UNIVERSITY OF SÃO PAULO
INSTITUTE OF CHEMISTRY

Graduate Program in Biological Sciences (Biochemistry)

VITOR HEIDRICH

Association of the oral microbiota dynamics with complications and outcomes in allogeneic hematopoietic stem cell transplant

Corrected version of the defended Thesis

São Paulo
Date of deposit at SPG:
13/02/2023

VITOR HEIDRICH

Associação da dinâmica da microbiota oral com complicações e desfechos do transplante alogênico de células-tronco hematopoiéticas

Tese apresentada ao Instituto de Química da Universidade de São Paulo para obtenção do Título de Doutor em Ciências no Programa: Ciências Biológicas (Bioquímica) Área: Bioquímica

Orientadora: Dra. Anamaria Aranha Camargo

São Paulo
2023

Universidade de São Paulo
Instituto de Química

"Associação da dinâmica da microbiota oral com complicações e desfechos do transplante alogênico de células-tronco hematopoiéticas"

VITOR HEIDRICH

Tese de Doutorado submetida ao Instituto de Química da Universidade de São Paulo como parte dos requisitos necessários à obtenção do grau de Doutor em Ciências obtido no Programa Ciências Biológicas (Bioquímica) - Área de Concentração: Bioquímica.

_____

Profa. Dra. Anamaria Aranha Camargo
(Orientadora e Presidente)

APROVADO(A) POR:

_____

Prof. Dr. João Carlos Setubal (por videoconferência)
IQ - USP

_____

Profa. Dra. Marcia Pinto Alves Mayer
ICB - USP

_____

Prof. Dr. Emmanuel Dias Neto (por videoconferência)
FAP-HACC

SÃO PAULO
18 de abril de 2023

Ficha Catalográfica elaborada eletronicamente pelo autor, utilizando o programa desenvolvido pela Seção Técnica de Informática do ICMC/USP e adaptado para a Divisão de Biblioteca e Documentação do Conjunto das Químicas da USP

Bibliotecária responsável pela orientação de catalogação da publicação:
Marlene Aparecida Vieira - CRB - 8/5562

H465a

Heidrich, Vitor
Association of the oral microbiota dynamics with complications and outcomes in allogeneic hematopoietic stem cell transplant / Vitor Heidrich. - São Paulo, 2023.
182 p.

Tese (doutorado) - Instituto de Química da Universidade de São Paulo. Departamento de Bioquímica.
Orientador: Camargo, Anamaria Aranha

1. Genomics. 2. Oral microbiome. 3. Hematopoietic cell transplant. 4. Hematological malignancies. 5. Biomarkers. I. T. II. Camargo, Anamaria Aranha, orientador.

AGRADECIMENTOS

Agradeço ao Instituto Ludwig de Pesquisa sobre o Câncer (ILPC), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, processo nº 13996-0/2018) pelo financiamento desta Tese.

Agradeço a todos os professores do Curso de Ciências Moleculares da USP, por me mostrarem diferentes perspectivas de carreira científica. Ao professor Paulo Nussenzveig, agradeço também pelas colheres de chá.

Agradeço à professora Daniela Sanchez Bassères, por ter aberto a primeira porta de um laboratório para mim. Agradeço também ao Sandro e à Thalita, por me ajudarem a dar os primeiros passos nele.

Agradeço aos pacientes e membros do corpo clínico do Hospital Sírio-Libanês, por tornarem este projeto possível. Em particular, pelas interações constantes que fizeram eu me fascinar pela hematologia e pela saúde oral, agradeço aos clínicos Celso, Vinícius e Eduardo, fundamentais para o caráter interdisciplinar desta Tese.

Agradeço aos pesquisadores e técnicos do Centro de Oncologia Molecular do Hospital Sírio-Libanês, pela colaboração e dedicação, que foram essenciais para a realização deste trabalho e para a minha formação.

Agradeço aos colegas de laboratório. Em especial, Julia e Franciele, que fizeram uma parte significativa do trabalho apresentado nesta Tese e me ensinaram coisas que eu nem sonhava em aprender quando entrei no Doutorado.

Agradeço à minha orientadora Anamaria Aranha Camargo. Levarei para sempre o que a Ana me ensinou e espero ser um orientador tão bom para alguém quanto ela foi para mim. Anos se passarão e seguirei incapaz de mensurar o impacto que ela teve na minha formação.

Agradeço aos meus meus familiares: meu irmão, minha cunhada, meus tios, minhas avós, e meu futuro afilhado Diego, por darem propósito a tudo o que faço.

Agradeço à minha noiva, Betina, pela compreensão e apoio nos momentos difíceis, e por compartilhar da minha alegria nos momentos de conquista. A jornada valeu a pena porque foi ao lado dela.

Por fim, os mais importantes, meus pais, Silvia e Gilmar, agradeço pelo carinho e pela educação que me deram. Não fosse por eles, não teria chegado onde cheguei.

O pensamento parece uma coisa à toa
Mas como é que a gente voa
Quando começa a pensar

Lupicínio Rodrigues

ABSTRACT

Heidrich, V. **Association of the oral microbiota dynamics with complications and outcomes in allogeneic hematopoietic stem cell transplant**. 182p. PhD Thesis - Graduate Program in Biological Sciences (Biochemistry). Institute of Chemistry, Universidade of São Paulo, São Paulo.

Allogeneic hematopoietic stem cell transplant (allo-HSCT) is a potentially curative therapy for several hematological disorders. Before stem cell infusion, recipients undergo a conditioning regimen with chemo/radiotherapy and immunosuppressants, requiring the use of antibiotics to treat and prevent infections. This regimen promotes drastic alterations in the recipient's gut microbiota, which have been associated with allo-HSCT complications and poor outcomes. Similar studies on the oral microbiota of allo-HSCT recipients are scarce and disregard the existence of distinct microbiotas within the oral cavity. Here, we used 16S rRNA gene sequencing to characterize the microbiota dynamics at three oral sites (gingival crevicular fluid, oral mucosa, and supragingival biofilm) during and after allo-HSCT. We used this data to associate the oral microbiota with an allo-HSCT toxicity (oral mucositis), allo-HSCT complications (graft-versus-host disease and bacterial infections), and allo-HSCT outcomes (progression-free survival, overall survival, and risk of underlying disease relapse, non-relapse death, and transplant-related death). In the first chapter, we analyzed the influence of the oral mucosa microbiota in the oral mucositis clinical course. We found taxa associated with higher oral mucositis grade (*Porphyromonas*), and lower time to oral mucositis healing (*Lactobacillus*). In the second chapter, we evaluated the association between supragingival biofilm microbiota and graft-versus-host disease risk. We identified taxa at preconditioning associated with higher (*Streptococcus* and *Corynebacterium*) and lower (*Veillonella*) risk of acute graft-versus-host disease, and observed that *Enterococcus faecalis* blooms during allo-HSCT were present in all patients developing this condition. In

the third chapter, we explored the association between oral mucosa microbiota and allo-HSCT outcomes. We noticed that preconditioning oral microbiota dysbiosis (low diversity or dominance by a single genus) was associated with poorer outcomes, such as shortened overall survival. Finally, in the fourth chapter, we analyzed samples from all oral sites. We observed that the microbiota of all three oral sites was damaged during allo-HSCT, which translated into a loss of differences between microbiota compositions of each site. Despite the loss of diversity and blooms of pathogenic genera observed during allo-HSCT (which preceded respiratory complications caused by the blooming bacteria in some cases), oral microbiotas were able to return to their initial state after engraftment, even though recovery levels varied between patients. After stratifying patients based on their ability to recover their preconditioning microbiota, we found that patients able to recover their oral mucosa microbiota composition showed earlier reconstitution of normal leukocyte counts in the bloodstream. Most notably, oral mucosa microbiota composition recovery was not associated with antibiotic usage and was an independent biomarker of better allo-HSCT outcomes. In summary, we identified clear patterns of dysbiosis in the oral microbiota during allo-HSCT. The oral microbiota of allo-HSCT recipients was associated with oral mucositis clinical course, allo-HSCT complications and allo-HSCT outcomes, highlighting the clinical value of tracking oral microbiota changes during allo-HSCT.

Keywords: oral microbiome; dysbiosis; bone marrow transplant; hematological malignancies; biomarkers; clinical outcomes

RESUMO

Heidrich, V. **Associação da dinâmica da microbiota oral com complicações e desfechos do transplante alogênico de células-tronco hematopoiéticas**. 182p. Tese de Doutorado - Programa de Pós-Graduação em Ciências Biológicas (Bioquímica). Instituto de Química, Universidade de São Paulo, São Paulo.

O transplante alogênico de células-tronco hematopoiéticas (TCTH-alo) é uma terapia potencialmente curativa para diversas doenças hematológicas. Antes da infusão das células-tronco, os receptores são submetidos a um regime de condicionamento com quimio/radioterapia e imunossupressores, exigindo o uso de antibióticos para tratar e prevenir infecções. Esse regime promove alterações drásticas na microbiota intestinal do receptor, que estão associadas com complicações e piores desfechos. Estudos similares da microbiota oral de receptores do TCTH-alo são escassos e desconsideram a existência de microbiotas distintas na cavidade oral. Neste estudo, através de sequenciamento do gene do RNA ribossomal 16S, caracterizamos a dinâmica da microbiota em três sítios orais (fluido crevicular gengival, mucosa oral e biofilme supragengival) durante e após o TCTH-alo. Utilizamos esses dados para associar a microbiota oral com uma toxicidade (mucosite oral), complicações (doença do enxerto-contra-hospedeiro e infecções bacterianas) e desfechos do TCTH-alo (sobrevida livre de progressão, sobrevida global, risco de recaída da doença de base, de morte não relacionada à recaída e de morte relacionada ao transplante). No primeiro capítulo, analisamos a influência da microbiota da mucosa oral no curso clínico de mucosite oral. Encontramos táxons associados com mucosite oral de maior grau (*Porphyromonas*) e menor tempo para resolução das lesões (*Lactobacillus*). No segundo capítulo, avaliamos a associação entre a microbiota do biofilme supragengival e o risco de doença do enxerto-contra-hospedeiro. Identificamos táxons no pré-condicionamento associados com maior (*Streptococcus* e *Corynebacterium*) e menor (*Veillonella*)

risco de doença do enxerto-contra-hospedeiro aguda, e observamos que todos os pacientes que apresentaram *blooms* de *Enterococcus faecalis* desenvolveram essa condição. No terceiro capítulo, exploramos a associação entre a microbiota da mucosa oral e desfechos do TCTH-alo. Notamos que disbiose (baixa diversidade ou dominância por um único gênero) da microbiota da mucosa oral do pré-condicionamento estava associada com piores desfechos, como menor sobrevida global. Finalmente, no quarto capítulo, analisamos amostras dos três sítios orais. Observamos que a microbiota dos três sítios foi danificada durante o TCTH-alo, o que traduziu numa perda das diferenças entre microbiotas de cada sítio. Apesar da queda de diversidade e de *blooms* de gêneros patogênicos observada durante o TCTH-alo (que precederam complicações respiratórias causadas pelas bactérias envolvidas nos *blooms* em alguns casos), as microbiotas foram capazes de retornar para seu estado inicial após a enxertia, embora os níveis de recuperação tenham variado entre os pacientes. Após classificar os pacientes com base na capacidade de recuperação de sua microbiota do pré-condicionamento, encontramos que pacientes que recuperaram a composição da microbiota da mucosa oral demonstraram reconstituição mais precoce da contagem normal de leucócitos na corrente sanguínea. Notavelmente, a recuperação da composição da microbiota da mucosa oral não apresentou associação com o uso de antibióticos e foi um biomarcador independente de melhores desfechos. De modo geral, identificamos claros padrões de disbiose na microbiota oral durante o TCTH-alo. A microbiota oral do recipiente do TCTH-alo mostrou associação com o curso clínico de mucosite oral, complicações do TCTH-alo e desfechos do TCTH-alo, salientando o valor clínico de rastrear mudanças na microbiota oral durante o TCTH-alo.

Palavras-chave: microbioma oral; disbiose; transplante de medula óssea; neoplasias hematológicas; biomarcadores; desfechos clínicos

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Allogeneic hematopoietic stem cell transplant

1.1.1. Indications and procedure

Allogeneic hematopoietic stem cell transplant (allo-HSCT) is a potentially curative treatment for several hematological disorders. Copelan (2006) lists 19 diseases commonly treated by allo-HSCT, out of which 10 are malignant blood disorders and 9 relate to nonmalignant hematopoietic defects, mainly different types of anemia (Copelan, 2006). This vast range of applications made allo-HSCT an extremely important therapeutic modality, with more than 38,000 allo-HSCTs performed worldwide yearly (Niederwieser et al., 2022).

The goal of allo-HSCT is twofold: eradicate malignant/defective blood cells and replace an abnormal hematopoietic system (Jenq & van den Brink, 2010). The former is accomplished by the conditioning regimen patients undergo before stem cell infusion and the latter by the infusion of hematopoietic stem cells derived from a healthy donor. A more detailed description of the procedure is possible by breaking it down into the following four phases (Fig. 1).

I. Conditioning regimen: course of chemotherapeutic drugs (may also include radiotherapy) that recipients undergo in preparation for the transplantation procedure. Its goal is to reduce disease burden and allow sufficient immunoablation so that donor stem cells can engraft after infusion (Gyurkocza & Sandmaier, 2014). There are three different types of conditioning regimens (myeloablative, reduced intensity, and non-myeloablative), which provoke different levels of cytopenia (Bacigalupo et al., 2009). The conditioning regimen used for a given recipient is selected based on several variables, such as

recipient age, donor age, and phase of underlying disease, among others (Bacigalupo et al., 2009). Besides chemo/radiotherapy, conditioning regimens often include immunosuppressant drugs to prevent graft-vs-host disease (GVHD) and prophylactic antibiotics (Welniak et al., 2007; Lehrnbecher et al., 2020). In addition to prophylaxis, the use of antibiotics is commonly necessary during and after the conditioning regimen to treat opportunistic infections that may arise throughout the hospitalization period (Omrani & Almaghrabi, 2017).

II. Stem cell infusion: intravenous infusion of hematopoietic stem cells derived from a healthy donor. Even though the ideal donor is a recipient's relative with identical human leukocyte antigen (HLA), donors can also be HLA-matched unrelated, HLA-mismatched related (includes the so-called haploidentical donor type, in which exactly half of HLA alleles are matched), and even HLA-mismatched unrelated, but with varying levels of success rates per donor type (Kekre & Antin, 2014). Possible sources of stem cells are the donor's bone marrow and peripheral blood (Russell et al., 1993). Additionally, umbilical cord blood can be used as a source of stem cells that do not require HLA-matching due to the extremely low number of T-cells in cord blood, but the limited hematopoietic cell doses in cord grafts is a barrier to their wider application in allo-HSCT (Barker et al., 2002).

III. Aplasia: due to the immunoablative effect of the conditioning regimen and the hematopoietic incompetence of the newly received stem cells, infusion is followed by a period of bone marrow aplasia, when the neutrophil blood count reaches its nadir (typically defined as starting in the first day of neutrophil blood count $<0.5 \times 10^3$/uL) (Solans et al., 2020).

IV. Engraftment: bone marrow function is restored 2–4 weeks after stem cell infusion in a successful allo-HSCT, and, at this point, the recipient is considered engrafted (Hatzimichael & Tuthill, 2010). Engraftment occurs in the first of three consecutive days of neutrophil blood count >0.5 × 10^9/uL (Wolff, 2002). In the post-engraftment period, the recipient is still watched closely, given the high mortality risk in the first 100 days following allo-HSCT (Styczyński et al., 2020).

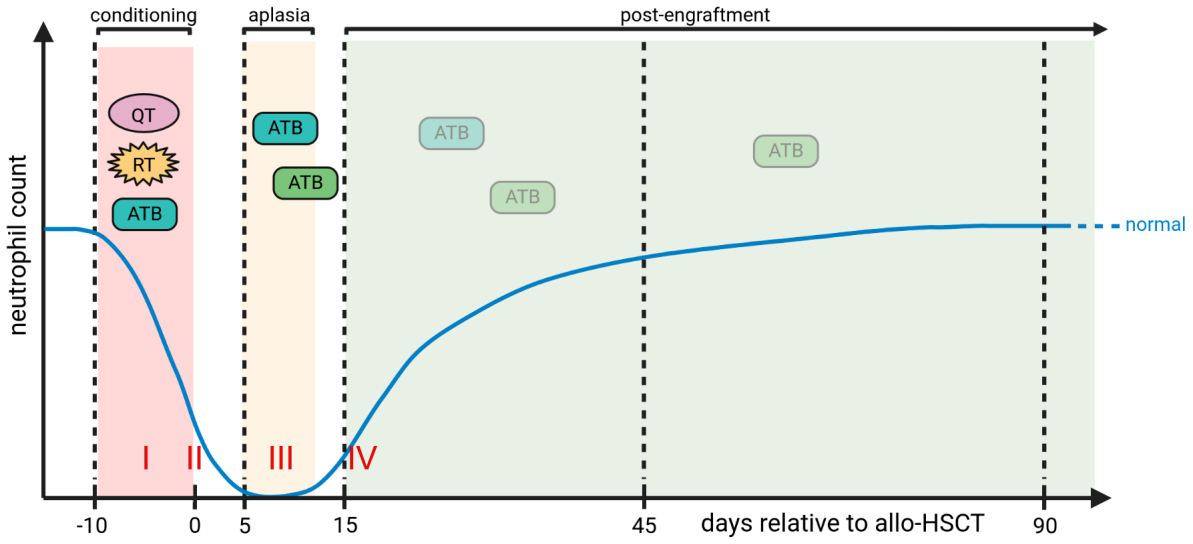


Fig. 1: Overview of the transplantation procedure. Illustration of the allo-HSCT procedure with neutrophil blood count in relation to days from infusion as reference. The bottom of the graph shows in red the four phases of the treatment mentioned above in the text. Day 0 refers to the day of stem cell infusion (II). QT, chemotherapy; RT, radiotherapy; ATB, antibiotics.

1.1.2. Outcomes, complications, and toxicities

Even though the clinical benefits of allo-HSCT have been thoroughly demonstrated in several scenarios (Devillier et al., 2022; Gonsalves et al., 2019; Scheinberg & Young, 2012), it remains associated with ~46% mortality rates 5-years after transplant (Styczyński et al., 2020). This is due to high rates of transplant-related

mortality and underlying disease relapse, with the latter being the leading death cause (43%) following allo-HSCT (Styczyński et al., 2020).

Transplant-related mortality concerns deaths caused by allo-HSCT complications rather than disease relapse. As such, most transplant-related fatalities occur within 100 days after allo-HSCT (Bunin et al., 2008; Cornelissen et al., 2001; Miano et al., 2007; Sierra et al., 2000; Styczyński et al., 2020). Infections and GVHD are the leading causes of transplant-related deaths among the potentially fatal complications associated with allo-HSCT (Styczyński et al., 2020).

Infections during allo-HSCT can be caused by different types of organisms (bacteria, viruses, fungi, protozoa, and helminths) and can affect various anatomical sites, including skin, bloodstream (bacteremia), and respiratory tract (Centers for Disease Control and Prevention, 2000). A scenario of extreme immunosuppression, together with the increasing presence of multidrug-resistant bacteria in the hospital environment (Wisplinghoff et al., 2003), makes these infections ever more challenging to treat, such that ~24% of deaths 5-years after allo-HSCT are due to infections (Styczyński et al., 2020).

While opportunistic infections may arise due to the overall immunosuppressed state of the recipient, GVHD relates to the overactivation of donor's reactive immune cells against the recipient's tissues (Ferrara & Chaudhry, 2018; Yehudai-Ofir et al., 2020). As a result, several of the recipient's organs may be damaged, leading to a potentially lethal (~19% 5-year mortality rate (Styczyński et al., 2020)) clinical syndrome classified as acute (aGVHD) or chronic (cGVHD), depending on the donor reactive immune cells involved, and the timing of symptoms' presentation (Toubai et al., 2008).

aGVHD manifestation usually occurs in the first 100 days following allo-HSCT due to overactivation of the donor's reactive T-cells, whilst cGVHD usually happens several months after allo-HSCT and with the participation of overactivated B-cells, with an immune profile resembling autoimmune diseases (Ferrara & Chaudhry, 2018; Yehudai-Ofir et al., 2020). Another difference between aGVHD and cGVHD relates to the organs affected. While aGVHD affects the skin (maculopapular rash), the liver (hyperbilirubinemia), and the gastrointestinal tract (anorexia, diarrhea, and abdominal pain), cGVHD affects joints (impaired ambulation) and eyes (dry eyes) (Lee et al., 2003; Schoemans et al., 2018).

Besides poor outcomes and complications, similar to other treatments based on the use of chemo and radiotherapy (De Ruysscher et al., 2019; Livshits et al., 2014), allo-HSCT is also associated with treatment toxicities, including cardiac arrhythmias and idiopathic pneumonia syndrome (López-Fernández et al., 2021; Panoskaltsis-Mortari et al., 2011). Complications caused by treatment toxicities are another source of transplant-related deaths.

Oral mucositis is an allo-HSCT toxicity with high incidence (60–85%) (Chaudhry et al., 2016; Villa & Sonis, 2015). Oral mucositis lesions, which usually show their first signs 3–4 days after infusion (Villa & Sonis, 2015), present as ulcers with reddish borders covered by a white pseudomembrane colonized by bacteria, which can cause further tissue damage via endotoxins release (Wysocka-Słowik et al., 2021). Patients cannot eat solids in the presence of severe oral mucositis lesions, demanding enteral or parenteral nutrition support (Chaudhry et al., 2016; Elad et al., 2020). Oral mucositis can also impact allo-HSCT clinical course by causing chemotherapy discontinuation, increase in analgesics usage, treatment delays, and prolonged hospitalization periods (Cinausero et al., 2017).

1.2. Human microbiota and allo-HSCT

1.2.1. Human microbiota

There are at least as many microbes as human cells in the human body (Sender et al., 2016). These microbes – viruses, archaea, bacteria, fungi, and protozoans – organize in communities (microbiotas) with different compositions in different body sites (Human Microbiome Project Consortium, 2012). The microbiotas that populate our bodies (collectively referred to as the human microbiota) possess several times more genes than the human genome (Gilbert et al., 2018). The collection of microbial genomes in a microbiota is often referred to as a metagenome (Berg et al., 2020).

Metagenomics refers to the simultaneous assessment of the genome contents of all organisms present in a sample (in opposition to isolate genome sequencing) (Quince et al., 2017). A metagenomics workflow comprises the collection of a sample containing a microbial community (e.g., feces), DNA extraction, and sequencing, totally bypassing the need to use cell culture assays to identify microbes (Quince et al., 2017). Since most human microbiota species are considered 'unculturable', the role of metagenomics in human microbiota studies is paramount (Nayfach et al., 2019).

Because of this, human microbiota studies gained traction only with the advent of cost-effective next-generation sequencing. There are mainly two sequencing strategies used in microbiota studies: shotgun metagenomic sequencing and amplicon sequencing (Franzosa et al., 2015). Shotgun metagenomic sequencing (also referred to simply as metagenomics) involves sequencing of all DNA content derived from a microbiota sample, which allows not only taxonomic profiling but also functional profiling, as the encoded functional potential can be determined by

analysis of the genes present in the community (Quince et al., 2017). This technique is considerably more expensive than amplicon sequencing, in which, instead of whole genomes, only a specific taxonomically informative gene is sequenced after PCR amplification using universal primers (Lundberg et al., 2013). Because amplicon sequencing can reveal taxonomic compositions by sequencing a single gene, this technique is also referred to as metabarcoding, metataxonomics, or metagenetics (Yap et al., 2022). For bacterial taxonomic profiling, the loci typically evaluated through amplicon sequencing are one or more hypervariable regions of the 16S rRNA gene (Tringe & Hugenholtz, 2008).

Most human microbiota studies so far have focused on the gut microbiota, the most abundant and diverse microbiota inhabiting our bodies (Human Microbiome Project Consortium, 2012). In the last two decades, the gut microbiota has been linked to multiple aspects of human health, ranging from type 1 diabetes to depression and cancer (de Vos et al., 2022; Li et al., 2022; Thomas et al., 2019; Yuan et al., 2022). Other microbiotas, such as the bladder and oral microbiota, also have demonstrated an impact on human health (Perez-Carrasco et al., 2021; Y. Zhang et al., 2018), but remain relatively understudied.

1.2.2. Oral microbiota

The oral microbiota is the second most abundant and diverse microbiota in the human body, comprising over 700 bacterial species (Human Microbiome Project Consortium, 2012; Palmer, 2014). This is partially due to the organization of oral microbes in biofilms, as oral biofilms create oxygen availability gradients that allow colonization by both anaerobic and aerobic bacteria (Mark Welch et al., 2020). Another contributor to oral microbiota diversity is the anatomical complexity of the oral cavity,

which presents particular topography, moisture levels, and tissue type (shedding vs. non-shedding) at each oral compartment, allowing colonization by microbes adapted to very different environments (Mark Welch et al., 2020; Proctor & Relman, 2017). Consequently, the oral microbiota is not a single entity. Instead, each oral site hosts distinct microbial communities, with main compositional differences existing between mucosa-associated and teeth-associated microbiotas (Segata et al., 2012).

Similarly to the gut microbiota, oral microbiotas play an essential role in regulating human health (Tuganbaev et al., 2022). The pathogenesis of oral diseases often has a microbial dimension (M. Zhang et al., 2022), including the role of supragingival biofilm in dental caries and subgingival bacteria in periodontal disease (Slots, 1979; Takahashi & Nyvad, 2011). But oral microbiotas can also influence systemic health, as recently highlighted in studies associating them with rheumatoid arthritis, Alzheimer's disease, and colorectal cancer (Dominy et al., 2019; Flemer et al., 2018; Tong et al., 2019).

1.2.3. Microbiota in the allo-HSCT setting

In the last decade, studies have shed light on several associations between the bacterial composition of the gut microbiota and the clinical course of allo-HSCT. They showed the gut microbiota undergoes remarkable changes during allo-HSCT, including loss of bacterial diversity and blooms of potentially pathogenic species (Shono & van den Brink, 2018). The extent of these alterations are associated with allo-HSCT complications and outcomes (Peled et al., 2020; Stein-Thoeringer et al., 2019).

Peled et al. (2020) showed that a low-diversity gut microbiota at peri-engraftment is associated with lower overall survival after allo-HSCT and a higher risk of transplant-related death, including deaths attributable to GVHD (Peled

et al., 2020). Stein-Thoeringer et al. (2019) analyzed the same data to show that this low-diversity scenario during allo-HSCT was often associated with gut dominance by enterococci, which was also associated with lower overall survival and increased incidence of GVHD (Stein-Thoeringer et al., 2019). An additional study suggests it is possible to predict aGVHD based on pre-allo-HSCT gut microbiota compositions (Liu et al., 2017).

Microbiota damage during allo-HSCT does not occur exclusively in the gut, with lung and oral microbiota also being affected (Sen & Thummer, 2022). Still, little is known about the association between microbiotas other than the gut and allo-HSCT clinical course (toxicities, complications, and outcomes). This would be especially interesting for the oral microbiota because, even though its abundance and diversity are second to the gut, the oral cavity is easy to sample and allows assessing bacterial presence at different compartments, making it the ideal place to look for allo-HSCT biomarkers.

1.3. Overview of the work described in this Thesis

We herein describe our findings regarding the evaluation of the oral microbiota in 31 allo-HSCT recipients treated at Hospital Sírio-Libanês between January 2016 and April 2018 (Fig. 2). To have an anatomically-aware view of the oral microbiota, we collected samples from three oral sites: gingival crevicular fluid (GCF), oral mucosa (OM), and supragingival biofilm (SB). To evaluate the dynamics of the oral microbiota during and after allo-HSCT, we collected samples at five timepoints: preconditioning (P), aplasia (A), engraftment (E), 30 days after engraftment (E30), and 75 days after engraftment (E75). Microbiotas were profiled by sequencing the V3V4 region of the taxonomic marker 16S rRNA gene. With this

data, we aimed to evaluate the association between the oral microbiota dynamics and allo-HSCT outcomes, complications, and toxicities.

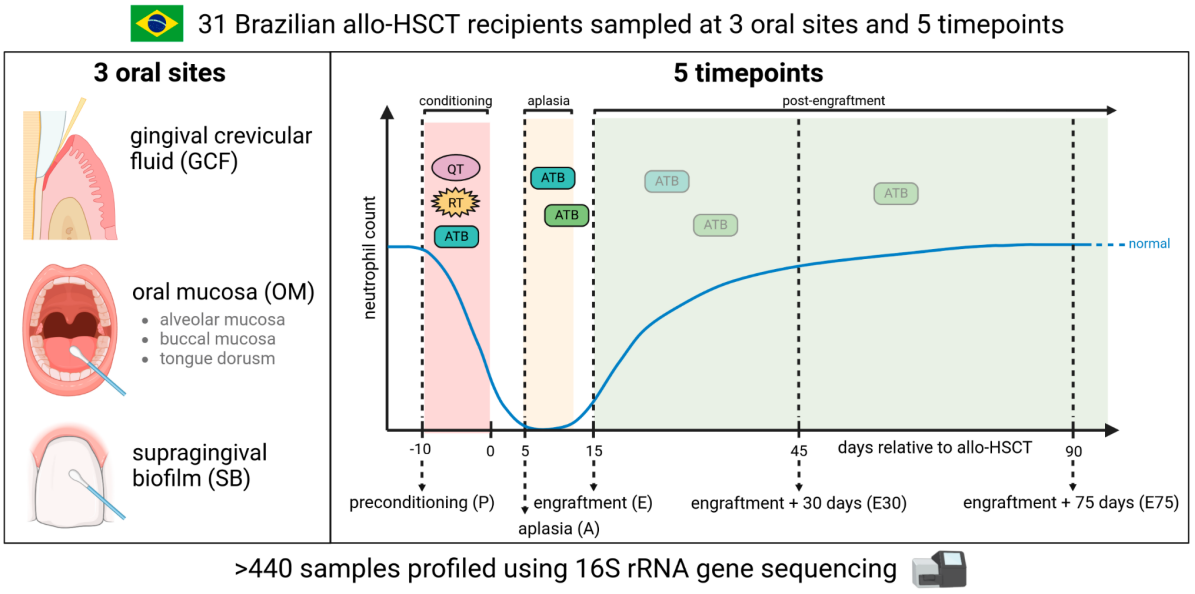


Fig. 2: Overview of the sample collection design. Overall, >440 samples spanning 3 oral sites and 5 allo-HSCT timepoints were collected and sequenced for microbiota profiling by 16S rRNA sequencing.

In Chapter 1 (DOI: 10.1038/s41598-022-21775-3; Appendix A: Chapter 1 supplementary tables and figures), we evaluated the association between OM microbiota and oral mucositis, a toxicity of the conditioning regimen affecting the OM (Chaudhry et al., 2016). Besides evaluating preconditioning samples of all patients, we evaluated additional OM samples collected from patients who developed oral mucositis at the onset and resolution of ulcerated lesions. This allowed us not only to predict the risk of oral mucositis based on preconditioning microbiotas but also to evaluate the impact of the OM microbiota in the clinical course of oral mucositis.

In Chapter 2 (DOI: 10.3389/fimmu.2021.692225; Appendix B: Chapter 2 supplementary tables and figures), we characterized SB microbiota damage caused by allo-HSCT. In addition, because SB bacteria can interact with host cells to modulate immune homeostasis (Moutsopoulos & Konkel, 2018), we evaluated the

potential of SB microbiota to predict the risk of aGVHD, a potentially lethal complication of allo-HSCT (Ferrara & Chaudhry, 2018). Since our goal was to predict aGVHD based on the information available before the completion of the transplantation procedure, we analyzed samples collected from preconditioning to engraftment.

In Chapter 3 (DOI: 10.1038/s41598-021-96939-8; Appendix C: Chapter 3 supplementary tables and figures), we explored OM microbiota damage caused by allo-HSCT and the potential of the OM microbiota to predict clinical outcomes of allo-HSCT. Similarly to Chapter 2, we assessed samples only up to engraftment, with a focus on the preconditioning timepoint.

In Chapter 4 (DOI: 10.1101/2022.11.18.22282520; Appendix D: Chapter 4 supplementary tables and figures; Appendix E: Chapter 4 supplementary methods; Appendix F: Chapter 4 antibiotics usage timeline), we did a comprehensive analysis of samples collected from all tree oral sites at all five timepoints. This allowed us to study the microbiota dynamics at different oral sites during and after allo-HSCT, as well as microbiota recovery patterns after allo-HSCT and its associations with clinical outcomes. Additionally, we evaluated oral microbiota dynamics and recovery under the light of extensive clinical metadata (antibiotics usage, blood cell counts, and the occurrence of bacterial infections during allo-HSCT) to uncover factors associated with microbiota recovery and better outcomes.

Finally, we include as attachments two additional reports related to this Thesis. In Attachment A (DOI: 10.3390/app112311473), we reported the implementation and expansion (through supplementary tools) of a previously proposed (Beule & Karlovsky, 2020) microbiome normalization algorithm (Scaling with Ranked Subsampling). The use of this normalization method was crucial to allow the

inclusion of relatively low-depth samples in the analysis, especially in Chapter 4. In attachment B (DOI: 10.1186/s12903-023-02777-7), we report, as far as we know, the first case to date of post-allo-HSCT gingival actinomycosis, a bacterial infection caused by some *Actinomyces* species (Valour et al., 2014), which afflicted one of the patients of our cohort. We showed how tracking the oral microbiota dynamics during allo-HSCT can aid in the diagnosis and management of aggressive local infections.

1.4. References

Bacigalupo, A., Ballen, K., Rizzo, D., Giralt, S., Lazarus, H., Ho, V., Apperley, J., Slavin, S., Pasquini, M., Sandmaier, B. M., Barrett, J., Blaise, D., Lowski, R., & Horowitz, M. (2009). Defining the intensity of conditioning regimens: Working definitions. *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation*, *15*(12), 1628–1633. https://doi.org/10.1016/j.bbmt.2009.07.004

Barker, J. N., Krepski, T. P., DeFor, T. E., Davies, S. M., Wagner, J. E., & Weisdorf, D. J. (2002). Searching for unrelated donor hematopoietic stem cells: Availability and speed of umbilical cord blood versus bone marrow. *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation*, *8*(5), 257–260. https://doi.org/10.1053/bbmt.2002.v8.pm12064362

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M. C. C., Charles, T., ... Schloter, M. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, *8*, 1-22. https://doi.org/10.1186/s40168-020-00875-0

Beule, L., & Karlovsky, P. (2020). Improved normalization of species count data in ecology by scaling with ranked subsampling (SRS): Application to microbial communities. *PeerJ*, *8*, e9593. https://doi.org/10.7717/peerj.9593

Bunin, N. J., Davies, S. M., Aplenc, R., Camitta, B. M., DeSantes, K. B., Goyal, R. K., Kapoor, N., Kernan, N. A., Rosenthal, J., Smith, F. O., & Eapen, M. (2008). Unrelated donor bone marrow transplantation for children with acute myeloid leukemia beyond first remission or refractory to chemotherapy. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *26*(26), 4326–4332. https://doi.org/10.1200/JCO.2008.16.4442

Centers for Disease Control and Prevention, Infectious Diseases Society of America, &

American Society of Blood and Marrow Transplantation. (2000). Guidelines for preventing opportunistic infections among hematopoietic stem cell transplant recipients. *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation*, *6*(6a), 7–83. https://doi.org/10.1016/s1083-8791(00)70002-4

Chaudhry, H. M., Bruce, A. J., Wolf, R. C., Litzow, M. R., Hogan, W. J., Patnaik, M. S., Kremers, W. K., Phillips, G. L., & Hashmi, S. K. (2016). The Incidence and Severity of Oral Mucositis among Allogeneic Hematopoietic Stem Cell Transplantation Patients: A Systematic Review. *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation*, *22*(4), 605–616. https://doi.org/10.1016/j.bbmt.2015.09.014

Cinausero, M., Aprile, G., Ermacora, P., Basile, D., Vitale, M. G., Fanotto, V., Parisi, G., Calvetti, L., & Sonis, S. T. (2017). New Frontiers in the Pathobiology and Treatment of Cancer Regimen-Related Mucosal Injury. *Frontiers in Pharmacology*, *8*, 354. https://doi.org/10.3389/fphar.2017.00354

Copelan, E. A. (2006). Hematopoietic stem-cell transplantation. *The New England Journal of Medicine*, *354*(17), 1813–1826. https://doi.org/10.1056/NEJMra052638

Cornelissen, J. J., Carston, M., Kollman, C., King, R., Dekker, A. W., Löwenberg, B., & Anasetti, C. (2001). Unrelated marrow transplantation for adult patients with poor-risk acute lymphoblastic leukemia: Strong graft-versus-leukemia effect and risk factors determining outcome. *Blood*, *97*(6), 1572–1577. https://doi.org/10.1182/blood.v97.6.1572

De Ruysscher, D., Niedermann, G., Burnet, N. G., Siva, S., Lee, A. W. M., & Hegi-Johnson, F. (2019). Radiotherapy toxicity. *Nature Reviews. Disease Primers*, *5*(1), 13. https://doi.org/10.1038/s41572-019-0064-5

de Vos, W. M., Tilg, H., Van Hul, M., & Cani, P. D. (2022). Gut microbiome and health: Mechanistic insights. *Gut*, *71*(5), 1020–1032. https://doi.org/10.1136/gutjnl-2021-326789

Devillier, R., Forcade, E., Garnier, A., Guenounou, S., Thepot, S., Guillerm, G., Ceballos, P., Hicheri, Y., Dumas, P.-Y., Peterlin, P., Hunault-Berger, M., Béné, M.-C., Bouvier, A., Chevallier, P., Blaise, D., Vey, N., Pigneux, A., Récher, C., & Huynh, A. (2022). In-depth time-dependent analysis of the benefit of allo-HSCT for elderly patients with CR1 AML: A FILO study. *Blood Advances*, *6*(6), 1804–1812. https://doi.org/10.1182/bloodadvances.2021004435

Dominy, S. S., Lynch, C., Ermini, F., Benedyk, M., Marczyk, A., Konradi, A., Nguyen, M.,

Haditsch, U., Raha, D., Griffin, C., Holsinger, L. J., Arastu-Kapur, S., Kaba, S., Lee, A., Ryder, M. I., Potempa, B., Mydel, P., Hellvard, A., Adamowicz, K., … Potempa, J. (2019). Porphyromonas gingivalis in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. *Science Advances*, *5*(1), eaau3333. https://doi.org/10.1126/sciadv.aau3333

Elad, S., Cheng, K. K. F., Lalla, R. V., Yarom, N., Hong, C., Logan, R. M., Bowen, J., Gibson, R., Saunders, D. P., Zadik, Y., Ariyawardana, A., Correa, M. E., Ranna, V., Bossi, P., & Mucositis Guidelines Leadership Group of the Multinational Association of Supportive Care in Cancer and International Society of Oral Oncology (MASCC/ISOO). (2020). MASCC/ISOO clinical practice guidelines for the management of mucositis secondary to cancer therapy. *Cancer*, *126*(19), 4423–4431. https://doi.org/10.1002/cncr.33100

Ferrara, J. L. M., & Chaudhry, M. S. (2018). GVHD: Biology matters. *Blood Advances*, *2*(22), 3411–3417. https://doi.org/10.1182/bloodadvances.2018020214

Flemer, B., Warren, R. D., Barrett, M. P., Cisek, K., Das, A., Jeffery, I. B., Hurley, E., O'Riordain, M., Shanahan, F., & O'Toole, P. W. (2018). The oral microbiota in colorectal cancer is distinctive and predictive. *Gut*, *67*(8), 1454–1463. https://doi.org/10.1136/gutjnl-2017-314814

Franzosa, E. A., Hsu, T., Sirota-Madi, A., Shafquat, A., Abu-Ali, G., Morgan, X. C., & Huttenhower, C. (2015). Sequencing and beyond: Integrating molecular "omics" for microbial community profiling. *Nature Reviews. Microbiology*, *13*(6), 360–372. https://doi.org/10.1038/nrmicro3451

Gilbert, J. A., Blaser, M. J., Caporaso, J. G., Jansson, J. K., Lynch, S. V., & Knight, R. (2018). Current understanding of the human microbiome. *Nature Medicine*, *24*(4), 392–400. https://doi.org/10.1038/nm.4517

Gonsalves, W. I., Buadi, F. K., Ailawadhi, S., Bergsagel, P. L., Chanan Khan, A. A., Dingli, D., Dispenzieri, A., Fonseca, R., Hayman, S. R., Kapoor, P., Kourelis, T. V., Lacy, M. Q., Larsen, J. T., Muchtar, E., Reeder, C. B., Sher, T., Stewart, A. K., Warsame, R., Go, R. S., … Roy, V. (2019). Utilization of hematopoietic stem cell transplantation for the treatment of multiple myeloma: A Mayo Stratification of Myeloma and Risk-Adapted Therapy (mSMART) consensus statement. *Bone Marrow Transplantation*, *54*(3), 353–367. https://doi.org/10.1038/s41409-018-0264-8

Gyurkocza, B., & Sandmaier, B. M. (2014). Conditioning regimens for hematopoietic cell transplantation: One size does not fit all. *Blood*, *124*(3), 344–353. https://doi.org/10.1182/blood-2014-02-514778

Hatzimichael, E., & Tuthill, M. (2010). Hematopoietic stem cell transplantation. *Stem Cells and Cloning: Advances and Applications*, *3*, 105–117. https://doi.org/10.2147/SCCAA.S6815

Human Microbiome Project Consortium. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, *486*(7402), 207–214. https://doi.org/10.1038/nature11234

Jenq, R. R., & van den Brink, M. R. M. (2010). Allogeneic haematopoietic stem cell transplantation: Individualized stem cell and immune therapy of cancer. *Nature Reviews. Cancer*, *10*(3), 213–221. https://doi.org/10.1038/nrc2804

Kekre, N., & Antin, J. H. (2014). Hematopoietic stem cell transplantation donor sources in the 21st century: Choosing the ideal donor when a perfect match does not exist. *Blood*, *124*(3), 334–343. https://doi.org/10.1182/blood-2014-02-514760

Lee, S. J., Vogelsang, G., & Flowers, M. E. D. (2003). Chronic graft-versus-host disease. *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation*, *9*(4), 215–233. https://doi.org/10.1053/bbmt.2003.50026

Lehrnbecher, T., Fisher, B. T., Phillips, B., Alexander, S., Ammann, R. A., Beauchemin, M., Carlesse, F., Castagnola, E., Davis, B. L., Dupuis, L. L., Egan, G., Groll, A. H., Haeusler, G. M., Santolaya, M., Steinbach, W. J., van de Wetering, M., Wolf, J., Cabral, S., Robinson, P. D., & Sung, L. (2020). Guideline for Antibacterial Prophylaxis Administration in Pediatric Cancer and Hematopoietic Stem Cell Transplantation. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *71*(1), 226–236. https://doi.org/10.1093/cid/ciz1082

Li, D., Liu, R., Wang, M., Peng, R., Fu, S., Fu, A., Le, J., Yao, Q., Yuan, T., Chi, H., Mu, X., Sun, T., Liu, H., Yan, P., Wang, S., Cheng, S., Deng, Z., Liu, Z., Wang, G., … Liu, T. (2022). 3β-Hydroxysteroid dehydrogenase expressed by gut microbes degrades testosterone and is linked to depression in males. *Cell Host & Microbe*, *30*(3), 329-339.e5. https://doi.org/10.1016/j.chom.2022.01.001

Liu, C., Frank, D. N., Horch, M., Chau, S., Ir, D., Horch, E. A., Tretina, K., van Besien, K., Lozupone, C. A., & Nguyen, V. H. (2017). Associations between acute gastrointestinal GvHD and the baseline gut microbiota of allogeneic hematopoietic stem cell transplant recipients and donors. *Bone Marrow Transplantation*, *52*(12), 1643–1650. https://doi.org/10.1038/bmt.2017.200

Livshits, Z., Rao, R. B., & Smith, S. W. (2014). An approach to chemotherapy-associated toxicity. *Emergency Medicine Clinics of North America*, *32*(1), 167–203.

https://doi.org/10.1016/j.emc.2013.09.002

López-Fernández, T., Vadillo, I. S., de la Guía, A. L., & Barbier, K. H. (2021). Cardiovascular Issues in Hematopoietic Stem Cell Transplantation (HSCT). *Current Treatment Options in Oncology*, *22*(6), 51. https://doi.org/10.1007/s11864-021-00850-3

Lundberg, D. S., Yourstone, S., Mieczkowski, P., Jones, C. D., & Dangl, J. L. (2013). Practical innovations for high-throughput amplicon sequencing. *Nature Methods*, *10*(10), 999–1002. https://doi.org/10.1038/nmeth.2634

Mark Welch, J. L., Ramírez-Puebla, S. T., & Borisy, G. G. (2020). Oral Microbiome Geography: Micron-Scale Habitat and Niche. *Cell Host & Microbe*, *28*(2), 160–168. https://doi.org/10.1016/j.chom.2020.07.009

Miano, M., Labopin, M., Hartmann, O., Angelucci, E., Cornish, J., Gluckman, E., Locatelli, F., Fischer, A., Egeler, R. M., Or, R., Peters, C., Ortega, J., Veys, P., Bordigoni, P., Iori, A. P., Niethammer, D., Rocha, V., Dini, G., & Paediatric Diseases Working Party of the European Group for Blood and Marrow Transplantation. (2007). Haematopoietic stem cell transplantation trends in children over the last three decades: A survey by the paediatric diseases working party of the European Group for Blood and Marrow Transplantation. *Bone Marrow Transplantation*, *39*(2), 89–99. https://doi.org/10.1038/sj.bmt.1705550

Moutsopoulos, N. M., & Konkel, J. E. (2018). Tissue-Specific Immunity at the Oral Mucosal Barrier. *Trends in Immunology*, *39*(4), 276–287. https://doi.org/10.1016/j.it.2017.08.005

Nayfach, S., Shi, Z. J., Seshadri, R., Pollard, K. S., & Kyrpides, N. C. (2019). New insights from uncultivated genomes of the global human gut microbiome. *Nature*, *568*(7753), 505–510. https://doi.org/10.1038/s41586-019-1058-x

Niederwieser, D., Baldomero, H., Bazuaye, N., Bupp, C., Chaudhri, N., Corbacioglu, S., Elhaddad, A., Frutos, C., Galeano, S., Hamad, N., Hamidieh, A. A., Hashmi, S., Ho, A., Horowitz, M. M., Iida, M., Jaimovich, G., Karduss, A., Kodera, Y., Kröger, N., … Saber, W. (2022). One and a half million hematopoietic stem cell transplants: Continuous and differential improvement in worldwide access with the use of non-identical family donors. *Haematologica*, *107*(5), 1045–1053. https://doi.org/10.3324/haematol.2021.279189

Omrani, A. S., & Almaghrabi, R. S. (2017). Complications of hematopoietic stem cell transplantation: Bacterial infections. *Hematology/Oncology and Stem Cell Therapy*, *10*(4), 228–232. https://doi.org/10.1016/j.hemonc.2017.05.018

Palmer, R. J. (2014). Composition and development of oral bacterial communities.

Periodontology 2000, *64*(1), 20–39. https://doi.org/10.1111/j.1600-0757.2012.00453.x

Panoskaltsis-Mortari, A., Griese, M., Madtes, D. K., Belperio, J. A., Haddad, I. Y., Folz, R. J., Cooke, K. R., & American Thoracic Society Committee on Idiopathic Pneumonia Syndrome. (2011). An official American Thoracic Society research statement: Noninfectious lung injury after hematopoietic stem cell transplantation: idiopathic pneumonia syndrome. *American Journal of Respiratory and Critical Care Medicine*, *183*(9), 1262–1279. https://doi.org/10.1164/rccm.2007-413ST

Peled, J. U., Gomes, A. L. C., Devlin, S. M., Littmann, E. R., Taur, Y., Sung, A. D., Weber, D., Hashimoto, D., Slingerland, A. E., Slingerland, J. B., Maloy, M., Clurman, A. G., Stein-Thoeringer, C. K., Markey, K. A., Docampo, M. D., Burgos da Silva, M., Khan, N., Gessner, A., Messina, J. A., … van den Brink, M. R. M. (2020). Microbiota as Predictor of Mortality in Allogeneic Hematopoietic-Cell Transplantation. *The New England Journal of Medicine*, *382*(9), 822–834. https://doi.org/10.1056/NEJMoa1900623

Perez-Carrasco, V., Soriano-Lerma, A., Soriano, M., Gutiérrez-Fernández, J., & Garcia-Salcedo, J. A. (2021). Urinary Microbiome: Yin and Yang of the Urinary Tract. *Frontiers in Cellular and Infection Microbiology*, *11*, 617002. https://doi.org/10.3389/fcimb.2021.617002

Proctor, D. M., & Relman, D. A. (2017). The Landscape Ecology and Microbiota of the Human Nose, Mouth, and Throat. *Cell Host & Microbe*, *21*(4), 421–432. https://doi.org/10.1016/j.chom.2017.03.011

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology*, *35*(9), 833–844. https://doi.org/10.1038/nbt.3935

Russell, N. H., Hunter, A., Rogers, S., Hanley, J., & Anderson, D. (1993). Peripheral blood stem cells as an alternative to marrow for allogeneic transplantation. *Lancet (London, England)*, *341*(8858), 1482. https://doi.org/10.1016/0140-6736(93)90929-b

Scheinberg, P., & Young, N. S. (2012). How I treat acquired aplastic anemia. *Blood*, *120*(6), 1185–1196. https://doi.org/10.1182/blood-2011-12-274019

Schoemans, H. M., Lee, S. J., Ferrara, J. L., Wolff, D., Levine, J. E., Schultz, K. R., Shaw, B. E., Flowers, M. E., Ruutu, T., Greinix, H., Holler, E., Basak, G., Duarte, R. F., Pavletic, S. Z., & EBMT (European Society for Blood and Marrow Transplantation) Transplant Complications Working Party and the “EBMT−NIH (National Institutes of Health)−CIBMTR (Center for International Blood and Marrow Transplant Research) GvHD Task Force.” (2018). EBMT-NIH-CIBMTR Task Force position statement on

standardized terminology & guidance for graft-versus-host disease assessment. *Bone Marrow Transplantation*, *53*(11), 1401–1415. https://doi.org/10.1038/s41409-018-0204-7

Segata, N., Haake, S. K., Mannon, P., Lemon, K. P., Waldron, L., Gevers, D., Huttenhower, C., & Izard, J. (2012). Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. *Genome Biology*, *13*(6), R42. https://doi.org/10.1186/gb-2012-13-6-r42

Sen, T., & Thummer, R. P. (2022). The Impact of Human Microbiotas in Hematopoietic Stem Cell and Organ Transplantation. *Frontiers in Immunology*, *13*, 932228. https://doi.org/10.3389/fimmu.2022.932228

Sender, R., Fuchs, S., & Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biology*, *14*(8), e1002533. https://doi.org/10.1371/journal.pbio.1002533

Shono, Y., & van den Brink, M. R. M. (2018). Gut microbiota injury in allogeneic haematopoietic stem cell transplantation. *Nature Reviews. Cancer*, *18*(5), 283–295. https://doi.org/10.1038/nrc.2018.10

Sierra, J., Storer, B., Hansen, J. A., Martin, P. J., Petersdorf, E. W., Woolfrey, A., Matthews, D., Sanders, J. E., Storb, R., Appelbaum, F. R., & Anasetti, C. (2000). Unrelated donor marrow transplantation for acute myeloid leukemia: An update of the Seattle experience. *Bone Marrow Transplantation*, *26*(4), 397–404. https://doi.org/10.1038/sj.bmt.1702519

Slots, J. (1979). Subgingival microflora and periodontal disease. *Journal of Clinical Periodontology*, *6*(5), 351–382. https://doi.org/10.1111/j.1600-051x.1979.tb01935.x

Solans, B. P., Chiesa, R., Doncheva, B., Prunty, H., Veys, P., Trocóniz, I. F., & Standing, J. F. (2020). Modelling of neutrophil dynamics in children receiving busulfan or treosulfan for haematopoietic stem cell transplant conditioning. *British Journal of Clinical Pharmacology*, *86*(8), 1537–1549. https://doi.org/10.1111/bcp.14260

Stein-Thoeringer, C. K., Nichols, K. B., Lazrak, A., Docampo, M. D., Slingerland, A. E., Slingerland, J. B., Clurman, A. G., Armijo, G., Gomes, A. L. C., Shono, Y., Staffas, A., Burgos da Silva, M., Devlin, S. M., Markey, K. A., Bajic, D., Pinedo, R., Tsakmaklis, A., Littmann, E. R., Pastore, A., … van den Brink, M. R. M. (2019). Lactose drives Enterococcus expansion to promote graft-versus-host disease. *Science (New York, N.Y.)*, *366*(6469), 1143–1149. https://doi.org/10.1126/science.aax3760

Styczyński, J., Tridello, G., Koster, L., Iacobelli, S., van Biezen, A., van der Werf, S., Mikulska, M., Gil, L., Cordonnier, C., Ljungman, P., Averbuch, D., Cesaro, S., de la

Camara, R., Baldomero, H., Bader, P., Basak, G., Bonini, C., Duarte, R., Dufour, C., … Infectious Diseases Working Party EBMT. (2020). Death after hematopoietic stem cell transplantation: Changes over calendar year time, infections and associated factors. *Bone Marrow Transplantation*, *55*(1), 126–136. https://doi.org/10.1038/s41409-019-0624-z

Takahashi, N., & Nyvad, B. (2011). The role of bacteria in the caries process: Ecological perspectives. *Journal of Dental Research*, *90*(3), 294–303. https://doi.org/10.1177/0022034510379602

Thomas, A. M., Manghi, P., Asnicar, F., Pasolli, E., Armanini, F., Zolfo, M., Beghini, F., Manara, S., Karcher, N., Pozzi, C., Gandini, S., Serrano, D., Tarallo, S., Francavilla, A., Gallo, G., Trompetto, M., Ferrero, G., Mizutani, S., Shiroma, H., … Segata, N. (2019). Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nature Medicine*, *25*(4), 667–678. https://doi.org/10.1038/s41591-019-0405-7

Tong, Y., Zheng, L., Qing, P., Zhao, H., Li, Y., Su, L., Zhang, Q., Zhao, Y., Luo, Y., & Liu, Y. (2019). Oral Microbiota Perturbations Are Linked to High Risk for Rheumatoid Arthritis. *Frontiers in Cellular and Infection Microbiology*, *9*, 475. https://doi.org/10.3389/fcimb.2019.00475

Toubai, T., Sun, Y., & Reddy, P. (2008). GVHD pathophysiology: Is acute different from chronic? *Best Practice & Research. Clinical Haematology*, *21*(2), 101–117. https://doi.org/10.1016/j.beha.2008.02.005

Tringe, S. G., & Hugenholtz, P. (2008). A renaissance for the pioneering 16S rRNA gene. *Current Opinion in Microbiology*, *11*(5), 442–446. https://doi.org/10.1016/j.mib.2008.09.011

Tuganbaev, T., Yoshida, K., & Honda, K. (2022). The effects of oral microbiota on health. *Science (New York, N.Y.)*, *376*(6596), 934–936. https://doi.org/10.1126/science.abn1890

Valour, F., Sénéchal, A., Dupieux, C., Karsenty, J., Lustig, S., Breton, P., Gleizal, A., Boussel, L., Laurent, F., Braun, E., Chidiac, C., Ader, F., & Ferry, T. (2014). Actinomycosis: Etiology, clinical features, diagnosis, treatment, and management. *Infection and Drug Resistance*, *7*, 183–197. https://doi.org/10.2147/IDR.S39601

Villa, A., & Sonis, S. T. (2015). Mucositis: Pathobiology and management. *Current Opinion in Oncology*, *27*(3), 159–164. https://doi.org/10.1097/CCO.0000000000000180

Welniak, L. A., Blazar, B. R., & Murphy, W. J. (2007). Immunobiology of allogeneic hematopoietic stem cell transplantation. *Annual Review of Immunology*, *25*, 139–170.

https://doi.org/10.1146/annurev.immunol.25.022106.141606

Wisplinghoff, H., Seifert, H., Wenzel, R. P., & Edmond, M. B. (2003). Current trends in the epidemiology of nosocomial bloodstream infections in patients with hematological malignancies and solid neoplasms in hospitals in the United States. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, *36*(9), 1103–1110. https://doi.org/10.1086/374339

Wolff, S. N. (2002). Second hematopoietic stem cell transplantation for the treatment of graft failure, graft rejection or relapse after allogeneic transplantation. *Bone Marrow Transplantation*, *29*(7), 545–552. https://doi.org/10.1038/sj.bmt.1703389

Wysocka-Słowik, A., Gil, L., Ślebioda, Z., Kręgielczak, A., & Dorocka-Bobkowska, B. (2021). Oral mucositis in patients with acute myeloid leukemia treated with allogeneic hematopoietic stem cell transplantation in relation to the conditioning used prior to transplantation. *Annals of Hematology*, *100*(8), 2079–2086. https://doi.org/10.1007/s00277-021-04568-y

Yap, M., Ercolini, D., Álvarez-Ordóñez, A., O'Toole, P. W., O'Sullivan, O., & Cotter, P. D. (2022). Next-Generation Food Research: Use of Meta-Omic Approaches for Characterizing Microbial Communities Along the Food Chain. *Annual Review of Food Science and Technology*, *13*, 361–384. https://doi.org/10.1146/annurev-food-052720-010751

Yehudai-Ofir, D., Henig, I., & Zuckerman, T. (2020). Aberrant B cells, autoimmunity and the benefit of targeting B cells in chronic graft-versus-host disease. *Autoimmunity Reviews*, *19*(4), 102493. https://doi.org/10.1016/j.autrev.2020.102493

Yuan, X., Wang, R., Han, B., Sun, C., Chen, R., Wei, H., Chen, L., Du, H., Li, G., Yang, Y., Chen, X., Cui, L., Xu, Z., Fu, J., Wu, J., Gu, W., Chen, Z., Fang, X., Yang, H., … Luo, F. (2022). Functional and metabolic alterations of gut microbiota in children with new-onset type 1 diabetes. *Nature Communications*, *13*(1), 6356. https://doi.org/10.1038/s41467-022-33656-4

Zhang, M., Whiteley, M., & Lewin, G. R. (2022). Polymicrobial Interactions of Oral Microbiota: A Historical Review and Current Perspective. *MBio*, *13*(3), e0023522. https://doi.org/10.1128/mbio.00235-22

Zhang, Y., Wang, X., Li, H., Ni, C., Du, Z., & Yan, F. (2018). Human oral microbiota and its modulation for oral health. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*, *99*, 883–893. https://doi.org/10.1016/j.biopha.2018.01.146

2. CHAPTER 1: Commensal oral microbiota impacts ulcerative oral mucositis clinical course in allogeneic stem cell transplant recipients

scientific reports

Check for updates

OPEN

Commensal oral microbiota impacts ulcerative oral mucositis clinical course in allogeneic stem cell transplant recipients

Julia S. Bruno[1,7], Vitor Heidrich[1,2,7], Franciele H. Knebel[1,7], Vinícius Campos de Molla[3,7], Claudia Joffily Parahyba[4], Wanessa Miranda-Silva[1], Paula F. Asprino[1], Luciana Tucunduva[4], Vanderson Rocha[4,5,6], Yana Novis[4], Celso Arrais-Rodrigues[3], Anamaria A. Camargo[1] & Eduardo R. Fregnani[1]✉

Oral mucositis (OM) is a complex acute cytotoxicity of antineoplastic treatment that affects 40–85% of patients undergoing hematopoietic stem-cell transplantation. OM is associated with prolonged hospitalization, increased extensive pharmacotherapy, need for parenteral nutrition, and elevated treatment costs. As OM onset relates to the mucosal microenvironment status, with a particular role for microbiota-driven inflammation, we aimed to investigate whether the oral mucosa microbiota was associated with the clinical course of OM in patients undergoing allogeneic hematopoietic stem-cell transplantation. We collected oral mucosa samples from 30 patients and analyzed the oral mucosa microbiota by 16S rRNA sequencing. A total of 13 patients (43%) developed ulcerative OM. We observed that specific taxa were associated with oral mucositis grade and time to oral mucositis healing. *Porphyromonas* relative abundance at preconditioning was positively correlated with ulcerative OM grade (Spearman ρ = 0.61, P = 0.028) and higher *Lactobacillus* relative abundance at ulcerative OM onset was associated with shortened ulcerative OM duration (P = 0.032). Additionally, we generated a machine-learning-based bacterial signature that uses pre-treatment microbial profiles to predict whether a patient will develop OM during treatment. Our findings suggest that further research should focus on host-microbiome interactions to better prevent and treat OM.

Allogeneic hematopoietic stem-cell transplantation (allo-HSCT) recipients undergo high doses of chemotherapy and, sometimes, total body irradiation during the conditioning regimen. During this period, they frequently experience treatment toxicities and immunity imbalance, affecting their quality of life[1]. Oral mucositis (OM) is a clinically relevant toxicity in the allo-HSCT setting, with incidences ranging from 15% (reduced intensity conditioning regimen) to 60–100% (myeloablative regimen)[2,3]. The reasons why OM is detrimental are manifold. It can cause treatment delay, early discontinuation of chemotherapy, prolonged hospitalization, extended use of analgesics, and even life-threatening complications[1,2].

Clinically, severe OM presents as an ulcer with reddish borders covered by a white pseudomembrane colonized by bacteria. OM onset in allo-HSCT recipients occurs 5–7 days after the start of the conditioning regimen[4]. Established therapies for OM involve promoting epithelial healing and reducing microbial load. Examples include basic oral care, anti-inflammatory agents, photobiomodulation, cryotherapy, and antimicrobial agents[1].

Although not yet fully elucidated, the pathophysiology of OM is multifactorial. It involves injuries to the epithelial and submucosal tissues through complex pro-inflammatory cascades. Besides, different factors can act directly on cell homeostasis affecting apoptosis and cell renewal, resulting in cell atrophy and ulceration[4].

Contributing to this complexity, there are many risk factors for OM. Genetic variables (e.g., immunogenetic variants), demographic data, tumor-related variables (e.g., malignant potential), and treatment history, among

[1]Centro de Oncologia Molecular, Hospital Sírio Libanês, Rua Prof. Daher Cutait, 69, São Paulo, SP, Brazil. [2]Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, SP, Brazil. [3]Hospital Nove de Julho, São Paulo, SP, Brazil. [4]Centro de Oncologia, Hospital Sírio Libanês, São Paulo, SP, Brazil. [5]Hospital das Clínicas da Faculdade de Medicina, Universidade de São Paulo/ICESP, São Paulo, SP, Brazil. [6]Churchill Hospital, NHS-BT, Oxford, UK. [7]These authors contributed equally: Julia S. Bruno, Vitor Heidrich, Franciele H. Knebel and Vinícius C. de Molla. ✉email: eduardofregnani@me.com

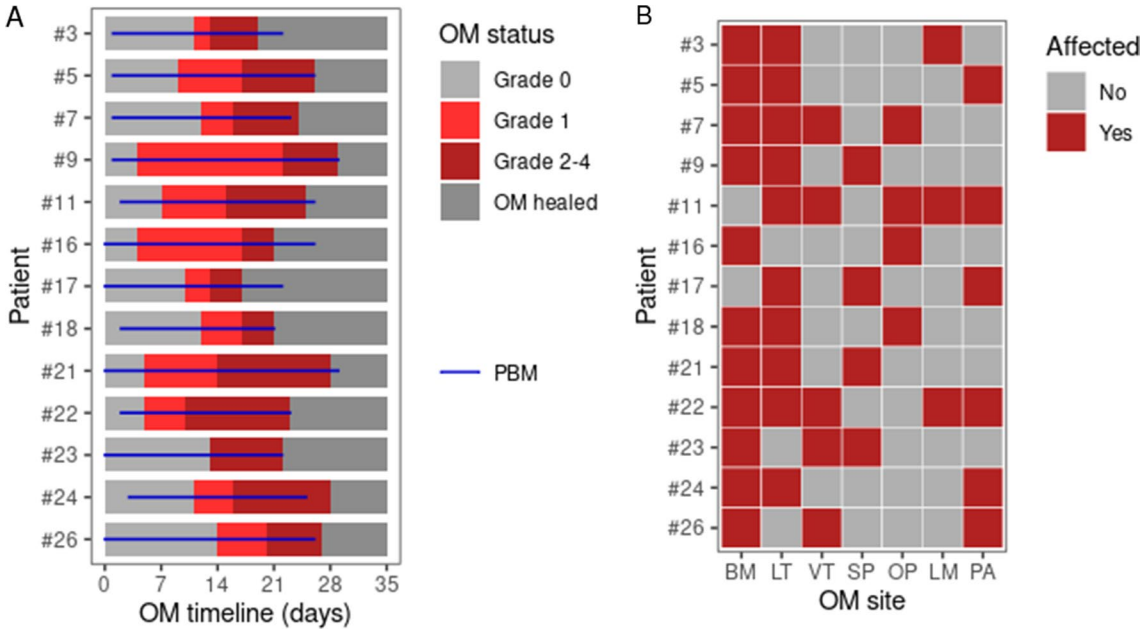


Figure 1. Oral mucositis (OM) timeline and sites affected by OM for each patient. (**A**) OM timeline in days for each OM patient. The OM grade along the timeline is indicated by a color scheme and the use of photobiomodulation (PBM) is indicated by a blue horizontal line. (**B**) Heatmap with the oral sites affected by OM for each patient. *BM* buccal mucosa, *LT* lateral tongue, *VT* ventral tongue, *SP* soft palate. *OP* oropharynx, *LM* labial mucosa, *PA* palatoglossal arches.

other factors, can affect the patient's risk of developing OM during allo-HSCT[5]. Although most risk factors associated with the incidence of OM cannot be changed, there are factors in the oral microenvironment that could be modulated—such as the oral microbiota[6–8].

In this study, we evaluated how the oral mucosal microbiota changes, from preconditioning to the OM healing, in addition to describing the changes in diversity and composition along the allo-HSCT, we also analyzed the microbiota of patients who had not developed OM. We found specific oral commensal bacterial genera associated with OM grade and duration, and generated a machine-learning-based bacterial signature to predict whether a patient will develop oral mucositis during treatment. Identifying modifiable OM risk factors can aid personalized oral care for OM prevention and treatment.

Results

Patient characteristics and OM clinical course. A total of 30 patients undergoing allo-HSCT in our institution between January 2016 and April 2018 were enrolled in this study (Table S1). Patients with periodontal disease were not included. Eighteen patients developed OM during the conditioning regimen, out of which 5 displayed only non-ulcerative OM (OM grade = 0 and 1) and 13 eventually displayed ulcerative OM (OM grade ≥ 2) during follow-up. Most patients (29/30) used broad-spectrum antibiotics during the conditioning regimen, so that there was no clear association between OM incidence and broad-spectrum antibiotic use before OM onset. The timeline of OM status for these patients, as well as the period of photobiomodulation treatment, is provided in Fig. 1A. The median number of photobiomodulation sessions were 25 (one session per day). The number of affected sites per patient varied between 2 and 5 (Fig. 1B), with buccal mucosa representing the most affected site (11/13 patients). Patients who developed ulcerative OM showed a non-significant trend ($P = 0.064$) towards showing non-ulcerative OM symptoms earlier during follow-up (Fig. S1). Most ulcerative OM patients (12/13) used broad-spectrum antibiotics during ulcerative OM. Due to the focus of this study on ulcerative OM, we will refer to it hereinafter simply as OM.

Characterization of the oral microbiota during OM. We evaluated the oral microbiota of the 13 OM patients during the OM clinical course. For each patient, 16S amplicon sequencing of oral samples was performed at preconditioning (P), oral mucositis onset (MO), and when oral mucositis was healed (MH). One sample did not achieve a satisfactory number of reads and was discarded (patient #5, MH).

Alpha-diversity significantly differed only between P and MH, although we observed a non-significant alpha-diversity decrease from P to MO and a further decrease from MO to MH (Fig. 2A). Moreover, beta-diversity significantly differed between timepoints, indicating that the oral microbiota possesses different bacterial compositions during OM clinical course (Fig. 2B).

To investigate which taxa were driving those differences in composition, we performed a differential abundance analysis at genus level with ANCOM-BC (Fig. 2C). The overall taxonomic composition at genus level for each patient during OM clinical course is provided in Fig. 2D.

