Integrated meta-omic analyses of the gastrointestinal tract microbiome in patients undergoing allogeneic hematopoietic stem cell transplantation

ANNE KAYSEN, ANNA HEINTZ-BUSCHART, EMILIE E. L. MULLER[1], SHAMAN NARAYANASAMY, LINDA WAMPACH, CÉDRIC C. LACZNY[2], NORBERT GRAF, ARNE SIMON, KATHARINA FRANKE, JÖRG BITTENBRING, PAUL WILMES, and JOCHEN G. SCHNEIDER

BELVAUX, LUXEMBOURG; HOMBURG, GERMANY

In patients undergoing allogeneic hematopoietic stem cell transplantation (allo-HSCT), treatment-induced changes to the gastrointestinal tract (GIT) microbiome have been linked to adverse outcomes, most notably graft-versus-host disease (GvHD). However, it is presently unknown whether this relationship is causal or consequential. Here, we performed an integrated meta-omic analysis to probe deeper into the GIT microbiome changes during allo-HSCT and its accompanying treatments. We used 16S and 18S rRNA gene amplicon sequencing to resolve archaea, bacteria, and eukaryotes within the GIT microbiomes of 16 patients undergoing allo-HSCT for the treatment of hematologic malignancies. These results revealed a major shift in the GIT microbiome after allo-HSCT including a marked reduction in bacterial diversity, accompanied by only limited changes in eukaryotes and archaea. An integrated analysis of metagenomic and metatranscriptomic data was performed on samples collected from a patient before and after allo-HSCT for acute myeloid leukemia. This patient developed severe GvHD, leading to death 9 months after allo-HSCT. In addition to drastically decreased bacterial diversity, the post-treatment microbiome showed a higher overall number and higher expression levels of antibiotic resistance genes (ARGs). One specific *Escherichia coli* strain causing a paravertebral abscess was linked to GIT dysbiosis, suggesting loss of intestinal barrier integrity. The apparent selection for bacteria expressing ARGs suggests that prophylactic antibiotic administration may adversely affect the overall treatment outcome. We therefore assert that such analyses including information about the selection of pathogenic bacteria expressing ARGs may assist clinicians in "personalizing" regimens for individual patients to improve overall outcomes. (Translational Research 2017;186:79–94)

[1] Present address: Department of Microbiology, Genomics and the Environment, UMR 7156 UNISTRA—CNRS, Université de Strasbourg, Strasbourg, France

[2] Present address: Chair for Clinical Bioinformatics, Saarland University, Saarbrücken, Germany

From the Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Belvaux, Luxembourg; Saarland University Medical Center, Klinik für Pädiatrische Onkologie und Hämatologie, Homburg, Germany; Saarland University Medical Center, Klinik für Innere Medizin I, Homburg, Germany; Saarland University Medical Center, Klinik für Innere Medizin II, Homburg, Germany.

Submitted for publication February 8, 2017; revision submitted May 17, 2017; accepted for publication June 12, 2017.

Reprint requests: Paul Wilmes and Jochen G. Schneider, University of Luxembourg, 4362 Esch-sur-Alzette, Luxembourg; e-mail: paul.wilmes@uni.lu or jg.schneider@outlook.com.

1931-5244

http://dx.doi.org/10.1016/j.trsl.2017.06.008

Abbreviations: aGvHD = acute graft-versus-host disease; allo-HSCT = allogeneic hematopoietic stem cell transplantation; ARG = antibiotic resistance gene; ATG = antithymocyte globulin; bp = base pair; cDNA = complementary DNA; RHM = reference healthy microbiome; contig = contiguous sequence; GIT = gastrointestinal tract; GvHD = graft-versus-host disease; IMP = Integrated Meta-omic Pipeline; MG = metagenomic; MT = metatranscriptomic; NCBI = National Center for Biotechnology Information; nt = nucleotide; OTU = operational taxonomic unit; PAMP = pathogen-associated molecular pattern; rRNA = ribosomal RNA; SCFA = short-chain fatty acid; SNV = single nucleotide variant; TP = time point

AT A GLANCE COMMENTARY

Kaysen A, et al.

Background

Allogeneic hematopoietic stem cell transplantation is a therapy for many hematological malignancies. However, its effects on the gastrointestinal tract microbiome remain poorly understood.

Translational Significance

- Allogeneic hematopoietic stem cell transplantation and its associated drug treatments have detrimental effects on the gastrointestinal tract microbiome.
- Antibiotic treatment of these patients gives rise to overgrowth of pathogenic microbes, possessing and expressing antibiotic-resistant genes.
- Loss of epithelial barrier function likely contributes to bacterial and fungal invasions and activation of inflammatory responses.
- Enrichment in potentially pathogenic strains through accompanying antibiotic treatment may further contribute to worsening overall treatment outcome through systemic infection.
- Pathogenic microbes might provide antigens to antigen-presenting cells, activating various immune effectors that associate with acute graft-versus-host disease.

INTRODUCTION

Humans live in a close ("symbiotic") relationship with an inherent "microbiome", comprised of bacteria, archaea, and unicellular eukaryotes. The most densely populated human body habitat is the gastrointestinal tract (GIT).[1] The GIT microbiome plays a myriad of roles vital to human physiology, including digestion of food, synthesis of vitamins, production of short-chain fatty acids, and the prevention of colonization by pathogens through exclusion.[2] In a healthy human GIT, microorganism homeostasis is tightly regulated by the host's immune system.[3-5] However, various perturbations, such as antibiotic attenuation of sensitive bacteria, may disrupt this balanced state, leading to a state typically referred to as "dysbiosis",[3,6] in which pathogenic microbes can overgrow the community.[6] Furthermore, reduced intestinal barrier function can facilitate translocation of microorganisms and microbial products from the GIT lumen to mesenteric lymph nodes and/or the bloodstream,[7] putting the host at risk for local infections and sepsis.[6,8]

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) represents an effective treatment for several hematological malignancies. Transplantation is preceded by a conditioning regimen, consisting of either total body immune ablative irradiation or high doses of chemotherapy, to facilitate engraftment of transplanted stem cells. Moreover, allo-HSCT is known to greatly impact the stability and integrity of the GIT microbiome, resulting in substantially reduced bacterial diversity and the emergence of dominance by (often drug-resistant) single bacterial taxa.[9]

The conditioning treatment for allo-HSCT may also lead to mucositis of the GIT, culminating in the formation of ulcers, dysphagia, and diarrhea.[10] One complication of allo-HSCT is acute graft-versus-host disease (aGvHD),[11] a systemic, inflammatory disease that is provoked by a complex allogeneic immune response, primarily affecting the skin, liver, and GIT.[12] In addition, the GIT microbiome has been implicated in the development and/or exaggeration of aGvHD. Specifically, the damaged GIT epithelial barrier in allo-HSCT patients allows translocation of microorganisms or pathogen-associated molecular patterns (PAMPs).[13] PAMPs can then activate antigen-presenting cells, leading to alloactivation and proliferation of donor T cells that trigger aGvHD.[13]

Supportive care of patients receiving allo-HSCT includes prophylactic broad-spectrum antibiotic treatment,[14] an intervention that also selects for potential pathogens carrying antibiotic resistance genes (ARGs),[15] within the GIT microbiome. Such antibiotics also drive horizontal transfer of ARGs among commensal bacteria, often including numerous opportunistic pathogens.[16] Antibiotic treatment has ambiguous effects on treatment outcome. On the one hand, a low bacterial diversity at engraftment, possibly caused by chemotherapy, total body irradiation, and/or broad-spectrum antibiotics, has been linked to

detrimental outcomes.[17] On the other hand, GIT decontamination using antimicrobials has been observed to lower the rate of aGvHD.[18,19]

Previous studies have demonstrated altered microbiome community structures directly after allo-HSCT or conditioning treatment.[17,20-22] Abnormal bacterial GIT communities, especially following antibiotic treatment, can facilitate expansion of yeasts, including potentially fatal, invasive *Candida albicans* infections.[23,24] However, it is not yet known how GIT microbial communities including archaea and eukaryotes evolve over longer periods of time. It is also uncertain how disruption of the microbiome, via antibiotic regimens and aGvHD, affects the human host and overall treatment outcome.

Here, we analyzed in detail the GIT microbiome of patients undergoing allo-HSCT. The main objectives were to describe the changes in the prokaryotic (bacterial and archaeal) community as well as in the eukaryotic subpopulation during treatment. Based on this data, the samples from one patient were subjected to detailed characterization using a combination of high-resolution metagenomic (MG) and metatranscriptomic (MT) analyses. Population- and strain-level resolutions highlighted higher expression of ARGs in samples following treatment driven by apparent selection of organisms encoding ARGs. This study serves as a proof-of-concept for future meta-omic studies of the GIT microbiome in the context of allo-HSCT and other intensive medical treatments.

MATERIAL AND METHODS

Study participants and fecal sample collection. This study was approved by the Ethics Review Board of the Saarland Amendments 1 and 2 (Reference Number 37/13) and by the Ethics Review Panel of the University of Luxembourg (Reference Number ERP-15-029). Following written informed consent, 16 patients undergoing allo-HSCT were enrolled in the study.

For microbial diversity and richness analyses, patients were included only if fecal samples were obtained from at least 2 of the following time points: (1) up to 8 days before allo-HSCT (designated time point-1, "TP1"); (2) directly after allo-HSCT (up to 4 days after allo-HSCT, "TP2"); and/or (3) shortly after the expected time of engraftment (between days 20 and day 33 after allo-HSCT, "TP3"). One additional patient was selected for in-depth analysis of treatment effects over an extended period of time. From this patient, fecal samples were collected 13 days before allo-HSCT, as well as 75 and 119 days after allo-HSCT. Fecal samples were immediately flash-frozen onsite and preserved at −80°C to ensure the integrity of the biomolecules of interest.

