The gut microbiota influences circulatory immune cell dynamics in humans

Jonas Schluter[1], Jonathan U. Peled[2,3], Bradford P. Taylor[1], Melody Smith[2,3], Ying Taur[4], Sejal Morjaria[4], Maly Fenelus[5], Melissa S. Pessin[5], Tobias M. Hohl[4], Miguel-Angel Perales[2,3], Marcel R.M. van den Brink[2,3], Joao B. Xavier[1]

1) Computational and Systems Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA

2) Adult Bone Marrow Transplantation Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

3) Weill Cornell Medical College, New York, NY, USA

4) Infectious Disease Service, Department of Medicine, and Immunology Program, Sloan Kettering Institute, New York, NY, USA.

5) Department of Laboratory Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Correspondence: schluetj@mskcc.org, xavierj@mskcc.org

ABSTRACT

The gut microbiota influences the development and homeostasis of the mammalian immune system[1–3] and can alter immune cell compositions in mice[4–7]. However, our understanding of how the microbiota modulates immunity remains limited, particularly in humans where a lack of manipulative experiments makes inference challenging. Here we overcome this challenge by studying hundreds of hospitalized—and closely monitored—bone marrow transplantation patients as they recover from chemotherapy-induced immune ablation. This aggressive treatment causes large shifts in both white blood cell and microbiota populations allowing the relationships between the two to be studied simultaneously over time, with unprecedented resolution. Our analysis shows that the ecological diversity of the gut microbiota had an immunosuppressive effect on circulating lymphocyte counts similar in magnitude to that of anti-inflammatory and immunosuppressive drugs administered to cancer patients. Moreover, by controlling for drug treatments and clinical metadata, we identified several microbiota components strongly associated with white blood cell dynamics: Streptococcaceae associated with lymphocyte increase and Actinomycetaceae with lymphocyte suppression, Prevotellaceae with monocyte increase and Enterobacteriaceae with monocyte suppression, Ruminococcaceae with neutrophil increase and Lachnospiraceae and Bacteroidaceae with neutrophil suppression. Our analysis establishes a direct link between the intestinal microbiota and systemic immunity in humans, with implications on clinical outcomes and microbiota-based therapies.

35 **MAIN TEXT**

36 Experiments in mice provide mounting evidence that the mammalian intestinal microbiome
37 influences the development[1–3] and homeostasis of its host's immune system[4–10]. Further evidence
38 comes from a recent study showing that white blood cell compositions differ between children born
39 preterm and at term along with their microbiome composition[11]. However, our understanding of the
40 influence of the microbiota on the adult immune system remains limited. To overcome this, we here
41 investigated the dynamics of immune reconstitution after allogeneic hematopoietic cell therapy (HCT)
42 in hundreds of patients at Memorial Sloan Kettering, treated for various hematological malignancies,
43 including leukemia (**Figure 1A, Table S1**).

44 The conditioning regimen of radiation and chemotherapy administered to HCT patients is the
45 most severe perturbation to the immune system deliberately performed in humans and presents a
46 unique opportunity to investigate the dynamic links between gut microbiota and immune system
47 directly in humans. Immune ablation during conditioning (**Figure 1A**) depletes circulating white
48 blood cell counts. This is followed by characteristic white blood cell trajectories with prolonged
49 periods of neutropenia (<500 neutrophils per μl blood) and reconstitution of white blood cell counts
50 once the transplanted cells engraft (**Figure 1A-C**). The blood of individual patients is monitored daily
51 throughout this recovery, and medications are administered to ensure successful immune
52 reconstitution and prevent graft-vs-host disease (**Figure 1 G,H**).

53 To investigate whether the composition of the gut microbiota could affect the dynamics of
54 circulating white blood cells in our patients, we combined high frequency, longitudinal microbiota
55 compositions measured for hundreds of patients[12] with the detailed blood and metadata from
56 thousands of patients. We compiled over 75,000 administrations of a range of medications, clinical
57 events such as blood stream infections, and over 100,000 host physiology measurements in the form
58 of the most abundant white blood cells—neutrophils, lymphocytes, monocytes, eosinophils—as well
59 as platelet counts (**Figure 1, S1**).

