diversity over time (Fig. 1j; $P < 0.01$).

After filtering to include evaluable microbiome samples with matching dietary data, we included 1,009 stool samples profiled by 16S rRNA gene sequencing from a subset of 158 patients (median, 5 samples per patient, range 1–28; Fig. 1c, Extended Data Fig. 1a and Supplementary Table 2). A temporal decline was observed in faecal microbiome α -diversity (Fig. 1k), as seen in other allo-HCT cohorts²⁶, that mirrored the trends in caloric intake and dietary α -diversity (Fig. 1i,j). Intake of the macronutrients carbohydrates, fat, fibres, protein and sugars all declined from hospitalization until day 12 (the median time to neutrophil engraftment in this cohort) (Fig. 1m; false-discovery rate (FDR)-adjusted $q < 0.001$). The average intake of all specific FNDDS food-group items except for ‘sugars, sweets and beverages’ also declined over this period (Fig. 1l; FDR-adjusted $q < 0.01$).

Shifts in microbiome composition (dissimilarity from baseline) were moderately, albeit significantly, associated with shifts in dietary composition from each patient’s baseline (Fig. 1n,o; slope = 0.6, $P < 0.001$). Substantial variability in microbiome perturbation remained unexplained by diet alone, suggesting the influence of other factors (Extended Data Fig. 1b). Despite these global dietary trends, inpatient and outpatient dietary intake was highly variable (Supplementary Table 6), calling for a granular investigation of dietary association with microbiome disruption patterns.

Low diversity after sugar and antibiotics

We developed a Bayesian model to quantify the contribution of dietary intake to microbiome composition in the context of potentially confounding clinical variables. We chose a dietary exposure period of 2 days preceding each faecal sample, as this window provided a reasonable correlation between microbiome composition and dietary composition (Fig. 2a) and was consistent with a previous report⁸. The model considers treatment regimen (conditioning) intensity³⁰ and exposures during the 2 days before sample collection: nutrition-support events (TPN and EN), antibiotics and dietary intake as fixed-effect predictors of microbiome composition; and an interaction term between diet and broad-spectrum-antibiotic exposure, which captures an additional dietary effect on the microbiome when antibiotics are administered (Fig. 2b).

All 173 patients received at least one antibiotic during their dietary collection period. They were typically treated initially with prophylactic antibiotics (fluoroquinolones and intravenous vancomycin). Importantly, 138 out of 173 (80%) patients were switched from prophylactic antibiotics to broader-spectrum antibiotics during this period when they developed fever or other signs of potential infection (most commonly piperacillin–tazobactam, cefepime, linezolid, or a carbapenem for neutropenic fever or bloodstream infection; and metronidazole or oral vancomycin for *Clostridioides difficile* diarrhoea). As they are major drivers of dysbiosis in HCT^{4,28,33}, we considered exposure to these broad-spectrum antibiotics explicitly in the model (Extended Data Fig. 2). To accommodate other exposures not explicitly captured

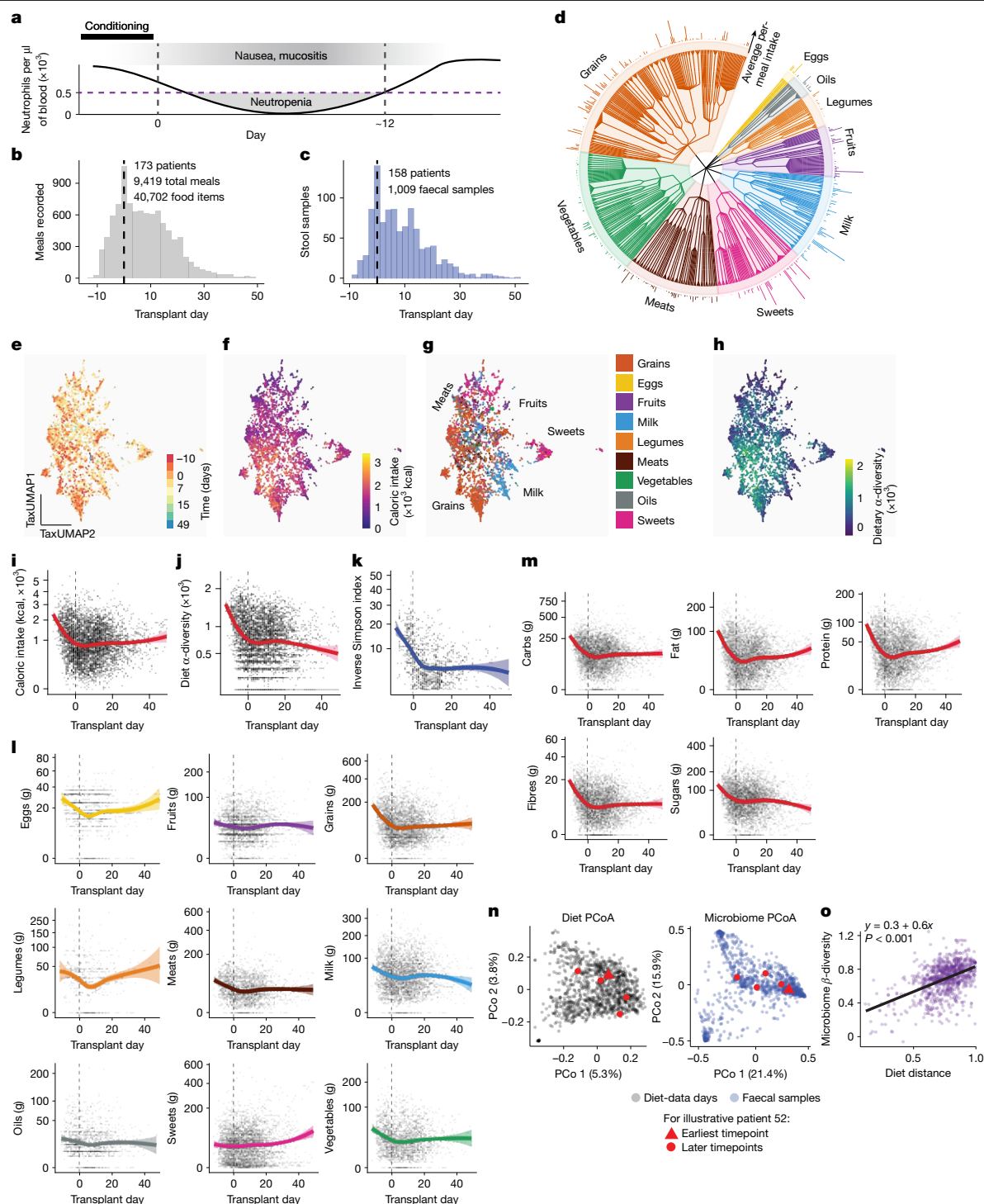


Fig. 1 | High-resolution nutritional intake data with longitudinal microbiota composition. **a**, Allo-HCT schematic. **b,c**, In total, 9,419 meal samples (**b**) and 1,009 evaluable stool samples (**c**) were collected from 173 patients. **d**, Food tree of 622 unique items according to FNDDS nomenclature; levels are derived from numeric food-code hierarchies. The tick marks indicate the average per-meal consumption of each item (imputed dehydrated mass). **e–h**, TaxUMAP⁴ visualization; each point is the food consumed by a patient on a single day, coloured by time relative to transplantation (**e**), caloric intake (**f**), most consumed food group (**g**) and dietary α -diversity (Faith's phylogenetic distance) (**h**). **i,j**, Caloric intake; each point is the daily food consumed by a patient (**i**) and daily diet α -diversity (**j**). **k**, Faecal microbiota α -diversity (inverse Simpson index); each point corresponds to a faecal sample. **l**, Daily intake of the nine FNDDS food groups (imputed dehydrated mass). **m**, daily consumption of macronutrients. Carbs, carbohydrates. The lines in **i–m** are locally estimated scatterplot

smoothing (LOESS) averages; the shading indicates the 95% confidence intervals. Analysis was performed using one-sided generalized estimating equation (GEE) Wald tests. See also Extended Data Fig. 4i. **n**, Principal coordinate analysis (PCoA) plots of dietary intake (unweighted UniFrac distances, whereby each grey point is the dietary composition in a 2-day window before collection of a faecal sample) (left), and of faecal microbiome composition (Bray–Curtis distances) (right). Longitudinal data from an illustrative patient (52) were replotted in red; the triangle and circles represent the earliest and subsequent timepoints, respectively. Baseline and subsequent data for all other patients are plotted as grey or blue circles. **o**, The relationship between diet distance (that is, dissimilarity of dietary composition relative to a patient's first recorded diet day) and faecal microbiome β -diversity relative to the first recorded sample. Fixed-effect slope of a linear mixed-effects model with per-patient random intercept; two-sided t -test $P = 6.7 \times 10^{-30}$, $n = 851$ sample pairs from 135 patients.

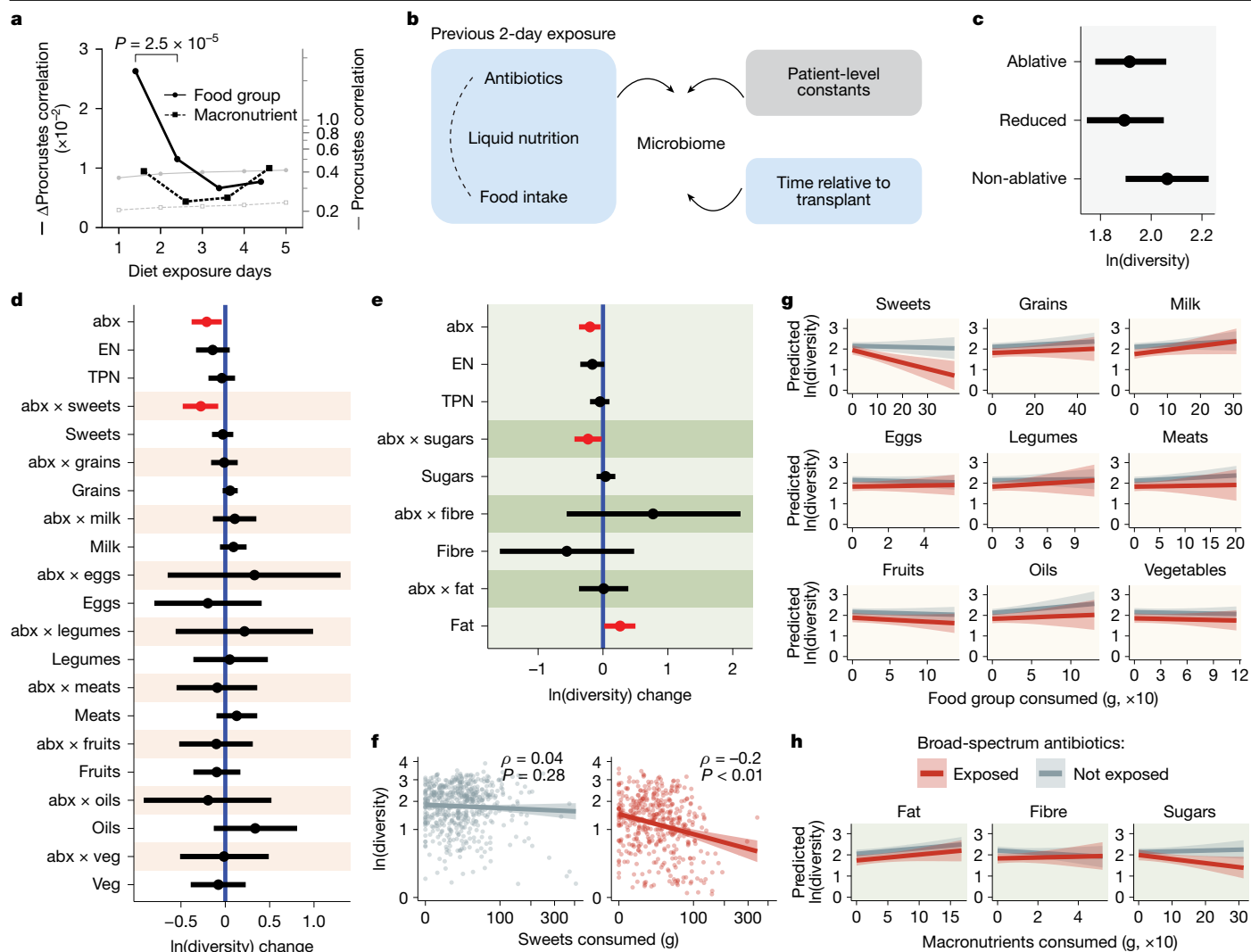


Fig. 2 | Consumption of sweets during broad-spectrum antibiotic treatment predicts low faecal microbiota α -diversity. **a**, Procrustes analyses to quantify changes (left axis, black) in the correlation (right axis, grey) between faecal microbiome composition and averaged dietary intake from preceding windows of varying length; by macronutrient composition (dashed) and by FNDDS food groups (solid). **b**, Diagram of the Bayesian model of microbiome features as a function of dietary intake during the preceding 2 days in the context of potential confounders, including TPN and tube feeds (EN), antibiotics and their interaction with diet (dashed line), patient-level exposures (conditioning intensity), and varying-effects terms per patient and weeks spent in hospital. The blue boxes show the time-varying quantities; the grey box shows constants. **c**, **d**, Posterior coefficient estimates for effects on microbiome α -diversity, assessed by the natural-log-transformed inverse Simpson index. **c**, Intercepts for conditioning intensity. **d**, Effects of the consumption of 100 g (dehydrated weight) of each

food group in the 2 days before faecal sample collection, antibiotics (abx) and antibiotic–food group interaction terms, and additional fixed effects for TPN and EN. **e**, Analogous regression using macronutrients; protein was excluded due to high correlation with fat. For **c–e**, the dots and lines indicate the posterior means and 95% CIs, respectively; red, 95% CIs not crossing zeros. $n = 1,009$ stool samples from 158 patients. **f**, Unadjusted correlation between sweets consumption during the previous 2 days and α -diversity, stratified by antibiotic exposure; lines were determined using linear regression with 95% confidence intervals. Two-sided Spearman test: exposed, $\rho = -0.195$, $P = 3.7 \times 10^{-5}$, 439 samples from 109 patients; unexposed, $\rho = 0.045$, $P = 0.28$, 570 samples from 128 patients. **g**, **h**, Posterior counterfactual: the predicted α -diversity across food group (**g**) or macronutrient intake (**h**), with or without antibiotic exposure, holding other variables constant. The lines show the posterior median and the shading shows the 95% CIs.

that might vary between patients and over time, we included varying intercepts for the week relative to transplantation and varying intercepts per patient (Extended Data Fig. 1c) to correct for differences between patients and to account for repeated samples from the same individuals.

We applied this model to 1,009 stool samples from 158 patients to quantify associations of antibiotics, clinical parameters and dietary food groups with microbiome α -diversity, a microbiome feature that we have previously correlated with mortality in this population²⁶. Consistent with our previous report³⁰, faecal samples from recipients of the mildest conditioning regimen (non-ablative) had the highest average α -diversity (Fig. 2c), and exposure to broad-spectrum antibiotics in the

previous 2 days was negatively associated with bacterial α -diversity, as expected (Fig. 2d; abx term, mean = -0.21 , 95% credible interval (CI) = -0.38 to -0.05). Notably, although intake of food items categorized as sugars, sweets and beverages (hereafter, sweets) alone was not associated with faecal microbiome α -diversity, their intake before the collection of antibiotic-exposed samples was associated with the largest diet-mediated decrease in α -diversity: for every 100 g increase of intake after broad-spectrum-antibiotic exposure, the model predicted an additional 24.1% decrease (95% CI = 7.7–38.1%) in mean α -diversity compared with samples without previous antibiotic exposure (Fig. 2d). Combined exposure to sweets and antibiotics was also associated with a shift toward compositions enriched for features such as community

domination, for example, by *Enterococcus faecium* (Extended Data Fig. 3). These associations suggested that dietary sugars could synergize to exacerbate the impact of antibiotics.

To verify this independently of the FNDDS classification of food items, we summarized the recorded meals by their macronutrient composition, performing an equivalent analysis on the now differently encoded dietary data. This confirmed that intake of sugars during antibiotic exposure (abx $\times$ sugars interaction term) was significantly additive to the antibiotic-associated α -diversity reduction (Fig. 2e), with a further mean decrease of 20.6% per 100 g (95% CI = 2–35.6% decrease). Although protein was omitted from the main macronutrient model as it was highly collinear with fat ($r = 0.82$, $P < 0.001$), a sensitivity analysis with terms for each of the major macronutrient categories confirmed the interaction between antibiotics and sugar intake predicting low α -diversity (Extended Data Fig. 4a,b).

These associations were surprising because sugar-enriched oral nutritional supplements are commonly recommended to transplant recipients³⁴ (Extended Data Fig. 4d). A higher-resolution model suggested that key contributors to the relationship between sugar-enriched foods, antibiotic exposure and low microbial α -diversity included formulated nutritional beverages (such as sports drinks and nutritional shakes; Extended Data Fig. 4e), as well as added, as opposed to naturally occurring, sugars (Extended Data Fig. 5a–d). To ascertain these findings, we examined several potential confounders in variants of the model. The interaction between sweets consumption and antibiotics remained a significant predictor of low microbiota α -diversity in models that considered underlying malignant disease, comorbidities or intensive analgesic exposure as an indicator of the most severely ill states (Extended Data Fig. 5e). We also confirmed the association was not simply explained by certain confounding factors, as it was still observed in subset analyses that excluded patients exposed to EN or TPN (Extended Data Fig. 5f), as well as in a subset that excluded samples that were collected during treatment of confirmed infections (Extended Data Fig. 5g). Finally, a similar finding was also observed in a subset analysis of samples that were collected early in transplantation, before the onset of severe mucositis, which can limit consumption of a regular diet (Extended Data Fig. 5h).

Having observed a robust relationship between microbiome disruption and concurrent exposure to antibiotics and sweets, we also analysed the raw data without adjustments for the many clinical variables encoded in the Bayesian model. Indeed, a statistically significant inverse relationship was apparent between grams of sweets consumed and faecal microbiota α -diversity in samples from patients who were exposed to antibiotics in the 2 days before sampling (Spearman correlation = -0.2 , $P < 0.01$), although substantial variation remained. However, this relationship was not observed in the absence of antibiotic exposure (Spearman correlation = 0.04 , $P = 0.28$) (Fig. 2f). Consistent with this, marginal-effects plots from the Bayesian model revealed an independent relationship between food group sweets intake and microbiota α -diversity only during periods of antibiotic exposure (Fig. 2g); the analogous model of macronutrients as dietary predictors also demonstrated a diverging effect of sugar intake on microbiota α -diversity dependent on antibiotic exposure (Fig. 2h). We also confirmed that these results did not depend on the way in which foods were classified (Extended Data Figs. 4f,g and 6a,b), nor driven by over-represented patients (Extended Data Fig. 5l), nor were they explained simply by total caloric intake in the 2 days before faecal sampling (Extended Data Fig. 4c).