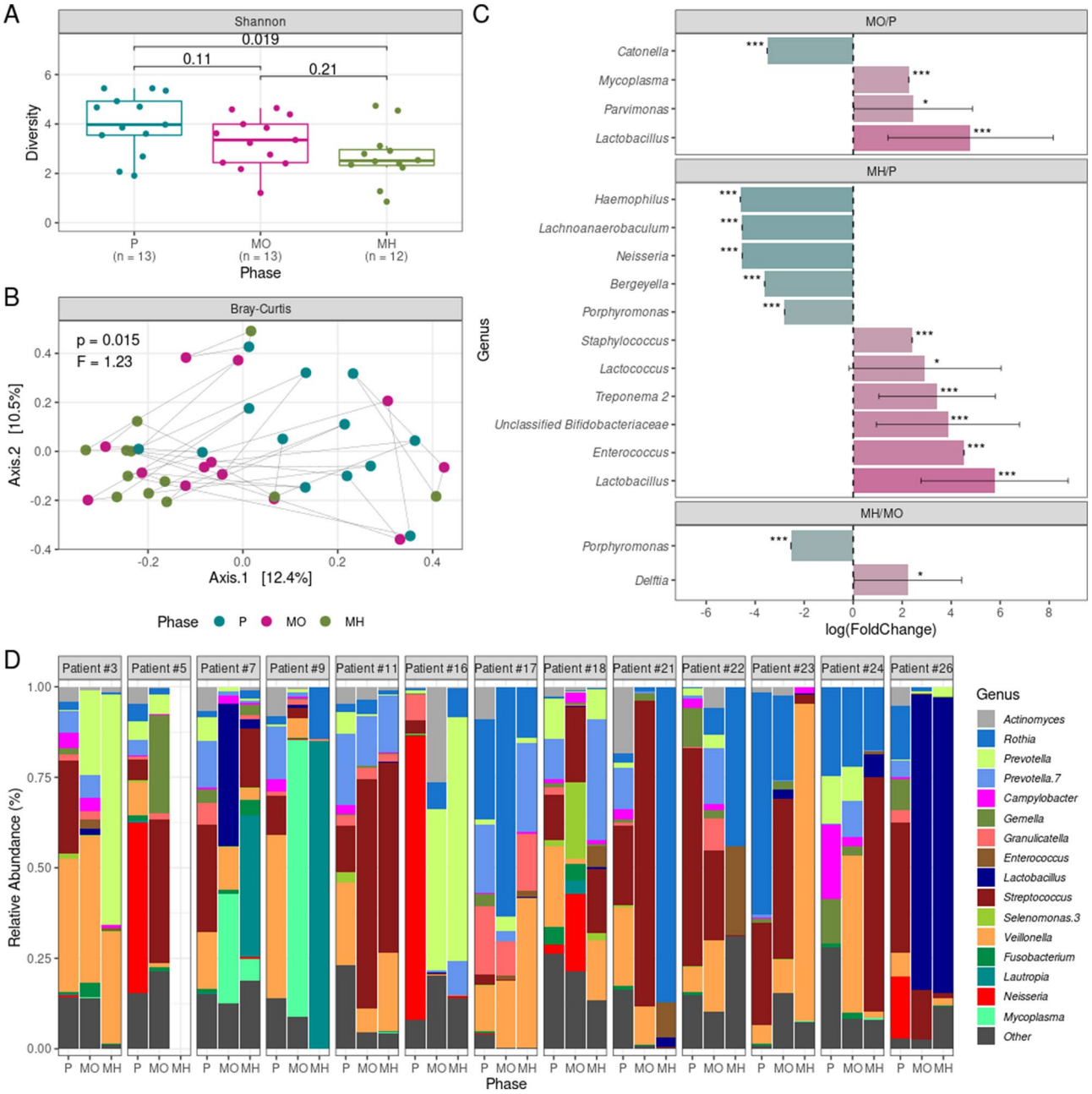


Figure 2. Changes in diversity and composition during oral mucositis (OM) clinical course. (**A**) Alpha-diversity boxplots at preconditioning (P), OM onset (MO), and OM healed (MH). Shannon was used as alpha-diversity metric. Statistical significance was evaluated by the Mann–Whitney U test, with P-values indicated. The boxes highlight the median value and cover the 25th and 75th percentiles, with whiskers extending to the more extreme value within 1.5 times the length of the box. (**B**) Principal coordinates analysis showing changes in composition during OM clinical course (beta-diversity). Bray–Curtis was used as beta-diversity metric. Samples from the same patient are linked by a gray line. Statistical significance was evaluated by the PERMANOVA test, with P- and F-values indicated. (**C**) Significant alterations in genera abundances between collection timepoints according to the ANCOM-BC test. *Adjusted P-value < 0.05; ***adjusted P-value < 0.001. (**D**) Genera relative abundances for each OM patient across collection timepoints. Only genera with > 1% relative abundance in > 25% of the samples or > 20% relative abundance in at least one sample are shown.

We identified several differentially abundant genera between timepoints. For instance, *Lactobacillus* is on average 120× more abundant in MO compared to P samples (Fig. 2C). This is also clear in terms of relative abundance, where patients #3, #7, and #26 show increased *Lactobacillus* relative abundance to the detriment of other genera in MO samples compared to P samples (Fig. 2D). A decrease in *Catonella* and increases in *Mycoplasma* and *Parvimonas* also marked the progression from P to MO (Fig. 2C).

Most of the differences were observed in the P vs. MH comparison, including increases in *Lactobacillus* and *Enterococcus* and decreases in *Haemophilus* and *Lachnoaerobaculum* (Fig. 2C). When comparing MO and MH samples, there were only two significantly differentially abundant genera between timepoints (Fig. 2C). While *Delftia* increased in abundance from MO to MH, *Porphyromonas* decreased. *Porphyromonas* is also more abundant at P in comparison to MH. In fact, in both comparisons, *Porphyromonas* was classified by ANCOM-BC as a structural zero, meaning it is not only more abundant in MO or P in comparison to MH, but that it is totally absent in MH samples.

Preconditioning oral microbiota and risk of OM development. Next, we evaluated whether the P oral microbiota was informative on the risk of OM development. To do so, we profiled the microbiota of P oral samples from all patients, which included 17 samples from patients that did not develop OM (OM-free) and 13 patients with OM. One sample from a patient of the OM-free group did not achieve a satisfactory number of reads and was discarded.

There was no difference in alpha-diversity between OM-free and OM patients at P (Fig. 3A). Accordingly, low and high alpha-diversity patients (stratified based on median Shannon index) showed no difference in OM cumulative incidence (Fig. 3B). Furthermore, oral microbiota compositions at P between OM-free and OM patients did not differ, as evaluated by a beta-diversity analysis (Fig. 3C). In line with this result, there were no significant differences in genera abundances between groups (Table S2). This can be visualized by relative abundance plots with patients sorted by OM incidence, where no signal of genera associated with OM-free or OM patients is apparent (Fig. 3D). We further confirmed that none of the P genera was associated with the OM risk using Cox regression analysis (Table S3).

To evaluate whether a signature of P genera was associated with the OM risk, we built a SVM model based on all P samples. A 96.6% accuracy (sensitivity: 92.3%; specificity: 100%) in predicting OM onset was achieved when evaluating a signature of eight genera (Fig. 3E). Differences in relative abundance and prevalence between groups for these eight genera are detailed in Fig. 3F. We also evaluated this model by leave-one-out cross-validation, showing good generalizability (82.8% mean accuracy).

Genera associated with OM clinical course. Finally, we investigated whether there were oral genera associated with OM grade and time to OM healing. When considering all patients (including OM grade < 2), even though *Streptococcus* relative abundance at P marginally correlated with lower grade OM ($P = 0.06$), none of the genus at P significantly correlated with OM grade during follow-up (Table S4). However, when considering only patients with OM grade ≥ 2, we observed that *Porphyromonas* relative abundance at P was significantly correlated with OM grade (Table S5, Fig. 4A). In fact, the top-three patients in terms of *Porphyromonas* relative abundance at P were the only patients that developed OM grade $= 4$ (Fig. 4B).

Next, we evaluated whether genera relative abundances at MO were associated with the time to OM healing using Cox regression analysis, with MO as the baseline. We found that *Lactobacillus* relative abundance at MO was significantly associated with time to OM healing (Table S6, Fig. 4C), with patients classified (based on median value) as having high *Lactobacillus* relative abundance at MO showing earlier OM healing (median time: 6 vs. 10 days; Fig. 4D).

Discussion

Initially, OM was considered a result of non-specific cell death. Currently, a series of biological events explains the progression of ulceration[9]. OM development can be divided into two stages. The initiation stage consists of chemoradiotherapy-induced DNA damage, prompting the generation of reactive oxygen species by basal epithelial cells. Consequently, inflammation-associated pathways are triggered. The most studied pathway in the pathophysiology of OM is the NF-κB signaling pathway, responsible for the expression of molecules that modulate stress, cell adhesion, apoptosis, and inflammation. Chemotherapy and radiotherapy also have indirect effects on the oral mucosa through activation of the ceramide pathway, leading to fibrolysis and production of metalloproteinases. In the second stage, named signal amplification phase, some pathways activated in the initiation stage promote higher levels of inflammation in the damaged epithelial tissue[9,10].

However, despite the huge impact of OM on the quality of life of cancer patients, it is still not clear how a patient's personal characteristics/markers can influence the incidence of OM[5]. In this work, we describe how the bacterial composition of the oral mucosa could be used as a predictive biomarker for OM in patients undergoing allo-HSCT. Oral commensals such as *Porphyromonas* and *Lactobacillus* are associated with the OM severity and healing period. Additionally, we provide a characterization of the oral mucosa microbiota dynamics during allo-HSCT with a detailed data collection of OM duration, grade, and anatomical sites affected.

There are no preventive strategies based on a patient's microenvironmental characteristics. OM preventive strategies are based on oral hydration to decrease mucosal friability, photobiomodulation to increase mucosal repair potential, and oral hygiene for unspecific microbial control. In this context, omics-based analyses can help elucidate the influence of the oral microbiota on OM onset and provide evidence to support future studies on microbial modulation as a preventive and curative strategy. Beyond oral side effects, our previous study showed an association between low bacterial diversity of oral mucosa microbiota at preconditioning and a higher risk of relapse[11].

Porphyromonas is known to be a key-pathogen of chronic periodontal disease, being found in 85% of periodontal pockets[12]. Additionally, its impact on systemic diseases has gained increased attention in the literature, including associations with inflammatory bowel disease[13] and Alzheimer's disease[14]. *Porphyromonas gingivalis* can manipulate the host's innate immune response, being able to adapt, invade and survive. Beyond the activation of inflammatory pathways, *Porphyromonas gingivalis* pathogenicity can be explained by its survival strategy

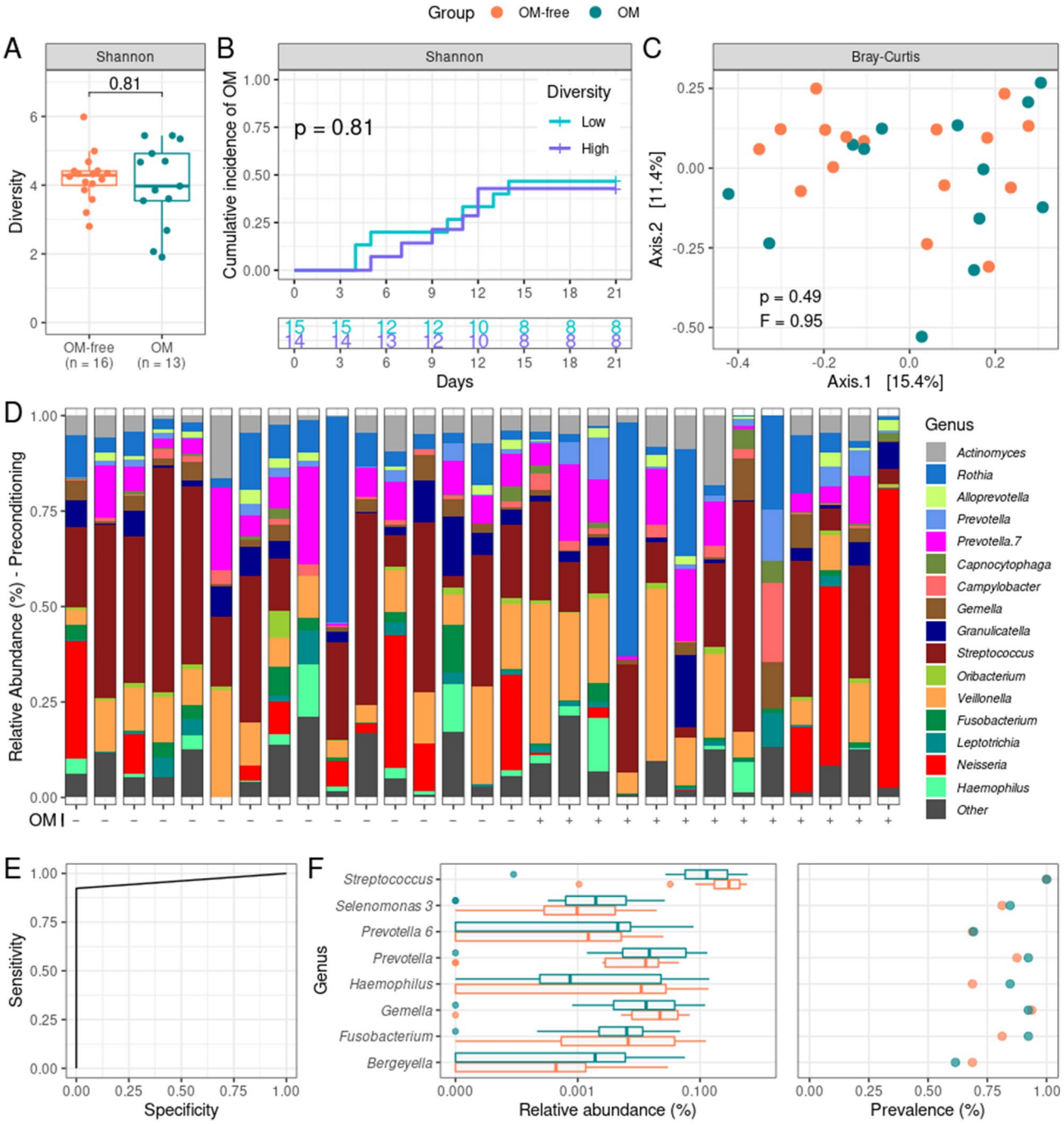


Figure 3. Comparisons between oral mucositis (OM) patients and OM-free patients at preconditioning (P). (**A**) Alpha-diversity boxplots at P for OM and OM-free patients. Shannon was used as alpha-diversity metric. Statistical significance was evaluated by the Mann–Whitney U test, with P-value indicated. The boxes highlight the median value and cover the 25th and 75th percentiles, with whiskers extending to the more extreme value within 1.5 times the length of the box. (**B**) Cumulative incidence curves of OM with patients stratified by alpha-diversity level (low/high, based on median Shannon index) at preconditioning. The number of patients at risk is shown. Statistical significance was evaluated by the log-rank test, with P-value indicated. (**C**) Principal coordinates analysis comparing compositions at P of OM and OM-free patients (beta-diversity). Bray–Curtis was used as a beta-diversity metric. Statistical significance was evaluated by the PERMANOVA test, with P- and F-values indicated. (**D**) Genera relative abundances at P for OM and OM-free patients. Patients are sorted based on OM categories (OM-free/OM: −/+), as indicated by x-axis labels. Only genera with > 1% relative abundance in > 25% of the samples or > 20% relative abundance in at least one sample are shown. (**E**) Receiver-operating characteristic curve for a support vector machine model (SVM) for classifying patients into OM and OM-free categories based on P oral microbiota data. The model was built based on the relative abundances of eight genera at P. (**F**) Relative abundances boxplots (left) and prevalence (right) for OM and OM-free patients of the eight genera at P used in the SVM model. A symlog scale was used in the x-axis of the relative abundances plot, with 10^{-5} as linearity threshold.

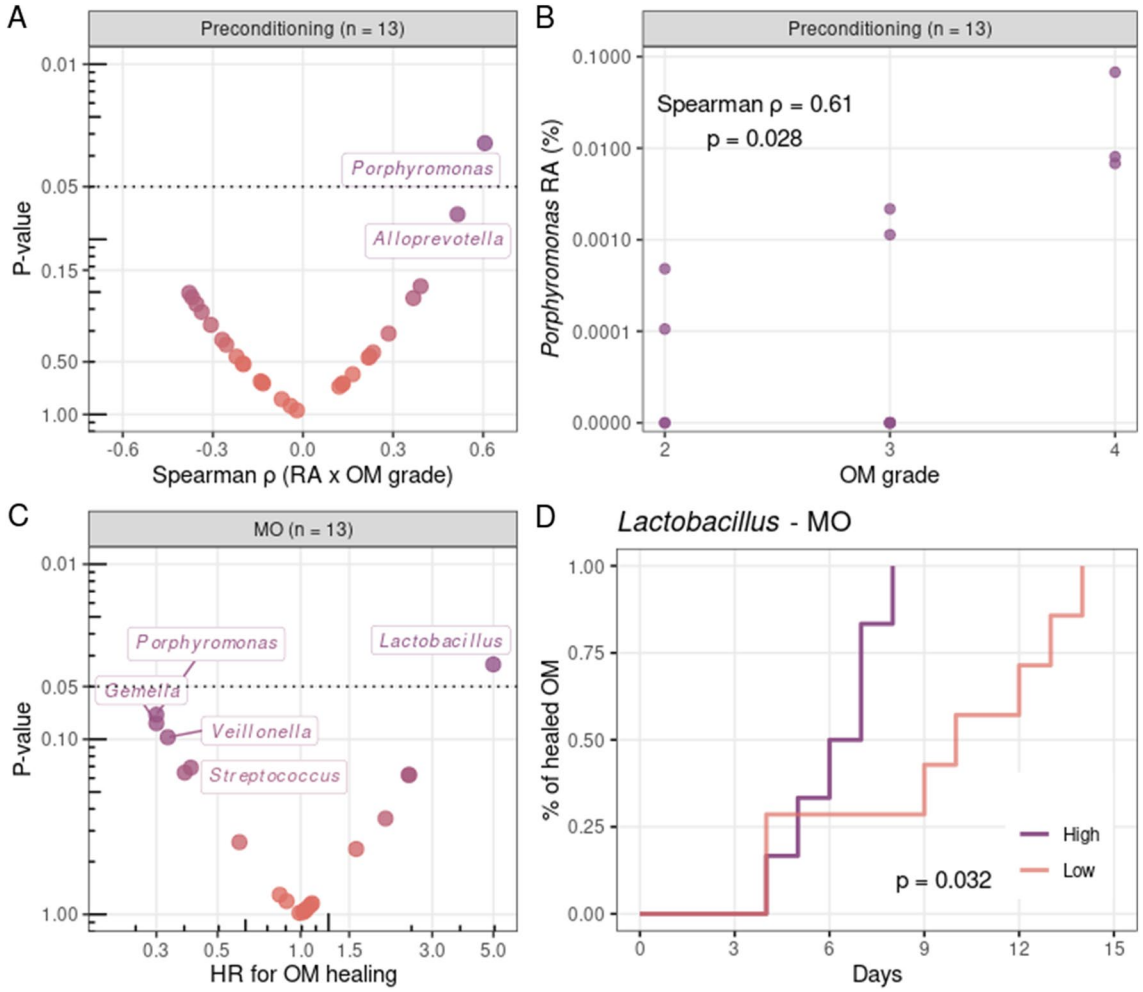


Figure 4. Genera associated with oral mucositis (OM) grade and OM healing. (**A**) Volcano plot (Spearman ρ vs. P-value) depicting correlations between the highest OM grade and genera relative abundances at preconditioning (P). (**B**) Spearman correlation between the highest OM grade and *Porphyromonas* relative abundance at P. Spearman ρ and P-value are indicated. *RA* relative abundance. (**C**) Volcano plot (Cox hazard ratio vs. P-value) of the risk analysis for the association of OM healing with genera relative abundance at OM onset (MO). (**D**) Cumulative incidence curves of healed OM with patients stratified by *Lactobacillus* relative abundance (low/high, based on median value) at MO. Statistical significance was evaluated by the log-rank test, with P-value indicated. In (**A,C**), only genera present (non-zero relative abundance) in > 50% of the samples were evaluated. Only genera with P-value < 0.15 are indicated explicitly.

that circumvents the immune system by invading host cells. Invasion occurs mainly through the interaction between the fimbriae and B1 integrins of host cells, which triggers cytoskeletal restructuring, allowing bacterial internalization. Noteworthy, invasion does not trigger cell apoptosis, allowing bacterial survival and replication within the host cell[12,15,16]. Our results showing that the relative abundance of *Porphyromonas* at preconditioning is correlated with the highest OM grade presented during follow-up reinforce the importance of studying this genus in the context of oral care in hospitalized cancer patients. Furthermore, we found that *Porphyromonas* is virtually absent in MH samples. Although causality cannot be evaluated, this result suggests *Porphyromonas* clearance may be necessary for OM healing, an intriguing hypothesis that also demands further investigation.

The use of probiotics containing *Lactobacillus* is being evaluated to prevent OM severity in head and neck cancer patients[6–8]. One phase II study prescribing *Lactobacillus brevis CD2* for HSCT recipients reported lower grades of OM. The putative mechanism of action involves the production of arginine deiminase by *Lactobacillus brevis CD2,* which downregulates the pro-inflammatory nitric oxide pathway[17]. Our results showing that the relative abundance of wild/natural *Lactobacillus* is associated with a faster ulcerative OM healing time supports future clinical trials in patients undergoing allo-HSCT.

Other studies analyzed the role of the oral microbiota in OM during oncohematologic treatment[18–22]. One-such study showed a decrease in bacterial diversity during transplantation and a greater abundance of specific genera only in patients who used methotrexate prior to allo-HSCT[19]. In another study, a decrease in diversity was noted in patients without ulcerative oral mucositis[18]. One work focused on patients undergoing allo-HSCT and OM severity, even though by evaluating saliva samples. They found associations between the relative abundance of *Kingella* and *Atopobium* in saliva and OM severity[19]. In our study, these genera were not associated with

OM parameters, possibly due to the evaluation herein of oral mucosa samples rather than saliva. A long-term analysis of saliva microbiome in allo-HSCT showed reestablishment of bacterial diversity months after stem-cell infusion. And patients who developed OM had lower diversity in the third week when compared with patients without OM[23].

Besides describing variations in the oral microbiota during OM clinical course, we also evaluated whether oral microbiota composition could be used as a biomarker for OM incidence. Among other results, we provide for the first time a machine-learning-based bacterial signature for predicting OM. This signature includes only eight genera: *Streptococcus*, *Selenomonas 3*, *Prevotella 6*, *Prevotella*, *Haemophilus*, *Gemella*, *Fusobacterium*, and *Bergeyella*—possible research targets for OM onset. Validation cohorts are needed to confirm the clinical value of this bacterial signature. Further studies will also be needed to overcome the limitations of our study, such as the lack of longitudinally collected samples from OM-free patients and small sample size.

Oral care is an essential part of the oncologic treatment, as it maintains patient's quality of life, decreases the use of analgesics and shortens hospitalization period. Predictive analysis is a fundamental part of precision medicine and supports the innovation of clinical guidelines. Our study highlights the role of commensal oral bacteria in OM clinical course. It also demonstrates the importance of characterizing the oral microbiota in oncologic patients for improving clinical care. Further, more powered studies will be necessary to evaluate the influence of commensals and pathogens in the pathophysiology of OM.

Materials and methods

Sample collection. Enrolled patients underwent allo-HSCT at Hospital Sírio-Libanês (São Paulo/Brazil) between 2016 and 2018. The study was approved by the local ethics committee (Comite de Ética em Pesquisa—Hospital Sírio-Libanês (#HSL 2016-08)), according to the Declaration of Helsinki, and all patients provided informed consent before sample collection. No tissue was procured from prisoners in this study.

The oral mucosa sample was collected with a sterile swab on bilateral buccal mucosa, alveolar mucosa of the jaws, and tongue dorsum. Samples were collected at preconditioning (before conditioning regimen), ulcerative OM onset, and when OM ulcerations were healed (no sign of ulceration). Patients did not perform oral hygiene for at least 6 h before sample collection.

Institutional standard antimicrobial prophylaxis. The standard antimicrobial prophylaxis in our institution included oral levofloxacin and/or sulfamethoxazole-trimethoprim, acyclovir, and antifungal prophylaxis according to the patient's risk of fungal infection (low risk: fluconazole; high risk: voriconazole).

Oral care and photobiomodulation. All patients were examined and treated by two trained professionals of the oral medicine department of our institution following the MASCC/ISOO Guideline for Cancer Patients[24]. The standard oral hygiene protocol was fluoride toothpaste and 0.12% chlorhexidine (CHX) mouthwash. The topical CHX was administered once a day. The photobiomodulation protocol was performed with low-level laser equipment (Laser XT Therapy, DMC, São Carlos, Brazil) at a wavelength of 660 nm (spot size = 0.028 cm^2; 100mW of power) irradiating 64 points of the oral mucosa, covering buccal mucosa, mucobuccal fold, palatoglossal arches, soft palate, labial mucosa, tongue (lateral and ventral). The irradiation ranged between 1 and 2 J/point for preventive and curative treatment for oral lesions, respectively.

DNA extraction and 16S rRNA amplicon sequencing. Bacterial cells were recovered from oral mucosa swabs using TE buffer and 6 µL PureLink RNAse A (20 mg/mL, Thermo Fisher Scientific, Waltham, MA, USA). DNA was extracted using QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol (DNA Purification from Blood or Body Fluids) and stored at − 80 °C. Pre-validated primers and 12.5 ng DNA were used to amplify the 16S rRNA hypervariable regions V3–V4[25]. Amplicons were sequenced as described elsewhere[26] on an Illumina MiSeq platform (Illumina, San Diego, CA, USA).

Bioinformatics pipeline. Reads were processed with QIIME 2[27] following the DADA2 pipeline[28] to generate Amplicon Sequencing Variants (ASVs). Chimeric ASVs were filtered out with VSEARCH[29] by using the SILVA database as reference[30]. The taxonomic assignment of ASVs was performed with VSEARCH and SILVA. ASVs not assigned to bacteria were removed. After read filtering steps, samples with < 1000 reads were discarded. Next, microbiota analysis was performed using custom R scripts[31].

Microbiota analyses. Libraries were normalized to 6256 reads by Scaling with Ranked Subsampling[32] with the R package *SRS*[33] to account for variable sequencing depth prior to diversity analysis. Alpha-diversity was calculated at ASV level with the QIIME 2 plugin *q2-diversity* using the Shannon index[34]. Differences in alpha-diversity between groups were evaluated using the Mann–Whitney U test. Beta-diversity was calculated at ASV level with the R package *phyloseq*[35] using Bray–Curtis dissimilarity index[36]. Compositional differences between groups were represented by Principal Coordinate Analysis and evaluated using the PERMANOVA test[37].

In genera relative abundance plots (generated with the R package *ggplot2*[38]) only genera with > 1% relative abundance in > 25% of the samples or > 20% relative abundance in at least one sample are shown. Differential abundance of genera between groups was evaluated with ANCOM-BC[39]. Genera with log (FoldChange) > 2 between groups and $P < 0.05$ after Bonferroni correction were considered statistically significant.

Only genera present (non-zero relative abundance) in > 50% of the samples were evaluated in the associations between genera relative abundance and OM clinical course. Associations between genera relative abundance and OM stage were evaluated using Spearman correlation. Associations between genera relative abundance and

time to OM development (with the starting day of the conditioning regimen as reference) or time to OM healing were evaluated by stratifying patients into low and high groups (based on median genus relative abundance) and estimating the Cox proportional hazards between groups with the R package survival[40]. The same approach was used to associate alpha-diversity with time to OM development, with patients stratified into low and high alpha-diversity groups based on the median Shannon index. Kaplan–Meier curves were generated with the R package *survminer*[41].

The support vector machine (SVM) model was generated with the R package *kernlab*[42]. All preconditioning samples were included and only genera present in > 50% of preconditioning samples were considered. The model was tested using the leave-one-out cross-validation approach. The final model was built with the number of genera (n = 8) and the cost (C = 10) that maximized cross-validation accuracy.

Data availability

Sequencing data were deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under Accession Number PRJEB49175.

Received: 12 April 2022; Accepted: 4 October 2022
Published online: 20 October 2022

References

1. Elad, S. *et al.* MASCC/ISOO clinical practice guidelines for the management of mucositis secondary to cancer therapy. *Cancer* **126**, 4423–4431 (2020).
2. Cinausero, M. *et al.* New frontiers in the pathobiology and treatment of cancer regimen-related mucosal injury. *Front. Pharmacol.* **8**, 354 (2017).
3. Wysocka-Słowik, A., Gil, L., Ślebioda, Z., Kręgielczak, A. & Dorocka-Bobkowska, B. Oral mucositis in patients with acute myeloid leukemia treated with allogeneic hematopoietic stem cell transplantation in relation to the conditioning used prior to transplantation. *Ann. Hematol.* **100**, 2079–2086 (2021).
4. Villa, A. & Sonis, S. T. Mucositis: Pathobiology and management. *Curr. Opin. Oncol.* **27**, 159–164 (2015).
5. Wardill, H. R. *et al.* Prediction of mucositis risk secondary to cancer therapy: A systematic review of current evidence and call to action. *Support Care Cancer* **28**, 5059–5073 (2020).
6. Xia, C. *et al.* A phase II randomized clinical trial and mechanistic studies using improved probiotics to prevent oral mucositis induced by concurrent radiotherapy and chemotherapy in nasopharyngeal carcinoma. *Front. Immunol.* **12**, 618150 (2021).
7. de Sanctis, V. *et al. Lactobacillus brevis* CD2 for prevention of oral mucositis in patients with head and neck tumors: A multicentric randomized study. *Anticancer Res.* **39**, 1935–1942 (2019).
8. Jiang, C. *et al.* A randomized, double-blind, placebo-controlled trial of probiotics to reduce the severity of oral mucositis induced by chemoradiotherapy for patients with nasopharyngeal carcinoma. *Cancer* **125**, 1081–1090 (2019).
9. Sonis, S. T. Mucositis: The impact, biology and therapeutic opportunities of oral mucositis. *Oral Oncol.* **45**, 1015–1020 (2009).
10. Laheij, A. M. G. A. *et al.* Oral bacteria and yeasts in relationship to oral ulcerations in hematopoietic stem cell transplant recipients. *Support Care Cancer* **20**, 3231–3240 (2012).
11. de Molla, V. C. *et al.* Disruption of the oral microbiota is associated with a higher risk of relapse after allogeneic hematopoietic stem cell transplantation. *Sci. Rep.* **11**, 1–11 (2021).
12. Bostanci, N. & Belibasakis, G. N. *Porphyromonas gingivalis*: An invasive and evasive opportunistic oral pathogen. *FEMS Microbiol. Lett.* **333**, 1–9 (2012).
13. Read, E., Curtis, M. A. & Neves, J. F. The role of oral bacteria in inflammatory bowel disease. *Nat. Rev. Gastroenterol. Hepatol.* **18**, 731–742 (2021).
14. Dominy, S. S. *et al. Porphyromonas gingivalis* in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. *Sci. Adv.* **5**, 3333 (2019).
15. Stringer, A. M. & Logan, R. M. The role of oral flora in the development of chemotherapy-induced oral mucositis. *J. Oral Pathol. Med.* **44**, 81–87 (2015).
16. Jia, L. *et al.* Pathogenesis of important virulence factors of *Porphyromonas gingivalis* via toll-like receptors. *Front. Cell Infect. Microbiol.* **9**, 262 (2019).
17. Sharma, A. *et al. Lactobacillus brevis* CD2 lozenges prevent oral mucositis in patients undergoing high dose chemotherapy followed by haematopoietic stem cell transplantation. *ESMO Open* **1**, e000138 (2016).
18. Laheij, A. M. G. A. *et al.* Microbial changes in relation to oral mucositis in autologous hematopoietic stem cell transplantation recipients. *Sci. Rep.* **9**, 16929 (2019).
19. Shouval, R. *et al.* Patterns of salivary microbiota injury and oral mucositis in recipients of allogeneic hematopoietic stem cell transplantation. *Blood Adv.* **4**, 2912–2917 (2020).
20. Lee, A. *et al.* Association of HSV-1 and reduced oral bacteriota diversity with chemotherapy-induced oral mucositis in patients undergoing autologous hematopoietic stem cell transplantation. *J. Clin. Med.* **9**, 1090 (2020).
21. Takahashi, M. *et al.* An analysis of oral microbial flora by T-RFLP in patients undergoing hematopoietic stem cell transplantation. *Int. J. Hematol.* **112**, 690–696 (2020).
22. Mougeot, J.-L.C. *et al.* Lasting Gammaproteobacteria profile changes characterized hematological cancer patients who developed oral mucositis following conditioning therapy. *J. Oral Microbiol.* **12**, 1761135 (2020).
23. Laheij, A. M. G. A. *et al.* Long-term analysis of resilience of the oral microbiome in allogeneic stem cell transplant recipients. *Microorganisms* **10**, 734 (2022).
24. Lalla, R. V. *et al.* MASCC/ISOO clinical practice guidelines for the management of mucositis secondary to cancer therapy. *Cancer* **120**, 1453–1461 (2014).
25. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1 (2013).
26. Wang, H. *et al.* Breast tissue, oral and urinary microbiomes in breast cancer. *Oncotarget* **8**, 88122–88138 (2017).
27. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
28. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
29. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: A versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
30. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
31. Team, R. *A Language and Environment for Statistical Computing*. https://www.r-project.org/ (2017).

32. Beule, L. & Karlovsky, P. Improved normalization of species count data in ecology by scaling with ranked subsampling (SRS): Application to microbial communities. *PeerJ* **8**, e9593 (2020).
33. Heidrich, V., Karlovsky, P. & Beule, L. "SRS" R Package and "q2-srs" QIIME 2 plugin: Normalization of microbiome data using scaling with ranked subsampling (SRS). *Appl. Sci.* **11**, 11473 (2021).
34. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
35. McMurdie, P. J. & Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).
36. Roger Bray, J. T. C. An ordination of the upland forest communities of southern Wisconsin. *Ecol. Soc. Am.* **27**, 325–349 (1957).
37. Anderson, B. M. J. *Permutational Multivariate Analysis of Variance (PERMANOVA)*, 1–15. https://doi.org/10.1002/9781118445112.stat07841.
38. Valero-Mora, P. M. ggplot2: Elegant graphics for data analysis. *J. Stat. Softw. Book Rev.* **35**, 1–3 (2010).
39. Lin, H. & Peddada, S. D. Analysis of compositions of microbiomes with bias correction. *Nat. Commun.* **11**, 3514 (2020).
40. Therneau, T. M. *A Package for Survival Analysis in R*. https://cran.r-project.org/package=survival (2021).
41. Kassambara, A., Kosinski, M. & Przemyslaw Biecek, S. F. *Survminer: Drawing Survival Curves Using "ggplot2"* (2021).
42. Karatzoglou, A., Smola, A., Hornik, K. & Zeileis, A. kernlab—An S4 package for Kernel methods in R. *J. Stat. Softw.* **11**, 1–20 (2004).

Author contributions

J.S.B., V.H., V.C.M., C.A.-R., A.A.C. and E.R.F. contributed to the analysis and interpretation of data; J.S.B., V.H., A.A.C. and E.R.F. contributed to writing the manuscript; F.H.K, P.F.A., V.H., J.S.B., C.J.P. and W.M.-S. performed the sequencing; A.A.C. and E.R.F. designed the study; V.H. and A.A.C. developed the bioinformatics pipeline; All authors have read and approved the final manuscript. All authors are accountable for all aspects of the work.

Funding

JB was supported by Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior (CAPES, process no.001). VH was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, process no. 13996-0/2018). VM was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, process n 141575/2018-2).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-21775-3.

Correspondence and requests for materials should be addressed to E.R.F.

Reprints and permissions information is available at www.nature.com/reprints.

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

3. CHAPTER 2: Dental biofilm microbiota dysbiosis is associated with the risk of acute graft-versus-host disease after allogeneic hematopoietic stem cell transplantation

frontiers in Immunology

ORIGINAL RESEARCH
published: 18 June 2021
doi: 10.3389/fimmu.2021.692225

Dental Biofilm Microbiota Dysbiosis Is Associated With the Risk of Acute Graft-*Versus*-Host Disease After Allogeneic Hematopoietic Stem Cell Transplantation

Vitor Heidrich[1,2†], Julia S. Bruno[1†], Franciele H. Knebel[1†], Vinícius C. de Molla[3,4†], Wanessa Miranda-Silva[1], Paula F. Asprino[1], Luciana Tucunduva[3], Vanderson Rocha[3,5,6], Yana Novis[3], Celso Arrais-Rodrigues[3,4], Eduardo R. Fregnani[3] and Anamaria A. Camargo[1]*

[1] Centro de Oncologia Molecular, Hospital Sírio-Libanês, São Paulo, Brazil, [2] Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, Brazil, [3] Centro de Oncologia, Hospital Sírio-Libanês, São Paulo, Brazil, [4] Departamento de Oncologia Clínica e Experimental, Universidade Federal de São Paulo, São Paulo, Brazil, [5] Hospital das Clínicas da Faculdade de Medicina, Universidade de São Paulo/Instituto do Câncer do Estado de São Paulo (ICESP), São Paulo, Brazil, [6] Churchill Hospital, National Health Service Blood and Transplant, Oxford, United Kingdom

Acute graft-*versus*-host disease (aGVHD) is one of the major causes of death after allogeneic hematopoietic stem cell transplantation (allo-HSCT). Recently, aGVHD onset was linked to intestinal microbiota (IM) dysbiosis. However, other bacterial-rich gastrointestinal sites, such as the mouth, which hosts several distinctive microbiotas, may also impact the risk of GVHD. The dental biofilm microbiota (DBM) is highly diverse and, like the IM, interacts with host cells and modulates immune homeostasis. We characterized changes in the DBM of patients during allo-HSCT and evaluated whether the DBM could be associated with the risk of aGVHD. DBM dysbiosis during allo-HSCT was marked by a gradual loss of bacterial diversity and changes in DBM genera composition, with commensal genera reductions and potentially pathogenic bacteria overgrowths. High *Streptococcus* and high *Corynebacterium* relative abundance at preconditioning were associated with a higher risk of aGVHD (67% *vs.* 33%; HR = 2.89, P = 0.04 and 73% *vs.* 37%; HR = 2.74, P = 0.04, respectively), while high *Veillonella* relative abundance was associated with a lower risk of aGVHD (27% *vs.* 73%; HR = 0.24, P < 0.01). *Enterococcus faecalis* bloom during allo-HSCT was observed in 17% of allo-HSCT recipients and was associated with a higher risk of aGVHD (100% *vs.* 40%; HR = 4.07, P < 0.001) and severe aGVHD (60% *vs.* 12%; HR = 6.82, P = 0.01). To the best of our knowledge, this is the first study demonstrating that DBM dysbiosis is associated with the aGVHD risk after allo-HSCT.

Keywords: oral microbiota, supragingival plaque, microbiome dysbiosis, acute GVHD, allogeneic HSCT, bone marrow transplant

OPEN ACCESS

Edited by:
Evelyn Ullrich, Goethe University Frankfurt, Germany

Reviewed by:
Michele Malagola, University of Brescia, Italy
Shigeo Fuji, Osaka International Cancer Institute, Japan

***Correspondence:**
Anamaria A. Camargo
anamaria.acamargo@hsl.org.br

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Alloimmunity and Transplantation, a section of the journal Frontiers in Immunology

Received: *13 April 2021*
Accepted: *02 June 2021*
Published: *18 June 2021*

Citation:
Heidrich V, Bruno JS, Knebel FH, de Molla VC, Miranda-Silva W, Asprino PF, Tucunduva L, Rocha V, Novis Y, Arrais-Rodrigues C, Fregnani ER and Camargo AA (2021) Dental Biofilm Microbiota Dysbiosis Is Associated With the Risk of Acute Graft-Versus-Host Disease After Allogeneic Hematopoietic Stem Cell Transplantation. Front. Immunol. 12:692225. doi: 10.3389/fimmu.2021.692225

INTRODUCTION

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) is the only curative treatment for several hematologic diseases. However, allo-HSCT recipients may experience potentially fatal complications, such as infections and graft-*versus*-host disease (GVHD) (1).

Acute GVHD (aGVHD) is a clinical syndrome characterized by maculopapular rash, hyperbilirubinemia, anorexia, diarrhea and abdominal pain (2). The incidence of aGVHD grade II-IV is 30-40% at day 100 (3). During transplantation, chemotherapy, radiotherapy, and infection can damage host cells, releasing sterile damage-associated molecular patterns (DAMPs) and pathogen-associated molecular patterns (PAMPs) into the extracellular milieu. DAMPs and PAMPs activate donor T cells leading to a proinflammatory state. Simultaneously, donor regulatory T cells, myeloid-derived suppressor cells and tolerogenic dendritic cells are activated, counterbalancing the inflammation as an anti-inflammatory response. An imbalance in these events towards the proinflammatory state may result in aGVHD (4).

In addition to the graft source and the intensity of the conditioning regimen (4), the intestinal microbiota (IM) composition was shown to be associated with the risk and intensity of aGVHD. Loss of IM diversity has been observed during the pre- and post-transplantation period (5), and low microbiota diversity at the time of stem cell engraftment has been associated with a higher risk of severe aGVHD (5) and transplant-related death (6).

Two non-exclusive ecological events can explain the link between loss of bacterial diversity and aGVHD risk: absence or loss of protective commensal bacterial species and sudden expansion (also known as bloom) of opportunistic pathogenic bacteria. Both events have been independently linked to aGVHD development. For instance, a higher abundance of commensal bacteria from the *Blautia* genus in the IM after allo-HSCT has been associated with reduced GVHD-related mortality and improved overall survival (7, 8). On the other hand, a shift in IM leading to the dominance of bacteria from the *Enterococcus* genus occurs more prominently in allo-HSCT recipients developing aGVHD (9), and it is associated with increased GVHD-related mortality (10).

Recent studies have shown that bacteria inhabiting the oral cavity can translocate to the gut (11) and drive IM dysbiosis (12). However, direct evaluation of the effect of allo-HSCT on the oral microbiota (OM) and the influence of OM dysbiosis on aGVHD risk have not been performed. To further understand the impact of gastrointestinal bacterial communities on aGVHD development following allo-HSCT, it would be crucial to extend the scope of these analyses to the OM.

The OM comprises over 700 bacterial species that stick to surfaces of the mouth, forming biofilms (13). The dental biofilm microbiota (DBM), in particular, is among the richest and most diverse and, like the IM, interacts with host cells and modulates immune homeostasis (14). In this study, we characterized changes of the DBM in patients during allo-HSCT and evaluated whether alterations in DBM diversity and composition could be associated with the risk of aGVHD.

MATERIALS AND METHODS

Sample Collection and Oral Care Protocol

Supragingival biofilm samples were collected from patients who underwent allo-HSCT. Samples were collected with sterile swabs at three phases during allo-HSCT: before the conditioning regimen (preconditioning), at aplasia and at engraftment. All patients were requested not to perform oral hygiene for at least 6h before sample collection. All patients were examined by an oral medicine specialist for potential infections and followed the same protocol for oral mucositis prophylaxis with photobiomodulation and oral hygiene with fluoride toothpaste and 0.12% chlorhexidine mouthwash. Informed consent was obtained from all participants prior to sample collection. The study was approved by the Institutional Ethics Committee (Protocol #1.414.217), in line with the Declaration of Helsinki.

DNA Extraction and Sequencing

Bacterial cells were recovered from swabs by vortexing in TE buffer supplemented with PureLink RNAse A (Thermo Fisher Scientific, Waltham, MA, USA). DNA was extracted using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Next, 12.5 ng of total DNA and pre-validated primers (15) were used to amplify 16S rRNA hypervariable regions V3–V4. Amplicons were sequenced as described elsewhere (16) on the MiSeq platform (Illumina, San Diego, CA, USA).

Bioinformatics Analyses

Reads were demultiplexed and primer sequences were removed using the MiSeq Reporter software. Read processing was carried out within the QIIME 2 (*Quantitative Insights Into Microbial Ecology 2*) framework (17). Briefly, forward and reverse sequences were filtered for quality and bimeras, denoised, and merged into consensus sequences with the DADA2 pipeline (18), generating unique amplicon sequencing variants (ASVs). ASVs were further filtered for chimeric sequences using the SILVA database (19) and UCHIME (20), resulting in a total of 6 434 516 high-quality 16S rRNA sequences, with the median number of sequences obtained per sample being 58 867 (range: 2 153 - 240 734). Afterwards, ASVs were taxonomically assigned using the SILVA database and VSEARCH tool (21).

Microbiota and Statistical Analyses

As determined by per sample alpha diversity rarefaction curves, <12 500 reads samples were considered defective and excluded. To adjust for differences in library sizes, the remaining samples were rarefied to 14 157 reads before calculating alpha diversity indexes (Shannon and Gini-Simpson indexes and the number of observed ASVs as a proxy for species richness) with the QIIME 2 *q2-diversity* plugin. Alpha diversity across transplantation phases was compared with the Mann-Whitney

U test. The relative abundance of each genus was calculated with the QIIME 2 *q2-taxa* plugin. Differentially abundant genera across transplantation phases were identified using ANCOM (22). ANCOM W represents the proportion of null hypotheses rejected when subtesting the differential abundance of a genus normalized by the abundance of each one of the genera in the dataset. W > 0.7 was considered as statistically significant. Cumulative incidence (CMI) rates for aGVHD (grade II to IV) and severe aGVHD (grade III and IV) were calculated with death as a competing event. Relative risks for developing aGVHD and severe aGVHD were estimated using the Fine-Gray risk regression model and adjusted for graft source and intensity of the conditioning regimen. Relative risks are presented as hazard ratios with 95% CIs and two-tailed P-values. R software (version 3.6.2) and the statistical package *cmprsk* (version 2.2.9) were used for statistical analyses.

RESULTS

Patient Characteristics

A total of 30 patients who underwent allo-HSCT for hematologic disorders at Hospital Sírio-Libanês between January 2016 and April 2018 were consecutively enrolled in our study. Patient clinical characteristics are summarized in **Table 1**. The most common underlying disease was acute leukemia (60%). The majority of patients received reduced-intensity conditioning (60%) and grafts from peripheral blood (67%).

TABLE 1 | Clinical characteristics of study patients.

	n = 30
Sex (Male)	16 (53%)
Age in years (median, range)	50 (19-73)
Underlying disease*	
Acute leukemia	18 (60%)
Other	12 (40%)
Conditioning intensity	
Reduced intensity	18 (60%)
Total body irradiation	11 (37%)
Pre-transplant T-cell depletion	15 (50%)
Graft source	
Bone marrow	10 (33%)
Peripheral blood	20 (67%)
Donor	
Matched sibling	9 (30%)
Haploidentical	10 (33%)
Matched unrelated	9 (30%)
Mismatched unrelated	2 (7%)
GVHD prophylaxis	
MMF + CsA	11 (37%)
MTX + CsA	10 (33%)
MMF + CsA + PTCy	9 (30%)
Follow-up in months (median, range)	37 (25-46)

*HCT-CI, Hematopoietic cell transplantation-specific comorbidity index; MMF, mycophenolate mofetil; MTX, methotrexate; CsA, cyclosporin A; PTCy, post-transplant cyclophosphamide. *Acute leukemia: 11 acute myeloid leukemia and 7 acute lymphocytic leukemia cases; other: 5 non-Hodgkin lymphoma, 4 myelodysplastic syndrome, 1 chronic myeloid leukemia, 1 chronic lymphocytic leukemia and 1 multiple myeloma cases.*

The standard antimicrobial prophylaxis in our institution included oral levofloxacin, antiviral prophylaxis with acyclovir or valacyclovir, and antifungal prophylaxis with echinocandins or azoles according to the patient's risk of fungal infection. In addition, cephalosporin and antibiotics for anaerobic bacteria (metronidazole, meropenem or piperacillin/tazobactam) were administered to 70% and 57% of patients, respectively.

aGVHD was diagnosed and classified according to the Glucksberg grading system (23). Fifteen patients developed grade II-IV aGVHD and, of those, 6 developed severe aGVHD (grade III-IV). None of this cohort's clinical characteristics, including graft source, conditioning regimen, GVHD prophylaxis and antibiotics usage, was significantly associated with the risk of aGVHD (**Table S1**).

Dental Biofilm Microbiota Dysbiosis During Allo-HSCT

Supragingival biofilm samples were collected for bacterial profiling at preconditioning, aplasia, and engraftment to characterize changes in DBM during allo-HSCT. Three engraftment samples were excluded from downstream analyses due to insufficient high-quality reads.

DBM alpha diversity was assessed using the Shannon index. We observed a statistically significant decrease in DBM alpha diversity during allo-HSCT, with engraftment samples presenting the lowest overall bacterial diversity (median at each collection phase: 4.15, 3.39, and 2.75, respectively; **Figure 1A**). A similar decrease in alpha diversity was observed when using the Gini-Simpson index (**Figure S1A**) or the number of observed ASVs as a proxy for species richness (**Figure S1B**).

Marked changes in DBM genera composition were observed for all patients during allo-HSCT (**Figure S2**). As expected, several dental biofilm commensal genera were detected at a high average relative abundance at preconditioning, including *Streptococcus* (19.5%), *Veillonella* (18.4%), *Actinomyces* (6.3%), and *Capnocytophaga* (6.1%) (**Figure 1B**). However, their average relative abundance decreased during allo-HSCT. Likewise, we observed an increase in the average relative abundance of potentially pathogenic genera, such as *Enterococcus* and *Lactobacillus* (**Figure 1B**).

For a more quantitative assessment of DBM changes during allo-HSCT, we compared genera abundances at preconditioning and engraftment using the ANCOM test (**Figure 1C**). The most statistically significant differences in abundance were observed for *Enterococcus*, *Lactobacillus*, and *Mycoplasma*, confirming the expansion of these potentially pathogenic genera in DBM during allo-HSCT. We also observed statistically significant (although less pronounced in terms of relative abundance change) decreases in commensal genera (**Figure 1C**).

Dental Biofilm Microbiota Diversity and aGVHD Risk

Patients were stratified into two equal-sized groups (high and low-diversity groups) by the entire cohort's median alpha diversity value to evaluate the association between DBM diversity and aGVHD risk. Using the Shannon diversity index,

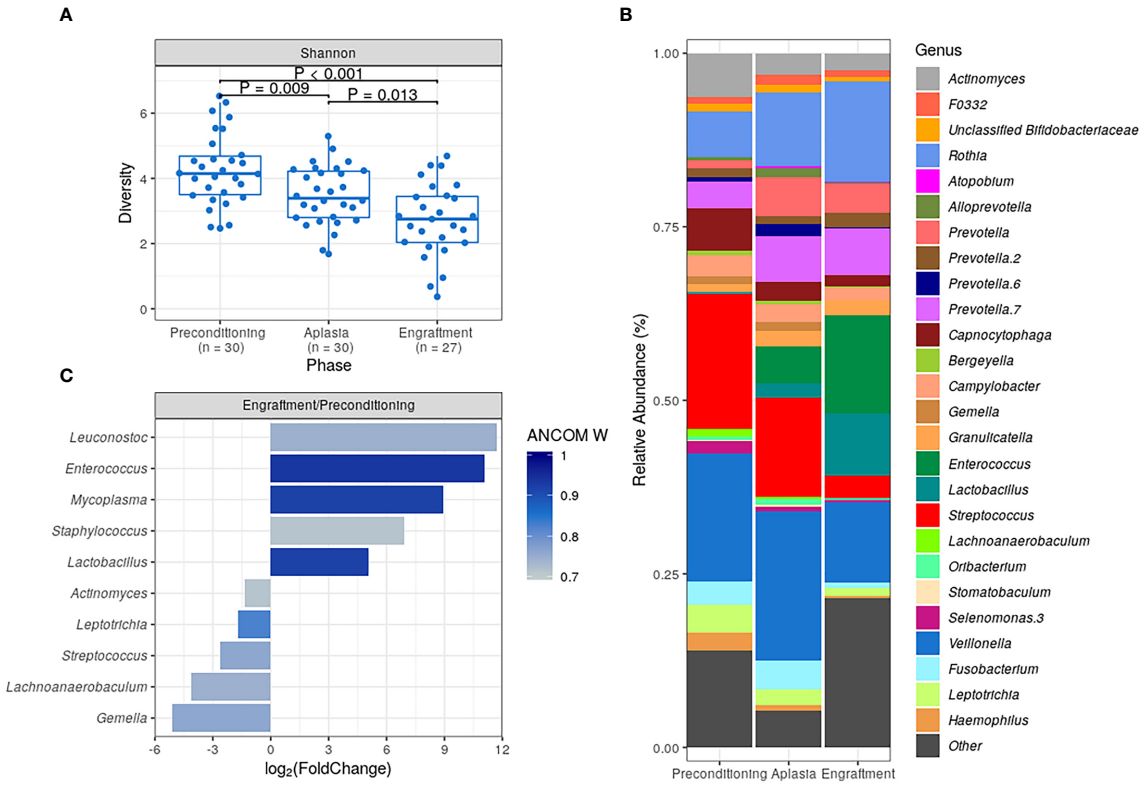


FIGURE 1 | Characterization of dental biofilm microbiota (DBM) during allogeneic hematopoietic stem cell transplantation. **(A)** DBM alpha diversity (Shannon) boxplots at preconditioning (n = 30), aplasia (n = 30) and engraftment (n = 27). Mann-Whitney U test was used with the preconditioning as the reference for comparisons. The boxes highlight the median value and cover the 25th and 75th percentiles, with whiskers extending to the more extreme value within 1.5 times the length of the box. **(B)** Average DBM genera relative abundance composition across transplantation phases. Only genera with at least 0.1% relative abundance in at least 25% study samples are shown. Taxa are sorted based on taxonomic relatedness. **(C)** Significant genera relative abundance variations from preconditioning to engraftment according to ANCOM test (W > 0.7). Log_2(Fold Change) for the average relative abundance variation (Engraftment/Preconditioning) is shown.

DBM diversity showed no association with the risk of aGVHD at preconditioning, aplasia, or engraftment (**Figures 2A–C** and **Table 2**). Similar results were obtained when using the Gini-Simpson diversity index or the number of observed ASVs as a proxy for species richness (**Figure S3**).

Dental Biofilm Microbiota Composition and aGVHD Risk

We then evaluated whether the abundance of specific genera at preconditioning, aplasia, or engraftment was associated with the risk of aGVHD (**Figure 3**). Only genera present at relative abundance ≥ 0.1% in at least 25% of the samples were considered for these analyses. Patients were stratified into two equal-sized groups (high and low relative abundance groups) by the median relative abundance observed in the entire cohort of each genus. *Veillonella*, *Streptococcus*, and *Corynebacterium* at preconditioning were significantly associated with the risk of aGVHD. We did not observe a similar association between the relative abundance of these or any other genus with the risk of aGVHD at aplasia or engraftment (**Figure 3A**).

Patients with high *Veillonella* relative abundance at preconditioning had a lower CMI of aGVHD (27% *vs.* 73%; HR = 0.24, 95% CI: 0.08–0.7, $P = 0.009$; **Figure 3B** and **Table 2**). This association remained significant after adjusting for graft source and intensity of the conditioning regimen (adjusted-HR = 0.21, 95% CI: 0.07–0.65, $P = 0.006$, **Table 2**). Patients with high *Streptococcus* or *Corynebacterium* relative abundance at preconditioning had a higher CMI of aGVHD (67% *vs.* 33%; HR = 2.89, 95% CI: 1.07–7.79, $P = 0.036$ and 73% *vs.* 37%; HR = 2.74, 95% CI: 1.05–7.15, $P = 0.04$, respectively; **Figures 3C, D** and **Table 2**). However, only *Streptococcus* remained significantly associated with the risk of aGVHD after adjusting for graft source and intensity of the conditioning regimen (adjusted-HR = 3.17, 95% CI: 1.12–9.01, $P = 0.03$, **Table 2**).

Veillonella and *Streptococcus* showed the highest average relative abundance at preconditioning (**Figure 1B**). Given their overall high relative abundance and an inverse association with the risk of aGVHD, we next evaluated whether the *Veillonella*/*Streptococcus* ratio at preconditioning was associated with the risk of aGVHD. Patients with a *Veillonella*/*Streptococcus* ratio >1 at preconditioning had a lower CMI of aGVHD (29% *vs.* 77%; HR = 0.23, 95% CI: 0.08–0.62, $P = 0.004$; **Figure 3E** and **Table 2**). Interestingly, the association between the *Veillonella*/*Streptococcus* ratio at preconditioning and aGVHD risk was

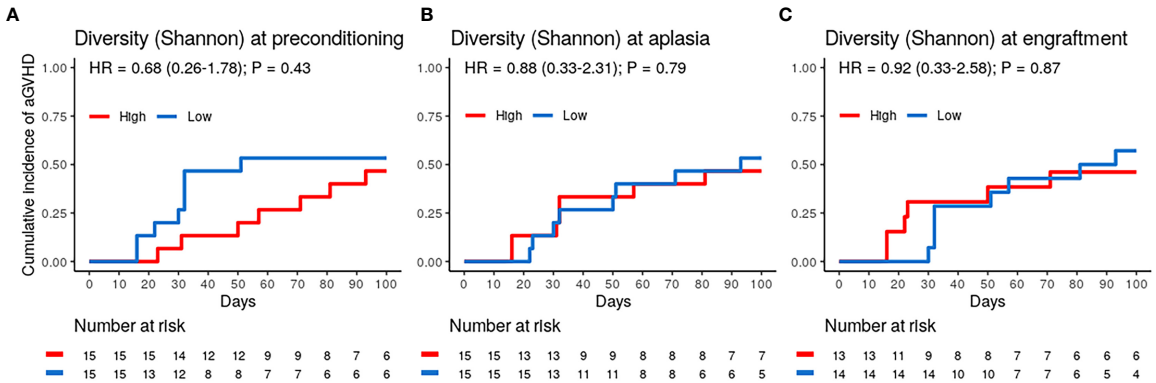


FIGURE 2 | Dental biofilm microbiota alpha diversity is not associated with the risk of acute graft-versus-host disease (aGVHD). **(A–C)** Cumulative incidence of aGVHD with patients stratified by Shannon diversity index (High *vs.* Low) at preconditioning (**A**; n = 30), aplasia (**B**; n = 30) or engraftment (**C**; n = 27).

TABLE 2 | Univariate (non-adjusted) and adjusted competing risk analyses for the association of acute graft-*versus*-host disease with relevant microbiota variables.

	Non-adjusted		Adjusted									
			***Veillonella* at P**		***Streptococcus* at P**		***Corynebacterium* at P**		**Ratio at P**		***E.faecalis* bloom**	
	HR (95% CI)	**P-value**	**HR (95% CI)**	**P-value**	**HR (95% CI)**	**P-value**	**HR (95% CI)**	**P-value**	**HR (95% CI)**	**P-value**	**HR (95% CI)**	**P-value**
Graft source (Bone Marrow)	0.95 (0.35-2.63)	0.92	1.42 (0.43-9.03)	0.38	0.75 (0.23-2.46)	0.64	1.42 (0.40-5.04)	0.59	0.78 (0.25-2.46)	0.67	1.63 (0.42-6.35)	0.49
Conditioning intensity (Myeloablative)	0.74 (0.26-2.17)	0.59	0.50 (0.11-2.32)	0.37	0.79 (0.24-2.61)	0.7	0.79 (0.20-3.04)	0.73	0.92 (0.27-3.16)	0.89	0.94 (0.24-3.61)	0.92
Diversity (Shannon) at P (High *vs.* Low)	0.68 (0.26-1.78)	0.43	–	–	–	–	–	–	–	–	–	–
Diversity (Shannon) at A (High *vs.* Low)	0.88 (0.33-2.31)	0.79	–	–	–	–	–	–	–	–	–	–
Diversity (Shannon) at E (High *vs.* Low)	0.92 (0.33-2.58)	0.87	–	–	–	–	–	–	–	–	–	–
***Veillonella* at P (High *vs.* Low)**	0.24 (0.08-0.70)	**0.009**	0.21 (0.07-0.65)	**0.006**	–	–	–	–	–	–	–	–
***Streptococcus* at P (High *vs.* Low)**	2.89 (1.07-7.79)	**0.036**	–	–	3.17 (1.12–9.01)	**0.03**	–	–	–	–	–	–
***Corynebacterium* at P (High *vs.* Low)**	2.74 (1.05-7.15)	**0.04**	–	–	–	–	2.79 (0.99-7.9)	0.053	–	–	–	–
Ratio at P (>1 *vs.* ≤1)	0.23 (0.08-0.62)	**0.004**	–	–	–	–	–	–	0.22 (0.08-0.64)	**0.005**	–	–
Ratio at A (>1 *vs.* ≤1)	0.45 (0.16-1.23)	0.12	–	–	–	–	–	–	–	–	–	–
Ratio at E (>1 *vs.* ≤1)	0.73 (0.27-1.98)	0.54	–	–	–	–	–	–	–	–	–	–
Any genus bloom (Yes *vs.* No)	2.29 (0.63-2.36)	0.21	–	–	–	–	–	–	–	–	–	–
***E. faecalis* bloom (Yes *vs.* No)**	4.07 (1.82-9.14)	**0.0007**	–	–	–	–	–	–	–	–	4.90 (1.66-14.5)	**0.004**

Each multivariate model adjusts for graft source and conditioning intensity. Statistically significant associations are marked in bold. HR, Hazard ratio; CI, Confidence interval; P, preconditioning; A, aplasia; E, engraftment.

stronger than the association observed for each genus separately and remained significant after adjusting for graft source and intensity of the conditioning regimen (adjusted-HR = 0.22, 95% CI: 0.08–0.64, *P*=0.005, **Table 2**). The *Veillonella*/*Streptococcus* ratio at aplasia or engraftment was not associated with the risk of aGVHD (**Table 2**).

Enterococcus faecalis Bloom and aGVHD Risk

Finally, we analyzed whether the blooming of potentially pathogenic genera observed during allo-HSCT was associated with the risk of aGVHD. For these analyses, bloom was defined as the sudden expansion of a particular genus from near absence

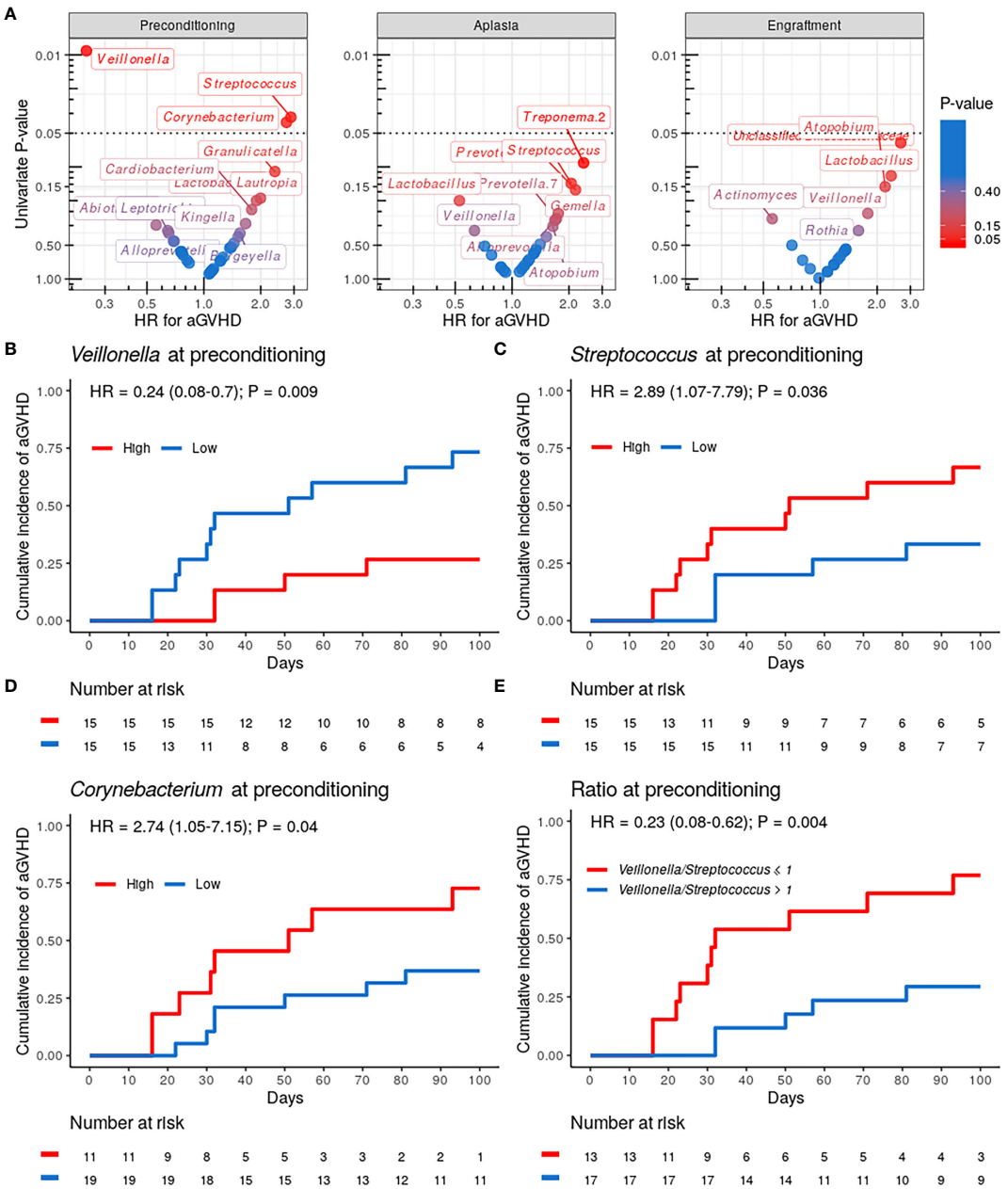


FIGURE 3 | Specific genera relative abundance at preconditioning are associated with the risk of acute graft-*versus*-host disease (aGVHD). **(A)** Volcano plot for the univariate competing risk analysis for the association of aGVHD with genera relative abundance (hazard ratio *vs.* P-value) at preconditioning (left), aplasia (center) and engraftment (right). Only genera with ≥0.1% relative abundance in at least 25% of samples at a given phase were evaluated. Genera with P-value < 0.4 for the association are indicated explicitly. **(B–D)** Cumulative incidence of aGVHD with patients (n = 30) stratified by either *Veillonella* **(B)**, *Streptococcus* **(C)** or *Corynebacterium* **(D)** relative abundance at preconditioning (High *vs.* Low). **(E)** Cumulative incidence of aGVHD with patients (n = 30) stratified by *Veillonella*/*Streptococcus* relative abundance ratio at preconditioning (>1 *vs.* ≤1).

(relative abundance <1% at preconditioning) to dominance (relative abundance ≥30% at aplasia or engraftment). Analyzing variations in genera relative abundance during allo-HSCT, we observed 23 blooms, involving 12 different genera and affecting a total of 20 patients. Three patients experienced more than one blooming event (**Figure S4**). Patients experiencing any genus bloom (n = 20) did not have altered aGVHD risk (**Table 2**). *Enterococcus* bloom was the most frequent event (**Figure 4A**), observed in 20% of the patients undergoing allo-HSCT. For all patients experiencing *Enterococcus* bloom except one, the phenomenon was attributed exclusively to *Enterococcus faecalis* expansion (**Figure 4B**). There was no association

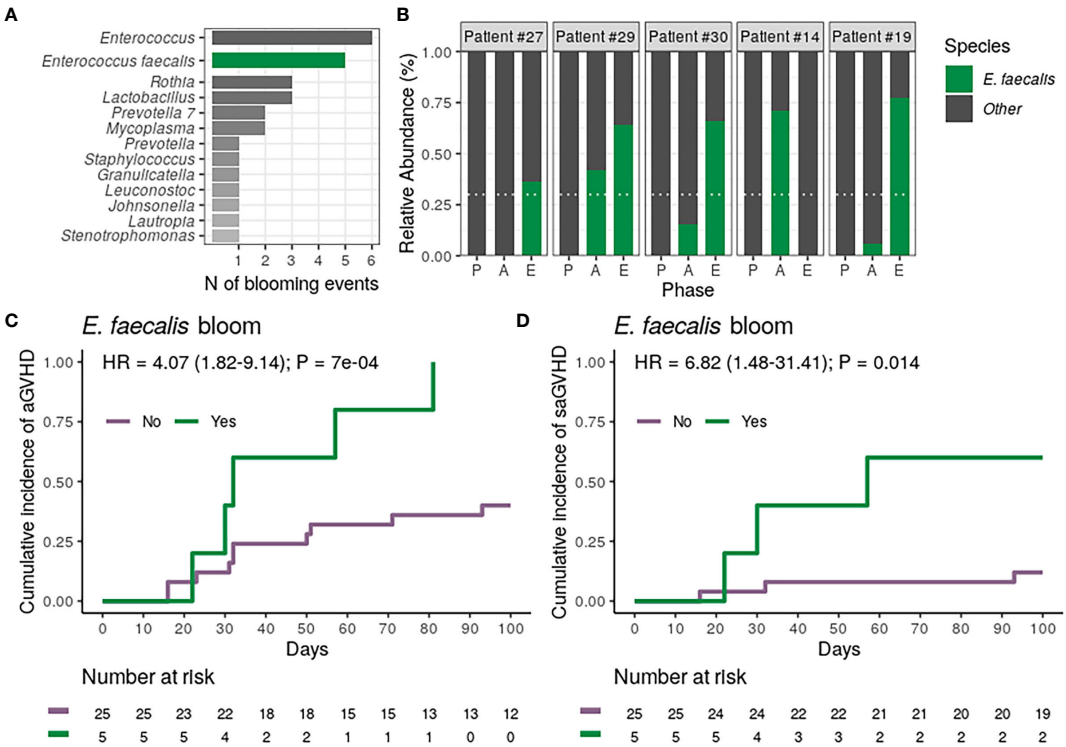


FIGURE 4 | Dental biofilm *Enterococcus faecalis* bloom during allogeneic hematopoietic stem cell transplantation is associated with a higher risk of acute graft-*versus*-host disease (aGVHD) and severe aGVHD (saGVHD). **(A)** Number of observed blooming events per genera in all patients (n = 30). The number of *Enterococcus* blooms caused exclusively by *Enterococcus faecalis* is indicated. **(B)** Relative abundance of *Enterococcus faecalis* across transplantation phases for all patients experiencing *Enterococcus faecalis* bloom (n = 5). Patients are sorted based on the highest *Enterococcus faecalis* relative abundance observed per patient. White horizontal dashed line indicates dominance threshold. P, Preconditioning; A, Aplasia; E, Engraftment. **(C, D)** Cumulative incidence of aGVHD **(C)** or saGVHD **(D)** with patients (n = 30) stratified by *Enterococcus faecalis* bloom occurrence (No *vs.* Yes).

between *E. faecalis* bloom and cephalosporin (Fisher's exact test, $P = 0.29$) or antibiotic for anaerobic bacteria usage (Fisher's exact test, $P = 1$).