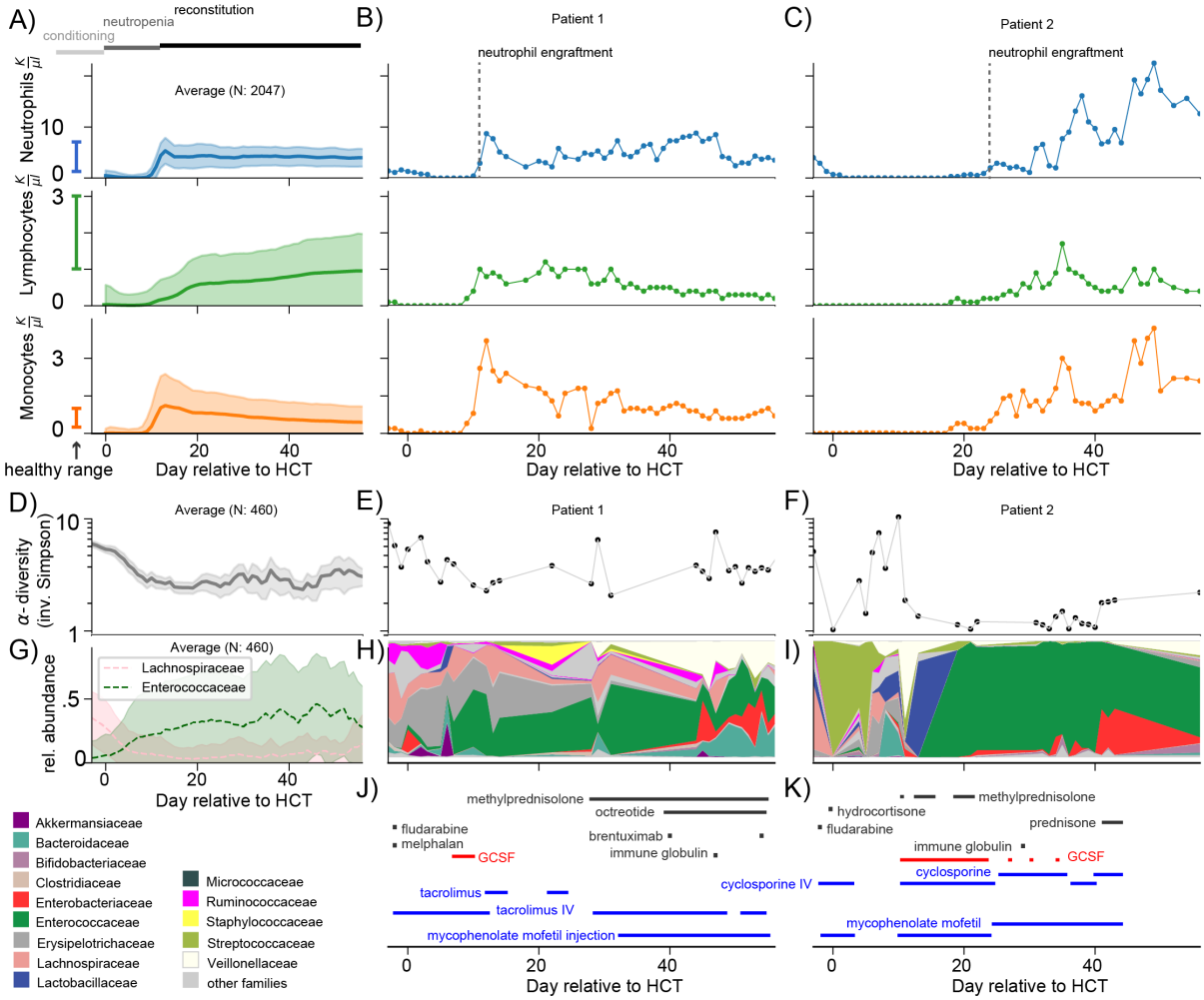
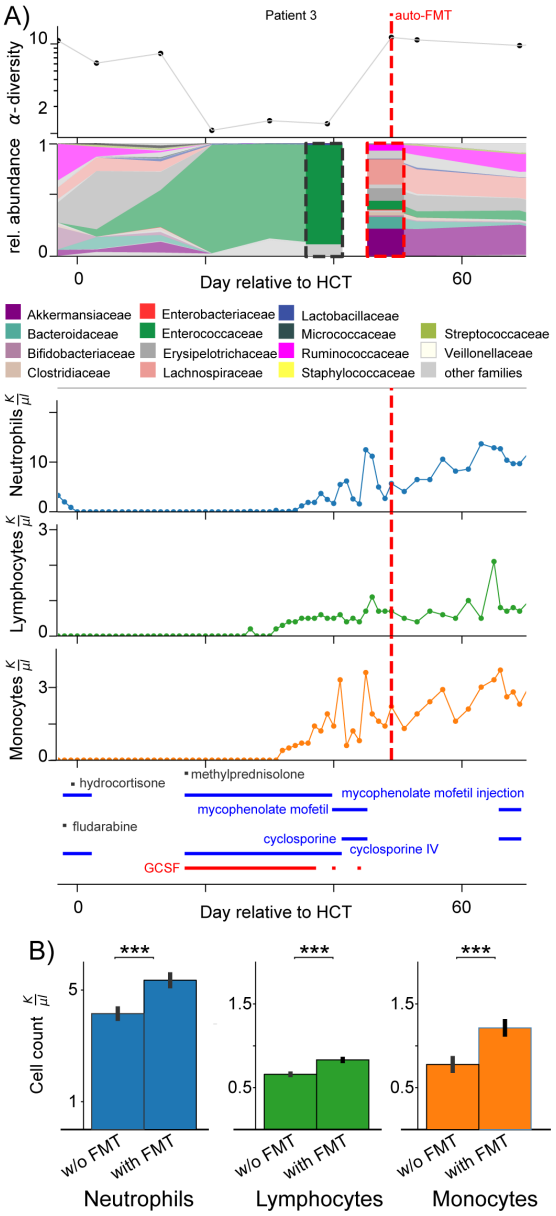


Figure 1: Immune reconstitution after hematopoietic cell transplantation (HCT). A) Three major periods of HCT: immunoablation during conditioning with chemotherapy and radiation before HCT, defined as day 0, post-HCT neutropenia, and reconstitution following neutrophil engraftment lead to repeatedly observed recovery trajectories. Shown are the average counts (shaded: ± std) of neutrophils, lymphocytes and monocytes per day relative to HCT from 2,047 patients transplanted between 2003 and 2018, with count ranges from healthy adults indicated (A), and two individual patients (B,C); patient 1 who received a peripheral blood stem cell graft, and patient 2 who received a graft of umbilical cord blood. For 460 patients, fecal samples were collected during HCT, with varying coverage, revealing loss of microbial diversity (D-F) and commensal families (G-I), often replaced by Enterococcaceae domination (shaded: ± std). J,K) Administration of immunomodulatory medications for the two example patients.

We have shown previously that HCT patients lose gut microbiota biodiversity and commensal microbial families during their treatment (**Figure 1D,G-I**)[12]. A recent study in mice demonstrated that such antibiotic-induced gut decontamination can negatively affect white blood cell recovery after HCT[13]. In our patients, microbial diversity usually recovers slowly during white blood cell reconstitution (**Figure 1D**); however, microbiota recovery as well as immune reconstitution can vary strongly between patients and treatment types. This variation is illustrated by the distinct trajectories

of two patients: patient 1 who received a graft of peripheral blood stem cells, retained high microbiota diversity and engrafted earlier (**Figure 1B,E,H**), and patient 2 who received a graft of umbilical cord blood, lost microbiota diversity and engrafted later (**Figure 1C,F,I**). Low microbiota diversity at the time of neutrophil engraftment has independently been associated with 5-fold increased transplant-related mortality[14], suggesting that that the joint recovery of microbiota diversity and white blood cells in circulation is critical for clinical outcomes.



To identify a causal link between the microbiota and white blood cell recovery dynamics, we first used data from a recent prospective clinical trial of autologous fecal microbiota transplantation (auto-FMT) in our patients[12]. We performed this deliberate microbiota manipulation in 24 patients, randomized to 10 untreated control and 14 treated patients, including patient 3 in **Figure 2A**. To ask if auto-FMT affected white blood cell reconstitution, we compared the 24 patients' neutrophil, lymphocyte and monocyte counts post-engraftment (i.e. when the transplanted hematopoietic cells begin producing new white blood cells). We observed higher counts of each white blood cell type in the period after randomization into control and treatment groups in patients who received an auto-FMT ($p<0.01$, **Figure 2B**).