Mortality linked to diet and antibiotics

As microbiome disruption is associated with poor prognoses after transplantation^{26–28}, we next examined whether the synergy between broad-spectrum antibiotics and high dietary sweets intake is predictive of clinical outcomes. We first dichotomized the cohort at the

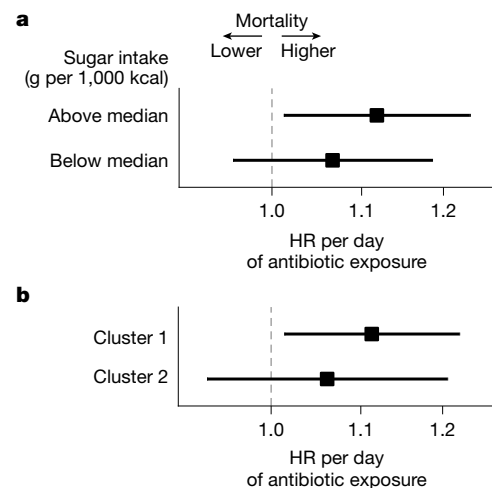


Fig. 3 | Antibiotic exposure correlates with mortality in patients consuming sugar-enriched diets. a, Adjusted HRs for overall survival associated with each additional day of broad-spectrum antibiotic exposure duration between days –7 and 12 relative to HCT, compared with antibiotic-unexposed patients, for those with above-median versus below-median sugar intake, calorie-adjusted and averaged between days –7 and 12. $n = 86$ (above-median sugar intake group) and $n = 86$ (below-median group who survived to landmark day 12). Survival analysis was landmarked at day 12 and adjusted for conditioning intensity, graft source and GVHD prophylaxis. Above-median sugar, HR = 1.12 (95% confidence interval = 1.01–1.23), two-sided Wald $P = 0.027$; below-median sugar, HR = 1.07 (95% confidence interval = 0.96–1.19), $P = 0.23$. Notably, without considering antibiotic exposure, these two groups had comparable mortality risk (multivariable HR = 0.72, 95% confidence interval = 0.18–2.85, $P = 0.6$). **b**, Adjusted HRs for overall survival associated with each additional day of broad-spectrum antibiotic exposure duration between days –7 and 12 relative to HCT, compared with antibiotic-unexposed patients, for the two dietary-patterns clusters. Cluster 1, HR = 1.11 (95% confidence interval = 1.01–1.22), two-sided Wald $P = 0.024$, $n = 114$ patients; cluster 2, HR = 1.06 (95% confidence interval = 0.93–1.21), $P = 0.36$, $n = 58$. The analysis includes an interaction term between dietary cluster and antibiotic exposure days, adjusting for conditioning regimen intensity, graft type and GVHD prophylaxis. $n = 172$ patients who survived to landmark day 12. In **a** and **b**, the squares are the adjusted HRs and the horizontal lines show the 95% confidence interval; the dashed vertical lines mark a HR of 1; two-sided Wald tests were not adjusted for multiple comparisons.

median sugar intake, normalized to total caloric intake and averaged across HCT days –7 to 12 (Supplementary Table 3). Among those who consumed an above-median intake of sugar, the duration of broad-spectrum antibiotic exposure correlated with a higher mortality risk (Fig. 3a; hazard ratio (HR) = 1.12, 95% confidence interval = 1.01–1.23, $P = 0.027$, for each additional day of antibiotic exposure in a model adjusted for conditioning intensity, graft source and GVHD prophylaxis; Supplementary Table 4 and Extended Data Fig. 6j). This was not observed in the below-median sugar-intake group (HR = 1.07, 95% confidence interval = 0.96–1.19, $P = 0.23$). A concordant result was observed in a separate analysis based on dietary timeseries by defining two clusters of dietary patterns using latent trajectory class analysis (Supplementary Table 5). Cluster 1 was characterized by lower average daily food intake (914 kcal per day less, $P < 0.001$) as compared to more sustained higher intake overall in cluster 2 (Extended Data Fig. 6e). Cluster 1 also had a higher proportion of sugar in the diet (10.5% higher, $P < 0.001$) compared with cluster 2 (Extended Data Fig. 6f). The duration of antibiotic exposure was significantly associated with higher mortality risk in cluster 1 (HR = 1.11, CI = 1.01–1.22, $P = 0.024$, for each additional day of antibiotic exposure) but not cluster 2 (HR = 1.06, CI = 0.93–1.21, $P = 0.36$) in a multivariable model accounting for conditioning intensity, graft source and GVHD prophylaxis (Fig. 3b and Supplementary Table 6). Patients in cluster 1 also

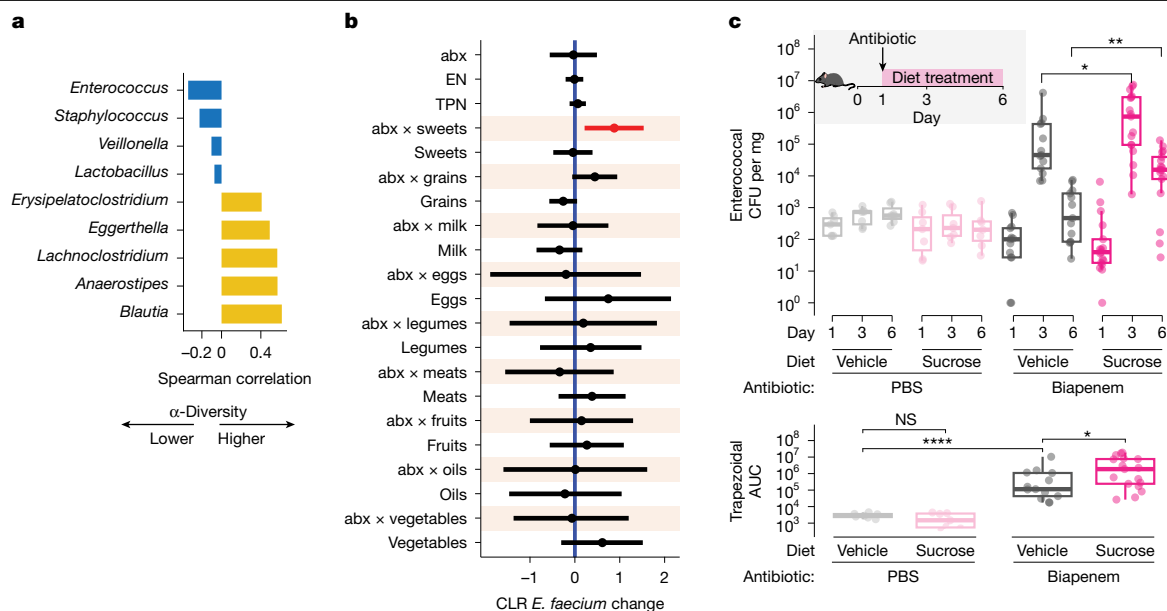


Fig. 4 | Sugar and antibiotics synergize to exacerbate *E. faecalis* expansion in the gut microbiome of humans and in mice. **a**, Genera most strongly correlated with faecal microbiome α -diversity (inverse Simpson index) in either direction. Benjamini–Hochberg-adjusted Spearman two-sided $P < 0.05$, among genera present in over 10% of samples with over 0.2% relative abundance. Blue and yellow shading represents negative and positive correlations between relative abundances and α -diversity, respectively. $n = 1,009$ samples from 158 patients; most correlated genera are plotted here (Extended Data Fig. 7b). **b**, Associations between *E. faecium* (amplicon sequence variant 1 (ASV_1)) relative abundance (CLR transformed) and broad-spectrum antibiotic exposure, TPN, EN and intake of food groups, plotted as in Fig. 2d. $n = 1,009$ stool samples from 158 patients (Extended Data Fig. 7a,c,d). **c**, Sucrose exacerbates *Enterococcus* expansion after antibiotic treatment in mice. Top, faecal enterococcal burden

as measured by dilution plating on selective medium. A schematic of the experimental timeline is shown at the top. Bottom, enterococcal abundance from the top panel analysed over time by trapezoidal area under the curve (AUC) to account for repeated measures from the same animals. The boxes represent the interquartile range, the centre line indicates the median, and the whiskers extend to the most extreme datapoint within $1.5 \times$ the interquartile range. $n = 7$ –17 mice per treatment group; 2 independent experiments. Two-sided Wilcoxon rank-sum test, unadjusted. P values were as follows: top, biapenem + vehicle versus biapenem + sucrose, $P = 0.048$ (day 3), $P = 0.0010$ (day 6); bottom, PBS + vehicle versus biapenem + vehicle, $P = 2.6 \times 10^{-5}$; biapenem + vehicle versus biapenem + sucrose, $P = 0.028$; PBS + vehicle versus PBS + sucrose, $P = 0.46$.

had longer hospitalizations after neutrophil engraftment (Extended Data Fig. 6i). Thus, mortality risk after allo-HCT correlated with duration of antibiotic exposure in patients who consumed dietary patterns characterized by relative sugar enrichment and overall lower intake.

Sugars exacerbate pathobiont expansion

We next examined whether sugar intake during antibiotic-induced microbiome perturbations might favour blooms of certain bacteria²⁶. Although a variety of taxa was correlated with food-group and antibiotic exposures (Extended Data Fig. 7a), we examined the genus *Enterococcus* and, specifically, *E. faecium* as the species that expanded most prominently during microbiome disruption ($r = -0.34$, FDR-adjusted $P < 0.001$; Fig. 4a and Extended Data Fig. 7b–d), consistent with our previous report²⁷. Notably, *Enterococcus* is an important cause of antibiotic-resistant bloodstream infections and is a microbiome surrogate of adverse clinical outcomes after HCT, including GVHD and mortality, and it exacerbates GVHD disease in mice^{4,27,35}. After antibiotic exposure, each 100 g of additional sweets intake was associated with a 0.89 U increase in the centred log ratio (CLR)-transformed abundance of *E. faecium* (95% CI = 0.22–1.58), compared with samples without previous antibiotic exposure (Fig. 4b and Extended Data Fig. 7e,f).

These observations in patients undergoing intensive cancer treatment suggested a model by which antibiotic-induced microbiome disruption, as manifested by α -diversity loss and pathobiont expansion, could be exacerbated by the consumption of sweets. To test this hypothesis experimentally, we established a system in which microbiome

disruption was induced in C57BL/6 mice with a single subcutaneous injection of the carbapenem antibiotic biapenem. This reproducibly induced a moderate expansion of endogenous enterococci by 2–3 orders of magnitude that peaked at day 3 and largely resolved by day 6, as assessed by dilution plating on enterococcal-selective agar (Fig. 4c). As predicted by the analysis of patient data, supplementation of standard diet with sucrose exacerbated the expansion of enterococci by 16.3-fold on day 3 ($P = 0.048$) and by 33.4-fold on day 6 ($P = 0.001$). Notably, sucrose treatment alone had no effect on enterococcal burden in the absence of antibiotics.

Recapitulating the patterns observed in patients, *Enterococcus* expansion was accompanied by a decline in α -diversity and a shift in global composition ($P < 0.001$; Extended Data Fig. 8a–f). We considered the possibility that sucrose supplementation might reduce the intake of chow, the fibre component of which might modulate the response of the microbiome to antibiotics^{12,13}. However, food consumption was not appreciably reduced by sucrose supplementation (Extended Data Fig. 8g,h), nor did these brief treatments significantly affect mouse weight (Extended Data Fig. 8m). Sucrose prolonged antibiotic-induced *Enterococcus* expansion in mice maintained on a fibre-free diet (Extended Data Fig. 8l), indicating a fibre-independent effect of sugar. Moreover, monosaccharide components of sucrose were sufficient on their own to augment the antibiotic-induced enterococcal bloom, as was blended fruit purée (Extended Data Fig. 9a–c). Sucrose supplementation had no significant effect on the kinetics of *Enterococcus faecalis* monocolonization in germ-free mice ($P = 0.7$, Extended Data Fig. 9d), suggesting that this sucrose effect requires the presence of a complex microbiota perturbed by antibiotics. Sucrose also prolonged the antibiotic-induced enterococcal expansion in mice treated with bone

marrow transplants (Extended Data Fig. 9h). Collectively, we provide evidence from real-world patients that consumption of dietary sugars during or after antibiotic exposure decreases microbiome α -diversity and drives *Enterococcus* expansion, a microbiome–diet dynamic that we validated in vivo.

Discussion

In patients undergoing allo-HCT, intake of sugar-rich foods during treatment with broad-spectrum antibiotics predicted microbiome disruptions that we previously linked to adverse outcomes: low α -diversity, and relative expansion of *Enterococcus*^{26,27}. Moreover, among patients in this cohort who consumed sugar-enriched diets, the duration of broad-spectrum antibiotic exposure predicted mortality risk. The increased mortality risk associated could also stem, in part, from lower food intake. The therapy-induced perturbations in this clinical setting, along with high-frequency serial dietary and faecal sampling, served as a natural experiment that enabled causal-inference approaches distinct from healthy-cohort or volunteer-feeding study designs^{1,8–10,12,15,18,32,36}. The hypothesis raised by this clinical observation—that sugars and antibiotics synergize to disrupt the microbiome—was tested experimentally in mice, in which we confirmed that sucrose exacerbated and prolonged antibiotic-induced expansion of endogenous *Enterococcus*. This is consistent with other reports that simple carbohydrates can induce dysbiosis and exacerbate experimental colitis or metabolic disease^{37–41}, and it expands the dietary determinants of antibiotic-induced microbiome perturbation beyond the well-described role of fibre^{12,13}.

Several mechanisms could underlie this synergy. Sucrose could have a positive effect on the growth of pathobionts such as *Enterococcus*⁴², or a suppressive effect on colonization resistance that is usually conferred by specific commensal organisms^{43,44}. Alternatively, as most simple sugars are absorbed in the small intestine, sucrose may modulate the host, which, in turn, shapes microbiota composition^{39,45}. Limitations in the analysis of the human cohort include the practice of near-universal prophylactic use of fluoroquinolones, so the reported interaction may more precisely reflect dietary intake in the context of escalation to broad-spectrum antibiotics, rather than exposure to any antibiotic per se. Furthermore, as the initiation of broad-spectrum antibiotics is clinically driven (for example, fever or suspected infection) there may be a confounding association with the severity of the condition necessitating antibiotic therapy. Although patients were invited to report dietary intake outside of the items provided by the hospital, records regarding outside food are incomplete. Moreover, transient symptoms can affect dietary preferences and introduce additional time-varying confounders.

Despite these limitations, our results identify dietary simple sugars as a modifiable amplifier of antibiotic-induced microbiome disruption. In the context of HCT, prospective trials are needed to test whether corresponding dietary modulation improves clinical outcomes. More broadly, limiting sugar intake during antibiotic therapy may help mitigate microbiome injury.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-026-11077-3>.

- Jumpertz von Schwartzberg, R. et al. Caloric restriction disrupts the microbiota and colonization resistance. *Nature* **595**, 272–277 (2021).
- Belkaid, Y. & Hand, T. W. Role of the microbiota in immunity and inflammation. *Cell* **157**, 121–141 (2014).

- Costea, P. I. et al. Enterotypes in the landscape of gut microbial community composition. *Nat. Microbiol.* **3**, 8–16 (2018).
- Schluter, J. et al. The TaxUMAP atlas: efficient display of large clinical microbiome data reveals ecological competition in protection against bacteremia. *Cell Host Microbe* **31**, 1126–1139 (2023).
- Litvak, Y. & Bäuml, A. J. Microbiota-nourishing immunity: a guide to understanding our microbial self. *Immunity* **51**, 214–224 (2019).
- Turnbaugh, P. J., Backhed, F., Fulton, L. & Gordon, J. I. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* **3**, 213–223 (2008).
- Gacesa, R. et al. Environmental factors shaping the gut microbiome in a Dutch population. *Nature* **604**, 732–739 (2022).
- Johnson, A. J. et al. Daily sampling reveals personalized diet-microbiome associations in humans. *Cell Host Microbe* **25**, 789–802 (2019).
- Wu, G. D. et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108 (2011).
- Wastyk, H. C. et al. Gut-microbiota-targeted diets modulate human immune status. *Cell* **184**, 4137–4153 (2021).
- Bernard-Raichon, L. et al. Gut microbiome dysbiosis in antibiotic-treated COVID-19 patients is associated with microbial translocation and bacteremia. *Nat. Commun.* **13**, 5926 (2022).
- Tanes, C. et al. Role of dietary fiber in the recovery of the human gut microbiome and its metabolome. *Cell Host Microbe* **29**, 394–407 (2021).
- Ng, K. M. et al. Recovery of the gut microbiota after antibiotics depends on host diet, community context, and environmental reservoirs. *Cell Host Microbe* **26**, 650–665 (2019).
- Tuganbaev, T. et al. Diet diurnally regulates small intestinal microbiome-epithelial-immune homeostasis and enteritis. *Cell* **182**, 1441–1459 (2020).
- David, L. A. et al. Host lifestyle affects human microbiota on daily timescales. *Genome Biol.* **15**, R89 (2014).
- Desai, M. S. et al. A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. *Cell* **167**, 1339–1353 (2016).
- Turnbaugh, P. J. et al. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci. Transl. Med.* **1**, 6ra14 (2009).
- Asnicar, F. et al. Gut micro-organisms associated with health, nutrition and dietary interventions. *Nature* <https://doi.org/10.1038/s41586-025-09854-7> (2025).
- Zhernakova, A. et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* **352**, 565–569 (2016).
- Rothschild, D. et al. Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**, 210–215 (2018).
- Smits, S. A. et al. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* **357**, 802–806 (2017).
- Vangay, P. et al. US immigration westernizes the human gut microbiome. *Cell* **175**, 962–972 (2018).
- Archer, E., Hand, G. A. & Blair, S. N. Validity of U.S. nutritional surveillance: National Health and Nutrition Examination Survey caloric energy intake data, 1971–2010. *PLoS ONE* **8**, e76632 (2013).
- Subar, A. F. et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN Study. *Am. J. Epidemiol.* **158**, 1–13 (2003).
- Baumgartner, A. et al. Revisiting nutritional support for allogeneic hematologic stem cell transplantation—a systematic review. *Bone Marrow Transpl.* **52**, 506–513 (2017).
- Peled, J. U. et al. Microbiota as predictor of mortality in allogeneic hematopoietic-cell transplantation. *N. Engl. J. Med.* **382**, 822–834 (2020).
- Stein-Thoeringer, C. K. et al. Lactose drives *Enterococcus* expansion to promote graft-versus-host disease. *Science* **366**, 1143–1149 (2019).
- Shono, Y. et al. Increased GVHD-related mortality with broad-spectrum antibiotic use after allogeneic hematopoietic stem cell transplantation in human patients and mice. *Sci. Transl. Med.* **8**, 339ra71 (2016).
- Jenq, R. R. et al. Intestinal blautia is associated with reduced death from graft-versus-host disease. *Biol. Blood Marrow Transpl.* **21**, 1373–1383 (2015).
- Shouval, R. et al. Conditioning regimens are associated with distinct patterns of microbiota injury in allogeneic hematopoietic cell transplantation. *Clin. Cancer Res.* **29**, 165–173 <https://doi.org/10.1158/1078-0432.CCR-22-1254> (2022).
- Craig, P., Katikireddi, S. V., Leyland, A. & Popham, F. Natural experiments: an overview of methods, approaches, and contributions to public health intervention research. *Annu. Rev. Publ. Health* **38**, 39–56 (2017).
- Schluter, J. et al. The gut microbiota is associated with immune cell dynamics in humans. *Nature* **588**, 303–307 (2020).
- Holler, E. et al. Metagenomic analysis of the stool microbiome in patients receiving allogeneic stem cell transplantation: loss of diversity is associated with use of systemic antibiotics and more pronounced in gastrointestinal graft-versus-host disease. *Biol. Blood Marrow Transpl.* **20**, 640–645 (2014).
- Baumgartner, A. & Schuetz, P. in *EBMT Handbook: Hematopoietic Stem Cell Transplantation and Cellular Therapies* Ch. 24 (Springer, 2019).
- Papanicolaou, G. A. et al. Bloodstream infection due to vancomycin-resistant *Enterococcus* is associated with increased mortality after hematopoietic cell transplantation for acute leukemia and myelodysplastic syndrome: a multicenter, retrospective cohort study. *Clin. Infect. Dis.* **69**, 1771–1779 (2019).
- Korem, T. et al. Bread affects clinical parameters and induces gut microbiome-associated personal glycemic responses. *Cell Metab.* **25**, 1243–1253 (2017).
- Montrose, D. C. et al. Dietary fructose alters the composition, localization, and metabolism of gut microbiota in association with worsening colitis. *Cell. Mol. Gastroenterol. Hepatol.* **11**, 525–550 (2021).
- Khan, S. et al. Dietary simple sugars alter microbial ecology in the gut and promote colitis in mice. *Sci. Transl. Med.* **12**, eaay6218 (2020).
- Kawano, Y. et al. Microbiota imbalance induced by dietary sugar disrupts immune-mediated protection from metabolic syndrome. *Cell* **185**, 3501–3519 (2022).

