haematologica
Journal of the Ferrata Storti Foundation

Dental plaque microbiota following allogeneic hematopoietic cell transplantation and risk of chronic graft-*versus*-host disease

by Hakan Gem, Maryam Ebadi, Gale Sebastian, Rania Abasaeed, Michele Lloid, Samuel S. Minot, David R. Dean and Armin Rashidi

Received: May 20, 2025.
Accepted: August 22, 2025.

Citation: Hakan Gem, Maryam Ebadi, Gale Sebastian, Rania Abasaeed, Michele Lloid, Samuel S. Minot, David R. Dean and Armin Rashidi. Dental plaque microbiota following allogeneic hematopoietic cell transplantation and risk of chronic graft-versus-host disease.
Haematologica. 2025 Sept 4. doi: 10.3324/haematol.2025.288279 [Epub ahead of print]

Dental plaque microbiota following allogeneic hematopoietic cell transplantation and risk of chronic graft-*versus*-host disease

Hakan Gem[1*], Maryam Ebadi[2*], Gale Sebastian[1], Rania Abasaeed[1], Michele Lloid[1], Samuel S. Minot[3], David R. Dean[1#], Armin Rashidi[4,5#]

[1]Department of Oral Medicine, University of Washington and Fred Hutchinson Cancer Center, Seattle, WA, USA

[2]Department of Radiation Oncology, University of Washington and Fred Hutchinson Cancer Center, Seattle, WA, USA

[3]Data Core, Shared Resources, Fred Hutchinson Cancer Center, Seattle, WA, USA

[4]Clinical Research Division, Fred Hutchinson Cancer Center, Seattle, WA, USA

[5]Division of Medical Oncology, Department of Medicine, University of Washington, Seattle, WA, USA

[*]M.E. and H.G. contributed equally as co-first authors. [#]D.R.D. and A.R. contributed equally as senior authors.

Corresponding author: Armin Rashidi, Email: arashidi@fredhutch.org

Running title: Dental plaque microbiota and chronic GVHD

Acknowledgments

We thank Nellie Ratsamee and Mikki Burns for coordination of study activities, Daniel Podlesny for help with SameStr, and Michael Guyumdzhyan, Pedrom Imankhan, and Rizmina Lathiff for help with kit preparation, sample collection, and sample processing. We thank the Microbial Interactions & Microbiome Center (University of Washington, Seattle, WA) and Clinical Research Specimen Processing (CRSP, Fred Hutchinson Cancer Center, Seattle, WA) for supporting this study.

Funding

This work was supported by the National Institutes of Health's award P30 CA015704 to the Fred Hutch/University of Washington/Seattle Children's Cancer Consortium. H.G. was supported by a resident research award from the American Academy of Oral Medicine Research Advancement Committee, and D.R.D. by the Dr. Douglass L. Morell Dentistry Research Fund from the Research Advisory Committee of the University of Washington School of Dentistry. The funders did not have any role in data collection, interpretation, or reporting. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions

H.G., D.R.D., M.L., and A.R. designed the study. H.G., M.E., and A.R. collected the data. H.G., G.S., R.A., D.R.D., and A.R. conducted the study. A.R. and S.S.M. analyzed

the data. H.G. and A.R. wrote the manuscript. All authors read and approved the final manuscript.

Competing Interests
A.R. has received consulting fees from Seres Therapeutics and serves as a member of an Emmes Data and Safety Monitoring Board, both outside of the scope of the present study.

Data sharing statement
The sequencing data reported in this paper are available from NCBI Sequence Read Archive (SRA) BioProject ID PRJNA1190407. De-identified patient-level data will be available from the corresponding author upon reasonable request.

Article Summary
We evaluated the relationship between oral microbiota early after allogeneic hematopoietic cell transplantation and subsequent chronic graft-versus-host disease (cGVHD). *Streptococcus sanguinis* abundance in day +28 dental plaque and two microbial metabolic pathways were associated with chronic GVHD months later, offering potential novel targets for early prophylactic intervention.

Abstract
Microbiota disruptions have been associated with short-term complications after allogeneic hematopoietic cell transplantation (alloHCT). However, only a few studies have examined the relationship between dysbiosis and chronic graft-versus-host disease (cGVHD), the main long-term immunologic toxicity of alloHCT. Considering the role of oral microbiota in systemic inflammatory diseases, we evaluated whether oral microbiota at day 28 post-HCT corresponding to clinical recovery from the acute events after transplantation is associated with subsequent cGVHD. Shotgun metagenomic sequencing of 207 saliva and supragingival plaque samples collected longitudinally at baseline (pre-conditioning), day +28, and day +84 from 37 patients (11 with subsequent moderate/severe cGVHD) revealed a significant association between day +28 plaque microbiota composition and cGVHD. Two orthogonal statistical approaches demonstrated *Streptococcus sanguinis* and *Prevotella loescheii* in day +28 plaque to be associated with cGVHD. Metagenome-based functional analysis identified 4 microbial metabolic pathways associated with future cGVHD, two of which were highly attributed to *S. sanguinis*. These pathways – ethanolamine utilization and glycerol metabolism – increase bacterial fitness by providing an alternative carbon/nitrogen source and improving survival in inflamed tissues. Our findings propose a novel mechanism by which the early post-transplant dental biofilm may contribute to cGVHD months later, offering a potential target for early prophylactic intervention.

Key words: chronic GVHD, microbiota, plaque, transplantation

Introduction

Allogeneic hematopoietic cell transplantation (alloHCT) is a potentially curative treatment for many patients with hematologic malignancies. Hematopoietic cells obtained from a donor are infused to the patient following a conditioning regimen – typically consisting of 1 or more chemotherapeutic drugs with or without total body irradiation -- varying in intensity from non-myeloablative to reduced-intensity to myeloablative. The immunologic attack mounted by the new immune system against allogeneic tumor antigens mediates the graft-versus-tumor effect, ultimately leading to eradication of cancer and cure in successful cases. Despite its high curative potential, alloHCT is associated with several short- and long-term toxicities. Common short-term toxicities include oral and intestinal mucositis related to conditioning-induced damage to epithelial tissues, infections, and acute graft-versus-host disease (aGVHD). The prototype chronic toxicity of alloHCT is chronic GVHD (cGVHD), a systemic inflammatory condition with significant morbidity that occurs in 20-40% of alloHCT recipients despite immunosuppressive prophylaxis[1]. In addition to various toxicities to host tissues, alloHCT is commonly associated with disruptions to the intestinal and oral microbiota[2-4]. Most of these disruptions occur in the first few weeks after alloHCT due to heavy antibiotic use to prevent and treat infections.

Prevention and treatment of cGVHD are currently based on the use of immunosuppressive medications which increase the risk of infections, among other short- and long-term toxicities. Known risk factors for cGVHD include donor/recipient HLA mismatch, use of peripheral blood as the stem cell source, female-to-male transplant, older patient age, and acute GVHD[5-8]. Other than acute GVHD, none of these risk factors can be modified post-transplantation. If modifiable clinical or laboratory features predictive of higher risk for cGVHD based on each patient's experience in the early post-transplant period could be identified, personalized preventative strategies may be developed to prevent cGVHD in high-risk patients.

Growing evidence suggests that early post-transplant disruptions to the microbiota (a.k.a. dysbiosis), especially the colonic microbiota, influence the risk for aGVHD[9-13]. Fewer studies have examined the potential involvement of the gut microbiota in cGVHD pathogenesis[14-19], and even fewer have focused on the oral microbiota[20,21]. The contribution of oral microbiota to systemic inflammatory diseases such as rheumatoid arthritis (RA) is increasingly recognized. Oral bacteria can reach distant organs and cause inflammation via two pathways: (i) direct hematogenous spread from the mouth[22,23] and (ii) translocation to the gut[24,25] via the swallowed saliva or the lymphatics[21,26]. In the setting of alloHCT, oral mucositis resulting from transplant conditioning facilitates bacterial translocation through mucosal breaks into the circulation. Indeed, oral streptococci are among the most common causes of bacteremias after alloHCT[27]. In addition, in a murine cGVHD model where mice with oral dysbiosis (induced by ligature) were used as allograft recipients, oral enterococci migrated to the draining cervical lymph nodes, activating local antigen-presenting cells and increasing alloreactive T-cell proliferation[21].

To advance this knowledge and examine the potential role of the oral microbiota in cGVHD pathogenesis in humans, we focused on supragingival plaque microbiota at day +28 of alloHCT in patients who received myeloablative conditioning. We made these choices because: (i) Day +28 represents an early post-engraftment timepoint when most acute toxicities of conditioning have resolved, antibacterial antibiotics have been discontinued, and enteral intake has been reestablished[28]. New biofilm with potentially different microbiota composition from baseline has started to form by this point[28]. (ii) Supragingival plaque can serve as a reservoir of pathogens for the subgingival plaque where the species are closer to the gingival mucosa and can more easily reach the submucosa[29]. (iii) Most patients receiving myeloablative conditioning develop oral mucositis[30,31], increasing the likelihood of systemic bacterial translocation. We assumed that the accumulation of putative pathogens in the supragingival plaque would partly mirror the evolution of the subgingival plaque where microbiota-host interaction may have implications for subsequent cGVHD. Thus, we chose supragingival plaque, hereafter referred to simply as plaque, as the focus of our microbiota analysis. In addition, we sequenced the salivary microbiota as a control. Longitudinal samples at two additional timepoints (baseline and day +84) were included to permit assessment of microbiota dynamics and evaluation of whether day +28 findings extend to earlier or later timepoints.

The specific hypothesis we tested in this work was whether the composition of plaque microbiota at day +28 is associated with subsequent risk for cGVHD. Using shotgun metagenomic sequencing, we find *Streptococcus sanguinis* and *Prevotella loescheii* to be associated with future moderate or severe cGVHD. Metagenome-based functional analysis identifies two microbial metabolic pathways increasing bacterial virulence and fitness, as putative mechanisms for *S. sanguinis* contribution to cGVHD pathogenesis.

Methods

Study design and procedures

We used samples from our umbrella protocol of longitudinal multi-site oral samples combined with mucositis assessment data, a prospective single-center collaborative study between HCT and Oral Medicine teams at Fred Hutchinson Cancer Center (FHCC, Seattle, WA). The protocol, summarized in **Fig. 1a**, was approved by FHCC's Institutional Review Board (#RG1123080). The recruitment period was between March 2023 and April 2024. All patients provided written informed consent, and study procedures were in accordance with the Helsinki Declaration. The only eligibility criteria were age >18 years and a myeloablative alloHCT. Day 0 was defined as the day of stem cell infusion. The baseline visit was performed by an Oral Medicine specialist during the standard transplant workup and before the initiation of conditioning. Details of the standard pre-transplant Oral Medicine evaluation in our center are provided in **Supplementary Methods**. Baseline samples were collected during this visit. We refer most patients without a recent dental cleaning for cleaning following the baseline visit. Supragingival plaque is thoroughly removed during the cleaning visit. For the present work, we selected the subset of patients with saliva and plaque samples at baseline, day +28, and day +84. Patients with relapse before cGVHD or non-GVHD death by 1 year post-transplant were not included.

cGVHD was diagnosed and graded according to the NIH Consensus Criteria[32]. Cases were defined as patients developing moderate or severe cGVHD by 1 year post-transplant; controls were defined as those with no or mild GVHD. Details of mucositis assessment and sample collection/processing are provided in **Supplementary Methods**. The microbiota was profiled to the species level using shotgun metagenomic sequencing. Details are provided in **Supplementary Methods**.