We next tested whether the occurrence of *E. faecalis* bloom was associated with the risk of aGVHD. All patients experiencing *E. faecalis* bloom developed aGVHD, and *E. faecalis* bloom was strongly associated with a higher CMI of aGVHD (100% *vs.* 40%; HR = 4.07, 95% CI: 1.82–9.14, $P = 0.0007$; **Figure 4C** and **Table 2**). This association remained significant after adjusting for graft source and intensity of the conditioning regimen (adjusted-HR = 4.90, 95% CI: 1.66–14.50, $P = 0.004$, **Table 2**). Notably, CMI of severe aGVHD (grade III-IV) was higher in patients experiencing *E. faecalis* bloom (60% *vs.* 12%; HR = 6.82, 95% CI: 1.48–31.41, $P = 0.014$; **Figure 4D**; **Table 2**), revealing a direct association between DBM *E. faecalis* bloom and aGVHD risk and grade.

DISCUSSION

In our study, we describe, for the first time using high-throughput 16S rRNA sequencing, changes in DBM diversity and composition in 30 patients undergoing allo-HSCT. As observed for IM, DBM dysbiosis during allo-HSCT was marked by a gradual loss of bacterial diversity, with engraftment samples presenting the lowest overall bacterial diversity. Like for the IM, we also observed significant changes in DBM genera composition, with a decrease in the abundance of commensal core DBM genera, such as *Streptococcus* and *Actinomyces* (the only genera that can adhere to the tooth surface to start ordinary DB formation) (24), and overgrowths of potentially pathogenic bacteria, such as *Enterococcus*, *Lactobacillus*, and *Mycoplasma*. Most importantly, we observed that DBM genera relative abundance at preconditioning and changes in DBM composition during allo-HSCT (namely, *E. faecalis* bloom) were both predictive of aGVHD risk after allo-HSCT. There was no association between these aGVHD-associated microbiota variables and other allo-HSCT outcomes, including chronic GVHD (**Table S2**), as diagnosed in accordance with the NIH 2014 consensus (25).

aGVHD is a major cause of non-relapse mortality following allo-HSCT, with a one-year survival rate for patients developing severe aGVHD of only 40% (26). First-line therapy for aGVHD is based on corticosteroids, with response rates that vary between

40 and 70% (27). In this scenario, identifying biomarkers capable of predicting aGVHD risk and developing preventive therapies are critical.

Recently, the IM composition has been analyzed as a biomarker for clinical outcomes in allo-HSCT recipients, including the development of aGVHD (5, 7). Moreover, microbiota-based therapeutic interventions, including microbiota-driven antibiotics selection, alternative dietary regimens (including probiotics/prebiotics usage) and fecal microbiota transplantation have been proposed to prevent and treat aGVHD (28–32).

Like the IM, the OM plays an essential role in maintaining local and systemic health. Dental biofilm (DB) bacteria, as opposed to other shedding surface-living bacteria in the oral cavity, can adhere to hard surfaces and coaggregate (33), allowing the assembly of an organized three-dimensional structure, which confers DBM its distinctive ecological properties. The DBM interacts directly with host immune cells and modulates immune homeostasis (14). Moreover, DBM can also act as a microbial reservoir for systemic diseases. DBM dysbiosis can trigger local inflammation, destruction of surrounding periodontal tissue, and systemic translocation of oral microbes (24). The influence of the OM in systemic diseases such as colorectal cancer (34) and arthritis (35) has been increasingly studied. However, in the allo-HSCT context, studies are still limited and have focused mainly on the saliva and the tongue microbiota (36–39).

Loss of bacterial diversity in the salivary microbiota of patients undergoing allo-HSCT has been previously described and associated with oral mucositis (36). Likewise, a steep decline in the tongue microbiota diversity was observed in severe aplastic anemia patients from preconditioning to the day of transplantation (37). On the other hand, no appreciable changes in OM during allo-HSCT were observed in an additional study evaluating 4 different oral sites (buccal mucosa, saliva, tongue, and DB) (38). However, this latter study used a low-resolution methodology (microarray) for microbiota characterization in a small number of patients (n = 11). Noteworthy, a single study evaluated the association between OM and allo-HSCT outcomes (39). Allo-HSCT recipients showed a less diverse and distinct tongue microbiota on the day of transplantation than that of community-dwelling adults. In this study, the presence of the non-commensal bacteria *Staphylococcus haemolyticus* and/or *Ralstonia pickettii* in the tongue microbiota was significantly associated with lower overall survival after allo-HSCT, but not with aGVHD.

Out of the many allo-HSCT outcomes evaluated so far (40), aGVHD onset has the clearest causal connection to the IM (28, 29, 40). Briefly, it has been shown that the loss of commensal bacteria (especially SCFA-producing Clostridia species) during the conditioning regimen reduces the intestinal concentration of butyrate and indole-3-aldehyde (41, 42). Low levels of these metabolites compromise mucosal integrity (42, 43), promoting extravasation of bacterial lipopolysaccharide and activation of donor reactive T cells (40). Additionally, *Enterococcus faecalis* might contribute to aGVHD development *via* production of metalloproteases that impair barrier function (44) and by stimulating macrophages to secrete TNF (45). Accordingly, low IM diversity at the time of stem cell engraftment (6, 7), low abundance of commensal bacteria from Clostridia class (7, 8), and intestinal enterococci dominance during allo-HSCT (10) have been all associated with worsened aGVHD-related outcomes in studies evaluating stool specimens from allo-HSCT recipients (28, 29, 40).

In our study, DBM diversity was not associated with the risk of aGVHD in any transplantation phase evaluated, which is in line with a recent IM study that did not find differences in IM diversity between aGVHD groups neither pre- nor post-transplantation (46). Also, despite the presence (as expected (47)) of many Clostridia genera in DBM (such as *Oribacterium*), we did not find DBM Clostridia class members significantly associated with the risk of aGVHD. However, as for the IM, we observed a decrease in the relative abundance of several DB commensal genera during allo-HSCT, such as *Streptococcus*, *Veillonella*, *Actinomyces*, and *Capnocytophaga*, and an increase in the relative abundance of potentially pathogenic genera such as *Enterococcus* and *Lactobacillus*. Most importantly, high *Streptococcus* and high *Corynebacterium* relative abundance at preconditioning were associated with a higher risk of aGVHD, while high *Veillonella* relative abundance at preconditioning was associated with a lower risk of aGVHD.

Streptococci, corynebacteria, and veillonellae are part of the core DBM (48) and represent the 1st, 2nd and 10th most important genera in terms of relative abundance in healthy volunteers DBM, respectively (47). In our study, streptococci and veillonellae showed the highest average relative abundance at preconditioning and were both associated with the risk of aGVHD. Given their overall high relative abundance and the relative nature of the data, higher *Veillonella* relative abundance imposes lower *Streptococcus* relative abundance and vice versa. Hence, it is not possible to determine whether both genera are genuinely associated with the risk of aGVHD. Interestingly, the association between the *Veillonella/Streptococcus* ratio at preconditioning and aGVHD risk, independently of the conditioning regimen and graft source, was stronger than the association observed for each genus separately, suggesting a partial role for both genera in the observed effect.

During DB formation, bacterial early colonizers, after adhering to teeth salivary pellicles, coaggregate with other early and late colonizers, and a repeatable microbial succession takes place on the tooth surface (33). Streptococci are the most abundant microbe in DB, representing a predominant early colonizer with broad coaggregation partnerships. Streptococci and veillonellae are in close physical contact during the early phases of DB maturation (33, 49) and can grow together in a metabolic cooperation-dependent manner (33, 49). Since this interaction occurs in the early phases of DB formation (and therefore are instrumental for DB maturation), the ratio *Veillonella/Streptococcus* might be a marker of early DBM disruption associated with a higher risk of aGVHD.

Corynebacteria bridge the early biofilm members to late colonizers (48). In contradiction with the documented in the

aforementioned healthy volunteers study (47), we did not observe a high corynebacteria average relative abundance in any of the allo-HSCT phases evaluated. It is possible that the overall lower relative abundance of corynebacteria in detriment of early colonizers (such as streptococci and veillonellae) in our study may be indicative of a basal DBM disruption afflicting all allo-HSCT recipients. Alternatively, the lower relative abundance of corynebacteria may be explained by the stricter oral hygiene protocol recommended to our patients.

Finally, in our study, *E. faecalis* bloom in the DBM was observed in 17% of allo-HSCT recipients and was significantly associated with a higher risk of aGVHD and saGVHD. Noteworthy, despite recent *in vitro* evidence suggesting that high-dose of cephalosporin may promote *E. faecalis* biofilm formation (50), there was no association between cephalosporin usage and DBM *E. faecalis* bloom in the evaluated cohort.

During allo-HSCT, intestinal enterococci expansion is well documented and is linked to both aGVHD development (10) and subsequent bacteremia (51). Notably, *E. faecalis* alone exacerbates aGVHD severity in gnotobiotic mouse models (10). Our study reveals an additional site with enterococci expansion that might have systemic impacts after allo-HSCT. We can speculate that, during allo-HSCT, the dysbiotic DBM may act as an enterococci reservoir, triggering translocation to the gut and intestinal enterococci domination. This possibility is corroborated by the fact that there is intense oral bacteria translocation to the gut in hepatic cirrhosis patients (52) and that such translocations in colorectal cancer patients are negatively correlated with intestinal Clostridia bacteria presence (34). Indeed, oral bacteria translocation to the gut has been described in allo-HSCT recipients, and the presence of oral Actinobacteria and oral Firmicutes in stool samples of these patients was positively correlated with subsequent aGVHD development (5). Alternatively, DBM enterococci may have an intestinal origin, since the injury to Goblet cells during conditioning regimen was shown to induce dissemination of dominant intestinal bacteria (28). Further studies evaluating synchronously IM and DBM are necessary to decipher whether IM and DBM enterococci bloom are linked and which event precedes the other. Importantly, enterococci are present in small amounts in the healthy OM (47) but may overgrow in pathogenic/dysbiotic settings, including after solid organ transplantation (53), in a biofilm-dependent manner (54). This may explain why previous microbiota studies on soft oral sites have not reported the expansion of *Enterococcus* in allo-HSCT recipients.

Our study has many limitations. As a pioneering and exploratory work, it is single-centered and has a limited sample size. Besides, the study patients analyzed are heterogeneous and encompass several underlying diseases. Therefore, validation cohorts and multicentric prospective studies are needed to confirm our findings. We also emphasize that the associations reported herein are correlative, so that further studies on DBM during allo-HSCT that include synchronous fecal sampling and metabolomics analyses are needed to associate DBM dysbiosis with aGVHD pathophysiology.

Although patients usually receive rigorous oral health care during allo-HSCT (55), OM dysbiosis has been overlooked. Common oral care protocols already used in allo-HSCT patients to prevent and counteract oral health decay can also be used to directly (e.g. chlorhexidine mouthwash) or indirectly (e.g. photobiomodulation) modulate the OM. However, as the role of oral microbes in allo-HSCT outcomes become more prominent, complementary odontologic/pharmacologic interventions targeting specific sites and bacteria of the OM will be necessary. For instance, DBM dysbiosis could be managed by antimicrobial photodynamic therapy, which can eliminate pathogens with no risk of the emergence of drug-resistant strains (56). DBM dysbiosis could also be countervailed with the use of nanoparticles that alters DBM composition by interfering in fundamental biofilm properties such as adhesion and quorum-sensing (57, 58). These innovative approaches will be instrumental to evaluate whether early interventions to correct DBM dysbiosis can prevent aGVHD onset.

In conclusion, to our knowledge, this is the first study evaluating the DBM during allo-HSCT using a high-resolution technique. We identified markers of DBM dysbiosis during allo-HSCT. Most importantly, we showed that DBM composition during allo-HSCT may be predictive of aGVHD onset after transplantation, providing a simple and reproducible protocol for collection and analysis of allo-HSCT recipients microbiota before transplantation that may substitute fecal sampling when evaluating gastrointestinal dysbiosis and *Enterococcus* bloom.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: European Nucleotide Archive (ENA), PRJEB42862.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Hospital Sírio-Libanês (Protocol #1.414.217). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AC and EF designed the study. FK, PA, VH, JB, and WM-S performed the sequencing. VH and AC developed the bioinformatics pipeline. VH, JB, VM, CA-R, EF, and AC contributed to the analysis and interpretation of data. VH, JB, VM, CA-R, and AC contributed to writing the manuscript.

All authors contributed to the article and approved the submitted version.

FUNDING

VH was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, process no. 13996-0/2018). VM was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, process no. 141575/2018-2). JB was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, process no. 001).

ACKNOWLEDGMENTS

A manuscript regarding this work has been previously submitted to medRxiv as a preprint (59).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2021.692225/full#supplementary-material

REFERENCES

1. Styczyński J, Tridello G, Koster L, Iacobelli S, van Biezen A, van der Werf S, et al. Death After Hematopoietic Stem Cell Transplantation: Changes Over Calendar Year Time, Infections and Associated Factors. *Bone Marrow Transpl* (2020) 55(1):126–36. doi: 10.1038/s41409-019-0624-z
2. Schoemans HM, Lee SJ, Ferrara JL, Wolff D, Levine JE, Schultz KR, et al. EBMT-NIH-CIBMTR Task Force Position Statement on Standardized Terminology & Guidance for Graft-Versus-Host Disease Assessment. *Bone Marrow Transpl* (2018) 53(11):1401–15. doi: 10.1038/s41409-018-0204-7
3. Greinix HT, Eikema D-J, Koster L, Penack O, Yakoub-Agha I, Montoto S, et al. Incidence of Acute Graft-Versus-Host Disease and Survival After Allogeneic Hematopoietic Cell Transplantation Over Time: A Study From the Transplant Complications and Chronic Malignancies Working Party of the EBMT. *Blood* (2018) 132(Supplement 1):2120–0. doi: 10.1182/blood-2018-99-111764
4. Zeiser R, Blazar BR. Acute Graft-versus-Host Disease - Biologic Process, Prevention, and Therapy. *N Engl J Med* (2017) 377(22):2167–79. doi: 10.1056/NEJMra1609337
5. Golob JL, Pergam SA, Srinivasan S, Fiedler TL, Liu C, Garcia K, et al. Stool Microbiota at Neutrophil Recovery Is Predictive for Severe Acute Graft vs Host Disease After Hematopoietic Cell Transplantation. *Clin Infect Dis Off Publ Infect Dis Soc Am* (2017) 65(12):1984–91. doi: 10.1093/cid/cix699
6. Taur Y, Jenq RR, Perales M-A, Littmann ER, Morjaria S, Ling L, et al. The Effects of Intestinal Tract Bacterial Diversity on Mortality Following Allogeneic Hematopoietic Stem Cell Transplantation. *Blood* (2014) 124 (7):1174–82. doi: 10.1182/blood-2014-02-554725
7. Jenq RR, Taur Y, Devlin SM, Ponce DM, Goldberg JD, Ahr KF, et al. Intestinal Blautia Is Associated With Reduced Death From Graft-Versus-Host Disease. *Biol Blood Marrow Transplant J Am Soc Blood Marrow Transpl* (2015) 21 (8):1373–83. doi: 10.1016/j.bbmt.2015.04.016
8. Payen M, Nicolis I, Robin M, Michonneau D, Delannoye J, Mayeur C, et al. Functional and Phylogenetic Alterations in Gut Microbiome Are Linked to Graft-Versus-Host Disease Severity. *Blood Adv* (2020) 4(9):1824–32. doi: 10.1182/bloodadvances.2020001531
9. Holler E, Butzhammer P, Schmid K, Hundsrucker C, Koestler J, Peter 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 J Am Soc Blood Marrow Transpl* (2014) 20(5):640–5. doi: 10.1016/j.bbmt.2014.01.030
10. Stein-Thoeringer CK, Nichols KB, Lazrak A, Docampo MD, Slingerland AE, Slingerland JB, et al. Lactose Drives Enterococcus Expansion to Promote Graft-Versus-Host Disease. *Science* (2019) 366(6469):1143–9. doi: 10.1126/science.aax3760
11. Olsen I, Yamazaki K. Can Oral Bacteria Affect the Microbiome of the Gut? *J Oral Microbiol* (2019) 11(1):1586422. doi: 10.1080/20002297.2019.1586422
12. Atarashi K, Suda W, Luo C, Kawaguchi T, Motoo I, Narushima S, et al. Ectopic Colonization of Oral Bacteria in the Intestine Drives TH1 Cell Induction and Inflammation. *Science* (2017) 358(6361):359–65. doi: 10.1126/science.aan4526
13. Sharma N, Bhatia S, Sodhi AS, Batra N. Oral Microbiome and Health. *AIMS Microbiol* (2018) 4(1):42–66. doi: 10.3934/microbiol.2018.1.42
14. Moutsopoulos NM, Konkel JE. Tissue-Specific Immunity at the Oral Mucosal Barrier. *Trends Immunol* (2018) 39(4):276–87. doi: 10.1016/j.it.2017.08.005
15. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of General 16S Ribosomal RNA Gene PCR Primers for Classical and Next-Generation Sequencing-Based Diversity Studies. *Nucleic Acids Res* (2013) 41(1):e1. doi: 10.1093/nar/gks808
16. Wang H, Altemus J, Niazi F, Green H, Calhoun BC, Sturgis C, et al. Breast Tissue, Oral and Urinary Microbiomes in Breast Cancer. *Oncotarget* (2017) 8 (50):88122–38. doi: 10.18632/oncotarget.21490
17. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2. *Nat Biotechnol* (2019) 37(8):852–7. doi: 10.1038/s41587-019-0209-9
18. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-Resolution Sample Inference From Illumina Amplicon Data. *Nat Methods* (2016) 13(7):581–3. doi: 10.1038/nmeth.3869
19. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools. *Nucleic Acids Res* (2013) 41(Database issue):D590–596. doi: 10.1093/nar/gks1219
20. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME Improves Sensitivity and Speed of Chimera Detection. *Bioinforma Oxf Engl* (2011) 27 (16):2194–200. doi: 10.1093/bioinformatics/btr381
21. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: A Versatile Open Source Tool for Metagenomics. *PeerJ* (2016) 4:e2584. doi: 10.7717/peerj.2584
22. Kaul A, Mandal S, Davidov O, Peddada SD. Analysis of Microbiome Data in the Presence of Excess Zeros. *Front Microbiol* (2017) 8:2114. doi: 10.3389/fmicb.2017.02114
23. Glucksberg H, Storb R, Fefer A, Buckner CD, Neiman PE, Clift RA, et al. Clinical Manifestations of Graft-Versus-Host Disease in Human Recipients of Marrow From HL-A-Matched Sibling Donors. *Transplantation* (1974) 18 (4):295–304. doi: 10.1097/00007890-197410000-00001
24. Vieira Colombo AP, Magalhães CB, Hartenbach FARR, Martins do Souto R, Maciel da Silva-Boghossian C. Periodontal-Disease-Associated Biofilm: A Reservoir for Pathogens of Medical Importance. *Microb Pathog* (2016) 94:27–34. doi: 10.1016/j.micpath.2015.09.009
25. Jagasia MH, Greinix HT, Arora M, Williams KM, Wolff D, Cowen EW, et al. National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-Versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group Report. *Biol Blood Marrow Transplant J Am Soc Blood Marrow Transpl* (2015) 21(3):389–401. doi: 10.1016/j.bbmt.2015.05.004
26. Khoury HJ, Wang T, Hemmer MT, Couriel D, Alousi A, Cutler C, et al. Improved Survival After Acute Graft-Versus-Host Disease Diagnosis in the Modern Era. *Haematologica* (2017) 102(5):958–66. doi: 10.3324/haematol.2016.156356
27. Garnett C, Apperley JF, Pavlů J. Treatment and Management of Graft-Versus-Host Disease: Improving Response and Survival. *Ther Adv Hematol* (2013) 4 (6):366–78. doi: 10.1177/2040620713489842

28. Kumari R, Palaniyandi S, Hildebrandt GC. Microbiome: An Emerging New Frontier in Graft-Versus-Host Disease. *Dig Dis Sci* (2019) 64(3):669–77. doi: 10.1007/s10620-018-5369-9
29. Staffas A, Burgos da Silva M, van den Brink MRM. The Intestinal Microbiota in Allogeneic Hematopoietic Cell Transplant and Graft-Versus-Host Disease. *Blood* (2017) 129(8):927–33. doi: 10.1182/blood-2016-09-691394
30. Taur Y, Coyte K, Schluter J, Robilotti E, Figueroa C, Gjonbalaj M, et al. Reconstitution of the Gut Microbiota of Antibiotic-Treated Patients by Autologous Fecal Microbiota Transplant. *Sci Transl Med* (2018) 10(460): eaap9489. doi: 10.1126/scitranslmed.aap9489
31. Spindelboeck W, Schulz E, Uhl B, Kashofer K, Aigelsreiter A, Zinke-Cerwenka W, et al. Repeated Fecal Microbiota Transplantations Attenuate Diarrhea and Lead to Sustained Changes in the Fecal Microbiota in Acute, Refractory Gastrointestinal Graft-Versus-Host-Disease. *Haematologica* (2017) 102(5): e210–3. doi: 10.3324/haematol.2016.154351
32. van Lier YF, Davids M, Haverkate NJE, de Groot PF, Donker ML, Meijer E, et al. Donor Fecal Microbiota Transplantation Ameliorates Intestinal Graft-Versus-Host Disease in Allogeneic Hematopoietic Cell Transplant Recipients. *Sci Transl Med* (2020) 12(556):eaaz8926. doi: 10.1126/scitranslmed.aaz8926
33. Kolenbrander PE, Palmer RJ, Periasamy S, Jakubovics NS. Oral Multispecies Biofilm Development and the Key Role of Cell-Cell Distance. *Nat Rev Microbiol* (2010) 8(7):471–80. doi: 10.1038/nrmicro2381
34. Flemer B, Warren RD, Barrett MP, Cisek K, Das A, Jeffery IB, et al. The Oral Microbiota in Colorectal Cancer Is Distinctive and Predictive. *Gut* (2018) 67 (8):1454–63. doi: 10.1136/gutjnl-2017-314814
35. Lorenzo D, GianVincenzo Z, Carlo Luca R, Karan G, Jorge V, Roberto M, et al. Oral-Gut Microbiota and Arthritis: Is There an Evidence-Based Axis? *J Clin Med* (2019) 8(10):1753. doi: 10.3390/jcm8101753
36. Shouval R, Eshel A, Dubovski B, Kuperman AA, Danylesko I, Fein JA, et al. Patterns of Salivary Microbiota Injury and Oral Mucositis in Recipients of Allogeneic Hematopoietic Stem Cell Transplantation. *Blood Adv* (2020) 4 (13):2912–7. doi: 10.1182/bloodadvances.2020001827
37. Ames NJ, Barb JJ, Ranucci A, Kim H, Mudra SE, Cashion AK, et al. The Oral Microbiome of Patients Undergoing Treatment for Severe Aplastic Anemia: A Pilot Study. *Ann Hematol* (2019) 98(6):1351–65. doi: 10.1007/s00277-019-03599-w
38. Ames NJ, Sulima P, Ngo T, Barb J, Munson PJ, Paster BJ, et al. A Characterization of the Oral Microbiome in Allogeneic Stem Cell Transplant Patients. *PloS One* (2012) 7(10):e47628. doi: 10.1371/journal.pone.0047628
39. Oku S, Takeshita T, Futatsuki T, Kageyama S, Asakawa M, Mori Y, et al. Disrupted Tongue Microbiota and Detection of Nonindigenous Bacteria on the Day of Allogeneic Hematopoietic Stem Cell Transplantation. *PloS Pathog* (2020) 16(3):e1008348. doi: 10.1371/journal.ppat.1008348
40. Shono Y, van den Brink MRM. Gut Microbiota Injury in Allogeneic Haematopoietic Stem Cell Transplantation. *Nat Rev Cancer* (2018) 18 (5):283–95. doi: 10.1038/nrc.2018.10
41. Ganapathy V, Thangaraju M, Prasad PD, Martin PM, Singh N. Transporters and Receptors for Short-Chain Fatty Acids as the Molecular Link Between Colonic Bacteria and the Host. *Curr Opin Pharmacol* (2013) 13(6):869–74. doi: 10.1016/j.coph.2013.08.006
42. Ferrara JLM, Chaudhry MS. GVHD: Biology Matters. *Blood Adv* (2018) 2 (22):3411–7. doi: 10.1182/bloodadvances.2018020214
43. Mathewson ND, Jenq R, Mathew AV, Koenigsknecht M, Hanash A, Toubai T, et al. Gut Microbiome-Derived Metabolites Modulate Intestinal Epithelial Cell Damage and Mitigate Graft-Versus-Host Disease. *Nat Immunol* (2016) 17 (5):505–13. doi: 10.1038/ni.3400
44. Steck N, Hoffmann M, Sava IG, Kim SC, Hahne H, Tonkonogy SL, et al. Enterococcus Faecalis Metalloprotease Compromises Epithelial Barrier and Contributes to Intestinal Inflammation. *Gastroenterology* (2011) 141(3):959–71. doi: 10.1053/j.gastro.2011.05.035
45. Kim SO, Sheikh HI, Ha S-D, Martins A, Reid G. G-CSF-mediated Inhibition of JNK Is a Key Mechanism for Lactobacillus Rhamnosus-Induced Suppression of TNF Production in Macrophages. *Cell Microbiol* (2006) 8 (12):1958–71. doi: 10.1111/j.1462-5822.2006.00763.x
46. Ilett EE, Jørgensen M, Noguera-Julian M, Nørgaard JC, Daugaard G, Helleberg M, et al. Associations of the Gut Microbiome and Clinical Factors With Acute GVHD in Allogeneic HSCT Recipients. *Blood Adv* (2020) 4(22):5797–809. doi: 10.1182/bloodadvances.2020002677
47. Segata N, Haake SK, Mannon P, Lemon KP, Waldron L, Gevers D, et al. Composition of the Adult Digestive Tract Bacterial Microbiome Based on Seven Mouth Surfaces, Tonsils, Throat and Stool Samples. *Genome Biol* (2012) 13(6):R42. doi: 10.1186/gb-2012-13-6-r42
48. Mark Welch JL, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG. Biogeography of a Human Oral Microbiome at the Micron Scale. *Proc Natl Acad Sci USA* (2016) 113(6):E791–800. doi: 10.1073/pnas.1522149113
49. Mashima I, Nakazawa F. The Interaction Between Streptococcus Spp. and Veillonella Tobetsuensis in the Early Stages of Oral Biofilm Formation. *J Bacteriol* (2015) 197(3):2104–11. doi: 10.1128/JB.02512-14
50. Thieme L, Klinger-Strobel M, Hartung A, Stein C, Makarewicz O, Pletz MW. In Vitro Synergism and Anti-Biofilm Activity of Ampicillin, Gentamicin, Ceftaroline and Ceftriaxone Against Enterococcus Faecalis. *J Antimicrob Chemother* (2018) 73(6):1553–61. doi: 10.1093/jac/dky051
51. Taur Y, Xavier JB, Lipuma L, Ubeda C, Goldberg J, Gobourne A, et al. Intestinal Domination and the Risk of Bacteremia in Patients Undergoing Allogeneic Hematopoietic Stem Cell Transplantation. *Clin Infect Dis Off Publ Infect Dis Soc Am* (2012) 55(7):905–14. doi: 10.1093/cid/cis580
52. Qin N, Yang F, Li A, Prifti E, Chen Y, Shao L, et al. Alterations of the Human Gut Microbiome in Liver Cirrhosis. *Nature* (2014) 513(7516):59–64. doi: 10.1038/nature13568
53. Diaz PI, Hong B-Y, Frias-Lopez J, Dupuy AK, Angeloni M, Abusleme L, et al. Transplantation-Associated Long-Term Immunosuppression Promotes Oral Colonization by Potentially Opportunistic Pathogens Without Impacting Other Members of the Salivary Bacteriome. *Clin Vaccine Immunol CVI* (2013) 20(6):920–30. doi: 10.1128/CVI.00734-12
54. Ch'ng J-H, Chong KKL, Lam LN, Wong JJ, Kline KA. Biofilm-Associated Infection by Enterococci. *Nat Rev Microbiol* (2019) 17(2):82–94. doi: 10.1038/s41579-018-0107-z
55. *Current Practice of Oral Care for Hematopoietic Stem Cell Transplant Patients: A Survey of the Eastern Mediterranean Blood and Marrow Transplantation Group - Pubmed*. Available at: https://pubmed.ncbi.nlm.nih.gov/33631114/.
56. Hu X, Huang Y-Y, Wang Y, Wang X, Hamblin MR. Antimicrobial Photodynamic Therapy to Control Clinically Relevant Biofilm Infections. *Front Microbiol* (2018) 9:1299. doi: 10.3389/fmicb.2018.01299
57. Zhang T, Kalimuthu S, Rajasekar V, Xu F, Yiu YC, Hui TKC, et al. Biofilm Inhibition in Oral Pathogens by Nanodiamonds. *Biomater Sci* (2021). doi: 10.1039/D1BM00608H
58. Shah S, Gaikwad S, Nagar S, Kulshrestha S, Vaidya V, Nawani N, et al. Biofilm Inhibition and Anti-Quorum Sensing Activity of Phytosynthesized Silver Nanoparticles Against the Nosocomial Pathogen Pseudomonas Aeruginosa. *Biofouling* (2019) 35(1):34–49. doi: 10.1080/08927014.2018.1563686
59. Heidrich V, Bruno JS, Knebel FH, Molla VC de, Miranda-Silva W, Asprino PF, et al. Dental Biofilm Microbiota Dysbiosis Is Associated With the Risk of Acute Graft-Versus-Host Disease After Allogeneic Hematopoietic Stem Cell Transplantation. *medRxiv* (2021). 2021.02.04.21251019. doi: 10.1101/2021.02.04.21251019

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

4. CHAPTER 3: Disruption of the oral microbiota is associated with a higher risk of relapse after allogeneic hematopoietic stem cell transplantation

scientific reports

Check for updates

OPEN

Disruption of the oral microbiota is associated with a higher risk of relapse after allogeneic hematopoietic stem cell transplantation

Vinícius Campos de Molla[1,2,7], Vitor Heidrich[3,4,7], Julia Stephanie Bruno[3,7], Franciele Hinterholz Knebel[3,7], Wanessa Miranda-Silva[3], Paula Fontes Asprino[3], Luciana Tucunduva[1], Vanderson Rocha[1,5,6], Yana Novis[1], Anamaria Aranha Camargo[3], Eduardo Rodrigues Fregnani[3] & Celso Arrais-Rodrigues[1,2]✉

Intestinal microbiota (IM) diversity and composition regulates host immunity and affects outcomes after allogeneic stem cell transplantation (allo-HSCT). We evaluated if the oral mucosa microbiota (OM) could impact the outcomes in patients who underwent allo-HSCT. Samples from the oral mucosa of 30 patients were collected at three time points: before the conditioning regimen, at aplasia, and at engraftment. We analyzed the associations of OM diversity and composition with allo-HSCT outcomes. Lower OM diversity at preconditioning was associated with a higher risk of relapse at 3 years (68% versus 33%, respectively; *P* = 0.04). Dominance (relative abundance ≥ 30%) by a single genus at preconditioning was also associated with a higher risk of relapse (63% versus 36% at 3 years, respectively; *P* = 0.04), as well as worse progression-free survival (PFS; 19% versus 55%, respectively; *P* = 0.01), and overall survival (OS) at 3 years (38% versus 81%, respectively; *P* = 0.02). In our study we observed that OM dysbiosis is associated with a higher risk of relapse and worse survival after allo-HSCT.

Abbreviations

95% Cis	Ninety-five percent confidence intervals
aGVHD	Acute GVHD
allo-HSCT	Allogeneic stem cell transplantation
ASVs	Amplicon sequencing variants
cGVHD	Chronic GVHD
DRI	Disease risk index
GVHD	Graft versus host disease
IM	Intestinal microbiota
NRM	Non-relapse mortality
OM	Oral mucosa microbiota
OS	Overall survival
PFS	Progression-free survival

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) remains the only therapeutic option for several hematological neoplasms[1]. Although transplant outcomes have markedly improved in recent decades, relapse

[1]Centro de Oncologia, Hospital Sírio Libanês, Rua Dona Adma Jafet, 91, São Paulo, SP 01308-050, Brazil. [2]Universidade Federal de São Paulo, São Paulo, SP, Brazil. [3]Centro de Oncologia Molecular, Hospital Sírio Libanês, São Paulo, SP, Brazil. [4]Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, SP, Brazil. [5]Hospital das Clínicas da Faculdade de Medicina, Universidade de São Paulo/ICESP, Sao Paulo, Brazil. [6]Churchill Hospital, NHS-BT, Oxford, UK. [7]These authors contributed equally: Vinícius Campos de Molla, Vitor Heidrich, Julia Stephanie Bruno and Franciele Hinterholz Knebel. ✉email: celsoarrais@gmail.com

of the underlying condition remains the leading cause of death after allo-HSCT[2]. Despite conflicting results, several risk factors have been shown to affect relapse, including the intensity of the conditioning regimen[3–5], pre-HSCT disease status[6], donor age[7,8], graft source[9], killer immunoglobulin-like receptor compatibility[10], graft versus host disease (GVHD) prophylaxis[11–13], and the occurrence of chronic GVHD (cGVHD)[14]. Infections, acute GVHD (aGVHD), cGVHD, and secondary neoplasia are the main causes of non-relapse mortality (NRM)[2,15]. The disease risk index (DRI) stratifies the risk of mortality in patients after allo-HSCT, according to diagnosis and disease status[16].

The intestinal microbiota (IM) has been shown to play a vital role in regulating host immunity[17] and improving antineoplastic activity[18,19]. In addition, IM disruption, characterized by significant changes in microbiota diversity and composition, is associated with allo-HSCT clinical outcomes. Common complications after allo-HSCT, such as infections, mucositis, and GVHD, are associated with significant changes in IM diversity and composition. In allo-HSCT, IM disruption is also associated with the incidence of GVHD[20–22], overall survival (OS)[23–26], and underlying disease relapse[27,28].

The human oral cavity harbors the second most abundant microbiota after the gastrointestinal tract. As observed for the IM, the oral microbiota (OM) directly influences human health[29]. OM disruption has been observed in several diseases, including diabetes, autoimmune diseases, endocarditis, gastrointestinal cancer, head and neck cancer[30–32], and acute lymphoblastic leukemia[33]. Changes in the OM in patients undergoing allo-HSCT are known to be associated with respiratory signs and symptoms[34] and oral mucositis[35]; however, no correlation between OM and allo-HSCT outcomes have been reported to date.

Accordingly, in this study, we evaluated whether the OM disruption is related to outcomes in patients who underwent allo-HSCT.

Methods

Patient characteristics and sample collection. We collected samples from the oral mucosa of patients who underwent allo-HSCT at Hospital Sírio Libanês, São Paulo, Brazil between January 2016 and April 2018.

Samples were collected by rubbing the dorsal tongue and andwith sterile swabs at three time points: before the conditioning regimen and before the oral medicine specialist intervention (preconditioning), at aplasia (defined as the first day of neutrophils under 0.5×10^3/uL), and at engraftment. All patients were requested not to perform oral hygiene for at least 6 h before collection. Informed consent was obtained from all patients before collection. The study was approved by the local ethics committee (Comite de Ética em Pesquisa—Hospital Sírio Libanês), according to the Declaration of Helsinki. No tissue was procured from prisoners in this study. All patients were examined by an oral medicine specialist for potential infections, and all followed the same protocol for oral mucositis prophylaxis with photobiomodulation and oral hygiene with fluoride toothpaste and 0.12% chlorhexidine mouthwash. The standard antimicrobial prophylaxis in our institution included oral levofloxacin, acyclovir, and antifungal prophylaxis according to the patient's risk of fungal infection (voriconazole for high-risk patients, and fluconazole for low risk patients).

DNA extraction. Bacterial cells were recovered from oral mucosa swabs through vortexing in TE buffer supplemented with 6 µL PureLink RNAse A (20 mg/mL; Thermo Fisher Scientific, Waltham, MA, USA). DNA was extracted using a QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol (DNA Purification from Blood or Body Fluids) and stored at − 80 °C.

16S rRNA amplicon sequencing. For 16S rRNA amplicon sequencing, 12.5 ng DNA and prevalidated primers[36] were used to amplify 16S rRNA hypervariable regions V3–V4. Amplicons were sequenced as described elsewhere[37] on an Illumina MiSeq platform (Illumina, San Diego, CA, USA).

Bioinformatics pipeline. Reads were demultiplexed, and primer sequences were removed using the MiSeq Reporter software. Within the QIIME 2 framework[38], using experiment-specific adaptive error models[39], forward and reverse sequences were filtered for quality and bimeras, denoised, and merged into consensus sequences with the DADA2 pipeline[40], generating unique amplicon sequencing variants (ASVs). ASVs were further filtered for chimeric sequences using the SILVA database[41] and UCHIME[42]. ASVs were taxonomically assigned using SILVA database and VSEARCH tool[43].

Statistical analyses. For alpha diversity analyses, the samples were rarefied to 12,500 reads before calculating the Shannon index, Simpson index, or the number of observed ASVs as bacterial diversity measures with the QIIME 2 *q2-diversity* plugin. Alpha diversity across groups was compared with the Mann–Whitney U test. OM diversity was classified based on the median Shannon index diversity measure across the study population at a given collection time point. Patients were classified as high diversity (above the Shannon index median) and low diversity (below Shannon index median). Fisher's exact tests and two-sided Student's t-tests were used to evaluate the associations between alpha diversity status and categorical and numerical clinical parameters, respectively. The relative abundance of each taxa was calculated with the QIIME 2 q2-taxa plugin. The taxa shown on relative abundance longitudinal plots are all those showing dominance (relative abundance ≥ 30%) in at least one study sample or relative abundance ≥ 5% in at least 25% of study samples. Differentially abundant genera across transplantation phases were identified using ANCOM test, with relative differences represented by the log-transformed average relative abundance fold change between groups. ANCOM W represents the proportion of null hypotheses rejected when sub-testing the differential abundance of a genus normalized by the abundance of each one of the genera in the dataset. W > 0.7 was considered as statistically significant. The relative abundance of a genus was considered to increase during allo-HSCT for a given patient when the relative

abundance at engraftment was greater than at preconditioning and the final relative abundance was ≥ 0.1%. The probabilities of progression-free survival (PFS) and OS were calculated using the Kaplan–Meier method and compared using log-rank tests. Cumulative incidence rates were calculated for aGVHD, cGVHD, NRM, and relapse/progression. Ninety-five percent confidence intervals (95% CIs) were estimated using the Greenwood formula. Adjusted probabilities for outcomes after transplantation were estimated using the Cox proportional hazards method (PFS and OS) and Fine-Gray risk regression model (aGVHD, cGVHD, NRM, and relapse/progression). The association between OM parameters and HSCT outcome was investigated in the final model after adjusting for the DRI. First-order interactions between OM parameters and each variable of interest were examined. The results are presented as relative risks of failure (adverse prognostic factors versus good prognostic factors), with 95% CIs and two-tailed *P* values. To examine the association between genus presence at preconditioning and relapse, only genera present in 25–75% of samples were evaluated, where presence was defined as relative abundance ≥ 0.1%. R software (version 3.5.0) and RStudio (version 1.2.5033) were used for statistical analyses. The statistical package *cmprsk* was used to evaluate relapse across groups with transplant-related death as the competing risk.

Ethics approval and consent to participate. The study was approved by the local ethics committee, according to the Declaration of Helsinki.

Results

Patient characteristics. Between January 2016 and April 2018, 30 patients who underwent allo-HSCT for hematologic malignancies and had oral mucosa samples collected were included in this study. The most common underlying diseases were acute myeloid leukemia and acute lymphoblastic leukemia (60%). Conditioning regimens and intensity, graft source, T-cell depletion, and other clinical characteristics are listed in Table 1. The underlying disease, disease status, and OM diversity at preconditioning are presented in Table S1. The median follow-up time for survivors was 41 (30–50) months.

Microbiota dynamics analyses. In total, 5,920,836 high-quality bacterial assigned sequencing reads were analyzed, representing 1723 unique ASVs. Out of the 90 samples sequenced, nine were excluded from diversity analyses owing to an insufficient number of high-quality reads (< 12,500 reads per sample, as determined using alpha diversity rarefaction curves) after the read-filtering steps employed in the pipeline. Therefore, adequate preconditioning samples were available for 27 of the 30 patients included in this study.

The intrasample bacterial diversity (Fig. 1A) and richness (Fig. S1) of OM samples decreased significantly during the clinical course. This drop in diversity is associated with changes in taxa relative abundance during the same period (Fig. S2). Notably, all patients showed bacterial dominance by a single genus after preconditioning. In Fig. 1B, we show three representative patients with major dominance (relative abundance > 80%) by a single genus (*Stenotrophomonas*, *Rothia*, and *Veillonella,* respectively) at engraftment.

For a broader assessment of the relative abundance changes from preconditioning to subsequent transplantation phases, we employed the ANCOM test at the genus level. We observed statistically significant variations in the abundance of both opportunistic pathogenic and commensal genera (Fig. S3). From preconditioning to aplasia, there was a significant increase in the abundance of the potentially pathogenic genera *Enterococcus* and *Lactobacillus,* which were even more increased in the engraftment phase in terms of relative abundance fold change from preconditioning. *Staphylococcus* and *Mycoplasma* were other potentially pathogenic genera increased at engraftment. Contrarily, there was a significant decrease in the abundance of the commensal genera *Haemophilus* (at aplasia) and *Gemella* (at engraftment).

A global increase of potentially pathogenic genera occurs during allo-HSCT. However, evaluating each patient individually, we noticed irregular changes in the relative abundance of those same genera from preconditioning to engraftment. An increase in the relative abundance of *Enterococcus*, *Lactobacillus*, *Staphylococcus,* and Mycoplasma was observed in 32%, 40%, 56%, and 68% of patients (Fig. S4). Patients who presented an increase in *Enterococcus* relative abundance had a higher incidence of cGVHD when compared with patients without the increase of relative abundance ($P = 0.03$). No other associations between the increase in the relative abundance of potentially pathogenic genera and allo-HSCT outcomes was observed (Table S2).

Impact of OM diversity on transplant outcomes. In order to elucidate the impact of OM bacterial diversity on allo-HSCT outcomes, we stratified patients into low or high diversity at each collection time (Table S3). A swimmer plot was used to illustrate these correlations at preconditioning (Fig. 2A). When we compared those with high or low OM diversity at preconditioning, no differences were found in PFS (36% versus 32%, respectively; hazard ratio [HR] 0.75, 95% CI 0.28–2.00, $P = 0.57$), or in OS at 3 years (54% versus 57%, respectively; HR 0.96, 95% CI 0.33–2.89, $P = 0.96$). We also did not observe any differences in aGVHD at 100 days (43% versus 62%, respectively; HR 1.77, 95% CI 0.66–4.81, $P = 0.26$). At 3 years, no difference between high and low diversity in the incidence of cGVHD (30% versus 7%, respectively; HR 4.79, 95% CI 0.56–40.8, $P = 0.15$), and NRM (18% vs. 0%, respectively, HR 4.12, 95% CI 0.86–19.32, $P = 0.07$). However, lower OM diversity at preconditioning was associated with a higher risk of relapse at 3 years when compared with higher diversity (68% versus 33%, respectively; HR, 95% CI, $P = 0.04$; Fig. 2B, Table S4).

Notably, 16 (59%) patients presented some type of bacterial dominance at preconditioning. Such events encompassed 4 different genera, all of which are oral commensal: *Streptococcus* (dominant in 9/16 patients) and *Veillonella* (dominant in 2/16 patients), both members of the Firmicutes phylum; *Neisseria* (dominant in 3/16 patients) and *Rothia* (dominant in 2/16 patients). Genus dominance was detected even among patients classified as having high diversity at preconditioning (Fig. 2A). The presence of dominance by any genus at preconditioning

	N = 30
Sex (male)	16 (53%)
Age in years (median, range)	50 (19–73)
Underlying disease	
Acute myeloid leukemia	18 (60%)
Acute lymphoblastic leukemia	7 (23%)
Non-hodgkin lymphoma	5 (17%)
Myelodysplastic syndrome	4 (13%)
Chronic lymphocytic leukemia	1 (3%)
Chronic myeloid leukemia	1 (3%)
Multiple Myeloma	1 (3%)
Conditioning intensity	
Reduced intensity	18 (60%)
Myeloablative	12 (40%)
Total body irradiation	11 (37%)
Pretransplant T-cell depletion	15 (50%)
Graft source	
Bone marrow	10 (33%)
Peripheral blood	20 (67%)
Donor	
Matched sibling	9 (30%)
Haploidentical	10 (33%)
Matched unrelated	9 (30%)
Mismatched unrelated	2 (7%)
Pretransplant comorbidity (HCT-CI)	
0	16 (53%)
1–2	8 (27%)
≥ 3	6 (20%)
Disease risk index	
Low–intermediate	17 (57%)
High	13 (43%)
Disease status at transplant	
First or second complete remission	22 (73%)
Third complete remission	2 (7%)
Partial remission or refractory disease	6 (20%)
GVHD prophylaxis	
MMF + CsA	11 (37%)
MTX + CsA	10 (33%)
MMF + CsA + PTCy	9 (30%)
Follow-up in months (median, range)	37 (25–46)

Table 1. Clinical characteristics of the study patients. *MMF* mycophenolate mofetil, *MTX* methotrexate, *CsA* cyclosporin A, *PTCy* post-transplant cyclophosphamide.

was also associated with an increased risk of relapse at 3 years when compared with the absence of dominance (63% versus 36%, respectively; HR 4.59, 95% CI 1.11–19, $P = 0.03$; Fig. 3A). When evaluating dominance by specific genera or types of genera at preconditioning, neither dominance by *Streptococcus* (56% versus 39%, respectively; HR 1.64, 95% CI 0.52–5.14, $P = 0.4$), nor dominance by facultative anaerobic genera (*Streptococcus* or *Rothia*; 56% versus 39%, respectively; HR 2.05, 95% CI 0.67–6.27, $P = 0.21$) were associated with an increased risk of relapse. Due to the very unequal group sizes, we could not evaluate the association between dominance by *Rothia* (2/27 patients), *Veillonella* (the only dominant anaerobe; 2/27 patients) or *Neisseria* (the only dominant aerobe; 3/27 patients) at preconditioning and the risk of relapse.

Additionally, the presence of dominance by any genus at preconditioning was associated with inferior PFS (19% versus 55%, respectively; HR 4.75, 95% CI 1.78–12.7, $P = 0.01$; Fig. 3B) and OS (38% versus 81%, respectively; HR 4.73, 95% CI 1.59–14.08, $P = 0.02$; Fig. 3C). No differences in aGVHD at 100 days (43% versus 63%, respectively; HR 0.50, 95% CI 0.18–1.37, $P = 0.18$), cGVHD at 3 years (19% versus 18%, respectively; HR 1.07, 95% CI 0.19–5.93, $P = 0.94$), or NRM at 3 years (20% versus 9%, respectively; HR 2.35, 95% CI 0.27–20.60, $P = 0.44$) were observed.

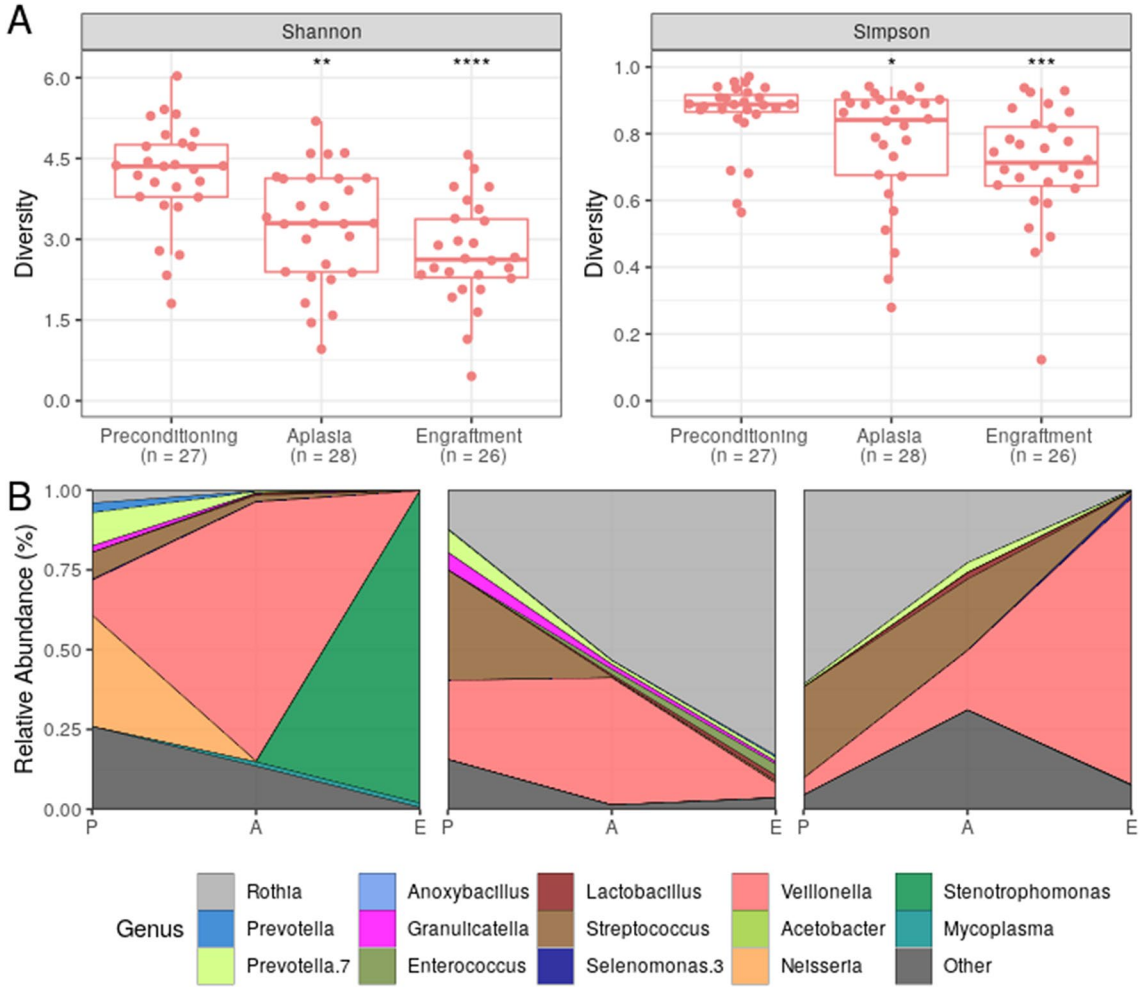


Figure 1. Bacterial diversity within the oral mucosa decreases during allo-HSCT. (**A**) Oral microbiota (OM) bacterial diversity boxplot at preconditioning (n = 27), aplasia (n = 28), and engraftment (n = 26) as measured by either Shannon index (left panel) or Simpson index (right panel). Mann–Whitney U tests were used with the preconditioning collection as the reference for comparisons. The boxes highlight the median values and cover the 25th and 75th percentiles, with whiskers extending to the more extreme value within 1.5 times the length of the box. Outliers are represented explicitly. Asterisks represent statistical significance: $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$. (**B**) OM genera relative abundance composition across transplantation phases for three representative patients showing the decrease in bacterial diversity. Only genera showing relative abundance ≥ 30% in at least one study sample or relative abundance ≥ 5% in at least 25% of study samples are shown. *P* preconditioning, *A* aplasia, *E* engraftment.

As expected, we also observed that patients with a high DRI had a significantly higher risk of relapse/progression, as compared with those with low-intermediate DRI at 3 years (62% versus 12%, respectively; HR 10.2, 95% CI 2.24–46.7, $P < 0.01$) and worse OS (77% versus 30%, respectively; HR 4.07, 95% CI 1.38–11.97, $P = 0.01$).

After adjusting analyses for the DRI, there was a trend toward a higher risk of relapse/progression in those with low OM diversity (HR 0.30, 95% CI 0.08–1.09, $P = 0.07$; Fig. 2C), and bacterial dominance of any genus remained significantly associated with the risk of relapse (HR 4.19, 95% CI 1.25–14.1, $P = 0.02$; Fig. 3D) and worse PFS (HR 4.14, 95% CI 1.15–14.89, $P = 0.03$; Fig. 3E). There was also a trend for bacterial dominance of any genus to be associated with worse OS (HR 4.12, 95% CI 0.89–19.13, $P = 0.07$; Fig. 3F).

Other relevant clinical parameters, such as conditioning intensity, underlying disease, and graft source, were not significantly associated with the risk of relapse (Fig. S5, Table S5).

Genus presence and transplant outcomes. As the genus level represents the most specific taxonomic level that still provides reliable taxonomic classification for V3–V4 amplicons, to further evaluate the association between preconditioning OM and transplant outcomes, we analyzed whether any non-core genus (those present in 25–75% of samples) was associated with a higher risk of relapse. In this exploratory analysis (without adjustment for multiple comparisons), of the 18 genera that matched the selection criteria tested in a univariate analysis for relapse (Fig. 4A, Fig. S6), only *Solobacterium* was significantly associated with lower relapse risk (9% versus 56%, respectively; HR 0.23, 95% CI 0.05–0.94, $P = 0.04$; Fig. 4B), and this association remained significant

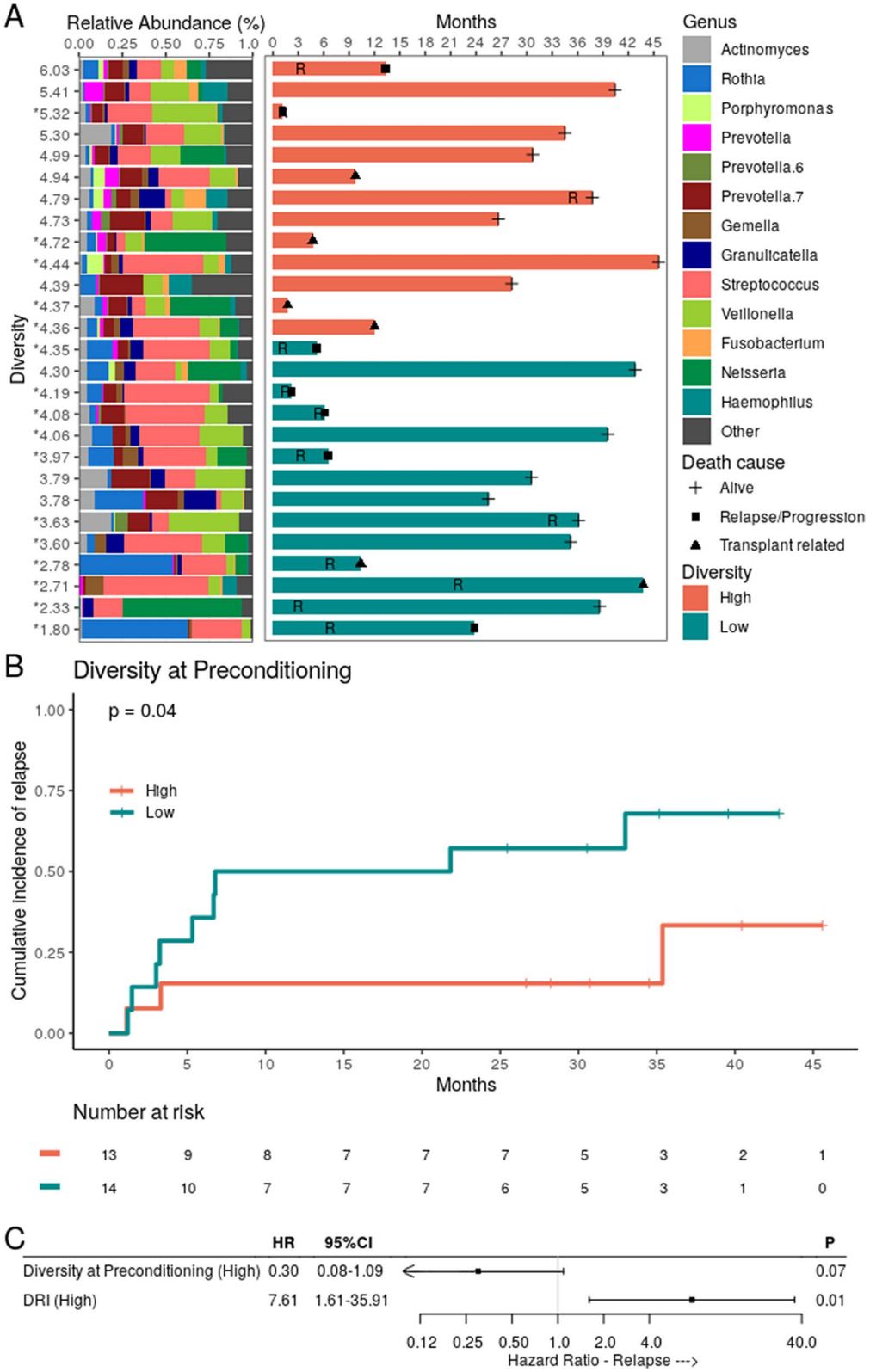


Figure 2. Oral microbiota bacterial dominance and bacterial diversity at preconditioning increased the risk of relapse in patients who underwent allo-HSCT. (**A**) Oral microbiota (OM) composition and diversity at preconditioning and the respective transplant course in each patient (n = 27). Patients are sorted based on descending Shannon diversity index, with the measures shown in the left subplot y-axis. The asterisk in the Shannon index indicates patients with at least one dominant (relative abundance > 30%) genus at preconditioning. Only genera showing relative abundance ≥ 30% in at least one preconditioning sample or relative abundance ≥ 5% in at least 10% of preconditioning samples are shown. Relevant outcomes (relapse and death) after infusion (aplasia) are shown in a timeline (in months) subplot (right). The plus sign represents censoring. *R* relapse. (**B**) Cumulative incidence of relapse with patients (n = 27) stratified by OM bacterial diversity at preconditioning (high versus low). (**C**) The DRI-adjusted hazard ratio for the association of OM bacterial diversity at preconditioning and relapse (n = 27).

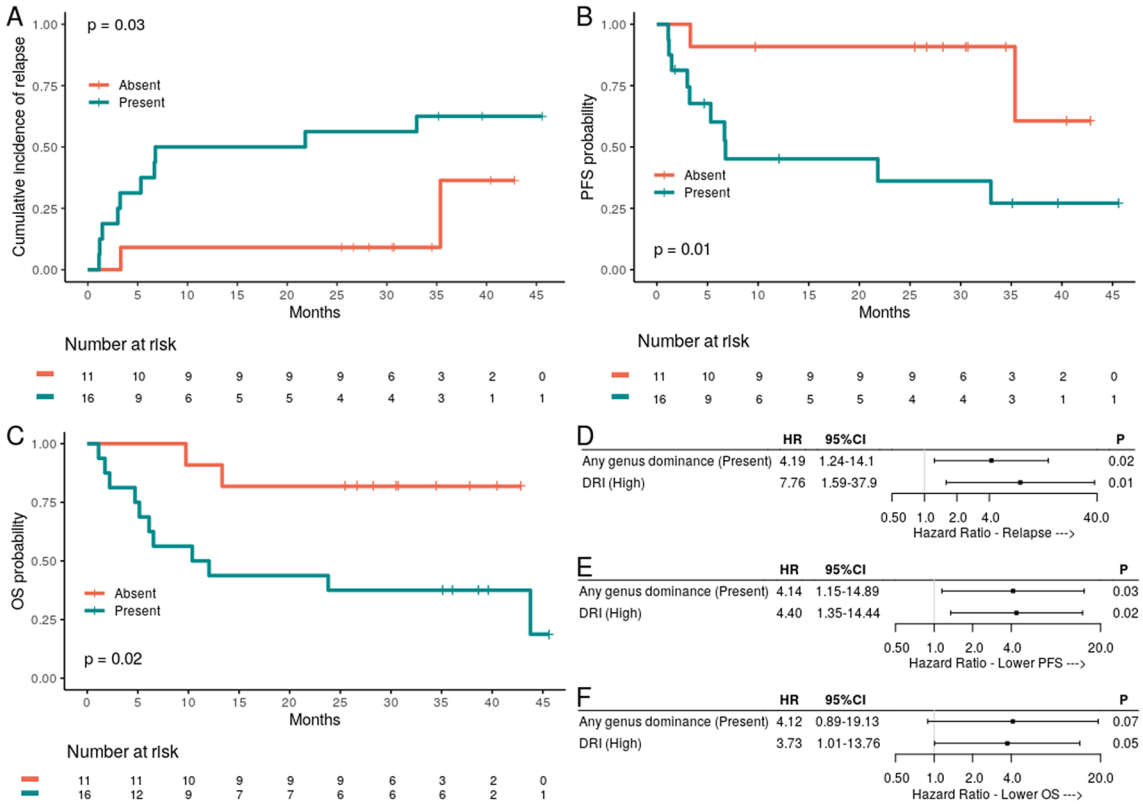


Figure 3. Association of any genus dominance with relapse, progression-free survival, and overall survival. (**A**) Cumulative incidence of relapse with patients (n = 27) stratified by any genus dominance at preconditioning. (**B**) Progression-free survival (PFS) with patients (n = 27) stratified by any genus dominance at preconditioning. (**C**) Overall survival (OS) with patients (n = 27) stratified by any genus dominance at preconditioning. (**D**) The DRI-adjusted hazard ratio for the association of dominance (relative abundance > 30%) of any genus at preconditioning and relapse (n = 27). (**E**) The DRI-adjusted hazard ratio for the association of dominance (relative abundance > 30%) of any genus at preconditioning and PFS (n = 27). (**F**) The DRI-adjusted hazard ratio for the association of dominance (relative abundance > 30%) of any genus at preconditioning and OS (n = 27).

after adjusting for DRI (HR 0.20, 95% CI 0.06–0.67, $P = 0.01$; Fig. 4C). However, after adjusting for multiple comparisons using the Bonferroni correction, because of the limited statistical power of this study, the univariate association between *Solobacterium* presence and lower relapse risk lost significance ($P = 0.72$). The relative abundance of *Solobacterium* at preconditioning per patient is depicted in Fig. S7. No differences in the presence of *Solobacterium* were found in other outcomes (aGVHD at 100 days: 64% versus 44%, respectively [HR 1.84, 95% CI 0.68–4.95, $P = 0.23$]; cGVHD: 27% versus 13%, respectively [HR 2.41, 95% CI 0.43–13.4, $P = 0.31$]; PFS: 55% versus 37%, respectively [HR 0.83, 95% CI 0.31–0.83, $P = 0.71$]; and OS at 3 years: 55% versus 28%, respectively [HR 0.99, 95% CI 0.32–3.08, $P = 0.99$]).

Antibiotic use. From one week before until the first day of the conditioning regimen, 4 (13%) patients received antibiotics. From the first day of the conditioning regimen until engraftment, 28 (93%) patients received antibiotics: 20 (67%) used cefepime, 16 (53%) meropenem, 14 (47%) vancomycin, and four (13%) piperacillin-tazobactam. The use of these antibiotics were not associated with the risk of relapse (Fig. S5). We could not analyze the association between the use of antibiotics before transplant (30 days before starting the conditioning regimen) and OM bacterial diversity because of the small number of patients who used antibiotics at that time point.

Discussion

In this single-center observational study, we prospectively collected samples from the oral mucosa of patients who underwent allo-HSCT. To the best of our knowledge, this is the first study to evaluate the possible impact of the OM using ASVs on allo-HSCT outcomes. ASVs, which are read sequences denoised to single-nucleotide resolution, is a more reproducible and comprehensive technique with higher sensitivity and specificity than operational taxonomic units (OTU) in analyzing microbiota[44–46]. The OTU can identify bacteria at the genus level, while ASVs allow to distinguish bacteria at the species level, which could explain discrepancies between our findings when compared to previous studies.

We observed that patients who presented low OM diversity or dominance of any genus before conditioning had a significantly increased risk of relapse. The dominance of any genus was also associated with worse PFS

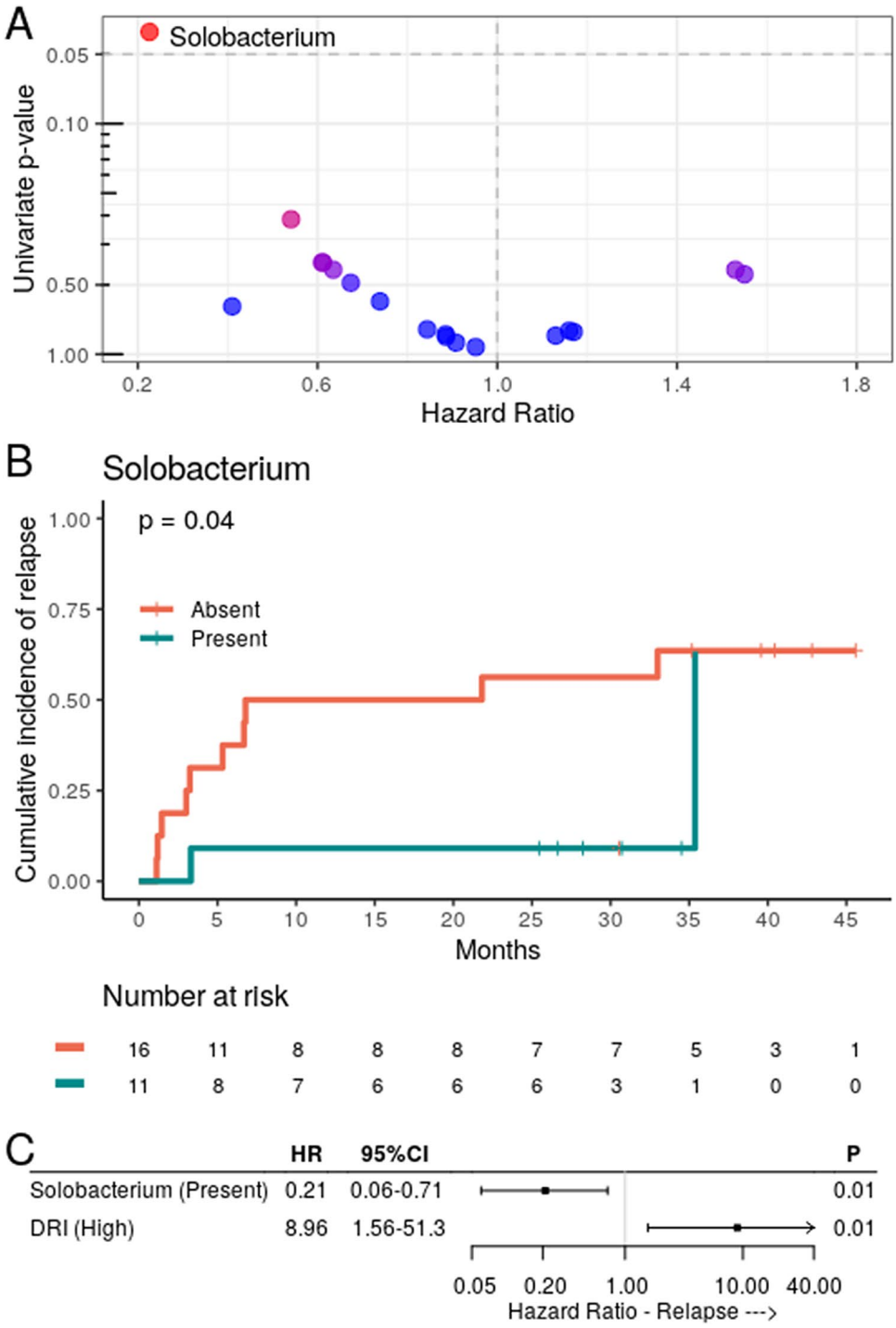


Figure 4. *Solobacterium* absence at preconditioning was associated with an increased risk of relapse in patients who underwent allo-HSCT. (**A**) Volcano plot for the univariate competing risk analysis of the association of relapse with the presence of specific genera at preconditioning (*P* value versus hazard ratio). The *Solobacterium* data point is indicated as it was the only genus significantly associated with relapse ($P < 0.05$). (**B**) Cumulative incidence of relapse with patients (n = 27) stratified by *Solobacterium* presence. (**C**) The DRI-adjusted hazard ratio for the association of *Solobacterium* presence at preconditioning and relapse (n = 27).

and OS. Although a low oral microbiota diversity and the dominance of any genus are proxies for microbiota dysbiosis, the former was not associated with worse PFS and OS. Only 7 (25%) patients share the binomial: low OM diversity and dominance of any genus, or high OM diversity and absence of dominance. The dominance

of a single genus may denote a deeper immune imbalance and could represent a more sensitive predictor of alloHSCT outcomes when compared with OM diversity.