40. Laffin, M. et al. A high-sugar diet rapidly enhances susceptibility to colitis via depletion of luminal short-chain fatty acids in mice. *Sci. Rep.* **9**, 12294 (2019).
41. Isaac, S. et al. Microbiome-mediated fructose depletion restricts murine gut colonization by vancomycin-resistant *Enterococcus*. *Nat. Commun.* **13**, 7718 (2022).
42. Kim, M.-A., Rosa, V. & Min, K.-S. Characterization of *Enterococcus faecalis* in different culture conditions. *Sci. Rep.* **10**, 21867 (2020).
43. Ng, K. M. et al. Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature* **502**, 96–99 (2013).
44. Spragge, F. et al. Microbiome diversity protects against pathogens by nutrient blocking. *Science* **382**, eadj3502 (2023).
45. Faber, F. et al. Host-mediated sugar oxidation promotes post-antibiotic pathogen expansion. *Nature* **534**, 697–699 (2016).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026

Methods

Patients

In this cohort study of adult allo-HCT recipients at Memorial Sloan Kettering Cancer Center in New York between 2017 and 2022, eligible participants had been admitted for Allo-HCT. Three times weekly, patients who had been recently admitted for allo-HCT were invited to participate in dietary-intake collection. Neutrophil engraftment was defined as the first of 3 days of neutrophil count ≥ 500 per μl . Five patients died without achieving engraftment and were excluded from the analysis of median time to engraftment. No patients were lost to follow-up. For mortality analysis, one patient died before landmark day 12 and was excluded from that analysis. The latest follow-up/data collection date was in April 2023.

Nutrition data collection and annotation

The hospital kitchen commercial computer system (Computrition) was configured to provide a printout that accompanied each meal tray to the bedside on which patients were asked to indicate, immediately after each meal, whether they consumed 0, 25%, 50%, 75% or 100% of each item they ordered for that meal. These instruments were collected by a dietitian or other research team members three times weekly, during which missing entries were completed at the bedside along with encouragement to sustain motivation with the project. Consumption data were entered into the kitchen software, which was linked to the recipe and mass of each item. Data were manually vetted by a research dietitian for quality control (for example, grossly implausible kcal values resulting from sporadic typographic errors in the hospital kitchen records). Habitual dietary patterns before admission were not surveyed; the analysis was focused on inpatient dietary intake. Per-patient intercept terms were included in the Bayesian model (see below), which may capture interpatient variability in unmeasured parameters such as patient-specific pretransplant dietary habits.

An eight-digit food code was assigned to each unique food item according to the classification of the FNDDS⁴⁶. The numeric codes are structures such that subsequent numeral positions differentiate foods within larger groups. For example, in the food code 14109010 for Swiss cheese, the first digit, 1, denotes milk and milk products; the second digit, 4, denotes cheeses; and each subsequent digit conveys progressively higher resolution classifications of foods. We removed the water content from the food total weight before analysis. To compute the dehydrated weight of consumed foods, we subtracted the water content, derived from FNDDS water fractions, from the total food weight. However, for EN, which is administered as a liquid and lacks a standardized total weight in grams, we used an additive approach. The dehydrated weight for these enteral formulations was calculated by summing the total gram weight of their protein, fat and carbohydrate components. This approach allows a more accurate comparison of nutrient intake across individuals and different food types, as it eliminates the variability introduced by water content. Therefore, when we refer to each 100 g increase in sweets, we are referring to 100 g of the dehydrated weight of foods within the 'sweets' category. This ensures that we are comparing the actual nutrient content, including sugar, rather than the total weight or volume of the food consumed. The analysed dataset therefore listed the dehydrated weight in grams of FNDDS food items consumed in each meal on each day of hospitalization.

Foodtree construction

A food tree was constructed with the 622 unique FNDDS items consumed by the patients in this cohort, as done previously⁷ (https://github.com/knights-lab/Food_Tree). The tree spans nine broad FNDDS food groups, namely, grain products (abbreviated here as grains); vegetables; meat, poultry, fish and mixtures (abbreviated here as meats); milk and milk products (abbreviated here as milk); sugars, sweets and

beverages (abbreviated here as sweets); fruits; dry beans, peas, other legumes, nuts and seeds (abbreviated here as legumes); fats, oils and salad dressings (abbreviated here as fats); and eggs. In some cases, food items not explicitly classified in FNDDS contained ingredients from multiple categories (such as milk-based mango smoothie). We addressed this by manually categorizing foods on the basis of which FNDDS food description they fit best to. For example, milk-based mango smoothie, despite having mango in it, fits the description of '11553110 fruit smoothie, with whole fruit and dairy'; it was therefore classified in the 'milk and milk products' group.

Dietary data analysis: TaxUMAP and diet α -diversity

The hierarchical organization of the FNDDS vocabulary facilitated application of α -diversities to diet data using Faith's phylogenetic distance⁴⁷, which was implemented in Qiime2 (qiime2-2021.11)⁴⁸ using the qiime diversity alpha-phylogenetic function with the faith_pd metric. The food tree taxonomy was used, as well as the dehydrated weight consumption of the food item represented by food code per patient per day. The TaxUMAP method⁴ was used to visualize compositional similarities between the patients' daily meals, similar to β -diversity. The fraction of each consumed food represented by a food code per patient per day was used to calculate the food tree taxonomy. Trends in microbiome and nutrition dynamics over time (Fig. li–k) were analysed by GEEs using the `geeglm` function in `geepack` (v.1.3.9) in R.

Human faecal microbiome analysis

Faecal sample inclusion criteria and flow through the study are detailed in Extended Data Fig. 1a. Samples were transported from the inpatient transplant unit to the laboratory through a pneumatic tube system at room temperature, promptly refrigerated and then aliquoted and frozen at -80°C within 1 business day. Microbiome profiling by 16S rRNA sequencing was performed in the MSK Molecular Microbiology Facility as described⁴⁹. In brief, bacterial cell walls were disrupted, and nucleic acids isolated using silica bead-beating and phenol–chloroform extraction, and the V4–V5 variable region of the 16S rRNA gene was amplified. Amplicons were purified using either the Qiagen PCR Purification Kit (Qiagen) or AMPure magnetic beads (Beckman Coulter) and quantified using the TapeStation instrument (Agilent). No enrichment of microbial DNA was performed. DNA was pooled to equal final concentrations for each sample and then sequenced on the Illumina platform. Of the 1,009 human faecal samples in this study, 42 were sequenced more than once. The median Bray–Curtis distance between distinct microbiome profiles generated from these same samples was 0.01, which was substantially smaller than the distance between samples from the same patients collected on different days (median value, 0.54; $P < 0.0001$) and also smaller than the distance between samples from different patients (median value, 0.89; $P < 0.0001$). This suggests that much more variation in composition is explained by biological differences than by technical sequencing variation. The median read count was 132,978 and the range was from 1,196 to 8,698,524. The 16S sequencing data were analysed using the R package DADA2 (v.1.16.0) pipeline with the default parameters except for `maxEE=2` and `truncQ=2` in the `filterandtrim()` function⁵⁰, 16S FASTQ files were capped at 10^5 reads per sample. ASVs were annotated according to NCBI 16S database using BLAST⁵¹. Microbiome α -diversity was evaluated using the inverse Simpson index—a summary statistic of both the richness and evenness of the bacterial community. Taxa abundances were summarized at the genus level.

Relationships between microbiome α -diversity and clinical factors such as conditioning intensity, exposures to antibiotics, sweets and parenteral nutrition were analysed mainly through Bayesian models that accounted for repeated sampling, time relative to transplantation and other patient-level variables (see the 'Bayesian multilevel model' section below). To facilitate visualization of relationships in the unadjusted

raw dataset, α -diversity values in various sample subsets are plotted as beeswarm plots in Extended Data Fig. 1f.

To address taxonomic heterogeneity within the *Enterococcus* genus and the pattern that *E. faecium* is the most commonly dominating species in this dataset (Extended Data Fig. 7d), we constructed a hybrid taxonomic feature table for the β -diversity analysis in Extended Data Fig. 3. In this taxonomic feature table, the *Enterococcus* genus was partitioned into two distinct features: (1) *Enterococcus* ASV_1 (identified as *E. faecium* through paired shotgun sequencing in a subset of samples (Extended Data Fig. 7c,d), and (2) a composite feature containing all remaining non-ASV_1 *Enterococcus* reads combined. All other taxa were aggregated at the genus level. β -Diversity was calculated using the Bray–Curtis dissimilarity metric on this hybrid feature table. For visualization, samples were classified as dominated if the relative abundance of any single feature (including the split *Enterococcus* features) exceeded 30%.

Mouse 16S rRNA gene sequencing and analysis

Mouse stool samples were sequenced using the ZymoBIOMICS Targeted Sequencing Service (Zymo Research). DNA was extracted using the ZymoBIOMICS-96 MagBead DNA Kit (Zymo Research) on an automated platform. The V3–V4 region of the bacterial 16S rRNA gene was amplified using the Quick-16S NGS Library Prep Kit (Zymo Research) with custom primers designed for optimal coverage and sensitivity. PCR reactions were performed in real-time PCR machines to control cycles and limit chimera formation. The final pooled library was cleaned, quantified and sequenced on the Illumina NextSeq system with the P1 reagent kit (600 cycles) using a 30% PhiX spike-in. ZymoBIOMICS Microbial Community Standards were included as positive controls for DNA extraction and library preparation. Negative controls were also included to assess for contamination. Bioinformatics analysis was performed as described in the following sections.

ASV inference. Unique ASVs were inferred from raw reads using the DADA2 pipeline⁵⁰, with removal of potential sequencing errors and chimeric sequences.

Taxonomic classification. Taxonomy assignment was performed using Uclust from QIIME (v.1.9.1)⁵² with the Zymo Research Database as a reference. α -Diversity metrics were calculated using QIIME v.1.9.1. PCoA was performed using the vegan (v.2.5-7) package⁵³ to visualize β -diversity. Differential abundance analysis was performed using MaAsLin2 (v.1.24.1)⁵⁴ to identify differentially abundant taxa between groups. The following parameters were used: normalization = “TSS”, transform = “LOG”, analysis_method = “LM”, max_significance = 0.05, min_abundance = 0.0001, min_prevalence = 0.10. Volcano plots were generated to visualize the results of the differential abundance analysis. The x axis of the volcano plot represents the coefficient from the MaAsLin2 model, indicating the direction and magnitude of the effect. For example, *E. faecalis* (seq2) exhibited the most significant increase in relative abundance in the biapenem + sucrose group compared with the biapenem-only group (coefficient = 1.58, FDR-adjusted $q < 0.05$). This coefficient of 1.58 means an approximately threefold change (or a threefold increase) in relative abundance, and the y axis represents the $-\log_{10}[q]$, indicating the statistical significance of the association.

Whole-metagenomic shotgun sequencing

For whole-metagenomic shotgun sequencing, profiles were analysed from 333 human and 24 mouse samples. DNA extraction and sequencing was performed as previously described⁴⁹. Metagenomic sequencing data were preprocessed and taxonomic profiles were generated using vdblab-shotgun (<https://github.com/vdblab/vdblab-shotgun/>). In brief, the samples were deduplicated and quality-filtered using clumpify and bbduk from the BBTools package. Removal of potential human host or mouse DNA was performed with the two-step process described previously⁵⁵. Taxonomic profilers

were generated with Metaphlan v.4.0.5, database version mpa_vJan21_CHOCOPhAnSGB_202103 (ref. 56).

Procrustes test

As we collected serial faecal samples and serial dietary intake data, a question arose as to how many previous days of dietary intake should be correlated with any given faecal sample. We therefore assessed concordance between microbiome and dietary datasets by Procrustes analysis, as done previously⁸ across dietary exposure windows of 1 to 5 days preceding each faecal sample. We considered dietary data in two alternative intake metrics: first, by macronutrients; and second, by specific named foods in FNDDS categories. For each window, food-group intake was represented by unweighted UniFrac distances on FNDDS food codes and macronutrient intake by Bray–Curtis distances on five macronutrients, each summarized as the mean daily intake over the window, and stool composition was represented by Bray–Curtis distances at the genus level. Each distance matrix was ordinated by PCoA, and the symmetric Procrustes correlation was computed with the Procrustes function from the vegan (v.2.5-7) package⁵³, with each ordination scaled to unit sum of squares so that the residual sum of squares of the symmetric Procrustes fit (M^2) lies between 0 and 1. For 801 stool samples, each of the 5 preceding days was recorded, either as food intake or as verified zero food intake. For 16 of these, the most recent preceding day was a day of verified zero food intake, which gives an empty food composition and therefore no defined UniFrac distance; these were removed, leaving a final analysis set of 785 samples from 143 patients. The marginal gain is the difference in the Procrustes correlation (Δ Procrustes) between successive windows. Confidence intervals and P values were obtained by a delete-one-patient jackknife; the P value is a two-sided test of whether the gain from a given 1-day extension differs from the gain from the next 1-day extension, a difference that was largest and significant for the extension from 1 day to 2 days on food group intake data (Fig. 2a). Moreover, a 2-day window aligns with typical human gut transit time and is consistent with a previous report⁵⁷.

Assessing the relationships between dietary changes and the microbiome changes

To assess the relationship between dietary shifts and microbiome changes (Fig. 1n,o), we calculated dissimilarities from baseline for both dietary composition and faecal microbiome profiles. Dietary dissimilarity was calculated using unweighted UniFrac distances between food-code-level intake for the 2-day period preceding each stool sample. Microbiome dissimilarity was calculated using Bray–Curtis distances at the genus level. For both stool and diet, we calculated the distance between each sampling timepoint and that patient’s baseline. We then constructed a linear mixed-effects model to assess the association between diet distance and stool distance, while accounting for individual patient differences. This model allowed us to test the hypothesis that, on average, as the dissimilarity increases between a patient’s initial and subsequent dietary composition, so does the dissimilarity between their initial and subsequent stool microbiome composition.

Bayesian multilevel model

Bayesian multilevel models were used to analyse these data due to their advantages in handling repeated measurements and providing understanding of uncertainty compared to traditional generalized linear models.

Data preparation. For macronutrient analysis, for each stool sample the included dietary data were summarized as the previous 2-day average intake of sugars, fibres and fat in grams. For the food-group analysis, for each stool sample, the included dietary data were summarized as the previous 2-day average intake of grains, vegetables, meats, milk, sweets, fruits, legumes, fats and eggs. Weights in grams were divided by 100 so that the resulting coefficients represent the expected change in

the outcome variable per 100 g intake of the dietary component. With respect to transplant characteristics, conditioning intensity exhibited considerable collinearity with graft source (peripheral blood stem cells versus bone marrow versus cord blood) and with GVHD prophylaxis regimen (Cramér's $V > 0.5$, $P < 0.001$ for both comparisons). As we have previously reported associations between conditioning intensity and microbiome disruption³⁰, we included this variable in the model as a three-level factor variable comprising (in increasing order of intensity) nonmyeloablative (nonablative), reduced intensity conditioning (reduced) and myeloablative (ablative). An alternative variant of the model that included a three-way interaction between sweets, antibiotics and conditioning intensity confirmed that the main sweets–antibiotic interaction predicting low microbial diversity was observed in both ablative and reduced-intensity conditioning (Extended Data Fig. 4j). EN, TPN and patient-controlled analgesia were binary variables encoded as true if the patient was exposed in the 2 days preceding faecal collection and otherwise false. Likewise, if the stool sample was collected after exposure to microbiome-damaging antibiotics in the 2-day window before it, it was set to true; otherwise, false. The microbiome-perturbing antibiotics considered were piperacillin–tazobactam, carbapenems, cefepime, linezolid, oral vancomycin and metronidazole; these were most commonly used as empiric therapy (most commonly for neutropenic fever) or in a pathogen-directed manner (most commonly for bloodstream infections or *C. difficile* diarrhoea)²⁶. Prophylactic antibiotics fluoroquinolones^{58,59} and intravenous vancomycin^{60,61} were not considered.

Association of foods and antibiotics with microbiome diversity were also analysed as interaction terms, with the assumption that food correlations with microbiome may be different depending on whether the patients had antibiotics in the previous 2-day window or not. We considered repeated measurements and therefore per-patient differences in microbiome features across samples by implementing a per-patient varying intercept. We also considered additional exposures during hospitalization that we did not capture explicitly and could therefore confound the analyses. Thus, we also accounted for time by grouping samples taken during similar timepoints along the therapy course into weekly categorical time bins, as done previously^{62,63}; these time bins were included in the model as varying intercepts. Our reasoning for this implementation was that we lack a biologically informed prior model for the effect of time on the microbiome over the short course of several weeks, and because the a priori known causal effectors of microbiome changes such as antibiotic exposures and diet were explicitly included in the model. We partially pooled the data using a varying-effects implementation because we believe that, while the microbiome could vary over time due to unmeasured confounders during the highly planned and controlled HCT treatment regimens, the time bins do contain mutual information justifying their cross-talk during Bayesian inference. The time bins were implemented in the format of weeks relative to transplant; for example, the sampling window expressed in days as [−7,0) is the week before transplant, and [7,14) is the second week after transplant.