Statistical and bioinformatic analysis

All eligible patients and their baseline, day +28, and day +84 saliva and plaque samples were included, with no power or sample size calculation. Microbiota analysis was performed using species-level data. Within-sample microbiota diversity (i.e., alpha diversity) was quantified by Shannon index[33]. Between-sample microbiota diversity (i.e., beta diversity) was quantified by Aitchison distance[34]. Ordination was visualized by principal coordinate analysis using the first two axes and differences between cases and controls in the overall microbiome composition were statistically examined using an adonis test with 999 permutations. To identify differentially abundant species between cases and controls, two orthogonal approaches were used: (i) general linear model-based Microbiome Multivariable Association with Linear Models (MaAsLin2)[35], using species relative abundance and prevalence thresholds of 1% and 10%, respectively. Other than a stricter significance threshold of 0.05 than the software's default for Benjamini-Hochberg corrected P values[36] (q values), default parameters were used. (ii) LinDA[37], which fits linear regression models on centered log-ratio transformed data and corrects the bias due to compositional effects. To identify differentially abundant pathways between cases and controls, CPM-normalized pathways were first filtered using relative abundance and prevalence thresholds of 0.01% and 10%, respectively, and then analyzed using MaAsLin2. A P value threshold of 0.01 and effect size (i.e., enrichment coefficient) threshold of 2 were used to identify the pathways most likely relevant for the species-cGVHD associations observed.

Continuous variables were compared between the groups using a Wilcoxon's test. Correlations were evaluated using a Spearman's correlation test. All analyses were performed in R 4.2.0.

Results

Thirty-seven eligible patients were included (**Table 1**), including 11 cases (moderate or severe cGVHD) and 26 controls (no or mild cGVHD). The only significant difference between the groups was a higher proportion of female to male transplants among cases (P = 0.02). The only death event was a cGVHD-related death at 10 months post-transplant. A total of 207 samples were profiled and analyzed (**Fig. 1a**). Antibacterial antibiotic exposures between day -7 and day +28 are shown in **Fig. 1b**. Fifteen (41%) patients developed grade II-IV acute GVHD. Five of these patients also developed cGVHD (cases). The remaining 6 patients with cGVHD (cases) did not have a history of grade II-IV aGVHD. The median time to onset of cGVHD was 6 months (range 3-12 months). In 4 patients with cGVHD, only 1 organ was involved, while in the remaining 7 patients, multiple organs were involved. Overall, the most commonly affected organ with

cGVHD was the mouth (9 patients), followed by the skin (5 patients), eyes (4 patients), and genitalia (4 patients). The median number of involved organs was 3 (range 1-4).

Day +28 plaque microbiota is associated with the risk of future cGVHD

A recent study reported a higher risk for cGVHD in patients with more severe oral mucositis in the first month after alloHCT[21]. Consistent with previous work by others, oral mucositis in our patients peaked in the second week after transplantation[19]. Therefore, we compared cases and controls for total mucositis score (derived from the comprehensive OMAS scoring system) at days +7 and +14, separately. No significant differences were observed between cases and controls in these comparisons (**Fig. S1a-b**). However, the maximum total score (over days +7 and +14) was higher among cases, though not statistically significantly (P = 0.15; **Fig. S1c**).

The number of high-quality microbial reads was >0.5M for all samples, with a median of 13.2M (mean 15.2M). A total of 452 species were identified (plaque: 373 species; saliva: 381 species). The distribution of the 15 most abundant genera in samples of each type at each timepoint, grouped by cases and controls, is shown in **Fig. S2**. No significant differences were found between cases and controls in saliva or plaque microbiota diversity at baseline (plaque: P = 0.83; saliva: P = 0.18) or at day +28 (plaque: 0.08; saliva: P = 0.11). At day +84, cases and controls were not different in plaque microbiota diversity (P = 0.74), but salivary microbiota diversity was higher in cases compared to controls (P = 0.02) (**Fig. 2a-f**).

Microbiota composition was significantly different between cases and controls for day +28 plaque (P = 0.003) but not for baseline (P = 0.57) or day +84 plaque (P = 0.43) or for saliva at any timepoint (baseline: P = 0.57; day +28: P = 0.24; day +84: P = 0.18) (**Fig. 2g-l**). In differential abundance analysis using MaAsLin2 to identify enriched and depleted species in day +28 plaque microbiota that drove the overall microbiota compositional difference between cases vs. controls, two species were enriched in cases ($q < 0.05$; **Fig. 3a**): *S. sanguinis* and *P. loescheii*. Differential abundance analysis using LinDA identified the same 2 species as the most significantly enriched species in day +28 plaque microbiota of cases, but also *Schaalia odontolytica* and *Granulicatella*_SGB8255 ($q < 0.05$ for all 4 species; **Fig. 3b**). Since *S. sanguinis* and *P. loescheii* were strongly associated with cGVHD in both models (**Fig. 3c-d**), we focused on these 2 species in further analyses. **Fig. S3** shows the results of differential abundance analysis at baseline and day +84; no differentially abundant species were identified in either site at either timepoint.

Metagenome-based functional analysis identifies a putative mechanism for microbiota effects on cGVHD risk

So far, our data demonstrated an association between day +28 plaque microbiota, especially *S. sanguinis* and *P. loescheii*, and future cGVHD. Next, we evaluated whether metagenome-based functional analysis can identify putative mechanisms for this novel association. HUMAnN-based functional analysis identified 5 pathways enriched in day +28 plaque (**Fig. 4a**): PWY0-1477 (ethanolamine utilization), GOLPDLCAT-PWY (superpathway of glycerol degradation to 1,3-propanediol), PWY-

5675 (nitrate reduction V (assimilatory)), PWY0-1479 (tRNA processing), and PWY-6969 (TCA cycle V). Per-sample species attributions of PWY0-1477 and GOLPDLCAT-PWY, the only pathways with *S. sanguinis* attribution, are shown in **Fig. 4b-c**. *S. sanguinis* was the second most significant contributor to PWY0-1477 (**Fig. 4d**) and the only classified contributing species to GOLPDLCAT-PWY (**Fig. 4e**). *P. loescheii* had no contribution to any of the enriched pathways. Species contributions to the other enriched pathways are shown in **Fig. S4**. Together, these data suggest PWY0-1477 and/or GOLPDLCAT-PWY pathways as putative mechanisms by which *S. sanguinis* in day +28 plaque may increase the risk of subsequent cGVHD. The area under the curve (AUC) in receiver operating characteristic curves analysis for day +28 plaque *S. sanguinis* to predict subsequent cGVHD was 0.79 (95%CI 0.62-0.97)(**Fig. 4f**), suggesting its potential value as a predictive biomarker. Importantly, the enriched pathways in the plaque were not enriched in the saliva (**Fig. S5**), supporting niche specificity of our findings.

S. sanguinis abundance in day +28 plaque depends on patient-intrinsic and post-transplant factors

Next, we asked whether specific patient or transplant characteristics including antibiotic exposures and mucositis severity were predictive of day +28 plaque *S. sanguinis* abundance. None of the variables analyzed were significant (**Fig. S6**). However, *S. sanguinis* abundance in baseline plaque was moderately correlated with *S. sanguinis* abundance in day +28 plaque from the same patient (Spearman's correlation coefficient 0.52, P = 0.002; **Fig. 4g**). As patients without recent dental cleaning (27 patients; 73%) underwent thorough cleaning including plaque removal following baseline research sample collection, day +28 plaques represented de novo biofilm formation. Therefore, the higher abundance of a given species in day +28 plaque suggests a patient-specific tendency to accumulate more of that species (consistent with individuality of plaque microbiota[38]) combined with post-transplant factors (e.g., diet, oral hygiene) as other determinants. Importantly, although baseline and day +28 plaque *S. sanguinis* correlated, baseline levels were not associated with cGVHD (P = 0.22; **Fig. 4h**) and had little predictive value (AUC 0.63, 95%CI 0.43-0.84; **Fig. 4i**). These findings suggest that day +28 plaque *S. sanguinis* abundance is personalized and depends on (unknown) post-transplant variables.

The dynamics of *S. sanguinis* and *P. loescheii* from baseline to day +28 to day +84 are shown in **Figs. 4j-k**. *S. sanguinis* abundance declined between baseline and day +28 in the cohort without cGVHD but remained stable in the cGVHD cohort. Between days +28 and +84, the abundance of this species declined in both cohorts, while remaining higher in the cGVHD cohort.

Discussion

The microbiota has been repeatedly associated with short- and long-term toxicities of alloHCT. Here, we used our detailed clinical database of longitudinal multi-site oral mucositis assessments associated with a biorepository including longitudinal salivary and dental plaque samples to examine whether the plaque microbiota at the time of recovery from the early phase of alloHCT (i.e. day +28) was associated with subsequent

cGVHD. We used saliva as a control for sample type. For comparison purposes, sample of both types at baseline and day +84 were also included. We demonstrated that only the day +28 plaque microbiota was associated with cGVHD. Specifically, enrichment of day +28 plaque in *S. sanguinis* and *P. loescheii* was associated with a higher risk for future cGVHD. Metagenome-based functional analysis revealed two metabolic pathways increasing bacterial fitness as potential mechanisms for *S. sanguinis* association with cGVHD.

The role of oral microbiota in systemic inflammatory conditions has been established, most compellingly in RA[39]. Consistent with our findings in cGVHD, recent studies in RA point to streptococci as the main potential microbial contributor to pathogenesis. Streptococci are abundant members of the dental plaque, with *S. sanguinis* being a pioneer species and an early colonizer of the biofilm[40]. Through adherence to the enamel, *S. sanguinis* permits other species to colonize the biofilm, forming a complex structure[41,42]. Streptococci are enriched in the oral cavity of patients with RA and their poorly degraded cell wall peptidoglycan-polysaccharide polymers can induce proinflammatory cytokine production and arthritis in mice[43]. Oral mucosal breaks seem to be necessary for pathogenicity of oral bacteria in patients with RA, a process facilitated by periodontitis, which is particularly common in these patients[44]. Oral streptococci, citrullinated by neutrophils in the periodontal tissues, enter the bloodstream through a fragile mucosal barrier and activate inflammatory monocytes found in inflamed RA synovia. In addition, these bacteria stimulate B cells to make anti-citrullinated protein antibodies with cross-reactivity against citrullinated host autoantigens found in target joints[45]. Conditioning-related oral mucositis after myeloablative alloHCT can play the same role as periodontitis in patients RA by facilitating bacterial translocation. Although the association between mucositis severity and cGVHD risk did not reach statistical significance in our patients, there was a clear trend, consistent with the finding of moderate-to-severe oral mucositis as an independent risk factor for cGVHD in a recent study with a larger sample size (230 patients)[21]. Whether plaque *S. sanguinis* is citrullinated in these patients and can be found in the circulation or cGVHD target organs is worth further research.

In a murine alloHCT model, and the only published mechanistic study of oral microbiota connection with cGVHD, ligature-induced oral dysbiosis and periodontitis exacerbated cGVHD. Expansion of the Enterococcaceae family in the mouth and its translocation to cervical lymph nodes activated antigen-presenting cells, promoting the expansion of donor-derived inflammatory T cells and worsening of cGVHD. Importantly, ligature removal or antibiotic therapy reversed the process, confirming causality[21]. In the companion human study reported along with this mechanistic murine work, buccal swab microbiota profiling of 31 alloHCT patients (16 with cGVHD) using short-amplicon sequencing showed a tendency for differences between the two groups in their family-level microbiota compositional changes from pre- to post-transplant timepoints. Using shotgun metagenomic sequencing which permitted species level resolution, we here found a significant association between day +28 plaque *S. sanguinis* and cGVHD. Associations between the oral microbiota and other post-alloHCT outcomes (e.g.

survival, immune reconstitution, aGVHD, relapse of the underlying disease) have been demonstrated in recent work[3,4,46].