The OM has different niches in the same environment and is as diverse as the IM. Previous studies evaluating OM and allo-HSCT have shown conflicting results, likely because of the use of low-resolution techniques for microbiota analysis and the small sample sizes. In one case series, there were no changes in OM after allo-HSCT. The most common oral organisms, e.g., *Streptococcus*, *Gemella*, and *Veillonella*, remained relatively stable after transplant[34]. However, another study showed a reduction in alpha diversity after allo-HSCT when compared with the pretransplant OM[35], and this reduction was more pronounced in patients who developed oral mucositis[47]. Besides, we did not find any direct correlation between the use of antibiotics after conditioning and transplant outcomes, as other studies have shown for IM diversity[23,24].

Recently, IM has attracted attention as a potential predictive marker for allo-HSCT outcomes. Previous studies have shown that low IM diversity is associated with a higher risk of mortality, but not with the risk of relapse[25–27], diverging from our findings.

Higher risk of aGVHD in patients with low IM diversity[48] and a higher risk of transplant-related mortality attributable to GVHD[26] were also reported. In the oral mucosa samples analyzed in the current study, low OM diversity was associated with an increased risk of relapse but did not change the risk of mortality, aGVHD or cGVHD.

The dominance of a specific bacterial group in IM, *Eubacterium limosum*, has also been shown to be related to relapse and disease progression. In our series, the dominance of any genus was associated with a higher risk of relapse.

As opposed to what has been observed for IM[27], all dominant genera at preconditioning reported herein are commensal organisms. Thus, it is unlikely that they all have detrimental roles in the allo-HSCT setting, being more plausible that the presence of dominance by any genus is a proxy for low diversity/dysbiotic OM.

Furthermore, the presence of *Solobacterium* in the OM before conditioning seems to have a protective effect against relapse. *S. moorei*, the only species in the *Solobacterium* genus, is normally associated with halitosis[49,50] and endodontic infection[51,52]. However, in the allo-HSCT scenario, the lack of *Solobacterium* could be a marker of dysbiosis, pretransplant disease status, or previous treatments. Alternatively, this genus may also play a role as an immune mediator by producing hydrogen sulfide[49], a metabolite associated with decreased oxidative stress and increased sensitivity to antibiotics[53]. Although, the low overall *Solobacterium* relative abundance even in patients where it was present makes the latter alternative more unlikely, this finding need to be validated in future studies.

A previous study analyzed the tongue microbiota in patients who underwent alloHSCT and compared it with community-dwelling adults. AlloHSCT patients have a lower tongue microbiota alpha diversity when compared to community adults. Moreover, the presence of *Staphylococcus haemolyticus* or *Ralstonia pickettii* was associated with a higher risk of mortality. Nevertheless, no relationship was observed between alpha diversity of the tongue microbiota and incidence of transplant complications[46]. A study of salivary microbiota showed a reduction in alpha diversity during the course of transplantation. Again, no correlation between salivary microbiota diversity and alloHSCT outcomes was found[54]. The discrepancies between these studies and our findings may be related to different sites of sample collections, and different distinct microbiome analysis techniques.

Our study had several limitations of a relatively small and heterogenous single-center transplant cohort. However, as observed in studies of IM, in our series, OM showed a significant correlation with relapse and may also provide valuable information on host-related microbial dysbiosis, providing a simple, reproducible technique for collection and analysis prior to transplantation.

In conclusion, in the current study, we focused on preconditioning samples in order to identify potential clinical effects of OM on allo-HSCT outcomes and observer that lower OM diversity was associated with a higher risk of relapse after allo-HSCT and dominance by a single genus was associated with a higher risk of relapse and worse survival after allo-HSCT.

Prospective trials and validation cohorts are needed to confirm these findings and to test whether early interventions to correct OM dysbiosis or more aggressive strategies to prevent relapse in OM dysbiotic patients, such as early immunosuppression withdrawal, maintenance therapy, or prophylactic donor lymphocyte infusions, could improve the predicted adverse outcome.

Received: 10 May 2021; Accepted: 11 August 2021
Published online: 02 September 2021

References

1. Gyurkocza, B. & Sandmaier, B. M. Conditioning regimens for hematopoietic cell transplantation: One size does not fit all. *Blood* **124**, 344–353 (2014).
2. Styczyński, J. *et al.* Death after hematopoietic stem cell transplantation: Changes over calendar year time, infections and associated factors. *Bone Marrow Transplant.* **55**, 126–136 (2020).
3. Ossenkoppele, G. J., Janssen, J. J. W. M. & van de Loosdrecht, A. A. Risk factors for relapse after allogeneic transplantation in acute myeloid leukemia. *Haematologica* **101**, 20–25 (2016).
4. Martino, R. *et al.* Comparison of conditioning regimens of various intensities for allogeneic hematopoietic SCT using HLA-identical sibling donors in AML and MDS with <10% BM blasts: A report from EBMT. *Bone Marrow Transplant.* **48**, 761–770 (2013).
5. Bornhäuser, M. *et al.* Reduced-intensity conditioning versus standard conditioning before allogeneic haemopoietic cell transplantation in patients with acute myeloid leukaemia in first complete remission: A prospective, open-label randomised phase 3 trial. *Lancet Oncol.* **13**, 1035–1044 (2012).
6. Shen, Z. *et al.* Influence of pre-transplant minimal residual disease on prognosis after Allo-SCT for patients with acute lymphoblastic leukemia: Systematic review and meta-analysis. *BMC Cancer* **18**, 1–12 (2018).

7. Rezvani, A. R. *et al.* Impact of donor age on outcome after allogeneic hematopoietic cell transplantation. *Biol. Blood Marrow Transplant.* **21**, 105–112 (2015).
8. Bastida, J. M. *et al.* Influence of donor age in allogeneic stem cell transplant outcome in acute myeloid leukemia and myelodisplastic syndrome. *Leuk. Res.* **39**, 828–834 (2015).
9. Milano, F. *et al.* Cord-blood transplantation in patients with minimal residual disease. *N. Engl. J. Med.* **375**, 944–953 (2016).
10. Willemze, R. *et al.* KIR-ligand incompatibility in the graft-versus-host direction improves outcomes after umbilical cord blood transplantation for acute leukemia. *Leukemia* **23**, 492–500 (2009).
11. Soiffer, R. J. *et al.* Prospective, randomized, double-blind, phase III clinical trial of anti-T-lymphocyte globulin to assess impact on chronic graft-versus-host disease-free survival in patients undergoing HLA-matched unrelated myeloablative hematopoietic cell transplantation. *J. Clin. Oncol.* **35**, 4003–4011 (2017).
12. Walker, I. *et al.* Addition of anti-thymocyte globulin to standard graft-versus-host disease prophylaxis versus standard treatment alone in patients with haematological malignancies undergoing transplantation from unrelated donors: Final analysis of a randomised, open-label. *Lancet Haematol.* **7**, e100–e111 (2020).
13. Nykolyszyn, C. *et al.* Posttransplantation cyclophosphamide vs antithymocyte globulin as GVHD prophylaxis for mismatched unrelated hematopoietic stem cell transplantation. *Bone Marrow Transplant.* **55**, 349–355 (2020).
14. Signori, A. *et al.* Chronic GVHD is associated with inferior relapse risk irrespective of stem cell source among patients receiving transplantation from unrelated donors. *Bone Marrow Transplant.* **47**, 1474–1478 (2012).
15. Tanaka, Y. *et al.* Analysis of non-relapse mortality and causes of death over 15 years following allogeneic hematopoietic stem cell transplantation. *Bone Marrow Transplant.* **51**, 553–559 (2016).
16. Armand, P. *et al.* Validation and refinement of the disease risk index for allogeneic stem cell transplantation. *Blood* **123**, 3664–3671 (2014).
17. Belkaid, Y. & Hand, T. W. Role of the microbiota in immunity and inflammation. *Cell* **157**, 121–141 (2014).
18. Zitvogel, L. *et al.* Cancer and the gut microbiota: An unexpected link. *Sci. Transl. Med.* **7**, 271 (2015).
19. Vetizou, M. *et al.* Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* **350**, 1079–1084 (2015).
20. Holler, E. *et al.* Metagenomic analysis of the stool microbiome in patients receiving allogeneic stem cell transplantation: Loss of diversity is associated with use of systemic antibiotics and more pronounced in gastrointestinal graft-versus-host disease. *Biol. Blood Marrow Transplant.* **20**, 640–645 (2014).
21. Mathewson, N. D. *et al.* Gut microbiome-derived metabolites modulate intestinal epithelial cell damage and mitigate graft-versus-host disease. *Nat. Immunol.* **17**, 505–513 (2016).
22. Eriguchi, Y. *et al.* Graft-versus-host disease disrupts intestinal microbial ecology by inhibiting Paneth cell production of α-defensins. *Blood* **120**, 223–231 (2012).
23. Jenq, R. R. *et al.* Intestinal blautia is associated with reduced death from graft-versus-host disease. *Biol. Blood Marrow Transplant.* **21**, 1373–1383 (2015).
24. Shono, Y. *et al.* Increased GVHD-related mortality with broad-spectrum antibiotic use after allogeneic hematopoietic stem cell transplantation in human patients and mice. *Sci. Transl. Med.* **8**, 339 (2016).
25. Taur, Y. *et al.* The effects of intestinal tract bacterial diversity on mortality following allogeneic hematopoietic stem cell transplantation. *Blood* **124**, 1174–1182 (2014).
26. Peled, J. U. *et al.* Microbiota as predictor of mortality in allogeneic hematopoietic-cell transplantation. *N. Engl. J. Med.* **382**, 822–834 (2020).
27. Peled, J. U. *et al.* Intestinal microbiota and relapse after hematopoietic-cell transplantation. *J. Clin. Oncol.* **35**, 1650–1659 (2017).
28. Shono, Y. & Van Den Brink, M. R. M. Gut microbiota injury in allogeneic haematopoietic stem cell transplantation. *Nat. Rev. Cancer* **18**, 283–295 (2018).
29. Idris, A., Hasnain, S. Z., Huat, L. Z. & Koh, D. Human diseases, immunity and the oral microbiota—Insights gained from metagenomic studies. *Oral Sci. Int.* **14**, 27–32 (2017).
30. Mascitti, M. *et al.* Beyond head and neck cancer: The relationship between oral microbiota and tumour development in distant organs. *Front. Cell Infect. Microbiol.* https://doi.org/10.3389/fcimb.2019.00232 (2019).
31. Zhao, H. *et al.* Variations in oral microbiota associated with oral cancer. *Sci. Rep.* **7**, 11773 (2017).
32. Zhang, Y., Niu, Q., Fan, W., Huang, F. & He, H. Oral microbiota and gastrointestinal cancer. *Onco Targets Ther.* **12**, 4721–4728 (2019).
33. Wang, Y. *et al.* Oral microbiota distinguishes acute lymphoblastic leukemia pediatric hosts from healthy populations. *PLoS ONE* **9**, e102116 (2014).
34. Ames, N. J. *et al.* A characterization of the oral microbiome in allogeneic stem cell transplant patients. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0047628 (2012).
35. Ames, N. J. *et al.* The oral microbiome of patients undergoing treatment for severe aplastic anemia: A pilot study. *Ann. Hematol.* **98**, 1351–1365 (2019).
36. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1 (2013).
37. Wang, H. *et al.* Breast tissue, oral and urinary microbiomes in breast cancer. *Oncotarget* **8**, 88122. https://doi.org/10.18632/oncotarget.21490 (2017).
38. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
39. Edgar, R. C. & Flyvbjerg, H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* **31**, 3476–3482 (2015).
40. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
41. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
42. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200 (2011).
43. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: A versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
44. Edgar, R. C. UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–998 (2013).
45. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**, 2639–2643 (2017).
46. Oku, S. *et al.* Disrupted tongue microbiota and detection of nonindigenous bacteria on the day of allogeneic hematopoietic stem cell transplantation. *PLoS Pathog.* **16**, e1008348 (2020).
47. Shouval, R. *et al.* Salivary microbial and metabolic determinants of oral mucositis in recipients of allogeneic hematopoietic stem cell transplantation. *Biol. Blood Marrow Transplant.* **26**, S50 (2020).
48. Golob, J. L. *et al.* Stool microbiota at neutrophil recovery is predictive for severe acute graft vs host disease after hematopoietic cell transplantation. *Clin. Infect. Dis.* **65**, 1984–1991 (2017).
49. Haraszthy, V. I. *et al.* Characterization and prevalence of *Solobacterium moorei* associated with oral halitosis. *J. Breath Res.* **2**, 017002 (2008).

50. Vancauwenberghe, F. *et al.* The role of *Solobacterium moorei* in oral malodour. *J. Breath Res.* **7**, 046006 (2013).
51. Rolph, H. J. *et al.* Molecular identification of microorganisms from endodontic infections. *J. Clin. Microbiol.* **39**, 3282–3289 (2001).
52. Rôças, I. N., Hülsmann, M. & Siqueira, J. F. Microorganisms in root canal-treated teeth from a German population. *J. Endod.* **34**, 926–931 (2008).
53. Yang, R., Yu, T., Liu, D., Shi, S. & Zhou, Y. Hydrogen sulfide promotes immunomodulation of gingiva-derived mesenchymal stem cells via the Fas/FasL coupling pathway. *Stem Cell Res. Ther.* **9**, 62 (2018).
54. Shouval, R. *et al.* Patterns of salivary microbiota injury and oral mucositis in recipients of allogeneic hematopoietic stem cell transplantation. *Blood Adv.* **4**, 2912–2917 (2020).

Acknowledgements

We are grateful to all patients from Hospital Sírio Libanês, São Paulo, Brazil. We thank our colleagues from Hospital Sírio Libanês.

Author contributions

Conception and design: A.A.C., E.R.F. and C.A.-R. Collection and assembly of data: V.H., J.S.B., F.H.K., W.M.-S., P.F.A., V.C.d.M. and C.A.-R. Data analysis and interpretation: V.C.d.M., V.H., J.S.B., A.A.C., E.R.F. and C.A.-R. Manuscript writing: All authors. Final approval of manuscript: All authors. Accountable for all aspects of the work: All authors.

Funding

V.C.M. was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq process no. 141575/2018-2). J.S.B. was supported by Coordenação de Aperfeiçoamento de pessoal de Nível Superior (CAPES process no. 001). V.H. was supported by Fundação de Amapro à Pesquisa do Estado de São Paulo (FAPESP process no. 13996-0/2018).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-96939-8.

Correspondence and requests for materials should be addressed to C.A.-R.

Reprints and permissions information is available at www.nature.com/reprints.

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

5. CHAPTER 4: Longitudinal analysis at three oral sites links oral microbiota to clinical outcomes in allogeneic hematopoietic stem-cell transplant

Longitudinal analysis at three oral sites links oral microbiota to clinical outcomes in allogeneic hematopoietic stem-cell transplant

Vitor Heidrich[1,2], Franciele H. Knebel[1], Julia S. Bruno[1], Vinícius C. de Molla[3,4], Wanessa Miranda-Silva[1], Paula F. Asprino[1], Luciana Tucunduva[5], Vanderson Rocha[6], Yana Novis[5], Eduardo R. Fregnani[1], Celso Arrais-Rodrigues[3,4], Anamaria A. Camargo[1,*]

[1]Centro de Oncologia Molecular, Hospital Sírio-Libanês, São Paulo, SP, Brazil

[2]Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, SP, Brazil

[3]Hospital Nove de Julho, Rede DASA, São Paulo, SP, Brazil

[4]Universidade Federal de São Paulo, São Paulo, SP, Brazil

[5]Centro de Oncologia, Hospital Sírio-Libanês, São Paulo, SP, Brazil

[6]Hospital das Clínicas da Faculdade de Medicina, Universidade de São Paulo/Instituto do Câncer do Estado de São Paulo (ICESP), São Paulo, SP, Brazil

*Corresponding author

Authors' email addresses

VH: vheidrich@mochsl.org.br

FHK: fhknebel@mochsl.org.br

JSB: juliasb9@gmail.com

VCM: viniciuscamposdemolla@gmail.com

WMS: wmswanessa@yahoo.com.br

PFA: pasprino@mochsl.org.br

LT: luciana.tucunduva@gmail.com

VR: rocha.vanderson@hotmail.fr

YN: yananovis@yahoo.com

EFR: eduardofregnani@me.com

CAR: celsoarrais@gmail.com

AAC: anamaria.acamargo@hsl.org.br

Abstract

Background

Allogeneic hematopoietic stem-cell transplant (allo-HSCT) is a potentially curative therapy for several hematological disorders. Before stem-cell infusion, recipients undergo a conditioning regimen with chemo/radiotherapy and immunosuppressants, requiring the use of antibiotics to treat and prevent infections. This regimen promotes drastic alterations in the recipient's microbiotas, including the oral microbiota, which have been associated with allo-HSCT complications and poor outcomes. However, long-term longitudinal studies on the oral microbiota of allo-HSCT recipients are scarce and disregard the existence of distinct microbiotas within the oral cavity. Here, we used 16S rRNA gene sequencing to characterize the microbiota dynamics (during and after allo-HSCT) of 31 allo-HSCT recipients at 3 oral sites (gingival crevicular fluid, oral mucosa, and supragingival biofilm).

Results

Analysis of the oral microbiota dynamics during allo-HSCT revealed a significant decline in bacterial diversity and major shifts in microbiota composition in all oral sites, including blooms of potentially pathogenic genera. These blooms in some cases preceded respiratory infections caused by the blooming genera. We also noticed that differences in microbiota diversity and composition between oral sites were lost during allo-HSCT. Overall, oral microbiotas returned to their preconditioning state after engraftment. However, the ability to recover the initial bacterial composition varied between patients. After stratifying patients based on their ability to recover their preconditioning microbiota composition, we found that recovery of the oral mucosa microbiota composition was not associated with antibiotic usage but was associated with higher preconditioning diversity and earlier reconstitution of normal leukocyte counts. Most notably, oral mucosa microbiota composition recovery was an independent biomarker of better allo-HSCT outcomes.

Conclusion

We observed clear patterns of microbiota dysbiosis in all three oral sites during allo-HSCT, however each oral site responded differently to the perturbations associated with allo-HSCT. Oral microbiota injury and recovery patterns were associated with allo-HSCT complications and outcomes. This study highlights the potential clinical impact of the oral microbiota in the allo-HSCT setting and the clinical value of tracking oral microbiota changes during allo-HSCT.

Keywords

Oral microbiome; 16S rRNA gene sequencing; allogeneic hematopoietic cell transplant; microbiome stability; blooming of bacteria; biomarkers; clinical outcomes.

Introduction

Countless microbes from food, air, and our physical/biological environment arrive in our mouths daily. However, only a small subset of these microbes can colonize the oral cavity to compose the oral microbiota [1]. This constant contact with non-resident microbes and frequent exposure to other insults (e.g., toothbrushing) made the human oral microbiota remarkably stable and resilient to external perturbations [2].

Residing oral microbes organize in biofilms, creating oxygen gradients that allow colonization by both anaerobic and aerobic bacteria [1]. Differences in moisture, topography, and tissue type (shedding vs. non-shedding), among others, make each oral site home to distinct bacterial communities [1, 3] with main compositional differences existing between mucosa-associated and teeth-associated microbiotas [4].

These distinct oral microbiotas are important regulators of human health, as they have been associated with different local and systemic disorders [5]. While the supragingival biofilm is causally linked to the pathogenesis of dental caries [6], bacteria at the gingival crevice, an oxygen-limited environment bathed in immune exudate (gingival crevicular fluid), are linked to periodontitis [7] and may cause bacteremia by translocation to the circulation

across the thin gingival crevice epithelium [8]. Oral bacteria can further facilitate systemic reach by producing molecules that increase vascular permeability [5]. Using this strategy, oral *Porphyromonas gingivalis* is able to colonize the brain, contributing to the pathogenesis of Alzheimer's disease [9].

Allogeneic hematopoietic stem-cell transplant (allo-HSCT) is used to treat malignant (e.g., acute myeloid leukemia) and non-malignant (e.g., aplastic anemia) hematological disorders [10]. The goal of allo-HSCT is to eradicate malignant/defective cells and to replace an abnormal hematopoietic and immune system [11]. Allo-HSCT recipients undergo a conditioning regimen with chemo/radiotherapy that reduces disease burden and provides sufficient immunoablation to allow donor stem-cell engraftment [12]. After engraftment, the graft-vs-tumor/autoimmunity effect further promotes disease eradication and the hematopoietic/immune function gradually reconstitutes [13]. Besides chemo/radiotherapy, allo-HSCT recipients are treated with immunosuppressants to prevent engraftment failure and graft-vs-host disease, and antibiotics to prevent and treat opportunistic infections during immunosuppression [13, 14].

Allo-HSCT is considered one of the most severe perturbations the immune system undergoes in the therapeutic setting [15]. Since the immune system regulates microbiota composition [16] and chemotherapy [17], radiotherapy [18], and antibiotics [19] have detrimental effects on the microbiota, drastic alterations in the gut microbiota have been reported in allo-HSCT recipients, including loss of bacterial diversity and blooms of potentially pathogenic species [20]. Recent evidence shows these alterations extend to other microbiotas [21], including the relatively more stable oral microbiota [22–26]. More importantly, the pre-transplant microbiota and the extent of microbiota damage during allo-HSCT are associated with allo-HSCT complications and outcomes, so that gut and oral microbiota provide biomarkers in the allo-HSCT setting [24, 25, 27–30].

The stability of the oral microbiota [5] and its associations with allo-HSCT outcomes offer a unique opportunity to identify predictive biomarkers and develop therapeutic interventions to promote oral health in allo-HSCT recipients, potentially improving allo-HSCT

4

safety and efficacy. However, so far, oral microbiota studies in allo-HSCT recipients evaluated single oral sites, not leveraging the ease of sampling of different oral compartments [22–26, 30]. In addition, although a causal link between post-transplant gut microbiota recovery and improved clinical responses to allo-HSCT has been suggested [15], oral microbiota recovery trajectories after allo-HSCT were not thoroughly characterized and their association with allo-HSCT outcomes remain unknown.

To obtain a more in-depth understanding of oral microbiota dynamics during and after allo-HSCT and to test whether oral microbiota recovery is associated with allo-HSCT outcomes, we profiled the oral microbiota of a Brazilian cohort of allo-HSCT recipients. We collected over 440 samples encompassing five timepoints and three oral sites: gingival crevicular fluid (GCF), oral mucosa (OM), and supragingival biofilm (SB), which allowed a longitudinal anatomically-aware analysis of the oral microbiota. We used 16S rRNA gene sequencing to characterize diversity, compositional, and taxonomical changes in oral microbiota during allo-HSCT and after engraftment. We associated these changes with antibiotic usage and allo-HSCT complications. Finally, we evaluated recovery trajectories after allo-HSCT to associate oral microbiota recovery with allo-HSCT outcomes.

Materials and methods

Patients' clinical characteristics

Thirty-one patients undergoing allo-HSCT at Hospital Sírio-Libanês (São Paulo, Brazil) were recruited between January 2016 and April 2018. The median age was 50 years, most patients were male (55%), and acute leukemia was the most common underlying disease (58%; 11 acute myeloid leukemia and 7 acute lymphocytic leukemia cases). Most patients underwent reduced intensity conditioning (61%) and received grafts from peripheral blood (68%). Patient clinical information is summarized in Table S1.

5

Antibiotic usage analysis

Antibiotic prescriptions were retrieved retrospectively from clinical records. Information spanning 30 days before preconditioning sampling and 100 days after stem-cell infusion was collected to build timelines of antibiotic usage for each patient (Additional file 1: Timelines of antibiotic usage). A ridgeline plot of antibiotic usage detailing all antibiotics and antibiotic classes used showed antibiotics prescription concentrates in the few weeks immediately after infusion (Fig. S1), with only 5/31 patients receiving antibiotics before preconditioning (Additional file 1). Due to the sparse use of antibiotics before preconditioning and the unlikely effect of antibiotics received months after allo-HSCT on clinical outcomes, antibiotic usage was analyzed considering only the time window between preconditioning and 30 days after engraftment (a patient deceased during this period was excluded from the analysis). For each patient, the length in days under antibiotic therapy (length of therapy, LOT) and the number of agent days under antibiotic therapy (days of therapy, DOT) was calculated, as defined previously [31]. To evaluate the impact of specific antibiotic classes on microbiota dynamics, patients were further classified according to antibiotic class usage during the period of interest. Only antibiotic classes received by at least 20% of our patients (6/30) were considered in this analysis. In addition to individual antibiotics prescriptions, all patients underwent standard antimicrobial prophylaxis with antibiotic, antiviral and antifungal drugs. Because the standard antibiotic prophylaxis protocol in our institution comprises oral levofloxacin and sulfamethoxazole-trimethoprim, their use was not considered in the antibiotic usage analysis.

Sample collection

Patients were examined frequently by an oral medicine specialist throughout the hospitalization period. The standard oral hygiene protocol comprised toothbrushing with fluoridated toothpaste and 0.12% chlorhexidine mouthwash. Samples were collected at least six hours after the last oral hygiene procedure by an oral medicine specialist at three oral

sites. GCF samples were collected by inserting absorbent paper points in the gingival crevice; OM samples were collected by swabbing bilateral buccal mucosa, alveolar mucosa of the jaws, and tongue dorsum; SB samples were collected by swabbing all vestibular enamel surface. Samples were dry-stored at -20ºC.

DNA extraction and 16S rRNA gene amplicon sequencing

Samples were brought to room temperature. Bacterial cells were recovered from swabs or paper points by vortexing in 600 µl or 800 µl TE buffer (10mM Tris; 1mM EDTA; pH 8,0), respectively. Samples were transferred to a new tube, supplemented with 6 µl (OM and SB) or 8 µl (GCF) PureLink™ RNAse A (20 mg/ml; Invitrogen), and DNA was extracted using the QIAamp DNA Mini Blood kit (Qiagen) following the manufacturer's protocol (*Buccal Swab Spin Protocol*).

Bacterial communities were profiled by 16S rRNA gene amplicon-sequencing as described in detail previously [32]. In short, amplicon libraries were prepared with 12.5 ng of total DNA and pre-validated V3V4 primers [33] following Illumina's protocol (*Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System*). Amplicons were sequenced on the Illumina MiSeq platform using the MiSeq Reagent Kit v3 (600-cycle) (Illumina).

Bioinformatics pipeline

Reads were demultiplexed using the MiSeq Reporter Software. Primers were removed and low-quality 3' ends were trimmed using *seqtk* [34]. Next, reads were processed using QIIME 2 (v2019.10.0) as schematized in Fig. S2a [35]. In detail, amplicon sequence variants (ASVs) were generated using the DADA2 pipeline (via *q2-dada2*), which includes removal of low-quality reads, denoising, merging, and removal of bimeras [36]. Chimeric ASVs were further filtered out using a reference-based approach with VSEARCH [37] (via *q2-vsearch*) and SILVA database (v132) [38]. Taxonomic assignment of ASVs was also performed with VSEARCH [37] (via *q2-feature-classifier*) and SILVA (v132) [38]. Finally,

7

non-bacterial ASVs were removed (via *q2-feature-table*). QIIME 2 outputs were transferred to the R environment [39] using the *qiime2R* R package [40] and analyzed for microbiota profiling with custom R scripts as detailed below.

Microbiota and statistical analyses

The total number of reads of the sample with the lowest number of reads (3,578 reads) among the samples included in the microbiota profiling analyses was used as C_{min} for Scaling with Ranked Subsampling (SRS) normalization prior to diversity analyses [41]. Diversity was measured by the Gini-Simpson index [42] using the *phyloseq* R package [43]. Longitudinal diversity variations were evaluated by calculating diversity resistance, resilience, and stability [44, 45] (see Additional file 3: Supplementary methods). Compositional dissimilarity between samples was measured by the weighted UniFrac distance [46] using the *rbiom* R package [47]. Longitudinal compositional variations were evaluated by calculating compositional stability (see Additional file 3). Multiple linear regression was used to evaluate whether antibiotic usage was associated with diversity stability and compositional stability (see Additional file 3). Recovery to baseline composition was defined as distance between samples collected at preconditioning and 30 days after engraftment <0.5.

Taxonomic nomenclature was homogenized prior to all taxonomic analyses (see Additional file 3). Taxa relative abundance plots included only the most relevant genera according to criteria specified in figure legends. Differential abundance analysis was performed using ANCOM-BC [48] with genera present in ≥25% of the samples being compared. Genera abundance differences between groups at q-value < 0.05 (Bonferroni correction) were considered statistically significant, including ANCOM-BC structural zeroes.

Associations between oral microbiota composition recovery or clinical parameters with allo-HSCT outcomes were determined using univariate Cox proportional-hazards regression [49] or univariate Fine-Gray competing risk regression [50]. Cox models were used to evaluate overall survival and progression-free survival, while Fine-Gray models were

8

used to evaluate the risk of transplant-related death (with relapse mortality as competing risk) and the risk of underlying disease relapse (with transplant-related mortality as competing risk). Multivariate analysis was used to evaluate oral microbiota composition recovery and correct for clinical parameters significantly associated with the outcome (P-value < 0.05) in the univariate analysis. Patients experiencing the event before oral microbiota composition recovery evaluation were excluded from univariate and multivariate analyses.

Results

Samples collected and sequencing output

We collected samples from three oral sites (GCF, OM, and SB) at five timepoints during allo-HSCT: preconditioning (P), aplasia (A), engraftment (E), 30 days after engraftment (E30), and 75 days after engraftment (E75). Since most patients were discharged shortly after engraftment, the exact date of sample collection varied for E30 (20–45 days after engraftment) and E75 (60–131 days after engraftment) samples, as indicated in Fig. S3. Premature death after allo-HSCT hampered the collection of the E30 sample for patient #3 and E75 samples for patients #1, #2, #3, #21, and #31 (Fig. S3). In addition, the E75 sample from patient #9 was excluded due to low DNA yield. Overall, 444 samples were successfully processed and sequenced for microbiota profiling.

We generated a total of 53,253,725 V3V4 16S rRNA reads (median per sample: 104,230.5; range: 2,059–502,409; Fig. S2b). After filtering, 31,343,619 reads (59%; Fig. S2c–d) were retained (median per sample: 63,075.5; range: 87–310,082; Fig. S2e), corresponding to 4,046 ASVs. Using SRS curves [51] (Fig. S4), we established a minimum sequencing depth cutoff of 3,000 reads and 4 low-depth samples were excluded from further analysis (patient #1, OM, P; #5, OM, E; #6, OM, E; #25, SB, E). We proceeded to profile the oral microbiota during allo-HSCT with 440 samples.

9

Compositional differences between oral microbiotas during allo-HSCT and after engraftment

We first assessed microbiota compositional differences between oral sites at each allo-HSCT timepoint. Visually, all three oral microbiotas occupied a similar compositional space throughout allo-HSCT (Fig. 1a). Nevertheless, similarly to what is observed in healthy adults [4], each oral site contained a significantly different microbiota composition at P (PERMANOVA, GCF vs. OM: P-value = 0.001; GCF vs. SB: P-value = 0.002; OM vs. SB: P-value = 0.018). Noteworthy, these differences progressively diminished in subsequent timepoints until E30 and were partially recovered at E75 (Fig. 1b). Calculation of the minimum compositional distance between oral sites for each patient confirmed lower compositional distance between sites after P (Fig. 1c).

Differential abundance analysis at genus level using ANCOM-BC revealed a similar picture (Fig. 1d). As expected, all three oral microbiotas showed many distinguishing genera at P. For example, we observed a higher abundance of *Actinomyces* in the SB as compared to GCF and a higher abundance of *Solobacterium* in the OM as compared to SB (Fig. S5). *Actinomyces spp.* are early colonizers of the SB with a crucial role in ecological succession during SB maturation [52]. On the other hand, *Solobacterium moorei*, the only known species in the *Solobacterium* genus, is a halitosis-associated bacteria typically found in the tongue dorsum [53], a site contemplated in OM samples. However, a smaller number of differentially abundant genera was observed in subsequent timepoints, with a slight increase in the number of differentially abundant genera between sites at E75, illustrated by the reappearance of *Solobacterium* as an OM-associated genus (Fig. S5).

In short, our data indicate that compositional differences between oral microbiotas are reduced during allo-HSCT, being only partially recovered several weeks after engraftment.

10

Oral microbiota dynamics during allo-HSCT and after engraftment

We next characterized microbiota diversity dynamics at each oral site during allo-HSCT and after engraftment. As previously shown for OM [25] and SB [24], GCF presented a stepped decline in diversity up to E (Fig. 2a). By extending this analysis to the post-engraftment period for all oral sites, we observed a gradual recovery of diversity, with baseline levels almost fully reestablished around E75.

We then applied key concepts from ecology [45] for a more in-depth characterization of diversity dynamics during allo-HSCT. By considering allo-HSCT as a perturbation relieved immediately after engraftment, we calculated for each patient diversity resistance (inversely proportional to the diversity loss up to E), resilience (rate of diversity gain after E), and stability (combined effect of resistance and resilience) to allo-HSCT (Fig. S6a; see Additional File 3). GCF showed higher diversity resistance than OM and SB (Fig. 2b), in line with the less pronounced loss of diversity observed in this oral site at E (Fig. S6b). All oral sites presented equivalent levels of diversity resilience and stability (Fig. 2b), in line with the similar levels of diversity after engraftment observed for all oral sites (Fig. S6b).

Next, we characterized compositional changes in each oral site during allo-HSCT and after engraftment. The compositional distance to P centroid increased up to E and decreased in the post-engraftment period, indicating a displacement from and posterior recovery to baseline compositions (Fig. 1c). However, when comparing the compositional distance from P to all other timepoints using PERMANOVA tests, we observed that GCF and SB post-engraftment samples still showed significantly different compositions after engraftment compared to P, while OM samples more fully recovered their preconditioning state (Fig. 1d). Finally, in analogy to diversity stability, we calculated the compositional stability for each patient (see Additional File 3). As observed for diversity stability, all oral sites showed equivalent levels of compositional stability (Fig. S6c).

11

Our data indicate that allo-HSCT transiently damages oral microbiotas diversity and composition, but each oral site responds differently to the perturbations associated with allo-HSCT.

Oral taxa abundances during allo-HSCT and after engraftment

The loss of differences between microbiotas of distinct oral sites and the displacement from initial compositions observed during allo-HSCT point out to a complex compositional dynamics that likely involves many bacterial taxa and thus can be better appreciated by longitudinal taxonomic composition analysis at each specific site. As expected, all oral sites presented high relative abundance of commensal bacteria at P (Fig. 3a; Fig. S7). For instance, *Veillonella* and *Streptococcus*, genera with high relative abundance in all oral sites of healthy adults [4], occupied either the first or second position in terms of mean relative abundance at P in all three oral sites (Fig. 3b). However, there were several changes in the ranking of the most abundant taxa (on average) across timepoints (Fig. 3b; Fig. S7), pointing to drastic taxonomic composition changes during allo-HSCT. There are some noteworthy examples, such as *Streptococcus* in SB, which went from first in the relative abundance ranking at P to the eleventh position at E. Interestingly, *Streptococcus* recovered its initial ranking position after engraftment (first position at E30 and E75). On the other hand, some non-commensals genera were close to absent in P and only emerged in the subsequent timepoints. For instance, *Enterococcus* and *Lactobacillus*, both potentially pathogenic genera in the oral microbiota [54, 55], showed low mean relative abundance at P but were among the most abundant genera in all sites at E.

Differential abundance analysis at genus level using ANCOM-BC with P as reference for comparisons confirmed these results and showed several additional differentially abundant genera (Fig. 3c). The number of differentially abundant genera at each timepoint was consistent with the compositional displacement and recovery aforementioned, with a maximum of differentially abundant genera at E (Fig. S8). Although there were considerably fewer differentially abundant genera after engraftment, some differences persisted. For

instance, we observed a decreased abundance of *Catonella* in OM and SB, and of Tannerella in GCF at E75, suggesting a long-lasting reduction of these genera caused by allo-HSCT.

In summary, we observed that the dynamics of some commensal bacteria reproduce the same pattern of displacement during allo-HSCT and recovery after engraftment observed for the overall community. We also observed the emergence of opportunistic potentially pathogenic genera during the most perturbed allo-HSCT phase which are capable of colonizing all three oral sites and likely contribute to the loss of compositional differences between oral microbiotas observed after preconditioning.

Emergence of opportunistic genera and allo-HSCT complications

The emergence of opportunistic genera during allo-HSCT can be more rigorously quantified by assessing taxa blooms, defined as a taxon relative abundance increase from <1% at P to dominance levels (≥30%) at any subsequent timepoint. We have previously shown, by analyzing this same cohort, blooms of specific genera occurring in SB during A and E [24]. We now extended this analysis to other oral sites and to the post-engraftment period. Overall, we detected 81 blooms, involving 22 genera and 27/31 patients. All oral sites showed several blooming events, but SB blooms were more frequent (SB: n = 35; GCF: n = 24; OM: n = 22; Fig. 4a) and significantly more prevalent (SB: 23/31; GCF: 14/31; MO: 16/30; chi-square test, P-value = 0.022). Blooms typically occurred at E (53% of events; Fig. 4b) and were rapidly resolved in the post-engraftment period.

Lactobacillus (15%), *Enterococcus* (12%), and *Staphylococcus* (10%) were the genera most frequently observed in blooming events in the oral microbiota during allo-HSCT (Fig. 4c). But oral sites differed in the genera typically associated with blooms (Fig. 4d). SB showed mainly *Enterococcus* (7 events) or *Lactobacillus* (6) blooms, while GCF showed mostly *Staphylococcus* (4) or *Lactobacillus* (4) blooms. In contrast, OM blooms showed a less clear signal of blooming genera. Nevertheless, some patients presented concomitant blooms of the same genus in all oral sites.

We noticed that many of the blooming genera are potentially pathogenic for allo-HSCT recipients. For instance, *Staphylococcus* genus contains species related to several infections, including hospital-acquired pneumonia [56], an allo-HSCT complication with 15-30% incidence [57]. Therefore, we evaluated whether blooming events in the oral microbiota were associated with respiratory infections in our cohort. Between P and E75, only 3/31 patients presented bacterial respiratory infections (patients #1, #2, and #7). All three patients showed blooms of genera in the oral microbiota during allo-HSCT. Specifically, patient #1 presented blooms of *Enterococcus* (in GCF and SB at E) and Acetobacter (in GCF and SB at E30), patient #2 presented blooms of *Stenotrophomonas* (in all oral sites at E) and *Mycoplasma* (in GCF at E), and patient #7 presented blooms of *Mycoplasma* (in OM and SB at E). Interestingly, patients #1 and #2 presented blooms of the same genus identified in the microbiological exam of their respiratory tract samples: *Enterococcus* and *Stenotrophomonas*, respectively. Importantly, these blooms preceded the clinical manifestation of the respiratory infection by one and two weeks, respectively, suggesting a potential oral origin for the bacteria associated with the respiratory infections in these cases. On the other hand, patient #7 developed a respiratory infection caused by *Escherichia coli* between E30 and E75, which was unrelated to the blooms detected for this patient.

Given the apparent translocation of abundant oral bacteria to the respiratory tract in our cohort and the well-known association between intestinal dominance and bacteremia during allo-HSCT [58], we also tested whether blooming events in the oral microbiota were associated with bacteremia events. Positive blood cultures for bacteria were detected for 15/31 patients between P and E75. We did not find an association between oral microbiota blooms and altered odds of bacteremia (Fisher's exact test, GCF bloom: OR = 3.17, P-value = 0.156; OM bloom: OR = 2.25, P-value = 0.299; SB bloom: OR = 0.92, P-value = 1; any site bloom: OR = 3.12, P-value = 0.600). We detected a single case in which the blooming of a genus in the oral microbiota preceded a bacteremia event with the same genus involved. In detail, patient #14 presented blooms of *Enterococcus* in GCF and SB at A, which preceded positive blood cultures for *Enterococcus* by 1.5 weeks.

Overall, we observed that blooms of opportunistic genera occur frequently in oral microbiota during allo-HSCT, especially in SB. The examples described suggest oral microbiota blooms during allo-HSCT may trigger translocation of oral microbes to the respiratory tract (as often happens during oral microbiome dysbiosis [56]) and cause respiratory infections in allo-HSCT recipients.

Impact of antibiotic usage on oral microbiota dynamics

To investigate the impact of antibiotic usage on oral microbiota dynamics and blooming events during allo-HSCT, we analyzed antibiotic usage data between P and E30 (see Materials and Methods). Antibiotic usage varied widely across patients in terms of length of therapy (LOT range: 0–58 days; median: 15.5 days) and days of therapy (DOT range: 0–112 days; median: 22 days) (Table S1). Overall, 17 antibiotic agents (range: 0-10; median: 3), spanning 12 distinct antibiotic classes (range: 0–9; median: 3 antibiotics) were administered to our patients. The antibiotics administered to each patient are illustrated in Fig. 5a. Most patients received cefepime (73%) and meropenem (63%), making cephalosporins and carbapenems the most frequently used antibiotic classes: 73% and 63%, respectively (Fig. S9a). Glycopeptides and penicillins were also used in a considerable proportion of patients: 60% and 23%, respectively. All other antibiotic classes were used by less than 17% of our patients (Fig. S9b).

First, to assess the effect of antibiotic usage in microbiota dynamics, we modeled diversity stability (which incorporates diversity resistance and resilience) and compositional stability using antibiotic usage information (Table S2). We found that DOT significantly predicted diversity stability during allo-HSCT for all oral sites, with prolonged use of antibiotic therapy associated with lower diversity stability. However, the use of specific antibiotic classes was not associated with altered diversity stability (Table S2). On the other hand, DOT was not a predictor of compositional stability, but glycopeptide usage was significantly associated with decreased SB compositional stability (Table S2). In addition, we found non-significant associations at P-value < 0.1 between other antibiotic classes and decreased

15

compositional stability in GCF (cephalosporins and penicillins) and SB (cephalosporins), while OM compositional stability was clearly less impacted by antibiotic usage during allo-HSCT (Table S2).

We next tested whether blooms at different oral sites were associated with antibiotic usage. E75 blooms were not considered in this analysis since our antibiotic usage survey focused on the period between P and E30 (see Materials and Methods). With one exception (glycopeptides and GCF blooms), the use of specific antibiotic classes was not associated with blooms, but patients experiencing blooms showed higher LOT and DOT (Fig. 5b), although it is not clear whether a more extended period under antibiotic therapy was the cause or consequence of the blooms.

GCF blooms were significantly associated not only with LOT and DOT but also with the use of glycopeptides (Fisher's exact test, odds ratio (OR) = 15.65, P-value = 0.006, P-adjusted = 0.025), which enabled the investigation of the relation between the timing of glycopeptide usage and GCF blooming events. GCF blooms occurred in 12 patients up to E30, out of which 11 used glycopeptides (vancomycin and/or teicoplanin) between P and E30. Notably, 10/11 patients that used glycopeptides and experienced GCF blooms received glycopeptides a few days before or during the interval in which the bloom was detected, indicating that glycopeptide usage during allo-HSCT may cause blooms of genera in the oral microbiota.

The relationship between glycopeptide usage and blooming events and its consequences can be illustrated by the genera composition trajectories and antibiotic usage timeline of patients #1 and #2. Patient #2 experienced *Stenotrophomonas* blooms in all sites at E, which occurred during the administration of vancomycin (Fig. 5c). Two weeks after these blooms, patient #2 developed a respiratory infection caused by *Stenotrophomonas maltophilia*, detected in microbiological exams of respiratory tract samples (e.g., bronchoalveolar lavage). Despite the intensification in the use of antibiotics, *Stenotrophomonas* levels only rose in the oral microbiota after E, reaching staggering levels at E30 (>95% relative abundance in all oral sites). Analysis at ASV level revealed that

Stenotrophomonas ASVs were absent in patient #2 at P (relative abundance = 0% in all oral sites). At A, during the first course of vancomycin (Fig. 5c), a *Stenotrophomonas maltophilia* ASV emerged in the SB (relative abundance = 0.02%). This ASV would later be responsible for the blooms at E and the domination observed at E30. Taken together, these results suggest that the use of vancomycin during allo-HSCT allowed the emergence and the bloom of pathogenic *Stenotrophomonas maltophilia* in oral microbiota, which later translocated to the respiratory tract, causing a respiratory infection. Patient #1 presented a similar picture (Fig. S10), with the use of vancomycin followed by *Enterococcus* blooms and a subsequent respiratory infection caused by *Enterococcus faecium*. Notably, patients #1 and #2 died before E75, with death causes at least partially associated with their respiratory infections.

In summary, greater time of antibiotic exposure was associated with lower microbiota diversity stability and blooms in all oral sites. Glycopeptide usage was associated with lower microbiota compositional stability in SB and, although direct evidence is lacking, it seems causally linked to some of the blooming events.

Inter-patient variability in oral microbiota dynamics during allo-HSCT and after engraftment

To investigate inter-patient variability in oral microbiota dynamics during allo-HSCT and after engraftment, we assessed longitudinal changes in oral microbiota in a patient-centered analysis. Although most patients presented high diversity stability, which was achieved either by having high resistance, high resilience, or a balance between the two, some patients presented low diversity stability and even negative resilience values (Fig. 6a), indicating loss of diversity after E. Curiously, this inter-patient variability was not due to different levels of baseline diversity, since diversity at P was not correlated with diversity resistance, resilience, nor stability (Fig. S11a). Compositional stability was also not correlated with diversity levels at P (Fig. S11b)

In addition, when representing samples from all timepoints using Principal Coordinate Analysis (PCoA), we noticed that confidence intervals for E samples were larger, indicating

considerable inter-patient compositional variability under perturbation (Fig. S11c). To confirm this observation, we determined the most perturbed timepoint by quantifying the extent of compositional shifts between timepoints. As presented in Fig. 6b, compositional changes were more pronounced between A and E. Next, we evaluated inter-patient compositional variability at each timepoint either by assessing the compositional distance between samples and the respective timepoint centroid (Fig. 6c) or by calculating for each timepoint all pairwise compositional distances (Fig. 6d). Both results confirmed maximum inter-patient compositional variability at E under maximized perturbation, underscoring that allo-HSCT modifies oral microbiota differently for each patient.

Finally, we investigated if this variability in oral microbiota dynamics during allo-HSCT influenced oral microbiota recovery after engraftment. Although our results indicate that post-engraftment samples overall occupy a similar compositional space in comparison to P, this does not necessarily imply that patients recover their respective initial oral microbiota compositions after engraftment. In order to evaluate oral microbiota compositional recovery per patient, we analyzed the compositional distance from P for each patient and each site during allo-HSCT and after engraftment. Interestingly, we noted that even though most patients showed a recovery trajectory after engraftment, some did not (Fig. 6e).

Our data indicate a marked inter-patient variability in oral microbiota dynamics in response to allo-HSCT. Despite oral microbiotas as a whole resembling preconditioning microbiotas after allo-HSCT, patients differ in their ability to recover their initial oral microbiota composition.

Recovery of oral microbiota composition and allo-HSCT outcomes

To investigate whether oral microbiota recovery after allo-HSCT was associated with allo-HSCT outcomes we grouped our patients based on their ability to recover their preconditioning composition. We calculated the compositional distance between P and E30, and classified patients as recoverers (distance <0.5) or non-recoverers (distance ≥0.5). We further illustrate these contrasting recovery behaviors using PCoA with compositional

18

trajectories of a representative OM recoverer and of an OM non-recoverer (Fig. 7a). PCoAs for each patient are presented in Fig. S12. Overall, 77, 69, and 77% of our patients recovered their initial GCF, OM, and SB microbiota composition after engraftment, respectively (Fig. 7b).

Next, we used univariate analysis to investigate whether oral microbiota recovery after allo-HSCT was associated with allo-HSCT outcomes (Table S3; Fig. S13). Interestingly, OM recovery was associated with prolonged overall survival (OS; hazard ratio (HR) [95% confidence interval (CI)] = 0.17 [0.05–0.52], P-value = 0.002; Fig. 7c), prolonged progression-free survival (PFS; HR [95% CI] = 0.06 [0.01–0.34], P-value = 0.001; Fig. 7d), and a lower risk of underlying disease relapse (HR [95% CI] = 0.20 [0.06–0.69], P-value = 0.011; Fig. 7e). OM recovery, however, was not associated with altered risk of transplant-related death and GCF recovery or SB recovery were not associated with allo-HSCT outcomes (Table S3; Fig. S13).

To identify possible confounding variables, we used univariate analysis to investigate whether clinical parameters (including antibiotic usage; Table S1) were associated with allo-HSCT outcomes (Table S4–7). We found that disease risk index (DRI), conditioning intensity, and DOT were significantly associated with OS (Table S4). DRI was also associated with PFS (Table S5) and the risk of underlying disease relapse (Table S6). We then used a multivariate analysis to assess whether OM recovery was an independent predictor of allo-HSCT outcomes (Table S8). In all cases, OM recovery remained significantly associated with prolonged OS (HR [95% CI] = 0.09 [0.02–0.35], P-value < 0.001; Fig. 7f), prolonged PFS (HR [95% CI] = 0.09 (0.02–0.49), P-value = 0.005; Fig. 7g), and with a lower risk of underlying disease relapse (HR [95% CI] = 0.19 [0.06–0.55], P-value = 0.003; Fig. 7h). Taken together, these results robustly indicate that OM recovery at E30 is an independent biomarker of better allo-HSCT outcomes.

19

Underlying factors associated with oral mucosa microbiota recovery

Given the relevant associations between OM recovery and allo-HSCT outcomes, we searched for underlying factors associated with OM recovery. OM recovery was not associated with clinical parameters such as age, underlying disease, and graft source (Table S9). The usage of specific antibiotic classes, LOT, and DOT between P and E30 were also not associated with OM recovery (Table S9; Fig. S14a). In addition, OM recoverers and non-recoverers showed similar intervals between stem-cell infusion and engraftment (Fig. S14b).

We also evaluated whether OM microbiota characteristics could be related to OM recovery. OM recoverers did not show higher OM diversity at E30 (Fig. 8a), indicating OM non-recoverers did not necessarily possess dysbiotic OM microbiotas at E30. In line with this, OM blooms throughout allo-HSCT were not more frequent among OM non-recoverers (Fisher's exact test, OR = 4.07, P-value = 0.13). On the other hand, OM recoverers showed higher OM diversity at P and E (Fig. 8a). In fact, there was a significant negative correlation between OM diversity at P and the compositional distance between P and E30 (Fig. 8b). This effect was not observed for GCF and SB (Fig. 8b).

Lastly, we investigated if earlier reconstitution of blood cell counts was associated with OM recovery (see Additional File 3; Fig. 8c). Blood cell counts at P or E were not associated with OM recovery. Interestingly, however, OM recoverers showed higher leukocyte counts at E30, which is mostly due to significantly higher neutrophil and lymphocyte counts in this group. Furthermore, normal (within reference values) leukocyte counts at E30 were more frequently observed among OM recoverers compared to OM non-recoverers (16/20 vs. 3/9, respectively; Fisher's exact test, P-value = 0.032) and OM recoverers presented higher leukocyte counts throughout one year after allo-HSCT compared to non-recoverers due to the combined contribution of higher neutrophil, lymphocyte, and monocyte counts (Fig. S14c).

In summary, we found independent (blood cell counts) and non-independent (OM microbiota at P) parameters to illuminate the differences between OM recoverers and non-

20

recoverers. OM recovery was associated with higher diversity at P, indicating more diverse OM communities are more competent in recovering their pre-perturbation compositions. In addition, OM recoverers showed higher leukocyte counts at E30, suggesting an association between OM microbiota composition recovery and earlier immune system reconstitution.

Discussion

The anatomical complexity of the oral cavity provides a multitude of physicochemical environments for microbes to thrive [1, 3]. Although several dozen core bacterial genera inhabit all oral compartments, different species occupy each oral niche, meaning oral microbes are site-specialists that compose distinct microbiotas in each oral environment [1, 59]. We and others have previously reported the impact of allo-HSCT in oral microbiotas and their associations with allo-HSCT complications and outcomes [22–26, 30]. However, these studies analyzed single oral sites and were mostly limited to the peri-engraftment period of allo-HSCT. To our knowledge, this is the first study to evaluate the impact of allo-HSCT in the microbiota of various oral sites simultaneously during and after allo-HSCT.

We found that the microbiota of all oral sites was severely damaged by allo-HSCT, but each site responded differently to the perturbations associated with allo-HSCT. Compositional differences between oral sites were lost during allo-HSCT and partially recovered after engraftment. Oral microbiota injury was marked by loss of diversity and emergence of opportunistic potentially pathogenic genera. Notably, these opportunistic genera could colonize all three oral sites and likely contributed to the loss of compositional differences between oral microbiotas observed after conditioning. Colonization by opportunistic genera was more common at E, explaining the higher compositional variability and lower diversity observed at E, which we found to be the most perturbed allo-HSCT phase for all oral sites. This is in line with the Anna Karenina Principle applied to host-associated microbiomes [60], which states that more diverse communities tend to be more

21

553 compositionally similar, while perturbed communities tend to occupy several alternative
554 dysbiotic states.

555 Blooms of opportunistic genera were associated with prolonged antibiotic exposure
556 and the use of glycopeptides. This association is clinically relevant in the allo-HSCT setting
557 since glycopeptide-resistant bacteria (e.g., vancomycin-resistant enterococci) are a common
558 cause of infections in the hospital environment [61], especially in immunosuppressed
559 individuals. In addition, we observed that, in some cases, oral microbiota blooms preceded
560 respiratory infections caused by the blooming bacteria, linking the oral microbiota dynamics
561 during allo-HSCT to a common allo-HSCT complication [62], probably due to translocation of
562 oral bacteria to the respiratory tract through aspiration [56]. Similarly to our study, Thänert et
563 al. (2019) showed pathobiont blooms in the gut microbiota often preceded urinary tract
564 infections, but, as observed here, not all blooms were associated with subsequent infection
565 [63]. Interestingly, even though the mouth is a highly vascularized organ and the existence of
566 an oral-blood translocation axis has been proposed [64], we did not find a clear association
567 between oral bacteria blooms and bacteremia events during allo-HSCT.

568 Respiratory infections following blooms were caused by *E. faecium* in patient #1 and
569 *S. maltophilia* in patient #2. *S. maltophilia* colonization has been reported in 7% of allo-HSCT
570 recipients and is associated with higher non-relapse mortality risk due to higher odds of
571 invasive *S. maltophilia* infections [65]. Our results highlight that nosocomial bacteria such as
572 *S. maltophilia* can colonize the oral cavity during allo-HSCT. These results point to the
573 importance of maintaining oral health during allo-HSCT not only to prevent oral but also
574 distal complications (e.g., hospital-acquired pneumonia) [56]. Furthermore, our results
575 suggest that tracking drastic oral microbiota changes during allo-HSCT may guide early
576 interventions to prevent infections. This will be especially useful when the causative agent is
577 not a common respiratory pathogen such as in the case of *E. faecium* [66].

578 Longitudinal analysis of oral microbiota diversity and composition showed post-
579 transplant oral microbiotas were overall similar to preconditioning microbiotas, but patient-
580 level analysis showed that 23-31% of the patients did not recover their preconditioning

581 microbiota composition. Variability in gut microbiota recovery following a perturbation has
582 been previously described [67, 68], including after allo-HSCT, where most patients (>90%)
583 were unable to recover their initial gut microbiota composition [68]. The higher proportion of
584 patients that recovered their preconditioning composition in our study suggests that the oral
585 microbiota is more resilient to the perturbations associated with allo-HSCT than the gut
586 microbiota. This result is in line with a previous study showing that the oral microbiota is
587 more resilient than the gut microbiota to antibiotic perturbation [69].

588 Pre-perturbation microbiota characteristics, such as the presence of keystone
589 bacteria, influence microbiota recovery [70]. Here, we found that patients that recovered their
590 OM microbiota composition after allo-HSCT showed higher preconditioning OM diversity,
591 indicating that more diverse OM microbiotas are more resilient to allo-HSCT. Our results
592 converge on the insurance hypothesis, which proposes that high-diversity communities are
593 less susceptible to perturbations [71]. Interestingly, in our study, OM compositional recovery
594 was not associated with the use of specific antibiotics nor with the duration of antibiotic
595 exposure. This is possibly because OM microbiota composition is not impacted by
596 antibiotics, as evidenced by the lack of associations between antibiotic usage and OM
597 compositional stability. Host genetics, reestablishment of normal diet, and reconstitution of
598 the immune system are other possible drivers of microbiota recovery after allo-HSCT. Here,
599 we showed that leukocyte blood counts at E30 were higher in patients that recovered their
600 OM microbiota composition, indicating a close link between early immune system
601 reconstitution and oral microbiota recovery. We can speculate that immune reconstitution
602 allows stricter control of microbiota compositions (e.g., via immunoglobulin A [72]), which,
603 along with reestablishment of microbial environment (e.g., normal diet), supports the
604 recovery of the initial OM microbiota composition [73, 74].

605 The ability to recover the OM initial microbiota composition was associated with
606 better allo-HSCT outcomes. However, it is unclear if OM microbiota recovery is just a
607 consequence or also a driver of early immune reconstitution, thus having a causal role in the
608 improved outcomes following allo-HSCT. Evidence from gut microbiota studies indicates that

23

609 the latter hypothesis is plausible [75]. For instance, recent studies have shown that specific
610 gut microbes are associated with immune cell dynamics post-allo-HSCT [15, 76]. Similarly,
611 Miltiadous et al. (2022) found that higher peri-engraftment gut microbiota diversity was
612 associated with higher lymphocyte counts 100 days after transplant [77]. In addition, murine
613 model experiments showed that gut microbiota supports immune reconstitution by allowing a
614 higher dietary energy uptake [78]. Most importantly, in a controlled randomized clinical trial,
615 patients who received autologous fecal microbiota transplant after allo-HSCT showed higher
616 leukocyte counts 100 days after engraftment, indicating recovery of the gut microbiota has a
617 causal role in facilitating immune system reconstitution [15]. If this causal relationship
618 extends to the oral microbiota, the use of therapeutic interventions to promote oral health
619 and microbiota recovery in allo-HSCT recipients, such as oral microbiota transplants [79],
620 could potentially improve allo-HSCT outcomes.

621 An important limitation of our study is its small sample size, which did not allow
622 underlying disease stratification to parse the effect of different diseases on oral microbiota
623 dynamics. Still, the longitudinal design, assessment of different oral sites, and evaluation of
624 a Brazilian cohort (a population underrepresented in human microbiome studies [80]) with
625 extensive metadata publicly available are strengths of our study that should be highlighted.
626 Also, to better address the influence of oral bacteria in immune cell dynamics, future studies
627 will have to combine high temporal resolution oral microbiota data with more deeply
628 phenotyped immune cell counts (e.g., flow cytometry data). In addition, since 16S rRNA
629 amplicon sequencing has limited taxonomic resolution, further studies should ideally be
630 performed using shotgun metagenomic sequencing, as this would allow strain-level
631 dynamics tracking. Finally, here and previously [24, 25], we showed that associations
632 between gut microbiota and allo-HSCT outcomes broadly extend to the oral microbiota.
633 However, studies with synchronous gut and oral microbiota profiling will be necessary to
634 decipher how these microbiotas are linked during allo-HSCT, especially considering the
635 higher translocation of oral bacteria along the oral-gut axis during disease [81].

24

Conclusions

The oral cavity is the ultimate doorway for microbes entering the human body. We analyzed oral microbiotas dynamics in allo-HSCT recipients and showed that microbiota injury and recovery patterns were highly informative on allo-HSCT complications and outcomes. Our results highlight the importance of tracking recipient's microbiotas changes during allo-HSCT to improve our understanding of allo-HSCT biology, safety, and efficacy.

Availability of data and materials

The bioinformatics pipeline used to process the sequencing data, the R scripts used to run the analyses and generate the figures, and all clinical metadata (anonymized) necessary to reproduce these results are available at https://github.com/vitorheidrich/oral-microbiota-hsct. Raw sequencing data have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB53914. Some samples (analyzed in past studies) were deposited previously in ENA at EMBL-EBI under accession numbers: PRJEB42862, PRJEB49175.

Abbreviations

A: Aplasia

Allo-HSCT: Allogeneic hematopoietic stem-cell transplant

ASV: Amplicon Sequence Variant

CI: Confidence interval

DOT: Days of therapy

DRI: Disease Risk Index

E: Engraftment

E30: 30 days after engraftment

E75: 75 days after engraftment

E. faecium: Enterococcus faecium

GCF: Gingival crevicular fluid

HR: Hazard ratio

LOT: Length of therapy

OM: Oral mucosa

OR: Odds ratio

OS: Overall survival

P: Preconditioning

PFS: Progression-free survival

SB: Supragingival biofilm

S. maltophilia: Stenotrophomonas maltophilia

SRS: Scaling with ranked subsampling

References

1. Welch JLM, Ramírez-Puebla ST, Borisy GG. Oral Microbiome Geography: Micron-Scale Habitat and Niche. Cell Host Microbe. 2020;28:160–8.

2. Rosier BT, Marsh PD, Mira A. Resilience of the Oral Microbiota in Health: Mechanisms That Prevent Dysbiosis. J Dent Res. 2018;97:371–80.

3. Proctor DM, Relman DA. The Landscape Ecology and Microbiota of the Human Nose, Mouth, and Throat. Cell Host Microbe. 2017;21:421–32.

4. Segata N, Haake SK, Mannon P, Lemon KP, Waldron L, Gevers D, et al. Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. Genome Biol. 2012;13:R42.

5. Tuganbaev T, Yoshida K, Honda K. The effects of oral microbiota on health. Science. 2022;376:934–6.

6. Takahashi N, Nyvad B. The Role of Bacteria in the Caries Process: Ecological Perspectives. J Dent Res. 2011;90:294–303.

7. Ng E, Tay JRH, Balan P, Ong MMA, Bostanci N, Belibasakis GN, et al. Metagenomic sequencing provides new insights into the subgingival bacteriome and aetiopathology of periodontitis. J Periodontal Res. 2021;56:205–18.

8. Gaffen SL, Moutsopoulos NM. Regulation of host-microbe interactions at oral mucosal barriers by type 17 immunity. Sci Immunol. 2020;5:eaau4594.

9. Dominy SS, Lynch C, Ermini F, Benedyk M, Marczyk A, Konradi A, et al. Porphyromonas gingivalis in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. Sci Adv. 2019;5:eaau3333.

10. Snowden JA, Sánchez-Ortega I, Corbacioglu S, Basak GW, Chabannon C, de la

695 Camara R, et al. Indications for haematopoietic cell transplantation for haematological
696 diseases, solid tumours and immune disorders: current practice in Europe, 2022. Bone
697 Marrow Transplant. 2022;57:1217–39.

698 11. Jenq RR, van den Brink MRM. Allogeneic haematopoietic stem cell transplantation:
699 individualized stem cell and immune therapy of cancer. Nat Rev Cancer. 2010;10:213–21.

700 12. Gyurkocza B, Sandmaier BM. Conditioning regimens for hematopoietic cell
701 transplantation: one size does not fit all. Blood. 2014;124:344–53.

702 13. Welniak LA, Blazar BR, Murphy WJ. Immunobiology of Allogeneic Hematopoietic Stem
703 Cell Transplantation. Annu Rev Immunol. 2007;25:139–70.

704 14. Lehrnbecher T, Fisher BT, Phillips B, Alexander S, Ammann RA, Beauchemin M, et al.
705 Guideline for Antibacterial Prophylaxis Administration in Pediatric Cancer and Hematopoietic
706 Stem Cell Transplantation. Clin Infect Dis. 2020;71:226–36.

707 15. Schluter J, Peled JU, Taylor BP, Markey KA, Smith M, Taur Y, et al. The gut microbiota
708 is associated with immune cell dynamics in humans. Nature. 2020;588:303–7.

709 16. Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health
710 and disease. Cell Res. 2020;30:492–506.

711 17. Ervin SM, Ramanan SV, Bhatt AP. Relationship Between the Gut Microbiome and
712 Systemic Chemotherapy. Dig Dis Sci. 2020;65:874–84.

713 18. Wang L, Wang X, Zhang G, Ma Y, Zhang Q, Li Z, et al. The impact of pelvic radiotherapy
714 on the gut microbiome and its role in radiation-induced diarrhoea: a systematic review.
715 Radiat Oncol Lond Engl. 2021;16:187.

716 19. Ramirez J, Guarner F, Bustos Fernandez L, Maruy A, Sdepanian VL, Cohen H.
717 Antibiotics as Major Disruptors of Gut Microbiota. Front Cell Infect Microbiol. 2020;10.

718 20. Shono Y, van den Brink MRM. Gut microbiota injury in allogeneic haematopoietic stem
719 cell transplantation. Nat Rev Cancer. 2018;18:283–95.

720 21. Sen T, Thummer RP. The Impact of Human Microbiotas in Hematopoietic Stem Cell and
721 Organ Transplantation. Front Immunol. 2022;13.

722 22. Ames NJ, Barb JJ, Ranucci A, Kim H, Mudra SE, Cashion AK, et al. The oral microbiome
723 of patients undergoing treatment for severe aplastic anemia: a pilot study. Ann Hematol.
724 2019;98:1351–65.

725 23. Shouval R, Eshel A, Dubovski B, Kuperman AA, Danylesko I, Fein JA, et al. Patterns of
726 salivary microbiota injury and oral mucositis in recipients of allogeneic hematopoietic stem
727 cell transplantation. Blood Adv. 2020;4:2912–7.

728 24. Heidrich V, Bruno JS, Knebel FH, de Molla VC, Miranda-Silva W, Asprino PF, et al.
729 Dental Biofilm Microbiota Dysbiosis Is Associated With the Risk of Acute Graft-Versus-Host
730 Disease After Allogeneic Hematopoietic Stem Cell Transplantation. Front Immunol. 2021;12.

731 25. de Molla VC, Heidrich V, Bruno JS, Knebel FH, Miranda-Silva W, Asprino PF, et al.
732 Disruption of the oral microbiota is associated with a higher risk of relapse after allogeneic
733 hematopoietic stem cell transplantation. Sci Rep. 2021;11:17552.

734 26. Laheij AMGA, Rozema FR, Brennan MT, von Bültzingslöwen I, van Leeuwen SJM,
735 Potting C, et al. Long-Term Analysis of Resilience of the Oral Microbiome in Allogeneic Stem
736 Cell Transplant Recipients. Microorganisms. 2022;10:734.

27. Liu C, Frank DN, Horch M, Chau S, Ir D, Horch EA, et al. Associations between acute gastrointestinal GvHD and the baseline gut microbiota of allogeneic hematopoietic stem cell transplant recipients and donors. Bone Marrow Transplant. 2017;52:1643–50.

28. Stein-Thoeringer CK, Nichols KB, Lazrak A, Docampo MD, Slingerland AE, Slingerland JB, et al. Lactose drives Enterococcus expansion to promote graft-versus-host disease. Science. 2019;366:1143–9.

29. Peled JU, Gomes ALC, Devlin SM, Littmann ER, Taur Y, Sung AD, et al. Microbiota as Predictor of Mortality in Allogeneic Hematopoietic-Cell Transplantation. N Engl J Med. 2020;382:822–34.

30. Oku S, Takeshita T, Futatsuki T, Kageyama S, Asakawa M, Mori Y, et al. Disrupted tongue microbiota and detection of nonindigenous bacteria on the day of allogeneic hematopoietic stem cell transplantation. PLOS Pathog. 2020;16:e1008348.

31. Stanić Benić M, Milanič R, Monnier AA, Gyssens IC, Adriaenssens N, Versporten A, et al. Metrics for quantifying antibiotic use in the hospital setting: results from a systematic review and international multidisciplinary consensus procedure. J Antimicrob Chemother. 2018;73 suppl_6:vi50–8.

32. Wang H, Altemus J, Niazi F, Green H, Calhoun BC, Sturgis C, et al. Breast tissue, oral and urinary microbiomes in breast cancer. Oncotarget. 2017;8:88122–38.

33. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. 2013;41:e1.

34. Heng Li. seqtk: Toolkit for processing sequences in FASTA/Q formats.

35. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37:852–7.

36. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–3.

37. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016;4:e2584.

38. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:D590–6.

39. R Core Team. R: A language and environment for statistical computing. R Found Stat Comput Vienna Austria. 2021.

40. Jordan E. Bisanz. qiime2R: Importing QIIME2 artifacts and associated data into R sessions. 2018.

41. Beule L, Karlovsky P. Improved normalization of species count data in ecology by scaling with ranked subsampling (SRS): application to microbial communities. PeerJ. 2020;8:e9593.

42. Simpson EH. Measurement of Diversity. Nature. 1949;163:688–688.

43. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE. 2013;8:e61217.

779 44. Orwin KH, Wardle DA. New indices for quantifying the resistance and resilience of soil
780 biota to exogenous disturbances. Soil Biol Biochem. 2004;36:1907–12.

781 45. Shade A, Peter H, Allison S, Baho D, Berga M, Buergmann H, et al. Fundamentals of
782 Microbial Community Resistance and Resilience. Front Microbiol. 2012;3.

783 46. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: an effective
784 distance metric for microbial community comparison. ISME J. 2011;5:169–72.

785 47. Daniel P. Smith. rbiom: Read/Write, Transform, and Summarize “BIOM” Data. 2022.

786 48. Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. Nat
787 Commun. 2020;11:3514.

788 49. Cox DR. Regression Models and Life-Tables. J R Stat Soc Ser B Methodol.
789 1972;34:187–202.

790 50. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing
791 Risk. J Am Stat Assoc. 1999;94:496–509.

792 51. Heidrich V, Karlovsky P, Beule L. ‘SRS’ R Package and ‘q2-srs’ QIIME 2 Plugin:
793 Normalization of Microbiome Data Using Scaling with Ranked Subsampling (SRS). Appl Sci.
794 2021;11:11473.

795 52. Kolenbrander PE, Palmer RJ, Periasamy S, Jakubovics NS. Oral multispecies biofilm
796 development and the key role of cell–cell distance. Nat Rev Microbiol. 2010;8:471–80.