When the outcome was microbiome α -diversity, it was analysed as the natural-log-transformed inverse Simpson's index. When the outcome was genus abundance (Extended Data Fig. 7a), it was analysed as CLR-transformed raw ASV counts of the genus after adding a pseudocount of 0.5 reads using the `clr` function in the `compositions` package (v.2.0-6)⁶⁴.

Model construction, sampling and result visualization. The `brms` (v.2.16.3)^{65–67} and `rstan` (v.2.26.4)⁶⁸ packages were used to build and run the model, with the model formula:

```
log(simpson_reciprocal) ~ 0 + ave_fiber + ave_fat + ave_Sugars
+ ave_fiber : abx + ave_fat : abx + ave_Sugars : abx + intensity
+ EN + TPN + abx + (1|pid) + (1|timebin)
```

for nutritional intake represented as the macronutrients, or:

```
log(simpson_reciprocal) ~ 0 + ave_fruit + ave_meat + ave_milk
+ ave_oils + ave_egg + ave_grain + ave_sweets + ave_legume
+ ave_veggie + ave_fruit : abx + ave_meat : abx + ave_milk : abx
+ ave_oils : abx + ave_egg : abx + ave_grain : abx + ave_sweets : abx
+ ave_legume : abx + ave_veggie : abx + intensity + EN + TPN
+ abx + (1|pid) + (1|timebin)
```

for nutritional intake represented as the food groups.

In the above model formula, `log(simpson_reciprocal)` represents the natural logarithm of the inverse Simpson index α -diversity metric. The `ave_` terms (such as `ave_fruit`, `ave_meat`) represent the average daily intake of each food group in the 2 days before the faecal sample collection. The interaction terms (such as `ave_fruit:abx`) assess how the effect of each food group on diversity differs depending on antibiotic exposure. The model also includes terms for conditioning intensity (`intensity`), EN, TPN and antibiotic exposure (`abx`). Finally, `(1|pid)` and `(1|timebin)` are varying intercept terms that account, respectively, for repeated measurements within patients (patient identifier) and for consistent temporal effects of hospitalization not captured by the explicitly modelled variables. This multilevel structure induces partial pooling, a regularization technique that adaptively weights information across participants^{69,70}. Partial pooling prevents participants with a high frequency of samples from dominating the population-level estimates (overfitting) while simultaneously improving estimates for participants with fewer samples by shrinking them toward the population mean, thereby robustly handling the unbalanced sampling structure inherent to clinical observational cohorts.

Three levels of the conditioning intensity were used as intercept terms with normal distributions with a mean of 2 and s.d. of 0.1 as priors, intercept food group effects in units of per 100 g with or without antibiotic exposure were set to normal distributions with a mean of 0 and s.d. of 1; the coefficient for the binary indicator for antibiotic exposure was set to a normal distributions with a mean of 0 and s.d. of 0.5; EN and TPN exposure were set to normal distributions with a mean of 0 and an s.d. of 0.1. Intercepts were a priori centred on the median diversity observed in previous HCT cohorts. We selected priors to induce a regularizing effect on the many coefficients, reducing the risk of overfitting, and we confirmed that the model was constrained to realistic data domains (Extended Data Fig. 1g). The inference parameters were set as follows: `warmup` = 1000, `iter` = 3000, `control` = `list(adapt_delta = 0.99)`, `cores` = 16, `chains` = 2, `seed` = 123, which means that the model will sample 1,000 steps during warm up, and 3,000 iterations in two chains, and `adapt_delta` is raised to 0.99 instead of 0.8 to avoid divergent transitions.

To quantify the effect of dietary predictors in concert, we leveraged the full posterior to conduct marginal-effects analyses (Fig. 2g,h), as well as posterior predictive checks (Extended Data Fig. 1h). Prior and posterior predictions were visualized with `ggplot` (v.3.3.5)⁷¹, `tidybayes` (v.3.0.2)⁷² and `ggpubr` (v.0.4.0) packages.

When modelling taxon abundance, the formula was changed to:

```
CLR(taxon) ~ 0 + ave_fruit + ave_meat + ave_milk + ave_oils
+ ave_egg + ave_grain + ave_sweets + ave_legume + ave_veggie
+ ave_fruit : abx + ave_meat : abx + ave_milk : abx + ave_oils : abx
+ ave_egg : abx + ave_grain : abx + ave_sweets : abx
+ ave_legume : abx + ave_veggie : abx + intensity + EN + TPN
+ abx + (1|pid) + (1|timebin)
```

with the same running parameters.

Instead of relying on P -value adjustments for multiple testing, our Bayesian approach with regularizing priors provides a conservative

Article

assessment of the associations between dietary components and taxon abundances. This approach leverages the full posterior probability distribution to quantify uncertainty, offering a more comprehensive understanding of the relationships.

To identify clinical and dietary correlates of global community shifts, we modelled the sample coordinates on the first two principal coordinates (PCoA 1 and PCoA 2) as continuous outcomes. We used the same Bayesian multilevel regression framework described above, with the formula:

```
PCoA_Axis ~ 0 + ave_fruit + ave_meat + ave_milk + ave_oils  
+ ave_egg + ave_grain + ave_sweets + ave_legume + ave_veggie  
+ ave_fruit : abx + ave_meat : abx + ave_milk : abx + ave_oils : abx  
+ ave_egg : abx + ave_grain : abx + ave_sweets : abx + ave_legume : abx  
+ ave_veggie : abx + intensity + EN + TPN + abx + (1|pid) + (1|timebin)
```

This allowed us to estimate the ‘vector’ of community movement (along axis 1 or axis 2) associated with specific dietary–antibiotic interactions while controlling for repeated measures and clinical confounders.

The marginal effects of each food group consumption on the predicted diversity are calculated with the conditional_effects function from brms package. The method used in the function was posterior_epred. While visualizing the predicted diversity as the intake of one food group changes, the rest of the food groups intake are held at average level, with intensity of the conditioning regimen set to be nonablative with no TPN or EN exposure.

For analysis of the association between diet, antibiotic exposures and α -diversity considering the reason broad-spectrum antibiotics (Extended Data Fig. 5g), samples were classified according to the reason for broad-spectrum antibiotics in the 2-day window before faecal sample collection. First, by pathogen-directed treatment either for bloodstream infection (91 samples from 27 patients) or for *C. difficile* (101 samples from 17 patients). Samples exposed to broad-spectrum antibiotics without either of these documented sources of infection were then classified as ‘empiric only’ (247 samples from 73 patients). The remaining 570 samples (from 128 patients) were classified as no broad-spectrum exposure.

In silico analysis of ASV-to-species specificity within *Enterococcus*.

To determine the theoretical species-level specificity of the *Enterococcus* ASVs, we compared them to a database of all possible V4–V5 amplicons derived from complete *Enterococcus* genomes in the RefSeq database. This analysis was performed using the SYTYCS pipeline (<https://github.com/nickp60/sytycs>). In brief, theoretical V4–V5 amplicons were identified from the genomes using PrimerSearch⁷³ with the 16S PCR primer sequences used herein (563F, AYTGGGYDTAAAGNG; and 926Rb, CCGTCAATTYHTTTTTRAGT). The resulting sequences were extracted using seqkit⁷⁴, oriented in the forward direction using vsearch⁷⁵, and then clustered at 100% identity, also with vsearch, to identify unique V4–V5 sequences. These unique sequences were then mapped back to their source genomes to determine which species could produce identical amplicons.

Clinical-outcome analysis

To identify dietary patterns among patients, we applied latent trajectory class analysis⁷⁶ to log-transformed daily intake of the macronutrients fat, fibre, protein and sugar between days –7 and 12 relative to HCT. Day –7 was selected as the starting point as most of the dietary data were collected after day –7 (Fig. 1b); day 12 was selected as the end point as this was the median engraftment day and demarcates the time at which most patients begin to recover. We used cubic natural splines with three nodes (defined at the 10th, 50th and 90th quantiles of the observed timepoints) to model the nonlinear trajectories of macronutrient intake. The optimal number of clusters was determined

using the classification entropy extended quasi-likelihood information criterion. We tested whether there was a significant difference in the caloric intake, as well as macronutrient fraction of the daily diet (as proportion of fat, protein, sugar, fibre and other carbohydrates) between the two patient clusters with a GEE model. The model considered the effect of time (days relative to transplant) using natural splines to flexibly model the potential nonlinear changes in sugar fraction over time. Knots were placed at the 20th and 80th percentiles of the time variable. The model also accounted for the correlation between repeated measurements within each patient, with patient identifier as the clustering variable. The model was fitted using the geeglm function from the geepack (v.1.3.11) package in R⁷⁷. The statistical significance was defined as $P < 0.05$. Next, we incorporated the trajectory cluster labels as a covariate in a multivariable Cox proportional hazards model to assess the association between dietary patterns and mortality end points. The model was adjusted for potential confounders, including the duration of broad-spectrum antibiotic exposure (between day –7 and day 12), conditioning intensity and graft source. Moreover, we included an interaction term between macronutrient intake clusters and antibiotic exposure to evaluate the combined effect of these factors.

Mice

Female C57BL/6 mice 6–8 weeks of age were obtained from Jackson Laboratory, specifically from rooms RB03, RB04, EM07 and RB08, or based on the presence of endogenous *Enterococcus*. The animals were housed under conditions that included a 06:00 to 18:00 light–dark cycle, at a temperature of 72 °F with a set point ± 2 °F and, relative humidity of 30–70%. Mice were single-housed in shoebox-style cages for 1–2 days before each experiment. Baseline faecal samples (day 1) were collected to confirm the presence of endogenous *Enterococcus* before any manipulation and were processed for enumeration and analysis as described below.

Stool collection and colony counting

Faecal pellets were collected in a biosafety hood directly from each mouse into a barcoded preweighed sterile tube and kept on ice until the faecal pellets were weighed and homogenized in 1 ml of PBS and serially diluted. Then, 20 μ l of each dilution was plated onto *Enterococcus*-selective agar (BD, 212205) in six-well plates and incubated at 37 °C under ambient air for 24 h. Colonies from a well with a countable density were manually enumerated.

Antibiotic intervention and diet preparation

We had noticed that the carbapenem antibiotic biapenem tends to facilitate expansion of endogenous enterococci after subcutaneous injection. We therefore established a model in which 2 mg of biapenem is administered subcutaneously to singly housed C57BL/6 female mice sourced from Jackson laboratories. Mice were treated with a single subcutaneous injection of antibiotic (Biapenem (Sigma-Aldrich; SML0306, 2 mg in 100 ml PBS)) or vehicle (PBS). In addition to standard mouse chow, the animals were fed plain hydrogel in cups (HydroGel, ClearH2O, 70-01-502) with or without added sucrose (Sigma-Aldrich, S0389; 5% sucrose supplement: 1 g per 20 ml hydrogel; and control supplement (diet vehicle): 20 ml plain hydrogel, respectively) or 5% glucose (Sigma-Aldrich, G7021), 5% fructose (Sigma-Aldrich, F0127) where indicated. Cups were replenished every 48 h. In some experiments, sucrose was administered in the drinking water (25 g per 500 ml). We also explored delaying the administration of sugar by 1 or 2 days (Extended Data Fig. 8i–k).

Trapezoidal AUC

The trapezoidal AUC was calculated between days 1 and 3, as well as between days 3 and 6, for each mouse in the experimental setting, following the trapezoidal rule. The day-1 raw counts were subtracted

from day 3 and day 6 for each mouse before applying the trapezoidal rule. The reported total trapezoidal AUCs for each treatment group were the sums of the above two separate time periods and they were compared using the Wilcoxon rank-sum test.

Chow consumption

Daily chow consumption in singly housed mice was calculated as the difference between the starting weight and the weight of the chow remaining in the hopper, divided by the number of days between measurements, providing an average daily consumption rate. The trapezoidal AUC was calculated for each mouse over the experimental period to account for repeated measurements. Differences in chow consumption between groups were assessed using the Wilcoxon rank-sum test.

Fibre-free diet experiment

Mice were fed a fibre-free diet (W.F. Fisher and Son; AIN-93G) 5 days before starting the interventions and throughout the whole experiment course, instead of the standard mouse chow used in our institutional vivarium (LabDiet PicoLab Rodent Diet 20, 5053) (diet composition is shown in Supplementary Table 7).

Mouse transplantation model

C57BL/6 (donors) and BALB/c (recipients) mice were procured from The Jackson Laboratory housed at five mice per cage under standard specific-pathogen-free conditions (chow and water ad libitum, 12 h–12 h light–dark cycle). BMT was performed as previously described²⁸. In brief, after split-dose lethal irradiation with 900 cGy, mice received 5×10^6 bone marrow cells that were depleted of T cells (anti-Thy-1.2 (BioXcell) and low-Tox-M rabbit complement (CEDERLANE Laboratories)) through retro-orbital injection. Mice were monitored daily for survival and weekly as described previously²⁸. Then, 3 days after transplant, the mice received 2 mg biapenem or injection vehicle subcutaneously. In addition to food pellets, mice received either control diet supplement of hydrogel (100 ml per cage) with or without 5 g (w/v) sucrose. Grafts were characterized by flow cytometry after staining for surface CD45, CD19, CD3, CD4 and CD8 for 20 min at 4 °C in PBS with 0.5% BSA and counterstaining with DAPI on the LSR-II Fortessa (BD Biosciences) system with FlowJo v.10.10.0 (Tree Star Software).

Monocolonization of germ-free mice

E. faecalis was isolated from a faecal pellet obtained from specific-pathogen-free C57BL/6 mice by plating on *Enterococcus*-selective agar under ambient air at 37 °C and 5% CO₂. A single colony was inoculated into Difco brain–heart infusion (BHI) broth (BD, 237500) and cultured overnight at 37 °C with shaking. A growth curve was generated to determine that optical density at 600 nm of 0.1 corresponded to about 10¹⁰ colony-forming units (CFU) per ml. For subsequent experiments, a fresh overnight culture was prepared under identical conditions, aliquots were frozen at –80 °C in 15% glycerol and thawed immediately before use for oral gavage. In the initial experiment, six germ-free female 8-week-old C57BL/6 mice housed in isocages at the Cornell University Comparative Cancer Biology Center, Belfer Research Building were orally gavaged with 10⁵ CFU *E. faecalis* in sterile PBS. Of these, three mice received 2.5% (w/v) sucrose in their autoclaved drinking water for the duration of the experiment. Four germ-free control mice received PBS gavage alone; two of these also received 2.5% sucrose in drinking water. In the repeat experiment, which did not include germ-free controls, two germ-free mice received *E. faecalis* with regular water and three germ-free mice received *E. faecalis* with 2.5% sucrose in drinking water. Faecal pellets were collected at the baseline (pre-gavage) and at 4, 8, 24, 48 and 144 h after gavage for colonized groups. For germ-free controls, faecal pellets were collected at the baseline and at 4, 48 and 144 h after gavage. Stool was plated and *Enterococcus* colonies were enumerated as described above for the colonized mice. To confirm germ-free status,

stool from germ-free mice was plated onto BHI agar under both anaerobic and aerobic conditions and incubated for 5 days to monitor for bacterial growth.

Isolation of intestinal epithelium and RNA extraction

The small and large intestines were excised, opened longitudinally on a sterile, moistened paper towel and cut into approximately 1 cm segments. Tissue pieces were transferred to a new Petri dish containing PBS and residual mucus was gently removed from both sides using a coverslip (two gentle passes per side). Cleaned intestinal segments were then transferred into a 2 ml Eppendorf tube containing PBS and minced into small fragments. The fragments were moved to a 50 ml Falcon tube with PBS and shaken at 80 rpm for 10 min at 4 °C. The supernatant was carefully removed, and the tissue was incubated in 12 ml of 10 mM EDTA for 20 min at 80 rpm in a cold room. After allowing the fragments to settle, the supernatant was discarded, and 10 ml of PBS was added. The suspension was shaken vigorously by hand (approximately 40 times) to detach epithelial cells. The resulting cell suspension was passed through a 70-µm cell strainer, centrifuged at 350g for 10 min at 4 °C to remove debris and the supernatant was discarded. The epithelial tissue pellet was resuspended in 1 ml of RNeasy lysis buffer (Qiagen, 74104) and stored at 4 °C until RNA extraction. Total RNA was extracted using the RNeasy Mini Kit (Qiagen, 74104) according to the manufacturer's instructions, including on-column DNase treatment to remove genomic DNA contamination.

RNA-seq library preparation and sequencing

For RNA-sequencing (RNA-seq) analyses, the RNA integrity number was assessed using the Agilent Bioanalyzer. Sequencing libraries were prepared and sequenced on an Illumina NovaSeq X to generate 100 bp paired-end reads at the MSKCC Integrated Genomics Operations core facility.

RNA-seq data analysis

Raw sequencing reads (FASTQ) were processed using the Illumina DRAGEN (Dynamic Read Analysis for GENomics) Bio-IT Platform, v.4.2.7. The DRAGEN pipeline was executed in RNA-aware mode (--enable-rna=true) against the mouse reference genome GRCm39; duplicate reads were marked, and the resulting alignments were delivered as coordinate-sorted, indexed BAM files. Gene counts were generated using featureCounts (v.2.1.1) against the GENCODE mouse basic gene annotation, release vM37 (GRCm39; https://www.gencodegenes.org/mouse/release_M37.html). Bioinformatic analysis was performed in R (v.4.4.1). Genes with low expression were removed by prefiltering; genes were retained only if they had a counts-per-million value of greater than 2.5 in at least three samples. Principal component analysis was performed to visualize sample clustering using the DESeq2 package. To ensure homoscedasticity, raw read counts were normalized using variance stabilizing transformation. This transformation was applied in a design-blind manner (blind = TRUE) to provide an unbiased view of sample variance. The principal component analysis coordinates were calculated using the top 500 genes with the highest row variance. The differential expression analysis was performed using the DESeq2 package (v.1.44.0), which incorporates normalization based on size factors, dispersion estimation and fitting a negative binomial generalized linear model. The significance threshold is defined as FDR < 0.05. To correct for the high variance observed in genes with low read counts, log₂[FC] estimates were moderated using the adaptive shrinkage estimator (ashr v.2.2-63) before generating volcano plots. Gene-set enrichment analysis was performed using the fgsea package (v.1.30.0). Genes were preranked using a metric combining the log₂[FC] and *P* value from the DESeq2 results, calculated as: sign(log₂[FC]) × –log₁₀[*P*]. The Hallmark (H) gene set collection (v2025.1.Mm) from the Molecular Signatures Database (MSigDB) was obtained through

Article

the msigdb package (v.25.1.1). Mouse gene symbols were mapped to Entrez IDs as required using the org.Mm.eg.db annotation database (v.3.19.1).