Our sequencing platform also informed on microbial gene content and functional pathways, allowing us to identify a potential mechanism for our finding. The ethanolamine utilization pathway, enriched in patients with cGVHD, was strongly attributed to *S. sanguinis*. The main source of ethanolamine in humans is diet and cell membranes (of both microbiota and the host). Some bacteria, with *Salmonella typhimurium* being the most extensively studied species, can break down ethanolamine into acetaldehyde and ammonia; the latter can be used as a non-fermentable source of carbon and/or nitrogen[47]. This ability confers a fitness advantage in inflamed tissues by reducing the competition with bacteria that rely entirely on fermentative metabolism[48]. As a plausible model, we hypothesize that ethanolamine utilization in *S. sanguinis* increases its likelihood of surviving in the bloodstream and reaching GVHD target tissues where it can either trigger *de novo* pathogenesis or worsen pre-existing inflammation. Further expansion of *S. sanguinis* in tissues affected by cGVHD due to ethanolamine release from dying cells creates a vicious cycle leading to more damage. The other pathway enriched in patients with cGVHD was the superpathway of glycerol degradation to 1,3-propanediol. In *in vitro* assays, glycerol metabolism using this pathway increased the competitive fitness of *S. sanguinis* against other streptococci[49]. Other than an ability to form biofilm *in vitro* and being present in dental plaque[50], there is little literature on *P. loescheii*, the other species associated with cGVHD in our analysis. Our pathway analysis was not informative with regard to this species. Further research is needed to validate our observation for *P. loescheii* and the potential mechanisms involved.

Limitations of the present work include its small sample size and single-center nature. Validation in other cohorts will be important to support the value of day +28 plaque *S. sanguinis* as a biomarker for cGVHD risk and a precision target for preventative intervention. Ultimately, demonstration of causality will require mechanistic research, either via a randomized trial showing decreased cGVHD risk by decreasing *S. sanguinis* abundance or controlled murine experiments. Whether *S. sanguinis* reaches cGVHD target tissues or stimulates the immune system locally (e.g. in draining cervical lymph nodes) cannot be determined from our work. In addition, we did not find the etiology of enrichment for *S. sanguinis* in day +28 plaque. Although we found evidence for patient-specific tendency to accumulate more *S. sanguinis* (consistent with individuality of plaque microbiota), diet, oral hygiene, and salivary biochemical composition are among other possible determinants of plaque microbiota composition that we did not collect data on. Thorough dietary and oral hygiene surveys and biochemical analysis of saliva could help identify patients more likely to experience expansion of *S. sanguinis* early after alloHCT. Lifestyle modification or simple oral health interventions in such patients might help reduce the risk of cGVHD.

In conclusion, we show an association between day +28 plaque microbiota composition and cGVHD risk. Using shotgun metagenomic sequencing, we identified 2 candidate species and 2 metabolic pathways as potential mechanisms of pathogenesis. If

validated in other cohorts, these findings could support simple topical interventions targeting the oral microbiota as a novel non-immunosuppressive approach to cGVHD prophylaxis. cGVHD is a leading cause of long-term morbidity after alloHCT, thus more effective prevention could significantly improve the quality of life in transplant survivors.

References

1. Zeiser R, Blazar BR. Pathophysiology of Chronic Graft-versus-Host Disease and Therapeutic Targets. N Engl J Med. 2017;377(26):2565-2579.
2. Shono Y, van den Brink MRM. Gut microbiota injury in allogeneic haematopoietic stem cell transplantation. Nat Rev Cancer. 2018;18(5):283-295.
3. Heidrich V, Knebel FH, Bruno JS, et al. Longitudinal analysis at three oral sites links oral microbiota to clinical outcomes in allogeneic hematopoietic stem-cell transplant. Microbiol Spectr. 2023;11(6):e0291023.
4. Heidrich V, Bruno JS, Knebel FH, 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.
5. Atkinson K, Horowitz MM, Gale RP, et al. Risk factors for chronic graft-versus-host disease after HLA-identical sibling bone marrow transplantation. Blood. 1990;75(12):2459-2464.
6. Anasetti C, Logan BR, Lee SJ, et al. Peripheral-blood stem cells versus bone marrow from unrelated donors. N Engl J Med. 2012;367(16):1487-1496.
7. Cutler C, Giri S, Jeyapalan S, Paniagua D, Viswanathan A, Antin JH. Acute and chronic graft-versus-host disease after allogeneic peripheral-blood stem-cell and bone marrow transplantation: a meta-analysis. J Clin Oncol. 2001;19(16):3685-3691.
8. Storb R, Prentice RL, Sullivan KM, et al. Predictive factors in chronic graft-versus-host disease in patients with aplastic anemia treated by marrow transplantation from HLA-identical siblings. Ann Intern Med. 1983;98(4):461-466.
9. Jenq RR, Taur Y, Devlin SM, et al. Intestinal Blautia Is Associated with Reduced Death from Graft-versus-Host Disease. Biol Blood Marrow Transplant. 2015;21(8):1373-1383.
10. Golob JL, Pergam SA, Srinivasan S, et al. Stool Microbiota at Neutrophil Recovery Is Predictive for Severe Acute Graft vs Host Disease After Hematopoietic Cell Transplantation. Clin Infect Dis. 2017;65(12):1984-1991.
11. Burgos da Silva M, Ponce DM, Dai A, et al. Preservation of fecal microbiome is associated with reduced severity of Graft-versus-Host Disease. Blood. 2022;140(22):2385-2397.
12. Stein-Thoeringer CK, Nichols KB, Lazrak A, et al. Lactose drives Enterococcus expansion to promote graft-versus-host disease. Science. 2019;366(6469):1143-1149.
13. Shono Y, Docampo MD, Peled JU, 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. 2016;8(339):339ra71.

14. Rashidi A, Liang L, Gooley T, et al. Microbiota signature of oral chronic graft-versus-host disease 6+ years after transplantation. Haematologica. 2024;109(11):3800-3805.
15. Golob JL, DeMeules MM, Loeffelholz T, et al. Butyrogenic bacteria after acute graft-versus-host disease (GVHD) are associated with the development of steroid-refractory GVHD. Blood Adv. 2019;3(19):2866-2869.
16. Markey KA, Schluter J, Gomes ALC, et al. The microbe-derived short-chain fatty acids butyrate and propionate are associated with protection from chronic GVHD. Blood. 2020;136(1):130-136.
17. Wang Y, Huang L, Huang T, et al. The Gut Bacteria Dysbiosis Contributes to Chronic Graft-Versus-Host Disease Associated With a Treg/Th1 Ratio Imbalance. Front Microbiol. 2022;13:813576.
18. El Jurdi N, Holtan SG, Hoeschen A, et al. Pre-transplant and longitudinal changes in faecal microbiome characteristics are associated with subsequent development of chronic graft-versus-host disease. Br J Haematol. 2023;203(2):288-294.
19. Ebadi M, Gem H, Sebastian G, et al. Different patterns of oral mucositis and Microbiota injury after total body irradiation- versus chemotherapy-based myeloablative allogeneic hematopoietic cell transplantation. Adv Radiat Oncol. 2025;10(6):101787.
20. Rashidi A, Pidala J, Hamilton BK, et al. Oral and gut microbiome alterations in oral chronic graft-versus-host disease: results from Close Assessment and Testing for Chronic GVHD (CATCH study). Clin Cancer Res. Clin Cancer Res. 2024;30(18):4240-4250.
21. Kambara Y, Fujiwara H, Yamamoto A, et al. Oral inflammation and microbiome dysbiosis exacerbate chronic graft-versus-host disease. Blood. 2025;145(8):881-896.
22. Jeong SH, Nam Y, Jung H, et al. Interrupting oral infection of Porphyromonas gingivalis with anti-FimA antibody attenuates bacterial dissemination to the arthritic joint and improves experimental arthritis. Exp Mol Med. 2018;50(3):e460.
23. Chukkapalli S, Rivera-Kweh M, Gehlot P, et al. Periodontal bacterial colonization in synovial tissues exacerbates collagen-induced arthritis in B10.RIII mice. Arthritis Res Ther. 2016;18(1):161.
24. Atarashi K, Suda W, Luo C, et al. Ectopic colonization of oral bacteria in the intestine drives TH1 cell induction and inflammation. Science. 2017;358(6361):359-365.
25. Read E, Curtis MA, Neves JF. The role of oral bacteria in inflammatory bowel disease. Nat Rev Gastroenterol Hepatol. 2021;18(10):731-742.
26. Kitamoto S, Nagao-Kitamoto H, Jiao Y, et al. The intermucosal connection between the mouth and gut in commensal pathobiont-driven colitis. Cell. 2020;182(2):447-462.e14.
27. Kimura M, Araoka H, Yoshida A, et al. Breakthrough viridans streptococcal bacteremia in allogeneic hematopoietic stem cell transplant recipients receiving levofloxacin prophylaxis in a Japanese hospital. BMC Infect Dis. 2016;16:372.
28. Gem H, Ebadi M, Sebastian G, et al. A sex-dependent salivary bacterium influences oral mucositis severity after allogeneic hematopoietic cell transplantation. NPJ Biofilms Microbiomes. 2024;10(1):140.

29. Ximénez-Fyvie LA, Haffajee AD, Socransky SS. Microbial composition of supra- and subgingival plaque in subjects with adult periodontitis. J Clin Periodontol. 2000;27(10):722-732.
30. Robien K, Schubert MM, Bruemmer B, Lloid ME, Potter JD, Ulrich CM. Predictors of oral mucositis in patients receiving hematopoietic cell transplants for chronic myelogenous leukemia. J Clin Oncol. 2004;22(7):1268-1275.
31. Wardley AM, Jayson GC, Swindell R, et al. Prospective evaluation of oral mucositis in patients receiving myeloablative conditioning regimens and haemopoietic progenitor rescue. Br J Haematol. 2000;110(2):292-299.
32. Jagasia MH, Greinix HT, Arora M, 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. 2015;21(3):389-401.e1.
33. Shannon CE, Weaver W. The mathematical theory of communication. Urbana, IL: The University of Illinois Press, 1-117.
34. Aitchison J, Barceló-Vidal C, Martín-Fernández JA, Pawlowsky-Glahn V. Logratio Analysis and Compositional Distance. Math Geol. 2000;32(3):271-275.
35. Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics studies. PLoS Comput Biol. 2021;17(11):e1009442.
36. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Series B Stat Methodol 1995;57(1):289-300.
37. Zhou H, He K, Chen J, Zhang X. LinDA: linear models for differential abundance analysis of microbiome compositional data. Genome Biol. 2022;23(1):95.
38. Utter DR, Mark Welch JL, Borisy GG. Individuality, stability, and variability of the plaque microbiome. Front Microbiol. 2016;7:564.
39. Zhang X, Zhang D, Jia H, et al. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. Nat Med. 2015;21(8):895-905.
40. Carlsson J, Grahnén H, Jonsson G, Wikner S. Establishment of Streptococcus sanguis in the mouths of infants. Arch Oral Biol. 1970;15(12):1143-1148.
41. de Aquino SG, Abdollahi-Roodsaz S, Koenders MI, et al. Periodontal pathogens directly promote autoimmune experimental arthritis by inducing a TLR2- and IL-1-driven Th17 response. J Immunol. 2014;192(9):4103-4111.
42. Kolenbrander PE, Palmer RJ Jr, Rickard AH, Jakubovics NS, Chalmers NI, Diaz PI. Bacterial interactions and successions during plaque development. Periodontol 2000. 2006;42(1):47-79.
43. Moentadj R, Wang Y, Bowerman K, et al. Streptococcus species enriched in the oral cavity of patients with RA are a source of peptidoglycan-polysaccharide polymers that can induce arthritis in mice. Ann Rheum Dis. 2021;80(5):573-581.
44. Bolstad AI, Fevang B-TS, Lie SA. Increased risk of periodontitis in patients with rheumatoid arthritis: A nationwide register study in Norway. J Clin Periodontol. 2023;50(8):1022-1032.
45. Brewer RC, Lanz TV, Hale CR, et al. Oral mucosal breaks trigger anti-citrullinated bacterial and human protein antibody responses in rheumatoid arthritis. Sci Transl Med. 2023;15(684):eabq8476.