Extraction of biomolecules from fecal samples. Fecal samples were pretreated as described in Heintz-Buschart et al.[25] DNA and RNA were extracted from this pretreated sample as described previously using the DNA/RNA/Protein kit (Qiagen, Hilden, Germany),[26,27] and additional DNA was obtained using the PowerSoil DNA isolation kit (MO BIO Laboratories, Carlsbad, Calif). Extracted biomolecules were stored at −80°C until sequencing.

16S and 18S rRNA gene amplicon sequencing and data analysis. Amplification and paired-end sequencing of extracted and purified DNA was performed on an Illumina MiSeq platform at the Groupe Interdisciplinaire de Génoprotéomique Appliquée (GIGA, Belgium). The V4 region of the 16S rRNA gene, used for identifying bacteria and archaea, and the 18S rRNA gene, used for resolving the eukaryotic community structure, were amplified and sequenced as previously described.[28] 16S rRNA gene sequencing reads were processed using the LotuS pipeline[29] (version 1.34) with default parameters. Processed reads were clustered into operational taxonomic units (OTUs), designating taxa with similar amplicon sequences, at an identity level of 97%. For taxonomic assignment of 16S rRNA gene amplicon sequencing data, we used the Ribosomal Database Project classifier.[30] OTUs with a confidence level below 0.8 at the domain level were filtered out as well as OTUs with no more than 10 reads in any given sample.

To process the 18S rRNA gene sequencing reads, a workflow specifically designed to process nonoverlapping reads was used.[31] For classification of 18S rRNA gene amplicon sequencing data, the Protist Ribosomal Reference database[32] was employed. After processing, OTUs with less than 10 reads in all samples were removed, as well as unclassified OTUs and OTUs belonging to the taxon *Craniata*. For analyses, the 16S and 18S rRNA gene sequencing data were rarefied to the lowest number of respective reads for any sample (16S: 71,051 reads and 18S: 1020 reads).

Statistical analyses and plots were generated in R (version 3.2.1).[33] Microbial alpha-diversity and richness were determined at the OTU level, by calculating the Shannon diversity index and the Chao1 index, using the R package "vegan".[34] Statistical comparison of diversity and richness was carried out using the Kruskal-Wallis test, the non-parametric Wilcoxon rank sum test, or, when applicable, the Wilcoxon signed-rank test. Comparisons with P values <0.05 were considered statistically significant. Differential analyses of taxa, based on 16S rRNA gene sequencing data, were performed using the Bioconductor package DESeq2,[35] combined with the Wald test, after multiple-testing adjustment.

Metagenomic and metatranscriptomic sequencing, processing, and assembly. MG and MT sequencing of the extracted DNA and RNA fractions was conducted by GATC Biotech AG (Konstanz, Germany). Ribosomal RNA (rRNA) was depleted from the RNA fractions using a Ribo-Zero Gold rRNA Removal kit (Illumina, San Diego, Calif). Libraries representing both nucleic acid fractions were sequenced using a 100-bp paired-end approach on an Illumina HiSeq 2500 using HiSeq V3 reagents (Illumina). MG and MT data sets were processed using the Integrated Meta-omic Pipeline (IMP) version 1.1.[36] Further information on this pipeline and on calculations used in this work can be found in Supplementary File 1.

Population-level binning of contigs from the co-assembly. To analyze and compare the population-level structure of the microbial communities based on the assembled genomic information, contiguous sequences (contigs) were binned. Using VizBin,[37,38] 2D embeddings based on BH-SNE of the contigs of at least 1000 nt were produced, as part of IMP. Population-level clusters[39] were selected as previously described.[25] The resulting bins are hereafter referred to as "population-level genomes."

Within a community, the relative size of a population-level genome was determined by dividing the number of MG reads mapping to the contigs forming this cluster by the total number of MG reads mapping to all the contigs used in the assembly.

Taxonomic affiliation, reassembly, and sequence comparisons of reconstructed population-level genomes. Taxonomic affiliation, reassembly, and sequence comparisons were performed using a modified workflow of previously described complementary tools and methods.[40-46] Details are described in Supplementary File 1.

Detection of antibiotic resistance genes. Antibiotic resistance genes (ARGs) within a community or population were searched against Resfams version 1.2[47] using HMMer version 3.1b2.[48] We used the core version of the Resfams database, which includes 119 protein families.[47] In accordance with the HMMer user manual, only genes identified with a bitscore higher than the binary logarithm of the total number of genes (of the community or population) were retained.

Variant identification. Variants were identified in population-level reassembled genomes using SAMtools mpileup[49] with default settings, which include the calling of single nucleotide variants as well as the identification of small insertions/deletions. The output of SAMtools mpileup was filtered using a conservative heuristic established in Eren et al.,[50] which reduces the effect of sequencing errors.

Extraction, sequencing, and analysis of bacterial DNA from a blood culture. DNA was extracted from a blood culture of an organism identified as a multidrug-resistant *Escherichia coli* and sequenced on an Illumina MiSeq, 300 bp paired-end at GIGA. To assess the relationships between different *E. coli* genomes, PanPhlAn[51] and the provided database, including 118 *E. coli* reference strains, was applied. In accordance with the PanPhlAn manual, only genes present in 10 or more genomes were further considered.

Availability of data and materials. Reassembled population-level genomes of *E. coli* (ID 6666666.166711) and *Enterococcus faecium* (ID 6666666.166708) are accessible via the RAST guest account (http://rast.nmpdr.org, login: guest; password: guest). Preprocessed MG and MT reads were submitted to the NCBI Sequence Read Archive (SRA) repository under the BioProject ID PRJNA317435 (http://www.ncbi.nlm.nih.gov/bioproject/317435). Supplementary tables are archived on Zenodo (https://doi.org/10.5281/zenodo.268914).

RESULTS

Patient characteristics and treatment. Anthropometric and clinical information of the 10 female and 6 male patients included in the study are provided in Table I. Five patients with relapsed or refractory lymphoma received FluBuCy (fludarabine, busulfan, cyclophosphamide) as conditioning treatment, 6 acute myeloid leukemia (AML) patients received BuCy (busulfan, cyclophosphamide), 1 myeloma, and 1 AML patient received Treo/Flu (treosulfan, fludarabine), 1 AML patient received FluBu (fludarabine, busulfan [total dose 9.6 mg/kg, 3 days of intravenous Bu]), and 2 refractory AML patients received FLAMSA-Bu (fludarabine, amsacrine, busulfan) conditioning treatments. Grafts from 8 fully matched unrelated, 3 single antigen mismatched unrelated, and 5 sibling donors were used. 1.5 years after allo-HSCT, 10 patients were still alive and 6 deceased. Twelve patients developed aGvHD and were treated with corticosteroids (0.5–2 mg/kg/d). Even so, 3 of them progressed to at least grade III aGvHD.[52]

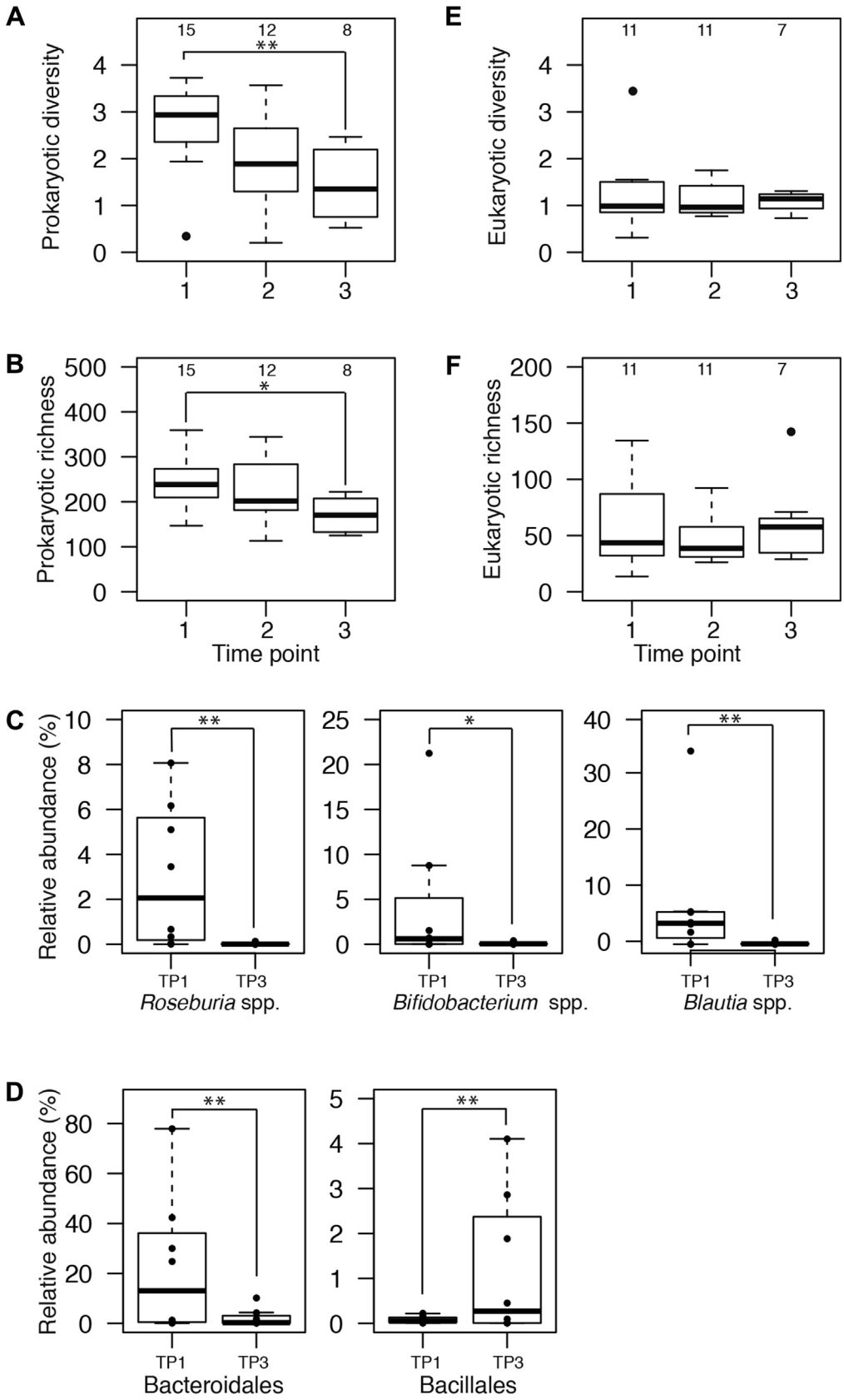
As a prophylactic treatment, patients received a fluoroquinolone antibiotic during leukopenia. On the occurrence of fever, patients were treated with piperacillin-tazobactam, followed by meropenem and subsequently vancomycin, if necessary. As antifungal prophylaxis, patients received 200 mg fluconazole. In case of suspected fungal infection, patients also received antifungal treatment with liposomal amphotericin B or caspofungin (Table I).