Figure 2: HCT patient who received an autologous fecal microbiota transplant (auto-FMT, dashed red line) that restored commensal microbial families and ecological diversity in the gut microbiota (A), with concurrent cell counts of peripheral neutrophils, lymphocytes and monocytes and immunomodulatory drug administration. B) Results from a linear mixed effects model with random effects per patient and per day relative to neutrophil engraftment show elevated neutrophil, lymphocyte and monocyte counts in patients receiving auto-FMT after their treatment as compared to control patients (***: $p<0.01$).

109 The increased white blood cell counts in patients receiving auto-FMT pointed towards a
110 causal role of this severe microbiota manipulation on circulatory immune cell dynamics. This increase
111 could be due to the re-establishment of a complex microbiota and the associated metabolic
112 capabilities[6,7,13], or it could be a systemic response of the body to the sudden influx of billions of
113 intestinal organisms. Moreover, while our mixed-effects analysis accounted for patient-specific HCT
114 treatments such as their graft source, chance differences in extrinsic factors such as
115 immunomodulatory drug administration could have affected this result due to the small cohort size.
116 Therefore, we next investigated the link between the gut microbiota composition and the dynamics of
117 white blood cell recovery using our large observational cohort of HCT patients.
118 Homeostasis of mammalian blood cell counts is complex: neutrophils, lymphocytes and
119 monocytes are formed and released into the blood *de novo* by differentiation of hematopoietic
120 progenitor cells from the bone marrow, and they can be mobilized from thymus and lymph nodes
121 (lymphocytes), spleen, liver and lungs (neutrophils); they can also migrate from the blood to other
122 tissues when needed[15]. These processes are dynamic sources and sinks of circulatory white blood
123 cells, and they can be modulated by drugs administered to patients receiving HCT. To identify drivers
124 of these dynamics—including the microbiota—we developed a two-stage approach. We first used a
125 dataset of patients *without* microbiome information to identify medications, clinical observations and
126 other metadata that drive changes in white blood cell counts from one day to the next. This prior
127 information fed into and informed our main analysis which we performed on set of different patients
128 from whom concurrent microbiome samples were available.
129 We first used the longitudinal data of over 1,500 patients from whom no microbiome samples
130 were collected (**Figure 3A**). We analyzed the counts of neutrophils, lymphocytes and monocytes
131 during patients' recovery (i.e. from 5 days before neutrophil engraftment until day 100), observing re-
132 establishment of homeostasis 'as it happened'. We used 3-fold cross-validated (lasso) regression to
133 identify factors consistently associated with increases and decreases in neutrophil, lymphocyte and
134 monocyte counts among these patients, including their stem cell graft source, medications as well as
135 various clinical and demographic data (**Figure S2,S3, Table S1**). We then used the identified effects
136 as prior information for a Bayesian inference to contextualize potential contributions of the gut

137 microbiota. The set of patients with microbiota data was smaller (446 patients, **Figure 3A**), but they
138 provided an unbiased sampling of the white blood cell dynamics observed across patients lacking
139 microbiota data (**Figure 3B**).

140 Notably, as a benchmark, this approach detected the expected strong effects of
141 immunomodulatory medications (**Figure 3C,S4**). Most obviously, we identified the administration of
142 granulocyte-stimulating factor (GCSF)—administered to accelerate recovery from chemotherapy-
143 induced neutropenia[16]—as the strongest driver of neutrophil and monocyte increases from one day to
144 the next, yielding an estimated increase of 97% ([76%, 121%], 90 percent probability density interval
145 [HDI90]) of neutrophils and an increase of 41% ([28%, 54%] HDI90) of monocytes. Lymphocytes,
146 too, increased during GCSF administration, although less so (9%, [-1%, 19%] HDI90). The strongest
147 associations with lymphocyte dynamics resulted from administration anti-inflammatory steroids that
148 are known to induce lymphocytopenia by removing circulatory lymphocytes (prednisone[17], -16%
149 ([-27%, -3%], HDI90).

150 Beyond correctly recognizing the causal roles of medications with known biological
151 mechanisms, our analysis detected feedbacks between white blood cell types. Specifically, we
152 identified a negative feedback loop between neutrophils and lymphocytes, a suppression of
153 monocytes by neutrophils and platelets (**Figure 3D**), and an additional suppression from platelet
154 transfusion events (**Figure 3C**). Monocytes positively fed-back on each of the investigated white
155 blood cell counts, including a positive feedback of monocytes onto themselves. Interestingly, this
156 suggests a homeostatic feedback network of circulatory white blood cell counts.