Ethics statement

Participants consented in writing to collection and analysis of biospecimens; specimens were analysed for this study with the approval of the Memorial Sloan Kettering Cancer Center Institutional Review Board under protocol 16-834, which has been previously published²⁶ but is not listed in a clinical trial registry. The study was conducted in accordance with the Declaration of Helsinki. Mouse experiments were conducted under the supervision of the Memorial Sloan Kettering Institutional Animal Care and Use Committee (protocols 23-03-006 and 99-07-025).

Reporting summary

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

Data availability

All sequencing data analysed in this study are publicly available through the NCBI Sequence Read Archive (SRA) across several BioProject accessions, as follows. Accession identifiers for individual samples, together with the sample identifiers used throughout this study, are listed in Supplementary Table 1. Human faecal 16S rRNA amplicon sequencing, 1,009 samples: PRJNA878528 (476 samples), PRJNA545312 (484 samples)²⁷, PRJNA810900 (38 samples), PRJNA972747 (5 samples), PRJNA979931 (5 samples) and PRJNA810891 (1 sample). Human faecal shotgun metagenomic sequencing, 333 samples: PRJNA878528, SRR35981832–SRR35982164. Mouse faecal 16S rRNA amplicon sequencing, 111 samples: PRJNA878528, SRR31624460–SRR31624570. Mouse caecal shallow shotgun metagenomic sequencing, 24 samples: PRJNA878528, SRR37009193–SRR37009216. Mouse intestinal epithelial RNA-seq, 20 libraries: PRJNA878528, large intestine (SRR36928799–SRR36928808) and small intestine (SRR39866831–SRR39866840). Source data are provided for animal experiments (Fig. 4c and Extended Data Figs. 8 and 9) and are also archived in a public Zenodo repository (<https://doi.org/10.5281/ZENODO.14538105>)⁷⁸ as data S7, along with de-identified tables of patient data: meal-level dietary intake records (data S2), stool-sample-level microbiome α -diversity, dietary exposures and clinical covariates (data S3), medication-exposure records (data S4), nutrient profiles of the foods consumed from FNDDS 2015–2016 (data S5), patient-level longitudinal summaries of caloric intake and dietary and faecal microbiota α -diversity (data S6), clinical metadata and processed microbiome feature tables (ASV counts and relative abundances, taxonomic annotation, genus-level relative abundances and shotgun-derived *Enterococcus* species abundances), along with derived analysis inputs and food-taxonomy files. Analysis code for these panels is included in the public repository. This study also used the following public resources: the USDA FNDDS 2015–2016 (<https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndds-download-databases/>); the USDA Food Patterns Equivalents Database (FPED) 2015–2016; the NCBI 16S ribosomal RNA database of Bacteria and Archaea type strains (RefSeq Targeted Loci, BioProject accessions PRJNA33175 and PRJNA33317), downloaded in 2020 and used to assign taxonomy to the human ASVs with BLASTN (BLAST+ v.2.11.0); and the Zymo Research Database to assign taxonomy to the mouse 16S data with Uclust in QIIME v.1.9.1; MetaPhlAn v.4.0.5 with database mpa_vJan21_CHOCOPHiAnSGB_202103; the GENCODE mouse basic gene annotation release vM37 (GRCm39; https://www.gencodegenes.org/mouse/release_M37.html), used for gene-level quantification of the mouse RNA-seq; and the MSigDB Hallmark gene set collection (v2025.1.Mm, 50 gene sets), accessed through the msigdb R package v.25.1.1. Source data are provided with this paper.

Code availability

Code used for the analysis and figure generation is publicly available at GitHub (https://github.com/Anqi-Dai/Nutrition_microbiome_reproducibility) and is archived at Zenodo⁷⁹ (<https://doi.org/10.5281/ZENODO.21290618>). R package versions are pinned with renv, and the external tools required to run the pipeline end to end are specified in the repository README file. The tables distributed with the repository are the same files archived in the Zenodo data deposit⁷⁸. The code is released under an MIT license.

46. USDA Food and Nutrient Database for Dietary Studies 2015–2016 (US Department of Agriculture, Agricultural Research Service, 2018).
47. Faith, D. P. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **61**, 1–10 (1992).
48. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
49. Nguyen, C. L. et al. High-resolution analyses of associations between medications, microbiome, and mortality in cancer patients. *Cell* **186**, 2705–2718 (2023).
50. Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
51. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
52. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
53. Oksanen, J. et al. vegan: community ecology package (2022).
54. Mallick, H. et al. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput. Biol.* **17**, e1009442 (2021).
55. Bush, S. J., Connor, T. R., Peto, T. E. A., Crook, D. W. & Walker, A. S. Evaluation of methods for detecting human reads in microbial sequencing datasets. *Microb. Genom.* **6**, mgen000393 (2020).
56. Blanco-Míguez, A. et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat. Biotechnol.* **41**, 1633–1644 (2023).
57. Gentile, C. L. & Weir, T. L. The gut microbiota at the intersection of diet and human health. *Science* **362**, 776–780 (2018).
58. Taur, Y. et al. The effects of intestinal tract bacterial diversity on mortality following allogeneic hematopoietic stem cell transplantation. *Blood* **124**, 1174–1182 (2014).
59. Rashidi, A. et al. Levaquin gets a pass. *Biol. Blood Marrow Transpl.* **26**, 778–781 (2020).
60. Seo, S. K. et al. Impact of peri-transplant vancomycin and fluoroquinolone administration on rates of bacteremia in allogeneic hematopoietic stem cell transplant (HSCT) recipients: a 12-year single institution study. *J. Infect.* **69**, 341–351 (2014).
61. Moellering, R. C. Jr. Pharmacokinetics of vancomycin. *J. Antimicrob. Chemother.* **14**, 43–52 (1984).
62. Taur, Y. et al. Reconstitution of the gut microbiota of antibiotic-treated patients by autologous fecal microbiota transplant. *Sci. Transl. Med.* **10**, eaap9489 (2018).
63. Button, J. E. et al. Precision modulation of dysbiotic adult microbiomes with a human-milk-derived synbiotic reshapes gut microbial composition and metabolites. *Cell Host Microbe* **31**, 1523–1538 (2023).
64. van den Boogaart, K. G., Tolosana-Delgado, R. & Bren, M. compositions: compositional data analysis. (2023).
65. Bürkner, P.-C. Bayesian item response modeling in R with brms and Stan. *J. Stat. Softw.* **100**, 1–54 (2021).
66. Bürkner, P.-C. Advanced Bayesian multilevel modeling with the r Package brms. *R J.* **10**, 395 (2018).
67. Bürkner, P.-C. brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 1–28 (2017).
68. Stan Development Team. RStan: the R interface to Stan <https://mc-stan.org/> (2022).
69. McElreath, R. *Statistical Rethinking: A Bayesian Course with Examples in R and STAN* <https://doi.org/10.1201/9780429029608> (Chapman and Hall/CRC, 2020).
70. Gelman, A., Hill, J. & Vehtari, A. *Regression and Other Stories* <https://doi.org/10.1017/9781139161879> (Cambridge Aspire, 2020).
71. Wickham, H. *Ggplot2: elegant graphics for data analysis* (Springer, 2016).
72. Kay, M. tidybayes: tidy data and geoms for Bayesian models (2022).
73. Rice, P., Longden, I. & Bleasby, A. EMBOS: the European Molecular Biology Open Software Suite. *Trends Genet.* **16**, 276–277 (2000).
74. Shen, W., Sipos, B. & Zhao, L. SeqKit2: a Swiss army knife for sequence and alignment processing. *iMeta* **3**, e191 (2024).
75. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
76. Hart, K. R., Fei, T. & Hanfelt, J. J. Scalable and robust latent trajectory class analysis using artificial likelihood. *Biometrics* **77**, 1118–1128 (2021).
77. Højsgaard, S., Halekoh, U. & Yan, J. The R package Geepack for generalized estimating equations. *J. Stat. Softw.* **15**, 1–11 (2006).
78. Dai, A. Supplementary data for ‘Sugar-rich foods exacerbate antibiotic-induced microbiome disruption’. *Zenodo* <https://doi.org/10.5281/ZENODO.14538105> (2026).
79. Dai, A. Nutrition_microbiome_reproducibility: reproduction of the public figures of ‘Sugar-rich foods exacerbate antibiotic-induced microbiome disruption’. *Zenodo* <https://doi.org/10.5281/ZENODO.21290618> (2026).

Acknowledgements We thank H. Rice for technical editing assistance. Extended Data Figs. 8 and 9 include art from Flaticon designed by Icongeek26.

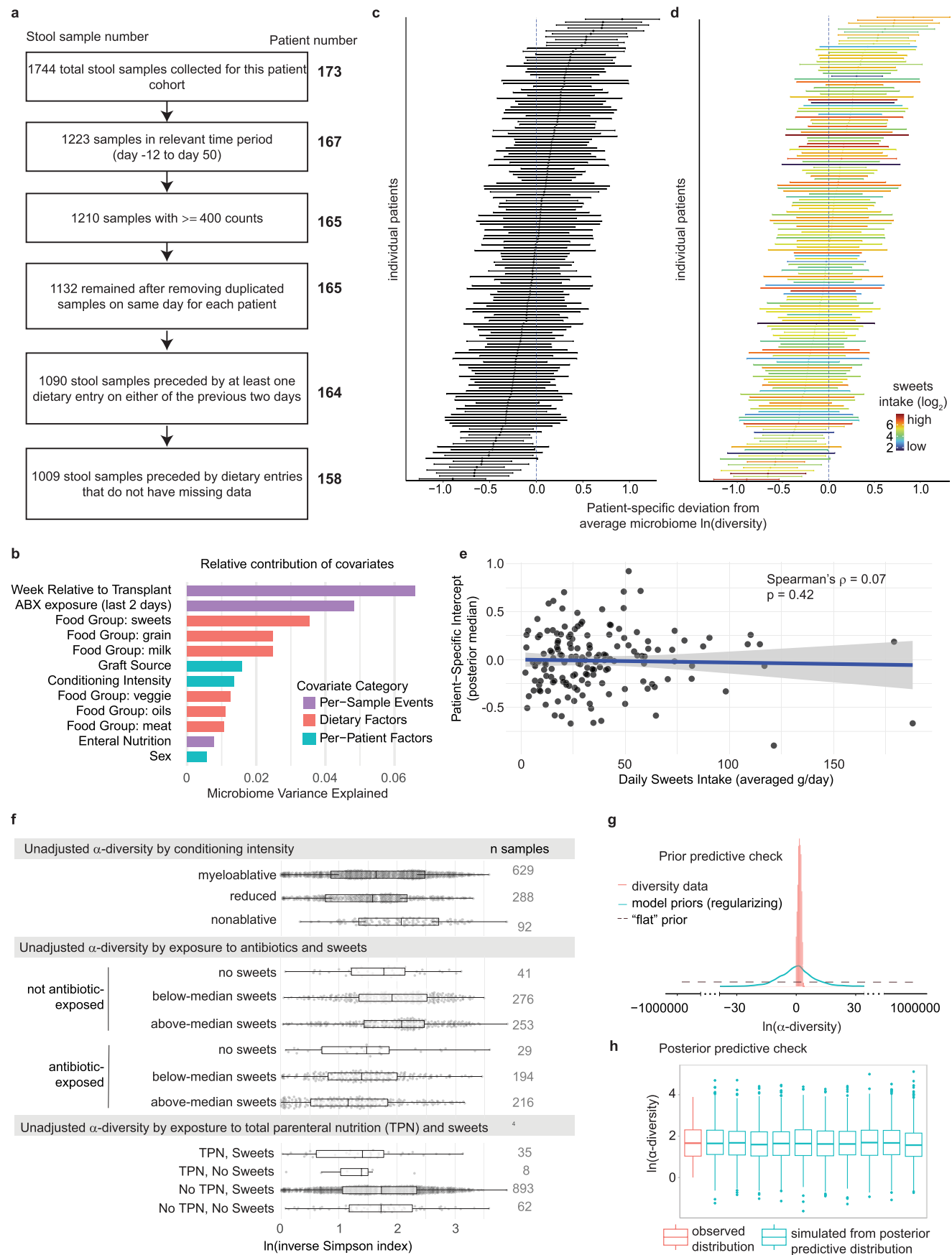
Author contributions W.J., M.R. and P.A.A. contributed equally as joint second authors. M.L.B., R.R.J., K.A.M., J.S., M.R.M.v.d.B. and J.U.P. conceptualized the project. A.D., A.B., M.R., P.A.A., A.S., S.S., B.G., S.S.R., E.H., Z.P., K.W., N.S., J.N., A.W., M.L.B., T.K.P., A.G., L.A.A., J.B.S., A.G.C., D.G.B. and P.A.G. collected data. A.D., W.J., M.B., T. Funnell, T. Fei, N.R.W., E.H., C.D., C.Z., F.M., A.P.S., A.L.C.G., A.J.J., D.K., J.S., M.R.M.v.d.B. and J.U.P. were involved in data analysis. A.D., A.B., W.J., M.R., P.A.A., T. Funnell, T. Fei, N.R.W., B.G., S.S.R., M.B.d.S., C.Z.B., M.L.B., A.P.S., A.L.C.G., A.J.J., D.K., R.R.J., J.S. and J.U.P. contributed to methodology. J.S., M.R.M.v.d.B. and J.U.P. helped with funding acquisition. A.G., J.B.S., M.-A.P., S.A.G., M.R.M.v.d.B. and J.U.P. helped with project administration. J.S., M.R.M.v.d.B. and J.U.P. supervised the project. A.D., J.S. and J.U.P. wrote the original draft. A.B., W.J., M.B., Z.P., K.W., O.M., K.A.M., J.S., M.R.M.v.d.B. and J.U.P. helped with the review and editing of the draft.

Funding J.U.P. was funded by The Susan and Peter Solomon Microbiome, Nutrition and Cancer Program at Memorial Sloan Kettering Cancer Center, NHLBI K08HL143189, and the MSKCC Cancer Center Core Grant NCI P30 CA008748. J.S. and NYU team members were funded by an NIH/NIAID DP2 grant DP2AI164318, R01AI199143, an NCI R01 grant R01CA269617 and an award by the Gabrielle's Angel Foundation and a center grant to the NYU Perlmutter Cancer Center, P30CA0160087. F.M. is supported by a Helen Hay Whitney postdoctoral fellowship. M.R.M.v.d.B. has been supported by the National Cancer Institute (R35CA284024, R01CA228308, P01CA023766, P30CA008748 and P30CA033572), the National Institute on Aging (P01AG052359), Starr Cancer Consortium, the Tri-Institutional Stem Cell Initiative, The Lymphoma Foundation, The Susan and Peter Solomon Divisional Genomics Program, Cycle for Survival and the Parker Institute for Cancer Immunotherapy, with research also supported in part by the Deana and Steve Campbell Chief Physician Executive Distinguished Chair endowment and the Morgan and Helen Chu Presidential Chair endowment at City of Hope. O.M. was supported by the American Society of Clinical Oncology Young Investigator Award, a Hyundai Hope on Wheels Young Investigator Award and a Tow Center for Developmental Oncology Career Development Award. C.Z.B. was supported by an AIRC fellowship for Abroad.

Competing interests J.S. has filed intellectual property applications related to the microbiome (reference numbers US2025/0064882A1, PCT/US2025/023826), serves on an advisory board and holds equity of Jona Health, and is cofounder of Postbiotics Plus Research. J.U.P. reports research funding, intellectual property fees and travel reimbursement from Seres Therapeutics, and consulting fees from DaVolterra, CSL Behring, Crestone, MaaT Pharma and Canaccord Genuity and RA Capital; he serves on an advisory board of and holds equity in Postbiotics Plus Research; he serves on an advisory board of and holds equity in Prodigy Biosciences; he has filed intellectual property applications related to the microbiome

(reference numbers 62/843,849, 62/977,908 and 15/756,845). Memorial Sloan Kettering Cancer Center (MSK) has financial interests relative to Seres Therapeutics. K.A.M. reports previous consulting for Incucyte and Crestone; she serves on an advisory board of and holds equity in Postbiotics Plus Research. M.R.M.v.d.B. has received research support and stock options from Seres Therapeutics and stock options from Notch Therapeutics, Pluto Therapeutics and TymoFox; he has received royalties from Wolters Kluwer; he has consulted, received honorarium from or participated in advisory boards for Seres Therapeutics, Vor Biopharma, WindMIL Therapeutics, Rheos Medicines, Merck, Magenta Therapeutics, Frazier Healthcare Partners, Nektar Therapeutics, Notch Therapeutics, Forty Seven, Priothera, Ceramedix, Lygenesis, Pluto Therapeutics, GlaskoSmithKline, Da Volterra, Garuda, Thymofox, smarNovartis (spouse), SyntheKine (spouse), Beigene (spouse), Kite (spouse), MustangBio (spouse) and Collectar (spouse); he has IP licensing with Seres Therapeutics and Juno Therapeutics; he holds a fiduciary role on the Foundation Board of the non-profit organization DKMS; and he is the chairman of the scientific advisory board for Smart Immune. M.-A.P. reports honoraria from Arog Pharmaceuticals, Bristol-Myers Squibb, Cartesian, Celgene, Incyte, Kite/Gilead, Merck, Miltenyi Biotec, Novartis, Omeros, OrcaBio, Ossium Health, Pierre Fabre, Sanofi and Takeda; he serves on data and safety monitoring boards for Cidara Therapeutics/ Mundipharma and Sellas Life Sciences; he has ownership interests in Omeros and OrcaBio; he has received institutional research support for clinical trials from Allogene, Genmab, Incyte, Kite/Gilead, Miltenyi Biotec, Novartis and Trlx. R.R.J. reports patent royalties and stock options from Seres Therapeutics, and consulting fees from MaaT Pharma and Nestle; he serves on the scientific advisory board of Postbiotics Plus Research. S.A.G. reports research funding from Amgen, Johnson & Johnson, Takeda, Celgene, Actinuum, Sanofi, Miltenyi, Kite, EUSA and Incyte; he serves on advisory boards at Amgen, Bristol-Myers Squibb, Actinuum, Celgene, Johnson & Johnson, Miltenyi, Takeda, Novartis, EUSA Omeros, Sanofi, Kite and Spectrum Pharma. A.L.C.G. reports stock/shares from Seres Therapeutics and equity ownership at Xbiome. M.L.B. reports employment and equity ownership at Savor Health. The other authors declare no competing interests.