797 53. Barrak I, Stájer A, Gajdács M, Urbán E. Small, but smelly: the importance of
798 Solobacterium moorei in halitosis and other human infections. Heliyon. 2020;6:e05371.

799 54. Komiyama EY, Lepesqueur LSS, Yassuda CG, Samaranayake LP, Parahitiyawa NB,
800 Balducci I, et al. Enterococcus Species in the Oral Cavity: Prevalence, Virulence Factors and
801 Antimicrobial Susceptibility. PLOS ONE. 2016;11:e0163001.

802 55. Caufield PW, Schön CN, Saraithong P, Li Y, Argimón S. Oral Lactobacilli and Dental
803 Caries: A Model for Niche Adaptation in Humans. J Dent Res. 2015;94 9_suppl:110S-118S.

804 56. Dong J, Li W, Wang Q, Chen J, Zu Y, Zhou X, et al. Relationships Between Oral
805 Microecosystem and Respiratory Diseases. Front Mol Biosci. 2022;8.

806 57. Gudiol C, Sabé N, Carratalà J. Is hospital-acquired pneumonia different in transplant
807 recipients? Clin Microbiol Infect. 2019;25:1186–94.

808 58. Taur Y, Xavier JB, Lipuma L, Ubeda C, Goldberg J, Gobourne A, et al. Intestinal
809 Domination and the Risk of Bacteremia in Patients Undergoing Allogeneic Hematopoietic
810 Stem Cell Transplantation. Clin Infect Dis. 2012;55:905–14.

811 59. Mark Welch JL, Dewhirst FE, Borisy GG. Biogeography of the Oral Microbiome: The
812 Site-Specialist Hypothesis. Annu Rev Microbiol. 2019;73:335–58.

813 60. Zaneveld JR, McMinds R, Vega Thurber R. Stress and stability: applying the Anna
814 Karenina principle to animal microbiomes. Nat Microbiol. 2017;2:1–8.

815 61. Cookson BD, Macrae MB, Barrett SP, Brown DFJ, Chadwick C, French GL, et al.
816 Guidelines for the control of glycopeptide-resistant enterococci in hospitals. J Hosp Infect.
817 2006;62:6–21.

818 62. Sahin U, Toprak SK, Atilla PA, Atilla E, Demirer T. An overview of infectious
819 complications after allogeneic hematopoietic stem cell transplantation. J Infect Chemother.
820 2016;22:505–14.

821 63. Thänert R, Reske KA, Hink T, Wallace MA, Wang B, Schwartz DJ, et al. Comparative
822 Genomics of Antibiotic-Resistant Uropathogens Implicates Three Routes for Recurrence of
823 Urinary Tract Infections. mBio. 2019;10:e01977-19.

824 64. Abed J, Maalouf N, Manson AL, Earl AM, Parhi L, Emgård JEM, et al. Colon Cancer-
825 Associated Fusobacterium nucleatum May Originate From the Oral Cavity and Reach Colon
826 Tumors via the Circulatory System. Front Cell Infect Microbiol. 2020;10.

827 65. Scheich S, Koenig R, Wilke AC, Lindner S, Reinheimer C, Wichelhaus TA, et al.
828 Stenotrophomonas maltophilia colonization during allogeneic hematopoietic stem cell
829 transplantation is associated with impaired survival. PLOS ONE. 2018;13:e0201169.

830 66. Li F, Wang Y, Sun L, Wang X. Vancomycin-resistant Enterococcus faecium pneumonia
831 in a uremic patient on hemodialysis: a case report and review of the literature. BMC Infect
832 Dis. 2020;20:167.

833 67. Dethlefsen L, Relman DA. Incomplete recovery and individualized responses of the
834 human distal gut microbiota to repeated antibiotic perturbation. Proc Natl Acad Sci. 2011;108
835 supplement_1:4554–61.

836 68. Vaitkute G, Panic G, Alber DG, Faizura-Yeop I, Cloutman-Green E, Swann J, et al.
837 Linking gastrointestinal microbiota and metabolome dynamics to clinical outcomes in
838 paediatric haematopoietic stem cell transplantation. Microbiome. 2022;10:89.

839 69. Zaura E, Brandt BW, Teixeira de Mattos MJ, Buijs MJ, Caspers MPM, Rashid M-U, et al.
840 Same Exposure but Two Radically Different Responses to Antibiotics: Resilience of the
841 Salivary Microbiome versus Long-Term Microbial Shifts in Feces. mBio. 2015;6:e01693-15.

842 70. Chng KR, Ghosh TS, Tan YH, Nandi T, Lee IR, Ng AHQ, et al. Metagenome-wide
843 association analysis identifies microbial determinants of post-antibiotic ecological recovery in
844 the gut. Nat Ecol Evol. 2020;4:1256–67.

845 71. Sommer F, Anderson JM, Bharti R, Raes J, Rosenstiel P. The resilience of the intestinal
846 microbiota influences health and disease. Nat Rev Microbiol. 2017;15:630–8.

847 72. Hooper LV, Littman DR, Macpherson AJ. Interactions Between the Microbiota and the
848 Immune System. Science. 2012;336:1268–73.

849 73. Martino C, Dilmore AH, Burcham ZM, Metcalf JL, Jeste D, Knight R. Microbiota
850 succession throughout life from the cradle to the grave. Nat Rev Microbiol. 2022;20:707–20.

851 74. Ng KM, Aranda-Díaz A, Tropini C, Frankel MR, Treuren WV, O'Loughlin CT, et al.
852 Recovery of the Gut Microbiota after Antibiotics Depends on Host Diet, Community Context,
853 and Environmental Reservoirs. Cell Host Microbe. 2019;26:650-665.e4.

854 75. Fiorenza S, Turtle CJ. Associations between the Gut Microbiota, Immune Reconstitution,
855 and Outcomes of Allogeneic Hematopoietic Stem Cell Transplantation. Immunometabolism.
856 2021;3:e210004.

857 76. Ingham AC, Kielsen K, Cilieborg MS, Lund O, Holmes S, Aarestrup FM, et al. Specific
858 gut microbiome members are associated with distinct immune markers in pediatric
859 allogeneic hematopoietic stem cell transplantation. Microbiome. 2019;7:131.

860 77. Miltiadous O, Waters NR, Andrlová H, Dai A, Nguyen CL, Burgos da Silva M, et al. Early
861 intestinal microbial features are associated with CD4 T-cell recovery after allogeneic
862 hematopoietic transplant. Blood. 2022;139:2758–69.

863 78. Staffas A, Silva MB da, Slingerland AE, Lazrak A, Bare CJ, Holman CD, et al. Nutritional

Support from the Intestinal Microbiota Improves Hematopoietic Reconstitution after Bone Marrow Transplantation in Mice. Cell Host Microbe. 2018;23:447-457.e4.

79. Nath S, Zilm P, Jamieson L, Kapellas K, Goswami N, Ketagoda K, et al. Development and characterization of an oral microbiome transplant among Australians for the treatment of dental caries and periodontal disease: A study protocol. PLOS ONE. 2021;16:e0260433.

80. Abdill RJ, Adamowicz EM, Blekhman R. Public human microbiome data are dominated by highly developed countries. PLOS Biol. 2022;20:e3001536.

81. Jin S, Wetzel D, Schirmer M. Deciphering mechanisms and implications of bacterial translocation in human health and disease. Curr Opin Microbiol. 2022;67:102147.

Acknowledgements

Not applicable.

Funding

VH was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, process no. 13996-0/2018). FHK was supported by FAPESP (process no. 16854-4/2015).

Author information

Authors and Affiliations

Centro de Oncologia Molecular, Hospital Sírio-Libanês, São Paulo, SP, Brazil

Vitor Heidrich, Franciele H. Knebel, Julia S. Bruno, Wanessa Miranda-Silva, Paula F. Asprino, Eduardo R. Fregnani, Anamaria A. Camargo

Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, SP, Brazil

Vitor Heidrich

Centro de Oncologia, Hospital Sírio-Libanês, São Paulo, SP, Brazil

Luciana Tucunduva, Yana Novis

Hospital Nove de Julho, Rede DASA, São Paulo, SP, Brazil

Vinícius C. de Molla, Celso Arrais-Rodrigues

Universidade Federal de São Paulo, São Paulo, SP, Brazil

31

Vinícius C. de Molla, Celso Arrais-Rodrigues

Hospital das Clínicas da Faculdade de Medicina, Universidade de São Paulo/Instituto do Câncer do Estado de São Paulo (ICESP), São Paulo, SP, Brazil

Vanderson Rocha

Contributions

EFR and AAC designed the study. VCM, LT, VR, YN, and CAR recruited and clinically evaluated volunteers. VCM collected data from clinical records. WMS collected oral samples. FHK processed most of the samples. FHK and PFA performed the sequencing. VH and AC conceptualized the analysis. VH performed all bioinformatics and statistical analyses. VH, JSB, and AAC contributed to the interpretation of results. VH and AAC wrote the original draft of the manuscript. VH, JSB, VCM, PFA, CAR, and AAC reviewed and edited the manuscript. All authors read and approved the final manuscript.

Corresponding author

Correspondence to Anamaria A. Camargo.

Ethics declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of Hospital Sírio-Libanês (#HSL 2016-08), in line with the Declaration of Helsinki. All patients provided their written informed consent to participate.

Consent for publication

Not applicable.

32

Competing interests

The authors declare that they have no competing interests.

Figure legends

Figure 1

a Principal Coordinate Analysis (PCoA) of microbiota distances (weighted UniFrac) between oral sites for each timepoint. Ellipsoids indicate 95% confidence intervals. **b** Magnitude (PERMANOVA F) of distances (weighted UniFrac) between oral sites per timepoint. **c** Minimum distance (weighted UniFrac) between oral sites within patients per timepoint. Mann-Whitney U test was used with preconditioning (P) as the reference for comparisons. **d** Number of differentially abundant genera (ANCOM-BC) between oral sites per timepoint. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days after engraftment; **, P-value < 0.01; ***, P-value < 0.001.

Figure 2

a Diversity (Gini-Simpson) per timepoint for each oral site. Mann-Whitney U test was used with preconditioning (P) as the reference for comparisons. **b** Diversity resistance, resilience, and stability (see Methods) per oral site. Mann-Whitney U test was used. Distance to P centroid (weighted UniFrac) per timepoint for each oral site. Mann-Whitney U test was used with P as the reference for comparisons. **d** Magnitude (PERMANOVA F) of distances (weighted UniFrac) between P and other timepoints for each site. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days after engraftment; *, P-value < 0.05; **, P-value < 0.01; ***, P-value < 0.001; ****, P-value < 0.0001.

Figure 3

a Mean genera relative abundances (RA) per timepoint for each oral site. Genera with >2% mean RA in any combination of oral site and timepoint are shown. **b** Mean genera RA ranking per timepoint for each oral site. Top-10 genera are shown. **c** Differentially abundant genera (ANCOM-BC) between P and other timepoints for each site. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; P, preconditioning; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days

33

after engraftment; *, q-value < 0.05; **, q-value < 0.01; ***, q-value < 0.001; z, ANCOM-BC structural zero.

Figure 4

a-c Proportion of blooming events per oral site (a), timepoint (b) and genus (c). **d** Number of blooming events per genus in each oral site. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days after engraftment.

Figure 5

a Antibiotic agents used by each patient between preconditioning (P) and 30 days after engraftment (E30). **b** Time of antibiotic administration (LOT: length of therapy; DOT: days of therapy) among patients showing and not showing blooms between P and E30. **c** Patient #2: genera relative abundance dynamics for each oral site (top) and antibiotic usage timeline (bottom). Genera with >1% mean relative abundance in any combination of oral site and timepoint are shown. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; A, aplasia; E, engraftment; E75, 75 days after engraftment; SC, stem-cell; vanc, vancomycin; tige, tigecycline; tazo, piperacillin tazobactam; poli, polymyxin B; mero, meropenem; line, linezolid.

Figure 6

a Relationship between diversity resistance, resilience, and stability values calculated for each patient. **b** Extent of compositional shifts (weighted UniFrac) between consecutive timepoints (adjusted for the time in days between timepoints) for each oral site. The line indicates the median value per interval. **c** Distance (weighted UniFrac) to timepoint centroid per timepoint for each oral site. Mann-Whitney U test was used with preconditioning (P) as the reference for comparisons. **d** Pairwise distances (weighted UniFrac) per timepoint (all-against-all) for each oral site. Mann-Whitney U test was used with P as the reference for comparisons. **e** Distance to P (weighted UniFrac) at engraftment (E) and 30 days after engraftment (E30) for each patient for each oral site. The thick line indicates the median value at each timepoint. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; A, aplasia; E75, 75 days after engraftment; **, P-value < 0.01; ***, P-value < 0.001; ****, P-value < 0.0001.

Figure 7

34

a Principal Coordinate Analysis (PCoA) with representative microbiota trajectories of an oral mucosa (OM) recoverer and non-recoverer. **b** Recovery classifications per site for each patient. Patient #1 OM recovery could not be evaluated due to missing samples. **c-d** Kaplan-Meier curves comparing overall survival (c) and progression-free survival (d) among OM recoverers (R) and non-recoverers (NR). **e** Cumulative incidence curves of relapse among OM R and OM NR. **f-h** multivariate analysis for overall survival (f), progression-free survival (g), and risk of relapse (h). Each model includes OM recovery and the clinical variables that are relevant for each outcome. P, preconditioning; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days after engraftment; HR, hazard ratio; DRI, disease risk index; DOT, days of antibiotic therapy; Cond Int, conditioning intensity.

Figure 8

a Diversity (Gini-Simpson) among oral mucosa (OM) recoverers and non-recoverers for each timepoint. Mann-Whitney U test was used. **b** Correlation between diversity (Gini-Simpson) at preconditioning (P) and the compositional distance (weighted UniFrac) between P and 30 days after engraftment (E30) for each oral site. Spearman's rank correlation test was used. **c** Blood cell counts among OM recoverers and non-recoverers per timepoint for each blood cell type. Red dotted horizontal lines indicate normal counts (within reference values). Mann-Whitney U test was used. A, aplasia; E, engraftment; E75, 75 days after engraftment; *, P-value < 0.05; **, P-value < 0.01.

Supplementary information

Additional file 1: Timelines of antibiotic usage.

Antibiotic usage timelines for each patient in relation to stem-cell infusion. Red dashed line indicates preconditioning sampling. Red solid line indicates stem-cell infusion. Blue solid line indicates stem-cell engraftment. Blue dashed line indicates 30 days after engraftment sampling. clav, amoxicillin clavulanate; tazo, piperacillin tazobactam; amox, amoxicillin; cefe, cefepime; mero, meropenem; metr, metronidazole; ceft, ceftriaxone; vanc, vancomycin; teic, teicoplanin; cipr, ciprofloxacin; levo, levofloxacin; doxi, doxycycline; ampi, ampicillin; clar, clarithromycin; bact, sulfamethoxazole trimethoprim; erta, ertapenem; poli, polymyxin b; dapt, daptomycin;line, linezolid; tige, tigecycline; amic, amikacin.

35

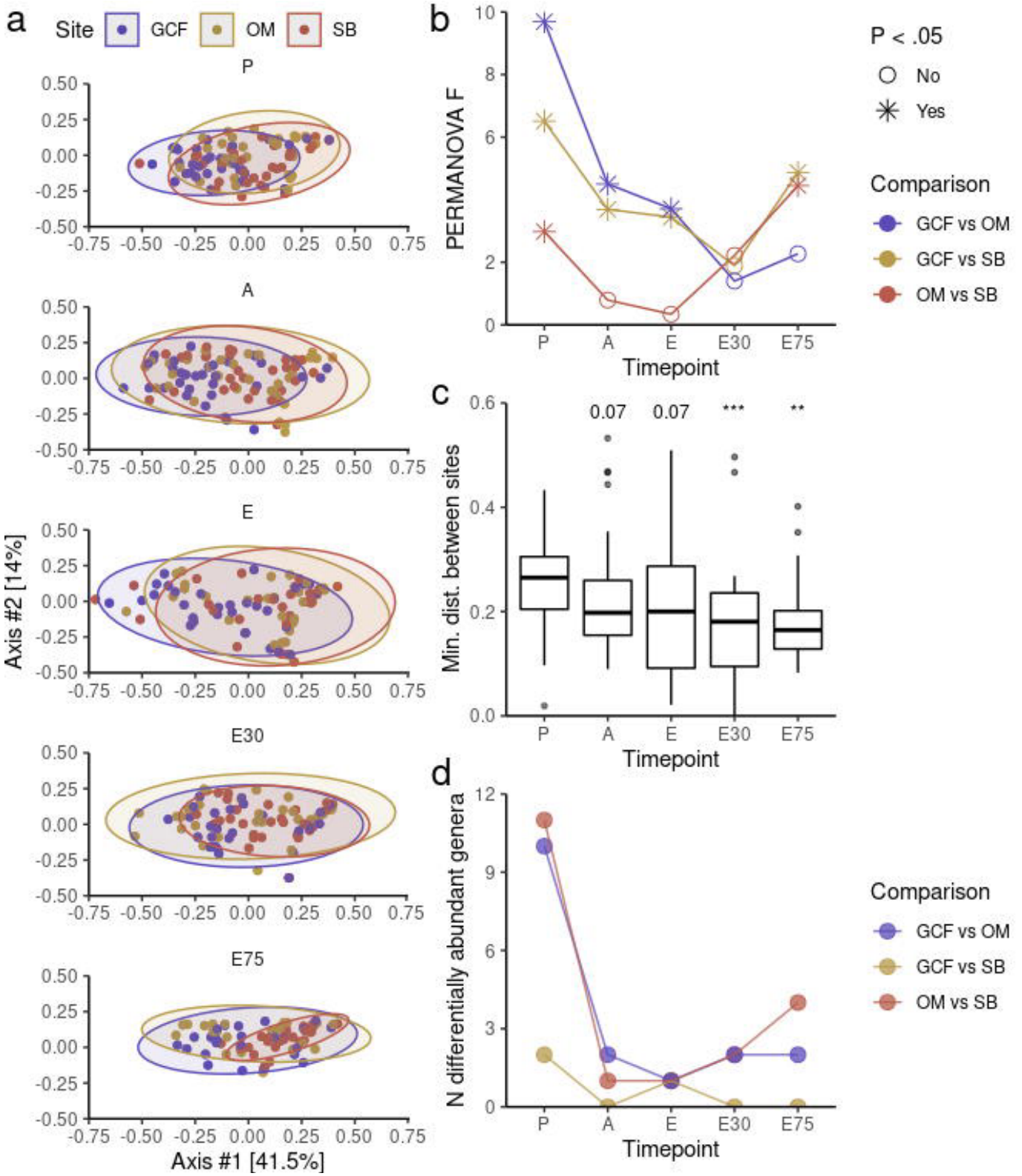
994 **Additional file 2: Supplementary tables and figures.**

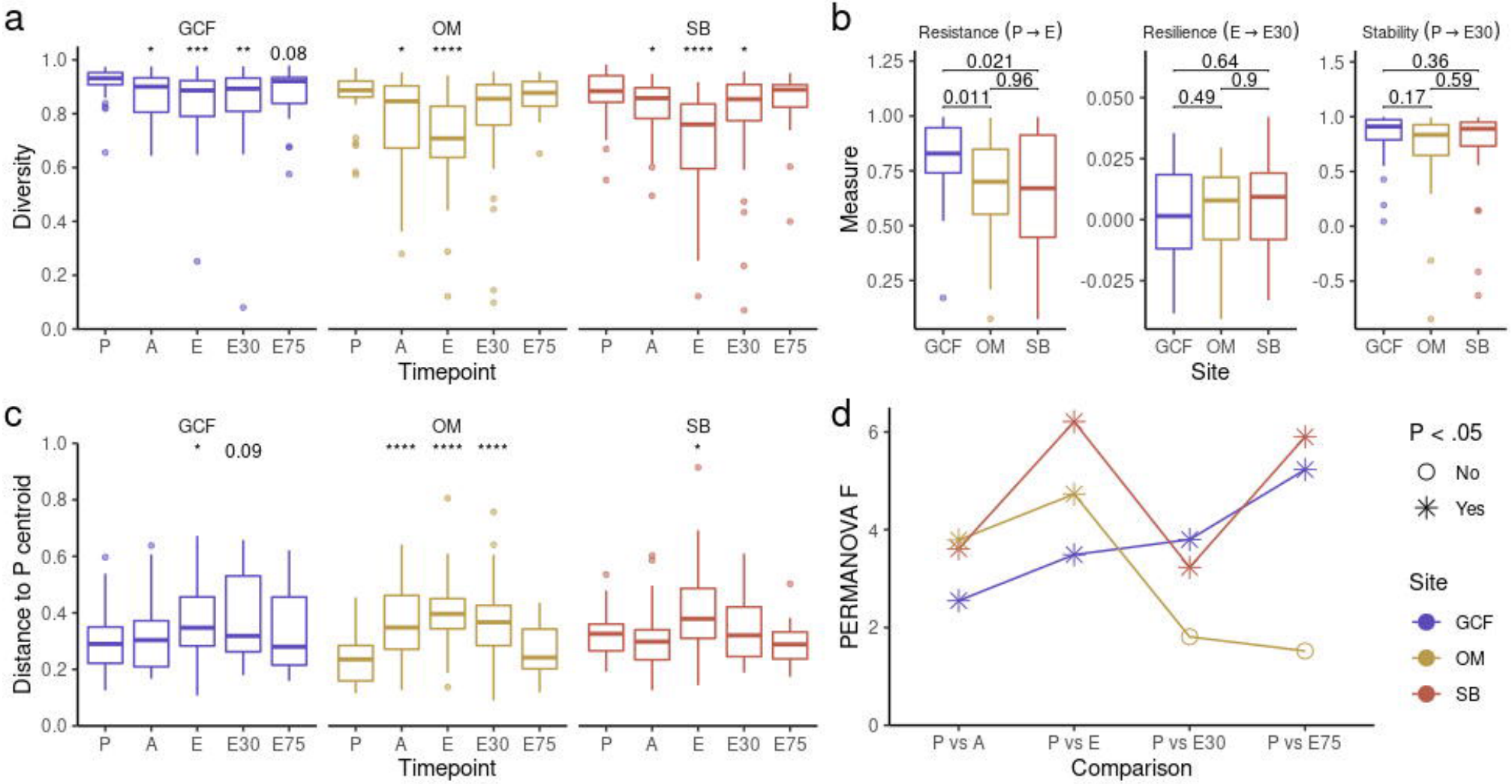
995 Supplementary material with 9 tables and 14 figures.

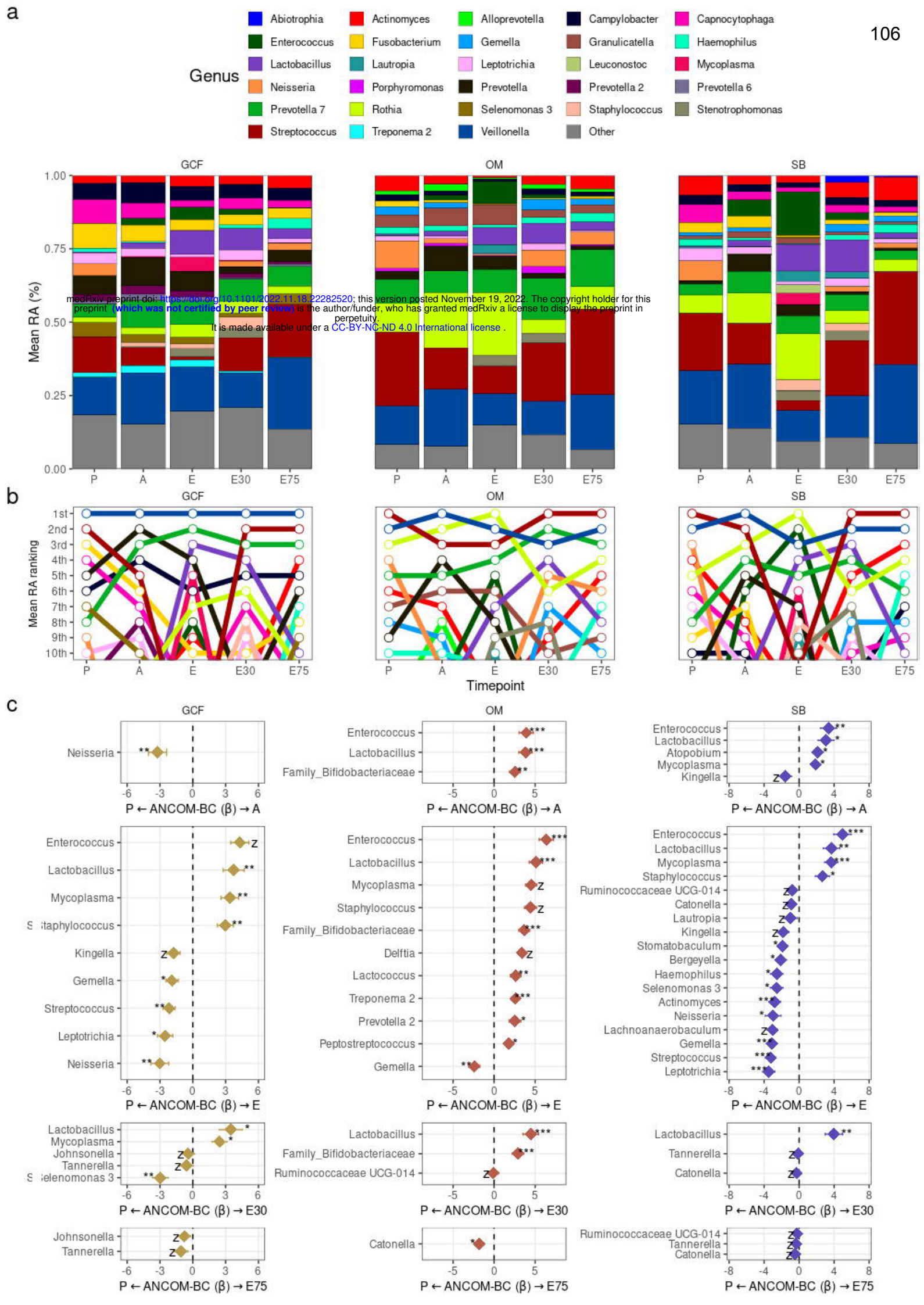
996 **Additional file 3: Supplementary methods.**

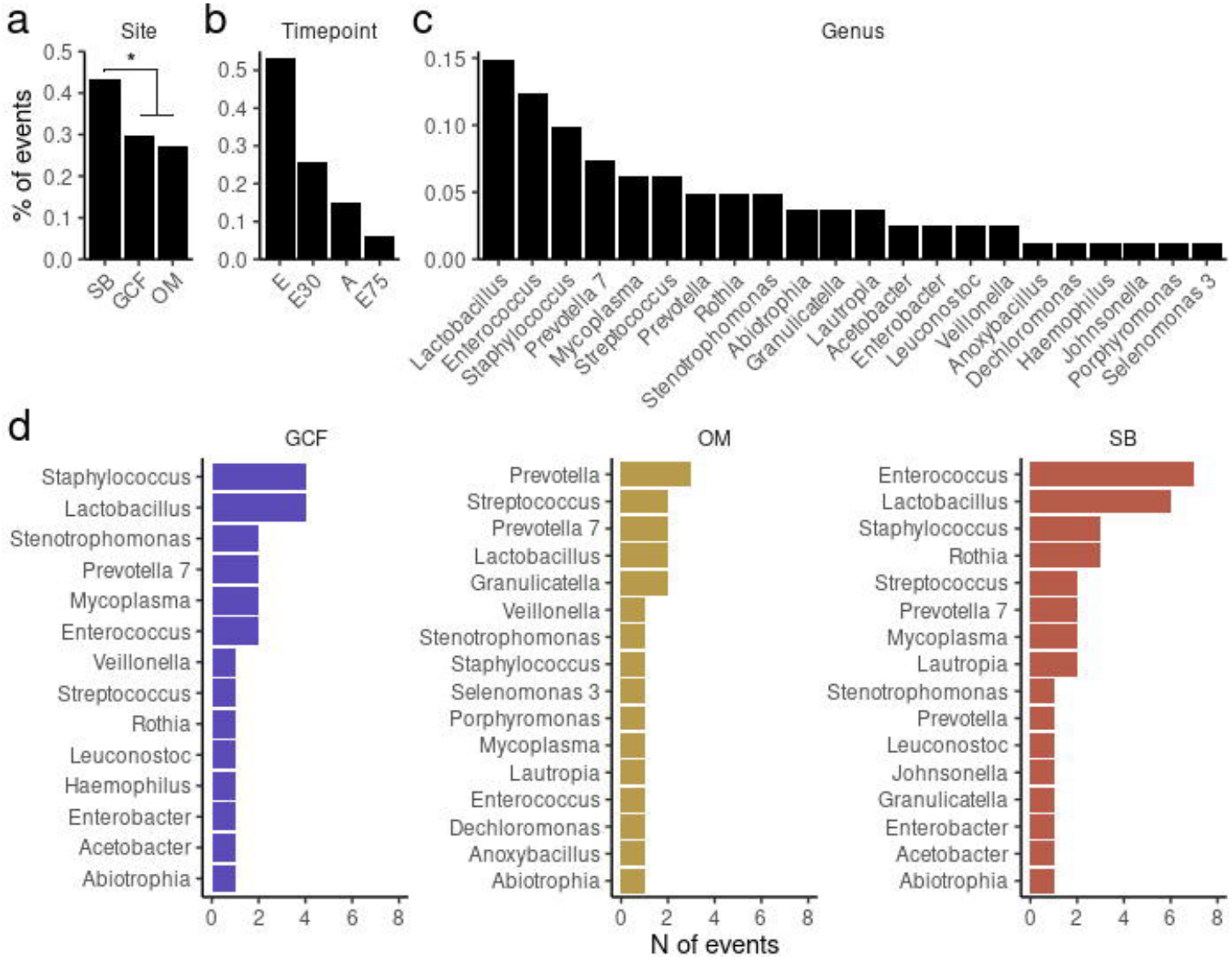
997 Supplementary text to the Materials and methods section.

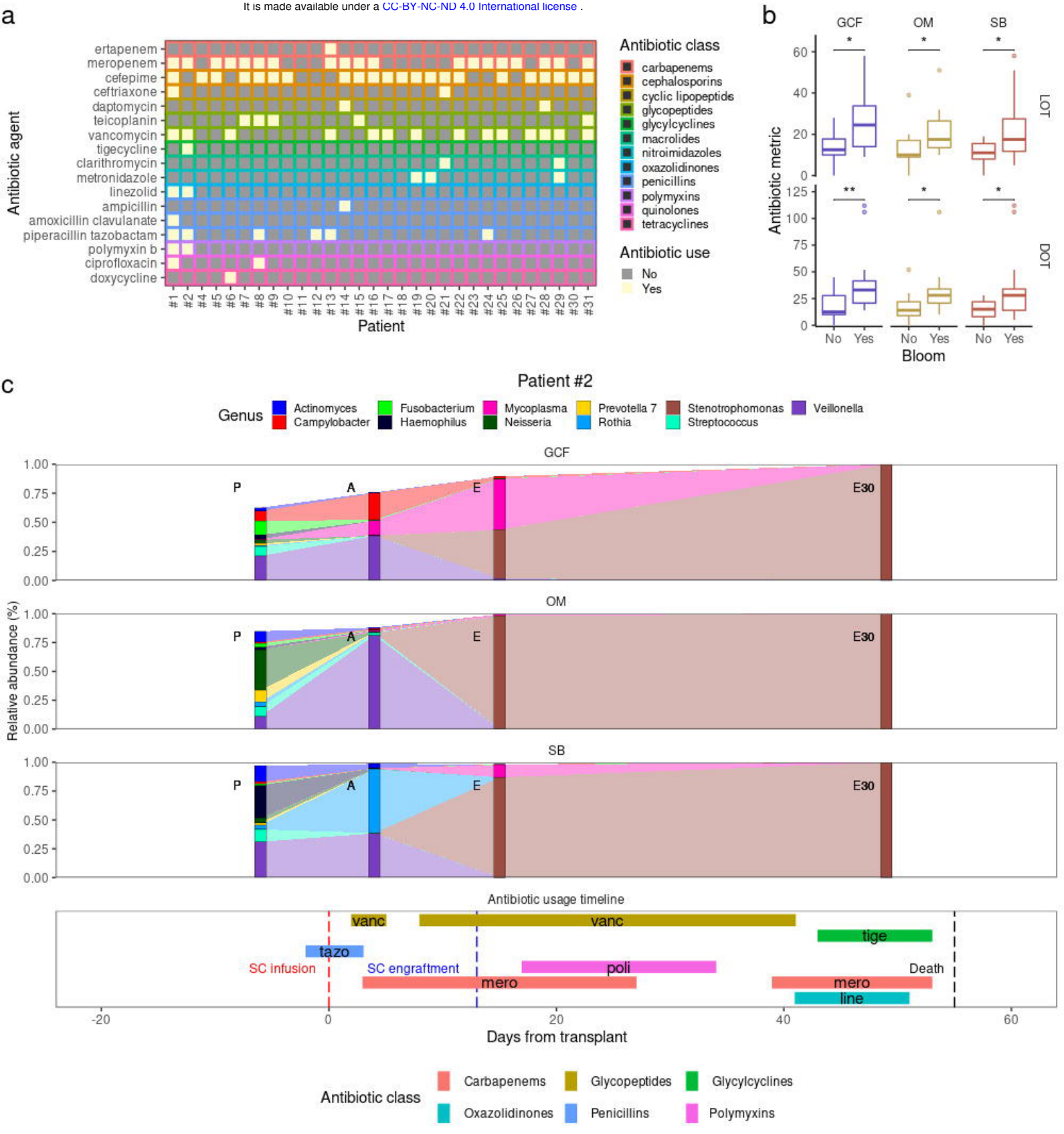
36

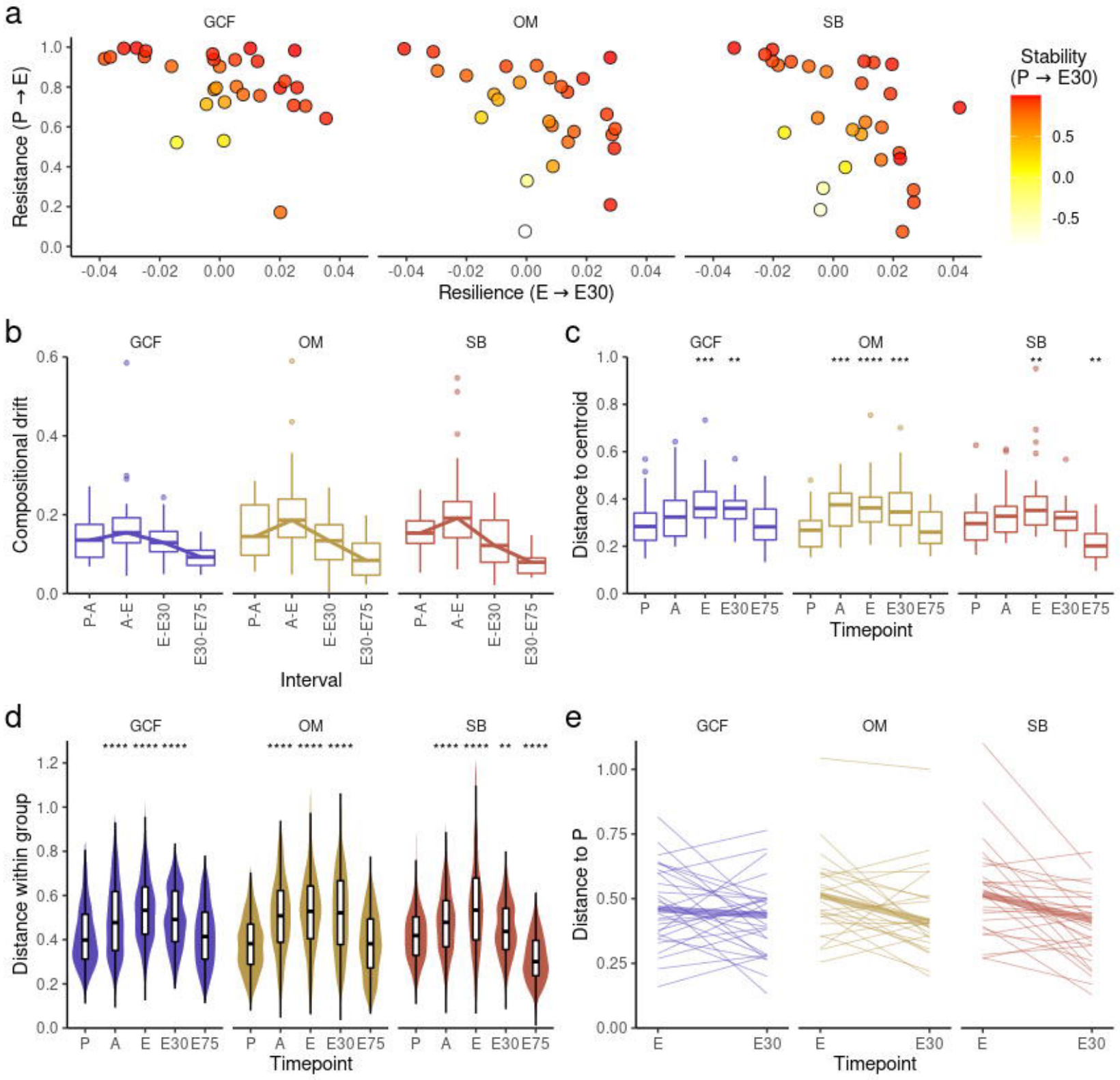

a
Site
GCF
OM
SB
P
A
E
E30
E75
Axis #2 [14%]
Axis #1 [41.5%]
b
PERMANOVA F
Timepoint
P < .05
No
Yes
Comparison
GCF vs OM
GCF vs SB
OM vs SB
c
Min. dist. between sites
0.07
0.07
***
**
Timepoint
d
N differentially abundant genera
Timepoint
Comparison
GCF vs OM
GCF vs SB
OM vs SB

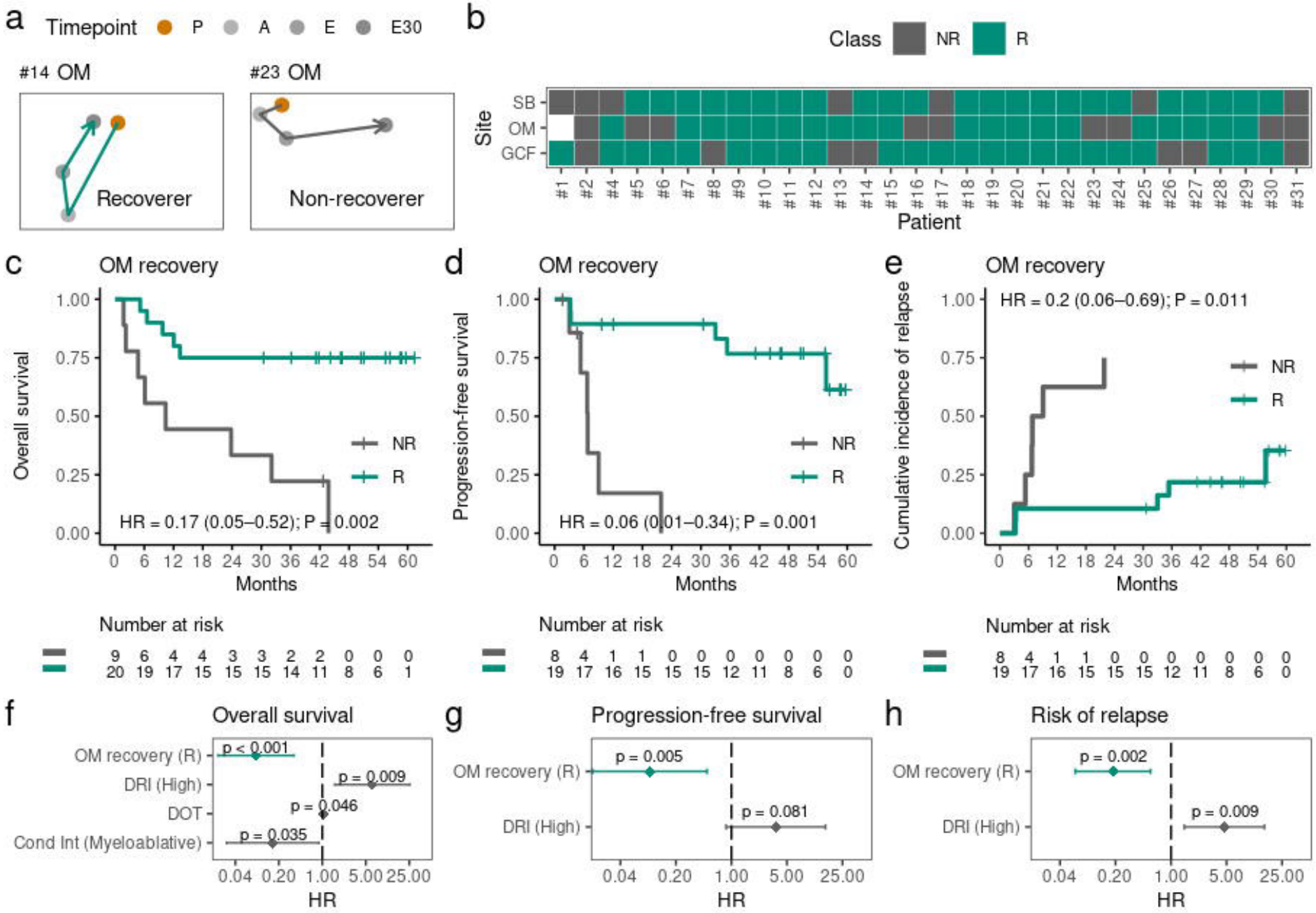

a
GCF
OM
SB
Diversity
Timepoint
P
A
E
E30
E75
0.08
b
Resistance (P → E)
Resilience (E → E30)
Stability (P → E30)
Measure
Site
GCF
OM
SB
0.021
0.96
0.011
0.64
0.9
0.49
0.36
0.59
0.17
c
Distance to P centroid
0.09
d
PERMANOVA F
Comparison
P vs A
P vs E
P vs E30
P vs E75
P < .05
No
Yes
Site
GCF
OM
SB

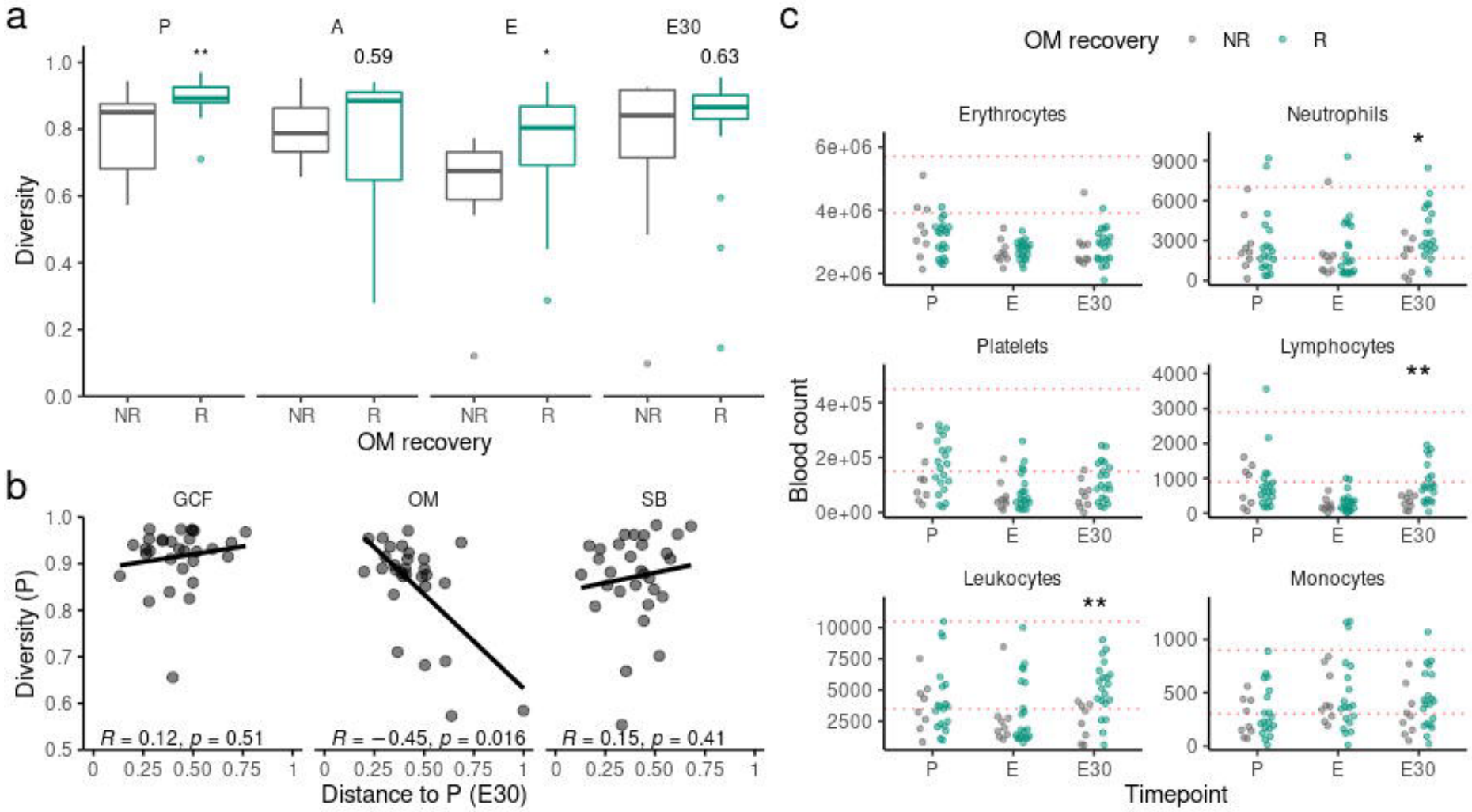



a
Site
% of events
0.5
0.4
0.3
0.2
0.1
0.0
*
SB
GCF
OM
b
Timepoint
0.5
0.4
0.3
0.2
0.1
0.0
E
E30
A
E75
c
Genus
0.15
0.10
0.05
0.00
Lactobacillus
Enterococcus
Staphylococcus
Prevotella 7
Mycoplasma
Streptococcus
Prevotella
Rothia
Stenotrophomonas
Abiotrophia
Granulicatella
Lautropia
Acetobacter
Enterobacter
Leuconostoc
Veillonella
Anoxybacillus
Dechloromonas
Haemophilus
Johnsonella
Porphyromonas
Selenomonas 3
d
GCF
Staphylococcus
Lactobacillus
Stenotrophomonas
Prevotella 7
Mycoplasma
Enterococcus
Veillonella
Streptococcus
Rothia
Leuconostoc
Haemophilus
Enterobacter
Acetobacter
Abiotrophia
0
2
4
6
8
OM
Prevotella
Streptococcus
Prevotella 7
Lactobacillus
Granulicatella
Veillonella
Stenotrophomonas
Staphylococcus
Selenomonas 3
Porphyromonas
Mycoplasma
Lautropia
Enterococcus
Dechloromonas
Anoxybacillus
Abiotrophia
0
2
4
6
8
N of events
SB
Enterococcus
Lactobacillus
Staphylococcus
Rothia
Streptococcus
Prevotella 7
Mycoplasma
Lautropia
Stenotrophomonas
Prevotella
Leuconostoc
Johnsonella
Granulicatella
Enterobacter
Acetobacter
Abiotrophia
0
2
4
6
8




a
GCF
OM
SB
Resistance (P → E)
Resilience (E → E30)
Stability
(P → E30)
b
Compositional drift
P-A
A-E
E-E30
E30-E75
Interval
c
Distance to centroid
P
A
E
E30
E75
Timepoint
d
Distance within group
e
Distance to P


a
Timepoint
P
A
E
E30
#14 OM
#23 OM
Recoverer
Non-recoverer
b
Class
NR
R
Site
SB
OM
GCF
Patient
c
OM recovery
Overall survival
HR = 0.17 (0.05–0.52); P = 0.002
Months
Number at risk
d
OM recovery
Progression-free survival
HR = 0.06 (0.01–0.34); P = 0.001
Months
Number at risk
e
OM recovery
Cumulative incidence of relapse
HR = 0.2 (0.06–0.69); P = 0.011
Months
Number at risk
f
Overall survival
OM recovery (R)
DRI (High)
DOT
Cond Int (Myeloablative)
p < 0.001
p = 0.009
p = 0.046
p = 0.035
HR
g
Progression-free survival
OM recovery (R)
DRI (High)
p = 0.005
p = 0.081
HR
h
Risk of relapse
OM recovery (R)
DRI (High)
p = 0.002
p = 0.009
HR


a
P
A
E
E30
**
0.59
*
0.63
Diversity
NR
R
OM recovery
b
GCF
OM
SB
Diversity (P)
R = 0.12, p = 0.51
R = −0.45, p = 0.016
R = 0.15, p = 0.41
Distance to P (E30)
c
OM recovery
NR
R
Erythrocytes
Neutrophils
Platelets
Lymphocytes
Leukocytes
Monocytes
Blood count
Timepoint

6. CONCLUSIONS

In this Thesis, we profiled the oral microbiota of allo-HSCT recipients during and after allo-HSCT (individualized clinical and microbiota data is provided in Attachment C). We leveraged the ease of sampling of the oral microbiota to profile bacterial composition changes at different oral compartments. We found that the microbiota of all oral sites was severely damaged during allo-HSCT. In particular, we found that this damage, which included loss of diversity and several blooms of potentially pathogenic species, made the microbiota at each oral site more similar to each other during allo-HSCT, despite the environmental richness of the oral cavity.

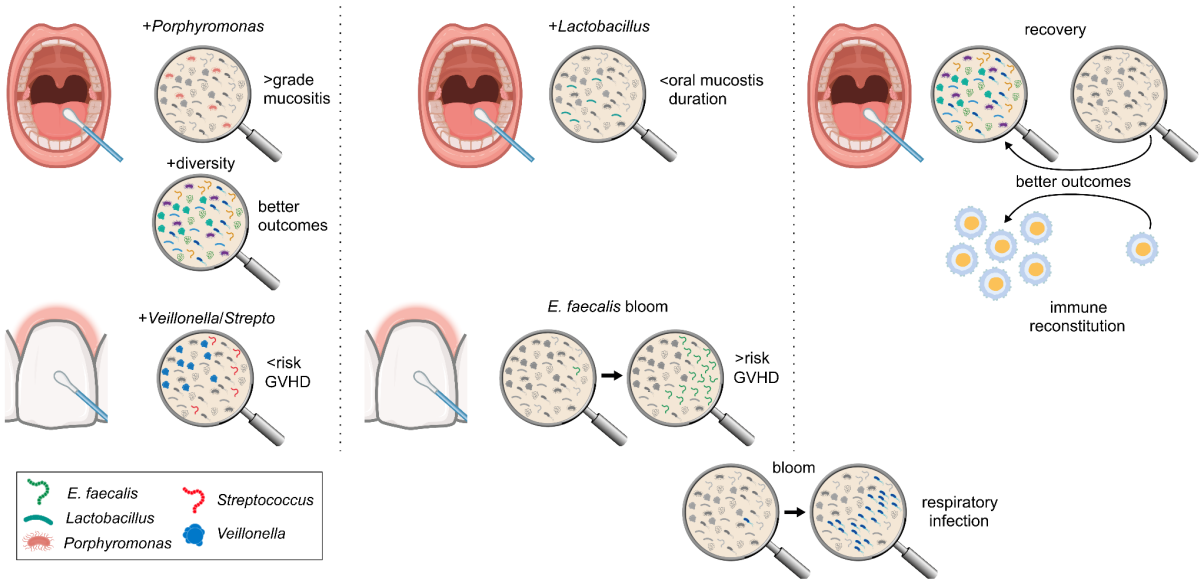


Fig. 3: Associations between oral microbiota and allo-HSCT found throughout this work. Non-exhaustive illustration of the associations found between the microbiota dynamics at different oral sites and the allo-HSCT clinical course.

The anatomically-aware analysis also allowed us to find site-specific associations with the allo-HSCT clinical course (Fig. 3), which is in line with the observation that the microbiota at each oral site can be associated with different conditions. For instance, although blooms were common in all oral sites during allo-HSCT, we found that only GCF blooms were associated with glycopeptide usage

in our cohort. In some cases, these blooms developed into respiratory infections, probably due to the translocation to the lungs of highly abundant oral microbes through aspiration.

Blooms were not only associated with infections but also with aGVHD. We found that *Enterococcus faecalis* SB blooms at A or E were associated with a higher risk of aGVHD and severe aGVHD. Preconditioning SB microbiota was also associated with aGVHD, as a high *Veillonella*/*Streptococcus* ratio at the preconditioning SB microbiota showed to be protective against aGVHD.

We showed that preconditioning OM microbiota diversity and composition might as well be clinically informative, as we observed a clear association between OM microbiota dysbiosis and poorer allo-HSCT outcomes. We also identified a preconditioning OM genera signature that may be used to predict the risk of developing ulcerative oral mucositis after the conditioning regimen. Taxonomic analysis further revealed that higher OM preconditioning levels of *Solobacterium* were associated with a lower risk of underlying disease relapse. In contrast, higher levels of *Porphyromonas* were associated with higher ulcerative oral mucositis grade.

We observed an overall microbiota recovery to the preconditioning state after allo-HSCT, but patient-level analysis revealed this recovery was only partial for some patients. After stratifying patients based on their ability to recover their preconditioning microbiota, we found that OM recovery was not associated with antibiotic usage but was associated with higher OM preconditioning diversity and earlier reconstitution of normal leukocyte counts. Most notably, OM recovery was associated with excellent outcomes, likely due to a more robust immune reconstitution in OM recoverers following allo-HSCT.

Despite the numerous and clinically relevant associations regarding different aspects of the allo-HSCT clinical course described in this Thesis, our studies have many limitations. The cohort analyzed is small, single-centered, and heterogeneous, encompassing several underlying diseases. Therefore, validation cohorts with greater sample sizes in multicentric prospective studies are needed to evaluate the generalizability and applicability of our findings. Such studies should ideally be performed using shotgun metagenomic sequencing, as this would enable a greater taxonomic resolution than the genus-level data produced here through 16S rRNA amplicon sequencing. Furthermore, we emphasize that, as expected from an exploratory work, the associations described here are correlative, so the causal role of the oral microbiota in the allo-HSCT clinical course remains to be elucidated. Nevertheless, we were able to show the potential of the oral microbiota to provide biomarkers in the allo-HSCT setting. We also provided a simple and reproducible protocol to evaluate the risk of allo-HSCT complications and outcomes based on a non-invasive scalable technique. Overall, we believe the strengths of this pioneer work, including the longitudinal design, assessment of different oral sites, and evaluation of a Brazilian cohort with extensive metadata publicly available, largely surpass its limitations.

In summary, we showed that the previously reported associations between gut microbiota and allo-HSCT systemic outcomes and complications broadly extend to the oral microbiota. In addition, we described associations with local (oral mucositis) and distal conditions (respiratory infections) that could only be analyzed by looking into oral microbiotas. These results and the easier access to oral samples indicate more attention should be given to the oral microbiota dynamics in patients undergoing allo-HSCT. Tracking oral microbiota injury and recovery in the allo-HSCT

setting may improve our understanding of allo-HSCT clinical course and help deliver a safer and more effective treatment for allo-HSCT recipients.

7. CONTRIBUTIONS

I contributed to several aspects of the work presented in this Thesis. As I joined the Doctorate Program in August 2018, most samples (collected between January 2016 and July 2018) had been previously processed and sequenced by co-authors of the chapters included in this Thesis (mainly, FHK and PFA). Still, I extracted DNA from ~10% and prepared sequencing libraries for ~20% of the oral samples analyzed. I collected antibiotic usage and blood count data from clinical records and submitted raw sequencing data and associated metadata to public databases. I was the sole responsible for the development and implementation of the bioinformatics pipeline used and the statistical analyses performed. More importantly, I was the main contributor to all the analytical decisions and the interpretation of the results. Furthermore, I contributed greatly to writing the chapters included in this Thesis. In detail, I single-handedly wrote all sections of chapters 2 and 4, which were later revised and approved by all co-authors. Besides, I wrote most of the Results and Methods sections of chapters 1 and 3. I also contributed to revising and editing sections written by co-authors (JSB and VCM). The full list of publications to which I contributed (including publications not included as chapters here) is available in Attachment D. Overall, I authored 10 publications during my Doctorate studies, including 7 first-author publications.

APPENDICES

Appendix A: Chapter 1 supplementary tables and figures

Supplementary material to Chapter 1 with 6 tables and 1 figure.

SUPPLEMENTAL MATERIAL

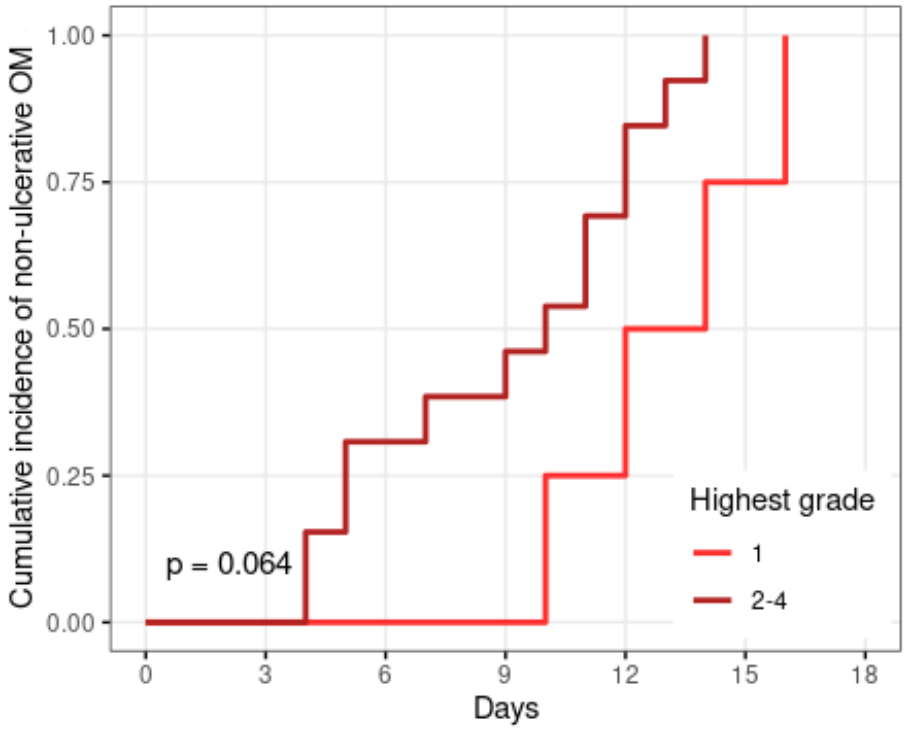


Figure S1: Cumulative incidence curves of non-ulcerative OM with patients stratified by the highest OM grade during OM clinical course (1/2-4). Statistical significance was evaluated by the log-rank test, with P-value indicated.

	n = 30
Age in years (median, range)	50 (19-73)
Sex (male)	16 (53%)
Oral Mucosits Grades (WHO)	
Grade 0	12 (40%)
Grade I	5 (17%)
Grade II	4 (13%)
Grade III	6 (20%)
Grade IV	3 (10%)
Conditioning Intensity	
Reduced Intensity	18 (60%)
Total Body Irradiation	11 (37%)
Alkylating Agents	
Busulfan	15 (50%)
Cyclophosphamide	10 (33%)
Melphalan	4 (13%)
Bendamustine	1 (3%)
Treosulfan	1 (3%)
Graft Source	
Bone Marrow	10 (33%)
Peripheal Blood	20 (67%)
Underlying Disease	
Acute Myeloid Leukemia	18 (60%)
Acute Lymbhoblastic Leukemia	7 (23%)
Non-hodgkin Lymphoma	5 (17%)
Myelodysplastic Syndrome	4 (13%)
Chronic Lymphocytic Leukemia	1 (3%)
Chronic Myeloid Leukemia	2 (3%)
Multiple Myeloma	3 (3%)
Graft-versus-Host Disease Prophylaxis	
Mycophenolate+Cyclosporin A	11 (37%)
Methotrexate+Cysclosporin A	10 (33%)
Mycophenolate+Cyclosporin A+Cyclophosphamide	9 (30%)

Table S1: Clinical characteristics of study patients.

Genus	log(FoldChange)	P-value	Adjusted P-value	Differentially abundant
D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Actinomycetales;D_4__Actinomycetaceae;D_5__Actinomyces	-0,59	0,46	1	No
D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Actinomycetales;D_4__Actinomycetaceae;D_5__F0332	0,79	0,00	0	No
D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Bifidobacteriales;D_4__Bifidobacteriaceae;D_5__Bifidobacterium	-0,85	0,00	0	No
D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Bifidobacteriales;D_4__Bifidobacteriaceae;__	0,59	0,00	0	No
D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Micrococcales;D_4__Micrococcaceae;D_5__Rothia	-0,24	0,76	1	No
D_0__Bacteria;D_1__Actinobacteria;D_2__Coriobacteriia;D_3__Coriobacteriales;D_4__Atopobiaceae;D_5__Atopobium	-0,29	0,74	1	No
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Porphyromonadaceae;D_5__Porphyromonas	-0,70	0,54	1	No
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Prevotellaceae;D_5__Alloprevotella	-0,07	0,95	1	No
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Prevotellaceae;D_5__Prevotella	0,85	0,28	1	No
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Prevotellaceae;D_5__Prevotella 2	-0,33	0,72	1	No
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Prevotellaceae;D_5__Prevotella 6	0,51	0,62	1	No
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Prevotellaceae;D_5__Prevotella 7	-0,02	0,98	1	No
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Tannerellaceae;D_5__Tannerella	1,19	0,00	0	No
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Flavobacteriales;D_4__Flavobacteriaceae;D_5__Capnocytophaga	1,00	0,33	1	No
D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Flavobacteriales;D_4__Weeksellaceae;D_5__Bergeyella	0,56	0,60	1	No
D_0__Bacteria;D_1__Epsilonbacteraeota;D_2__Campylobacteria;D_3__Campylobacterales;D_4__Campylobacteraceae;D_5__Campylobacter	0,38	0,60	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Bacillales;D_4__Family XI;D_5__Gemella	-0,31	0,75	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Carnobacteriaceae;D_5__Granulicatella	-0,95	0,37	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Lactobacillaceae;D_5__Lactobacillus	-0,17	0,85	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Streptococcaceae;D_5__Lactococcus	1,38	0,00	0	No
D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Streptococcaceae;D_5__Streptococcus	-0,70	0,40	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Family XI;D_5__Parvimonas	-0,71	0,31	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Family XIII;D_5__Mogibacterium	0,96	0,00	0	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Family XIII;D_5__[Eubacterium] nodatum group	0,16	0,83	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Lachnospiraceae;D_5__Butyrivibrio 2	-0,21	0,77	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Lachnospiraceae;D_5__Catonella	0,33	0,66	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Lachnospiraceae;D_5__Lachnoanaerobaculum	1,09	0,21	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Lachnospiraceae;D_5__Oribacterium	0,55	0,42	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Lachnospiraceae;D_5__Stomatobaculum	0,46	0,57	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Lachnospiraceae;__	0,06	0,95	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Peptostreptococcaceae;D_5__Peptostreptococcus	0,14	0,81	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Ruminococcaceae;D_5__Ruminococcaceae UCG-014	-1,04	0,27	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Erysipelotrichia;D_3__Erysipelotrichales;D_4__Erysipelotrichaceae;D_5__Solobacterium	0,28	0,69	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Negativicutes;D_3__Selenomonadales;D_4__Veillonellaceae;D_5__Dialister	0,67	0,00	0	No
D_0__Bacteria;D_1__Firmicutes;D_2__Negativicutes;D_3__Selenomonadales;D_4__Veillonellaceae;D_5__Megasphaera	-0,29	0,78	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Negativicutes;D_3__Selenomonadales;D_4__Veillonellaceae;D_5__Selenomonas	0,16	0,84	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Negativicutes;D_3__Selenomonadales;D_4__Veillonellaceae;D_5__Selenomonas 3	0,71	0,37	1	No
D_0__Bacteria;D_1__Firmicutes;D_2__Negativicutes;D_3__Selenomonadales;D_4__Veillonellaceae;D_5__Veillonella	-0,26	0,71	1	No
D_0__Bacteria;D_1__Fusobacteria;D_2__Fusobacteriia;D_3__Fusobacteriales;D_4__Fusobacteriaceae;D_5__Fusobacterium	0,19	0,83	1	No
D_0__Bacteria;D_1__Fusobacteria;D_2__Fusobacteriia;D_3__Fusobacteriales;D_4__Leptotrichiaceae;D_5__Leptotrichia	0,25	0,76	1	No
D_0__Bacteria;D_1__Patescibacteria;D_2__Saccharimonadia;D_3__Saccharimonadales;D_4__Saccharimonadaceae;D_5__Candidatus Saccharimonas	1,76	0,00	0	No
D_0__Bacteria;D_1__Patescibacteria;D_2__Saccharimonadia;D_3__Saccharimonadales;D_4__Saccharimonadaceae;D_5__uncultured bacterium	-0,65	0,53	1	No
D_0__Bacteria;D_1__Patescibacteria;D_2__Saccharimonadia;D_3__Saccharimonadales;D_4__uncultured bacterium;D_5__uncultured bacterium	0,89	0,00	0	No
D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Betaproteobacteriales;D_4__Burkholderiaceae;D_5__Lautropia	0,60	0,42	1	No
D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Betaproteobacteriales;D_4__Neisseriaceae;D_5__Kingella	1,10	0,00	0	No
D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Betaproteobacteriales;D_4__Neisseriaceae;D_5__Neisseria	-0,64	0,68	1	No
D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Cardiobacteriales;D_4__Cardiobacteriaceae;D_5__Cardiobacterium	-0,48	0,40	1	No
D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Pasteurellales;D_4__Pasteurellaceae;D_5__Haemophilus	-0,59	0,61	1	No
D_0__Bacteria;D_1__Spirochaetes;D_2__Spirochaetia;D_3__Spirochaetales;D_4__Spirochaetaceae;D_5__Treponema 2	-0,33	0,64	1	No

Table S2: Differential abundance analysis between oral mucositis (OM) and OM-free patients at preconditioning. The analysis was performed with ANCOM-BC at genus level. No genera passed the log(FoldChange) and adjusted P-value cutoffs to be considered as significantly differentially abundant between groups.

Genus	P-value	HR	HR (2.5 percentile)	HR (97.5 percentile)	% Group1	% Group0
D_0__Bacteria.D_1__Actinobacteria.D_2__Actinobacteria.D_3__Actinomycetales.D_4__Actinomycetaceae.D_5__Actinomyces	0,56	1,38	0,46	4,12	50	40
D_0__Bacteria.D_1__Actinobacteria.D_2__Actinobacteria.D_3__Micrococcales.D_4__Micrococcaceae.D_5__Rothia	0,23	0,51	0,17	1,55	36	53
D_0__Bacteria.D_1__Actinobacteria.D_2__Coriobacteriia.D_3__Coriobacteriales.D_4__Atopobiaceae.D_5__Atopobium	0,53	1,41	0,47	4,22	50	40
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Porphyromonadaceae.D_5__Porphyromonas	0,95	1,04	0,35	3,08	43	47
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Alloprevotella	0,48	0,67	0,22	2,05	36	53
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella	0,56	1,38	0,46	4,13	50	40
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella.6	0,53	1,41	0,47	4,22	50	40
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella.7	0,93	0,95	0,32	2,83	43	47
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Flavobacteriales.D_4__Flavobacteriaceae.D_5__Capnocytophaga	0,27	1,85	0,6	5,68	57	33
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Flavobacteriales.D_4__Weeksellaceae.D_5__Bergeyella	0,59	1,35	0,45	4,02	50	40
D_0__Bacteria.D_1__Epsilonbacteraeota.D_2__Campylobacteria.D_3__Campylobacterales.D_4__Campylobacteraceae.D_5__Campylobacter	0,46	1,51	0,51	4,5	50	40
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Bacillales.D_4__Family.XI.D_5__Gemella	0,27	0,54	0,18	1,64	36	53
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Carnobacteriaceae.D_5__Granulicatella	0,08	0,36	0,11	1,17	29	60
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Lactobacillaceae.D_5__Lactobacillus	0,21	2,01	0,65	6,16	57	33
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Streptococcaceae.D_5__Streptococcus	0,21	0,49	0,16	1,52	36	53
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Catonella	0,74	1,2	0,4	3,58	50	40
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Lachnoanaerobaculum	0,61	1,32	0,44	3,95	50	40
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Oribacterium	0,69	1,25	0,42	3,72	50	40
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Stomatobaculum	0,74	1,2	0,4	3,58	50	40
D_0__Bacteria.D_1__Firmicutes.D_2__Erysipelotrichia.D_3__Erysipelotrichales.D_4__Erysipelotrichaceae.D_5__Solobacterium	0,89	0,93	0,31	2,77	43	47
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Megasphaera	0,39	0,62	0,2	1,88	36	53
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Selenomonas.3	0,50	1,46	0,49	4,36	50	40
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Veillonella	0,55	1,4	0,47	4,16	50	40
D_0__Bacteria.D_1__Fusobacteria.D_2__Fusobacteriia.D_3__Fusobacteriales.D_4__Fusobacteriaceae.D_5__Fusobacterium	0,85	0,9	0,3	2,67	43	47
D_0__Bacteria.D_1__Fusobacteria.D_2__Fusobacteriia.D_3__Fusobacteriales.D_4__Leptotrichiaceae.D_5__Leptotrichia	0,86	0,91	0,31	2,71	43	47
D_0__Bacteria.D_1__Patescibacteria.D_2__Saccharimonadia.D_3__Saccharimonadales.D_4__Saccharimonadaceae.D_5__uncultured.bacterium	0,41	0,63	0,21	1,93	36	53
D_0__Bacteria.D_1__Proteobacteria.D_2__Gammaproteobacteria.D_3__Betaproteobacteriales.D_4__Neisseriaceae.D_5__Neisseria	0,30	0,56	0,18	1,7	36	53
D_0__Bacteria.D_1__Proteobacteria.D_2__Gammaproteobacteria.D_3__Pasteurellales.D_4__Pasteurellaceae.D_5__Haemophilus	0,42	0,63	0,21	1,94	36	53

Table S3: Association between genera relative abundance at preconditioning and oral mucositis (OM) risk. Groups were stratified based on median genus relative abundance and association with the risk of OM was assessed by estimating the Cox proportional hazards between groups. HR, Hazard Ratio; % Group1, percentage of patients classified as having high genus relative abundance at preconditioning that developed OM; % Group0, percentage of patients classified as having low genus relative abundance at preconditioning that developed OM.