46. de Molla VC, Heidrich V, Bruno JS, et al. Disruption of the oral microbiota is associated with a higher risk of relapse after allogeneic hematopoietic stem cell transplantation. Sci Rep. 2021;11(1):17552.
47. Garsin DA. Ethanolamine utilization in bacterial pathogens: roles and regulation. Nat Rev Microbiol. 2010;8(4):290-295.
48. Thiennimitr P, Winter SE, Winter MG, et al. Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota. Proc Natl Acad Sci U S A. 2011;108(42):17480-17485.
49. Taylor ZA, Chen P, Noeparvar P, et al. Glycerol metabolism contributes to competition by oral streptococci through production of hydrogen peroxide. J Bacteriol. 2024;206(9):e0022724.
50. Albaghdadi SZ, Altaher JB, Drobiova H, Bhardwaj RG, Karched M. In vitro Characterization of Biofilm Formation in Prevotella Species. Front Oral Health. 2021;2:724194.

Table 1: Patient characteristics

Characteristic	Cases	Controls	Total	P
Total, N	11	26	37	
Sex, N (%)				0.73
Male	5 (45%)	10 (38%)	15 (41%)	
Female	6 (55%)	16 (62%)	22 (59%)	
Age, years; median (range)	43 (20-60)	41 (24-55)	43 (20-60)	0.40
Female to male transplant, N (%)	5 (45%)	2 (8%)	7 (19%)	0.02
Underlying disease, N (%)				0.40
Acute leukemia	8 (73%)	22 (85%)	30 (81%)	
Other	3 (27%)	4 (15%)	7 (19%)	
Conditioning regimen, N (%)				1.00
Cy/TBI (12 Gy)	5 (46%)	10 (38%)	15 (41%)	
Flu/Cy/TBI (12-13.2 Gy)	0	1 (4%)	1 (3%)	
Flu/Thiotepa/TBI (13.2 Gy)	0	1 (4%)	1 (3%)	
Bu/Cy	4 (36%)	4 (15%)	8 (22%)	
Flu/Bu4	2 (18%)	4 (15%)	6 (16%)	
Flu/Bu4/TBI (4 Gy)	0	3 (12%)	3 (8%)	
Flu/Cy/Thiotepa/TBI (4 Gy)	0	3 (12%)	3 (8%)	
Donor type, N (%)				0.07
HLA-matched sibling	6 (55%)	5 (19%)	11 (30%)	
HLA-matched unrelated	5 (45%)	17 (65%)	22 (59%)	
Cord blood	0	4 (15%)	4 (11%)	
Graft source, N (%)				0.30
Peripheral blood	11 (100%)	22 (85%)	33 (89%)	
Cord blood	0	4 (15%)	4 (11%)	
GVHD prophylaxis, N (%)				0.36
Tacrolimus/Methotrexate	10 (91%)	17 (65%)	27 (73%)	
PTCy-based	1 (9%)	5 (19%)	6 (16%)	
Cyclosporine/MMF	0	4 (15%)	4 (11%)	
Grade II-IV acute GVHD, N (%)	6 (55%)	9 (35%)	15 (41%)	0.30
Onset by day +28	2	4	6	
Use of total parenteral nutrition, N (%)	3 (27%)	14 (54%)	17 (46%)	0.17

Cases and controls are defined as patients with moderate/severe and mild/no chronic GVHD, respectively. Some of the percentages do not sum up precisely to 100% due to rounding. P values were derived from a Fisher's exact test for categorical variables and a Wilcoxon's test for continuous variables. Bu: busulfan; Cy: cyclophosphamide; Flu: fludarabine; GVHD: graft-versus-host disease; HLA: human leukocyte antigen; MMF: mycophenolate mofetil; PTCy: post-transplantation cyclophosphamide; TBI: total body irradiation

Fig. 1: Study schema and antibiotic exposures
(**a**) Study schema for the umbrella protocol. Samples used in the present analysis are shown in red. (**b**) Antibacterial antibiotic exposures between days -7 and +28. The 6 most commonly used antibiotics are shown. Ceph 1-2: first or second generation cephalosporins; Ceph 3+: third or higher generation cephalosporins; FQN: fluoroquinolones; PCN: penicillins; TMP-SMX: trimethoprim-sulfamethoxazole

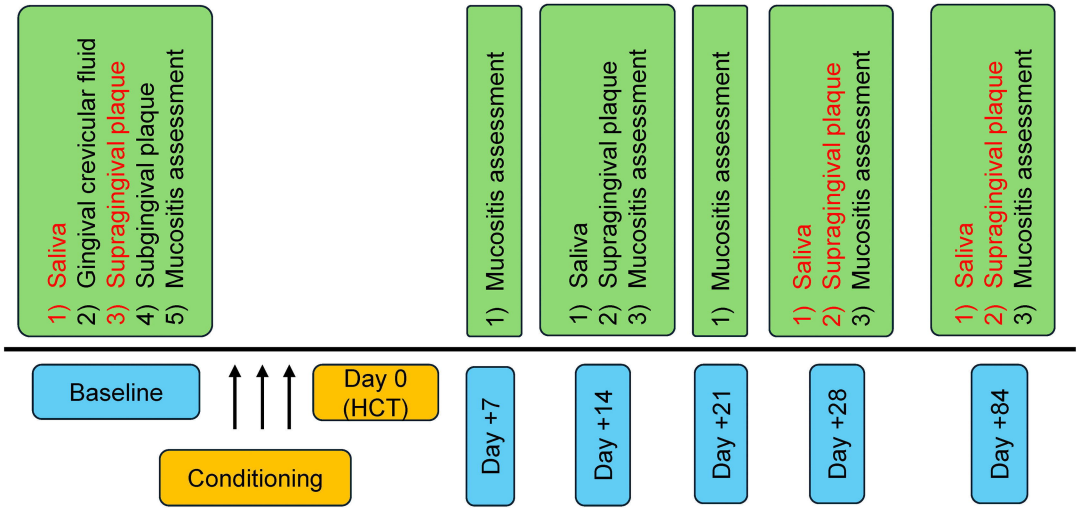
Fig. 2: Microbiota composition and chronic GVHD
(**a-f**) Alpha diversity as measured by the Shannon's index, compared between cases and controls at baseline (supragingival plaque in a; saliva in d), day +28 (supragingival plaque in b; saliva in e), and day +84 (supragingival plaque in c; saliva in f). P values are from a Wilcoxon's test. Each box shows the median (horizontal middle line) and interquartile range. Whisker lines indicate non-outlier maximum and minimum values. A small jitter is included for better visualization. (**g-l**) Principal coordinate analysis by applying principal component analysis on species-level Aitchison distances. Supragingival plaque microbiota is compared between cases and controls at baseline (g), day +28 (h), and day +84 (i); Salivary microbiota is compared between cases and controls at baseline (j), day +28 (k), and day +84 (l). The first two principal coordinates (PC1 and PC2) are shown, with numbers in parentheses indicating percent variation explained by the corresponding axis. P values are from an adonis test with 999 permutations. Cases and controls in all panels are defined as patients with developing moderate/severe and mild/no chronic GVHD, respectively.

Fig. 3: Differential abundance analysis for day +28 plaque microbiota
(**a-b**) Volcano plots showing the results of differential abundance analysis for day +28 plaque microbiota using MaAsLin2 (a) and LinDA (b). Each point shows a species. Points above the horizontal line (-log10(0.05) = 1.30 and corresponding to q = 0.05) are statistically significant. Species to the right (left) are enriched among cases (controls). To derive q values, P values were corrected using the Benjamini-Hochberg method. (**c-d**) Comparison between cases and controls for the relative abundance (in day +28 plaque microbiota) of the 2 species identified by both methods in (a-b). P values are from a Wilcoxon's test. Each box shows the median (horizontal middle line) and interquartile range. Whisker lines indicate non-outlier maximum and minimum values. A small jitter is included for better visualization. Cases and controls are defined as patients with moderate/severe and mild/no chronic GVHD, respectively.

Fig. 4: Microbiota functional pathways and chronic GVHD
(**a**) Volcano plot showing the results of differential abundance analysis for day +28 plaque microbiota using MaAsLin2. Each point shows a pathway. Points above the horizontal red line (-log10(0.01) = 2 and corresponding to P = 0.01) are statistically significant. Species to the right (left) are enriched among cases (controls). The vertical red line narrows down the enriched species to those with more substantial enrichment, defined as having more than two-fold abundance among cases. (**b-c**) Two enriched

pathways (among cases) in panel (a) were found in *S. sanguinis*. All species contributions to these pathways are shown. HUMAnN's pathway abundance tables (normalized to copies per million) include the number of copies of the entire pathway of interest assigned to each species. These numbers were used to quantify species contributions. Each column is a sample. The top panels show controls and the bottom panels show cases. *S. sanguinis* made the second top contributor to PWY0-1477 and the only classified contributor to GOLPDLCAT-PWY. (**d-e**) Contribution of *S. sanguinis* to the 2 pathways in panels (b-c). (**f**) Receiver operating curve characteristic analysis for day +28 plaque *S. sanguinis* relative abundance predicting case vs. control. The area under the curve (AUC) and its 95% confidence interval (CI) are shown. (**g**) Correlation between baseline and day +28 plaque *S. sanguinis* relative abundance in the same patient. The Spearman's correlation coefficient and its corresponding P value are shown. (**h**) Comparison between cases and controls for the relative abundance of *S. sanguinis* in baseline plaque microbiota. (**i**) Receiver operating curve characteristic analysis for baseline plaque *S. sanguinis* relative abundance predicting case vs. control. The AUC and its 95%CI are shown. (**j**) *S. sanguinis* dynamics. Relative abundances of this species at each timepoint is shown, with samples from the same patient connected. A loess smoother is added for each cohort as a thicker line. (**k**) *P. loescheii* dynamics. Relative abundances of this species at each timepoint is shown, with samples from the same patient connected. A loess smoother is added for each cohort as a thicker line. Each box in panels (d-e, h) shows the median (horizontal middle line) and interquartile range. Whisker lines indicate non-outlier maximum and minimum values. A small jitter is included for better visualization. P values in these panels are from a Wilcoxon's test. Cases and controls are defined as patients with moderate/severe and mild/no chronic GVHD, respectively.