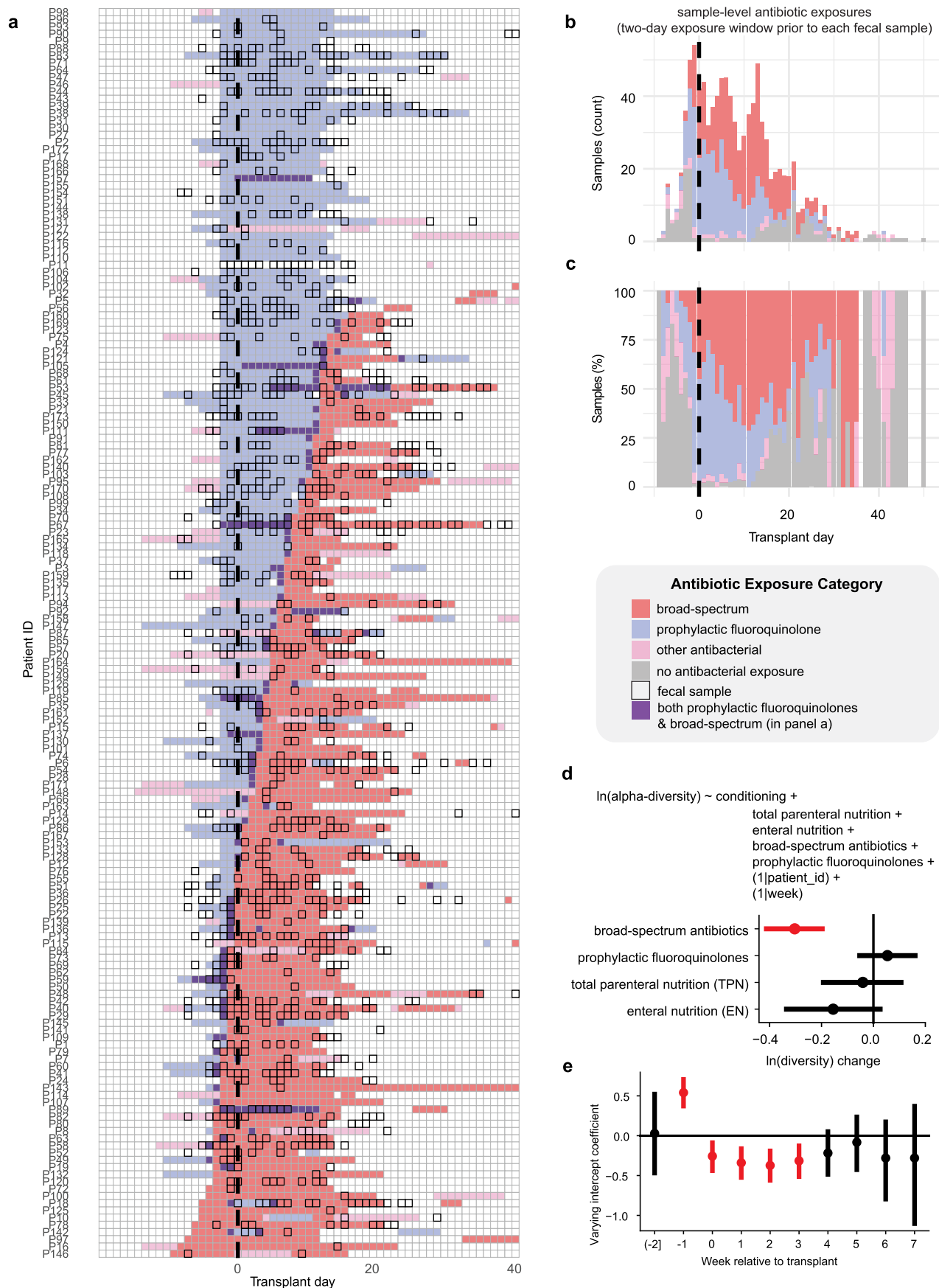
Additional information
Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-026-11077-3>.
Correspondence and requests for materials should be addressed to Marcel R. M. van den Brink, Jonas Schluter or Jonathan U. Peled.
Peer review information *Nature* thanks Andrew Chan, Mahesh Desai, Jeremiah Faith and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Sample selection, covariate contributions, varying intercepts, and Bayesian model checks. (a) Patient and faecal sample selection. (b) Relative explanation of microbiome variation by covariates. Proportion of variance R^2 in ASV-level Bray-Curtis dissimilarity for covariates with statistically significant association ($FDR < 0.05$); four patients (14 samples) were excluded because disease could not be categorized as myeloid or lymphoid. (c) Patient-specific deviations (random intercept posterior) from the average log-microbiome-diversity; each line represents a patient; dots indicate estimated deviation; lines are 95% Bayesian credible intervals. Patients are ordered by estimated deviation; vertical line marks the overall average microbiome diversity, $n = 158$. (d) Same as (c) with lines colour-coded by each patient's averaged sweets intake ($\log_2$ -transformed g/day + 1 pseudocount), $n = 158$. (e) Scatterplot of each patient's averaged sweets intake versus posterior

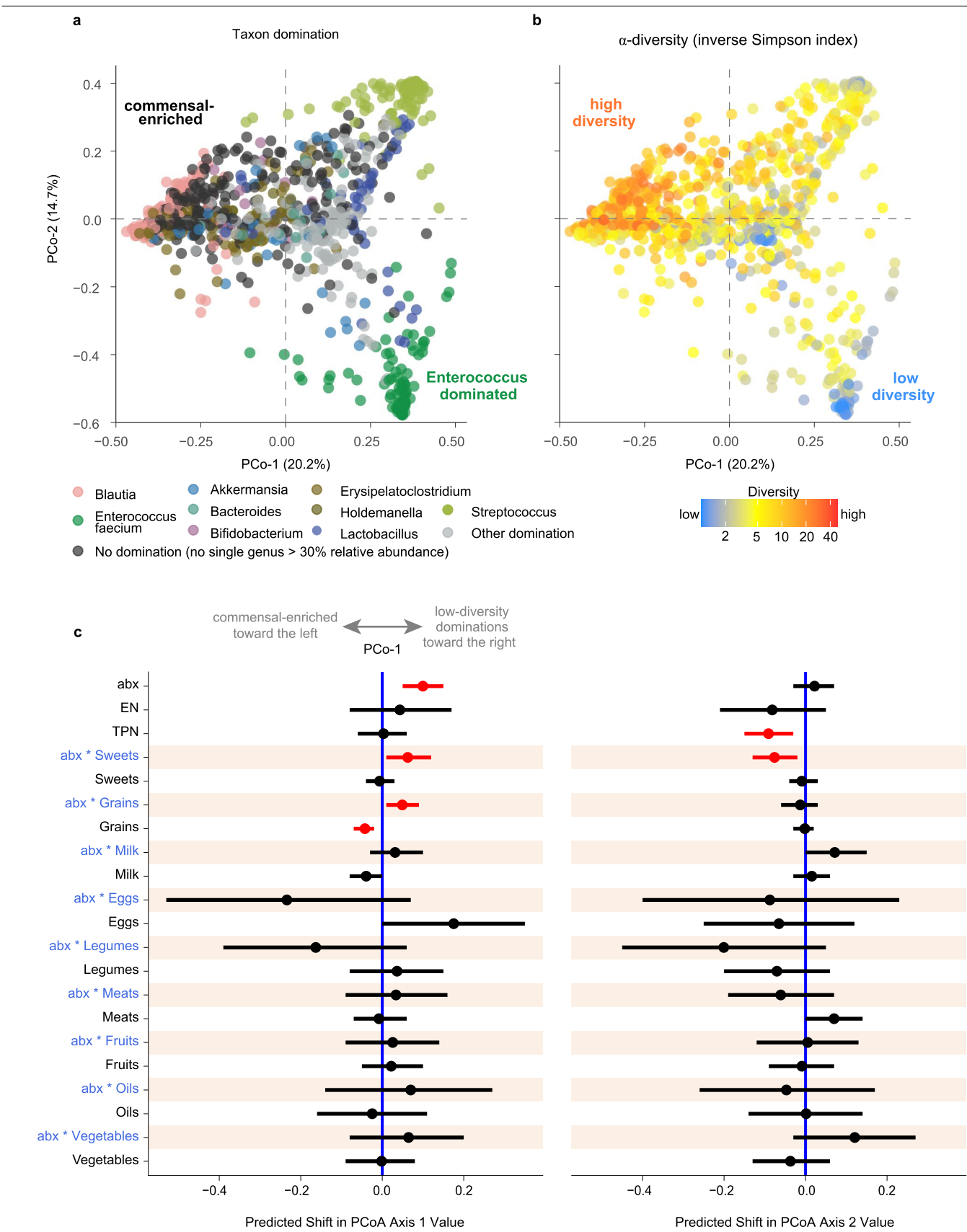
mean patient-specific intercept; line = linear fit; shading = 95% confidence interval. Two-sided Spearman $\rho = 0.065$, $P = 0.42$, $n = 158$. (f) Unadjusted faecal α -diversity stratified as indicated; 11 EN-exposed samples excluded to isolate dietary sweets from liquid nutrition. These raw-data visualizations were not adjusted for the patient-level clinical variables, sample-level exposure other than those indicated, timing of sample collection relative to transplantation, nor for repeated measures from the same patients. (g) Prior predictive check: mildly regularizing priors (teal) constrain the model toward the observed diversity data (red); non-regularizing flat priors (dashed) are shown for illustration. (h) Bayesian posterior predictive check: observed data (red) versus ten datasets independently simulated from the posterior predictive distribution (teal); boxes = interquartile range (IQR), line = median, whiskers = most extreme value within $1.5 \times IQR$.



Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Antibiotic exposure trends in allo-HCT patients. In this cohort, fluoroquinolones (ciprofloxacin or levofloxacin) were administered as prophylaxis against Gram-negative bloodstream infections, typically starting at HCT day -2 or at the onset of neutropenia, whichever was sooner. As patients developed neutropenic fevers, *Clostridioides difficile* diarrhoea, or other signs of potential infection, they were switched to broad-spectrum antibiotics, as classified in this study: piperacillin/tazobactam, carbapenems, cefepime, linezolid, oral vancomycin, and metronidazole. Broad-spectrum antibiotics were typically discontinued upon neutrophil engraftment unless specific infections necessitate longer treatment. **(a)** Individual patient time courses of antibiotic exposures. Days on which faecal samples were collected are outlines in black. **(b)** Histogram displaying the count of stool samples collected during the hospitalizations, colour coded according to antibiotic exposures in the two days preceding sample collection. Any broad-spectrum exposure in this

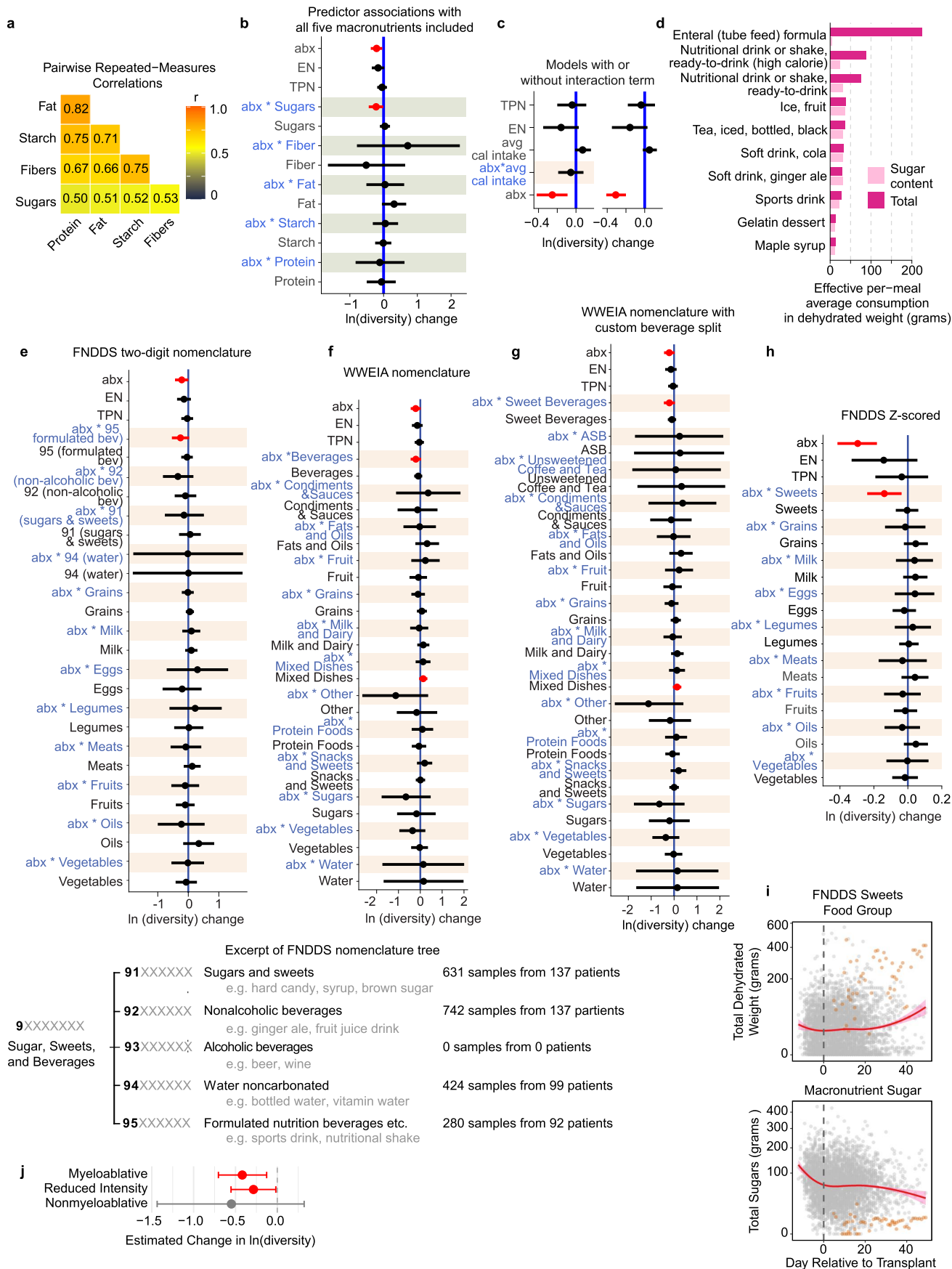
window was coloured red, even if fluoroquinolone exposure also occurred. See Supplementary Table 4. **(c)** Percentage of faecal samples exposed to each medication group on each day following transplantation. The dashed vertical line indicates the day of transplant (day 0). The category "other antibacterial" included systemic (oral or intravenous) but no other routes of administration such as topical or ophthalmic. **(d)** A simplified variant of the main model that explicitly includes a term for exposure to prophylactic fluoroquinolones revealed no strong association with α -diversity independently from those of early hospitalization; **(e)** the varying-intercept term for time (per week relative to transplant) captured α -diversity dynamics early in hospitalization, in the period during which fluoroquinolone exposure is most prevalent. In (d,e): dots = posterior mean; lines = 95% credible intervals (CIs); CIs not crossing zero in red; n = 1,009 stool samples from 158 patients.



Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Dietary predictors of microbiome composition by Bray-Curtis dissimilarity. (a) Principal Coordinates Analysis (PCoA) of microbiota compositional Bray-Curtis dissimilarity. Each point is a microbiome sample, coloured by dominant taxon (> 30% of reads), if present; the nine most common dominating taxa each have their own colour; samples dominated by less-commonly dominating taxa are light grey; non-dominated samples are black. The input was a hybrid feature table summarized at the genus level, except for *Enterococcus*, which was partitioned into ASV_1 (*E. faecium*) and a

second composite feature comprising all other *Enterococcus* ASVs. (b) The same PCoA plot coloured by α -diversity (inverse Simpson index). (c) Bayesian multilevel regression results considering position along PCoA-Axis 1 (left) and Axis 2 (right) as continuous outcomes; posterior mean estimates (dots) and 95% credible intervals (lines) for each predictor; red dots highlight significant predictors where the 95% CI does not cross zero (blue vertical line). n = 1,009 stool samples from 158 patients.

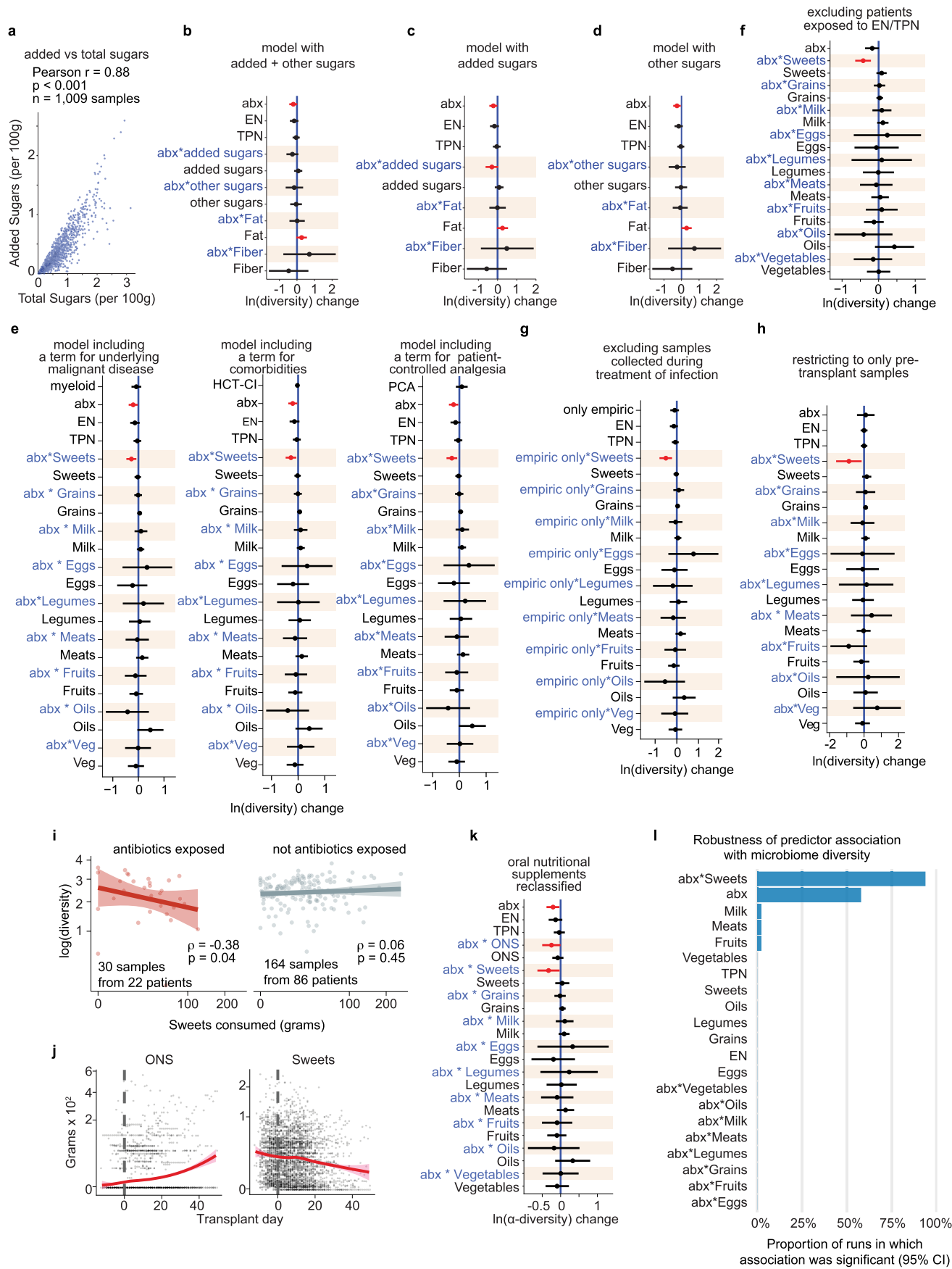


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Macronutrient and food-group robustness checks.