Changes within the GIT microbiome of patients undergoing allo-HSCT. We assessed the diversity and richness of the microbial community, separately, for

Table I. Anthropometric and clinical information of the study cohort

Patient	Sex	Age	Underlying disease*	Donor relationship and HLA†	Conditioning regimen‡	Antimicrobials§	GvHD‖¶	Outcome 1.5 y after allo-HSCT
A01	m	43	lymphoma	MRD	FluBuCy	F, M, P-T, V	Skin I°	alive
A03	m	56	lymphoma	MRD	FluBuCy	AF, F, M, P-T, other	-	deceased day 66, relapse
A04	f	43	AML	MUD	BuCy	AF, F, M, V	Skin I°	alive
A05	m	49	lymphoma	MMUD	FluBuCy	AF, F, M, P-T, V	Skin II°	deceased day 275, pneumonia
A06	m	52	AML	MRD	BuCy	AF, F, M, P-T, V, other	-	alive
A07	f	63	AML	MMUD	FLAMSA-Bu	AF, F, M, P-T, V, other	**Skin II°, GIT III°**	deceased day 268, GvHD
A08	f	50	AML	MUD	BuCy	AF, F, M, P-T, V	Skin I°	alive
A09	m	30	lymphoma	MUD	FluBuCy	F, M, P-T	-	deceased day 212, pneumonia
A10	m	54	AML	MRD	BuCy	F, M, P-T	Skin I°, GIT II°	alive
A12	m	57	lymphoma	MUD	FluBuCy	F, M, P-T, V, other	Skin III°	alive
A13	m	57	AML	MRD	BuCy	AF, F, M, V	Skin I°, lung II°	alive
A17	m	66	AML	MUD	BuCy	F, M, V	Skin II°	alive
A18	f	67	AML	MUD	FluBu	F, M, P-T, V, other	**Skin III°, GIT III°**	deceased day 184, GvHD
A19	f	58	myeloma	MUD	Treo/Flu	F, M, P-T	-	deceased day 39, relapse
A20	m	51	AML	MMUD	FLAMSA-Bu	AF, F, M, P-T, V, other	**Skin II°, GIT II°**	alive
A21	f	64	AML	MUD	Treo/Flu	AF, M, P-T, V, other	Skin II°	alive

*AML: acute myeloid leukemia.
†MRD: matched related, MUD: matched unrelated, MMUD: mismatched unrelated, HLA: human leukocyte antigen.
‡Bu: busulfan, Cy: cyclophosphamide, Flu: fludarabine, FLAMSA: fludarabine, amsacrine, Treo: treosulfan.
§AF: antifungal (other than prophylaxis), F: fluoroquinolone, M: meropenem; P-T: piperacillin-tazobactam, V: vancomycin.
‖Organ involvement, stages according to Glucksberg et al.[52]
¶Bold: aGvHD with summed stages ≥ 4 considered as severe aGvHD.


A
Prokaryotic diversity
15
12
8
**
B
Prokaryotic richness
*
Time point
E
Eukaryotic diversity
11
11
7
F
Eukaryotic richness
Time point
C
Relative abundance (%)
TP1
TP3
Roseburia spp.
Bifidobacterium spp.
Blautia spp.
D
Bacteroidales
Bacillales

prokaryotic (bacteria and archaea; 16S rRNA gene sequencing) and eukaryotic (18S rRNA gene sequencing) community structures. From TP1 to TP3, the prokaryotic subpopulation drastically and significantly decreased in diversity (2.2-fold reduction, Fig 1, *A*) and richness (1.3-fold reduction, Fig 1, *B*). On the genus level, average decreases of 119-, 47-, and 44-fold relative abundances of the genera *Roseburia*, *Bifidobacterium* and *Blautia*, respectively, were observed from TP1 to TP3 (Fig 1, *C*). At the taxonomic rank of order, decreased Bacteroidales relative abundance was observed in parallel with increased Bacillales (Fig 1, *D*). Moreover, we identified only 1 OTU belonging to the domain archaea, the methanogen *Methanobrevibacter smithii*,[53] that was detected in 13 of 35 total samples (and 10 of 15 patients) with a total of 914 reads.[54]

Analysis of the eukaryotic community did not reveal statistically significant differences in Shannon diversity (Fig 1, *E*) or Chao1 richness (Fig 1, *F*), between the different TPs, with no apparent statistically significant difference being observed for the 8 patients undergoing specific antifungal treatments. Overall, around 99% of classified eukaryotic OTUs belonged to the fungal domain, with the majority representing the genera *Saccharomyces, Candida*, and *Kluyveromyces*. Only a few different and lowly abundant protists could be identified including *Vorticella* sp., *Prorodon teres*, and *Phytophthora* sp.[54]

In summary, we found generally decreased bacterial diversity after allo-HSCT, while the eukaryotic community stayed relatively stable. To further explore the effects of treatment on the structure and function of the GIT microbiome, we applied a detailed meta-omic approach on 1 patient.

Patient A07: description of treatment and status of the patient. Due to the uniqueness of the case, we chose to focus on 1 patient, patient A07, who displayed a marked reduction in bacterial diversity, with high relative abundances of opportunistic pathogens (Fig 2, *A* and *B*) and a fatal treatment outcome. This 63-year-old female patient, being treated for AML (with deletion 7q), was refractory to conventional induction (3 + 7) and salvage chemotherapy with high-dose cytarabine and mitoxantrone and, therefore, needed further treatment. As a second-line therapy, FLAMSA-Bu,[55] a modified sequential conditioning regimen for allo-HSCT treatment of refractory AML, was implemented (fludarabine 30 mg/m^2 on days −11 to −8, cytarabine 2000 mg/m^2 on days −11 to −8, amsacrine 100 mg/m^2 on days −11 to −8 and busulfan 3.2 mg/kg on days −7 to −4) for remission induction and transplantation (Fig 2, *D*). She then received peripheral hematopoietic stem cells from a single HLA-C antigen mismatched, unrelated donor. After engraftment on day 26, her bone marrow was hypocellular, but free of leukemia. Further immunosuppression therapy consisted of antithymocyte globulin (anti–T-cell antibodies) from days −4 to −2, mycophenolate mofetil until day 28, and cyclosporine until day 100.

Indicating elevated inflammation, a high level of C-reactive protein before and around allo-HSCT was observed, which decreased slightly, but stayed considerably high throughout the entire observation period[54] (Fig 2, *C*). After initial onset of neutropenia after allo-HSCT, the leukocyte count increased to 3500/μl 20 days after allo-HSCT, and then further increased to a normal value 80 days after allo-HSCT. However, considerable fluctuations and later a decrease in the leukocyte count was observed[54] (Fig 2, *C*).

Since this patient had prolonged neutropenia due to refractory leukemia and intensive chemotherapy, various antibiotics and antifungals were used to treat infectious complications before and during transplantation including piperacillin/tazobactam, meropenem, vancomycin, tigecycline, levofloxacin, ceftazidime and amphotericin B (Fig 2, *D*).

74 days after allo-HSCT, the patient developed aGvHD overall grade III, skin stage 2, and GIT stage 3.

Fig 1. Changes in the gastrointestinal microbial community structure in patients receiving allo-HSCT. Boxplots depicting (**A** and **E**) diversity (Shannon diversity index) and (**B** and **F**) richness (Chao1 richness estimator) per collection time point (TP), for (**A** and **B**) prokaryotes and (**E** and **F**) eukaryotes, respectively. The number of samples per collection TP is indicated at the top of each box. Diversity and richness were determined after rarefaction of the data set. Statistically significant decreases in prokaryotic diversity, between TP1 and TP3 (*P* value 0.014 in Kruskal-Wallis rank sum test), and in prokaryotic richness, between TP1 and TP3 (*P* value 0.026, Wilcoxon rank sum test), were observed. (**C**) Changes in the relative abundances of 3 symbiotic bacterial genera between TP1 and TP3. Genera with ≥1.5-fold decrease, adjusted *P* values < 0.05, and a relative abundance ≥5%, in one sample, were included (adjusted *P* values of 0.0025, 0.026, and 3.68×10^{-5}, Wald test). (**D**) Changes in the relative abundance, of two bacterial orders, between TP1 and TP3 (adjusted *P* values of 0.0054 and 0.009, Wald test). (**C** and **D**) Data from all 8 patients who had samples collected at TP1 and TP3 are displayed. TP1 included samples taken (up to 8 days) before allo-HSCT. TP2 included samples taken up to 4 days after transplantation. TP3 included samples taken between days 20 and 33, after transplantation. Significant differences between TPs are indicated by asterisks (*P value < 0.05, **P value < 0.01). allo-HSCT, allogeneic hematopoietic stem cell transplantation.