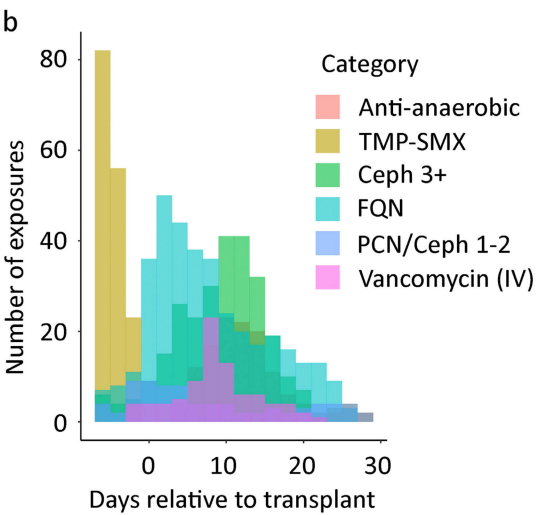
a

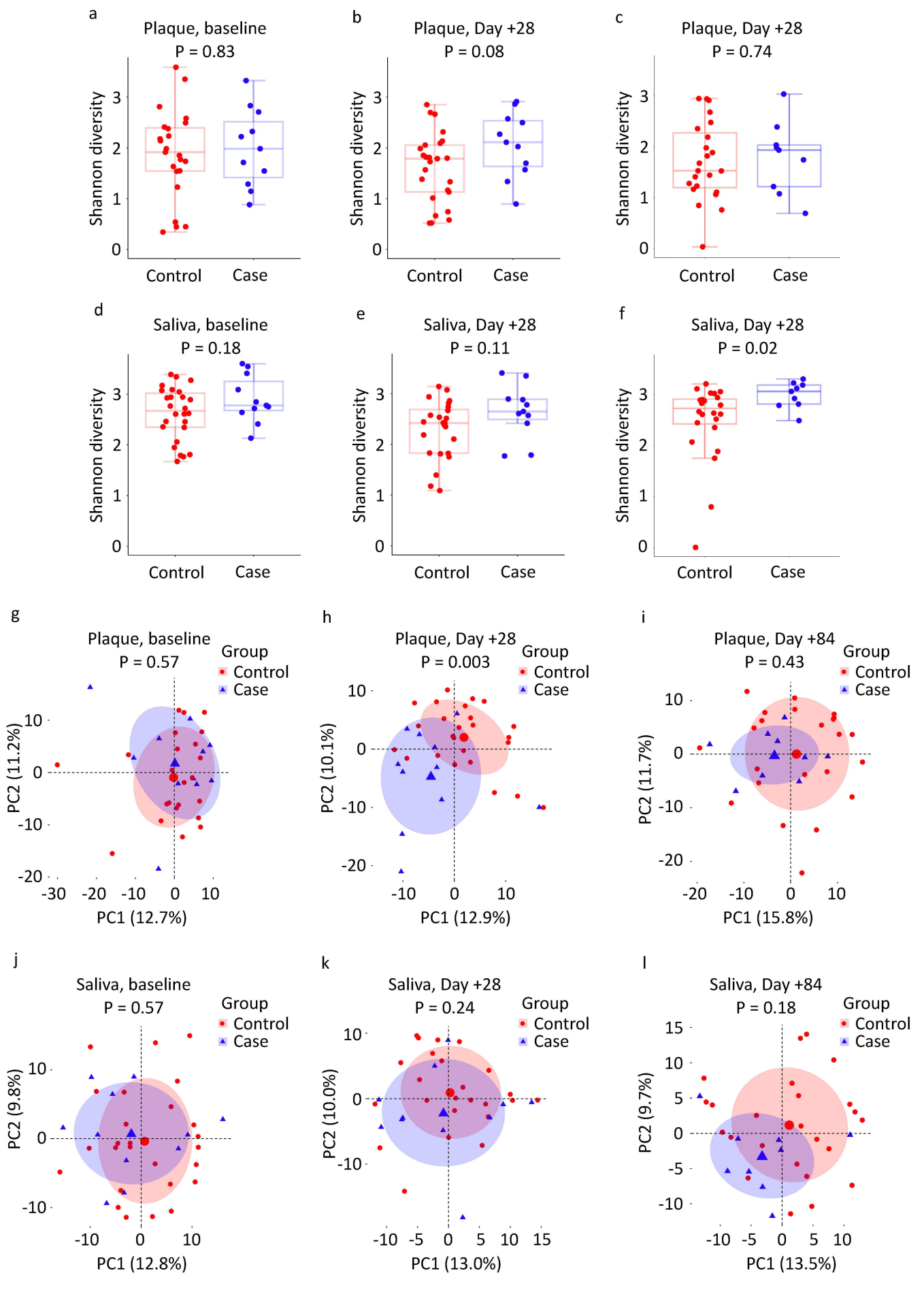


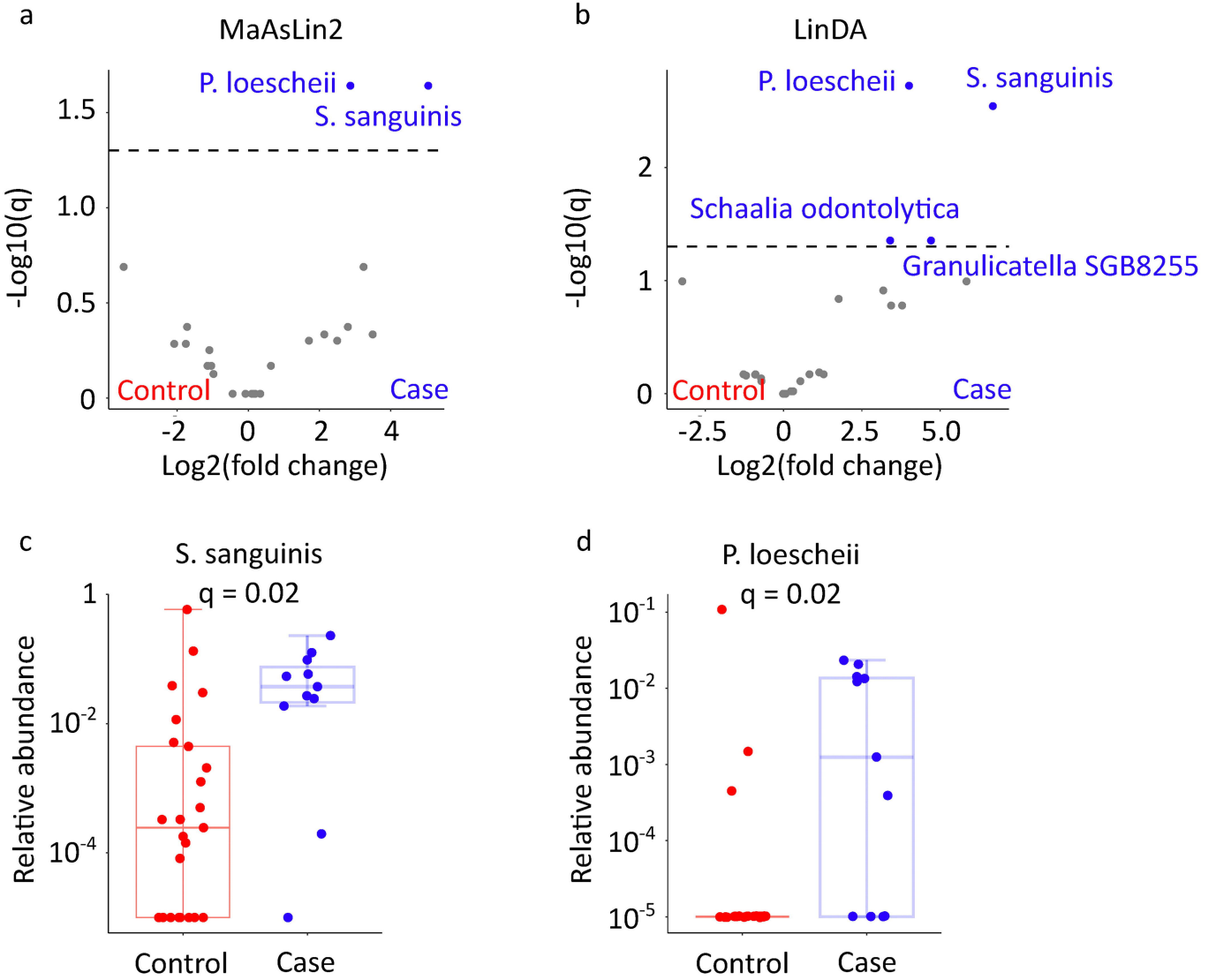
Samples used in the present work

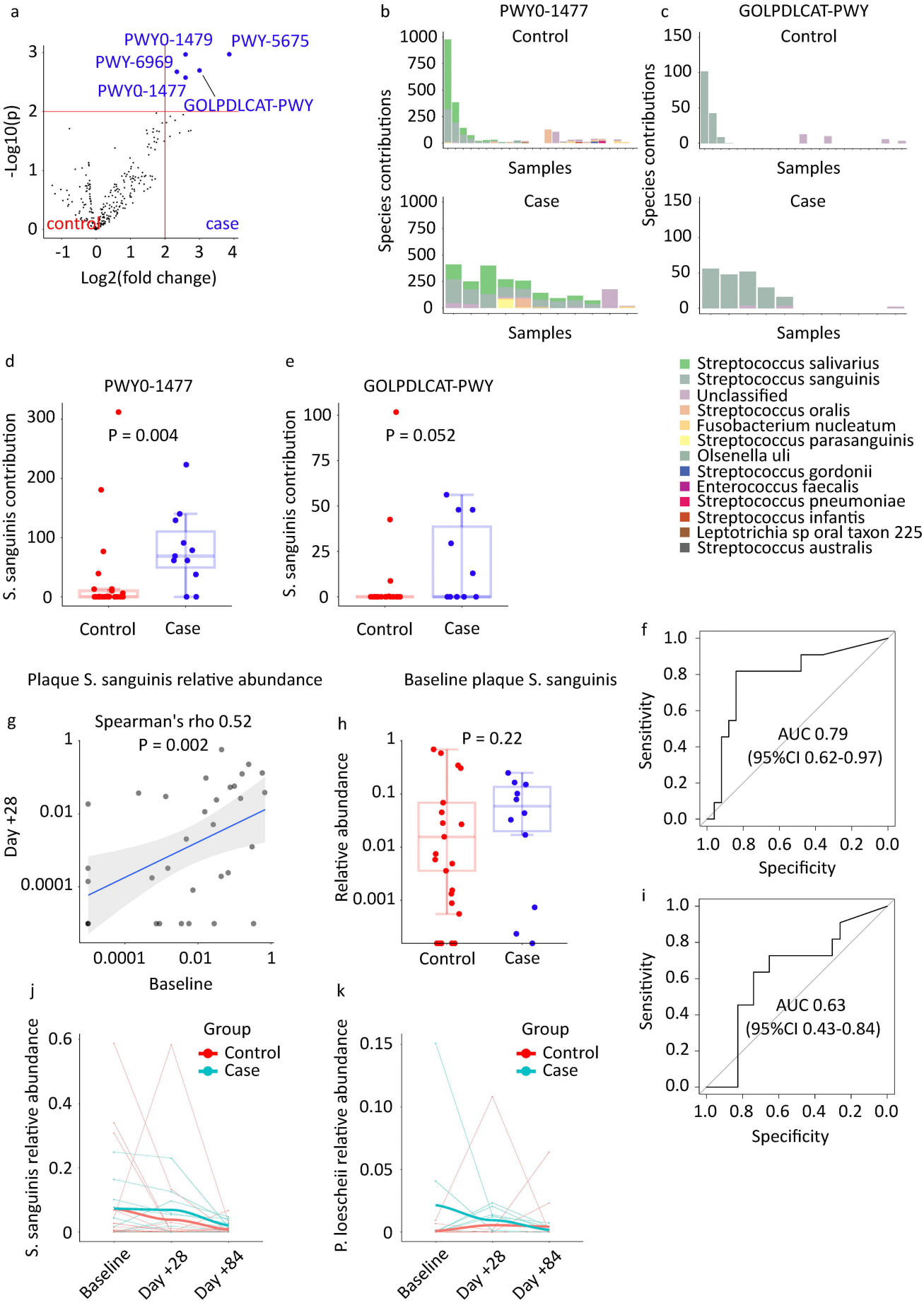
	Baseline	Day +28	Day +84
Saliva	37 (11 cases, 26 controls)	35 (11 cases, 24 controls)	33 (9 cases, 24 controls)
Plaque	34 (11 cases, 23 controls)	36 (11 cases, 25 controls)	32 (9 cases, 23 controls)

b




a
Plaque, baseline
P = 0.83
b
Plaque, Day +28
P = 0.08
c
Plaque, Day +28
P = 0.74
d
Saliva, baseline
P = 0.18
e
Saliva, Day +28
P = 0.11
f
Saliva, Day +28
P = 0.02
Shannon diversity
Control
Case
g
Plaque, baseline
P = 0.57
h
Plaque, Day +28
P = 0.003
i
Plaque, Day +84
P = 0.43
j
Saliva, baseline
P = 0.57
k
Saliva, Day +28
P = 0.24
l
Saliva, Day +84
P = 0.18
Group
Control
Case
PC1 (12.7%)
PC2 (11.2%)
PC1 (12.9%)
PC2 (10.1%)
PC1 (15.8%)
PC2 (11.7%)
PC1 (12.8%)
PC2 (9.8%)
PC1 (13.0%)
PC2 (10.0%)
PC1 (13.5%)
PC2 (9.7%)