Genus	P-value	Spearman rho
D_0__Bacteria.D_1__Actinobacteria.D_2__Actinobacteria.D_3__Actinomycetales.D_4__Actinomycetaceae.D_5__Actinomyces	0,57	0,11
D_0__Bacteria.D_1__Actinobacteria.D_2__Actinobacteria.D_3__Micrococcales.D_4__Micrococcaceae.D_5__Rothia	0,11	-0,30
D_0__Bacteria.D_1__Actinobacteria.D_2__Coriobacteriia.D_3__Coriobacteriales.D_4__Atopobiaceae.D_5__Atopobium	0,86	-0,03
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Porphyromonadaceae.D_5__Porphyromonas	0,96	0,01
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Alloprevotella	0,78	0,05
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella	0,15	0,27
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella.6	0,45	0,15
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella.7	0,82	-0,04
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Flavobacteriales.D_4__Flavobacteriaceae.D_5__Capnocytophaga	0,37	0,17
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Flavobacteriales.D_4__Weeksellaceae.D_5__Bergeyella	0,46	0,14
D_0__Bacteria.D_1__Epsilonbacteraeota.D_2__Campylobacteria.D_3__Campylobacterales.D_4__Campylobacteraceae.D_5__Campylobacter	0,95	-0,01
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Bacillales.D_4__Family.XI.D_5__Gemella	0,71	-0,07
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Carnobacteriaceae.D_5__Granulicatella	0,76	-0,06
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Lactobacillaceae.D_5__Lactobacillus	0,95	0,01
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Streptococcaceae.D_5__Streptococcus	0,06	-0,35
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Catonella	0,64	0,09
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Lachnoanaerobaculum	0,74	0,06
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Oribacterium	0,61	0,10
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Stomatobaculum	0,58	0,11
D_0__Bacteria.D_1__Firmicutes.D_2__Erysipelotrichia.D_3__Erysipelotrichales.D_4__Erysipelotrichaceae.D_5__Solobacterium	0,25	0,22
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Megasphaera	0,73	-0,07
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Selenomonas.3	0,90	-0,03
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Veillonella	0,79	0,05
D_0__Bacteria.D_1__Fusobacteria.D_2__Fusobacteriia.D_3__Fusobacteriales.D_4__Fusobacteriaceae.D_5__Fusobacterium	0,76	-0,06
D_0__Bacteria.D_1__Fusobacteria.D_2__Fusobacteriia.D_3__Fusobacteriales.D_4__Leptotrichiaceae.D_5__Leptotrichia	0,99	0,00
D_0__Bacteria.D_1__Patescibacteria.D_2__Saccharimonadia.D_3__Saccharimonadales.D_4__Saccharimonadaceae.D_5__uncultured.bacterium	0,94	0,02
D_0__Bacteria.D_1__Proteobacteria.D_2__Gammaproteobacteria.D_3__Betaproteobacteriales.D_4__Neisseriaceae.D_5__Neisseria	0,60	0,10
D_0__Bacteria.D_1__Proteobacteria.D_2__Gammaproteobacteria.D_3__Pasteurellales.D_4__Pasteurellaceae.D_5__Haemophilus	0,54	-0,12

Table S4: Correlation between genera relative abundance at preconditioning and ulcerative oral mucositis (OM) grade. All OM grades (0-4) are considered in this analysis. Spearman correlation was used.

Genus	P-value	Spearman rho
D_0__Bacteria.D_1__Actinobacteria.D_2__Actinobacteria.D_3__Actinomycetales.D_4__Actinomycetaceae.D_5__Actinomyces	0,69	0,12
D_0__Bacteria.D_1__Actinobacteria.D_2__Actinobacteria.D_3__Micrococcales.D_4__Micrococcaceae.D_5__Rothia	0,67	-0,13
D_0__Bacteria.D_1__Actinobacteria.D_2__Coriobacteriia.D_3__Coriobacteriales.D_4__Atopobiaceae.D_5__Atopobium	0,89	-0,04
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Porphyromonadaceae.D_5__Porphyromonas	0,03	0,61
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Alloprevotella	0,07	0,51
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella	0,59	0,17
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella.6	0,66	-0,14
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella.7	0,47	-0,22
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Flavobacteriales.D_4__Flavobacteriaceae.D_5__Capnocytophaga	0,46	0,22
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Flavobacteriales.D_4__Weeksellaceae.D_5__Bergeyella	0,22	0,37
D_0__Bacteria.D_1__Epsilonbacteraeota.D_2__Campylobacteria.D_3__Campylobacterales.D_4__Campylobacteraceae.D_5__Campylobacter	0,37	-0,27
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Bacillales.D_4__Family.XI.D_5__Gemella	0,82	-0,07
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Carnobacteriaceae.D_5__Granulicatella	0,44	0,23
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Lactobacillaceae.D_5__Lactobacillus	0,31	-0,31
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Streptococcaceae.D_5__Streptococcus	0,52	-0,20
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Catonella	0,65	-0,14
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Lachnoanaerobaculum	0,40	-0,26
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Oribacterium	0,47	0,22
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Stomatobaculum	0,67	0,13
D_0__Bacteria.D_1__Firmicutes.D_2__Erysipelotrichia.D_3__Erysipelotrichales.D_4__Erysipelotrichaceae.D_5__Solobacterium	0,35	0,29
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Megasphaera	0,26	-0,34
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Selenomonas.3	0,20	-0,38
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Veillonella	0,21	-0,37
D_0__Bacteria.D_1__Fusobacteria.D_2__Fusobacteriia.D_3__Fusobacteriales.D_4__Fusobacteriaceae.D_5__Fusobacterium	0,67	0,13
D_0__Bacteria.D_1__Fusobacteria.D_2__Fusobacteriia.D_3__Fusobacteriales.D_4__Leptotrichiaceae.D_5__Leptotrichia	0,51	-0,20
D_0__Bacteria.D_1__Patescibacteria.D_2__Saccharimonadia.D_3__Saccharimonadales.D_4__Saccharimonadaceae.D_5__uncultured.bacterium	0,95	-0,02
D_0__Bacteria.D_1__Proteobacteria.D_2__Gammaproteobacteria.D_3__Betaproteobacteriales.D_4__Neisseriaceae.D_5__Neisseria	0,19	0,39
D_0__Bacteria.D_1__Proteobacteria.D_2__Gammaproteobacteria.D_3__Pasteurellales.D_4__Pasteurellaceae.D_5__Haemophilus	0,23	-0,35

Table S5: Correlation between genera relative abundance at preconditioning and ulcerative oral mucositis (OM) grade. Only ulcerative OM grades (2-4) are considered in this analysis. Spearman correlation was used.

Genus	P-value	HR	HR (2.5 percentile)	HR (97.5 percentile)
D_0__Bacteria.D_1__Actinobacteria.D_2__Actinobacteria.D_3__Actinomycetales.D_4__Actinomycetaceae.D_5__Actinomyces	0,42	1,59	0,51	5,01
D_0__Bacteria.D_1__Actinobacteria.D_2__Actinobacteria.D_3__Bifidobacteriales.D_4__Bifidobacteriaceae.__	0,16	2,47	0,67	9,01
D_0__Bacteria.D_1__Actinobacteria.D_2__Actinobacteria.D_3__Micrococcales.D_4__Micrococcaceae.D_5__Rothia	0,94	1,04	0,33	3,34
D_0__Bacteria.D_1__Actinobacteria.D_2__Coriobacteriia.D_3__Coriobacteriales.D_4__Atopobiaceae.D_5__Atopobium	0,95	1,04	0,33	3,25
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Porphyromonadaceae.D_5__Porphyromonas	0,07	0,3	0,08	1,2
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Alloprevotella	0,86	1,1	0,35	3,45
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella	0,88	1,09	0,35	3,43
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella.6	0,28	2,03	0,54	7,55
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Bacteroidales.D_4__Prevotellaceae.D_5__Prevotella.7	0,77	0,84	0,27	2,66
D_0__Bacteria.D_1__Bacteroidetes.D_2__Bacteroidia.D_3__Flavobacteriales.D_4__Flavobacteriaceae.D_5__Capnocytophaga	0,84	0,89	0,28	2,81
D_0__Bacteria.D_1__Epsilonbacteraeota.D_2__Campylobacteria.D_3__Campylobacterales.D_4__Campylobacteraceae.D_5__Campylobacter	0,98	0,99	0,32	3,08
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Bacillales.D_4__Family.XI.D_5__Gemella	0,08	0,3	0,08	1,24
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Carnobacteriaceae.D_5__Granulicatella	0,96	1,03	0,33	3,23
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Lactobacillaceae.D_5__Lactobacillus	0,04	4,99	0,96	26,04
D_0__Bacteria.D_1__Firmicutes.D_2__Bacilli.D_3__Lactobacillales.D_4__Streptococcaceae.D_5__Streptococcus	0,15	0,4	0,11	1,42
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Lachnoanaerobaculum	0,16	0,38	0,1	1,51
D_0__Bacteria.D_1__Firmicutes.D_2__Clostridia.D_3__Clostridiales.D_4__Lachnospiraceae.D_5__Oribacterium	0,91	1,07	0,34	3,35
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Selenomonas.3	0,16	2,47	0,67	9,01
D_0__Bacteria.D_1__Firmicutes.D_2__Negativicutes.D_3__Selenomonadales.D_4__Veillonellaceae.D_5__Veillonella	0,10	0,33	0,08	1,3
D_0__Bacteria.D_1__Fusobacteria.D_2__Fusobacteriia.D_3__Fusobacteriales.D_4__Fusobacteriaceae.D_5__Fusobacterium	0,39	0,6	0,19	1,92
D_0__Bacteria.D_1__Fusobacteria.D_2__Fusobacteriia.D_3__Fusobacteriales.D_4__Leptotrichiaceae.D_5__Leptotrichia	0,97	1,02	0,33	3,2

Table S6: Association between genera relative abundance at oral mucositis (OM) onset and time to OM healing. Groups were stratified based on median genus relative abundance and association with time to OM healing was assessed by estimating the Cox proportional hazards between groups. HR, Hazard Ratio.

Appendix B: Chapter 2 supplementary tables and figures

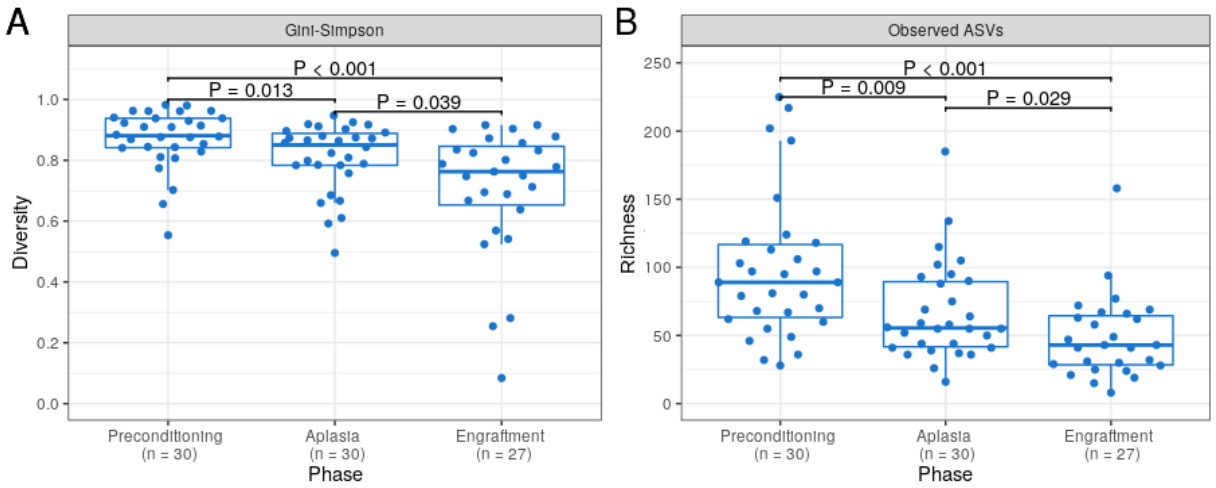
Supplementary material to Chapter 2 with 2 tables and 4 figures.

frontiers

Supplementary Material

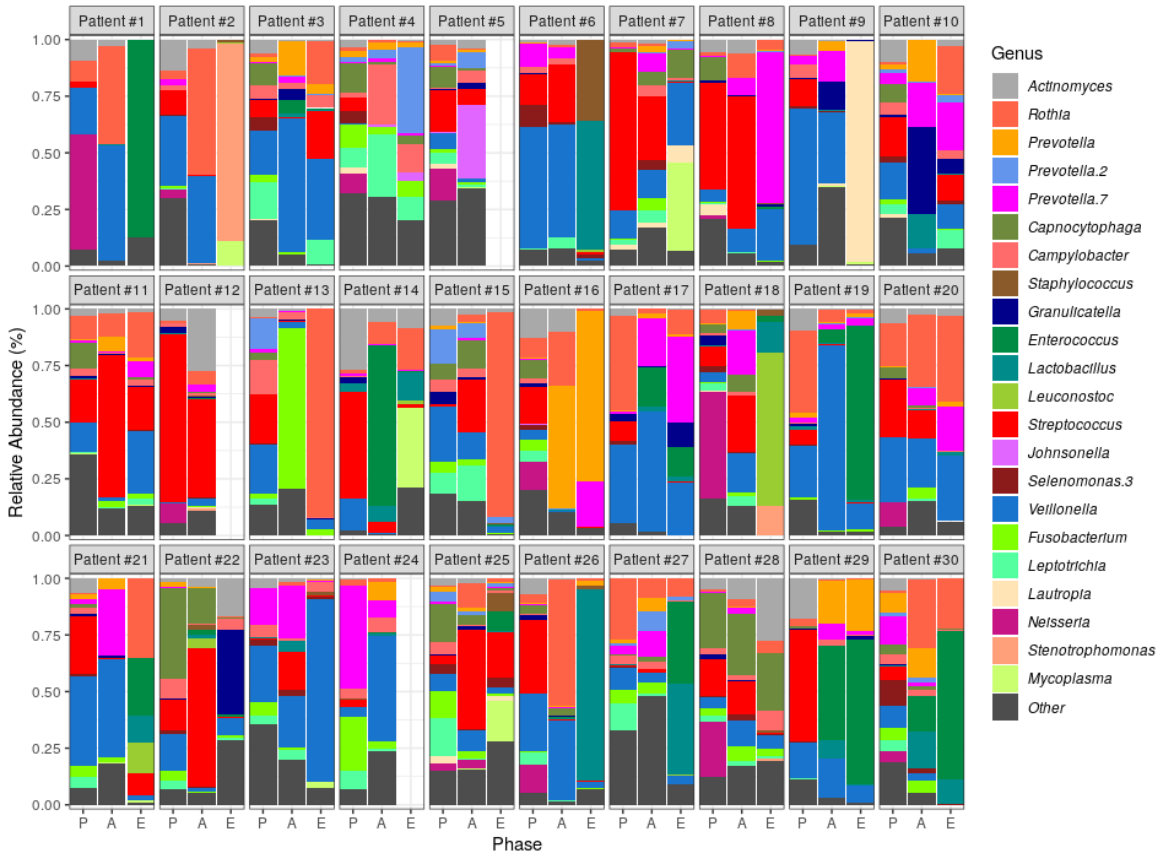
1 Supplementary Figures and Tables

1.1 Supplementary Figures

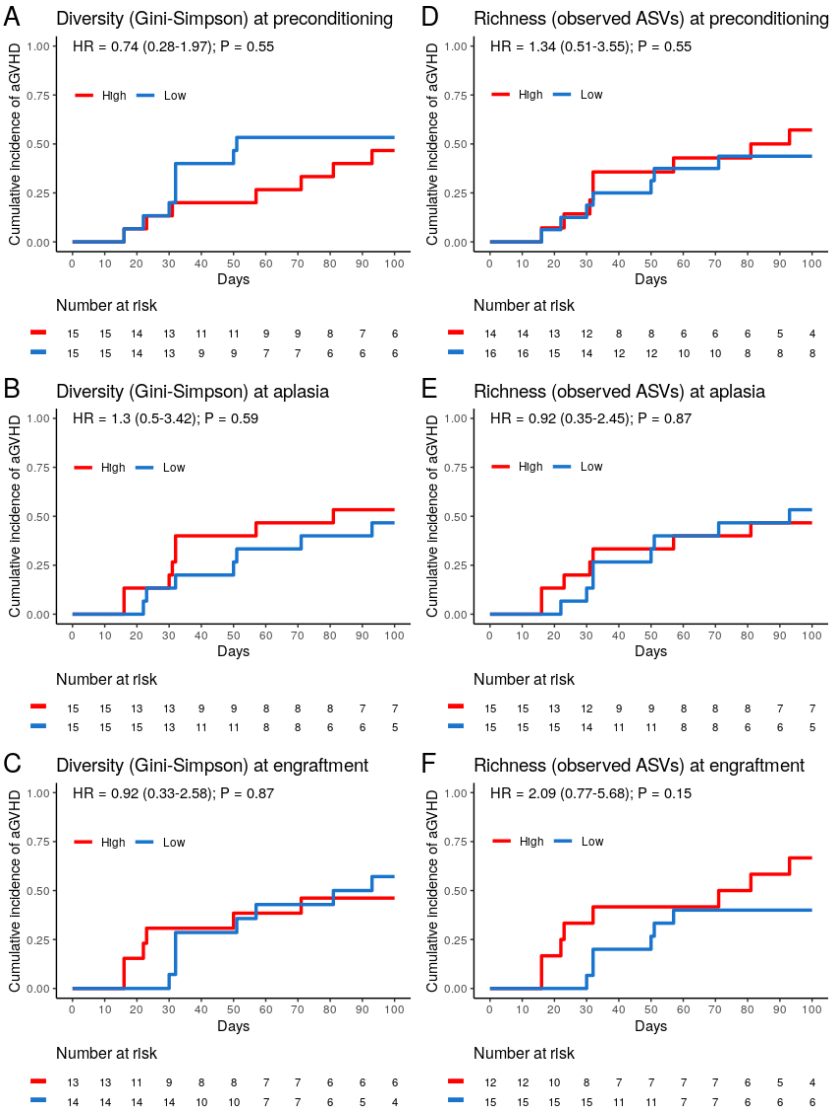


Supplementary Figure 1: Dental biofilm microbiota (DBM) alpha diversity decreases during allogeneic hematopoietic stem cell transplantation. (A-B) DBM alpha diversity boxplots at preconditioning (n = 30), aplasia (n = 30) and engraftment (n = 27) as measured by either Gini-Simpson index (A) or the number of observed ASVs as a proxy for species richness (B). Mann-Whitney U test was used with the preconditioning as the reference for comparisons. The boxes highlight the median value and cover the 25th and 75th percentiles, with whiskers extending to the more extreme value within 1.5 times the length of the box.

Supplementary Material

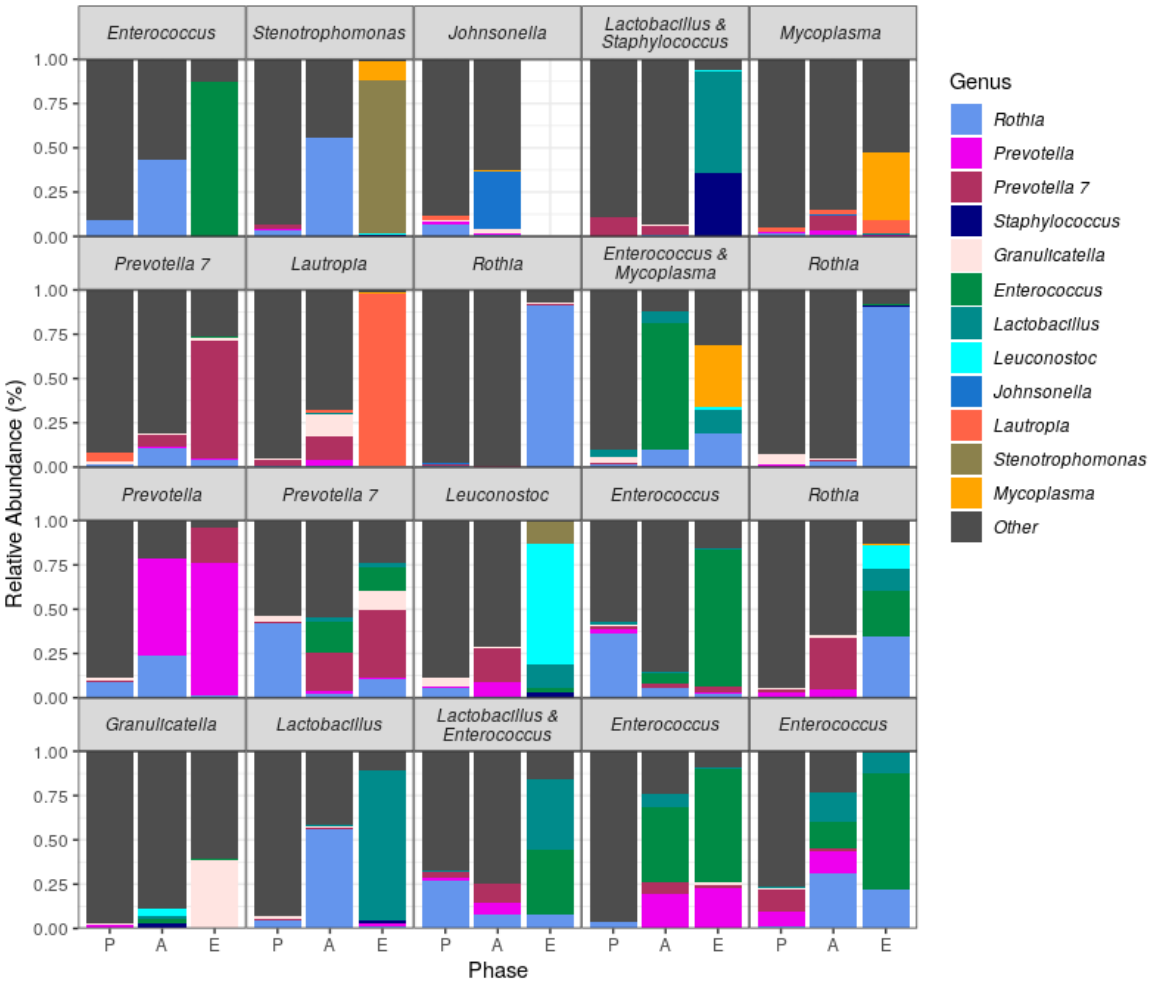


Supplementary Figure 2: Bacterial genera relative abundance changes in dental biofilm microbiota during allogeneic hematopoietic stem cell transplantation. Genera relative abundance composition across transplantation phases for all patients (n = 30). Missing samples did not reach quality criteria for analyses. Only genera with at least 1% relative abundance in at least 25% study samples or dominant genera are shown. Taxa are sorted based on taxonomic relatedness. P, Preconditioning; A, Aplasia; E, Engraftment.



Supplementary Figure 3: Dental biofilm microbiota alpha diversity is not associated with the risk of acute graft-versus-host disease (aGVHD). (A-C) Cumulative incidence of aGVHD with patients stratified by Gini-Simpson diversity index (High vs. Low) at preconditioning (A; n = 30), aplasia (B; n = 30) or engraftment (C; n = 27). (D-F) Cumulative incidence of aGVHD with patients stratified by the number of observed ASVs as a proxy for species richness (High vs. Low) at preconditioning (A; n = 30), aplasia (B; n = 30) or engraftment (C; n = 27).

Supplementary Material



Supplementary Figure 4: Relative abundance of genera across transplantation phases for all patients experiencing genus blooming events (n = 20). Only blooming genera are shown. Each subplot represents one patient experiencing some genus bloom, with subplot titles indicating the genera observed to bloom in such patient. P, Preconditioning; A, Aplasia; E, Engraftment.

1.2 Supplementary Tables

	HR (95% CI)	P-value
Age (in years)	0.99 (0.96-1.03)	0.64
Underlying disease* (AL vs. Other)	1.00 (0.37-2.69)	1
Conditioning intensity (Myeloablative)	0.74 (0.26-2.17)	0.59
TBI (Yes)	0.79 (0.29-2.20)	0.66
T-cell depletion (Yes)	0.78 (0.30-2.06)	0.61
Graft source (Bone Marrow)	0.95 (0.35-2.63)	0.92
Donor (MSD vs. Haploidentical)	1.14 (0.34-3.77)	0.83
Donor (MUD vs. Haploidentical)	0.80 (0.24-2.73)	0.72
Donor (MMUD vs. Haploidentical)	0.96 (0.14-6.30)	0.96
GVHD prophylaxis (MMF vs. MTX)	1.32 (0.47-3.70)	0.6
GVHD prophylaxis (MMF+CyPT vs. MTX)	1.18 (0.32-4.38)	0.81
AAB (Yes)	0.92 (0.35-2.40)	0.86
Cephalosporin (Yes)	3.44 (0.75-15.9)	0.11

Supplementary Table 1: **Univariate competing risk analysis for the association of acute graft-versus-host disease with clinical parameters**. All GVHD prophylaxis protocols include cyclosporin A. HCT-CI, Hematopoietic cell transplantation-specific comorbidity index; MMF, Mycophenolate mofetil; MTX, Methotrexate; TBI, Total body irradiation; AL, Acute leukemia; DRI, Disease relapse index; MSD, Matched sibling donor; MUD, Matched unrelated donor; MMUD, Mismatched unrelated donor; AAB, antibiotic for anaerobic bacteria; HR, Hazard ratio, CI, Confidence interval. *Acute leukemia: 11 acute myeloid leukemia and 7 acute lymphocytic leukemia cases**;** other: 5 non-Hodgkin lymphoma, 4 myelodysplastic syndrome, 1 chronic myeloid leukemia, 1 chronic lymphocytic leukemia and 1 multiple myeloma cases.

Supplementary Material

	HR (95% CI)	**P-value**
Diversity (Shannon) at P (High vs. Low)	0.89 (0.19-4.21)	0.89
Diversity (Shannon) at A (High vs. Low)	0.18 (0.02-1.58)	0.12
Diversity (Shannon) at E (High vs. Low)	0.92 (0.33-2.58)	0.96
Veillonella at P (High vs. Low)	1.93 (0.35-10.6)	0.45
Streptococcus at P (High vs. Low)	5.61 (0.67-47.1)	0.11
Corynebacterium at P (High vs. Low)	0.95 (0.17-5.21)	0.95
Ratio at P (>1 vs. ≤1)	0.68 (0.14-3.20)	0.63
Ratio at A (>1 vs. ≤1)	1.12 (0.20-6.15)	0.90
Ratio at E (>1 vs. ≤1)	0.60 (0.13-2.81)	0.52
Any genus bloom (Yes vs. No)	0.97 (0.19-5.09)	0.97
E. faecalis bloom (Yes vs. No)	1.02 (0.12-8.44)	0.98

Supplementary Table 2: **Univariate competing risk analysis for the association of chronic graft-versus-host disease with relevant microbiota variables**. HR, Hazard ratio; CI, Confidence interval; P, preconditioning; A, aplasia; E, engraftment.

Appendix C: Chapter 3 supplementary tables and figures

Supplementary material to Chapter 3 with 5 tables and 7 figures.

SUPPLEMENTARY MATERIAL

Tables

Table S1: Underlying disease, disease status at preconditioning, and oral microbiota diversity at preconditioning of 31 patients who underwent an allo-HSCT.

	OM diversity at preconditioning	**Diagnosis**	**Diasease status at preconditioning**
1	Low	ALL	Refractory
2	High	ALL	CR2
3	NA	ALL	CR1
4	Low	ALL	CR2
5	Low	ALL	CR3
6	Low	ALL	CR1
7	High	ALL	CR1
8	NA	AML	CR1
9	Low	AML	CR1
10	High	AML	CR1
11	Low	AML	CR1
12	Low	AML	CR1
13	High	AML	CR1
14	Low	AML	CR1
15	NA	AML	CR1
16	High	AML	CR1
17	High	AML	CR1
18	Low	AML	Refractory
19	High	CLL	PR
20	Low	CML	CR2
21	Low	MDS	CR1
22	High	MDS	CR1
23	High	MDS	CR1
24	High	MDS	CR1
25	High	MM	CR2
26	High	NHL	Refractory
27	High	NHL	CR3
28	Low	NHL	Refractory
29	Low	NHL	CR2
30	Low	NHL	PR

OM, oral microbiota; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; CLL, chronic lymphocytic leukemia; CML, chronic myeloid leukemia; MDS, myelodysplastic syndrome; MM, multiple myeloma; NHL, non-hodgkin lymphoma; CR1, first complete remission; CR2, second complete remission; CR3, third complete remission; PR, partial remission; NA, not available.

Table S2: Correlation of relative abundance changes from preconditioning to engraftment and allo-HSCT outcomes.

	N	Present	Absent	P value
aGVHD				
Enterococcus	25	14	11	0.54
Lactobacillus	25	17	8	0.14
Mycoplasma	25	8	17	0.96
Staphylococcus	25	10	15	0.26
saGVHD				
Enterococcus	25	14	11	0.52
Lactobacillus	25	17	8	1.00
Mycoplasma	25	8	17	0.42
Staphylococcus	25	10	15	0.58
cGVHD				
Enterococcus	25	14	11	0.03
Lactobacillus	25	17	8	0.63
Mycoplasma	25	8	17	0.46
Staphylococcus	25	10	15	0.36
NRM				
Enterococcus	25	14	11	0.19
Lactobacillus	25	17	8	0.10
Mycoplasma	25	8	17	0.12
Staphylococcus	25	10	15	0.67
Relapse				
Enterococcus	25	14	11	0.78
Lactobacillus	25	17	8	0.74
Mycoplasma	25	8	17	0.19
Staphylococcus	25	10	15	0.09
PFS				

Enterococcus	25	14	11	0.53
Lactobacillus	25	17	8	0.61
Mycoplasma	25	8	17	0.36
Staphylococcus	25	10	15	0.06
OS				
Enterococcus	25	14	11	0.31
Lactobacillus	25	17	8	0.80
Mycoplasma	25	8	17	0.43
Staphylococcus	25	10	15	0.34

Only patients with preconditioning and engraftment samples were included (n = 25). The relative abundance of a genus was considered to increase during allo-HSCT for a given patient when the relative abundance at engraftment was greater than at preconditioning and the final relative abundance was ≥0.1%. GVHD, graft versus host disease; aGVHD, acute GVHD; saGVHD, severe acute GVHD; cGVHD, chronic GVHD; NRM, non-relapse mortality; PFS, progression-free survival; OS, overall survival.

Table S3: Correlations of bacterial diversity with clinical features and transplant outcomes.

	Preconditioning (n = 27)			Aplasia (n = 28)			Engraftment (n = 26)		
	Low	**High**	**P value**	**Low**	**High**	**P value**	**Low**	**High**	**P value**
Sex									
Female	9	4		8	6		4	8	
Male	5	9	0.13	6	8	0.71	9	5	0.24
Underlying disease									
Acute leukemia	9	6		7	9		9	6	
Other	5	7	1	7	5	0.70	4	7	0.43
HCT-CI									
0	7	9		8	8		4	9	
1–2	5	1		3	3		5	3	
≥3	2	3	0.24	3	3	1	4	1	0.15
Disease risk index									
Low–intermediate	6	8		8	7		9	6	
High	8	5	0.45	6	7	1	4	7	0.43
Conditioning intensity									
Reduced intensity	8	8		10	7		7	9	
Myeloablative	6	5	1	4	7	0.44	6	4	0.69
Total body irradiation									
No	8	10		8	10		7	9	
Yes	6	3	0.42	6	4	0.69	6	4	0.69
T-cell depletion									
No	6	6		7	6		8	6	
Yes	8	7	1	7	8	1	5	7	0.70
Graft source									
Peripheral blood	10	9		10	10		8	8	
Bone marrow	4	4	1	4	4	1	5	5	1
Donor									
Haploidentical	5	3		5	4		4	5	
Matched sibling	3	5		4	4		3	5	

Matched unrelated	5	4		4	5		4	3	
Mismatched unrelated	1	1	0.84	1	1	1	1	1	0.88
GVHD prophylaxis									
CsA + MTX	3	6		4	5		6	2	
CsA + MMF	7	4		6	5		2	8	
CsA + MMF + CyPT	4	3	0.49	4	4	1	5	3	0.09
aGVHD									
No	8	5		7	7		7	6	
Yes	6	8	0.45	7	7	1	6	7	1
saGVHD									
No	9	12		11	11		10	11	
Yes	5	1	0.16	3	3	1	3	2	1
cGVHD									
No	13	9		12	11		10	10	
Yes	1	4	0.16	2	3	1	3	3	1
Oral mucositis									
No	8	7		8	8		8	7	
Yes	6	6	1	6	6	1	5	6	1
Febrile neutropenia									
No	0	2		2	0		0	2	
Yes	14	11	0.22	12	14	0.48	13	11	0.48
Age (years)			0.30			0.87			0.68

Diversity classification was based on the median Shannon index diversity measured across the study population at each collection time. Association of diversity classification and clinical parameters was evaluated by Fisher's exact tests for categorical variables and two-sided Student's t-tests for age. HCT-CI, hematopoietic cell transplantation-specific comorbidity index; CsA, cyclosporin A; MMF, mycophenolate mofetil; MTX, methotrexate; aGVHD, acute GVHD; saGVHD, severe acute GVHD; cGVHD, chronic GVHD.

Table S4: Univariate competing risk analysis for the association of relapse with oral mucosa diversity.

	HR (95% CI)	***P* value**
Diversity at preconditioning (High)	0.27 (0.07–0.97)	0.04
Diversity at aplasia (High)	1.30 (0.43–3.90)	0.64
Diversity at engraftment (High)	0.73 (0.21–2.53)	0.62

HR, hazard ratio; CI, confidence interval.

Table S5: Univariate competing risk analysis for the association of relapse with clinical parameters.

	HR (95% CI)	*P* value
Age (years)	0.97 (0.94–1.01)	0.14
Underlying disease (AL versus other)	0.80 (0.25–2.56)	0.70
HCT-CI (1–2 versus 0)	1.97 (0.63–6.21)	0.25
HCT-CI (≥ 3 versus 0)	0.35 (0.05–2.76)	0.32
DRI (High)	10.2 (2.24–46.7)	< 0.01
Conditioning intensity (Myeloablative)	0.95 (0.32–2.84)	0.93
TBI (Yes)	2.04 (0.68–6.16)	0.21
T-cell depletion (Yes)	2.43 (0.79–7.53)	0.12
Graft source (Bone marrow)	0.96 (0.31–2.95)	0.94
Donor (MSD versus haploidentical)	0.71 (0.13–3.91)	0.69
Donor (MUD versus haploidentical)	1.97 (0.52–7.47)	0.32
Donor (MMUD versus haploidentical)	15.6 (2.21–110)	< 0.01
GVHD prophylaxis (MMF versus MTX)	0.81 (0.23–2.90)	0.75
GVHD prophylaxis (MMF+CyPT versus MTX)	0.69 (0.18–2.73)	0.60
Tazobactam (Yes)	0.48 (0.07–3.28)	0.46
Cefepime (Yes)	0.51 (0.17–1.53)	0.23
Meropenem (Yes)	1.11 (0.35–3.49)	0.86
Oral mucositis	1.41 (0.47–4.17)	0.54

All GVHD prophylaxis protocols included CsA. HCT-CI, hematopoietic cell transplantation-specific comorbidity index; MMF, mycophenolate mofetil; MTX, methotrexate; TBI, total body irradiation; AL, acute leukemia; DRI, Disease Risk Index; MSD, matched sibling donor; MUD, matched unrelated donor; MMUD, mismatched unrelated donor; HR, hazard ratio; CI, confidence interval.

Figures legends

Figure S1: Bacterial richness within the oral mucosa decreases during allo-HSCT. (A) Oral mucosa bacterial richness boxplot at preconditioning (n = 27), aplasia (n = 28), and engraftment (n = 26), as measured by the number of observed ASVs. Mann-Whitney U tests were used with the preconditioning collection as the reference for comparisons. The boxes highlight the median values and cover the 25th and 75th percentiles, with whiskers extending to the more extreme value within 1.5 times the length of the box. Outliers are represented explicitly. Asterisks represent statistical significance: *, $P < 0.05$; **, $P < 0.01$. ASV, amplicon sequencing variant.

Figure S2: Changes in bacterial taxa during allo-HSCT. Relative abundances of phyla (A), classes (B), orders (C), families (D) and genera (E) in the oral mucosa across transplantation phases for all patients (n = 30). Only taxa showing relative abundance ≥30% in at least one study sample or relative abundance ≥5% in at least 25% of study samples are shown. P, preconditioning; A, aplasia; E, engraftment.

Figure S3: Significant changes in bacterial genera during allo-HSCT. Significant genera relative abundance variations from preconditioning to aplasia and from preconditioning to engraftment according to ANCOM test (W > 0.7). Relative differences are represented by the $\log_2$-transformed average relative abundance fold change between groups.

Figure S4: Relative abundance changes from preconditioning to engraftment of potentially pathogenic genera. Each line represents a study patient. Only patients with preconditioning and engraftment samples were included (n = 25). The relative abundance of a genus was considered to increase during allo-HSCT for a given patient when the relative abundance at engraftment was greater than at preconditioning and the final relative abundance was ≥0.1%.

Figure S5: Univariate competing risk analysis for the association of relapse with clinical parameters. The variables are sorted in ascending order according to the hazard ratio. All GVHD prophylaxis protocols included CsA. HCT-CI, hematopoietic cell transplantation-specific comorbidity index; MMF, mycophenolate mofetil; MTX, methotrexate; TBI, total body irradiation; AL, acute leukemia; DRI, Disease Risk Index; CI, conditioning intensity; MSD, matched sibling donor; MUD, matched unrelated donor; MMUD, mismatched unrelated donor.

Figure S6: Univariate competing risk analysis for the association of relapse with specific genus presence at preconditioning. The variables are sorted in ascending order according to hazard ratio. Only genera present in at least 25% of samples and absent in at least 25% of samples were evaluated.

Figure S7: OM Solobacterium relative abundance in preconditioning samples. Each bar represents a study patient.

Figures

Figure S1

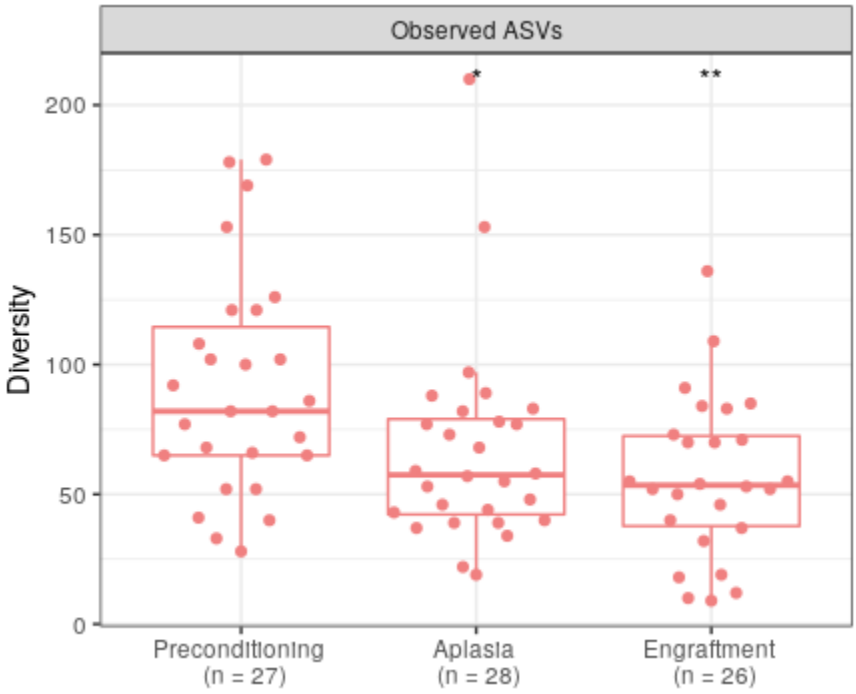


Figure S2

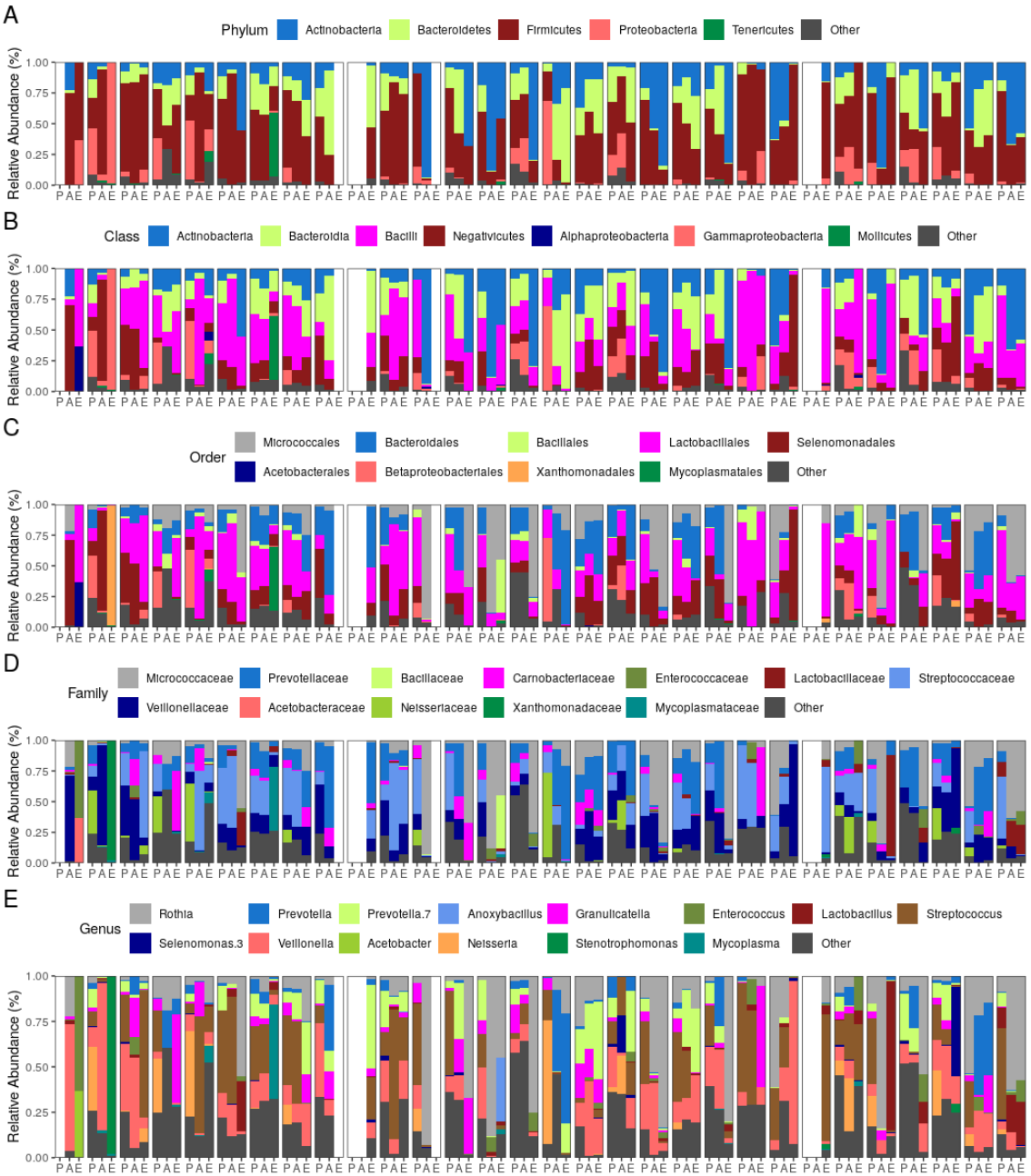

A
Phylum
Actinobacteria
Bacteroidetes
Firmicutes
Proteobacteria
Tenericutes
Other
Relative Abundance (%)
1.00
0.75
0.50
0.25
0.00
PAE
B
Class
Actinobacteria
Bacteroidia
Bacilli
Negativicutes
Alphaproteobacteria
Gammaproteobacteria
Mollicutes
Other
C
Order
Micrococcales
Bacteroidales
Bacillales
Lactobacillales
Selenomonadales
Acetobacterales
Betaproteobacterlales
Xanthomonadales
Mycoplasmatales
Other
D
Family
Micrococcaceae
Prevotellaceae
Bacillaceae
Carnobacteriaceae
Enterococcaceae
Lactobacillaceae
Streptococcaceae
Veillonellaceae
Acetobacteraceae
Neisseriaceae
Xanthomonadaceae
Mycoplasmataceae
Other
E
Genus
Rothia
Prevotella
Prevotella.7
Anoxybacillus
Granulicatella
Enterococcus
Lactobacillus
Streptococcus
Selenomonas.3
Veillonella
Acetobacter
Neisseria
Stenotrophomonas
Mycoplasma
Other

Figure S3

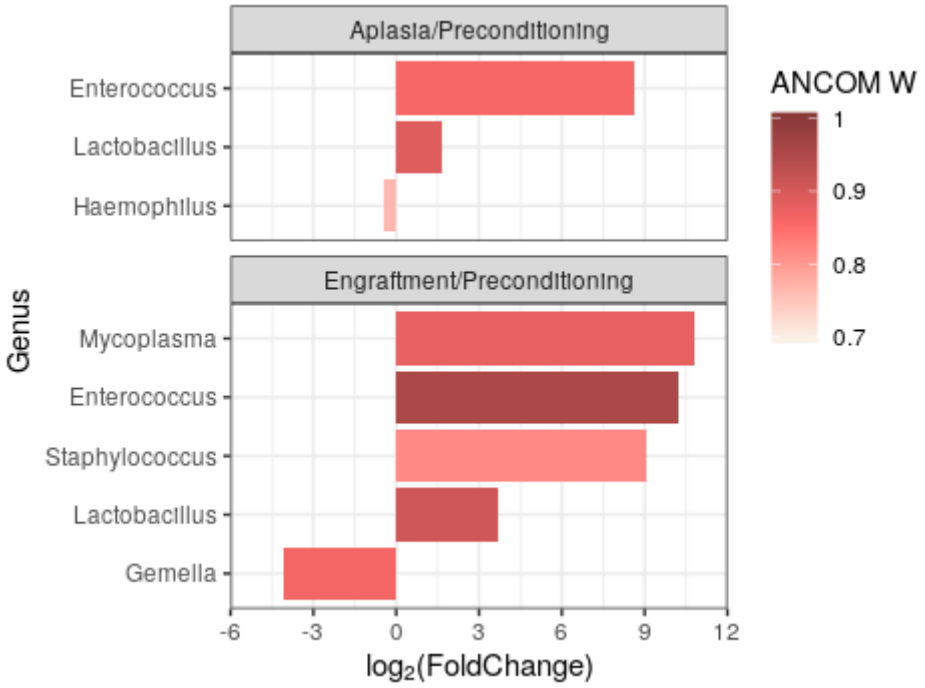

Aplasia/Preconditioning
Enterococcus
Lactobacillus
Haemophilus
Engraftment/Preconditioning
Mycoplasma
Enterococcus
Staphylococcus
Lactobacillus
Gemella
Genus
-6
-3
0
3
6
9
12
log2(FoldChange)
ANCOM W
1
0.9
0.8
0.7

Figure S4

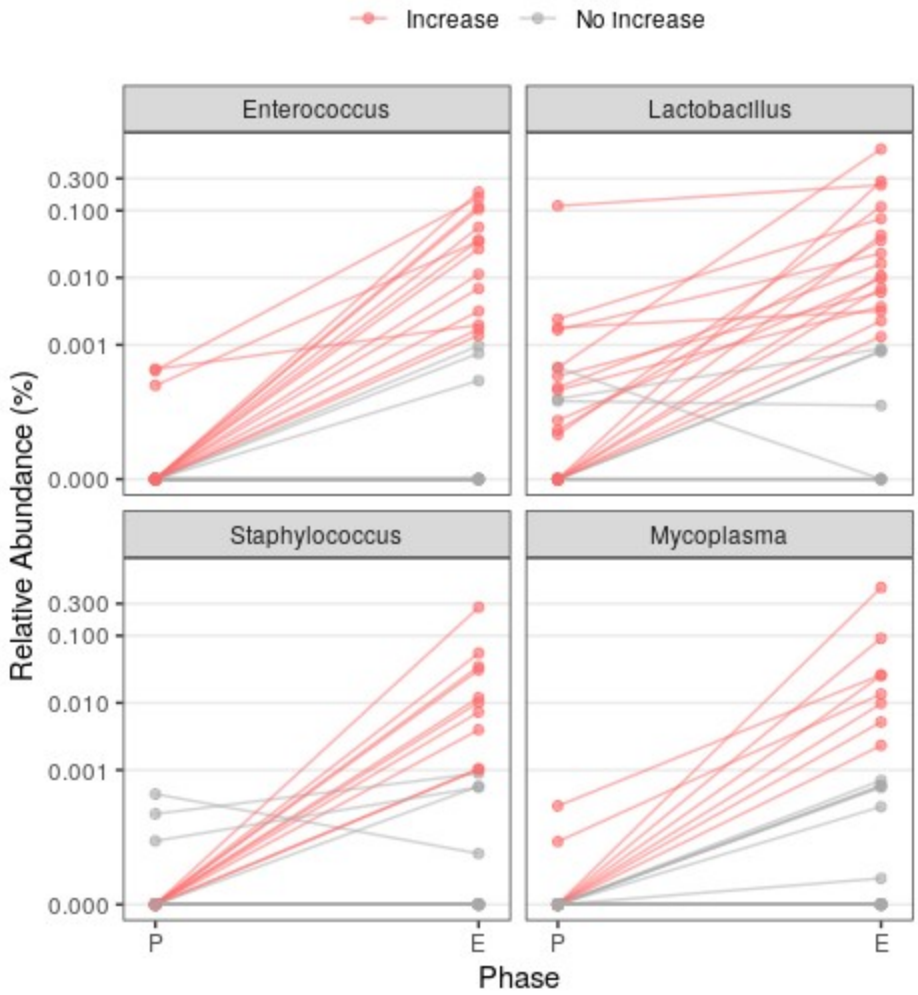

Increase
No Increase
Enterococcus
Lactobacillus
Staphylococcus
Mycoplasma
Relative Abundance (%)
0.300
0.100
0.010
0.001
0.000
P
E
Phase

Figure S5

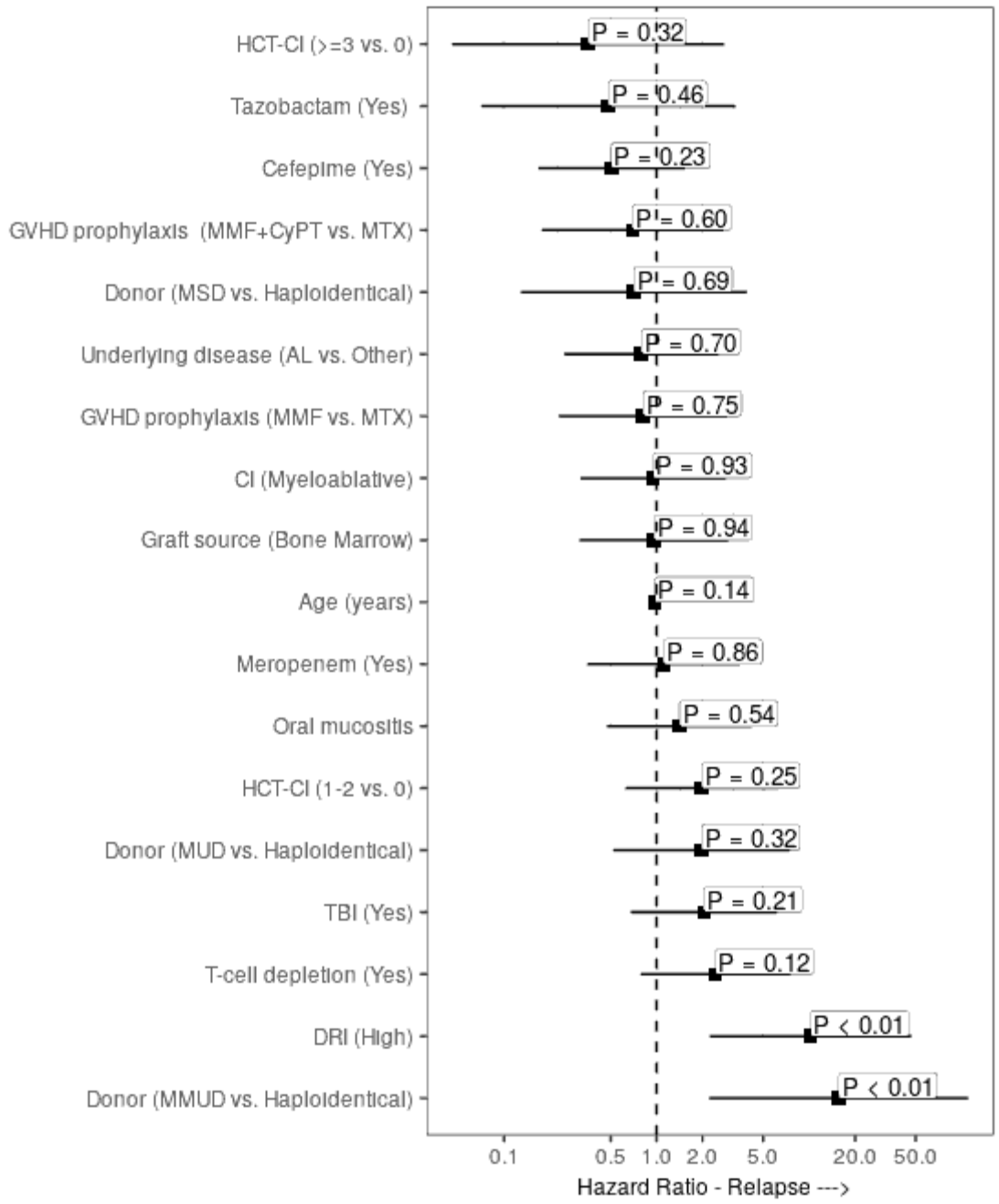

HCT-CI (>=3 vs. 0)
Tazobactam (Yes)
Cefepime (Yes)
GVHD prophylaxis (MMF+CyPT vs. MTX)
Donor (MSD vs. Haploidentical)
Underlying disease (AL vs. Other)
GVHD prophylaxis (MMF vs. MTX)
CI (Myeloablative)
Graft source (Bone Marrow)
Age (years)
Meropenem (Yes)
Oral mucositis
HCT-CI (1-2 vs. 0)
Donor (MUD vs. Haploidentical)
TBI (Yes)
T-cell depletion (Yes)
DRI (High)
Donor (MMUD vs. Haploidentical)
P = 0.32
P = 0.46
P = 0.23
P = 0.60
P = 0.69
P = 0.70
P = 0.75
P = 0.93
P = 0.94
P = 0.14
P = 0.86
P = 0.54
P = 0.25
P = 0.32
P = 0.21
P = 0.12
P < 0.01
P < 0.01
0.1
0.5
1.0
2.0
5.0
20.0
50.0
Hazard Ratio - Relapse --->

Figure S6

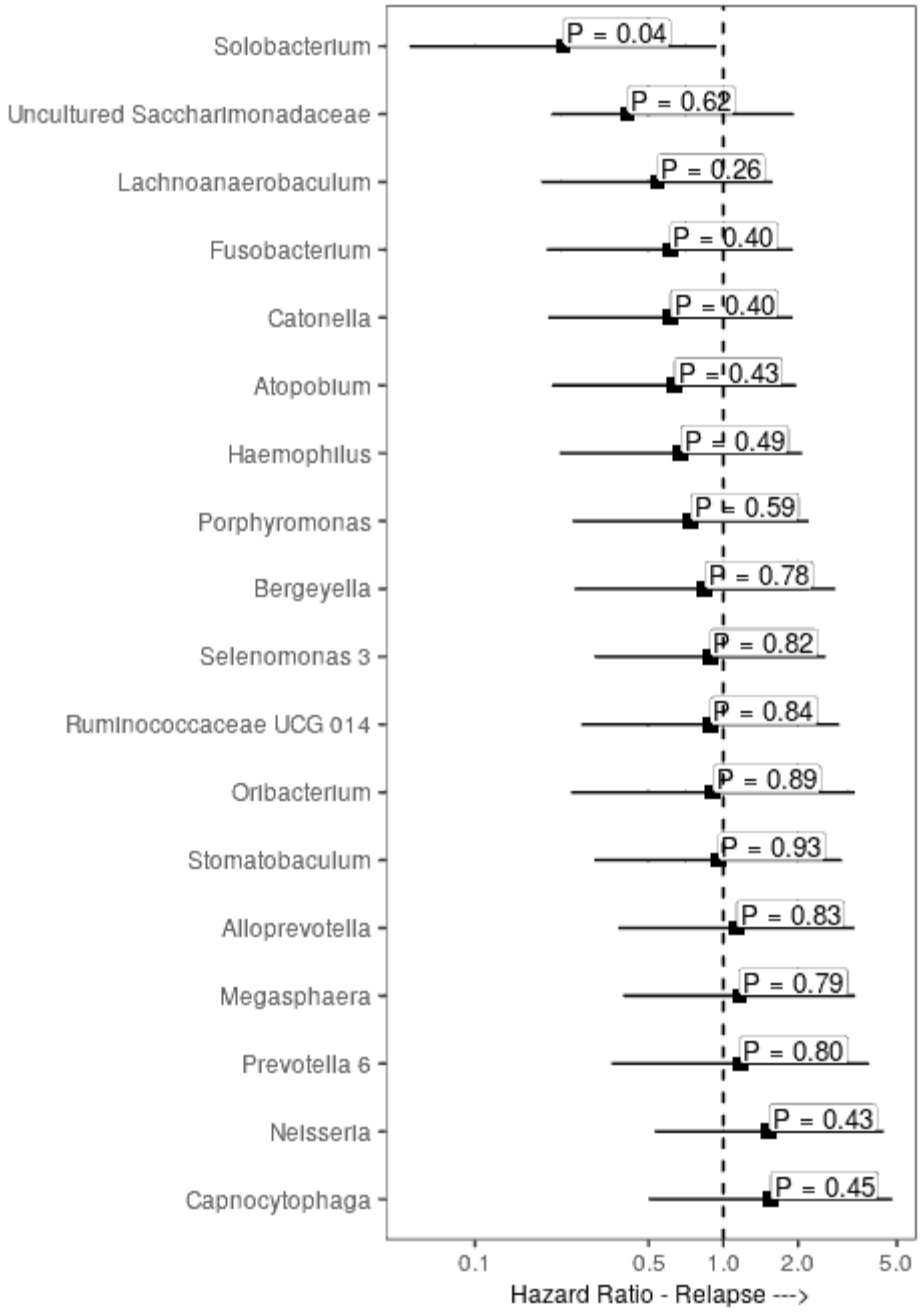
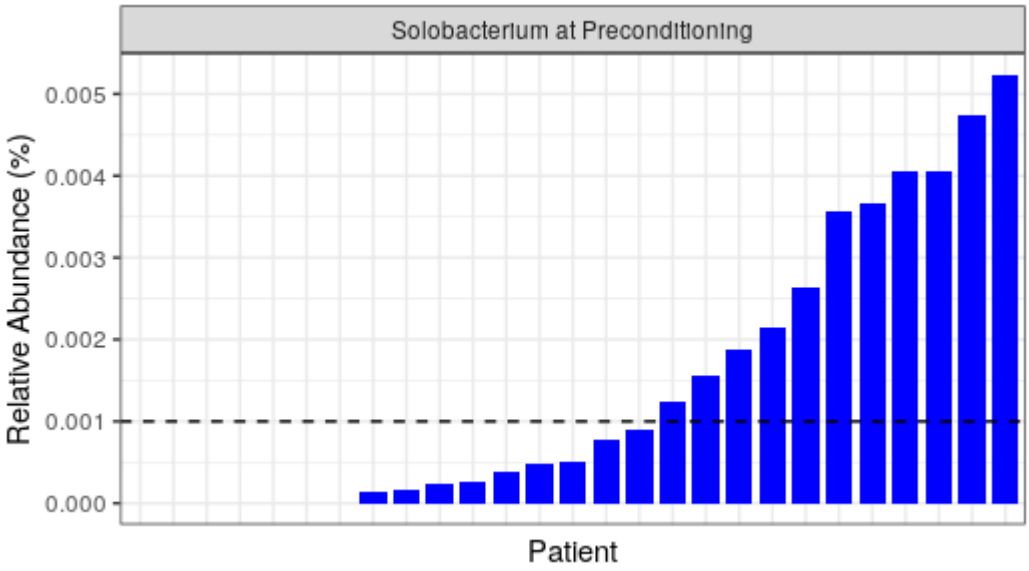
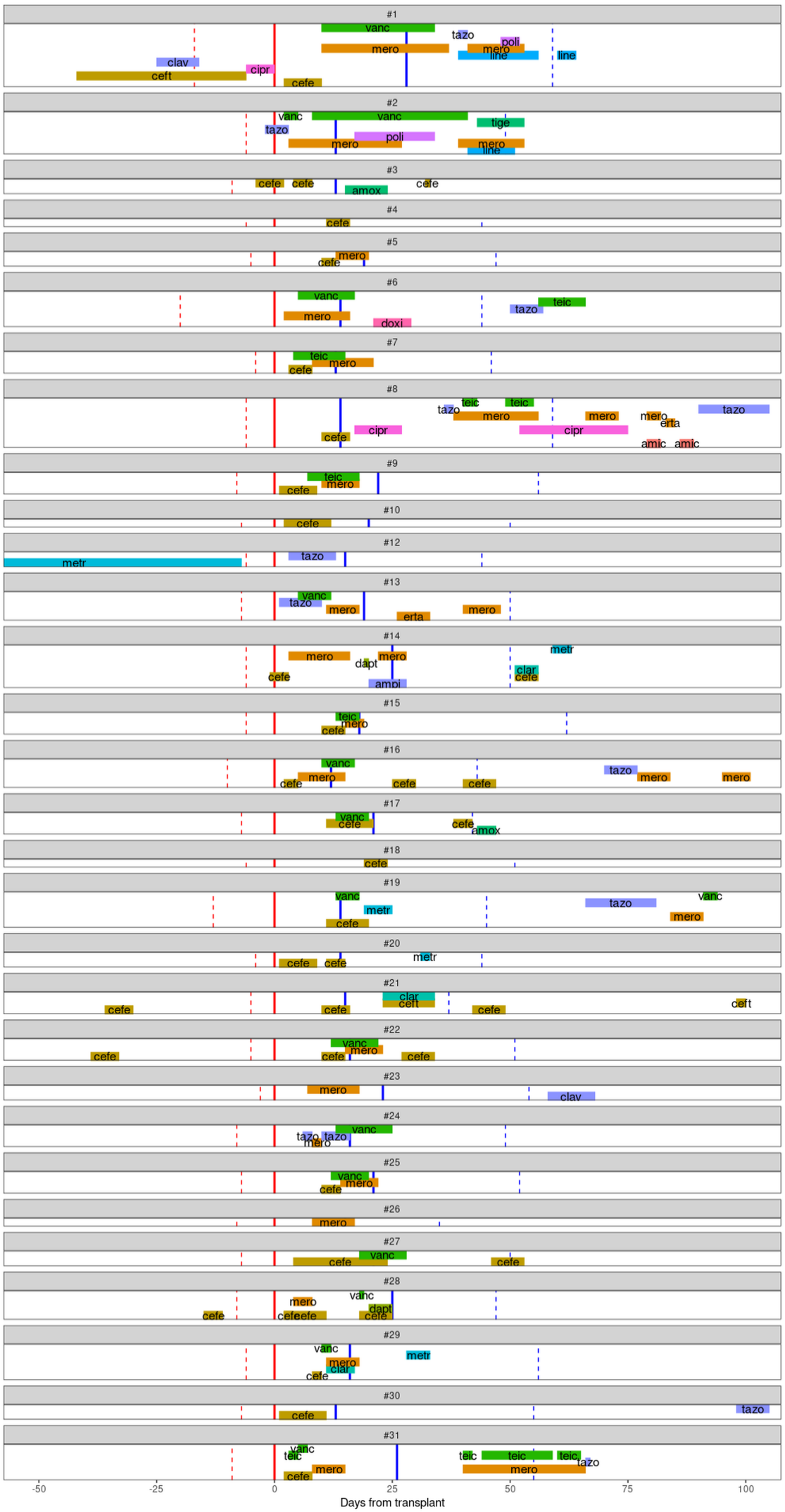

Solobacterium
Uncultured Saccharimonadaceae
Lachnoanaerobaculum
Fusobacterium
Catonella
Atopobium
Haemophilus
Porphyromonas
Bergeyella
Selenomonas 3
Ruminococcaceae UCG 014
Oribacterium
Stomatobaculum
Alloprevotella
Megasphaera
Prevotella 6
Neisseria
Capnocytophaga
P = 0.04
P = 0.62
P = 0.26
P = 0.40
P = 0.40
P = 0.43
P = 0.49
P = 0.59
P = 0.78
P = 0.82
P = 0.84
P = 0.89
P = 0.93
P = 0.83
P = 0.79
P = 0.80
P = 0.43
P = 0.45
0.1
0.5
1.0
2.0
5.0
Hazard Ratio - Relapse --->

Figure S7


Solobacterium at Preconditioning
Relative Abundance (%)
0.005
0.004
0.003
0.002
0.001
0.000
Patient

Appendix D: Chapter 4 timelines of antibiotic usage

Antibiotic usage timelines for each patient analyzed in Chapter 4 in relation to stem cell infusion. Red dashed line indicates preconditioning sampling. Red solid line indicates stem cell infusion. Blue solid line indicates stem cell engraftment. Blue dashed line indicates 30 days after engraftment sampling. clav, amoxicillin clavulanate; tazo, piperacillin tazobactam; amox, amoxicillin; cefe, cefepime; mero, meropenem; metr, metronidazole; ceft, ceftriaxone; vanc, vancomycin; teic, teicoplanin; cipr, ciprofloxacin; levo, levofloxacin; doxi, doxycycline; ampi, ampicillin; clar, clarithromycin; bact, sulfamethoxazole trimethoprim; erta, ertapenem; poli, polymyxin b; dapt, daptomycin; line, linezolid; tige, tigecycline; amic, amikacin.


Antibiotic class
aminoglycosides
carbapenems
cephalosporins
cyclic lipopeptides
glycopeptides
glycylcyclines
macrolides
nitroimidazoles
oxazolidinones
penicillins
polymyxins
quinolones
tetracyclines
#1
vanc
tazo
poli
mero
line
clav
cipr
ceft
cefe
#2
vanc
tige
tazo
poli
mero
line
#3
cefe
amox
#4
cefe
#5
mero
cefe
#6
vanc
teic
tazo
mero
doxi
#7
teic
mero
cefe
#8
teic
tazo
mero
erta
cipr
amic
#9
teic
mero
cefe
#10
cefe
#12
metr
tazo
#13
vanc
tazo
mero
erta
#14
metr
mero
dapt
clar
cefe
ambi
#15
teic
mero
cefe
#16
vanc
tazo
mero
cefe
#17
vanc
cefe
amox
#18
cefe
#19
vanc
metr
tazo
mero
cefe
#20
metr
cefe
#21
clar
ceft
cefe
#22
vanc
mero
cefe
#23
mero
clav
#24
vanc
tazo
mero
#25
vanc
mero
cefe
#26
mero
#27
vanc
cefe
#28
vanc
mero
dapt
cefe
#29
vanc
metr
mero
clar
cefe
#30
cefe
tazo
#31
vanc
teic
mero
cefe
tazo
-50
-25
0
25
50
75
100
Days from transplant

Appendix E: Chapter 4 supplementary tables and figures

Supplementary material to Chapter 4 with 9 tables and 14 figures.