(a) Pairwise repeated-measures correlation matrix of daily macronutrient intake; numbers are coefficients of repeated-measures correlations from 173 patients. (b) Bayesian model of all five major macronutrients and α -diversity including clinical confounders. (c) Microbiome α -diversity is not predicted by total caloric intake in the two days prior to sampling; two model versions, with or without a caloric-intake * abx interaction (interaction terms in blue). (d) Effective per-meal average consumption of the top ten foods in “sugars, sweets, and beverages” group (referred to as “Sweets”). Dark pink = dehydrated weight; light pink = sugar content. “Nutritional drink or shake” includes oral nutritional supplements; “Sports drink” includes, e.g., Gatorade. Bars are the mean per-meal amount for each food across every meal in which it was recorded among n = 173 patients. (e) Alternative version of the model in Fig. 2d with FNDDS “Sweets” replaced by its higher-resolution two-digit subgroups. (f) As an alternative to the FNDDS nomenclature, dietary data were classified according to “What We Eat In America” (WWEIA) categories. FNDDS “Sweets”

beverages map to the WWEIA “Beverages” category; the beverages-antibiotic interaction term was a significant predictor of diversity. (g) WWEIA with custom sub-categories for beverages. (h) To ascertain that the use of dehydrated food weights as predictors did not inflate the estimated effect of sweet beverages, z-scoring the food categories individually rendered them unitless. (i) As patients were discharged from the hospital, the dataset became more sparse and enriched for sicker patients with longer hospital courses, including those admitted to intensive care and fed via gastric tubes, which are classified as FNDDS “Sweets”, accounting for up-trending “Sweets,” but contain relatively lower sugar content, accounting for down-trending sugar intake; n = 173 (j) Inferred sweets-antibiotic interaction coefficients stratified by conditioning intensity (n = 629 myeloablative, 288 reduced-intensity, 92 nonmyeloablative samples). In (b, c, e-h, j) dots = posterior mean; lines = 95% credible intervals (CIs); CIs not crossing zero in red. n = 1,009 stool samples from 158 patients.



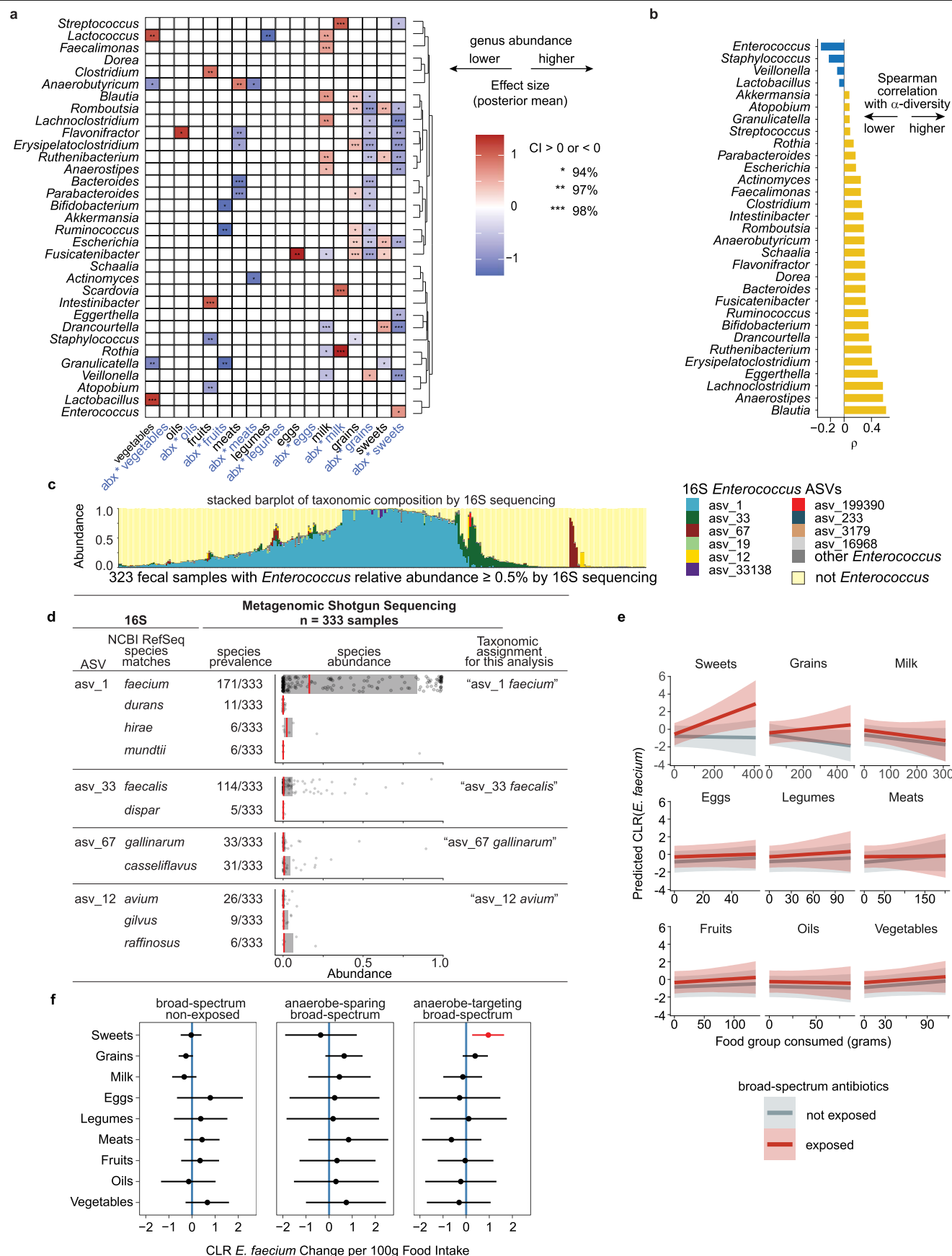
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Robustness of the sweets-antibiotic-diversity association across sensitivity analyses. (a) Correlation between “added” vs “total” sugars, two-sided Pearson $r = 0.88$, $P < 2.2 \times 10^{-16}$; $n = 1,009$ samples from 158 patients). (b) Bayesian inference on macronutrient data separating sugars into “added” vs “other.” (c) model including only “added sugars”; (d) model including only “other” sugars. (e) The sweets*antibiotic interaction remained a significant predictor of microbiome diversity in models adjusting for clinical covariates: left, cancer type (myeloid vs. lymphoid); centre, hematopoietic-cell-transplantation comorbidity index (HCT-CI); right, patient-controlled analgesia (PCA) as a proxy for mucositis severity (52 samples, 5.2% exposed). (f) Subset excluding 139 samples exposed to EN or TPN. (g) Subset analysis excluding samples collected during treatment of documented bloodstream infection or *C. difficile* diarrhoea; analysed are samples exposed to empiric broad-spectrum antibiotics (empiric only, 247 samples) and without any broad-spectrum exposure (570 samples), see Online Methods. (h) Subset restricted to 194 pre-transplant samples. (i) Correlation between Sweets and microbiome α -diversity in pre-transplant samples, stratified by antibiotic

exposure: 30 antibiotic-exposed samples from 22 patients (Spearman $R = -0.38$, two-sided $p = 0.04$) versus 164 unexposed samples from 86 patients (Spearman $R = 0.06$, two-sided $p = 0.45$); lines = linear regression with 95% confidence interval. (j) Time courses of ONS (oral nutritional supplements: nutritional shakes, drinks, powders, and tube feeds) versus non-ONS Sweets across the study period (907 samples after non-ONS Sweets; 155 after ONS). (k) Reclassification placing ONS into a separate group reveals that both ONS and non-ONS Sweets consumed during antibiotic exposure are each associated with reduced microbiome diversity. (l) Sensitivity analysis on 50 non-overlapping random subsets of samples; bars = frequency with which each model term significantly predicted microbiome diversity; the $\text{abx} * \text{Sweets}$ interaction was significant in 94 % of runs. In (b-h, k) dots are the posterior mean of each coefficient and horizontal lines the 95% credible interval; credible intervals that exclude zero are shown in red. $n = 1,009$ stool samples from $N = 158$ patients for (b-d, k); (e) 995 samples from 154 patients (left); 1,003 samples from 155 patients (centre); 1,009 samples from 158 patients (right); (f) 870 samples from 139 patients; (g) 817 samples from 147 patients; (h) 194 samples from 99 patients.

Extended Data Fig. 6 | Dietary patterns and antibiotic exposure predict mortality following allo-HCT. (a-b) Bayesian model of microbiome diversity with a split of the FNDDS “Grains” category into custom groupings in (a) “Sweet Grains” that contain >1.5 teaspoon-equivalents of added sugar per 100 g and “Other Grains” that contain ≤1.5 tsp equivalents of added sugar per 100 g or in (b) into “Sweet Grains” that contain >10 g macronutrient sugar per 100 g. In (a, b) dots, lines indicate the posterior means, and 95% CIs, respectively; red: 95% CIs not crossing zeros. n = 1,009 stool samples from 158 patients. (c) Sugar and overall calorie intake were highly correlated (Spearman R = 0.77, two-sided $P < 2.2 \times 10^{-16}$, n = 173 patients). Each point is a patient’s intake averaged across HCT days -7 to 12; linear regression in red. (d) Calorie-adjusted sugar intake is anti-correlated with overall calorie intake (two-sided Spearman R = -0.62, $P < 2.2 \times 10^{-16}$, n = 173 patients). (e-g) Latent trajectory class analysis identified two distinct dietary patterns of food intake between day -7 to 12; survival probabilities of the two clusters were not significantly different in univariable analysis (logrank p = 0.088, landmarked at HCT day 12). Dietary patterns with respect to (e) calories, (f) sugar, and (g) other macronutrients. In e-g, the centre marks are the median, the box bounds are the IQR and the whiskers extend to the most extreme value within $1.5 \times \text{IQR}$ of the nearest box bound; values beyond the whiskers are not plotted. The red lines join the daily medians.

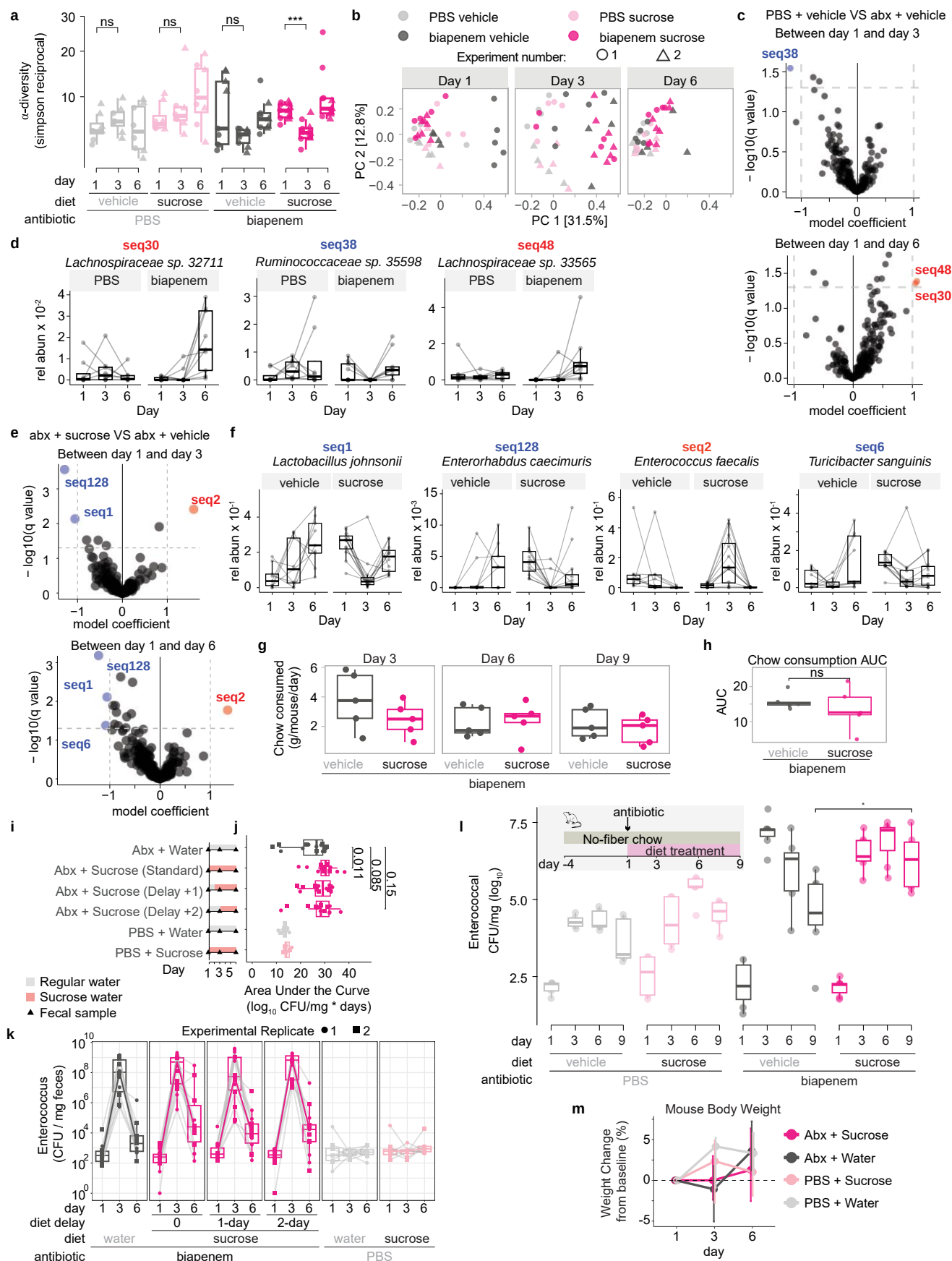
Clusters were compared by two-sided Wald tests from generalized estimating equations, with an exchangeable working correlation, modelling transplant day as a natural spline: calories, $P < 2.2 \times 10^{-16}$; sugar $P = 1.1 \times 10^{-11}$; other carbohydrate $P = 0.0089$; fat, $P = 1.6 \times 10^{-12}$; protein, $P = 4.7 \times 10^{-9}$. n = 173 patients. (h) Faecal microbiota α -diversity in the clusters; cluster 1 trend was modestly but significantly lower than cluster 2. Two-sided Wald test from a generalized estimating equation, $P = 0.0022$; n = 1,009 samples from 158 patients. (i) Cumulative incidence of hospital discharge following neutrophil engraftment in a Cox model adjusted for duration of broad-spectrum antibiotic exposure (day -7 to 12), graft source, conditioning intensity, and GVHD prophylaxis, adjusted HR 1.54 (95% confidence interval 1.06–2.23), two-sided Wald $P = 0.023$; n = 168. (j) Predicted mortality-risk surface from a Cox proportional hazards model as a joint function of average daily sugar intake per 1,000 kcal during HCT day -7 to 12; y-axis) and duration of broad-spectrum antibiotic exposure duration (in days, x-axis) adjusted for graft source and conditioning intensity. grey dashed curve represents the combinations of antibiotic exposure and energy-adjusted average sugar intake that correspond to a mortality hazard ratio of 1. Shades of red and blue indicate higher and lower mortality risk as compared to the light grey dashed curve, respectively. Marginal histograms depict the distribution of patient numbers across the cohort.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Microbiome composition, taxa and diversity associated with food groups. (a) Heatmap of associations between food-group and antibiotic exposures and CLR-transformed genus relative abundances from Bayesian models; posterior coefficient means are shown for significant associations (red = positive, blue = negative; asterisks denote posterior CI exclusion of zero at 94%/97%/98%). Genera observed in >10% of samples at >0.2% relative abundance are included. (b) Correlations between faecal microbiome α -diversity (inverse Simpson index) and genera present in >10% of samples with >0.2% relative abundance; Benjamini-Hochberg adjusted Spearman two-sided $p < 0.05$ in either direction. Blue and yellow represent negative and positive correlations between relative abundances and α -diversity, respectively; $n = 1,009$ samples from 158 patients. (c) Shotgun metagenomics-based species annotation of dominant *Enterococcus* 16S amplicon sequence variants (ASVs). Compositional bar plot of 323 faecal samples with *Enterococcus* relative abundance $\geq 0.5\%$ by 16S; bars coloured by the 10 most abundant *Enterococcus* ASVs; yellow indicates non-*Enterococcus* organisms. ASV_1