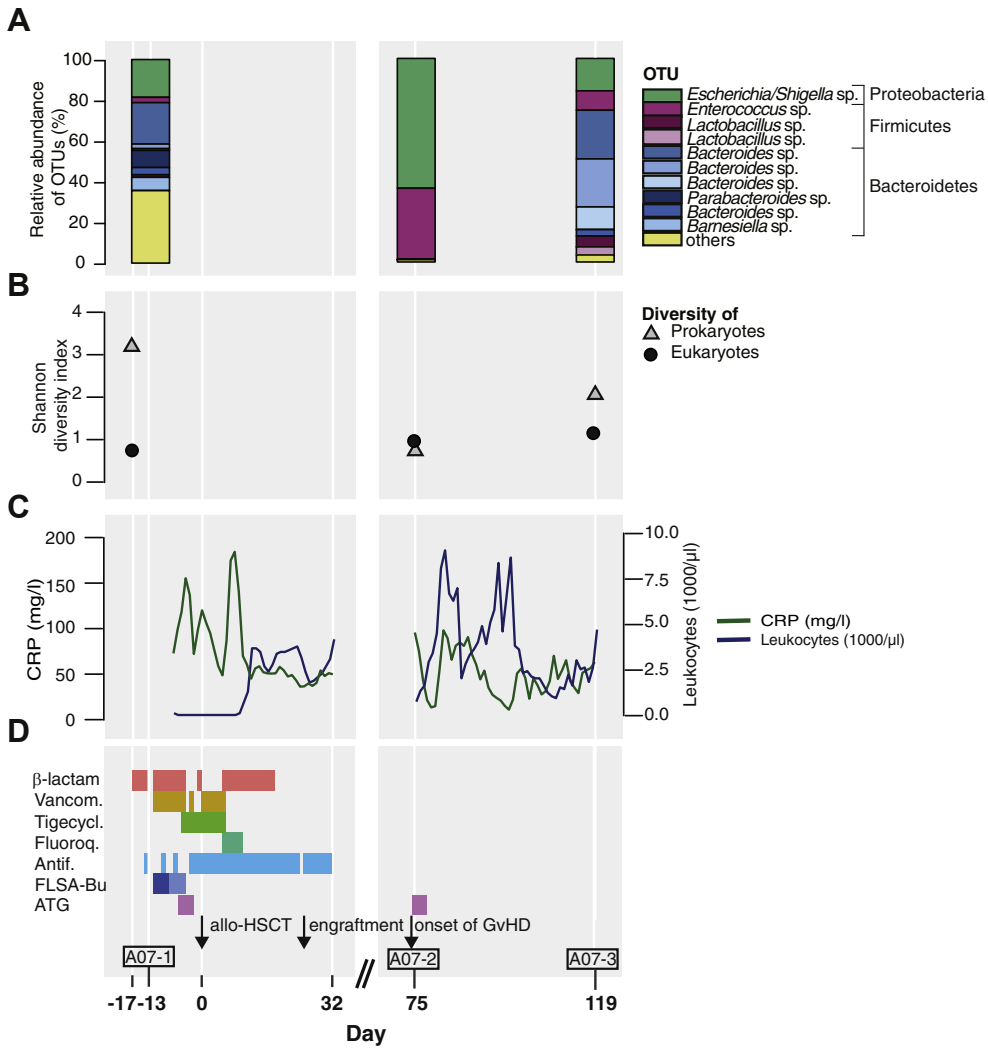


Fig 2. Variation of the microbial community structure over the course of allo-HSCT treatment in patient A07. (**A**) Relative proportions of the 10 most abundant operational taxonomic units (OTUs) based on 16S rRNA gene sequencing. The remaining OTUs are summarized as "others." Similar shades of the colors represent genera belonging to the same phylum. (**B**) Prokaryotic (triangles) and eukaryotic (circles) diversity represented by Shannon diversity indices at sampling TPs throughout the treatment. (**C**) C-reactive protein (CRP) blood levels (green lines) and leukocyte blood count (blue lines). (**D**) Drugs (antibiotics, antifungals, chemotherapeutics, and antithymocyte globulin) administered throughout the treatment. X-axis indicates days relative to the day of transplantation. Vancom, vancomycin; Tigecycl, tigecycline; Fluoroq, fluoroquinolone; Antif, antifungal; FLSA-Bu, FLAMSA-Bu (fludarabine, cytarabine, amsacrine, and busulfan); ATG, antithymocyte globulin; TP, time point; allo-HSCT, allogeneic hematopoietic stem cell transplantation; GvHD; graft-versus-host disease.

As the patient did not respond to 2 mg/kg prednisolone and deteriorated rapidly, antithymocyte globulin (5 mg/kg) was administered for 4 days as second-line GvHD treatment. Partial remission of intestinal GvHD was noted, with a reduction of diarrhea from >20 stools per day (3500–4500 ml) to 4–5 (formed but soft) stools per day. She was bedridden with general fatigue and malaise. Moreover, signs of infection continued, and due to lower back pain, an MRI scan of the spine was performed, showing a paravertebral abscess, which was surgically removed on day 126.

A multidrug-resistant *E. coli* strain was isolated, both from the abscess and from a blood culture, for further analysis. The patient's health status improved, she was able to walk again, and was discharged from hospital on day 209. However, she was readmitted on day 260 and deceased on day 268, due to GvHD and systemic inflammatory response syndrome, suspected to be

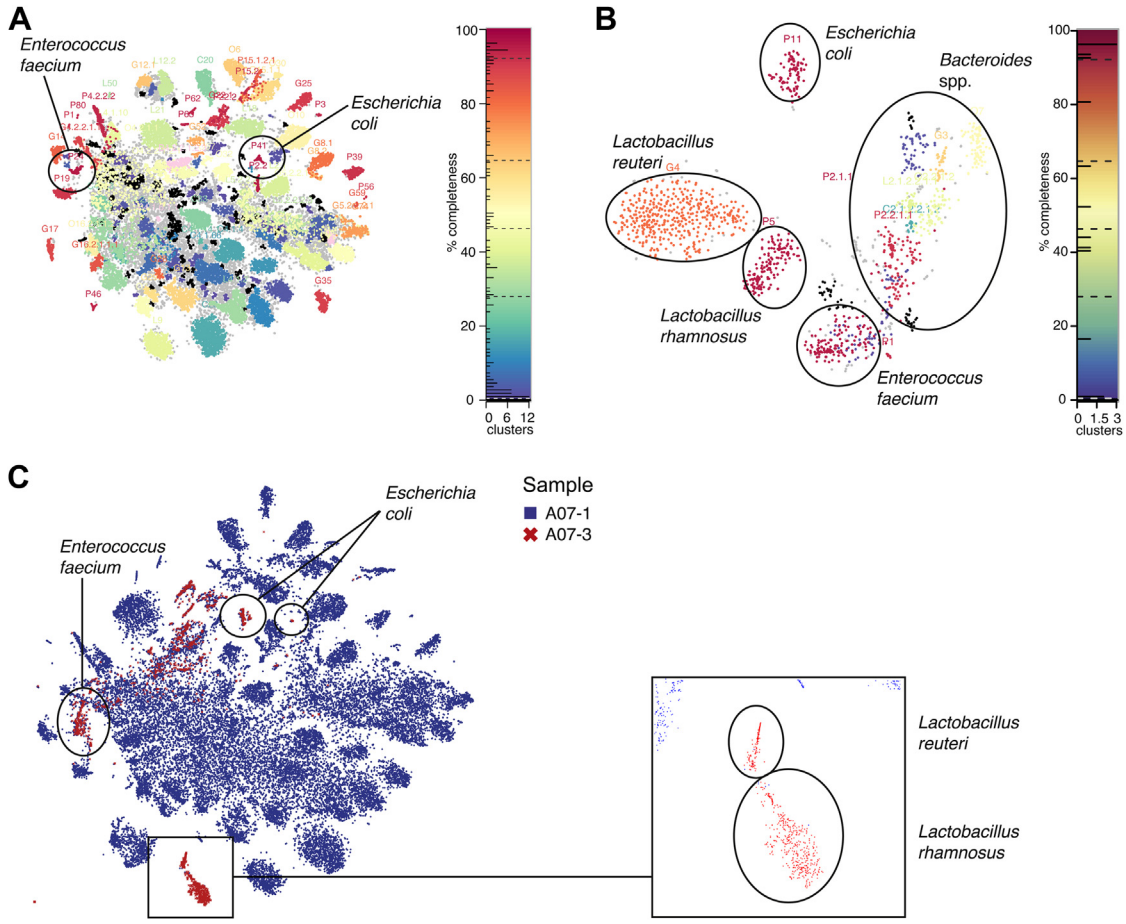


Fig 3. Barnes-Hut Stochastic Neighbor Embedding (BH-SNE)-based visualization of genomic fragment signatures of microbial communities present in samples from patient A07. Points represent contigs ≥1000 nt. Clusters are formed by contigs with similar genomic signatures. (**A**) Visualization of pretreatment sample contigs. (**B**) Visualization of post-treatment sample contigs. (**A** and **B**) Points within clusters are colored according to the reconstructed genomes' completeness based on the number of unique essential genes. Lines within the colored bars indicate the number of clusters at each percentage of completeness. (**C**) Combined visualization of contigs derived from pretreatment (A07-1, blue symbols) and post-treatment (A07-3, red symbols) samples. The inset displays a magnification of a section of the plot representing 2 populations (*Lactobacillus reuteri* and *Lactobacillus rhamnosus*), which are only present in the post-treatment sample. In each representation, clusters representing *Escherichia coli* and *Enterococcus faecium* are indicated. contigs, contiguous sequences.

bacterial sepsis. However, no pathogen could be recovered from blood cultures.