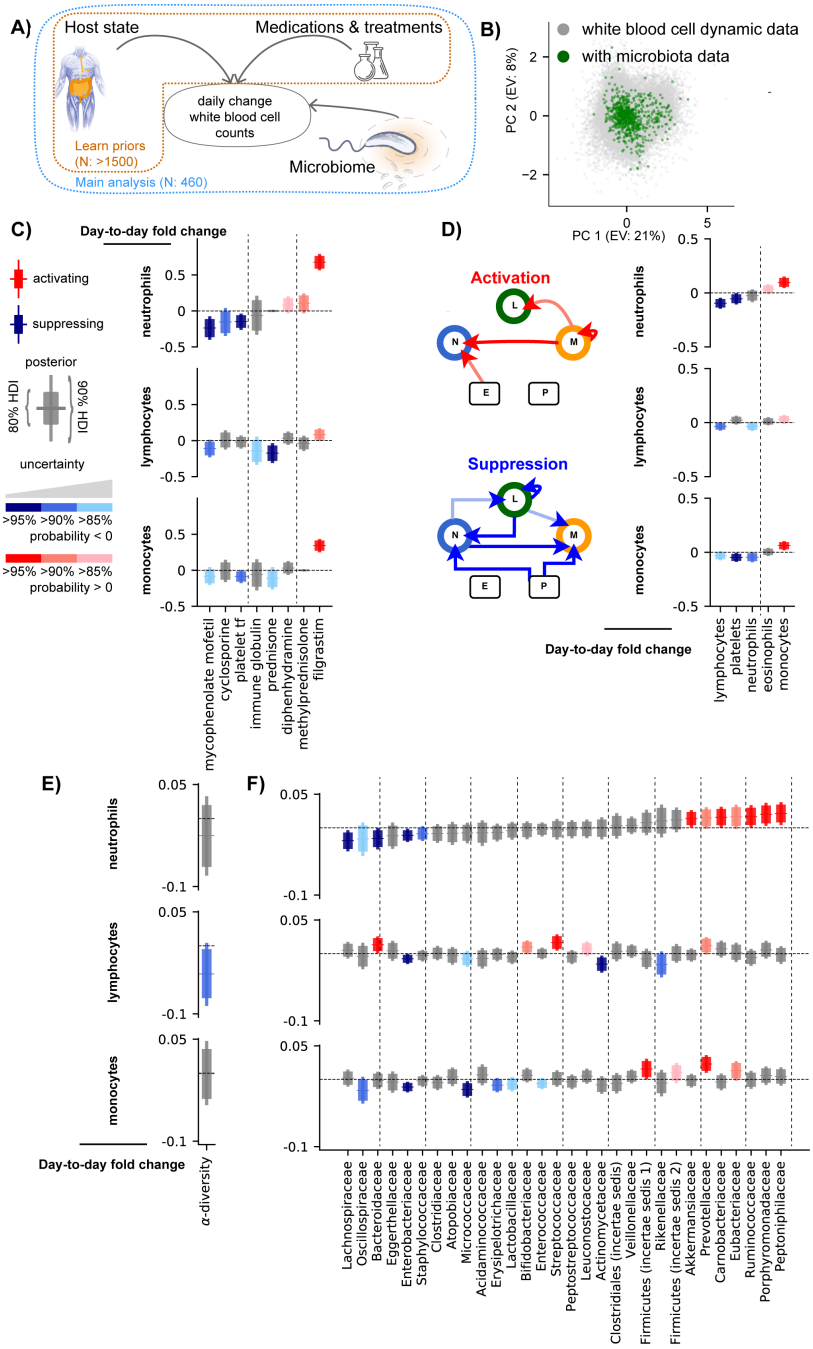


Figure 3: Results from Bayesian inference of coefficients in a model of peripheral white blood cell dynamics. A) Cartoon of the two-stage model inferred from patients' data: observed changes in white blood cell counts between two days are driven by the current state of the host, including the current counts of white blood cells in circulation, the effects of immunomodulatory medications administered, patient metadata and the state of the microbiome. Host state and medications were used in a prior analysis on patients without microbiota samples (brown outline, results in Figure S3), and informed the main analysis presented in panels C-F on patients with concurrent microbiome information. B) Visualization of the dynamic white blood cell data used; scatter plot of the principal components (PC) of observed daily changes of neutrophils, lymphocytes and monocytes without (grey) and with (green) concurrent longitudinal microbiome data. C-F) Regression results; bars show the posterior parameter estimate distributions (highest density intervals, HDI, color intensity indicate credibility of results) for the jointly inferred effects of immunomodulatory medications and platelet transfusion (C), homeostatic regulation network of white blood cells counts and associated posteriors (D), microbiota alpha diversity (inverse Simpson index, E) and log-microbial family abundances (F) on the observed daily changes in neutrophils, lymphocytes and monocytes.

To investigate the immunomodulatory role of the gut microbiota ecosystem, we included the inverse Simpson diversity index—an ecological index calculated at the OTU level. Additionally, we also included the relative abundances of OTUs combined at the family level where many important physiological traits are shared across member species[18]. We found that high diversity suppressed circulatory lymphocyte counts (**Figure 3E**). A raise in microbiota diversity of 5 inverse Simpson

208 index units corresponded to a -6% ([-11%, 1%] HDI90) suppression of lymphocytes. This is similar in
209 magnitude to the effect of immunosuppressive drug administrations to prevent graft-vs-host disease
210 (mycophenolate mofetil, -10% [-20%, 1%] HDI90) and that of anti-inflammatory steroids (**Figure**
211 **3C**). Whether this apparent immunosuppressive effect of a complex microbiota is an indicator for
212 good recovery and re-acquisition of homeostatic regulation among our patients, or indeed, may be
213 problematic in the special situation of an immunosuppressed population, requires further
214 investigation. Nonetheless, high microbiota diversity has a strong positive association with better
215 survival of HCT patients[14].

216 Our analysis revealed several commensal microbial families that modulated white blood cell
217 dynamics (**Figure 3F**). Bacteroidaceae increased lymphocytes, Prevotellaceae increased monocytes,
218 and Ruminococcaceae increased neutrophils. The effect sizes of individual families were relatively
219 small, suggesting that potential immunomodulatory microbiota treatments would need to be
220 prolonged or simultaneously target multiple microbial families. A notable exception would be a
221 reduction in Streptococcaceae: Streptococcaceae dominations reached maximum relative abundances
222 of over 90% in our patient cohort while their median abundance was <4%. Our results suggest that an
223 increase in Streptococcaceae from this median level to >90% domination results in daily lymphocyte
224 increases of ~6% or, conversely, that resolving a Streptococcaceae domination could have an
225 immunosuppressive effect. This response of circulatory lymphocyte counts to Streptococcaceae
226 domination could reflect reduced lymphocyte migration to the gut epithelium or indicate a gut
227 dysbiosis-induced acute immune response that may influence the risk of gut domination-born
228 infections[7,19].

229 Microbiota homeostasis is characterized by high abundances of obligate anaerobes. Our
230 analysis revealed that two major families of commensal obligate anaerobes, Bacteroidaceae and
231 Lachnospiraceae, suppressed neutrophils (**Figure 3F**). Neutrophil-associated immunity in mice
232 responds to gut bacteria, which could affect tumor progression in distal organs with reduced overall
233 tumor burden upon systemic neutrophil depletion[20]. Intriguingly, our finding that common commensal
234 families can modulate circulatory immune cells in human patients supports the idea established in
235 mice that microbiota modulation could influence systemic immunity to fight tumor progression[20,21].