Longitudinal analysis at three oral sites links oral microbiota to clinical outcomes in allogeneic hematopoietic stem-cell transplant

SUPPLEMENTARY TABLES

Table S1

	n = 31
Age in years (median, range)	50 (19–73)
Sex (male)	17 (55%)
Underlying disease	
Acute leukemia	18 (58%)
Myeloid	11 (35%)
Lymphocytic	7 (23%)
Other	13 (42%)
Non-Hodgkin lymphoma	5 (16%)
Myelodysplastic syndrome	5 (16%)
Chronic myeloid leukemia	1 (3%)
Chronic lymphocytic leukemia	1 (3%)
Multiple myeloma	1 (3%)
HCT-CI	
0	16 (52%)
1-2	9 (29%)
≥3	6 (19%)
DRI	
Low-intermediate	18 (58%)
High	13 (42%)
Conditioning intensity	
Reduced intensity	19 (61%)
Myeloablative	12 (39%)
Total body irradiation	11 (35%)
T-cell depletion	16 (52%)
Graft source	
Bone marrow	10 (32%)
Peripheral blood	21 (68%)
Donor	

MRD	9 (29%)
Haploidentical	10 (32%)
MUD/MMUD	12 (39%)
Antibiotic usage*	
Main classes	
Cephalosporins	22 (73%)
Carbapenems	19 (63%)
Glycopeptides	18 (60%)
Penicillins	7 (23%)
Metrics	
LOT (median, range)	15.5 (0–58)
DOT (median, range)	22 (0–112)

Table S1: Clinical characteristics of study patients. *Antibiotics usage percent values are calculated considering a total of 30 patients (see Materials and methods). HCT-CI, Hematopoietic Cell Transplantation-specific Comorbidity Index; DRI, Disease Risk Index; MRD, matched related donor; MUD, matched unrelated donor; MMUD, mismatched unrelated donor; LOT, length of therapy; DOT, days of therapy.

Table S2

	Diversity stability			Composition stability		
	Coefficient	**SE**	**P-value**	**Coefficient**	**SE**	**P-value**
GCF						
Intercept	0,9660	0,1007	<0,0001	0,4468	0,1697	0,0146
Cephalosporins	0,0427	0,0969	0,6636	-0,2937	0,1633	0,0848
Carbapenems	0,0088	0,0951	0,9269	0,0206	0,1603	0,8990
Glycopeptides	-0,0895	0,1011	0,3848	-0,1952	0,1704	0,2633
Penicillins	0,0872	0,1224	0,4830	-0,3756	0,2063	0,0812
DOT	-0,0057	0,0022	**0,0172**	0,0031	0,0037	0,4136
OM						
Intercept	0,6794	0,1852	0,0014	0,0623	0,1978	0,7560
Cephalosporins	0,2011	0,1858	0,2913	-0,0658	0,1985	0,7440
Carbapenems	0,1687	0,1706	0,3340	-0,1948	0,1822	0,2970
Glycopeptides	0,2534	0,1877	0,1914	0,0171	0,2005	0,9330
Penicillins	0,1655	0,2158	0,4516	0,1565	0,2306	0,5050
DOT	-0,0167	0,0046	**0,0015**	0,0019	0,0049	0,6970
SB						
Intercept	0,7131	0,1829	0,0007	0,3928	0,1888	0,0488
Cephalosporins	0,2947	0,1764	0,1084	-0,3677	0,1821	0,0553
Carbapenems	0,1024	0,1739	0,5617	0,1713	0,1795	0,3500
Glycopeptides	-0,1365	0,1846	0,4673	-0,4624	0,1906	**0,0235**
Penicillins	0,1890	0,2219	0,4031	-0,1793	0,2290	0,4417
DOT	-0,0085	0,0041	**0,0467**	0,0050	0,0042	0,2456

Table S2: Multiple linear models testing prediction of oral microbiota stability by antibiotic usage. Significant predictors are highlighted in bold. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; DOT, days of therapy; SE, standard error.

Table S3

	Outcome	N total	N group R	N group NR	% event R	% event NR	HR (95% CI)	P-value
GCF recovery	OS	30	23	7	39	71	0,39 (0,13–1,16)	0,0897
	PFS	28	22	6	36	50	0,51 (0,13–1,93)	0,3197
	Relapse	28	22	6	36	50	0,66 (0,18–2,36)	0,5200
	TRD	30	23	7	22	29	0,71 (0,14–3,66)	0,6900
OM recovery	OS	29	20	9	25	89	0,17 (0,05–0,52)	**0,0020**
	PFS	27	19	8	26	75	0,06 (0,01–0,34)	**0,0012**
	Relapse	27	19	8	26	75	0,20 (0,06–0,69)	**0,0110**
	TRD	29	20	9	10	44	0,19 (0,04–1,00)	0,0500
SB recovery	OS	30	23	7	48	43	0,87 (0,24–3,14)	0,8328
	PFS	28	22	6	41	33	1,15 (0,23–5,83)	0,8622
	Relapse	28	22	6	41	33	1,54 (0,36–6,66)	0,5600
	TRD	30	23	7	22	29	0,64 (0,12–3,50)	0,6000

Table S3: Univariate associations between oral microbiota recovery and clinical outcomes. The total number of patients considered in each association is indicated (N total). The variation in N total per associations is caused by the unavailability of a sample essential for recovery evaluation or the exclusion of patients experiencing the event before recovery evaluation. Patients were grouped into recoverers (R) and non-recoverers (NR). The percentage of patients in each group experiencing the event is indicated. Q-value refers to the P-value adjusted for the number of oral sites tested. Significant associations are highlighted in bold. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; OS, overall survival; PFS, progression-free survival; TRD, transplant-related death; HR, hazard ratio; CI, confidence interval.

Table S4

Univariate associations with OS		
	HR (95% CI)	**P-value**
Age in years	1,01 (0,97–1,05)	0,6223
Sex (Female vs Male)	1,28 (0,46–3,54)	0,6399
Underlying disease (AL vs Other)	0,48 (0,17–1,32)	0,1557
HCT-CI (1-2 vs 0)	1,67 (0,54–5,19)	0,3746
HCT-CI (≥3 vs 0)	1,62 (0,40–6,49)	0,4984
DRI (H vs LI)	3,93 (1,32–11,7)	**0,0139**
Conditioning intensity (M vs RI)	0,27 (0,08–0,97)	**0,0443**
Total body irradiation (Yes vs No)	1,76 (0,64–4,85)	0,2779
T-cell depletion (Yes vs No)	2,79 (0,93–8,32)	0,0659
Graft source (BM vs PB)	0,23 (0,05–1,02)	0,0539
Donor (HI vs MRD)	0,89 (0,22–3,56)	0,8675
Donor (MUD/MMUD vs MRD)	1,79 (0,50–6,37)	0,3678
Cephalosporins (Yes vs No)	0,63 (0,21–1,88)	0,4050
Carbapenems (Yes vs No)	2,46 (0,69–8,82)	0,1674
Glycopeptides (Yes vs No)	1,36 (0,46–4,07)	0,5803
Penicillins (Yes vs No)	1,49 (0,47–4,77)	0,4984
DOT	1,04 (1,02–1,07)	**0,0006**

Table S4: Univariate associations between clinical parameters and overall survival (OS). Significant associations are highlighted in bold. AL, acute leukemia; HCT-CI, Hematopoietic Cell Transplantation-specific Comorbidity Index; DRI, Disease Risk Index; H, high; LI, low-intermediate; M, myeloablative; RI, reduced intensity; BM, bone marrow; PB, peripheral blood; HI, haploidentical; MRD, matched related donor; MUD, matched unrelated donor; MMUD, mismatched unrelated donor; DOT, days of therapy; HR, hazard ratio; CI, confidence interval.

Table S5

Univariate associations with PFS		
	HR (95% CI)	**P-value**
Age in years	0,99 (0,95–1,02)	0,4348
Sex (Female vs Male)	2,41 (0,83–7,01)	0,1079
Underlying disease (AL vs Other)	0,51 (0,18–1,50)	0,2233
HCT-CI (1-2 vs 0)	1,57 (0,52–4,70)	0,4219
HCT-CI (≥3 vs 0)	0,38 (0,05–3,08)	0,3632
DRI (H vs LI)	7,12 (1,92–26,5)	**0,0034**
Conditioning intensity (M vs RI)	0,65 (0,22–1,98)	0,4533
Total body irradiation (Yes vs No)	2,03 (0,71–5,79)	0,1877
T-cell depletion (Yes vs No)	2,75 (0,90–8,36)	0,0745
Graft source (BM vs PB)	0,63 (0,20–2,01)	0,4331
Donor (HI vs MRD)	2,03 (0,37–11,1)	0,4166
Donor (MUD/MMUD vs MRD)	4,03 (0,85–19,1)	0,0792
Cephalosporins (Yes vs No)	0,50 (0,16–1,53)	0,2230
Carbapenems (Yes vs No)	1,50 (0,46–4,94)	0,5023
Glycopeptides (Yes vs No)	1,68 (0,52–5,46)	0,3892
Penicillins (Yes vs No)	0,65 (0,14–2,94)	0,5779
DOT	1,01 (0,98–1,05)	0,3782

Table S5: Univariate associations between clinical parameters and progression-free survival (PFS). Significant associations are highlighted in bold. AL, acute leukemia; HCT-CI, Hematopoietic Cell Transplantation-specific Comorbidity Index; DRI, Disease Risk Index; H, high; LI, low-intermediate; M, myeloablative; RI, reduced intensity; BM, bone marrow; PB, peripheral blood; HI, haploidentical; MRD, matched related donor; MUD, matched unrelated donor; MMUD, mismatched unrelated donor; DOT, days of therapy; HR, hazard ratio; CI, confidence interval.

Table S6

Univariate associations with relapse		
	HR (95% CI)	**P-value**
Age in years	0,98 (0,95–1,02)	0,3000
Sex (Female vs Male)	2,48 (0,89–6,85)	0,0810
Underlying disease (AL vs Other)	0,61 (0,22–1,68)	0,3400
HCT-CI (1-2 vs 0)	1,77 (0,61–5,18)	0,2900
HCT-CI (≥3 vs 0)	0,31 (0,04–2,40)	0,2600
DRI (H vs LI)	5,96 (1,94–18,3)	**0,0018**
Conditioning intensity (M vs RI)	0,79 (0,28–2,20)	0,6500
Total body irradiation (Yes vs No)	2,09 (0,77–5,68)	0,1500
T-cell depletion (Yes vs No)	2,20 (0,79–6,14)	0,1300
Graft source (BM vs PB)	0,79 (0,26–2,39)	0,6800
Donor (HI vs MRD)	2,07 (0,39–11,1)	0,4000
Donor (MUD/MMUD vs MRD)	3,84 (0,83–17,7)	0,0840
Cephalosporins (Yes vs No)	0,52 (0,19–1,41)	0,2000
Carbapenems (Yes vs No)	1,16 (0,34–3,93)	0,8200
Glycopeptides (Yes vs No)	1,48 (0,46–4,72)	0,5100
Penicillins (Yes vs No)	0,47 (0,12–1,76)	0,2600
DOT	0,99 (0,98–1,01)	0,5200

Table S6: Univariate associations between clinical parameters and risk of relapse. Significant associations are highlighted in bold. AL, acute leukemia; HCT-CI, Hematopoietic Cell Transplantation-specific Comorbidity Index; DRI, Disease Risk Index; H, high; LI, low-intermediate; M, myeloablative; RI, reduced intensity; BM, bone marrow; PB, peripheral blood; HI, haploidentical; MRD, matched related donor; MUD, matched unrelated donor; MMUD, mismatched unrelated donor; DOT, days of therapy; HR, hazard ratio; CI, confidence interval.

Table S7

Univariate associations with TRD		
	HR (95% CI)	**P-value**
Age in years	1,00 (0,95–1,05)	0,9600
Sex (Female vs Male)	0,51 (0,10–2,5)	0,4000
Underlying disease (AL vs Other)	0,49 (0,11–2,07)	0,3300
HCT-CI (1-2 vs 0)	2,45 (0,46–13,2)	0,3000
HCT-CI (≥3 vs 0)	3,36 (0,52–21,8)	0,2000
DRI (H vs LI)	1,00 (0,24–4,25)	1
Conditioning intensity (M vs RI)	0,63 (0,12–3,16)	0,5700
Total body irradiation (Yes vs No)	1,30 (0,32–5,39)	0,7100
T-cell depletion (Yes vs No)	1,40 (0,34–5,73)	0,6400
Graft source (BM vs PB)	0,32 (0,04–2,42)	0,2700
Donor (HI vs MRD)	0,97 (0,15–6,40)	0,9700
Donor (MUD/MMUD vs MRD)	1,43 (0,29–6,94)	0,6600
Cephalosporins (Yes vs No)	0,89 (0,19–4,22)	0,8800
Carbapenems (Yes vs No)	3,74 (0,47–29,6)	0,2100
Glycopeptides (Yes vs No)	1,88 (0,41–8,68)	0,4200
Penicillins (Yes vs No)	3,06 (0,69–13,6)	0,1400
DOT	1,06 (1,03–1,08)	**<0.0001**

Table S7: Univariate associations between clinical parameters and risk of transplant-related death (TRD). Significant associations are highlighted in bold. AL, acute leukemia; HCT-CI, Hematopoietic Cell Transplantation-specific Comorbidity Index; DRI, Disease Risk Index; H, high; LI, low-intermediate; M, myeloablative; RI, reduced intensity; BM, bone marrow; PB, peripheral blood; HI, haploidentical; MRD, matched related donor; MUD, matched unrelated donor; MMUD, mismatched unrelated donor; DOT, days of therapy; HR, hazard ratio; CI, confidence interval.

Table S8

Multivariate associations		
	HR (95% CI)	**P-value**
OS		
OM recovery (R vs NR)	0,09 (0,02–0,35)	**0,0006**
DRI (H vs LI)	6,34 (1,58–25,5)	**0,0092**
Conditioning intensity (M vs RI)	0,16 (0,03–0,88)	**0,0353**
DOT	1,04 (1,00–1,09)	**0,0457**
PFS		
OM recovery (R vs NR)	0,09 (0,02–0,49)	**0,0052**
DRI (H vs LI)	3,61 (0,85–15,2)	0,0807
Relapse		
OM recovery (R vs NR)	0,19 (0,06–0,55)	**0,0025**
DRI (H vs LI)	4,69 (1,47–15,0)	**0,0090**

Table S8: Multivariate associations of oral microbiota recovery and clinical parameters with clinical outcomes. Clinical parameters significantly associated with the outcome in the univariate models (Table S4–6) were used to adjust the significant univariate associations between oral microbiota recovery and clinical outcomes (Table S3). Significant associations are highlighted in bold. OM, oral mucosa; R, recoverers; NR, non-recoverers; DRI, Disease Risk Index; H, high; LI, low-intermediate; M, myeloablative; RI, reduced intensity; DOT, days of therapy; OS, overall survival; PFS, progression-free survival; HR, hazard ratio; CI, confidence interval.

Table S9

	N or median (IQR)		P-value
	R	**NR**	
Age in years	51,1 (15,1)	51,1 (21,5)	0,6268
Sex			0,6942
Female	9	3	
Male	11	6	
Underlying disease			0,4223
Acute leukemia	13	4	
Other	7	5	
HCT-CI			0,4758
0	12	3	
1-2	5	4	
≥3	3	2	
DRI			1
Low-intermediate	12	5	
High	8	4	
Conditioning intensity			0,6942
Reduced intensity	11	6	
Myeloablative	9	3	
Total body irradiation			0,3962
Yes	5	4	
No	15	5	
T-cell depletion			0,427
Yes	9	6	
No	11	3	
Graft source			0,4311
Bone marrow	8	2	
Peripheral blood	12	7	
Donor			0,535
MRD	7	2	
Haploidentical	7	2	
MUD/MMUD	6	5	
Cephalosporins			0,2089
Yes	16	5	
No	4	4	

Carbapenems			0,4118
Yes	11	7	
No	9	2	
Glycopeptides			0,6942
Yes	11	6	
No	9	3	
Penicillins			1
Yes	4	2	
No	16	7	
LOT	14,5 (8,75)	19 (12)	0,2986
DOT	21 (19,25)	22 (23)	0,4639

Table S9: Associations between clinical parameters and oral mucosa (OM) microbiota recovery. The Fisher's exact test and the Mann-Whitney U test were used for categorical and continuous variables, respectively. For categorical variables, the contingency table is shown. For numerical variables, the median value and the interquartile range (IQR) for each group are shown. R, OM recoverers; NR, OM non-recoverers; HCT-CI, Hematopoietic Cell Transplantation-specific Comorbidity Index; DRI, Disease Risk Index; MRD, matched related donor; MUD, matched unrelated donor; MMUD, mismatched unrelated donor; LOT, length of therapy; DOT, days of therapy.

SUPPLEMENTARY FIGURES

Figure S1

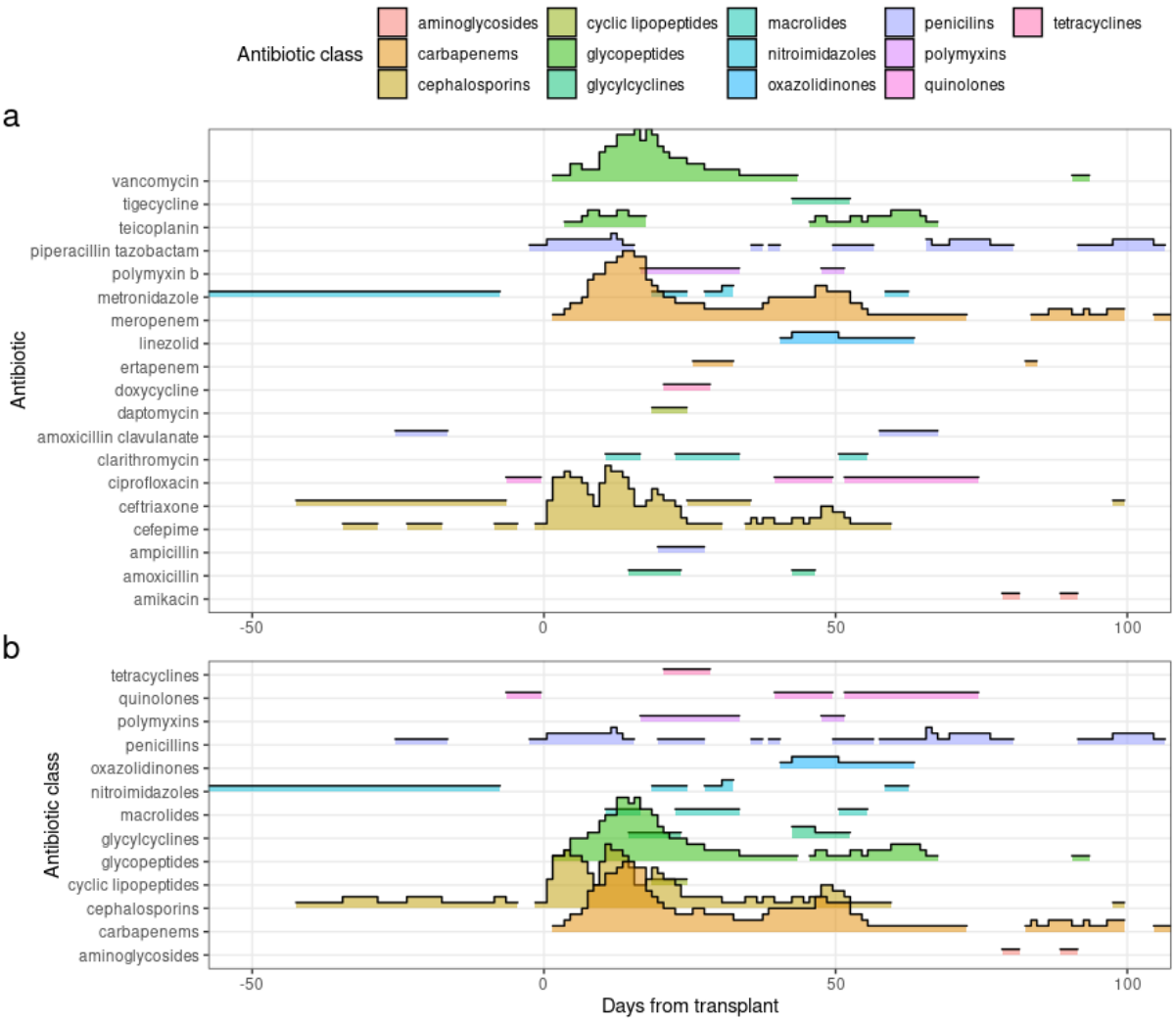


Figure S1: a Ridgeline plot of the antibiotic agents used by the cohort in relation to stem-cell infusion. **b** Ridgeline plot of the antibiotic classes used by the cohort in relation to stem-cell infusion.

Figure S2

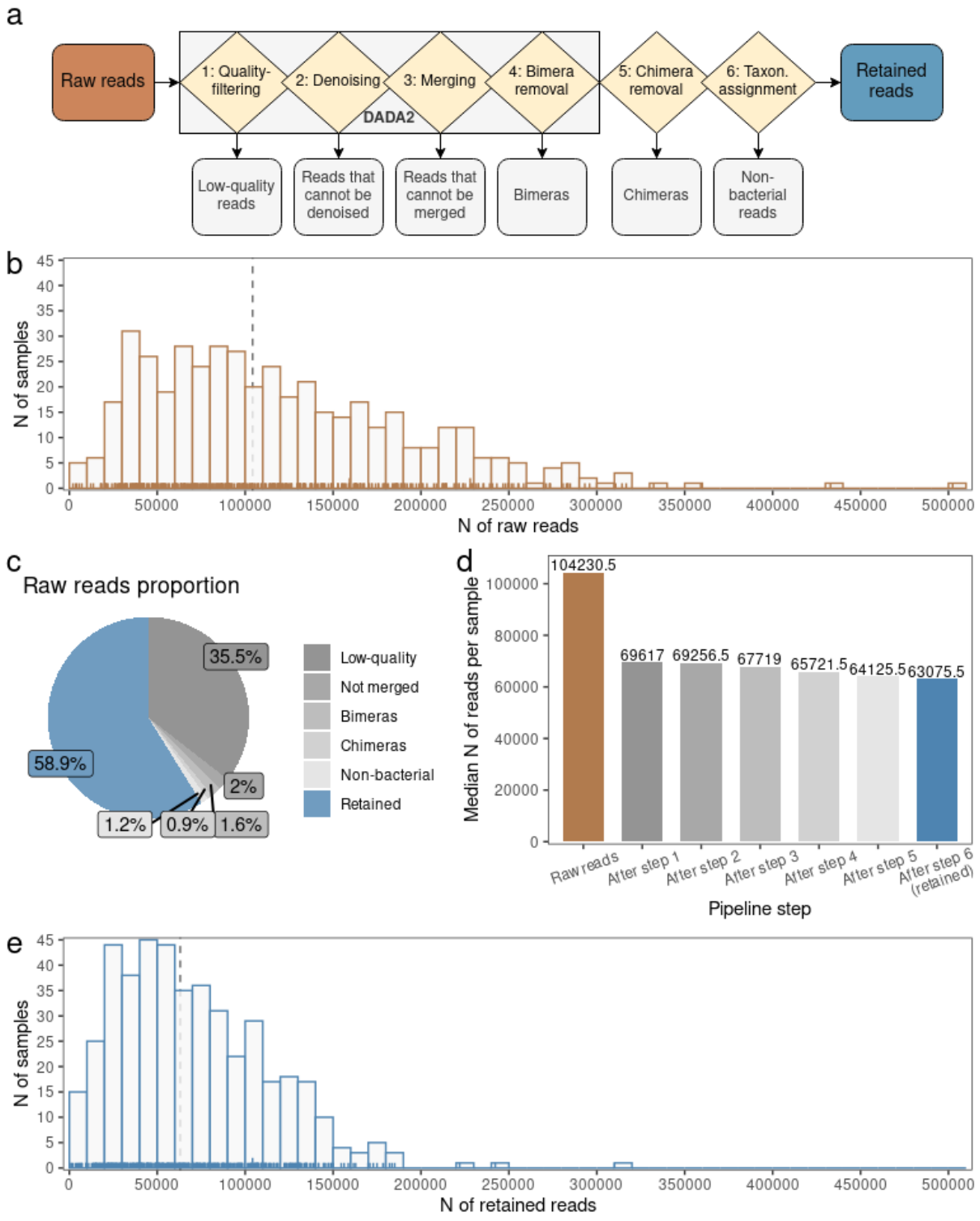


Figure S2: **a** Read processing pipeline scheme. **b** Histogram with the number of raw reads per sample. **c** Proportion of reads discarded at each pipeline step. **d** Median number of reads per sample at each pipeline step. **e** Histogram with the number of retained reads after running the read processing pipeline.

Figure S3

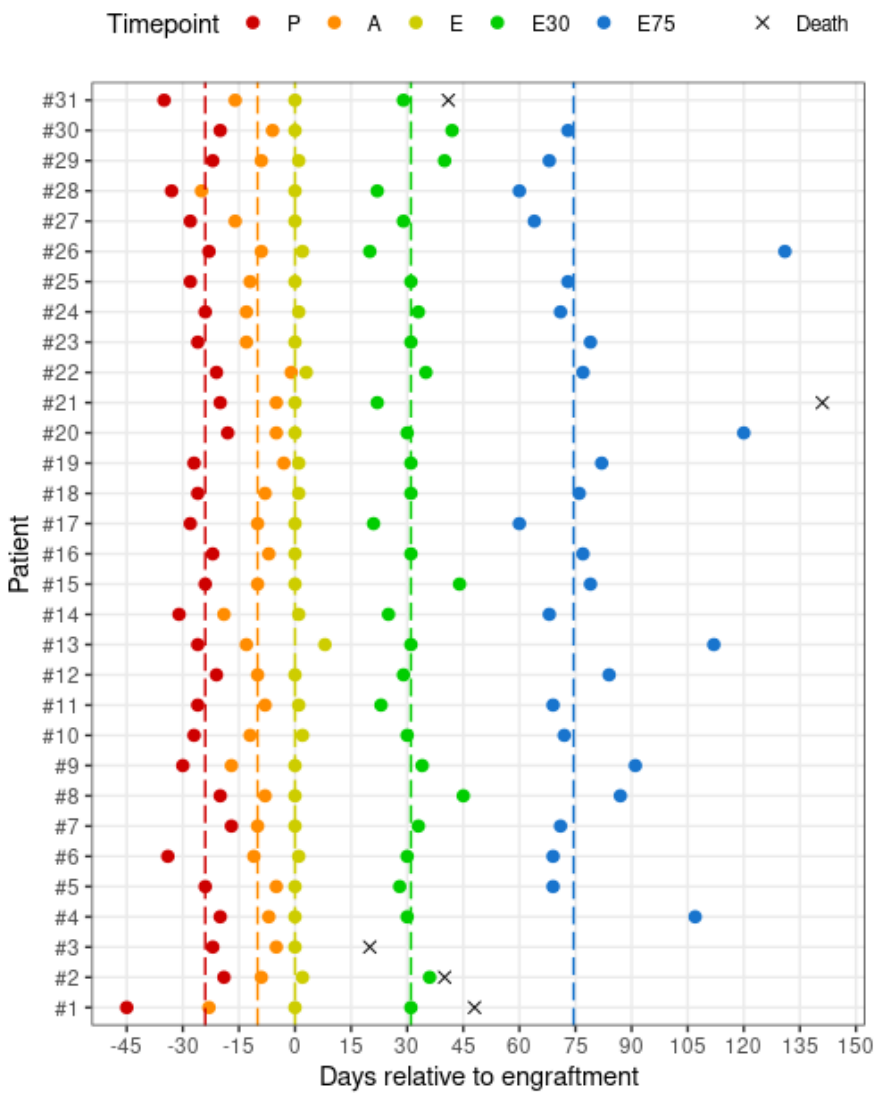


Figure S3: Sampling times for each patient in relation to engraftment day. Vertical dashed lines indicate the median sampling time per timepoint. P, preconditioning; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days after engraftment.

Figure S4

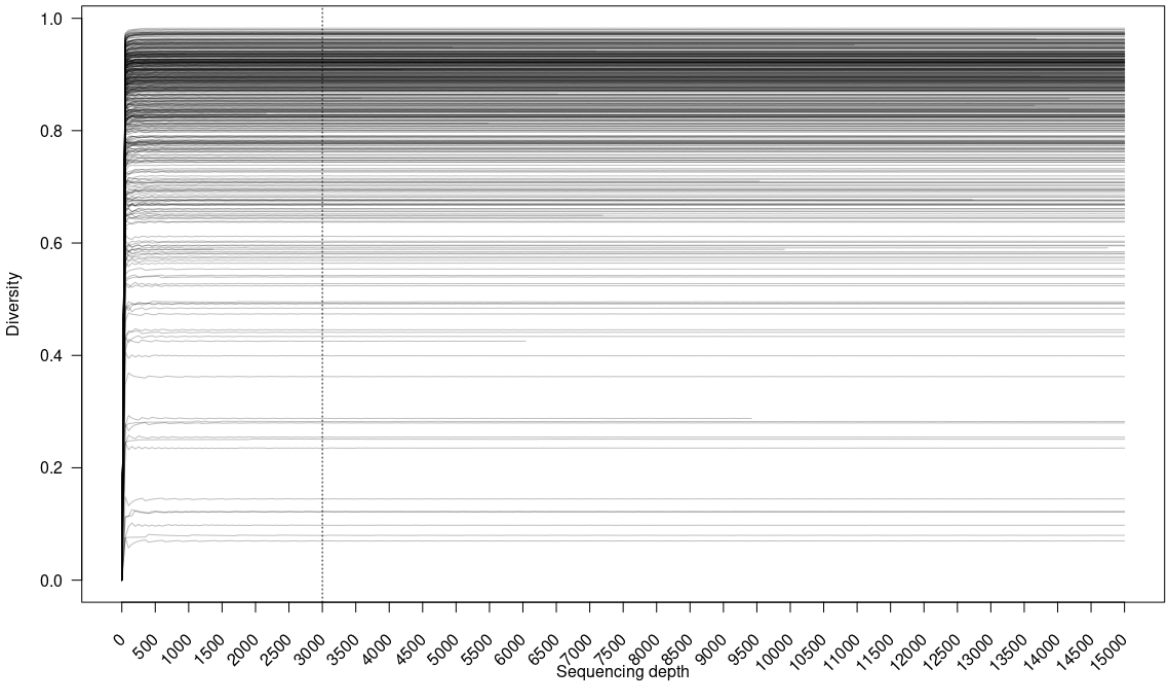


Figure S4: Gini-Simpson rarefaction curves per sample. Reads were selected by scaling with ranked subsampling (SRS) at incremental steps of 50 reads. The plot was limited to 15,000 reads.

Figure S5

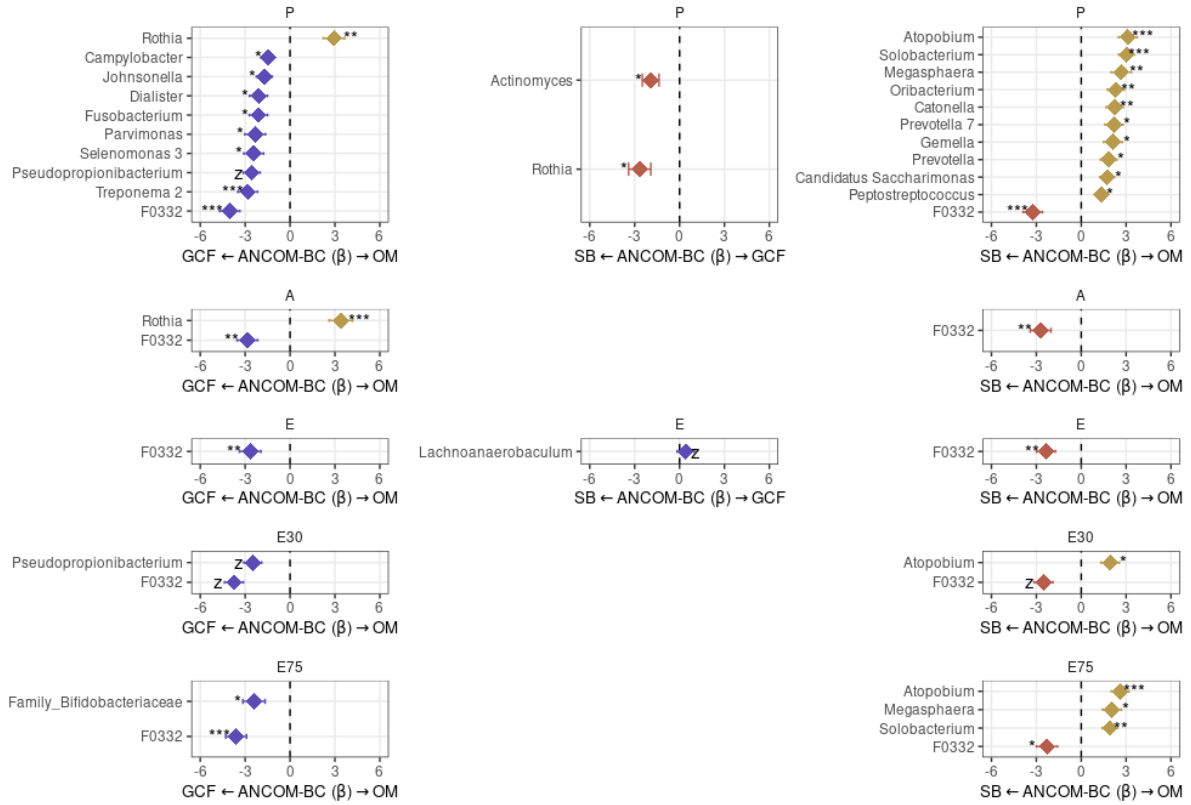


Figure S5: Differentially abundant genera (ANCOM-BC) between oral sites at each timepoint. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; P, preconditioning; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days after engraftment; *, q-value < 0.05; **, q-value < 0.01; ***, q-value < 0.001; z, ANCOM-BC structural zero.

.

Figure S6

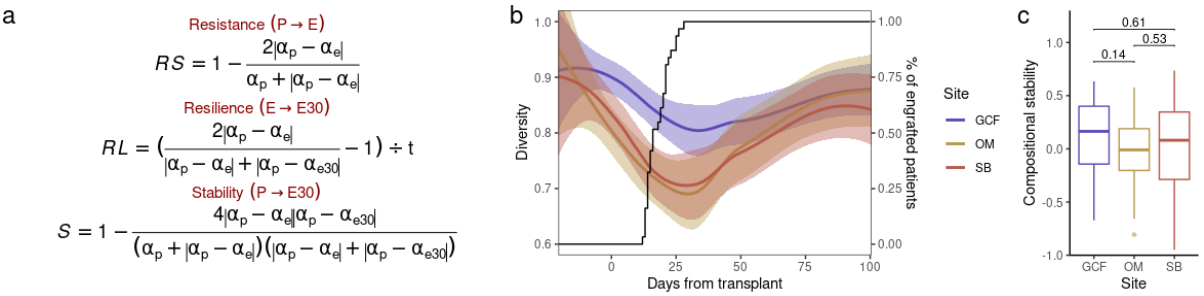


Figure S6: a Diversity resistance, resilience, and stability expressions (see Methods). **b** Smoothed trend-line of diversity (Gini-Simpson) in each oral site (left y-axis) and percentage of engrafted patients (right y-axis) per day from stem-cell infusion. Shaded areas represent 95% confidence intervals. **c** Compositional stability (see Methods) per oral site. Mann-Whitney U test was used. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm.

Figure S7

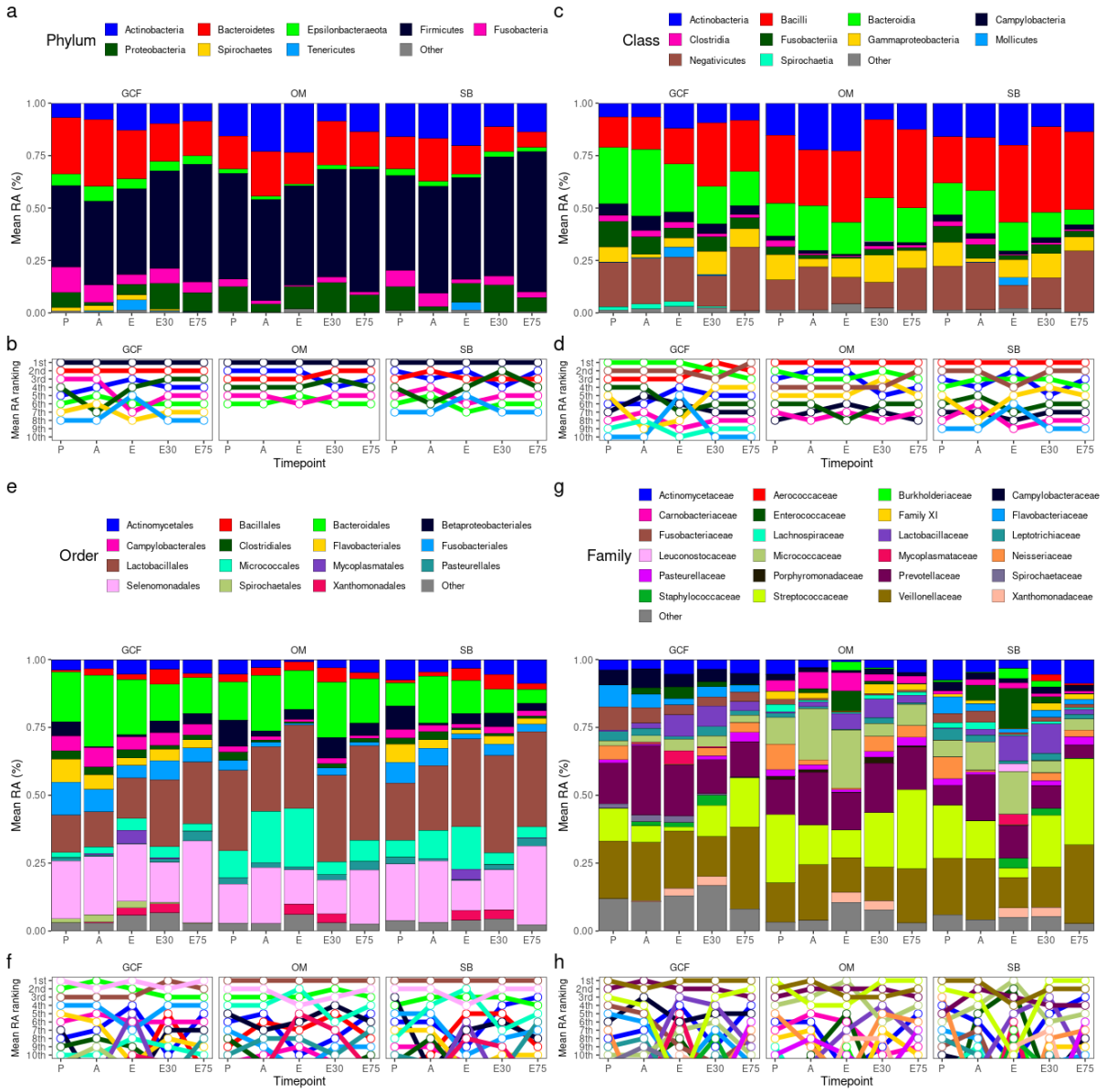


Figure S7: a, c, e, g Mean phyla (a), class (c), order (e), and family (g) relative abundances (RA) per timepoint for each oral site. Taxa with >2% mean RA in any combination of oral site and timepoint are shown. **b, d, f, h** Mean phyla (b), class (d), order (f), and family (g) RA ranking per timepoint for each oral site. Top-10 taxa are shown. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; P, preconditioning; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days after engraftment.

Figure S8

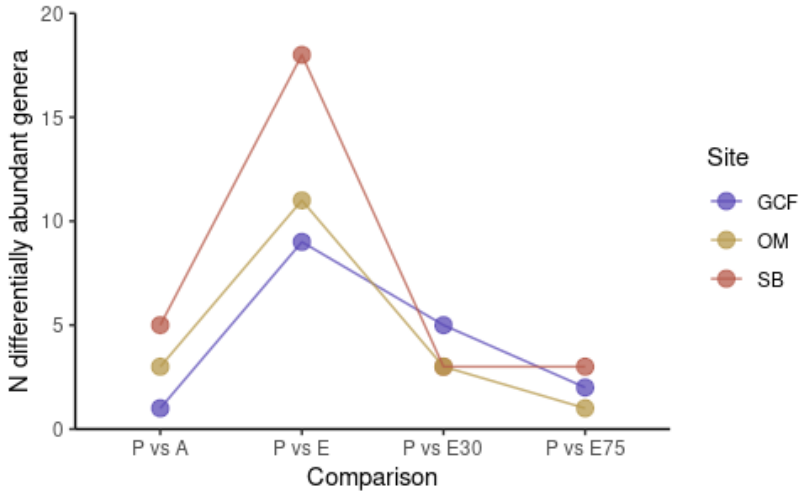


Figure S8: Number of differentially abundant genera (ANCOM-BC) between preconditioning (P) and other timepoints for each site. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days after engraftment.

Figure S9

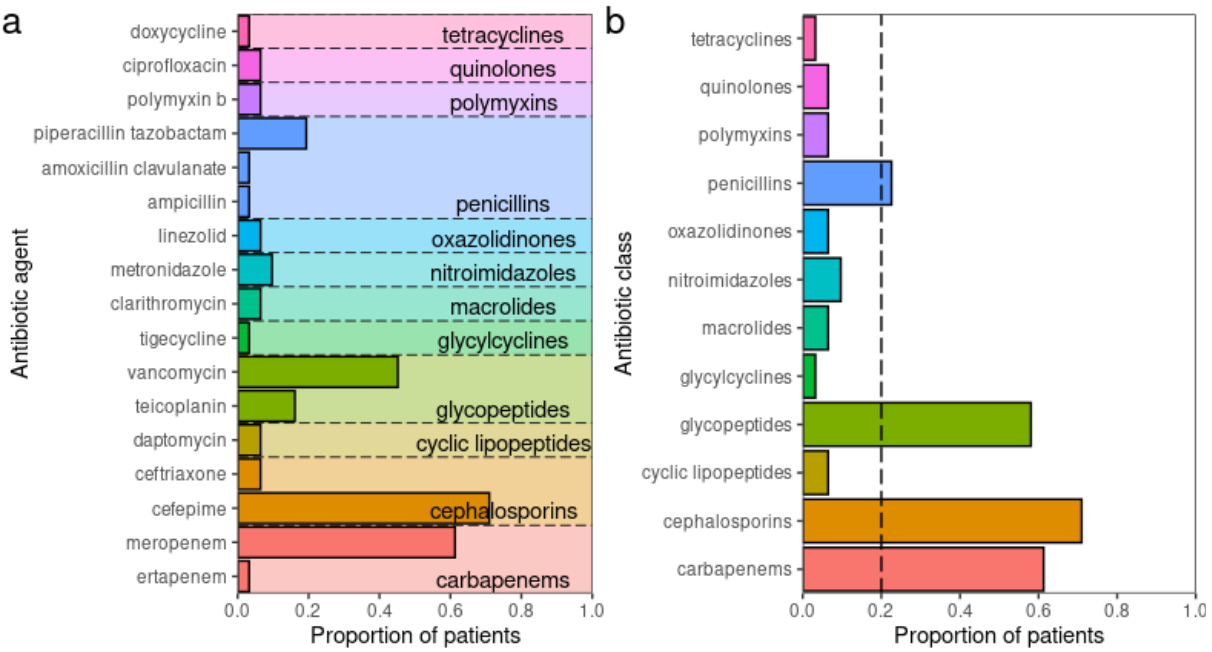


Figure S9: a Proportion of patients using each antibiotic agent between preconditioning (P) and 30 days after engraftment (E30). Respective antibiotic classes are indicated. **b** Proportion of patients using each antibiotic class between P and E30. Vertical dashed line indicates the proportion of 20%.

Figure S10

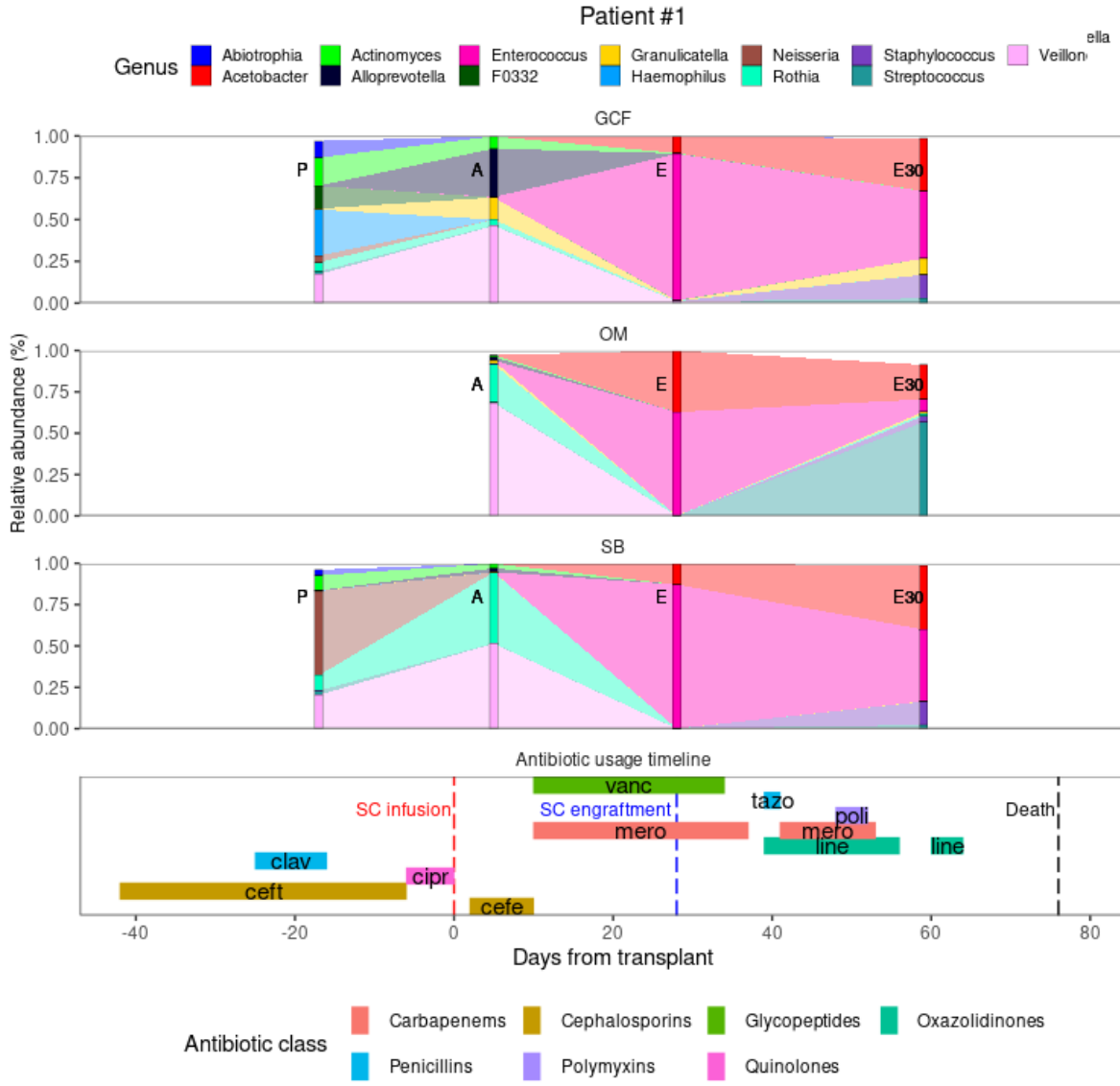


Figure S10: Patient #1: genera relative abundance dynamics for each oral site (top) and antibiotic usage timeline (bottom). Genera with >1% mean relative abundance in any combination of oral site and timepoint are shown. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; A, aplasia; E, engraftment; E75, 75 days after engraftment; SC, stem-cell; vanc, vancomycin; tazo, piperacillin tazobactam; poli, polymyxin B; mero, meropenem; line, linezolid; clav, amoxicillin clavulanate; cipr, ciprofloxacin; ceft, ceftriaxone; cefe, cefepime.

.

Figure S11

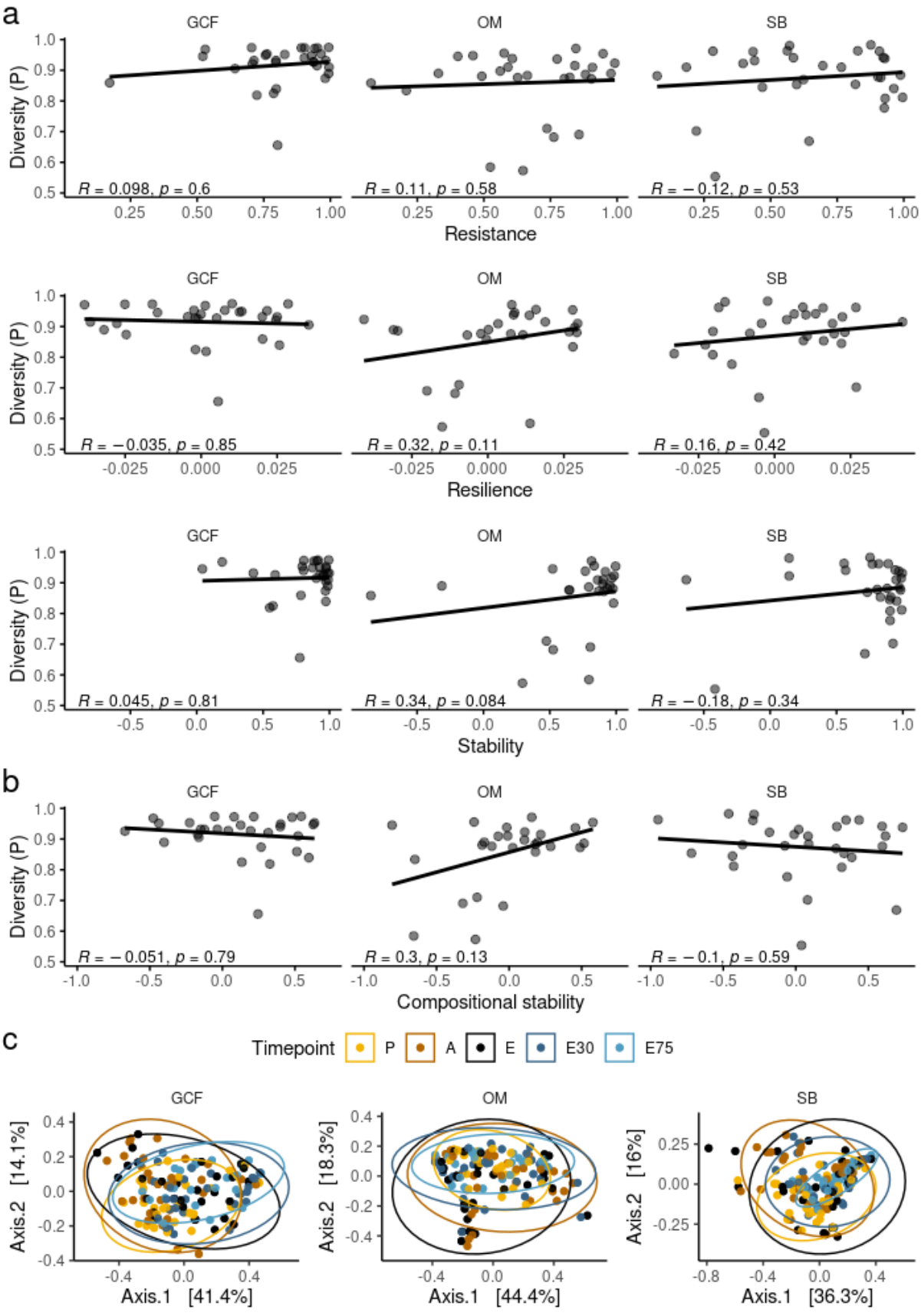

a
GCF
OM
SB
Diversity (P)
R = 0.098, p = 0.6
R = 0.11, p = 0.58
R = −0.12, p = 0.53
Resistance
R = −0.035, p = 0.85
R = 0.32, p = 0.11
R = 0.16, p = 0.42
Resilience
R = 0.045, p = 0.81
R = 0.34, p = 0.084
R = −0.18, p = 0.34
Stability
b
R = −0.051, p = 0.79
R = 0.3, p = 0.13
R = −0.1, p = 0.59
Compositional stability
c
Timepoint
P
A
E
E30
E75
Axis.2 [14.1%]
Axis.1 [41.4%]
Axis.2 [18.3%]
Axis.1 [44.4%]
Axis.2 [16%]
Axis.1 [36.3%]

Figure S11: a Correlation between diversity (Gini-Simpson) at preconditioning (P) and diversity resistance, resilience, or stability for each oral site. Spearman's rank correlation test was used. **b** Correlation between diversity (Gini-Simpson) at P and compositional stability for each oral site. Spearman's rank correlation test was used. **c** Principal Coordinate Analysis (PCoA) of microbiota distances (weighted UniFrac) between timepoints for each oral site. Ellipsoids indicate 95% confidence intervals. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; P, preconditioning; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days after engraftment.

Figure S12

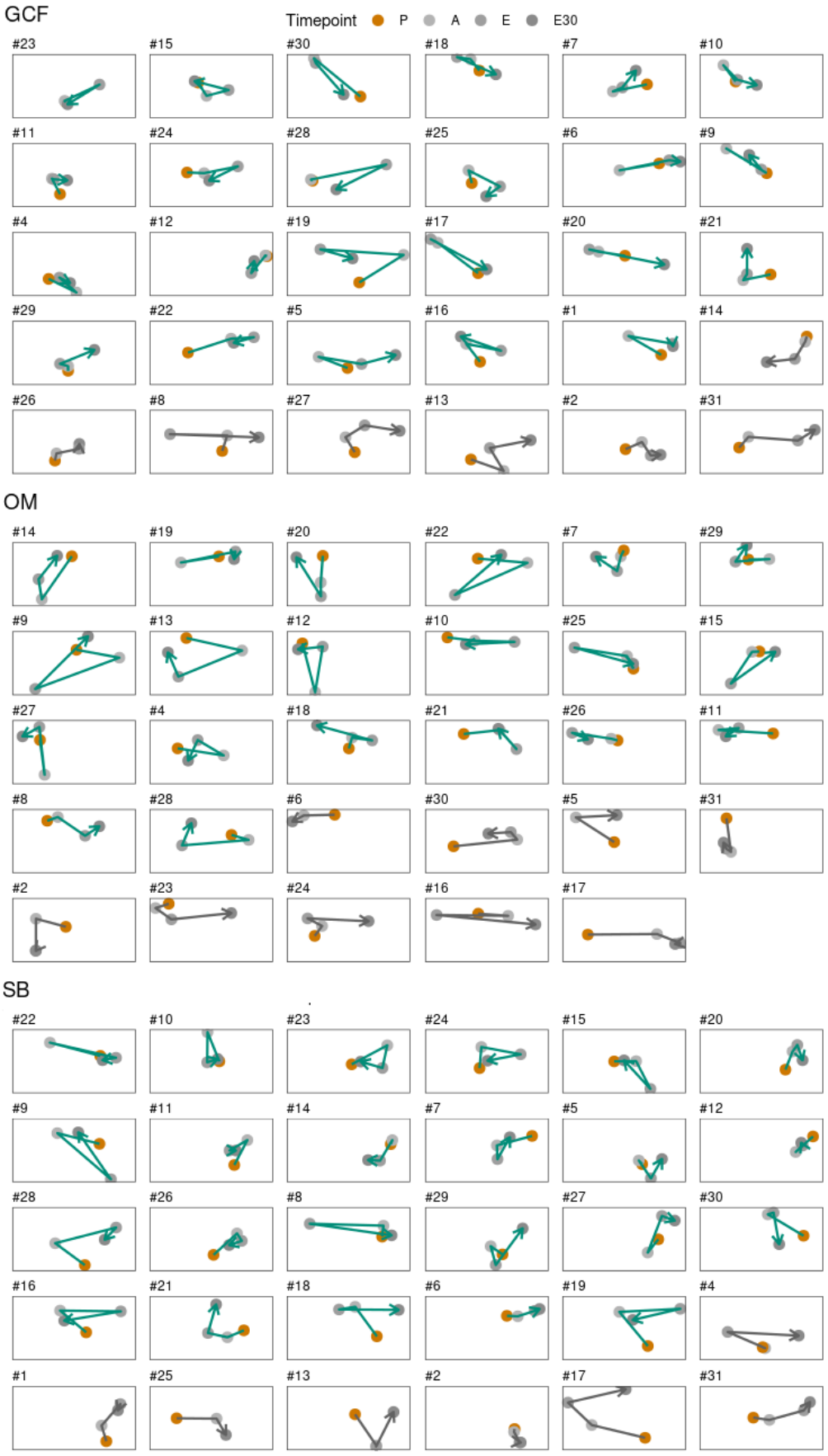

GCF
Timepoint P A E E30
#23 #15 #30 #18 #7 #10
#11 #24 #28 #25 #6 #9
#4 #12 #19 #17 #20 #21
#29 #22 #5 #16 #1 #14
#26 #8 #27 #13 #2 #31
OM
#14 #19 #20 #22 #7 #29
#9 #13 #12 #10 #25 #15
#27 #4 #18 #21 #26 #11
#8 #28 #6 #30 #5 #31
#2 #23 #24 #16 #17
SB
#22 #10 #23 #24 #15 #20
#9 #11 #14 #7 #5 #12
#28 #26 #8 #29 #27 #30
#16 #21 #18 #6 #19 #4
#1 #25 #13 #2 #17 #31

Figure S12: Principal Coordinate Analysis (PCoA) with microbiota trajectories for each patient in each oral site. Recovery trajectories are shown in teal and non-recovery in grey. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm; P, preconditioning; A, aplasia; E, engraftment; E30, 30 days after engraftment; E75, 75 days after engraftment.

Figure S13

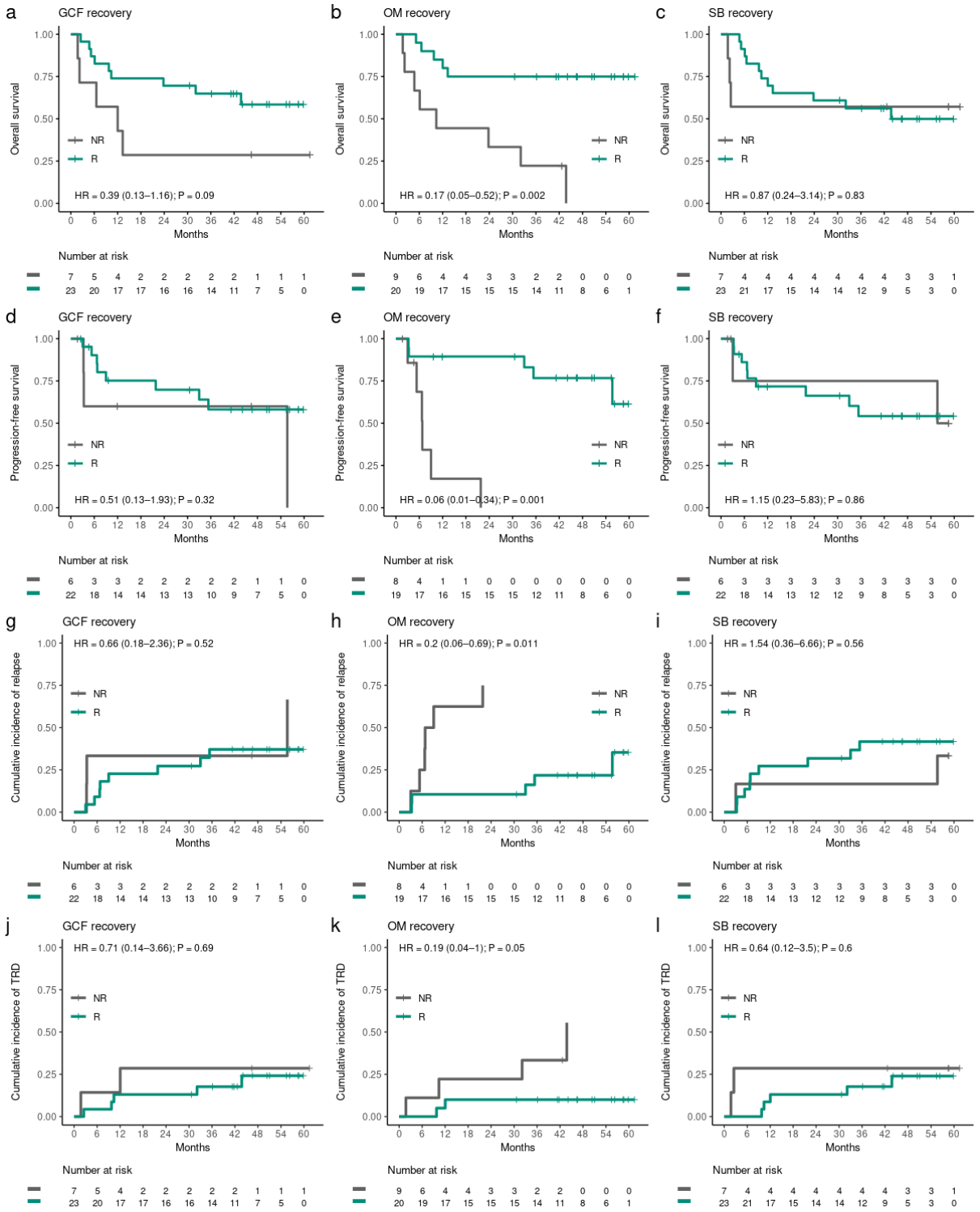


Figure S13: a-c Kaplan-Meier curves comparing overall survival among recoverers (R) and non-recoverers (NR) for each oral site. **d-f** Kaplan-Meier curves comparing progression-free survival among R and NR for each oral site. **g-i** Cumulative incidence curves of relapse among R and NR for each oral site. **j-l** Cumulative incidence curves of transplant-related death (TRD) among R and NR for each oral site. GCF, gingival crevicular fluid; OM, oral mucosa; SB, supragingival biofilm. HR, hazard ratio.

Figure S14

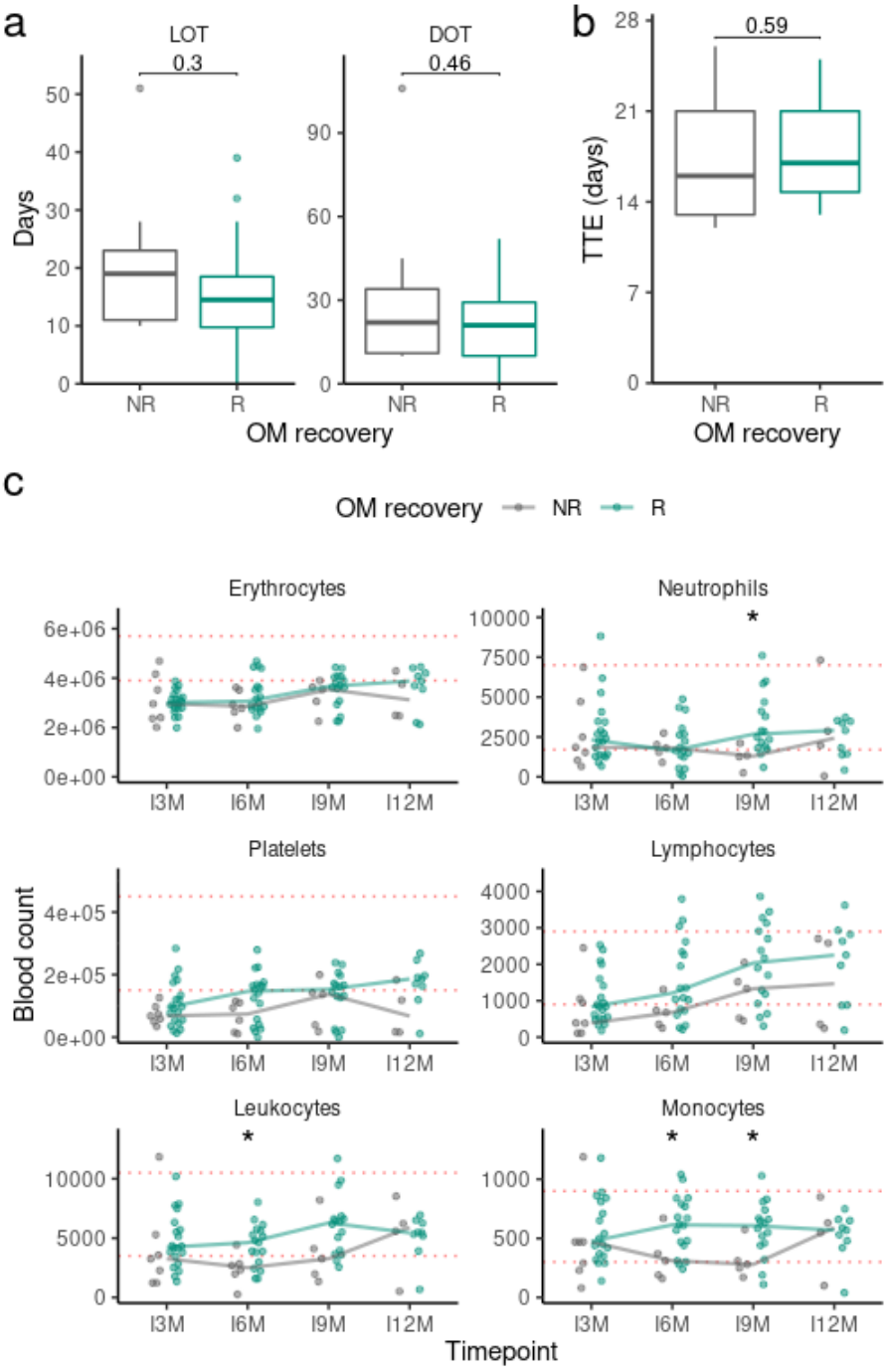


Figure S14: a Time of antibiotic administration (LOT: length of therapy; DOT: days of therapy) among oral mucosa (OM) recoverers (R) and non-recoverers (NR). **b** Time to engraftment (TTE) in days among OM R and NR. **c** Blood cell counts among OM R and NR per timepoint (described below) for each blood cell type. Red dotted horizontal lines indicate normal counts (within reference values). Solid lines indicate median values at each timepoint. Mann-Whitney U test was used. I3M, 3 months after stem-cell infusion; I6M, 6 months after stem-cell infusion; I9M, 9 months after stem-cell infusion; I12M, 12 months after stem-cell infusion; *, P-value < 0.05.

Appendix F: Chapter 4 supplementary methods

Supplementary text to the Materials and methods section of Chapter 4.

Longitudinal analysis at three oral sites links oral microbiota to clinical outcomes in allogeneic hematopoietic stem-cell transplant

SUPPLEMENTARY METHODS

Diversity resistance, resilience, and stability

The resistance (RS) of a microbiota parameter y measures the level of alteration undergone by y during a perturbation l. As proposed by Orwin & Wardle (2004), let y_0 be y at baseline and y_l be y immediately after l, y resistance to l can be measured by:

$$RS = 1 - \frac{2|y_0 - y_l|}{y_0 + |y_0 - y_l|} \quad (1)$$

Rewriting this expression in the context of this study, with y_0 being the diversity at preconditioning (α_p) and y_l the diversity at engraftment (α_e), diversity resistance to allo-HSCT can be calculated for each patient as follows:

$$RS = 1 - \frac{2|\alpha_p - \alpha_e|}{\alpha_p + |\alpha_p - \alpha_e|} \quad (2)$$

Orwin & Wardle further proposed an expression for y resilience (RL), which refers to the rate of change of y towards y_0 after l. Let y_f be y after a period of time t after l (i.e., $t = t_f - t_l$), y resilience at t can be measured by:

$$RL = \left(\frac{2|y_0 - y_l|}{|y_0 - y_l| + |y_0 - y_f|} - 1\right) \div t \quad (3)$$

Rewriting this expression in the context of this study, with y_f being the diversity at 30 days after engraftment (α_{e30}) and t being the interval in days between the engraftment and the 30 days after engraftment sampling times ($t = t_{e30} - t_e$), diversity resilience to allo-HSCT at 30 days after engraftment can be calculated for each patient as follows:

$$RL = \left(\frac{2|\alpha_p - \alpha_e|}{|\alpha_p - \alpha_e| + |\alpha_p - \alpha_{e30}|} - 1\right) \div t \quad (4)$$

Orwin & Wardle did not propose an expression for y stability (S). However, because stability is by definition composed of resistance and resilience (Shade et al., 2012), it follows that $S \propto RS + RL$, so that an expression for y stability can be algebraically derived.

First, it must be noted that RS is a unitless quantity, while RL is a rate ($time^{-1}$). Therefore, in order to combine RS and RL, RL must be multiplied by a factor $b \propto t$ in the new stability expression.

$$S = RS + RL \times b \tag{5}$$

In order to find b, let us expand (5) using (2) and (4):

$$S = 1 - \frac{2|y_0 - y_l|}{y_0 + |y_0 - y_l|} + \left(\frac{2|y_0 - y_l|}{|y_0 - y_l| + |y_0 - y_f|} - 1\right) \times \frac{b}{t} \tag{6}$$

Now, it is reasonable to impose that when y recovers to baseline levels after t, y stability is maximum: $y_f = y_0 \Rightarrow S = 1$. Using this information in (6), it is possible to solve the equation for b:

$$1 = 1 - \frac{2|y_0 - y_l|}{y_0 + |y_0 - y_l|} + \left(\frac{2|y_0 - y_l|}{|y_0 - y_l| + |y_0 - y_0|} - 1\right) \times \frac{b}{t} \tag{7}$$

Resulting that

$$b = \frac{2t|y_0 - y_l|}{y_0 + |y_0 - y_l|} \tag{8}$$

As desired, $b \propto t$. Replacing b in (6), it results that y stability is given by:

$$S = 1 - \frac{4|y_0 - y_l||y_0 - y_f|}{(y_0 + |y_0 - y_l|)(|y_0 - y_l| + |y_0 - y_f|)} \tag{9}$$

Which in the context of this study can be rewritten as follows:

$$S = 1 - \frac{4|\alpha_p - \alpha_e||\alpha_p - \alpha_{e30}|}{(\alpha_p + |\alpha_p - \alpha_e|)(|\alpha_p - \alpha_e| + |\alpha_p - \alpha_{e30}|)} \tag{10}$$

Compositional stability

In analogy to diversity stability, which is calculated based on α_p, α_e, and α_{e30}, compositional stability is evaluated by considering preconditioning, engraftment, and 30 days

after engraftment samples. Specifically, let C be the area in the compositional space enclosed by the convex hull of these samples. Compositional stability is calculated as $1 - C$.