Patient A07: changes in the microbial community structure during treatment. Fecal samples were taken, as indicated in Fig 2, *D*, at days −13 (sample "A07-1"), day 75 (sample "A07-2"), and day 119 (sample "A07-3"). Fecal prokaryotic diversity decreased markedly after allo-HSCT (Fig 2, *B*). Similarly, in sample A07-1, 177 different OTUs were detected, while A07-2 and A07-3 contained only 62 and 79 OTUs, respectively.

Dominant OTUs of sample A07-1, including several OTUs representing *Bacteroides* spp., *Escherichia/Shigella* sp. and *Enterococcus* sp., reappeared in A07-3 (Fig 2, *A*). However, many less-abundant OTUs, belonging to 25 different normal gut genera, disappeared entirely, including *Anaerostipes* and *Clostridium* cluster IV.[54] OTUs that were decreased in sample A07-3 (compared with sample A07-1) represented 50 genera, including *Alistipes*, *Barnesiella*, *Blautia*, *Clostridium* (clusters XIVa and XI), *Prevotella*, *Roseburia*, and *Ruminococcus*. In addition, OTUs belonging to the genus *Lactobacillus* exhibited a 10-fold increase in relative abundance, while those of *Bacteroides* spp. increased from 27% to 63% in A07-3 (Fig 2, *A*). In total, 19 different OTUs belonging to the genus *Bacteroides*

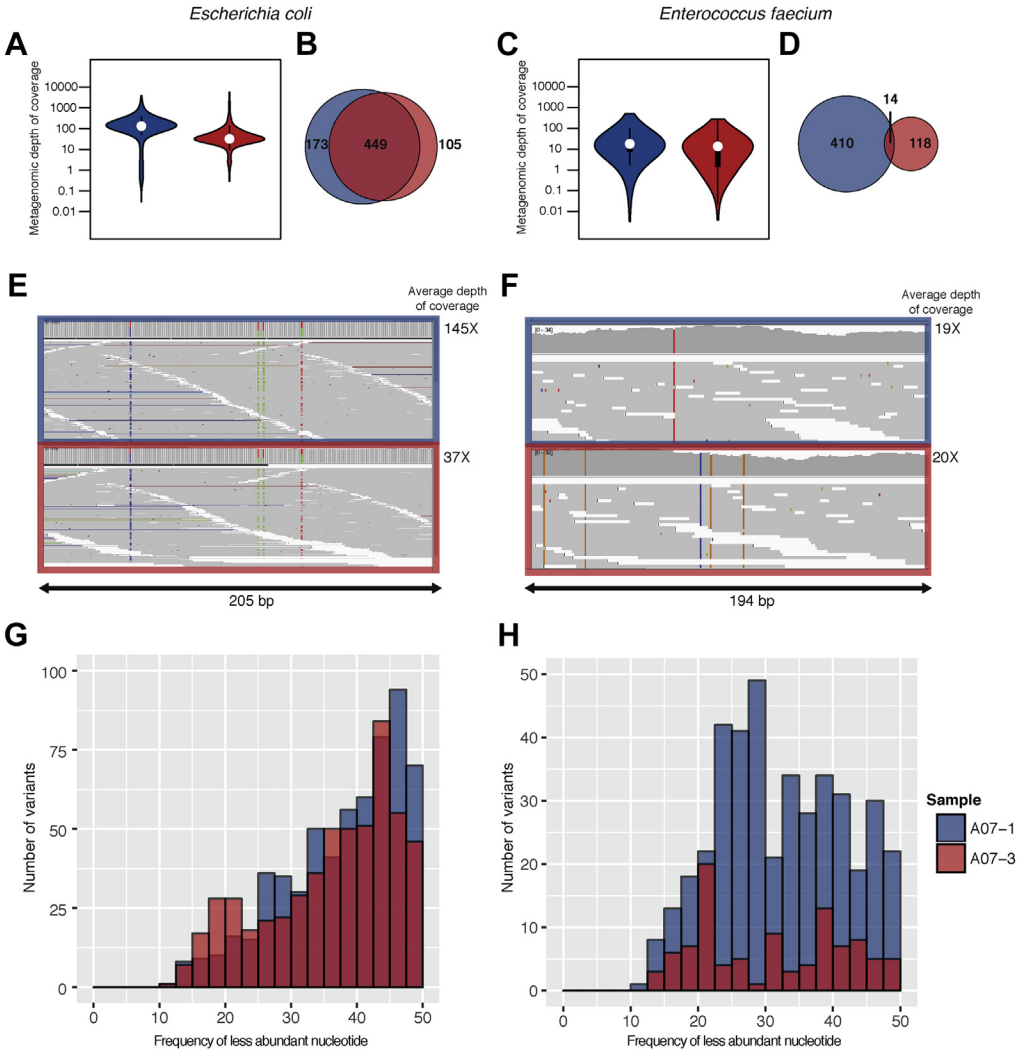


Fig 4. Number and distribution of variants in *Escherichia coli* and *Enterococcus faecium*. (**A** and **C**) Violin plots representing distributions of depth of coverage of the contigs contained in each population-level genome. (**B** and **D**) Venn diagrams indicating the number of variant positions exclusive to each sample, and respectively, the number of variant positions found in both samples. (**E** and **F**) Representation of exemplary sections of the reassembled population-level genomes with aligned reads of both samples, thus highlighting the occurrences of variants in each population, visualized with the Integrative Genomics Viewer. Lengths of the represented sections are indicated, as well as the average MG depths of coverage, of each reconstructed population-level genome. (**G** and **H**) Histograms of the variant frequencies of the minor nucleotide at all variant positions. Panels on the left represent results for *E. coli* and panels on the right represent results for *E. faecium*. Colored figure elements refer to the pretreatment sample (blue; A07-1) and the post-treatment sample (red; A07-3).

were detected in the first sample, 23 in the last sample, and only 5 in A07-2, accounting for 0.07% overall. Similar to the short-term changes in the whole cohort, and as described above, the eukaryotic microbial community exhibited no pronounced changes over time (Fig 2, *B*). Taken together, we observed a drastic decrease in prokaryotic diversity, with relative expansion of few bacteria, including potential pathogens.

Metagenomics- and metatranscriptomics-based analysis of the pre- and post-treatment microbiome. To analyze in detail the changes in the GIT microbiome with emphasis on identification of ARGs and strain-level analysis of specific bacterial populations, coupled MG and MT data sets of samples A07-1 (pretreatment) and A07-3 (post-treatment) were generated. This allows to focus more distinctly on the

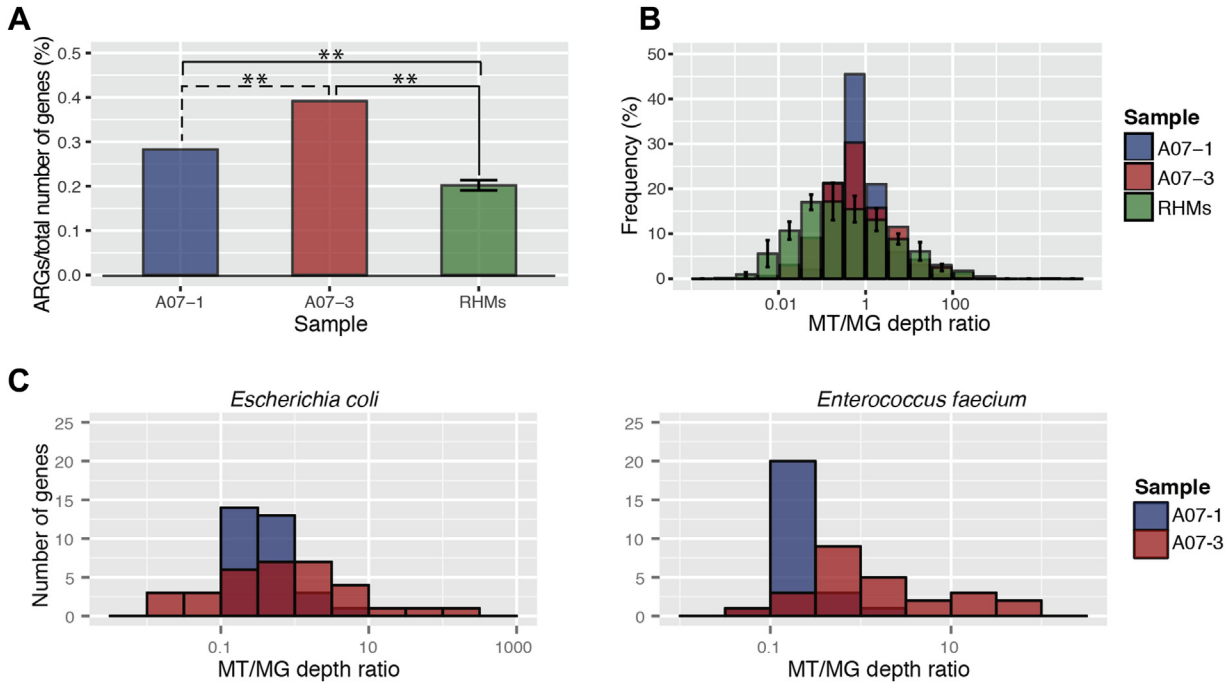


Fig 5. Expression levels and relative abundances of antibiotic resistance genes (ARGs). (**A**) Percentages of identified ARGs (in relation to the total number of genes) in the pretreatment (A07-1) and post-treatment (A07-3) samples, and in the GIT microbiome, of 4 untreated reference healthy microbiomes (RHMs; $^{**}P$ value < 0.01, Fisher's exact test). (**B**) Histogram of the ratios of metatranscriptomic (MT) to metagenomic (MG) depths of coverage of ARGs in the pretreatment and post-treatment samples and in the untreated RHMs. (**C**) Histograms of the ratios of MT to MG depths of coverage of ARGs in population-level genomes of *Escherichia coli* and of *Enterococcus faecium*, in pretreatment and post-treatment samples. Bars represent the number of ARGs, at a specific expression rate, in the pretreatment sample (blue), the post-treatment sample (red), and in the untreated RHMs (green). For the RHMs, the average of 4 data sets is represented with standard deviation as error bars.

effects of allo-HSCT and concurrent antibiotic use on the GIT microbiome.