236 The most immediate interface between mammalian hosts and their microbiota are the various
237 epithelia of the body. In mice, individual microbial species trigger specific immune responses at the
238 gut epithelium[9,22,23]. Also, various microbial traits, including the production of metabolic byproducts
239 such as short-chain fatty acids[24,25] and bacterial cell wall molecules[1,26,27] modulate immune responses,
240 including granulopoiesis[6]. These traits are shared by many different microbial species, and therefore
241 plausible candidates for the human body to sense gut homeostasis or dysbiosis. High-resolution,
242 longitudinal microbiome data[12] coupled with a large observational dataset of immune reconstitution in
243 cancer patients enabled us to detect immunomodulatory effects of the microbiota ecosystem. Our
244 approach accounted for patient differences such as the source of transplanted stem cells and yielded
245 features that consistently associated with increases or decreases of white blood cells from one day to
246 the next. Intriguingly, systemic effects of microbiota ecosystem features support predictions from
247 evolutionary theory that hosts evolved to sense, respond to and control broad ecological features such
248 as diversity and family abundances, rather than specific strains[28,29].
249 In sum, our work shows evidence of a link between the gut microbiota and the human
250 immune system via peripheral white blood cell dynamics. We were able to infer this link because we
251 leveraged an unprecedented resource: a vast dataset of high-resolution, longitudinal human phenotype
252 data in the form of blood cell counts, compiled with detailed metadata. The effects inferred between
253 the microbiota and the temporal changes in circulatory white blood counts indicate a systemic,
254 homeostatic link that could help understand the modulating role of a 'healthy' microbiota during
255 immunotherapy[30–33], and holds potential for microbiota therapies[20,21,34].
256

257 **Bibliography**

258

259 1. Mazmanian, S. K., Liu, C. H., Tzianabos, A. O. & Kasper, D. L. An immunomodulatory molecule of
260 symbiotic bacteria directs maturation of the host immune system. *Cell* **122,** 107–118 (2005).
261 2. Gomez de Agüero, M. *et al.* The maternal microbiota drives early postnatal innate immune
262 development. *Science* **351,** 1296–1302 (2016).
263 3. Sonnenberg, G. F. & Artis, D. Novel connections and precision approaches. *Nat Rev Immunol* **19,**
264 75–76 (2019).
265 4. Geva-Zatorsky, N. *et al.* Mining the human gut microbiota for immunomodulatory organisms.
266 *Cell* **168,** 928–943.e11 (2017).
267 5. Tan, T. G. *et al.* Identifying species of symbiont bacteria from the human gut that, alone, can
268 induce intestinal Th17 cells in mice. *Proc Natl Acad Sci U S A* **113,** E8141–E8150 (2016).
269 6. Balmer, M. L. *et al.* Microbiota-derived compounds drive steady-state granulopoiesis via
270 MyD88/TICAM signaling. *J Immunol* **193,** 5273–5283 (2014).
271 7. Deshmukh, H. S. *et al.* The microbiota regulates neutrophil homeostasis and host resistance to
272 Escherichia coli K1 sepsis in neonatal mice. *Nat Med* **20,** 524–530 (2014).
273 8. Atarashi, K. *et al.* ATP drives lamina propria T(H)17 cell differentiation. *Nature* **455,** 808–812
274 (2008).
275 9. Ivanov, I. I. *et al.* Specific microbiota direct the differentiation of IL-17-producing T-helper cells
276 in the mucosa of the small intestine. *Cell Host Microbe* **4,** 337–349 (2008).
277 10. Cahenzli, J., Köller, Y., Wyss, M., Geuking, M. B. & McCoy, K. D. Intestinal microbial diversity
278 during early-life colonization shapes long-term IgE levels. *Cell Host Microbe* **14,** 559–570 (2013).
279 11. Olin, A. *et al.* Stereotypic immune system development in newborn children. *Cell* **174,** 1277–
280 1292.e14 (2018).
281 12. Taur, Y. *et al.* Reconstitution of the gut microbiota of antibiotic-treated patients by autologous
282 fecal microbiota transplant. *Sci Transl Med* **10,** (2018).
283 13. Staffas, A. *et al.* Nutritional Support from the Intestinal Microbiota Improves Hematopoietic
284 Reconstitution after Bone Marrow Transplantation in Mice. *Cell Host Microbe* **23,** 447–457.e4
285 (2018).
286 14. Taur, Y. *et al.* The effects of intestinal tract bacterial diversity on mortality following allogeneic
287 hematopoietic stem cell transplantation. *Blood* **124,** 1174–1182 (2014).
288 15. Scheiermann, C., Frenette, P. S. & Hidalgo, A. Regulation of leucocyte homeostasis in the
289 circulation. *Cardiovasc Res* **107,** 340–351 (2015).
290 16. Gabrilove, J. L. *et al.* Effect of granulocyte colony-stimulating factor on neutropenia and
291 associated morbidity due to chemotherapy for transitional-cell carcinoma of the urothelium. *N*
292 *Engl J Med* **318,** 1414–1422 (1988).
293 17. Fauci, A. S. & Dale, D. C. Alternate-day prednisone therapy and human lymphocyte
294 subpopulations. *J Clin Invest* **55,** 22–32 (1975).
295 18. Goldford, J. E. *et al.* Emergent simplicity in microbial community assembly. *Science* **361,** 469–
296 474 (2018).
297 19. Taur, Y. *et al.* Intestinal domination and the risk of bacteremia in patients undergoing allogeneic
298 hematopoietic stem cell transplantation. *Clin Infect Dis* **55,** 905–914 (2012).
299 20. Lakritz, J. R. *et al.* Gut bacteria require neutrophils to promote mammary tumorigenesis.
300 *Oncotarget* **6,** 9387–9396 (2015).
301 21. Lakritz, J. R. *et al.* Beneficial bacteria stimulate host immune cells to counteract dietary and
302 genetic predisposition to mammary cancer in mice. *Int J Cancer* **135,** 529–540 (2014).
303 22. Ivanov, I. I. *et al.* Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139,**
304 485–498 (2009).
305 23. Tanoue, T. *et al.* A defined commensal consortium elicits CD8 T cells and anti-cancer immunity.
306 *Nature* **565,** 600–605 (2019).