a
MaAsLin2
P. loescheii
S. sanguinis
-Log10(q)
1.5
1.0
0.5
0
Control
Case
-2
0
2
4
Log2(fold change)
b
LinDA
P. loescheii
S. sanguinis
Schaalia odontolytica
Granulicatella SGB8255
-Log10(q)
2
1
0
Control
Case
-2.5
0
2.5
5.0
Log2(fold change)
c
S. sanguinis
q = 0.02
Relative abundance
1
10^{-2}
10^{-4}
Control
Case
d
P. loescheii
q = 0.02
Relative abundance
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
Control
Case


a
PWY0-1479
PWY-5675
PWY-6969
PWY0-1477
GOLPDLCAT-PWY
control
case
-Log10(p)
Log2(fold change)
b
PWY0-1477
Control
Case
Species contributions
Samples
c
GOLPDLCAT-PWY
Control
Case
Species contributions
Samples
Streptococcus salivarius
Streptococcus sanguinis
Unclassified
Streptococcus oralis
Fusobacterium nucleatum
Streptococcus parasanguinis
Olsenella uli
Streptococcus gordonii
Enterococcus faecalis
Streptococcus pneumoniae
Streptococcus infantis
Leptotrichia sp oral taxon 225
Streptococcus australis
d
PWY0-1477
P = 0.004
S. sanguinis contribution
Control
Case
e
GOLPDLCAT-PWY
P = 0.052
S. sanguinis contribution
Control
Case
f
Sensitivity
Specificity
AUC 0.79
(95%CI 0.62-0.97)
Plaque S. sanguinis relative abundance
g
Spearman's rho 0.52
P = 0.002
Day +28
Baseline
Baseline plaque S. sanguinis
h
P = 0.22
Relative abundance
Control
Case
i
Sensitivity
Specificity
AUC 0.63
(95%CI 0.43-0.84)
j
Group
Control
Case
S. sanguinis relative abundance
Baseline
Day +28
Day +84
k
Group
Control
Case
P. loescheii relative abundance
Baseline
Day +28
Day +84

Supplementary material

Supplementary Methods

Standard pre-transplant dental examination at Fred Hutchinson Cancer Center
The primary purpose of the standard pre-transplant dental examination is to prevent oral complications of alloHCT with a particular focus on recognizing and eliminating any sources of possible opportunistic infection[1]. The protocol for this examination includes a comprehensive review of dental and medical history, a head and neck examination (i.e., palpation of structures such as the lymph nodes, salivary glands, sinuses, muscles of mastication, and the temporomandibular joints), dental radiography (e.g., a panoramic radiograph with supplemental bitewing and/or periapical radiographs or a full mouth series), oral mucosal examination, caries assessment, and periodontal evaluation[1,2]. Patients are also counseled on proper oral care throughout the transplant timeline (i.e., proper techniques for brushing, flossing, and rinsing if necessary)[3,4]. This examination is typically completed 2-4 weeks prior to transplantation, and upon completion, a full dental treatment plan is completed (i.e., cleaning, restoration, root canal therapy, extraction, etc.) with enough time for recovery prior to transplant. Dental treatment planning is consistent with the Multinational Association of Supportive Care in Cancer/International Society of Oral Oncology (MASCC/ISOO) and the European Society for Blood and Marrow Transplantation (EBMT) Position Paper on Basic Oral Care for hematology-oncology patients and hematopoietic stem cell transplantation recipients[4].

Dental cleanings in our center include dental prophylaxis, scaling and root planing, and periodontal maintenance based on a patient's periodontal status (which is determined by an Oral Medicine specialist working closely with a registered dental hygienist who performs periodontal evaluation as part of all pre-transplant dental evaluations). Dental cleaning is performed to decrease risk for oral source infection in the peri-transplant period and are treatment planned for patients who are judged to benefit from periodontal intervention (e.g., those with gingival inflammation/infection in relation to local plaque and calculus accumulation, periodontal pocketing ≥4mm, or chronic unmanaged periodontal disease). Dental cleaning is performed by a registered dental hygienist using an ultrasonic scaler, supplemented with hand instrumentation. High volume suction is utilized to minimize risk for aspiration. Anti-infective protocols, including antibiotic prophylaxis and pre-operative rinsing with chlorhexidine gluconate, are utilized when patients have an absolute neutrophil count (ANC) <1.0, if a Hickman line or Port-au-Cath is in place, and/or if the patient has cardiac indications for antibiotic prophylaxis based on American Heart Association guidelines for prevention of infective endocarditis. Antibiotic choice and timing are based on the same AHA guidelines[5,6]. Local anesthesia is utilized when scaling and root planing is performed or when patient comfort would otherwise limit calculus removal. Oral sedation with a one-time pre-treatment dose of a benzodiazepine is used in a limited capacity for patients with high levels of dental anxiety (in coordination with WA State Law related to oral sedation)[7]. Oral hygiene instruction is also performed by a registered dental hygienist to help optimize local plaque control following dental procedures.

Mucositis assessment

Mucositis assessments were performed by 4 calibrated oral medicine specialists using the Oral Mucositis Assessment Scale (OMAS)[8] at baseline, day +7, day +14, day +21 (if the patient was still in the hospital), day +28, and day +84. OMAS evaluates 9 intraoral sites (upper and lower lip, left and right buccal mucosa, left and right ventrolateral tongue, floor of mouth, and hard and soft palates) for severity of ulceration/pseudomembrane (range 0-3) and erythema (range 0-2). Total mucositis score (range 0-45) at each timepoint, defined as the sum of the scores (ulceration/pseudomembrane and erythema) across all sites[9], was used as a quantitative continuous variable estimating mucositis severity at the corresponding timepoint.

Sample collection and processing

Samples were collected after at least 30 minutes of no oral intake or oral hygiene. Up to 5 mL of saliva was collected by passive drooling into a sterile tube containing 5 mL of sterile 95% ethanol. The procedure was stopped after a maximum of 15 min. The tube was pulse vortexed for 5 seconds to ensure homogenous mixture of saliva and ethanol before storage at -80°C. Plaque samples were taken from 3 teeth and transferred from the scaler to the tip of a sterilized plastic pick. Samples contaminated with blood were discarded and taken again from an adjacent tooth. The pick was submerged into a sterile tube containing 500 µL of sterile 95% ethanol and agitated until the clump of plaque fell off the instrument into ethanol. All 3 plaque samples from the same patient were pooled into the same tube. Samples were transferred to -80°C on the day of collection. The use of 95% ethanol for storing oral samples has been validated and allows long-term DNA preservation at room temperature[10,11].

Microbiota sequencing and pathway analysis

DNA was extracted using the ZymoBIOMICS®-96 MagBead DNA Kit (Zymo Research, Irvine, CA). Sequencing libraries were prepared using the Illumina® DNA Library Preparation Kit (Illumina, San Diego, CA) following the manufacturer's protocol and with unique dual-index 10 bp barcodes with Nextera® adapters. All libraries were pooled in equal abundance and the final pool was quantified using qPCR and TapeStation® (Agilent Technologies, Santa Clara, CA). The final library was sequenced on an Illumina NovaSeq 6000 using a S2-300 flow cell and a PE150 configuration. The ZymoBIOMICS® Microbial Community DNA Standard (Zymo Research, Irvine, CA) was used as a positive control for each library preparation. Multiple negative controls (i.e., blank extraction control, blank library preparation control) were included to assess the level of bioburden carried by the wet-lab process. The criteria for the negative controls for a pass run and findings in the present study are summarized in the Table below:

Metric	Criteria	Result
Extracted DNA concentration	<0.5 ng/uL	Negative extraction control #1 = 0.146 ng/uL Negative extraction control #2 = 0.29 ng/uL
Sublibrary DNA concentration	<3 ng/uL	Sublibrary- negative extraction control #1 = 0.5 ng/uL Sublibrary- negative extraction control #2 = 0.8 ng/uL
Read count	<100,000	Negative extraction control #1 = 856 reads Negative extraction control #2 = 51 reads

Raw paired-end metagenomic sequence reads were quality-processed using the integrated pipeline provided in KneadData v.0.12.0. This sequence-level procedure included two main steps: (*i*) removal of reads mapped to the human reference genome GRCh37 (hg19) using Bowtie2 v.2.4.5[12] and (*ii*) removal of adapter sequences and low-quality reads using Trimmomatic v.0.39[13] with default settings. Output files consisting of surviving paired and orphan reads were concatenated and used as input to MetaPhlAn4[14]. MetaPhlAn4 with default parameters was used for species-level taxonomic assignment. MetaPhlAn4 uses a set of species-level genome bins (SGBs)[15] as primary taxonomic units and accurately profiles their presence and abundance in metagenomes. The latest version (version 4) of MetaPhlAn uses a database containing ~5.1 million unique clade-specific marker genes for 21,978 existing SGBs and 4,992 yet-to-be-characterized SGBs (defined solely based on metagenome-assembled genomes).

Pathways and gene family abundances were profiled using HUMAnN v3.6[16]. HUMAnN's tiered search occurs in 3 phases. First, it identifies community species using MetaPhlAn and its clade-specific marker genes. Then, it maps KneadData-processed reads against the pangenomes of the identified species using Bowtie2[36]. Finally, it aligns unmapped reads to a comprehensive, non-redundant protein database (EC-filtered UniRef90) using DIAMOND[17]. HUMAnN performs read-count-based quantification of the microbial gene families and functional pathways present within each sample on both per-species and community-level basis. Pathways were annotated using MetaCyc v24.0 definitions[18] and gene families using UniRef90 definitions[19]. HUMAnN's default reads per kilobase values for gene family and pathway abundances were normalized to copies per million (CPM).

References

1. Centers for Disease Control and Prevention; Infectious Diseases Society of America; American Society of Blood and Marrow Transplantation. Guidelines for preventing opportunistic infections among hematopoietic stem cell transplant recipients. Biol Blood Marrow Transplant 2000;6(6a):7-83.
2. Dean D, Lee SJ, Cutler C, et al. Dental evaluation and clearance prior to allogeneic hematopoietic cell transplantation. Oral Dis 2024;30(4):2635-2644.
3. Raber-Durlacher JE, Zadik Y, Treister NS, et al. MASCC/ISOO clinical practice statement: Current understanding on controversies in basic oral care in hemato-oncology and hematopoietic cell transplantation. Support Care Cancer 2024 25;32(8):550.
4. Elad S, Raber-Durlacher JE, Brennan MT, 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 (MASCC/ISOO) and the European Society for Blood and Marrow Transplantation (EBMT). Support Care Cancer 2015;23(1):223-236.
5. Wilson WR, Gewitz M, Lockhart PB, et al; American Heart Association Young Hearts Rheumatic Fever, Endocarditis and Kawasaki Disease Committee of the Council on Lifelong Congenital Heart Disease and Heart Health in the Young; Council on Cardiovascular and Stroke Nursing; and the Council on Quality of Care and Outcomes Research. Prevention of Viridans Group Streptococcal Infective Endocarditis: A

Scientific Statement From the American Heart Association. Circulation 2021;143(20):e963-e978.
6. Wilson W, Taubert KA, Gewitz M, et al; American Heart Association Rheumatic Fever, Endocarditis, and Kawasaki Disease Committee; American Heart Association Council on Cardiovascular Disease in the Young; American Heart Association Council on Clinical Cardiology; American Heart Association Council on Cardiovascular Surgery and Anesthesia; Quality of Care and Outcomes Research Interdisciplinary Working Group. Prevention of infective endocarditis: guidelines from the American Heart Association: a guideline from the American Heart Association Rheumatic Fever, Endocarditis, and Kawasaki Disease Committee, Council on Cardiovascular Disease in the Young, and the Council on Clinical Cardiology, Council on Cardiovascular Surgery and Anesthesia, and the Quality of Care and Outcomes Research Interdisciplinary Working Group. Circulation 2007;116(15):1736-54.
7. Washington Administrative Code. (n.d.). WAC 246-817-745: Continuing education requirements.
8. Sonis ST, Eilers JP, Epstein JB, et al. Validation of a new scoring system for the assessment of clinical trial research of oral mucositis induced by radiation or chemotherapy. Mucositis Study Group. Cancer 1999;85(10):2103–2113.
9. Schubert MM, Williams BE, Lloid ME, Donaldson G, Chapko MK. Clinical assessment scale for the rating of oral mucosal changes associated with bone marrow transplantation. Development of an oral mucositis index. Cancer 1992;69(10):2469–2477.
10. Marotz C, Cavagnero KJ, Song SJ, et al. Evaluation of the Effect of Storage Methods on Fecal, Saliva, and Skin Microbiome Composition. mSystems;6(2.):
11. Zouiouich S, Byrd DA, Hua X, et al. Stability of the Fecal and Oral Microbiome over 2 Years at -80°C for Multiple Collection Methods. Cancer Epidemiol Biomarkers Prev 2023;32(3):444–451.
12. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9(4):357–359.
13. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30(15):2114–2120.
14. Blanco-Míguez A, Beghini F, Cumbo F, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nat Biotechnol 2023;41(11):1633–1644.
15. Pasolli E, Asnicar F, Manara S, et al. Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. Cell 2019;176(3):649–662.e20.
16. Beghini F, McIver LJ, Blanco-Míguez A, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife;10.
17. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods 2015;12(1):59–60.
18. Caspi R, Billington R, Keseler IM, et al. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. Nucleic Acids Res 2020;48(D1):D445–D453.
19. Suzek BE, Huang H, McGarvey P, Mazumder R, Wu CH. UniRef: comprehensive and non-redundant UniProt reference clusters. Bioinformatics 2007;23(10):1282–1288.