The visualizations of the two BH-SNE embeddings (Fig 3, *A* and *B*) reflect the drastic change in the GIT microbiome; in particular, the decrease in diversity with the representation of the post-treatment sample A07-3 being exceptionally sparse (Fig 3, *B*). In agreement with our 16S rRNA gene sequencing–based results (Fig 2, *A*), the most abundant populations in A07-3 were identified as *Escherichia coli*, *Enterococcus faecium*, *Lactobacillus reuteri*, *Lactobacillus rhamnosus*, and several species from the genus *Bacteroides*. Also, in accord with those results, most of the clusters were only found in the pretreatment sample (representation of both samples within a single plot, Fig 3, *C*).

Given the potential role of opportunistic pathogens in aGvHD,[13] we were specifically interested in 2 population-level reconstructed genomes that were found in both samples, and whose genomes could be recovered with high completeness, namely *E. coli* and *E. faecium*.

Evidence for selective pressure at the strain level. The average MG depths of coverage indicated a decreased population size of *E. coli* after allo-HSCT (in sample A07-3, Fig 4, *A*), while that of *E. faecium* remained rather constant (Fig 4, *C*). To reveal possible selective "sweeps" in the populations of interest, caused by antibiotics, we performed a gene-wise protein sequence comparison of the different population-level genomes. This analysis revealed that 97.4% of the genes found in the different population-level genomes of *E. coli*, reconstructed from samples A07–1 and A07-3, were 100% identical, and only 1.1% of the genes were less than 95% identical. In *E. faecium*, only 76% of the genes were completely identical and 13.2% of the genes showed less than 95% identity. In *E. coli*, a similarly high number of variants were identified in the pre- and post-treatment samples, with an important overlap of variants identified in both populations (Fig 4, *B* and *E*), whereas only a few variants were present in *E. faecium* of both samples (Fig 4, *D* and *F*).

Table II. Antibiotic resistance genes identified in population-level genomes of GIT *E. coli* from patient A07

Resfams_ID	Number of genes	Resfam Family name	Mechanism
RF0005	1	AAC6-Ib	Aminoglycoside modifying enzyme
RF0007	3	ABCAntibioticEffluxPump	ABC Transporter
RF0027	1	ANT3	Aminoglycoside modifying enzyme
RF0035	1	baeR	Gene modulating resistance
RF0053	1	ClassA	Beta-Lactamase
RF0055	1	ClassC-AmpC	Beta-Lactamase
RF0056	1	ClassD	Beta-Lactamase
RF0065	1	emrB	MFS transporter
RF0088	1	macA	ABC Transporter
RF0089	1	macB	ABC Transporter
RF0091	1	marA	Gene modulating resistance
RF0098	1	MexE	RND antibiotic efflux
RF0101	1	MexX	RND antibiotic efflux
RF0112	1	phoQ	Gene modulating resistance
RF0115	6	RNDAntibioticEffluxPump	RND Antibiotic Efflux
RF0121	1	soxR	Gene modulating resistance
RF0147	1	tolC	ABC transporter
RF0168	6	TE_Inactivator	Antibiotic inactivation
RF0172	1	APH3″	Phosphotransferase
RF0173	1	APH3′	Phosphotransferase
RF0174	1	ArmA_Rmt	rRNA Methyltransferase

Observed nucleotide variant frequencies (Fig 4, *G* and *H*) and patterns of variant distributions (Fig 4, *E* and *F*) indicated that the treatment may have constituted a genetic bottleneck for *E. faecium*, culminating in the observed lower genetic diversity. This also suggests that 2 different mechanisms influenced the respective compositions of *E. coli* and *E. faecium* populations. While the *E. coli* population (composed of different strains) remained relatively unaffected, the *E. faecium* population (mainly represented by a single strain) underwent a selective sweep in response to the antibiotic treatment with selection of a specific genotype expressing ARGs.

Coupled metagenomic and metatranscriptomic analysis of antibiotic resistance genes, in pre- and post-treatment samples from patient A07. ARGs were more often detected post-treatment (0.39% of all genes) than pretreatment (0.28% ARGs, P value $= 6.9 \times 10^{-4}$, Fisher's exact test). The relative abundances of ARGs of both the pre- and post-treatment samples were higher (P value 5.601×10^{-7} and 3.278×10^{-10}) than those in samples from healthy donors (reference healthy microbiomes (RHMs), mean $0.20\% \pm 0.01\%$ [standard deviation]). Moreover, ARG expression was higher in both samples from patient A07 compared with ARG expression in RHMs (Fig 5, *B*).

We were then interested in whether higher numbers of ARGs could also be detected in the specific populations of *E. coli* and *E. faecium*. Within the population-level genome of *E. coli*, 31 ARGs were identified in both samples and 2 additional genes were detected in the post-treatment sample only. In *E. faecium*, 25 ARGs were identified in both samples of which 21 genes were identical (summaries of the ARGs identified in each population-level genome are listed in Tables II and III).[54] In *E. coli*, 20 of the 31 ARGs that were found in both samples exhibited higher levels of expression in the post-treatment sample, while in *E. faecium*, 18 out of 21 ARGs showed higher expression post-HSCT.[54] Although patient A07 was only treated with antibiotics until day 18 (Fig 2, *D*), ARG expression was, in general, higher in the post-treatment sample, both in the whole sample (Fig 5, *B*) as well as in specific populations (Fig 5, *C*).

Genomic characterization of a blood culture *E. coli* isolate and comparison to GIT populations. The genomes of a blood culture isolate, and GIT population-level genomes of *E. coli* from patient A07, exhibited an average nucleotide identity of 99.995%. A heatmap and its corresponding dendrogram reflecting the presence and absence of genes in different strains of *E. coli* indicated that all the *E. coli* genomes from patient A07, the isolate and the genomes from the GIT MG, were more closely related to each other than to the reference genomes (Supplementary Fig. S1). In the genome of the *E. coli* isolate, the same ARGs as in the pre- and post-treatment GIT *E. coli* could be identified, with 4 additional ARGs, compared with the post-treatment GIT *E. coli*.

DISCUSSION

Short-term structural changes in the gastrointestinal microbiome following allogeneic hematopoietic stem cell transplantation. In this study, we observed that

Table III. Antibiotic resistance genes identified in population-level genomes of GIT *E. faecium* from patient A07

Resfams_ID	Number of genes	Resfam Family name	Mechanism
RF0004	1	AAC6-I	Aminoglycoside modifying enzyme
RF0007	9	ABCAntibioticEffluxPump	ABC transporter
RF0033	1	APH3	Aminoglycoside modifying enzyme
RF0066	1	emrE	Other efflux
RF0067	1	Erm23SRibosomalRNAMethyltransferase	rRNA Methyltransferase
RF0104	1	MFSAntibioticEffluxPump	MFS transporter
RF0134	1	Tetracycline_Resistance_MFS_Efflux_Pump	Tetracycline MFS efflux
RF0154	1	vanR	Gylcopeptide resistance
RF0155	2	vanS	Gylcopeptide resistance
RF0168	1	TE_Inactivator	Antibiotic inactivation
RF0172	2	APH3″	Aminoglycoside modifying enzyme
RF0173	2	APH3′	Aminoglycoside modifying enzyme
RF0174	6	ArmA_Rmt	Aminoglycoside resistance

allo-HSCT and its accompanying treatment (including antibiotics), strongly impacted the GIT microbiome, markedly decreasing its bacterial diversity. Moreover, our observed decreased diversity indices agree with values found in an earlier study.[9]

Also in accordance with a study focusing on bacterial diversity around engraftment and its association with survival, we found a reduced bacterial diversity shortly after engraftment in patients who did not survive.[17] A significant decrease in important short-chain fatty acid (SCFA) producers[56-58] (the 3 bacterial genera *Roseburia*, *Bifidobacterium*, and *Blautia*, Fig 1, *C*) was observed. SCFAs, especially the histone deacetylase inhibitor butyrate, are the main energy source for colonocytes, as well as anti-inflammatory agents which regulate NF-κB activation in colonic epithelial cells and cytokine release.[56] In addition, butyrate induces differentiation of regulatory T cells[59] and enhances the intestinal barrier function by regulating the assembly of epithelial tight junctions,[60] and a recent study showed that local administration of exogenous butyrate mitigated GvHD in mice.[61] Depletion of these important SCFA-producers in the GIT enhances the risk for developing GvHD after allo-HSCT.[22,62] The ensuing loss of intestinal barrier integrity facilitates translocation of pathogenic bacteria and PAMPs, activating antigen-presenting cells and consequently $CD8^+$ T cells.

We further found that fungi were the most prominent eukaryotes in the GIT microbiome, and that the eukaryotic diversity was stable during treatment (including antibiotic and antifungal treatment), despite changes in bacterial subpopulations. However, antibiotic treatment may indirectly increase the risk for invasive fungal infections, by opening niches to these organisms, which were previously occupied by commensal bacteria.[23,63] In this patient cohort, we did not observe strong treatment-induced effects on eukaryotic communities. Nevertheless, these organisms remain important for future studies as their overgrowth has previously been linked to adverse treatment outcomes.[24]

Long-term effect of allogeneic stem cell transplantation on the gastrointestinal microbiome. Employing detailed integrated meta-omic analyses, we demonstrated the effects of allo-HSCT and accompanying treatment on the GIT microbiome over an extended period of time. Only one study to date has followed the GIT microbiome trajectory up to 3 months after allo-HSCT.[64] Contrary to that study, which showed that the richness and metabolic capacity of the microbial community recovered after 2 months,[64] we found that the GIT microbial community never fully regained its initial composition, even 4 months after allo-HSCT, with dysbiosis likely conducing to detrimental treatment outcomes. Diversity remained decreased and many bacterial taxa remained absent or drastically decreased, including bacteria whose presence in the human GIT associate with health-promoting properties (such as butyrate production) and whose absence has been linked to negative consequences (such as inflammation).[65-67] The genus *Blautia*, for instance, has been linked to reduced aGvHD-associated mortality and improved overall survival,[22] while *Barnesiella* effectively blocked intestinal domination by vancomycin-resistant enterococci in allo-HSCT patients.[68] On the other hand, potential pathogens, like *Fusobacterium* sp. and *Proteus* sp., appeared only in the post-treatment sample. Also, loss in intestinal barrier integrity may have allowed a GIT-borne *E. coli* to cause a paravertebral abscess (Supplementary Fig. S1).