Multiple linear regression

Multiple linear regression was used to evaluate whether antibiotic usage parameters predicted diversity and compositional stability (S). Separate models for each site were run with the *lm* function from the *stat* R package. Due to the high collinearity between DOT and LOT, only DOT and the main antibiotic classes were included as predictors in the models: $S \sim cephalosporins + carbapenems + glycopeptides + penicillins + DOT$.

Taxonomic nomenclature homogenization

Either due to a lack of taxonomic resolution or incomplete annotated taxonomic information in the 16S rRNA database used, assigned taxonomies may contain generic proxies at low taxonomic ranks, such as “uncultured” or “s_”, with the latter indicating a lack of taxonomic resolution to identify the taxon at species level. As these generic proxies do not contribute with taxonomic information, these and other similar entries (namely, “uncultured”, “sp.”, “metagenome” and “human_gut”) were homogenized by replacing them with the lowest taxonomic rank with complete nomenclature and the corresponding taxon.

Blood count data

Complete blood count data spanning the first year after transplant was collected retrospectively from the blood test results database of our institution. Results from dates close to the oral sampling phases (aplasia not included due to the necessarily low counts for all patients) and 3, 6, 9, and 12 months after stem-cell infusion were collected. The median gap between these target periods and the actual blood test dates was 0 for all target periods. Outlier counts were identified by running the results collected altogether (independently of the period) through the Grubbs’ test and removed. Cell counts too low to be accurately quantified were set to 1.

LIST OF ATTACHMENTS

Attachment A – Scaling with ranked subsampling paper

Heidrich et al., 2021, *Applied Sciences*.

Attachment B – Actinomycosis case report

Bruno et al., 2023, *BMC oral health*.

Attachment C – Cohort's clinical and microbiota characteristics

Attachment D – Academic *Curriculum Vitae*

ATTACHMENTS

Attachment A – Scaling with ranked subsampling paper

Heidrich et al., 2021, *Applied Sciences*.

applied sciences

MDPI

Article

'SRS' R Package and 'q2-srs' QIIME 2 Plugin: Normalization of Microbiome Data Using Scaling with Ranked Subsampling (SRS)

Vitor Heidrich [1,2,*,†], Petr Karlovsky [3,†] and Lukas Beule [4,*,†]

1 Centro de Oncologia Molecular, Hospital Sírio-Libanês, São Paulo 01308-060, Brazil
2 Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo 05508-900, Brazil
3 Molecular Phytopathology and Mycotoxin Research, Faculty of Agricultural Sciences, University of Goettingen, 37077 Goettingen, Germany; pkarlov@gwdg.de
4 Julius Kühn Institute (JKI)—Federal Research Centre for Cultivated Plants, Institute for Ecological Chemistry, Plant Analysis and Stored Product Protection, 14195 Berlin, Germany
* Correspondence: vheidrich@mochsl.org.br (V.H.); lukas.beule@julius-kuehn.de (L.B.)
† These authors contributed equally to this work.

Abstract: Several ecological data types, especially microbiome count data, are commonly sample-wise normalized before analysis to correct for sampling bias and other technical artifacts. Recently, we developed an algorithm for the normalization of ecological count data called 'scaling with ranked subsampling (SRS)', which surpasses the widely adopted 'rarefying' (random subsampling without replacement) in reproducibility and in safeguarding the original community structure. Here, we describe an implementation of the SRS algorithm in the 'SRS' R package and the 'q2-srs' QIIME 2 plugin. We also provide accessory functions for dataset exploration to guide the choice of parameters for SRS.

Keywords: scaling with ranked subsampling (SRS); R package; QIIME 2 plugin; microbial ecology; microbiome analysis; bioinformatics; normalization

Citation: Heidrich, V.; Karlovsky, P.; Beule, L. 'SRS' R Package and 'q2-srs' QIIME 2 Plugin: Normalization of Microbiome Data Using Scaling with Ranked Subsampling (SRS). *Appl. Sci.* **2021**, *11*, 11473. https://doi.org/10.3390/app112311473

Academic Editor: Giovanna Donnarumma

Received: 1 November 2021
Accepted: 29 November 2021
Published: 3 December 2021

1. Introduction

High-throughput sequencing of taxonomically informative loci of microbial genomes by amplicon sequencing dramatically improved our understanding of microbial communities. Microbiome research expanded into all microbial habitats on earth, including the human intestine (e.g., [1]), soils (e.g., [2]), and deep-sea sediments (e.g., [3]). A range of bioinformatic tools and platforms as well as reference databases have been developed to enable the extraction of biological insight from the large amounts of data generated by multiplexed amplicon sequencing. The number of sequence counts per sample (sequencing depth) obtained from such sequencing runs can vary by orders of magnitude [4]. Those variations are technical artifacts caused by unequal pooling of samples prior to multiplexed sequencing runs and varying sequencing efficiencies. This contributes to biased estimates of several parameters assessed in microbiome analysis, such as alpha and beta diversity, and relative abundances of taxa.

Fortunately, variations in sequencing depth can be computationally compensated by normalization of sequence counts per sample, a step that has become essential in processing amplicon sequencing data. Traditionally, rarefying was used for this. In 2014, however, McMurdie and Holmes [4] demonstrated that rarefying is statistically inadmissible for the normalization of microbiome count data. Although the work of McMurdie and Holmes [4] received a lot of attention, rarefying is still frequently used in current microbiome studies, likely due to a lack of suitable alternatives. This motivated us to develop the scaling with ranked subsampling (SRS) algorithm, which outperforms rarefying for diversity analysis and relative abundance estimates, as recently shown [5].

Because unequal sampling depth is a problem inherent not only to microbiome research but to all studies based on ecological count data, we introduced SRS as a tool for the normalization of ecological count data and successfully applied it to microbiome analysis [5]. Yet, the implementation of SRS in bioinformatic platforms was missing.

In this work, we introduce an R package ('SRS') and a QIIME 2 plugin ('q2-srs') for the normalization of microbiome count data using SRS. Furthermore, we improve the original SRS algorithm and add features to visualize and evaluate the results. Finally, we provide an example for microbial ecologists that aim to normalize microbiome count data obtained by amplicon sequencing.

2. Theory

Ecological surveys and microbiome analysis by amplicon sequencing yield so-called species count data, which typically populate matrices with species represented by rows and samples represented by columns. Species are taxa (e.g., genera or binomial names), nucleotide sequences (ASVs), or sets of sequences grouped by similarity (OTUs). Samples are specimens of material (e.g., water or soil) or individuals or their parts (e.g., a plant or a bird intestine) distinguished by space-time attributes or treatments. The matrices are filled with nonnegative integers, which are designated counts. Analysis of count data is also used in other research fields such as bibliographic analysis, sociology of crime, and epidemiology of rare diseases. We suggest that study areas unrelated to ecology may also benefit from concepts developed for species count data in ecology.

The purpose of normalization is to convert a species count matrix into a normalized matrix, which has an equal dimension and is filled with integers such that the sum of counts of all species in each sample equals a pre-defined value, which we designate $C_{\min}$, and the structure of the normalized matrix approximates the structure of the original matrix. The criteria for the approximation may differ but a key principle is that relative frequencies of counts of the normalized matrix are as close as possible to the relative frequencies of counts in the original matrix. A relative frequency is obtained by dividing the count for a particular species in a particular sample by the sum of counts for all species in that sample. Different implementations of the criterion of matching relative frequencies are conceivable. The simplest option is to construct a normalization matrix minimizing the sum of absolute values of pairwise differences between the relative frequencies. This approach, however, ignores the effect of sampling error on the accuracy of relative frequencies. In the first approximation, the coefficient of variation of a count is proportional to the inverse of the square root of the count. Therefore, frequencies may be weighted by the inverse square root of counts. Depending on the purpose of the study, for instance, regarding the importance of rare species, other weighing may be more adequate.

Regardless of the criterion used to minimize the differences among sets of relative frequencies of species, which are colloquially referred to as "population structure", the task is an optimization problem under integer constraint, which is a special kind of integer programming problem. Let assume sampling data for J species in K samples with counts collected in a $J \times K$ matrix. Let $C_{(j,k)}$ denote the count of species j in sample k and $F_{(j,k)}$ the relative frequency of species j in sample k:

$$F_{(j,k)} = \frac{C_{(j,k)}}{\sum_{i=1}^{J} C_{(i,k)}}.$$

Let $C_{(j,k)norm}$ denote the normalized count of species j in sample k. The constraint of equal total species count per sample implies

$$\sum_{i=1}^{J} C_{(i,1)norm} = \sum_{i=1}^{J} C_{(i,2)norm} = \ldots = \sum_{i=1}^{J} C_{(i,K)norm} = C_{\min}.$$

Conversion of $C_{(j,k)}$ into $C_{(j,k)norm}$ satisfying this constraint and leading to frequencies derived from the normalized matrix

$$F_{(j,k)norm} = \frac{C_{(j,k)norm}}{\sum_{i=1}^{J} C_{(i,k)norm}}$$

as close as possible to the original frequencies $F_{(j,k)}$ is the purpose of normalization. The normalized matrix minimizes the sum of differences between original frequencies and frequencies derived from the normalized counts, while frequencies may be weighted by factor r and the differences may be raised to power s:

$$\sum_{i=1}^{J} r\left|F_{(i,k)} - F_{(i,k)norm}\right|^{s}.$$

As a weighting factor, 1 can be used for equal weights or $\sqrt{C_{(i,k)}}$ to compensate for differences in the sampling error. As a power s, 1 can be used for absolute differences or 2 in line with the least-square concept. Weighing or raising the difference to a power, however, rarely affects the results, as shown by the following example. Let $C_{(k)}$ be a column vector of species counts for sample k and $C_{(k)}^{T}$ its transposition into a row vector:

$$C_{(k)}^{T} = (2,\ 4,\ 30,\ 600,\ 0,\ 27,\ 231).$$

The total species count in sample k is 894. After normalization to C_{min} = 100, the same normalized counts are obtained for all combinations of optimization parameters:

$$r \in \left\{1, \sqrt{C_{(i,k)}}\right\}, s \in \{1,2\}: \quad C_{(k)}^{T} = (0,\ 1,\ 3,\ 67,\ 0,\ 3,\ 26).$$

The normalization was conducted by comparing 7-tuples of nonnegative integers such that each term varied from zero to

$$C_{(j,k)}\frac{100}{894} + 5 \tag{1}$$

while the sum of terms was C_{min}. Exhaustive enumeration of this kind is not feasible for real-world data. In 2014, Cont and Heidari suggested an algorithm solving this optimization problem with the complexity O(n log n), n being the number of species, but their preprint has not been subjected to a peer review yet [6]. The SRS algorithm [5], which has the complexity of O(n), generated the same results in this example.

SRS is an empirical algorithm that does not rely on comparison of relative frequencies of raw and normalized counts. On real as well as simulated count data, SRS was, however, shown to perform substantially better than normalization by rarefying [5].

3. Method

3.1. Principle of SRS

The SRS algorithm performs scaling followed by ranked subsampling.

1. Scaling: feature counts (such as OTUs (operational taxonomic units), ASVs (amplicon sequence variants), or clades) are scaled sample-wise so that the sum of the scaled counts (C_{scaled}) for each sample is equal to the desired number of counts (C_{min}).
2. Ranked subsampling: the scaling step produces fractional values that must be converted into counts (integers). To do this, the C_{scaled} for each feature is split into the floor (C_{int}) and fractional part (C_{frac}) of C_{scaled}. Because $C_{min} = \Sigma C_{scaled} = \Sigma C_{int} + \Sigma C_{frac}$, it follows that $C_{min} \geq \Sigma C_{int}$. Therefore, ΔC C_{frac} values (where $\Delta C = C_{min} - \Sigma C_{int}$) must be converted into additional counts (integers) so that C_{min} can be reached. To do so, C_{frac} values are ranked. Next, from the highest to the lowest rank, a count for

each feature is added until ΔC counts have been added. After this step, all samples will have been normalized to C_{min} counts.

3. Special cases: (i) when C_{frac} values involved in picking ΔC counts share the same rank across features, the counts are added for features based on the respective C_{int} ranks; (ii) when both C_{frac} and its respective C_{int} values involved in picking ΔC counts share the same ranks across features, the counts are assigned randomly (without replacement). The specification of the seed that initializes the random process enables reproducible results.

3.2. 'SRS' R Package

3.2.1. *SRS*-Function

The SRS algorithm was implemented as the *SRS*-function in the 'SRS' R package (https://CRAN.R-project.org/package=SRS (accessed on 1 November 2021)). As an extension of the original SRS algorithm published by Beule and Karlovsky [5], SRS as implemented in version 0.2.2 of the package enables reproducible results in case SRS uses random subsampling without replacement by specifying the seed that initializes the random process (*set.seed*). The default settings of the *SRS*-function (as of version 0.2.2) are:

$$SRS(data, C_{min}, set\_seed = TRUE, seed = 1)$$

where *data* is the input data (e.g., an OTU table), with samples distributed column-wise, C_{min} is the number of counts to which all samples will be normalized (C_{min}), *set_seed* enables the use of the *set.seed*-function, and seed specifies the seed used by *set.seed* to initialize the random process.

3.2.2. *SRScurve*-Function

In analogy to rarefaction curves, the *SRScurve*-function of the 'SRS' R package plots the number of observed unique features (observed richness) against the number of sampled counts utilizing the *SRS*-function (SRS curves). In addition to observed richness, different alpha diversity metrics (Shannon, Simpson, and inverse Simpson indices as implemented in the *diversity*-function of the 'vegan' R package [7]) can be selected to generate SRS curves. Furthermore, *SRScurve* allows a direct comparison to averaged repeated rarefying. The default settings of the *SRScurve*-function (as of version 0.2.2) are:

$$\begin{aligned} SRScurve(&data, metric = \text{“}richness\text{”}, step = 50, sample = 0,\\ &max.sample.size = 0, rarefy.comparison = FALSE,\\ &rarefy.repeats = 10, rarefy.comparison.legend = FALSE,\\ &xlab = \text{“}sample\ size\text{”}, ylab = \text{“}richness\text{”}, label = FALSE,\\ &col, lty, \ldots\ ) \end{aligned}$$

where *data* is the input data (e.g., an OTU table), metric selects the alpha diversity metric to be plotted ("*richness*" = observed richness; "*shannon*" = Shannon index; "*simpson*" = Simpson index; "*invsimpson*" = inverse Simpson index), *step* specifies the step size at which the alpha diversity metric are sampled, *sample* specifies the cutoff-level to visualize trade-offs between cutoff-level and alpha diversity, *max.sample.size* specifies the maximum sample size to which SRS curves are drawn (the default does not limit the maximum sample size), *rarefy.comparison* enables comparison of SRS curves to rarefying, *rarefy.repeats* specifies the number of repeats used for rarefying, *rarefy.comparison.legend*, *xlab*, *ylab*, *label*, *col*, *lty*, and ... are graphical parameters.

3.2.3. *SRS.shiny.app*-Function

The *SRS.shiny.app*-function of the 'SRS' R package launches a Shiny app for SRS in the default web browser to determine C_{min}. The app utilizes the *SRScurve*-function and enables the selection of four diversity metrics (see metric in *SRScurve*) that will be returned at different C_{min}. The selection of C_{min} is interactive through a slider or an interconnected numeric text field. In response to the selected C_{min}, the app returns

1. a rug plot that shows the distribution of the number of counts per sample and displays discarded samples as well as summary statistics (including a list of discarded samples and descriptive statistics of the global feature richness and selected alpha diversity metric of the input dataset) in response to the selected C_{min} (Figure 1A),
2. a plot of SRS curves (*SRScurve*-function) that respond to the selected step size (*step*) and maximum sample size (*max.sample.size*) (Figure 1B), and
3. an interactive table with sample names and the number of counts per sample as well as the initial diversity (non-normalized), retained diversity (normalized), %retained diversity (normalized), and %discarded diversity (normalized) of the selected alpha diversity metric in response to the selected C_{min} (Figure 1C).

The default C_{min} of the app is the lowest total number of counts per sample in the input data (no samples are discarded by default), which can be restored within the app using the *reset* C_{min}-button. The default maximum sample size equals the default setting of C_{min} and can be restored using the *reset max. sample size*-button. The default step size for SRS curves is 1000. The default setting of the *SRS.shiny.app*-function (as of version 0.2.2) is:

SRS.shiny.app(*data*)

where *data* is the input data (e.g., an OTU table).

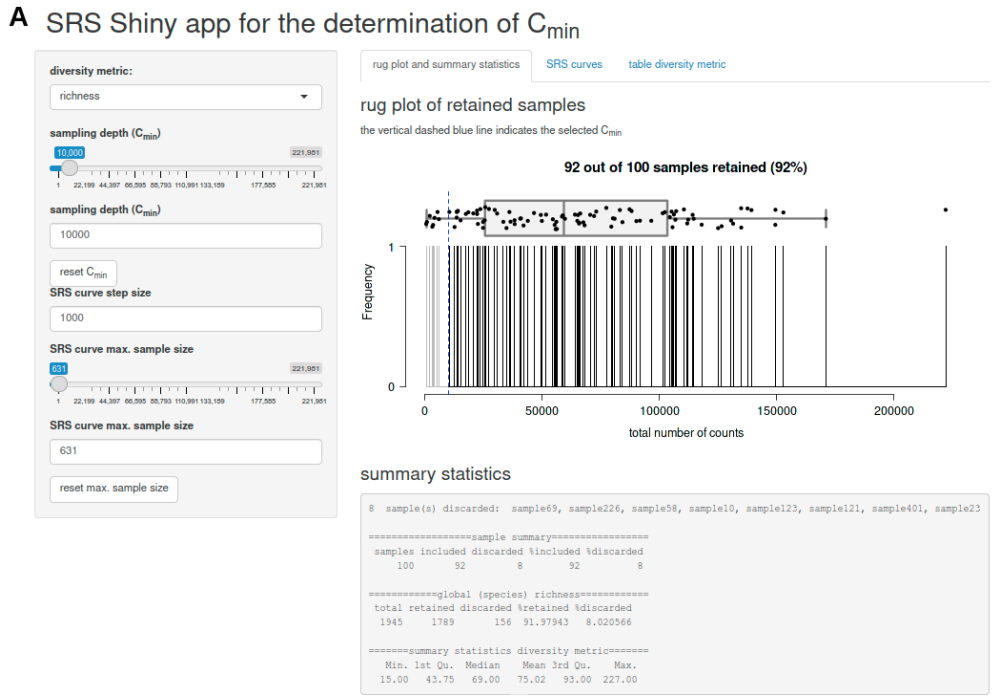


Figure 1. *Cont.*

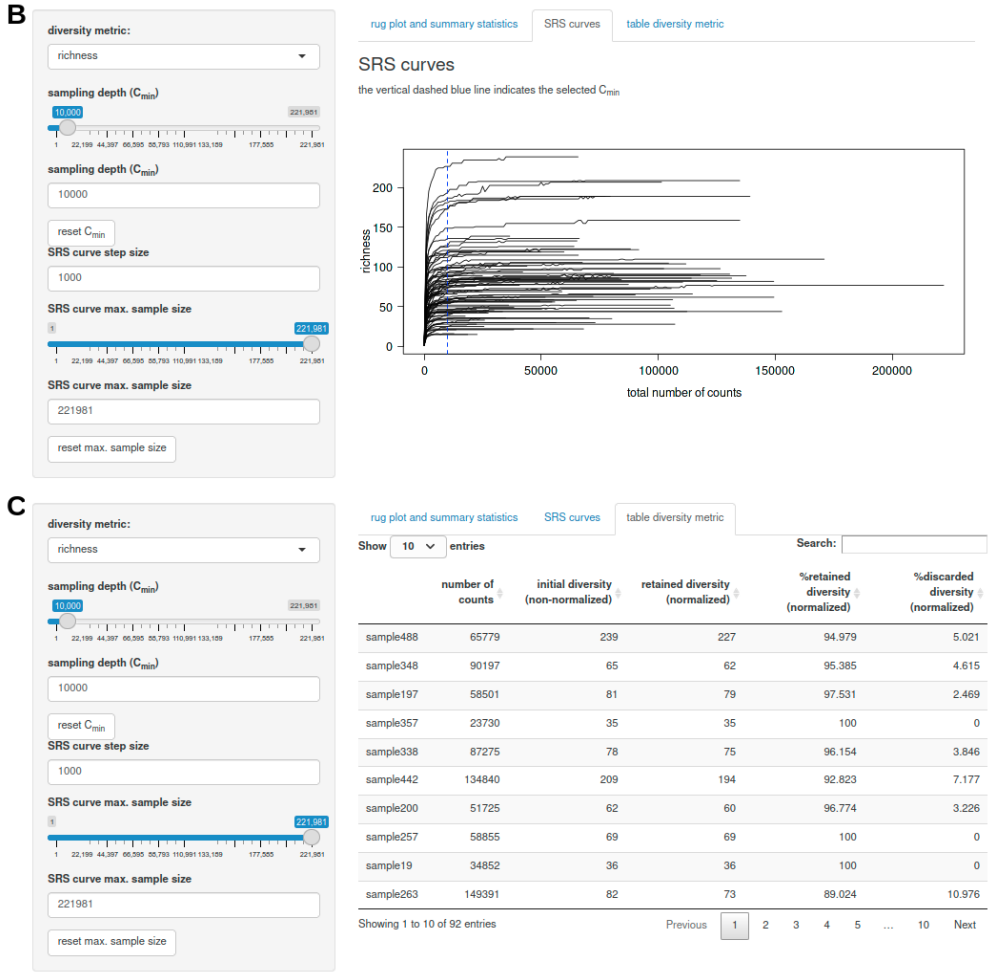


Figure 1. User interface of the Shiny app for SRS (*SRS.shiny.app*-function of the 'SRS' R package version 0.2.2). (**A**) Rug plot showing the distribution of the number of counts per sample, discarded samples, and summary statistics; (**B**) plot showing SRS curves; (**C**) interactive table with sample names, the number of counts per sample, and summary statistics for the diversity metric.

3.3. 'q2-srs' QIIME 2 Plugin

The 'q2-srs' QIIME 2 plugin (https://library.qiime2.org/plugins/q2-srs (accessed on 1 November 2021)) allows straightforward SRS algorithm incorporation into QIIME 2 pipelines. Because its implementation wraps up the 'SRS' R package, its functionalities are analogous to those presented in the previous section.

Specifically, 'q2-srs' features the QIIME 2 actions *SRS* and *SRScurve*, which mirror the 'SRS' R package *SRS*-function and *SRScurve*-function, respectively, with the same behaviour and default parameters as presented in the previous section. The command-line interface commands for the use of the *SRS*- and *SRScurve*-functions within QIIME 2 environment are, respectively, *qiime srs SRS* and *qiime srs SRScurve*. Finally, despite the 'q2-srs' QIIME 2 plugin not having a *SRS.shiny.app*-function counterpart, an online version of the SRS Shiny app (https://vitorheidrich.shinyapps.io/srsshinyapp/ (accessed on 1 November 2021)) is provided for 'q2-srs' users.

4. Results and Discussion

In both the R package as well as the QIIME 2 plugin, we modified the original SRS algorithm by specifying a seed that initializes the random process (*set.seed*) in cases where the SRS uses random subsampling without replacement of the lowest C_{frac}. The random step in SRS is rare and negligible for complex microbiome data, as noted previously [5].

This rather minor modification, however, ensures the reproducibility of SRS, which is essential for microbiome analysis [8].

As an example of microbiome count data normalization using SRS, we utilized a bacterial 16S rRNA gene amplicon sequencing dataset comprising 494 samples derived from an ongoing oral microbiome study. The dataset was processed in QIIME 2 [9]. After anonymization of samples and ASVs, an ASV table comprising a random subset of 100 samples was analyzed. The visualization of SRS curves revealed that the observed ASVs did not decay steadily with decreasing number of reads (Figure 2A). This is due to the way the ranked fractional values (C_{frac}) are chosen: depending on the scaling factor, an ASV with an integer value (C_{int}) of zero may or may not be chosen by ranked subsampling due to its C_{frac}, causing a reproducible zigzag behaviour in the observed number of species. The magnitude of the zigzag observed in SRS curves depends on the data structure (balance between rare and abundant ASVs). Despite the zigzag behaviour, the observed ASV richness was frequently observed to be higher after SRS as compared to rarefying (Figure 2B). Therefore, we recommend the use of the SRS Shiny app (*SRS.shiny.app*-function) prior to SRS for the determination of C_{min} for users working in the R environment. QIIME 2 users are also encouraged to utilize .qza files in the SRS Shiny app (https://vitorheidrich.shinyapps.io/srsshinyapp/ (accessed on 1 November 2021)).

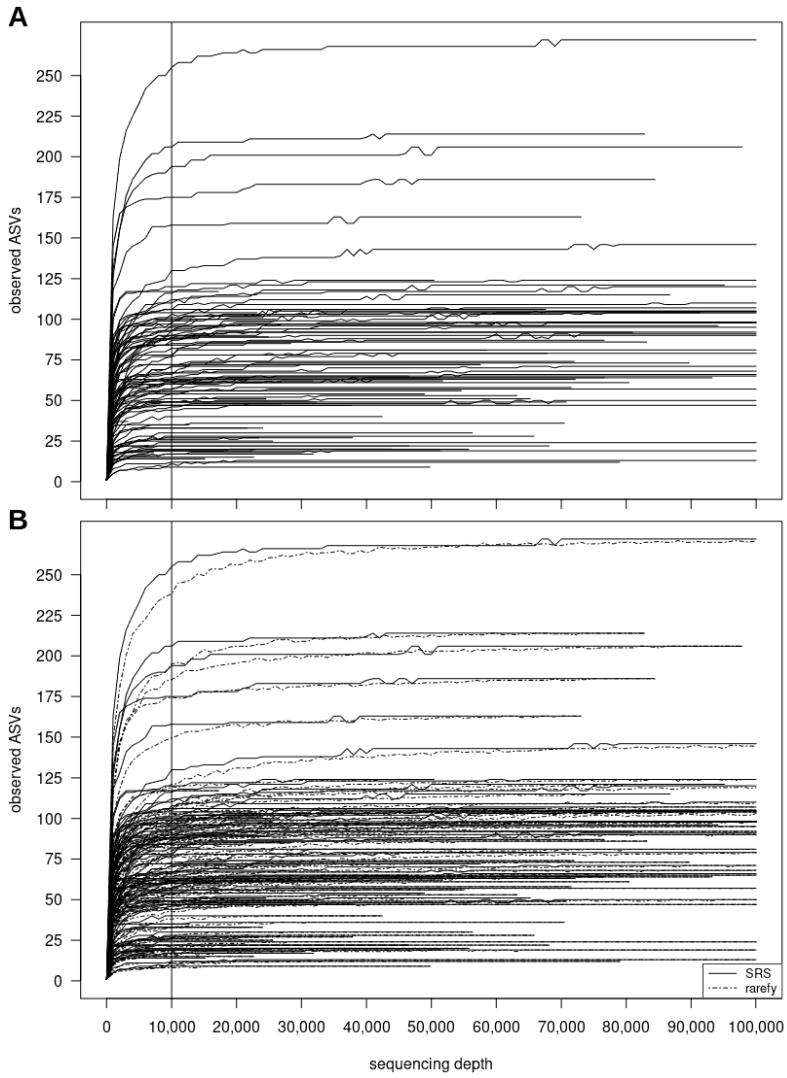


Figure 2. (**A**) SRS curves and (**B**) comparison of SRS curves and repeated rarefying (10 repeats) using the "*richness*" metric (*SRScurve*-function of the 'SRS' R package version 0.2.2). The vertical black solid line indicates the chosen number of counts (10,000) to which all samples will be normalized (C_{min}).

Since its implementation in accessible platforms, SRS has been used to normalize several microbiome datasets obtained from different environments such as animal guts [10], soils [11], oceans [12], and laboratory cultures [13]. McMurdie and Holmes [4] clearly demonstrated that rarefying should not be used to normalize microbiome count data; thus, we suggest that future studies should compare SRS to commonly used normalization techniques other than rarefying.

Author Contributions: Conceptualization, V.H., P.K. and L.B.; methodology, V.H., P.K. and L.B.; software, V.H. and L.B.; validation, V.H., P.K. and L.B.; formal analysis, V.H. and L.B.; investigation, V.H., P.K. and L.B.; resources, V.H., P.K. and L.B.; data curation, V.H. and L.B.; writing—original draft preparation, V.H. and L.B.; writing—review and editing, P.K.; visualization, V.H. and L.B.; supervision, P.K.; project administration, V.H., P.K. and L.B.; funding acquisition, V.H., P.K. and L.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the German Federal Ministry of Education and Research (BMBF) in the framework of the Bonares-SIGNAL project (funding codes: 031A562A and 031B0510A). V.H. was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, process No. 13996-0/2018).

Institutional Review Board Statement: The dataset analyzed here is provenient of a study that was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Hospital Sírio-Libanês (protocol code: HSL 2016-08; date of approval: 18 February 2016).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The ASV table as well as the R script underlying the results of this study are available on GitHub (https://github.com/vitorheidrich/SRS_q2-srs_info (accessed on 1 November 2021)).

Acknowledgments: The authors thank Devon O'rourke for his suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yatsunenko, T.; Rey, F.E.; Manary, M.J.; Trehan, I.; Dominguez-Bello, M.G.; Contreras, M.; Magris, M.; Hidalgo, G.; Baldassano, R.N.; Anokhin, A.P.; et al. Human Gut Microbiome Viewed across Age and Geography. *Nature* **2012**, *486*, 222–227. [CrossRef] [PubMed]
2. Fierer, N. Embracing the Unknown: Disentangling the Complexities of the Soil Microbiome. *Nat. Rev. Microbiol.* **2017**, *15*, 579–590. [CrossRef]
3. Orsi, W.D. Ecology and Evolution of Seafloor and Subseafloor Microbial Communities. *Nat. Rev. Microbiol.* **2018**, *16*, 671–683. [CrossRef]
4. McMurdie, P.J.; Holmes, S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Comput. Biol.* **2014**, *10*, e1003531. [CrossRef]
5. Beule, L.; Karlovsky, P. Improved Normalization of Species Count Data in Ecology by Scaling with Ranked Subsampling (SRS): Application to Microbial Communities. *PeerJ* **2020**, *8*, e9593. [CrossRef] [PubMed]
6. Cont, R.; Heidari, M. Optimal Rounding under Integer Constraints. *arXiv* **2014**, arXiv:1501.00014.
7. Oksanen, J.; Blanchet, F.G.; Friendly, M.; Kindt, R.; Legendre, P.; McGlinn, D.; Minchin, P.R.; O'Hara, R.B.; Simpson, G.L.; Solymos, P.; et al. Vegan: Community Ecology Package. R Package Version 2.5-7. 2020. Available online: https://CRAN.R-project.org/package=vegan (accessed on 1 November 2021).
8. Schloss, P.D. Identifying and Overcoming Threats to Reproducibility, Replicability, Robustness, and Generalizability in Microbiome Research. *mBio* **2018**, *9*, e00525-18. [CrossRef] [PubMed]
9. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857. [CrossRef]
10. Yang, J.; Park, J.; Jung, Y.; Chun, J. AMDB: A Database of Animal Gut Microbial Communities with Manually Curated Metadata. *Nucleic Acids Res.* **2021**, gkab1009. [CrossRef] [PubMed]
11. Beule, L.; Arndt, M.; Karlovsky, P. Relative Abundances of Species or Sequence Variants Can Be Misleading: Soil Fungal Communities as an Example. *Microorganisms* **2021**, *9*, 589. [CrossRef]
12. Pontiller, B.; Pérez-Martínez, C.; Bunse, C.; Osbeck, C.M.G.; González, J.M.; Lundin, D.; Pinhassi, J. Taxon-Specific Shifts in Bacterial and Archaeal Transcription of Dissolved Organic Matter Cycling Genes in a Stratified Fjord. *bioRxiv* **2021**. [CrossRef]
13. Barreto Filho, M.M.; Walker, M.; Ashworth, M.P.; Morris, J.J. Structure and Long-Term Stability of the Microbiome in Diverse Diatom Cultures. *Microbiol. Spectr.* **2021**, *9*, e00269-21. [CrossRef] [PubMed]

Attachment B – Actinomycosis case report

Bruno et al., 2023, *BMC oral health*.

BMC Oral Health

CASE REPORT

Open Access

Unusual gingival actinomycosis post allogeneic hematopoietic stem-cell transplant: case report

Julia Stephanie Bruno[1], Wanessa Miranda-Silva[1], Vitor Heidrich[1,2], Marianne de Castro Gonçalves[3], Yana Novis[3], Celso Arrais-Rodrigues[4,5], Anamaria Aranha Camargo[1] and Eduardo Rodrigues Fregnani[1*]

Abstract

Background Allogeneic hematopoietic stem cell transplant (allo-HSCT) is used to treat several hematological diseases, but immunosuppression during allo-HSCT facilitates opportunistic microbial growth in tissues, such as actinomycosis. An effective diagnosis of opportunistic diseases is essential for correct management of the disease and preservation of the immunosuppressed patient's life.

Case description A 57-year-old female patient was diagnosed with extranodal nasal type NK/T cell lymphoma and underwent curative treatment with allo-HSCT. Twenty-one days after the last clinical follow-up, the patient presented a necrotizing lesion in the papilla region between the first and second molars of the second quadrant. Histopathological analysis showed the presence of a bacterial cluster consistent with Actinomyces infection, and a dense lymphoid infiltrate was also observed. Immunohistochemistry for CD20, CD3, and CD56 was performed to exclude the possibility of the recurrence of extranodal NK/T cell lymphoma. Oral microbiota profiling showed a huge increase in the abundance of Actinomyces bacteria in the subgingival region three weeks prior to appearance of the lesion.

Conclusions Opportunistic infections with an unusual clinical appearance are confounding factors in therapeutic decision-making. We present for the first time a case of actinomycosis in the gingival papilla region following allo-HSCT. We also highlight how microbiota profiling through next-generation sequencing could be used to anticipate bacterial infection diagnosis.

Keywords Oral diseases, Actinomycosis, Oral microbiota, Allogeneic hematopoietic stem-cell transplant

Background

Allogeneic hematopoietic stem-cell transplant (allo-HSCT) recipients undergo a conditioning regimen to induce immunosuppression and prevent graft rejection. However, a less competent immune system puts patients at risk of opportunistic microbial infections. In the oral cavity, patients can present herpes simplex infection, candidiasis, and bacterial infections after allo-HSCT [1, 2].

*Correspondence:
Eduardo Rodrigues Fregnani
eduardofregnani@me.com
[1] Centro de Oncologia Molecular, Hospital Sírio-Libanês, São Paulo, SP, Brazil
[2] Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, SP, Brazil
[3] Centro de Oncologia, Hospital Sírio-Libanês, São Paulo, SP, Brazil
[4] Hospital Nove de Julho, Rede DASA, São Paulo, SP, Brazil
[5] Universidade Federal de São Paulo, São Paulo, SP, Brazil

Actinomyces is a prevalent bacterial genus in the oral cavity, mainly found in the periodontal region [3]. As *Actinomyces sp*. do not possess genes coding for decomposition enzymes, such as hyaluronidases, no tissue degradation is inflicted on the host under normal conditions. However, epithelial barrier injury caused by dental procedures or mucosal ulceration makes the oral cavity susceptible to opportunistic infections caused by *Actinomyces sp.* (i.e., actinomycosis). In line with this, the oral cavity accounts for 50–60% of all actinomycosis cases, afflicting the jaw region primarily. It is characterized by the presence of abscess and mandibular osteomyelitis [4]. Other susceptible sites include the pelvic-abdominal region and lungs, comprising 20% and 10% of actinomycosis cases, respectively [4–6]. Few cases report actinomycosis in the allo-HSCT setting [6, 7]. To the best of our knowledge, the case described here is the first with necrotizing aspects in the papillary region of the gingiva.

A rapid diagnosis of opportunistic infections during cancer treatment is essential for better management of the infection and prevention of treatment discontinuation. In addition to reporting a case with a unique appearance in the soft tissue of the oral cavity, we demonstrate that oral bacteria tracking through 16S rRNA sequencing can prompt an early diagnosis, anticipating clinical tissue disorder.

Case presentation

A female patient aged 57 years was diagnosed with extranodal nasal type NK/T cell lymphoma, cancer staging (EC) IVA, with cutaneous and central nervous system infiltration. The patient underwent allo-HSCT. The graft source was the first-degree sister's bone marrow and the conditioning regimen was performed with Fludarabine-Melphalan and Total Body Irradiation at 400 cGy. Graft-versus-host disease protocol included cyclophosphamide + mycophenolate mofetil (D90) + cyclosporin A. Patient's blood counts (erythrocytes, platelets, lymphocytes, neutrophils, monocytes) at critical allo-HSCT timepoints and antibiotics usage are represented in Fig. 1.

Dental evaluation before hematological treatment did not identify lesions in the oral mucosa. Dental elements were also in good condition with adapted amalgam restorations. The patient reported discomfort in the oral mucosa during allo-HSCT and developed xerostomia grade II (NCI CTCAE v 3.0) and mucositis grade II (WHO grade) in the retropharyngeal region. In addition to primary oral care, the patient received photobiomodulation with low-level laser equipment (Laser XT Therapy, DMC, São Carlos, Brazil) at a wavelength of 660 nm (spot-size = 0.028 cm2; 100 mW of power) to aid in the healing of oral mucositis. Healing occurred within 12 days after the onset of ulceration.

Oroscopy

Twenty-one days after the last clinical follow-up, the patient presented a necrotizing lesion in the papilla region between the first and second molars of the second quadrant and no involvement of the rest of the hard palate or alveolar ridge. A biopsy was performed to confirm the diagnosis of an opportunistic infection in the region of the erythematous border and necrotizing area (Fig. 2).

Histopathological analysis

Histopathological analysis showed ulcerated mucosa with dense lymphoid infiltrate with intermediate size lymphocytes, fibrosis, some plasma cells, and rare eosinophils (Fig. 3A). The presence of a bacterial cluster consistent with *Actinomyces* infection was also observed (Fig. 3B, C). Immunohistochemistry for CD20, CD3, and CD56 was performed to exclude the possibility that the lesion in the gingiva was a recurrence of extranodal NK/T cell lymphoma. There was a mixed infiltrate of mature B and T lymphocytes without atypia. As evaluated by in situ hybridization, the tissue was negative for Epstein-Barr virus (Fig. 4). After the histopathological analysis results, the patient was treated with amoxicillin (875 mg)-potassium clavulanate (125 mg), every 12 h for 4 weeks. There is no photographic record of the healed oral mucosa. The patient died due to aggressive lymphoma recurrence 93 days after the Actinomycosis diagnosis.

Oral microbiota analysis

This patient was enrolled in a research protocol for the longitudinal profiling of the oral microbiota during allo-HSCT (Research Ethics Committee-Hospital Sírio-Libanês: #HSL 2016-08). Oral mucosa (OM), dental biofilm (DB), and gingival crevicular fluid (GCF) samples were collected at: D-5 (preconditioning), D5 (aplasia), D21 (engraftment), D50 (~4 weeks after engraftment), and D85 (~9 weeks after engraftment). Samples were processed and sequenced for microbiota profiling as described previously [8, 9]. Oral microbiota profiling showed drastic diversity changes at all oral sites during allo-HSCT with higher decrease in OM, but similar diversity levels were maintained across all oral sites during follow-up (Fig. 3D). As expected, *Actinomyces* was detected at high relative abundance in all oral sites before allo-HSCT, but abundance decreased during transplant. In the last sample analysed, there was an increase in

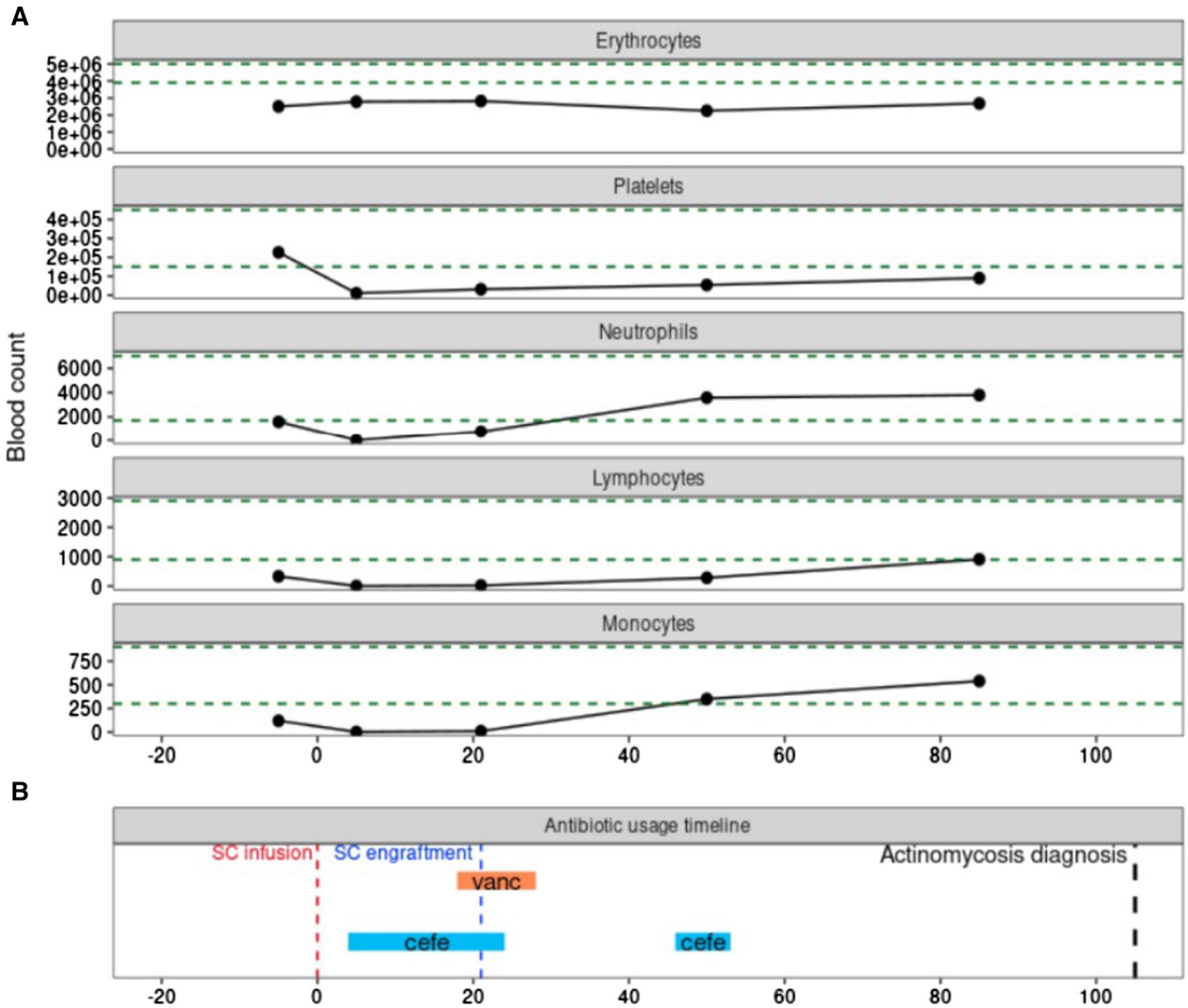


Fig. 1 Blood cell counts and antibiotics administered during allogeneic hematopoietic stem-cell transplant (allo-HSCT). **A** Blood cell counts during allo-HSCT. Blood cell count data was collected at the following allo-HSCT timepoints: D-5 (preconditioning), D5 (aplasia), D21 (engraftment), D50 (~4 weeks after engraftment), and D85 (~9 weeks after engraftment). Dashed green horizontal lines represent normal reference values. **B** Antibiotic usage timeline during allo-HSCT. Antibiotics used between stem-cell (SC) infusion and actinomycosis diagnosis: cefepime (cefe) and vancomycin (vanc)

Actinomyces abundance at all oral sites, particularly in GCF. *Actinomyces* relative abundance in the gingival crevicular fluid was 890% higher in D85 than in D-5, while other oral sites showed slight decreases compared to pre-allo-HSCT (− 20% in DB and − 49% in OM). The increase in *Actinomyces* abundance in GCF preceded the appearance of oral lesions by 20 days. Noteworthy, a single sequence variant of *Actinomyces* (100% identity with *Actinomyces oris*, *Actinomyces naeslundii,* and *Actinomyces viscosus*) was mainly responsible for this abundance increase, suggesting the involvement of a single or a few closely related opportunistic species.

Discussion

Actinomyces bacteria are gram-positive, filamentous bacilli. They compose the commensal microbiota of the oral cavity, genitourinary tract, and gastrointestinal tract in humans. In opportunistic situations, *Actinomyces*

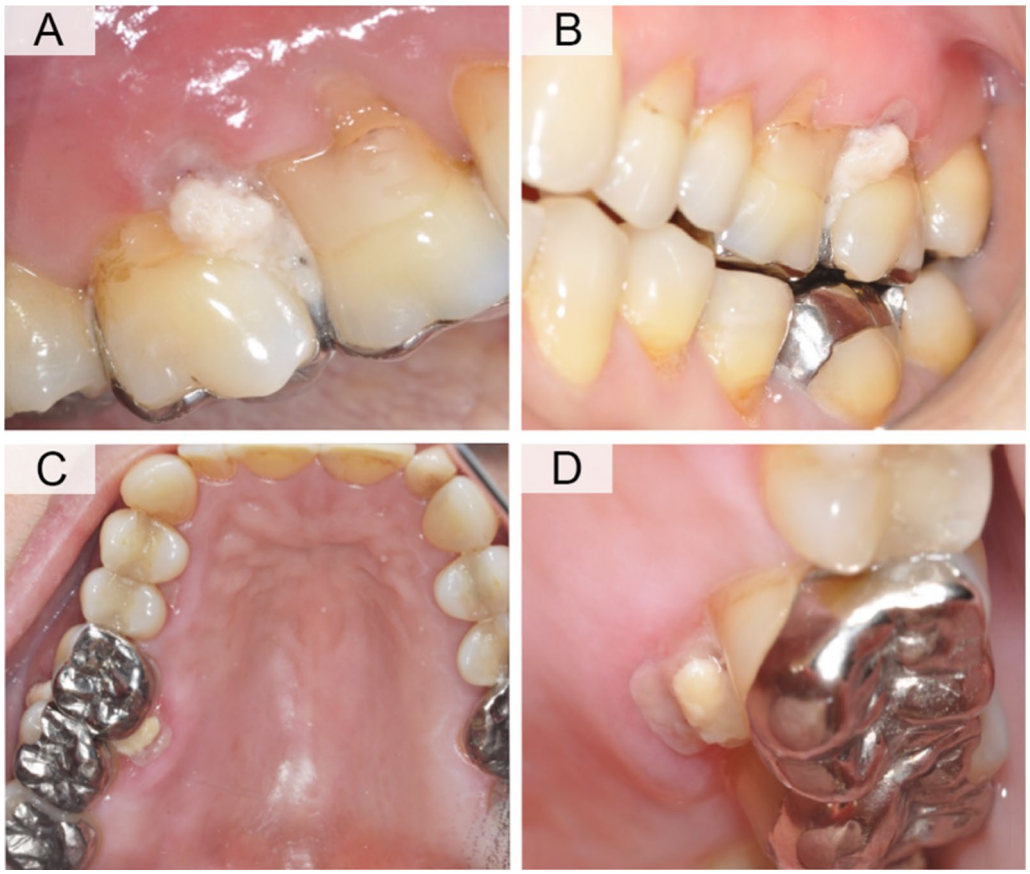


Fig. 2 Different perspectives of oral lesion clinical aspects. **A** mesial, **B** vestibulo-mesial, **C** palatal and **D** mesio-palatal

causes tissue destruction in the lung, bone, genitourinary tract, digestive tract, central nervous system, skin, and mucosa [4, 10]. Cervicofacial actinomycosis, represents 60% of cases [4]. Immunosuppression, poor oral hygiene, age, male sex, and malnutrition are risk factors for the progression of opportunistic lesions [4]. Strategies for maintaining oral microbial eubiosis should be further studied to contain the growth of periodontopathogens [11].

A correct actinomycosis diagnosis is crucial to contain the progression of abscesses and osteomyelitis [4] and can be life-saving for immunosuppressed patients. However, actinomycosis is often hard to diagnose, especially when it mimics other diseases, such as cancer, tuberculosis, and nocardiosis. The diagnosis by bacterial cultivation can be challenging. Since *Actinomyces* bacteria are part of the commensal polymicrobial community of the oral cavity and periodontal regions, contamination during swab collection can result in false-positive results. Therefore, sample collection for detecting oral actinomycosis must be performed carefully, preferably by sampling the inflammatory exudate or the tissue biopsy. Also, following sample collection, due to *Actinomyces* microaerophilic or strict anaerobic characteristic, samples need to be quickly processed in a controlled laboratory environment to prevent false-negative results [4]. False-negatives are even more common in the oncological setting because immunosuppressed patients receive continuous prophylactic antibiotics, which, in addition to a long *Actinomyces* incubation time (average of two weeks), can mislead cultivation results. Histological studies may help the diagnosis in some cases by the detection of sulfur granules, which are responsible for keeping the bacterial colony protected from phagocytosis [4, 10].

Opportunistic infections with an unusual clinical appearance are confounding factors in therapeutic decision-making. Drug therapy in immunosuppressed patients, as well as rapid diagnosis, can be decisive to maintain the patient's systemic stability and oncological

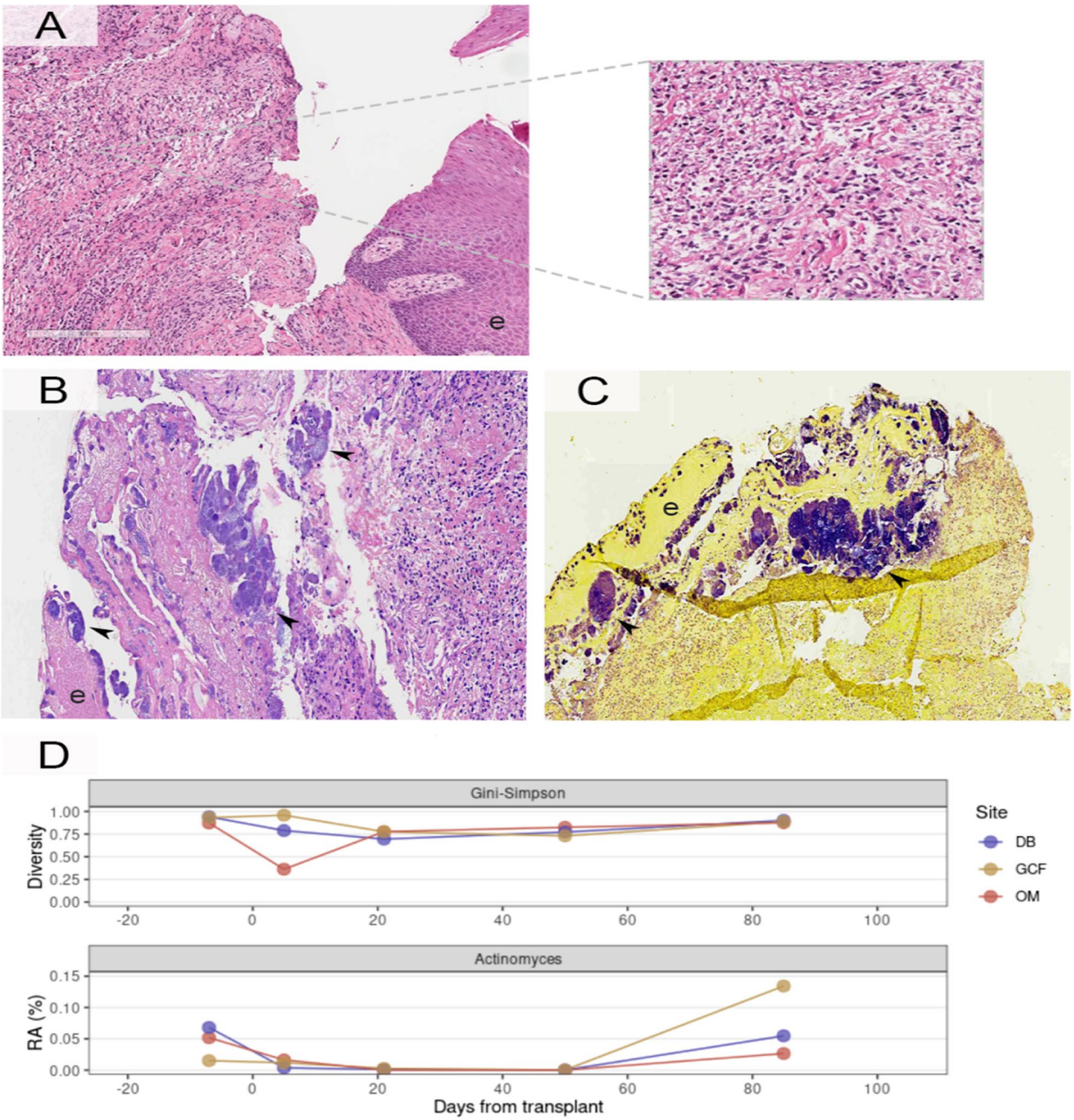


Fig. 3 Pathological and molecular studies. **A** Histological study between the erythematous border and deepithelialization demonstrating inflammatory reaction with lymphocytes and collagen fibres (HE, left × 100 and right × 300); **B** Histological study in the central portion of the lesion with necrotizing inflammation and accumulation of colonies of gram-positive bacteria (arrow); **C** Gram stain histology demonstrating positivities for gram-positive bacteria with coccoid-shaped colonies (GS, × 100); **D** Oral microbiota profiling. Upper graph: alpha diversity (Gini-Simpson index) throughout transplantation and follow-up. Bottom graph: with relative abundance (RA) of the genus *Actinomyces*. *DB* dental biofilm, *GCF* gingival crevicular fluid, *OM* oral mucosa (*HE* hematoxylin and eosin, *GS* gram-staining, *NGS* next-generation sequencing)

follow-up [4, 10]. We here detail a case of an unusual necrosing manifestation of actinomycosis in the gingival papilla. We showcase how microbiota profiling through next-generation sequencing can be an essential tool to anticipate tissue necrosis progression by tracking changes in oral microbes abundance.

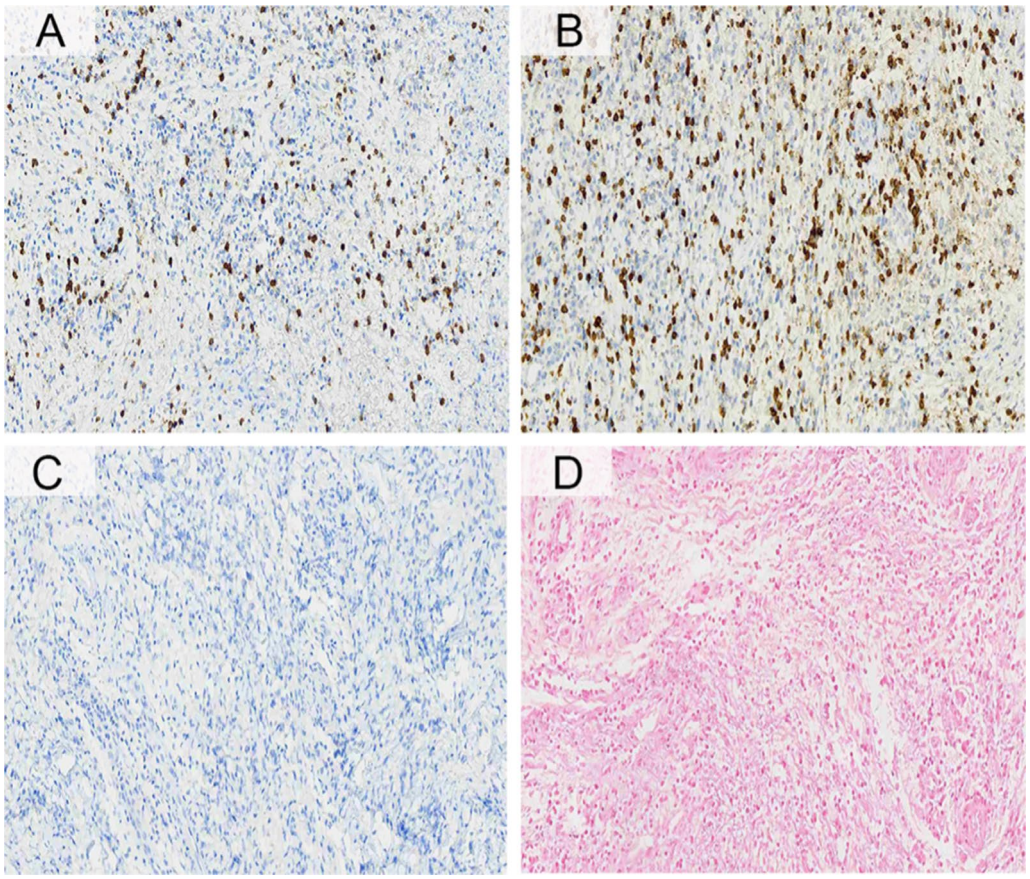


Fig. 4 Study of extranodal NK/T-cell lymphoma markers. Immunohistochemistry method for **A**–**C** and, in situ hybridization for viral status. **A** CD20, **B** CD3, **C** CD56 and **D** Epstein–Barr virus (EBV)

Abbreviations

allo-HSCT	Allogeneic hematopoietic stem-cell transplant
OM	Oral mucosa
DB	Dental biofilm
GCF	Gingival crevicular fluid

Acknowledgements
Not applicable.

Author contributions
JSB, VH and AAC was responsible for writing and revision of the manuscript. ERF, WM-S, YN and CA-R were responsible for conduction of the case. MCG was responsible for histopathological analysis of the specimens. All authors discussed the results and contributed to the final manuscript. All authors read and approved the final manuscript.

Funding
This research did not receive any funding.

Availability of data and materials
The raw sequencing data analysed during the current study are available at https://github.com/vitorheidrich/oral-microbiota-actinomycosis.

Declarations

Ethics approval and consent to participate
This report has been approved by Research Ethics Committee (#HSL 2016-08), Hospital Sírio-Libanês, São Paulo, SP, Brazil. The patient signed a written consent stating their approval for participation in this report.

Consent for publication
A written informed consent form was obtained from the patient for publication of their clinical data. The patient understands that their name or initials will not be mentioned in this article.

Competing interests
The authors declare no competing interests.

Received: 7 December 2022 Accepted: 30 January 2023
Published online: 02 February 2023

References

1. Kawajiri J, Nagata K, Nakamura A, Fujieda A, Ino K, Nomura J, et al. Clinical utility of oral management in allogeneic hematopoietic stem cell transplantation recipients: microbiological evidence based on molecular analysis of oral bacteria. Support Care Cancer. 2022;30(1):757–64.

2. Elad S, Raber-Durlacher JE, Brennan MT, Saunders DP, Mank AP, Zadik Y, et al. Basic oral care for hematology–oncology patients and hematopoietic stem cell transplantation recipients: a position paper from the joint task force of the Multinational Association of Supportive Care in Cancer/International Society of Oral Oncology (MASC. Support Care Cancer. 2015;23(1):223–36.
3. Segata N, Haake SK, Mannon P, Lemon KP, Waldron L, Gevers D, et al. Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. Genome Biol. 2012;13(6):R42.
4. Valour F, Sénéchal A, Dupieux C, Karsenty J, Lustig S, Breton P, et al. Actinomycosis: etiology, clinical features, diagnosis, treatment, and management. Infect Drug Resist. 2014;7:183–97.
5. Choi SG, Lee EY, Lee OJ, Kim S, Kang JY, Lim JS. Prediction models for early diagnosis of actinomycotic osteomyelitis of the jaw using machine learning techniques: a preliminary study. BMC Oral Health. 2022;22(1):164.
6. Yagi T, Fujino H, Hirai M, Inoue T, Sako M, Teshima H, et al. Esophageal actinomycosis after allogeneic peripheral blood stem cell transplantation for extranodal natural killer/T cell lymphoma, nasal type. Bone Marrow Transplant. 2003;32(4):451–3.
7. Barraco F, Labussière-Wallet H, Valour F, Ducastelle-Leprêtre S, Nicolini FE, Thomas X, et al. Actinomycosis after allogeneic hematopoietic stem cell transplantation despite penicillin prophylaxis. Transpl Infect Dis. 2016;18(4):595–600.
8. Bruno JS, Heidrich V, Knebel FH, de Molla VC, Parahyba CJ, Miranda-Silva W, et al. Commensal oral microbiota impacts ulcerative oral mucositis clinical course in allogeneic stem cell transplant recipients. Sci Rep. 2022;12(1):17527.
9. Heidrich V, Bruno JS, Knebel FH, de Molla VC, Miranda-Silva W, Asprino PF, et al. Dental biofilm microbiota dysbiosis is associated with the risk of acute graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. Front Immunol. 2021;12:692225.
10. Yuan Y, Hou Z, Peng D, Xing Z, Wang J, Zhang S. Pulmonary *Actinomyces graevenitzii* infection: case report and review of the literature. Front Med. 2022;9:916817.
11. Butera A, Gallo S, Maiorani C, Molino D, Chiesa A, Preda C, et al. Probiotic alternative to chlorhexidine in periodontal therapy: evaluation of clinical and microbiological parameters. Microorganisms. 2020;9(1):69.

Publisher's Note

Attachment C – Cohort's clinical and microbiota characteristics

Individualized clinical and oral microbiota characteristics of the study patients are available in https://github.com/vitorheidrich/oral-microbiota-hsct. Antibiotic usage data refers to the antibiotic use between preconditioning (P) and engraftment + 30 days (E30). Only categorical microbiota variables significantly associated with the allo-HSCT clinical course are included in the table. Patient #16 was analyzed only in the last chapter, therefore some columns miss information ("NA") for this patient. GVHD, graft-versus-host disease; HCT-CI, hematopoietic cell transplantation specific comorbidity index; MO, oral mucositis onset; OM, oral mucosa; SB, supragingival biofilm; TRD, transplant-related death.

Attachment D – Academic *Curriculum Vitae*

Vitor Heidrich

Place of Birth: São Paulo, Brazil • Date of Birth: 21 Jul 1996
vheidrich@mochsl.org.br • Website
ORCID ID: 0000-0001-6617-9187 • Linktree

Education

2018 – Present **Universidade de São Paulo** – São Paulo, Brazil
PhD in Biological Sciences (Biochemistry)
Supervisor: Dr. Anamaria Aranha Camargo
Title: Longitudinal analysis at three oral sites links oral microbiota to clinical outcomes in allogeneic hematopoietic stem-cell transplant.

2014 – 2018 **Universidade de São Paulo** – São Paulo, Brazil
Bachelor of Molecular Sciences
Supervisor: Dr. Daniela Sanchez Bassères
Title: Identification of long non-coding RNAs regulated by KRAS in pancreatic cancer.

Selected coursework

2021 *Biology*: Advances in microbial biotechnology and use of next-generation sequencing (45h) – Universidad Nacional de Tucumán, San Miguel de Tucumán, Argentina.

2018 *Physics*: School on Physics Applications in Biology (50h) – Instituto de Física Teórica, São Paulo, Brazil.

2017 *Mathematics*: Southern-Summer School on Mathematical Biology (120h) – Instituto de Física Teórica, São Paulo, Brazil.

Research experience

2018 – Present **Molecular Oncology Center (MOC)**
Hospital Sírio-Libanês – São Paulo, Brazil
Supervisor: Dr. Anamaria Aranha Camargo
Funding: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, process no. 13996-0/2018)

2016 – 2018 **Laboratory of Molecular Biology of Cancer**
Departamento de Bioquímica, Universidade de São Paulo – São Paulo, Brazil
Supervisor: Dr. Daniela Sanchez Bassères
Funding: Conselho Nacional de Desenvolvimento Científico e Tecnológico (PIBIC)

Technical skills

Programming languages
R (proficient), Python (familiar), Bash (familiar)

Techniques
Cell culture, DNA/RNA extraction, real-time PCR, next-generation sequencing

Languages
Portuguese (native), English (fluent)

Publications

Peer-reviewed publications

2023 **Unusual gingival actinomycosis post allogeneic hematopoietic stem-cell transplant: case report**
Bruno, J. S., Miranda-Silva, W., Heidrich, V., Gonçalves, M. C., Novis, Y., Rodrigues-Arrais, C., ..., & Fregnani, E. R. *BMC Oral Health*.

2022 **Dynamics of Gut Microbiota and Clinical Variables after Ketogenic and Mediterranean Diets in Drug-Naïve Patients with Type 2 Diabetes Mellitus and Obesity**
Deledda, A., Palmas, V., Heidrich, V., Fosci, M., Lombardo, M., ..., & Velluzzi, F. *Metabolites*.

Commensal oral microbiota impacts ulcerative oral mucositis clinical course in allogeneic stem cell transplant recipients
Bruno, J., Heidrich, V., Knebel, F. H., de Molla, V. C., Parahyba, C. J., Miranda-Silva, W., ... & Fregnani, E. R. *Scientific Reports*.

Are short-read amplicons suitable for the prediction of microbiome functional potential? A critical perspective
Heidrich, V., & Beule, L. *iMeta*.

Choice of 16S Ribosomal RNA Primers Impacts Male Urinary Microbiota Profiling
Heidrich, V., Inoue, L. T., Asprino, P. F., Bettoni, F., Mariotti, A. C., Bastos, D. A., ... & Camargo, A. A. *Frontiers in cellular and infection microbiology*.

2021 **'SRS' R Package and 'q2-srs' QIIME 2 Plugin: normalization of microbiome data using scaling with ranked subsampling (SRS)**
Heidrich, V., Karlovsky, P., & Beule, L. *Applied Sciences*.

Disruption of the oral microbiota is associated with a higher risk of relapse after allogeneic hematopoietic stem cell transplantation
de Molla, V. C., Heidrich, V., Bruno, J. S., Knebel, F. H., Miranda-Silva, W., Asprino, P. F., … & Arrais-Rodrigues, C. *Scientific Reports*.

Dental biofilm microbiota dysbiosis is associated with the risk of acute graft-versus-host disease after allogeneic hematopoietic stem cell transplantation
Heidrich, V., Bruno, J. S., Knebel, F. H., de Molla, V. C., Miranda-Silva, W., Asprino, P. F., … & Camargo, A. A. *Frontiers in Immunology*.

2020 **Aurora A kinase and its activator TPX2 are potential therapeutic targets in KRAS-induced pancreatic cancer**
Gomes-Filho, S. M., Dos Santos, E. O., Bertoldi, E. R. M., Scalabrini, L. C., Heidrich, V., Dazzani, B., … & Bassères, D. S. *Cellular Oncology*.

Preprints

2022 **Longitudinal analysis at three oral sites links oral microbiota to clinical outcomes in allogeneic hematopoietic stem-cell transplant**
Heidrich, V., Knebel, F. H., Bruno, J. S., de Molla, V. C., Miranda-Silva, W., Asprino, P. F., …, & Camargo, A. A. *medRxiv*.

Software

Scaling with Ranked Subsampling (SRS)

- R package (CRAN; GitHub)
- QIIME 2 plugin (q2-library; GitHub)

Selected presentations

Oral presentations

Oct 2021 Ecological resilience of the oral microbiota during allogeneic hematopoietic stem-cell transplantation
22nd Brazilian Congress of Clinical Oncology (virtual)

Sep 2021 Dental biofilm microbiota dysbiosis is associated with the risk of acute graft-versus-host disease after allogeneic hematopoietic stem cell transplantation
Microbiome Virtual International Forum n.1 – Open-access paper highlight (virtual)

Poster presentations

Nov 2022 Longitudinal analysis at three oral sites links oral microbiota to clinical outcomes in allogeneic hematopoietic stem-cell transplant
International Human Microbiome Consortium Congress 2022 (Kobe, Japan)

Oct 2020 Dental biofilm microbiota dysbiosis predicts aGVHD risk after hematopoietic stem-cell transplantation
SBOC-AACR Joint Congress: A Translational Approach to Clinical Oncology (virtual)

May 2018 Identification of long non-coding RNAs regulated by KRAS in pancreatic cancer
47th Annual Meeting of the Brazilian Society of Biochemistry and Molecular Biology (Joinville, Brazil)

Grants and awards

2018–Present Doctoral scholarship – FAPESP (process no. 13996-0/2018), São Paulo, Brazil.

2022 Best oral presentation – 2nd Translational Research Forum, Hospital Sírio-Libanês, São Paulo, Brazil.
Ecological resilience of the oral microbiota during allogeneic hematopoietic stem-cell transplantation.

2022 Best research highlight – Microbiome Virtual International Forum n.14.
Longitudinal analysis at three oral sites links oral microbiota to clinical outcomes in allogeneic hematopoietic stem-cell transplant.