24. Ulven, T. Short-chain free fatty acid receptors FFA2/GPR43 and FFA3/GPR41 as new potential therapeutic targets. *Front Endocrinol* **3,** 111 (2012).
25. Smith, P. M. *et al.* The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341,** 569–573 (2013).
26. Hergott, C. B. *et al.* Peptidoglycan from the gut microbiota governs the lifespan of circulating phagocytes at homeostasis. *Blood* **127,** 2460–2471 (2016).
27. Ladinsky, M. S. *et al.* Endocytosis of commensal antigens by intestinal epithelial cells regulates mucosal T cell homeostasis. *Science* **363,** (2019).
28. Schluter, J. & Foster, K. R. The evolution of mutualism in gut microbiota via host epithelial selection. *PLoS Biol* **10,** e1001424 (2012).
29. Foster, K. R., Schluter, J., Coyte, K. Z. & Rakoff-Nahoum, S. The evolution of the host microbiome as an ecosystem on a leash. *Nature* **548,** 43–51 (2017).
30. Zitvogel, L., Ma, Y., Raoult, D., Kroemer, G. & Gajewski, T. F. The microbiome in cancer immunotherapy: Diagnostic tools and therapeutic strategies. *Science* **359,** 1366–1370 (2018).
31. Routy, B. *et al.* Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359,** 91–97 (2018).
32. Gopalakrishnan, V. *et al.* Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359,** 97–103 (2018).
33. Hsiao, E. Y. *et al.* Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* **155,** 1451–1463 (2013).
34. Chen, H. *et al.* A forward chemical genetic screen reveals gut microbiota metabolites that modulate host physiology. *Cell* (2019). doi:10.1016/j.cell.2019.03.036
35. Wheeler, D. L. *et al.* Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **35,** D5-12 (2007).
36. Anasetti, C. *et al.* Peripheral-blood stem cells versus bone marrow from unrelated donors. *N Engl J Med* **367,** 1487–1496 (2012).
37. Davies, S. M. *et al.* Engraftment and survival after unrelated-donor bone marrow transplantation: a report from the national marrow donor program. *Blood* **96,** 4096–4102 (2000).
38. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. Royal Stat. Soc. (Methodological)* **58,** 267–288 (1996).
39. Salvatier, J., Wiecki, T. V. & Fonnesbeck, C. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science* **2,** e55 (2016).

342 **Supplementary Information**

343 Appended and available online.

344

345 **Acknowledgments and Conflicts of Interest**

346 We thank Kevin R. Foster, Sandra B. Andersen, Rene Niehus, Katharine Coyte, Jonathan Kevin Sia,
347 Eric G. Pamer, and the Xavier lab for helpful discussion and comments on the manuscript. This work
348 was supported by the National Institutes of Health (NIH) grant U01 AI124275 and grant R01
349 AI137269 to JBX. This research was further supported by the MSKCC Cancer Center Core Grant P30
350 CA008748, the Parker Institute for Cancer Immunotherapy at Memorial Sloan Kettering Cancer
351 Center, the Sawiris Foundation, the Society of Memorial Sloan Kettering Cancer Center, MSKCC
352 Cancer Systems Immunology Pilot Grant, Empire Clinical Research Investigator Program. MS
353 received funding from the Burroughs Wellcome Fund Postdoctoral Enrichment Program, the Damon
354 Runyon Physician-Scientist Award, and the Robert Wood Johnson Foundation. MRMvdB and JUP
355 received financial support from Seres Therapeutics. TMH is investigator in the Pathogenesis of
356 Infectious Diseases from the Burroughs Wellcome Fund, and funded via an award from Geoffrey
357 Beene Foundation, and NIH RO1 AI093808. M-AP has received honoraria from AbbVie, Bellicum,
358 Bristol-Myers Squibb, Incyte, Merck, Novartis, Nektar Therapeutics, and Takeda; has received
359 research support for clinical trials from Incyte, Kite (Gilead) and Miltenyi Biotec; and serves on data
360 and safety monitoring boards for Servier and Medigene and scientific advisory boards for MolMed
361 and NexImmune. The funders had no role in study design, data collection and analysis, decision to
362 publish, or preparation of the manuscript.

363

364 **Author Contributions**

365 J.S., B.P.T. and J.B.X. wrote the manuscript. J.S. and J.B.X. designed the analyses. J.U.P. and Y.T.
366 contributed to the data preparation. All authors contributed to the writing and interpretation of the
367 results.

368

SUPPLEMENTARY INFORMATION

Methods

Complete blood count collection and characterization

Absolute white blood cells count data were obtained from routine complete blood counts ordered by clinicians during normal clinical practice, used to obtain informative diagnostic and monitoring information. Blood samples received in the clinical hematology laboratory were analyzed using Sysmex XN automated hematology analyzers (Sysmex, Lincolnshire, IL) and, when needed based on specific flags and parameters as per MSKCC standard operating procedures, were validated manually using the Sysmex DI-60 Slide Processing System or CellaVision DM9600 Automated Digital Morphology System (Sysmex, Lincolnshire, IL).

Microbial sample collection, data availability and ecological diversity calculation

The microbiota data analyzed here was obtained from a previous publication [12]. Briefly, stool samples were collected from patients and sequenced using primers for the V4 and V5 region of the 16S gene. OTU clustering was performed using UPARSE and classification was conducted using NCBI BLAST on the refseqRNA database[35]. Diversity as measured by the inverse Simpson index of a sample was calculated by $IS = \frac{1}{\sum_{j=1}^{N} p_{ij}^2}$, where p is the relative abundance of the jth OTU out of N total OTUs in sample i.

Linear mixed-effects model of white blood cell counts

We used a linear mixed effects model to analyze white blood cell counts before and after auto-FMT with the following model equation: *value ~ armpost + (1|day) +(1|patientid)*, where *value* is the count of a given white blood cell on a given *day* relative to neutrophil engraftment, and *armpost* is a binary variable that is 1 when a patient was from the FMT treated arm of the trial and *day* was greater than or equal to the day of FMT procedure.

396 **Dynamic systems analyses**

397 We analyzed factors associated with the observed changes of absolute counts of neutrophils,
398 lymphocytes and monocytes between two days.

399

400 *Interval data calculation*

401 For a given daily interval, a medication was considered present for at least part of the interval when it
402 was administered on either endpoint. For the microbiota and blood stream infection data, data from
403 the end day were considered for the interval. Homeostatic feedback calculations used the geometric
404 mean of the white blood cell counts between the two endpoints. We included only those covariates
405 that were present during at least 10 intervals. The stem cell graft source is a major determinant of
406 engraftment times and can affect recovery dynamics[36,37], and we therefore included intercept terms for
407 unmodified peripheral blood stem cell grafts (PBSC), bone marrow (BM), T-cell depleted graft
408 (*ex-vivo*) by CD34+selection (TCD) and cord blood (cord).