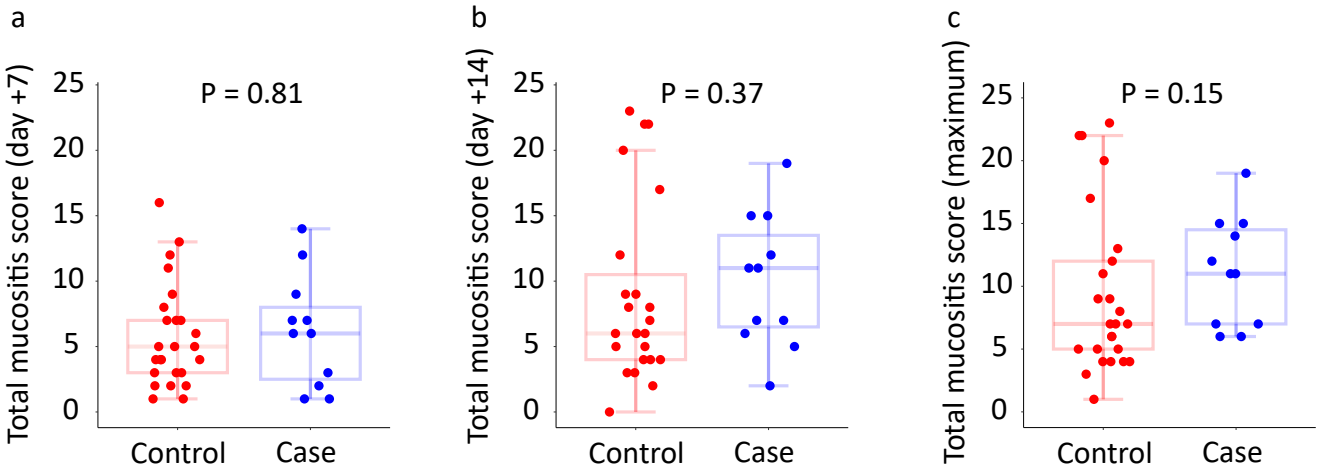

a
Total mucositis score (day +7)
25
20
15
10
5
0
P = 0.81
Control
Case
b
Total mucositis score (day +14)
25
20
15
10
5
0
P = 0.37
Control
Case
c
Total mucositis score (maximum)
25
20
15
10
5
0
P = 0.15
Control
Case

Fig. S1: Oral mucositis severity and chronic GVHD
Comparison between cases and controls for oral mucositis severity, using total mucositis score on the OMAS scale, at day +7 (**a**), day +14 (**b**), and maximum of the two (**c**). P values are from a Wilcoxon's test. Each box shows the median (horizontal middle line) and interquartile range. Whisker lines indicate non-outlier maximum and minimum values. A small jitter is included for better visualization. Cases and controls are defined as patients with moderate/severe and mild/no chronic GVHD, respectively.

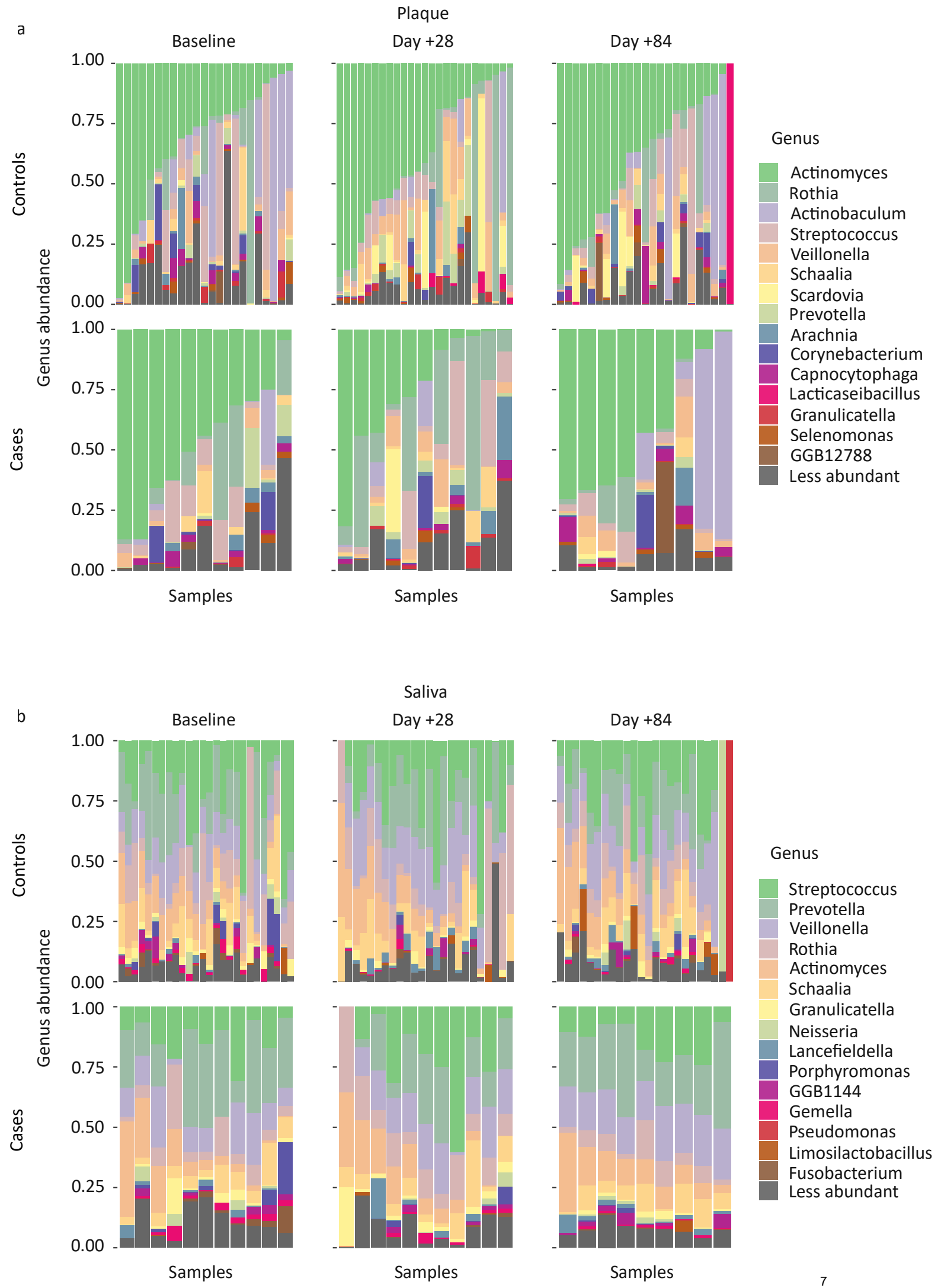

a
Plaque
Baseline
Day +28
Day +84
Controls
Cases
Genus abundance
1.00
0.75
0.50
0.25
0.00
Samples
Genus
Actinomyces
Rothia
Actinobaculum
Streptococcus
Veillonella
Schaalia
Scardovia
Prevotella
Arachnia
Corynebacterium
Capnocytophaga
Lacticaseibacillus
Granulicatella
Selenomonas
GGB12788
Less abundant
b
Saliva
Baseline
Day +28
Day +84
Controls
Cases
Genus abundance
1.00
0.75
0.50
0.25
0.00
Samples
Genus
Streptococcus
Prevotella
Veillonella
Rothia
Actinomyces
Schaalia
Granulicatella
Neisseria
Lancefieldella
Porphyromonas
GGB1144
Gemella
Pseudomonas
Limosilactobacillus
Fusobacterium
Less abundant

Fig. S2: Microbiota composition

(**a**) Relative abundance of the 15 most abundant genera in baseline (left), day +28 (middle), and day +84 (right) supragingival plaque microbiota among controls (top) and cases (bottom). (**b**) Relative abundance of the 15 most abundant genera in baseline (left), day +28 (middle), and day +84 (right) salivary microbiota among controls (top) and cases (bottom). Each column indicates a sample. The bars show relative abundances of different genera. Cases and controls are defined as patients subsequently developing moderate/severe and mild/no chronic GVHD, respectively.