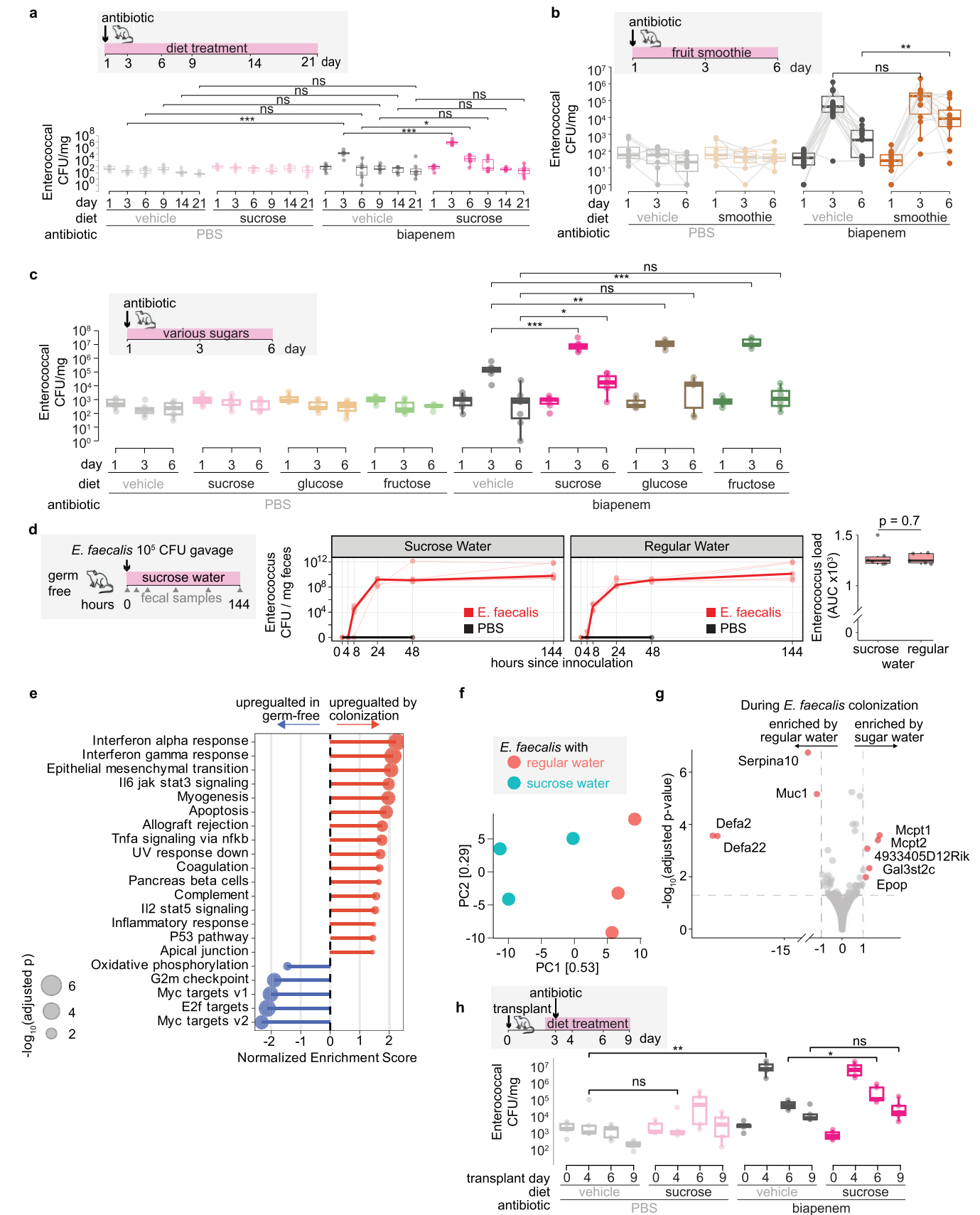
(putatively *E. faecium*) and ASV_33 (putatively *E. faecalis*) together accounted for a median of 96.2% of all *Enterococcus* reads, with ASV_1 by far the most prevalent and frequently dominating. (d) Matching V4-V5 amplicons were extracted in silico from *Enterococcus* genomes in RefSeq. Left: potential species matches for four ASVs. Middle: species prevalence and abundance in the matched shotgun metagenomes ($n = 333$ samples); *E. faecalis* and *E. faecium* can be unambiguously resolved using these V4-V5 primers. Right: taxonomic assignment used throughout this analysis (ASV_1 = *E. faecium*; ASV_33 = *E. faecalis*). (e) Posterior counterfactual prediction of *E. faecium* CLR-transformed abundances across varying food-group intake, holding other variables constant; lines = posterior median; shading = 95% interval; antibiotic-exposed (red) and unexposed (grey). (f) Posterior associations between *E. faecium* CLR and food-group intake under three antibiotic-exposure categories: broad-spectrum non-exposed (left), anaerobe-sparing broad-spectrum (middle), and anaerobe-targeting broad-spectrum (right). Dots = posterior mean; lines = 95% credible intervals; CIs not crossing zero in red.



Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Effects of antibiotic and sucrose treatments on the mouse gut microbiome. (a) α -diversity (inverse Simpson) across two independent experiments (triangles vs. circles); two-sided paired Wilcoxon signed-rank test, not adjusted for multiple comparisons: PBS + vehicle $P = 0.25$, PBS + sucrose $P = 0.20$, biapenem + vehicle $P = 0.30$, biapenem + sucrose $P = 4.9 \times 10^{-4}$, ($n = 37$ mice from 2 independent experiments). (b) PCoA of Bray-Curtis distances faceted by day; on day 6, sucrose vs. vehicle differed among the antibiotic-treated animals (PERMANOVA on Bray-Curtis distances, diet term adjusted for experiment: $P < 1 \times 10^{-4}$); $n = 9$ –12/group. (c, e) Volcano plots: (c) PBS+vehicle vs. biapenem+vehicle and (e) biapenem+sucrose vs. biapenem+vehicle, day 1 vs. 3 (top) and day 1 vs. 6 (bottom); $x = \text{MaAsLin2 coefficient}$; $y = -\log_{10}(q)$; horizontal dashed = FDR $q = 0.05$; vertical = $|\text{coefficient}| > 1$; red enriched, blue depleted. (d) Biapenem-enriched taxa, $n = 8$ –12/group (f) Sucrose-enriched taxa following biapenem treatment. (g) Daily chow

consumption (g/mouse/day) on days 3, 6, 9 in biapenem-treated mice with vehicle vs. sucrose; $n = 5$ /group, in a single experiment. Two-sided Wilcoxon rank-sum, vehicle vs sucrose: $P = 0.31$ (day 3), $P = 0.84$ (day 6), $P = 0.55$ (day 9) (h) AUC of data from panel g: two-sided Wilcoxon rank-sum $P = 0.55$. (i–k) Delayed sugar administration (0, 1, or 2 days after antibiotic): (i) schematic, (j) AUC of *Enterococcus* abundance, (k) raw data of *Enterococcus* CFU/mg faeces. $n = 4$ –7/group 2 experiments. (l) Mice on fibre-free chow (Supplementary Table 7) starting five days before treatment; 0% fibre alone permits a modest *Enterococcus* bloom, but sucrose still prolonged antibiotic-induced expansion through day 9 (two-sided t-test; *: $p = 0.019$; $N = 24$ mice, 5–7 per group in a single experiment). (m) mouse body weight, $n = 18$ –31/group, error bars span IQR. box plots throughout: boxes = IQR; line = median; whiskers = most extreme value within $1.5 \times \text{IQR}$. The mouse icon in l was designed by lcongeek26, Flaticon (Freepik Company).



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Mouse studies of sugar modulation of *Enterococcus* dynamics. (a) *Enterococcus* CFU/mg faeces following biapenem or PBS with sucrose or vehicle for a prolonged 21-day time course. Two-sided Wilcoxon rank-sum, biapenem + vehicle vs biapenem + sucrose: $P = 5.8 \times 10^{-4}$ (day 3), $P = 0.026$ (day 6), $P = 0.54$ (day 9), $P = 0.80$ (day 14), $P = 0.80$ (day 21); biapenem + vehicle vs PBS + vehicle: $P = 5.8 \times 10^{-4}$ (day 3), $P = 0.62$ (day 6), $P = 0.71$ (day 9), $P = 0.16$ (day 14), $P = 0.38$ (day 21), $n = 7$ in a single experiment. (b) Commercial fruit smoothie (mixed-sugar exposure commonly consumed by patients) or vehicle, following PBS or biapenem; $n = 8$ –19/group pooled across 3 experiments, two-sided Wilcoxon rank-sum: biapenem + vehicle vs biapenem + smoothie: $P = 0.63$ (day 3) and $P = 0.0013$ (day 6). (c) *Enterococcus* CFU/mg faeces in mice treated with biapenem with different sugars or vehicle. two-sided Wilcoxon rank-sum vs biapenem + vehicle at day 3: sucrose $P = 5.8 \times 10^{-4}$, glucose $P = 1.2 \times 10^{-3}$, fructose $P = 5.8 \times 10^{-4}$; at day 6: sucrose $P = 0.026$, glucose $P = 0.29$, fructose $P = 0.45$. $n = 7$ /group in a single experiment. (d) Germ-free C57BL/6 mice inoculated with 10^5 CFU *E. faecalis* (isolated from an Jax C57BL/6 specific-pathogen-free mouse, speciated on a Bruker MALDI-TOF Biotyper) or PBS, with regular of sucrose water (2.5% w/v) $n = 15$ mice across 2 experiments, with $n = 5$ –6/group in mono-colonized groups and $n = 2$ in germfree controls. Germ-free status was verified by plating faeces throughout. *Centre panel*: each thin line is an individual mouse; thick lines connect group medians. *Right panel*: In AUC boxplots the centre line is the median, the box bounds are the 25th and

75th percentiles (interquartile range, IQR), and the whiskers extend to the most extreme observation within $1.5 \times \text{IQR}$. two-way ANOVA with experimental replicate as a covariate, $P = 0.7$. (e–g) Host transcriptional response (bulk RNA-seq, colonic epithelium) (e) GSEA Hallmark pathways enriched in *E. faecalis*-colonized vs. uncolonized animals, including interferon- α - γ ; $n = 3$ colonized vs. 2 PBS-gavaged, with regular water; point size = $-\log_{10}(\text{adjusted } p)$. (f) PCA of colonic epithelial gene expression in colonized mice with regular vs. sucrose water; PC1 separated the groups (two-sided Welch t-test, $P = 0.045$, $n = 3$ /group). (g) Volcano plot of gene transcripts enriched in regular- vs. sucrose-water fed animals during *E. faecalis* colonization; adjusted $p < 0.05$ and $|\log_2\text{-fold change}| > 1$ in red; $n = 3$ /group in a single experiment. (h) *Enterococcus* CFU/mg faeces in BALB/c mice transplanted with C57BL/6 T-cell-depleted bone marrow; three days post-transplant they received biapenem or PBS with sucrose or vehicle; two-sided Wilcoxon rank-sum: PBS + vehicle vs PBS + sucrose $P = 0.55$ and PBS + vehicle vs biapenem + vehicle $P = 0.0079$ at day 4; biapenem + vehicle vs biapenem + sucrose $P = 0.032$ at day 6 and $P = 0.55$ at day 9. $n = 5$ /group. In all box plots (a, b, c, d right-hand panel, and h), the centre line is the median, the box bounds are the 25th and 75th percentiles (interquartile range, IQR), and the whiskers extend to the most extreme observation within $1.5 \times \text{IQR}$ of the nearest box bound. The mouse icon in a–d and h was designed by Icongeek26, Flaticon (Freepik Company).

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

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

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

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Meal-tray records were captured in the hospital kitchen commercial computer system (Computriton, Bedford, MA), which is linked to the recipe of each item; no version is published for this system. No custom software was used for data collection.

Data analysis

Analyses were performed in R (v4.4.1).

Human 16S and shotgun sequencing: DADA2 v1.16.0; BLAST+ v2.11.0 against the NCBI 16S ribosomal RNA reference database (Bacteria and Archaea, downloaded 2020); QIIME 2 v2021.11; vegan v2.5-7; MetaPhlAn v4.0.5 (database mpa_vJan21_CHOCOPHiAnSGB_202103), run through the vdblab-shotgun pipeline (<https://github.com/vdblab/vdblab-shotgun/>) with clumpify and bbduk from the BBTools package.

Mouse 16S sequencing: DADA2 and UCLUST as implemented in QIIME v1.9.1 within the ZymoBIOMICS Targeted Sequencing Service; QIIME v1.9.1 for alpha diversity; vegan v2.5-7 for principal coordinates analysis; MaAsLin2 v1.24.1 for differential abundance.

Dietary data: food tree built with the Food_Tree pipeline (https://github.com/knights-lab/Food_Tree); Faith's phylogenetic diversity computed in QIIME 2 v2021.11; TaxUMAP (<https://github.com/jsevo/taxumap>).

Enterococcus amplicon specificity: the SYTYCS pipeline (<https://github.com/nickp60/sytycs>), using EMBOS PrimerSearch, seqkit and vsearch.

Statistical modeling: brms v2.16.3 with rstan v2.26.4 for the Bayesian multilevel models; compositions v2.0-6 for the centered-log-ratio transformation; geepack v1.3.9 (dietary and microbiome trend analyses) and v1.3.11 (clinical-outcome analyses) for generalized estimating equations.

RNA-seq: Illumina DRAGEN Bio-IT Platform v3.10; featureCounts v2.1.1; DESeq2 v1.44.0; ashR v2.2-63 for log2 fold-change shrinkage; fgsea v1.30.0 with the MSigDB Hallmark collection v2025.1.Mm via msigdb v25.1.1; org.Mm.eg.db v3.19.1.

Flow cytometry: FlowJo v10.10.0 (Tree Star).

Visualization: ggplot2 v3.3.5, tidybayes v3.0.2, ggpubr v0.4.0.

Code used for the analysis and figure generation is publicly available at https://github.com/Anqi-Dai/Nutrition_microbiome_reproducibility and is archived at Zenodo (<https://doi.org/10.5281/zenodo.21290618>). R package versions are pinned with renv, and the external tools required to run the pipeline end to end are specified in the repository README. The tables distributed with the repository are the same files archived in the Zenodo data deposit⁷⁸. The code is released under an MIT license.

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

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All sequencing data analyzed in this study are publicly available through the NCBI Sequence Read Archive (SRA) across several BioProjects, as follows. Accession identifiers for individual samples, together with the sample identifiers used throughout this study, are listed in Data S1. Human fecal 16S rRNA amplicon sequencing, 1,009 samples: PRJNA878528 (476 samples), PRJNA545312 (484 samples), PRJNA810900 (38 samples), PRJNA972747 (5 samples), PRJNA979931 (5 samples) and PRJNA810891 (1 sample). Human fecal shotgun metagenomic sequencing, 333 samples: PRJNA878528, SRR35981832–SRR35982164. Mouse fecal 16S rRNA amplicon sequencing, 111 samples: PRJNA878528, SRR31624460–SRR31624570. Mouse cecal shallow shotgun metagenomic sequencing, 24 samples: PRJNA878528, SRR37009193–SRR37009216. Mouse intestinal epithelial RNA sequencing, 20 libraries: PRJNA878528, large intestine SRR36928799–SRR36928808 and small intestine SRR39866831–SRR39866840.

Source data are provided for animal experiments (Fig. 4c and Extended Data Figs. 8 and 9) as a spreadsheet workbook file herein and are also archived in a public Zenodo repository (<https://doi.org/10.5281/zenodo.14538105>)⁷⁸ as Data S7, along with de-identified tables of patient data: meal-level dietary intake records (Data S2), stool-sample-level microbiome α -diversity, dietary exposures and clinical covariates (Data S3), medication-exposure records (Data S4), nutrient profiles of the foods consumed from FNDSS 2015–2016 (Data S5), patient-level longitudinal summaries of caloric intake and dietary and fecal microbiota α -diversity (Data S6), clinical metadata, and processed microbiome feature tables (ASV counts and relative abundances, taxonomic annotation, genus-level relative abundances, and shotgun-derived Enterococcus species abundances), along with derived analysis inputs and food-taxonomy files. Analysis code for these panels is included in the public repository.

This study also used the following public resources: the USDA Food and Nutrient Database for Dietary Studies (FNDSS) 2015–2016 (<https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/fndss-download-databases/>); the USDA Food Patterns Equivalents Database (FPED) 2015–2016; the NCBI 16S ribosomal RNA database of Bacteria and Archaea type strains (RefSeq Targeted Loci, BioProjects PRJNA33175 and PRJNA33317; at <https://ftp.ncbi.nlm.nih.gov/blast/db/>), downloaded in 2020 and used to assign taxonomy to the human amplicon sequence variants with blastn (BLAST+ v2.11.0); and the Zymo Research Database to assign taxonomy to the mouse 16S data with Uclust in QIIME v1.9.1; MetaPhlAn v4.0.5 with database mpa_vJan21_CHOCOPhlanSGB_202103; the GENCODE mouse basic gene annotation release vM37 (GRCm39; https://www.gencodegenes.org/mouse/release_M37.html), used for gene-level quantification of the mouse RNA sequencing; and the MSigDB Hallmark gene set collection (v2025.1.Mm, 50 gene sets), accessed through the msigdb R package v25.1.1.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	Supplementary Table 1
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	Supplementary Table 1
Recruitment	In this cohort study conducted at Memorial Sloan Kettering Cancer Center in New York between 2017 and 2022, eligible participants had been admitted for an allogeneic hematopoietic cell transplantation. Participants consented in writing to collection and analysis of biospecimens; specimens were analyzed for this study with the approval of the Memorial Sloan Kettering Cancer Center Institutional Review Board under protocol #16-834, which has been previously published (PMID: 32101664) but is not listed in a clinical trial registry. The study was conducted in accordance with the Declaration of Helsinki. Thrice weekly, patients who had been recently admitted for allo-HCT were invited to participate in dietary-intake collection.

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

Field-specific reporting

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

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

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

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

Sample size	For human studies, eligible patients were invited to participate over a 5-year period (2017-2022) during which staffing and resources were available; the cohort size was the number of patients who agreed to participate during that period; a formal power calculation was not done. For mouse studies, the minimum number of animals for statistical significance were used based on preliminary data; no statistical method was used to formally predetermine sample size.
Data exclusions	Detailed in Extended Fig. E1a
Replication	Experiments were repeated at least twice as successful, independent experiments. Details as to the exact number of replicates are included in the Figure Legends.
Randomization	For antibiotic and diet experiments under conventional (specific-pathogen-free) conditions, experiments were conducted in singly housed mice. Batches of vendor-supplied mice were separated into individual cages and assigned randomly to treatment arms. For bone marrow transplants, mice were segregated into treatment groups by cage, 4-5 mice per cage, randomly, except with care that mice at the extremes of body mass were not allocated to the same treatment cages. For gnotobiotic experiments, germ-free mice were randomized to treatment arm and co-housed by treatment assignment in gnotobiotic isolator cages.
Blinding	Blinding was not done in mouse experiments as the outcomes were objective quantitative measures (e.g., abundance of Enterococcal colony forming units, next-generation sequencing data). Additionally, treatment-arm allocations, treatments, and processing of samples for a given experiment were often handled by the same individual, so blinding was not logistically practical.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	C57BL/6 mice, aged 6–8 weeks. Animals were housed under conditions including 6:00 am to 6:00 pm light/ dark cycle, at a temperature of 72 degrees Fahrenheit with a set point +/- 2 degrees Fahrenheit and relative humidity of 30-70%.
Wild animals	No wild animals were used in the study.
Reporting on sex	Females
Field-collected samples	No field-collected samples were used in the study.

Ethics oversight

Memorial Sloan Kettering Institutional Animal Care and Use Committee (protocols #23-03-006 and #99-07-025).

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

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

NA

Study protocol

In this cohort study of adult allo-HCT recipients at Memorial Sloan Kettering Cancer Center in New York between 2017 and 2022, eligible participants had been admitted for an allogeneic hematopoietic cell transplantation. Participants consented in writing to collection and analysis of biospecimens; specimens were analyzed for this study with the approval of the Memorial Sloan Kettering Cancer Center Institutional Review Board under protocol #16-834, which has been previously published (PMID: 32101664) but is not listed in a clinical trial registry. The study was conducted in accordance with the Declaration of Helsinki.

Data collection

New York, 2017-2022

Outcomes

NA

Plants

Seed stocks

NA

Novel plant genotypes

NA

Authentication

NA