Identification of antibiotic resistance genes in population-level genomes of opportunistic pathogens and evidence for selective pressure at the strain level. A higher ratio and expression of ARGs within the microbial community was observed post-treatment, even several months following antibiotic treatment (Fig 5,

A and *B*). Strains that carry or horizontally acquired mutations which lead to higher expression of ARGs might have been selected for by the antibiotic treatment.[69] Overall, our observations indicate that antibiotic pressure, and its associated selection of bacteria encoding ARGs, is likely an essential factor in governing the observed expansion of opportunistic, drug-resistant, pathogens.

Interestingly, the multidrug-resistant *E. coli*, isolated from a post–allo-HSCT blood culture, was closely related to the GIT-borne *E. coli* populations. The overlap of ARGs identified in each genome further indicated their association. These findings demonstrate the potential fatal effects of dysbiosis-associated pathogen dominance of the GIT microbiome, and subsequent systemic infections on allo-HSCT patient survival.

CONCLUSION

We observed drastic changes in the composition of the gastrointestinal microbiome of patients after allo-HSCT and supportive care, driven mainly by a decrease in bacterial diversity, but only limited changes in eukaryotes and archaea. Pronounced changes in community structure, especially decreased diversity and expansion of potential pathogens, along with translocation of pathogens and PAMPs could affect overall treatment outcome. Here, we applied novel high-resolution molecular methods to describe the different ways in which GIT bacterial populations respond to antibiotic stress. We identified an increased number and expression of ARGs linked to specific potentially pathogenic strains of clinical relevance during allo-HSCT. Such information could be used to individually tailor patient treatment. Sustaining or restoring diversity, either by: (1) limiting the usage of broad-spectrum antibiotics; (2) fecal microbiome transplantation; and/or (3) administration of specific probiotics could help to increase tolerance or improve the overall efficacy of the therapy.

In summary, we assert that detailed microbiome analyses may be established as a diversified tool for enhancing personalized treatment of immune dyshomeostasis.

ACKNOWLEDGMENTS

Conflicts of Interest: All authors have read the journal's authorship agreement and policy on disclosure of potential conflicts of interest, and have none to declare.

This work was partially funded by the University of Luxembourg (ImMicroDyn1) as well as by an ATTRACT program grant (ATTRACT/A09/03), a European Union Joint Programming in Neurodegenerative Disease grant (INTER/JPND/12/01), a CORE grant (CORE/15/BM/10404093), and a proof of concept grant (PoC/13/02) awarded to PW as well as Aide à la Formation Recherche grants to CCL (AFR PHD/4964712), LW (AFR PHD-2013-5824125), and to SN (AFR PHD-2014-1/7934898) and a CORE junior to EELM (C15/SR/10404839), all funded by the Luxembourg National Research Fund (FNR).

The authors thank the medical staff of the Hematology Department and the Pediatric Department of the Saarland University Medical Center in Homburg, Germany, especially Manuela Faust, Eyad Torfah, and Michael Ehrhardt for sample and data collection. We are deeply grateful for all the patients that participated in the study. Laura A. Lebrun and Claire Battin are thanked for their assistance with fecal sample extractions. Stephanie Kreis and Giulia Cesi are thanked for valuable discussions in the establishment of the project. In silico analyses presented in this article were carried out using the HPC facilities of the University of Luxembourg.[70]

Authors' contributions are as follow: JGS, PW, NG, AS, and JB initiated and designed the study. JB, NG, and AS recruited patients and collected samples. KF collected clinical data. AK performed experiments. AHB, SN, and CCL developed the bioinformatic analysis methods. AK, AHB, EELM, and PW contributed to analysis and interpretation of the data. AK, AHB, EELM, JB, JGS, LW, and PW wrote and revised the manuscript. All authors read and approved the final manuscript.

Supplementary Data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.trsl.2017.06.008.

REFERENCES

1. Hooper LV, Gordon J. Commensal host-bacterial relationships in the gut. Science 2001;292:1115–8.
2. Sekirov I, Russell SL, Antunes LC, Finlay BB. Gut microbiota in health and disease. Physiol Rev 2010;90:859–904.
3. Sommer F, Bäckhed F. The gut microbiota-masters of host development and physiology. Nat Rev Microbiol 2013;11:227–38.
4. Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. Nat Rev Immunol 2009;9:313–23.
5. Atarashi K, Tanoue T, Oshima K, et al. Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. Nature 2013;500:232–6.
6. Stecher B, Maier L, Hardt WD. "Blooming" in the gut: how dysbiosis might contribute to pathogen evolution. Nat Rev Microbiol 2013;11:277–84.
7. Yu LC, Shih YA, Wu LL, et al. Enteric dysbiosis promotes antibiotic-resistant bacterial infection: systemic dissemination of resistant and commensal bacteria through epithelial transcytosis. Am J Physiol Gastrointest Liver Physiol 2014;307:824–35.
8. Khosravi A, Mazmanian SK. Disruption of the gut microbiome as a risk factor for microbial infections. Curr Opin Microbiol 2013;16:221–7.

9. Taur Y, Xavier JB, Lipuma L, et al. Intestinal domination and the risk of bacteremia in patients undergoing allogeneic hematopoietic stem cell transplantation. Clin Infect Dis 2012;55:905–14.
10. Tuncer HH, Rana N, Milani C, Darko A, Al-Homsi SA. Gastrointestinal and hepatic complications of hematopoietic stem cell transplantation. World J Gastroenterol 2012;18:1851–60.
11. Couriel D, Caldera H, Champlin R, Komanduri K. Acute graft-versus-host disease: pathophysiology, clinical manifestations, and management. Cancer 2004;101:1936–46.
12. Jacobsohn DA, Vogelsang GB. Acute graft versus host disease. Orphanet J Rare Dis 2007;2:35.
13. Penack O, Holler E, van den Brink MR. Graft-versus-host disease: regulation by microbe-associated molecules and innate immune receptors. Blood 2010;115:1865–72.
14. Einsele H, Bertz H, Beyer J, et al. Infectious complications after allogeneic stem cell transplantation: epidemiology and interventional therapy strategies - Guidelines of the infectious diseases working party (AGIHO) of the German Society of Hematology and Oncology (DGHO). Ann Hematol 2003;82:175–85.
15. Aminov RI, Mackie RI. Evolution and ecology of antibiotic resistance genes. FEMS Microbiol Lett 2007;271:147–61.
16. Salyers A, Gupta A, Wang Y. Human intestinal bacteria as reservoirs for antibiotic resistance genes. Trends Microbiol 2004;12:412–6.
17. Taur Y, Jenq RR, Perales M, et al. The effects of intestinal tract bacterial diversity on mortality following allogeneic hematopoietic stem cell transplantation. Blood 2014;124:1174–82.
18. Beelen DW, Elmaagacli A, Müller K, et al. Influence of intestinal bacterial decontamination using metronidazole and ciprofloxacin or ciprofloxacin alone on the development of acute graft-versus-host disease after marrow transplantation in patients with hematologic malignancies: final results and. Blood 1999;15:3267–75.
19. Vossen JM, Guiot HF, Lankester AC, et al. Complete suppression of the gut microbiome prevents acute graft-versus-host disease following allogeneic bone marrow transplantation. PLoS One 2014;9:e105706.
20. Holler E, Butzhammer P, Schmid K, 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 Transplant 2014;20:640–5.
21. Montassier E, Batard E, Massart S, et al. 16S rRNA gene pyrosequencing reveals shift in patient faecal microbiota during high-dose chemotherapy as conditioning regimen for bone marrow transplantation. Microb Ecol 2014;67:690–9.
22. Jenq RR, Taur Y, Devlin SM, et al. Intestinal *Blautia* is associated with reduced death from graft-versus-host disease. Biol Blood Marrow Transplant 2015;21:1373–83.
23. Samonis G, Kofteridis DP, Maraki S, et al. Levofloxacin and moxifloxacin increase human gut colonization by *Candida* species. Antimicrob Agents Chemother 2005;49:5189.
24. Zollner-Schwetz I, Auner HW, Paulitsch A, et al. Oral and intestinal *Candida* colonization in patients undergoing hematopoietic stem-cell transplantation. J Infect Dis 2008;198:150–3.
25. Heintz-Buschart A, May P, Laczny CC, et al. Integrated multi-omics of the human gut microbiome in a case study of familial type 1 diabetes. Nat Microbiol 2016;2:16180.
26. Roume H, Heintz-Buschart A, Muller EEL, Wilmes P. Sequential isolation of metabolites, RNA, DNA, and proteins from the same unique sample. Methods Enzymol 2013;531:219–36.
27. Roume H, Muller EEL, Cordes T, Renaut J, Hiller K, Wilmes P. A biomolecular isolation framework for eco-systems biology. ISME J 2013;7:110–21.
28. Wampach L, Heintz-Buschart A, Hogan A, et al. Colonization and succession within the human gut microbiome by archaea, bacteria and microeukaryotes during the first year of life. Front Microbiol 2017;8:738.
29. Hildebrand F, Tadeo R, Voigt A, Bork P, Raes J. LotuS: an efficient and user-friendly OTU processing pipeline. Microbiome 2014;2:30.
30. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 2007;73:5261–7.
31. Hugerth L. Processing amplicons with non-overlapping reads [Internet]. Available at: https://github.com/EnvGen/Tutorials/blob/master/amplicons-no_overlap.rst; 2015. Accessed April 19, 2016.
32. Guillou L, Bachar D, Audic S, et al. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res 2013;41:D597–604.
33. R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2008:5.
34. Oksanen AJ, Blanchet FG, Kindt R, et al. Package "vegan", Community Ecology Package. R package version 2.3–0; 2015;1–285.
35. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
36. Narayanasamy S, Jarosz Y, Muller EEL, et al. IMP: a pipeline for reproducible reference-independent integrated metagenomic and metatranscriptomic analyses. Genome Biol 2016;17:260.
37. Laczny CC, Pinel N, Vlassis N, Wilmes P. Alignment-free visualization of metagenomic data by nonlinear dimension reduction. Sci Rep 2014;4:4516.
38. Laczny CC, Sternal T, Plugaru V, et al. VizBin–an application for reference-independent visualization and human-augmented binning of metagenomic data. Microbiome 2015;3:7.
39. Muller EEL, Pinel N, Laczny CC, et al. Community-integrated omics links dominance of a microbial generalist to fine-tuned resource usage. Nat Commun 2014;5:5603.
40. Madden T. Chapter 16: the BLAST sequence analysis tool. In: The NCBI Handbook [internet]; 2002;:1–15.
41. Huson D, Mitra S, Ruscheweyh H. Integrative analysis of environmental sequences using MEGAN4. Genome Res 2011;21:1552–60.
42. MOLE-BLAST webservice [Internet]. Available at: https://blast.ncbi.nlm.nih.gov/moleblast/moleblast.cgi. Accessed April 19, 2016.
43. Wu M, Scott AJ. Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. Bioinformatics 2012;28:1033–4.
44. Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 2012;19:455–77.
45. ANI Average Nucleotide Identity [Internet]. Available at: http://enve-omics.ce.gatech.edu/ani/. Accessed April 19, 2016.
46. Aziz RK, Bartels D, Best AA, et al. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 2008;9:75.
47. Gibson MK, Forsberg KJ, Dantas G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. ISME J 2014;9:207–16.
48. Eddy SR. Accelerated profile HMM searches. PLoS Comput Biol 2011;7:e1002195.
49. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25:2078–9.
50. Eren AM, Esen ÖC, Quince C, et al. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 2015;3:e1319.
51. Scholz M, Ward DV, Pasolli E, et al. Strain-level microbial epidemiology and population genomics from shotgun metagenomics. Nat Methods 2016;13:435–8.