409

410 *Identification of effectors on white blood cell dynamics from patients without microbiome data*

411 To reveal factors that affect day-to-day white blood cell count changes, we used L1-regularized
412 (lasso) regression[38]. Stronger regularization leads to fewer predictors in the final model; we here used
413 3-fold cross validation leaving out 33% of patients at each cross-validation step to choose the
414 regularization strength, yielding a sparse coefficient matrix of white blood cell changes. We used this
415 to parameterize a linearized differential equation of daily blood cell count changes,

416 $$\frac{\Delta \ln(W_i)}{\Delta t} = gr + \sum_{j=0}^{WBCs} s_{ij} W_j + \sum_{i=0}^{drugs,clinical\ obs,HCT\ parameters,patient\ demographics} \beta_i X_i$$

417

418 where $\Delta\ln(W_i)$ is the log difference between single days of neutrophils, lymphocytes or monocytes
419 counts, and $\Delta t=1$ for all intervals. g*r* is the intercept and represents the base line rate of change, and *s*
420 is the coefficient for homeostatic feedback of the current blood state, i.e. the counts of blood cells W_j
421 (neutrophils, lymphocytes, monocytes, eosinophils, platelets) on the rate of change of W_i. X_i are
422 indicator variables for clinical observations (e.g. bacteremia), patient gender, the drug administrations
423 and the type of cells that were transplanted (BM, PBSC, TCD, cord). Continuous variables were zero

424 centered and standardized to unit variance. After cross-validated regression, the resulting coefficient
425 matrix was used to inform priors of a hierarchical Bayesian model.
426
427 *Expanded analysis on patients with microbiome data*
428 We inferred effects of medications, clinical parameters and patients' demographics onto white blood
429 cell rates of change in an analogous Bayesian regression using the package PyMC3 for the Python
430 programming language[39]. We incorporated the above results as prior information, now additionally
431 including a linear effect from log-transformed family relative abundances (F_i). We combined the
432 relative OTU abundances at the family level, and only included families that at least once reached a
433 relative abundance of >1% and had relative abundances >0.1% in more than 10% of all samples. We
434 also included a linear effect of the zero-centered inverse Simpson diversity index (*IVS*) calculated at
435 the OTU level and standardized to unit variance. This data set was smaller, and we therefore used a
436 hierarchical model set up with random intercepts and partial pooling per transplanted cell type (*HCT*)
437 and patient gender (*SEX*), leading to following model (other parameters as before)

438 $$\frac{\Delta \ln(W_i)}{\Delta t} = gr + (1|HCT) + (1|SEX) + \sum_{j=0}^{WBCs} s_{ij} W_j + \sum_{i=0}^{drugs,patient\ age} \beta_i X_i + \sum_{k=0}^{families} \beta_i F_i + \beta_i IVS$$

439 The prior parameter distributions of the resulting computational graph were informed with the
440 regression results from the cross-validated analysis of patients without microbiota data (no reuse of
441 data). When a coefficient was non-zero, it's prior distribution was a permissive, wide Normal
442 distribution centered on the coefficient (e.g. for the activation of neutrophils from monocytes, we used
443 Normal(μ=6.4e-02, std=100^2) as prior. A coefficient that yielded a coefficient of zero after the cross-
444 validated regression of non-microbiota patients would be assigned a regularizing prior centered
445 around zero (e.g. Normal(μ=0, std=0.01)). Hierarchy was included by assigning regularized random
446 intercepts for transplanted cell types and patient gender. For example, the reduced baseline rate of
447 monocytes among cord patients identified in the non-microbiota regression would lead to a prior of
448 Normal(μ=-3.5e-02, std=1) with partially pooled information across all transplanted cell types via a
449 shared variance with prior Exponential(lambda=10). The resulting graph of the model is visualized in
450 **Figure S2**.

Supplementary Figures

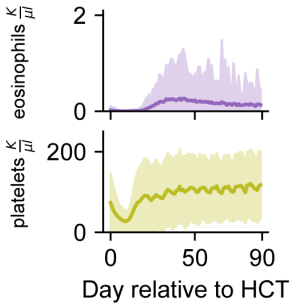


Figure S1: Eosinophil counts and platelet counts per day relative to HCT; line: mean, shaded: ± std)

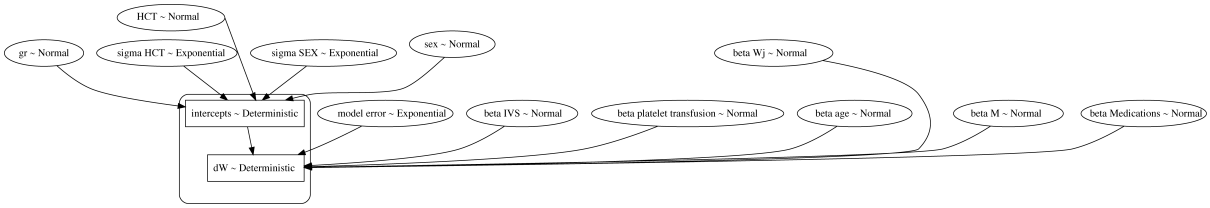


Figure S2: Computational graph of the hierarchical linear model of daily changes in white blood cell counts (*dW*) modeled as a function of graft source (HCT), patient characteristics (sex, age), medications (*Medications, platelet transfusions*), homeostatic feedbacks (*Wj*), log-relative abundances of microbiota families (*M*), and ecological diversity (*IVS*).