52. Glucksberg H, Storb R, Fefer A, et al. Clinical manifestations of graft-versus-host disease in human recipients of marrow from HL-A-matched sibling donors. Transplantation 1974;18:295–304.
53. Scanlan PD, Shanahan F, Marchesi JR. Human methanogen diversity and incidence in healthy and diseased colonic groups using *mcrA* gene analysis. BMC Microbiol 2008;8:79.
54. Table S1-Table S7: Integrated meta-omic analyses of the gastrointestinal tract microbiome in patients undergoing allogeneic stem cell transplantation [Internet]. Available at: https://zenodo.org/record/268914. Accessed February 6, 2017.
55. Schmid C, Schleuning M, Ledderose G, Tischer J, Kolb HJ. Sequential regimen of chemotherapy, reduced-intensity conditioning for allogeneic stem-cell transplantation, and prophylactic donor lymphocyte transfusion in high-risk acute myeloid leukemia and myelodysplastic syndrome. J Clin Oncol 2005;23:5675–87.
56. Canani RB, Costanzo MD, Leone L, Pedata M, Meli R, Calignano A. Potential beneficial effects of butyrate in intestinal and extraintestinal diseases. World J Gastroenterol 2011;17:1519–28.
57. Zwielehner J, Lassl C, Hippe B, et al. Changes in human fecal microbiota due to chemotherapy analyzed by TaqMan-PCR, 454 sequencing and PCR-DGGE fingerprinting. PLoS One 2011;6:e28654.
58. Montassier E, Gastinne T, Vangay P, et al. Chemotherapy-driven dysbiosis in the intestinal microbiome. Aliment Pharmacol Ther 2015;42:515–28.
59. Furusawa Y, Obata Y, Fukuda S, et al. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature 2013;504:446–50.
60. Peng L, Li Z, Green RS, Holzman IR, Lin J. Butyrate enhances the intestinal barrier by facilitating tight junction assembly via activation of AMP-activated protein kinase. J Nutr 2009;139:1619–25.
61. Mathewson ND, Jenq R, Mathew AV, et al. Gut microbiome–derived metabolites modulate intestinal epithelial cell damage and mitigate graft-versus-host disease. Nat Immunol 2016;17:505–13.
62. Docampo MD, Auletta JJ, Jenq RR. The emerging influence of the intestinal microbiota during allogeneic hematopoietic cell transplantation: Control the gut and the body will follow. Biol Blood Marrow Transplant 2015;21:1360–6.
63. Dollive S, Chen YY, Grunberg S, et al. Fungi of the murine gut: Episodic variation and proliferation during antibiotic treatment. PLoS One 2013;8:e71806.
64. Biagi E, Zama D, Nastasi C, et al. Gut microbiota trajectory in pediatric patients undergoing hematopoietic SCT. Bone Marrow Transplant 2015;50:992–8.
65. Abreu MT, Peek RM. Gastrointestinal malignancy and the microbiome. Gastroenterology 2014;146:1534–46.
66. Perez-Chanona E, Jobin C. From promotion to management: the wide impact of bacteria on cancer and its treatment. Bioessays 2014;36:658–64.
67. Jiang W, Wu N, Wang X, et al. Dysbiosis gut microbiota associated with inflammation and impaired mucosal immune function in intestine of humans with non-alcoholic fatty liver disease. Sci Rep 2015;5.
68. Ubeda C, Bucci V, Caballero S, et al. Intestinal microbiota containing *Barnesiella* species cures vancomycin-resistant *Enterococcus faecium* colonization. Infect Immun 2013;81:965–73.
69. Webber MA, Piddock LJ. The importance of efflux pumps in bacterial antibiotic resistance. J Antimicrob Chemother 2003;51:9–11.
70. Varrette S, Bouvry P, Cartiaux H, Georgatos F. Management of an academic HPC cluster: the UL experience. Proc 2014 Int Conf High Perform Comput Simulation, HPCS 2014. 2014;959-967.

SUPPLEMENTARY FILE 1

Processing and assembly of metagenomic and metatranscriptomic data sets. Within IMP, the average depth of coverage of a gene or contig was determined both for the metagenome and the metatranscriptome, by calculating the average number of reads mapping to each nucleotide within a gene, respectively, to a contig. Here, gene expression was calculated as the ratio of the average MT depth of coverage to the average MG depth of coverage for individual genes.

Published human GIT microbiome MG and MT read data from 4 healthy individuals was obtained from the NCBI Sequence Read Archive of data sets [MG: SRX247379, SRX247391, SRX247401, SRX247405; MT: SRX247335, SRX247345, SRX247349, SRX247340].[1] Sequencing reads were processed using IMP version 1.2.1.[2] Data from the individuals "X310763260," "X316192082," "X317690558," and "X316701492" are referred to as the "reference healthy microbiome" (RHM).

Taxonomic affiliation of reconstructed population-level genomes. Taxonomic affiliation of population-level genomes was determined using complementary methods. Contigs forming the population-level genomes were first aligned to the NCBI nucleotide collection (nr/nt) database using the BLAST webservice,[3] with default parameters for megablast. The output was analyzed using the MEtaGenome ANalyzer (MEGAN version 5.10.5).[4] Whenever the *rpoB* gene could be recovered within a population-level genome, the closest neighbor (in terms of sequence identity) was determined in the nucleotide collection (nr/nt) database using the MOLE-BLAST web service.[5] In addition, AMPHORA2[6] was used to identify the taxonomic affiliation of up to 31 bacterial or 104 archaeal phylogenetic marker genes.

Reassembly. Population-level genomes were reassembled using all MG and MT reads mapping to the contigs of the population-level genomes with the same taxonomic assignment. Reassembly of all recruited reads was carried out using SPAdes[46] (version 3.5.0) using standard parameters. MG and MT reads were subsequently mapped to the contigs forming this reassembly to determine expression levels and variant density.

Sequence comparison of population-level genomes. The average nucleotide identity calculator[47] was used with standard settings to compare the reassembly from population-level genomes to publicly available reference genomes. A gene-wise protein sequence comparison of different population-level genomes was performed using the RAST server[48] using standard parameters.

REFERENCES

1. Franzosa EA, Morgan XC, Segata N, et al. Relating the metatranscriptome and metagenome of the human gut. Proc Natl Acad Sci U S A 2014;111:E2329–38.
2. Narayanasamy S, Jarosz Y, Muller EEL, et al. IMP: a pipeline for reproducible reference-independent integrated metagenomic and metatranscriptomic analyses. Genome Biol 2016;17:260.
3. Madden T. Chapter 16: the BLAST sequence analysis tool. In: The NCBI Handbook [internet]. 2002:1–15.
4. Huson D, Mitra S, Ruscheweyh H. Integrative analysis of environmental sequences using MEGAN4. Genome Res 2011;21:1552–60.
5. MOLE-BLAST webservice [Internet]. Available at: https://blast.ncbi.nlm.nih.gov/moleblast/moleblast.cgi. Accessed April 19, 2016.
6. Wu M, Scott AJ. Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. Bioinformatics 2012;28:1033–4.