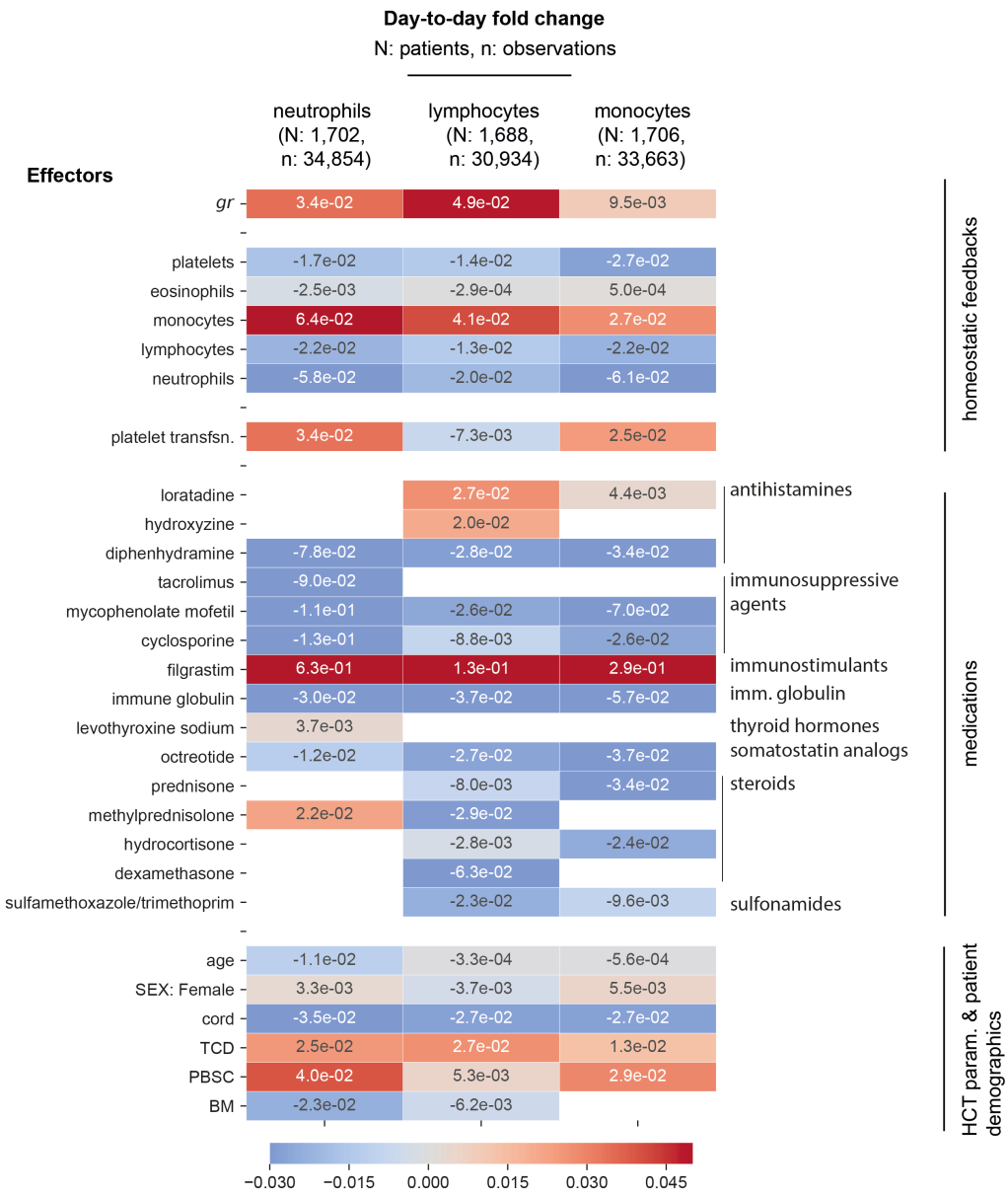


Figure S3: Coefficients from independent, L1-penalized (lasso), 3-fold cross-validated regressions of daily changes in neutrophils, lymphocytes and monocytes. gr: intercept; TCD: T-cell depleted graft (ex-vivo) by CD34+selection; PBSC: peripheral blood stem cells; BM: bone marrow; cord: umbilical cord blood.

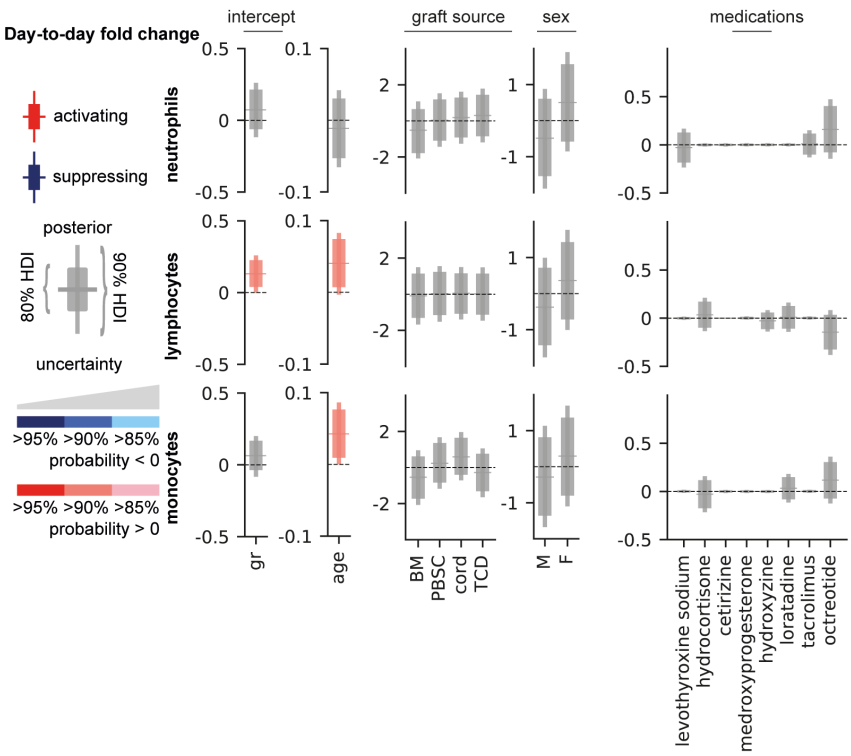


Figure S4: Additional coefficient estimates from the Bayesian regression, see also Figure 3.

Table S1: Patient characteristics. TCD: T-cell depleted graft (*ex-vivo*) by CD34+selection; PBSC: peripheral blood stem cells; BM: bone marrow; cord: umbilical cord blood.

Graft sources	**%-age**
PBSC	39
TCD	38
cord	14
BM	8

Age	**years**
mean	51
std	14
min	18
max	79

Sex	**%-age**
Female	59
Male	41

Conditioning regimen	**%-age**
ABLATIVE	56
REDUCE	29
NONABL	15

480
481 **Supplementary Data**

482 The data used in our study is organized in supplementary tables (data-tables.zip), with corresponding
483 filenames (italic):
484 1. *alphadiversity.csv*: inverse Simpson diversity index per sample
485 2. *hctmeta.csv*: HCT metadata
486 3. *infections.csv*: positive blood culture results
487 4. *microbiome.csv*: relative abundances of major microbial families
488 5. *sample_id_link.csv*: lookup table to raw microbiota sequencing data[12]
489 6. *pidmeta.csv*: anonymized patient demographics
490 7. *treatments.csv*: medication administration/treatment events
491 8. *wbc.csv*: absolute counts of neutrophils, lymphocytes, monocytes, eosinophils, and platelets
492