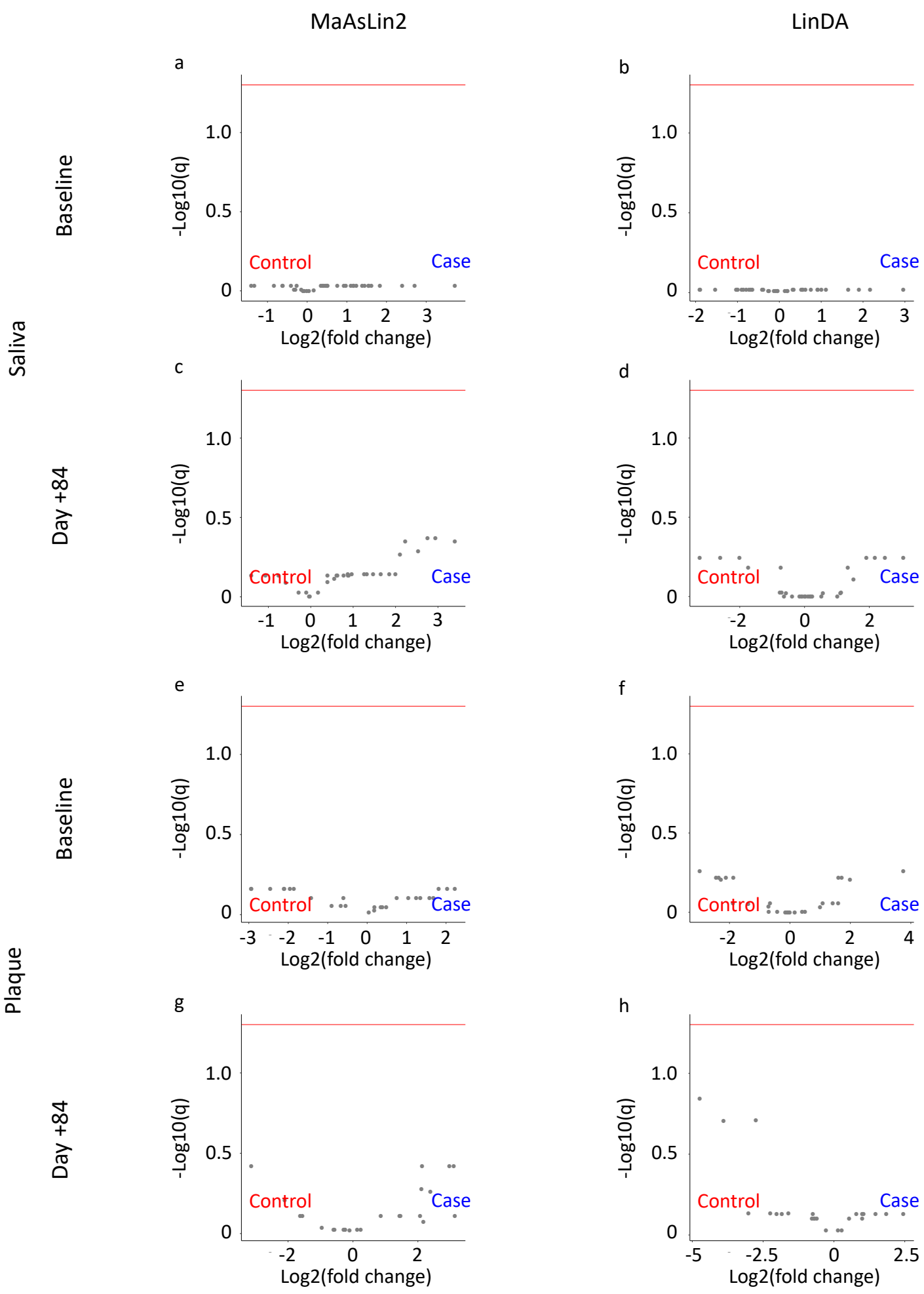

MaAsLin2
LinDA
Saliva
Plaque
Baseline
Day +84
a
b
c
d
e
f
g
h
-Log10(q)
Log2(fold change)
Control
Case

Fig. S3: Differential abundance analysis for salivary and plaque microbiota at baseline and day +84

(**a-d**) Volcano plots showing the results of differential abundance analysis for salivary microbiota at baseline (panels a-b) and day +84 (panels c-d). (**e-h**) Volcano plots showing the results of differential abundance analysis for supragingival plaque microbiota at baseline (panels e-f) and day +84 (panels g-h). Results from MaAsLin2 are shown on the left (panels a, c, e, and g); those from LinDA are shown on the right (panels b, d, f, and h). Each point shows a species. Points above the horizontal line (-log10(0.05) = 1.30 and corresponding to q = 0.05) are statistically significant. Species to the right (left) are enriched among cases (controls). To derive q values, P values were corrected using the Benjamini-Hochberg method. Cases and controls are defined as patients with moderate/severe and mild/no chronic GVHD, respectively.

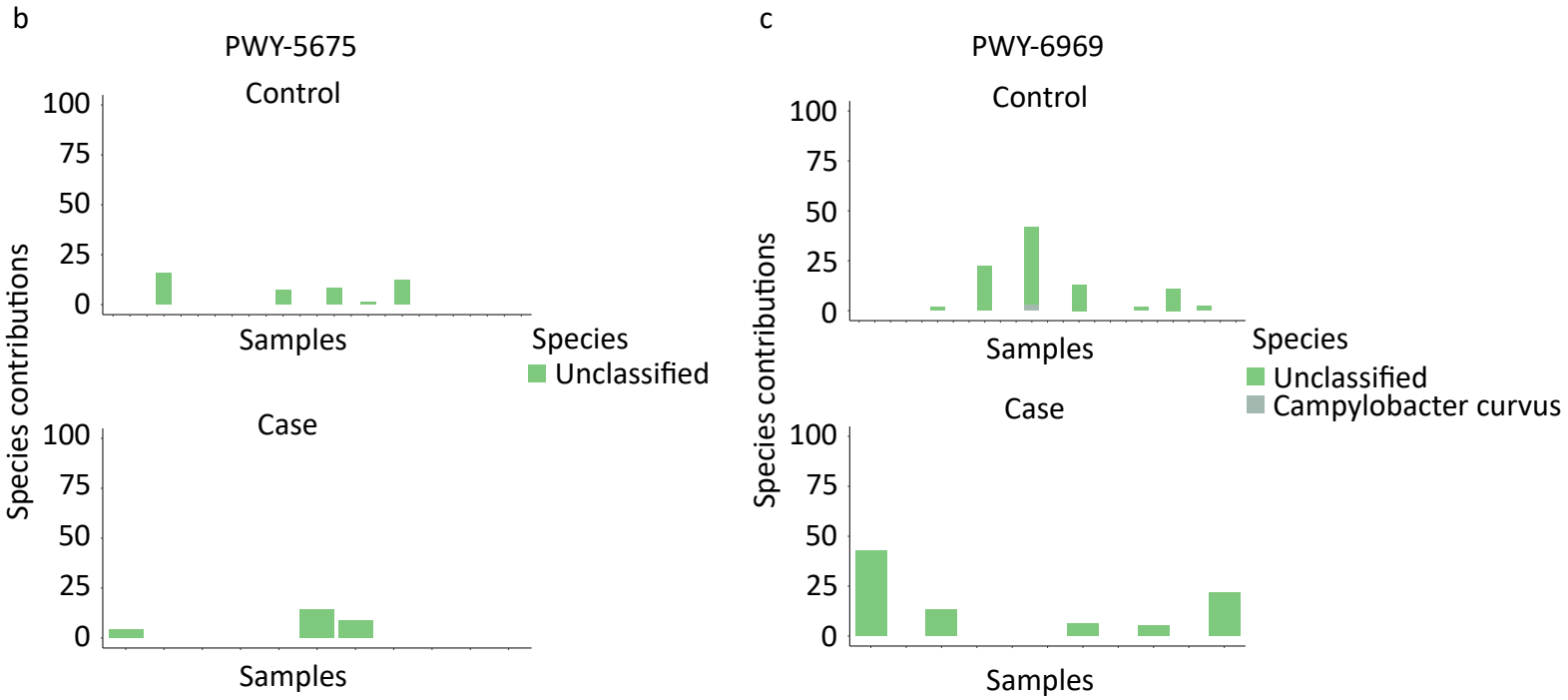

a
PWY0-1479
Control
Case
Species contributions
Samples
100
75
50
25
0
Species
Unclassified
Neisseria elongata
Klebsiella oralis
Neisseria sp oral taxon 014
Neisseria macacae
Haemophilus parainfluenzae
Eikenella corrodens
Neisseria sicca
Neisseria mucosa
Morococcus cerebosus
Haemophilus sp C1
Neisseria cinerea
b
PWY-5675
Control
Case
Species contributions
Samples
Species
Unclassified
c
PWY-6969
Control
Case
Species contributions
Samples
Species
Unclassified
Campylobacter curvus

Fig. S4: Species contributions to the other enriched pathways in Figure 4a
Of the 5 pathways enriched in day +28 plaque microbiota of cases (Fig. 4a), species contributions to two (PWY0-1477 and GOLPDLCAT-PWY) are shown in Fig. 4b-c. Species contributions to the other 3 pathways are shown here. Each panel shows data for 1 pathway. HUMAnN's pathway abundance tables (normalized to copies per million) include the number of copies of the entire pathway of interest assigned to each species. These numbers were used to quantify species contributions. Each column is a sample. The top panels show controls and the bottom panels show cases. Cases and controls are defined as patients with moderate/severe and mild/no chronic GVHD, respectively.

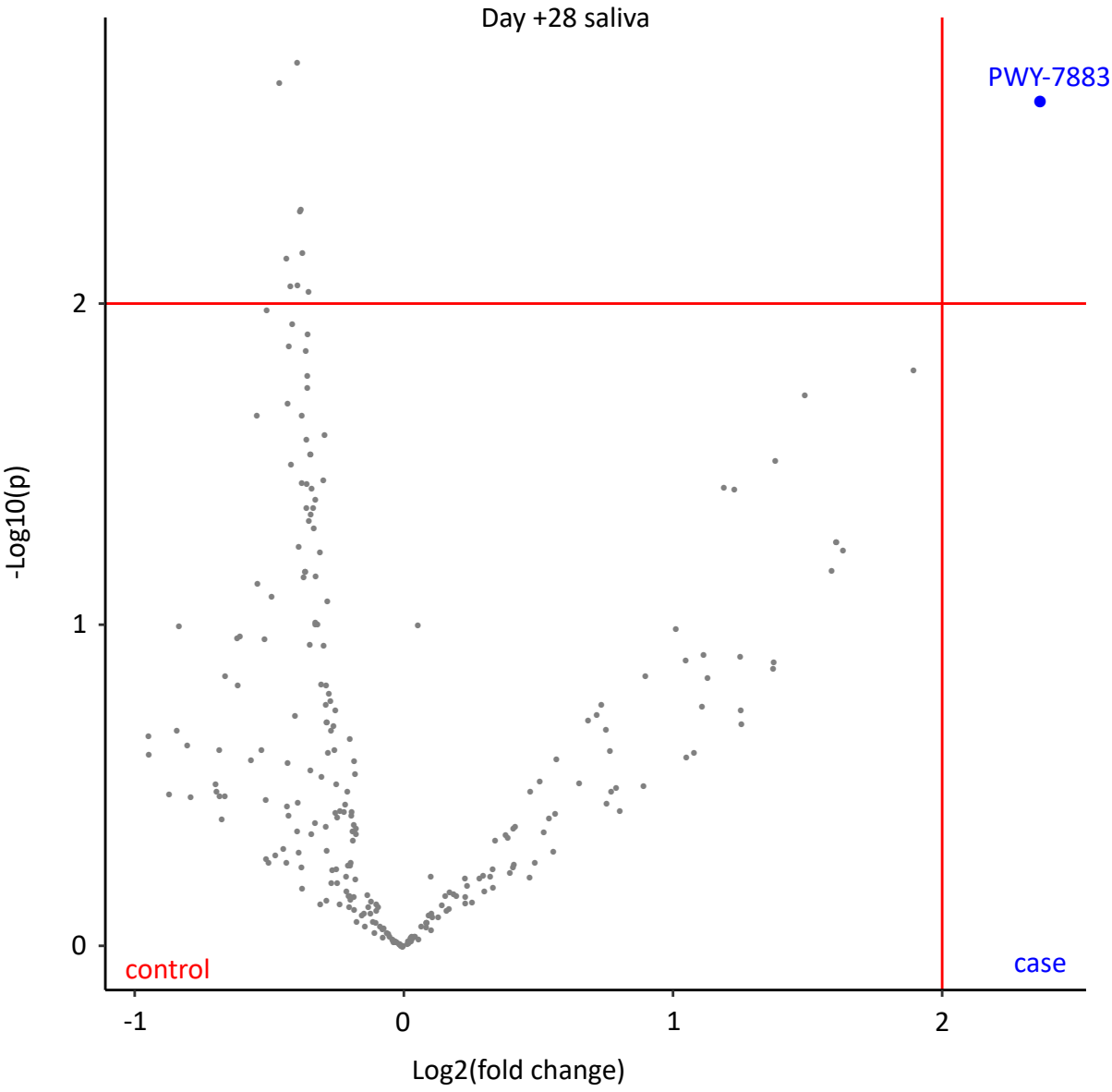

Day +28 saliva
PWY-7883
2
1
0
-Log10(p)
control
case
-1
0
1
2
Log2(fold change)

Fig. S5: Microbiota functional pathways in day +28 saliva and chronic GVHD
Volcano plot showing the results of differential abundance analysis for day +28 salivary microbiota using MaAsLin2. Each point shows a pathway. Points above the horizontal red line (-log10(0.01) = 2 and corresponding to P = 0.01) are statistically significant. Species to the right (left) are enriched among cases (controls). The vertical red line narrows down the enriched species to those with more substantial enrichment, defined as having more than two-fold abundance among cases. Cases and controls are defined as patients with moderate/severe and mild/no chronic GVHD, respectively.

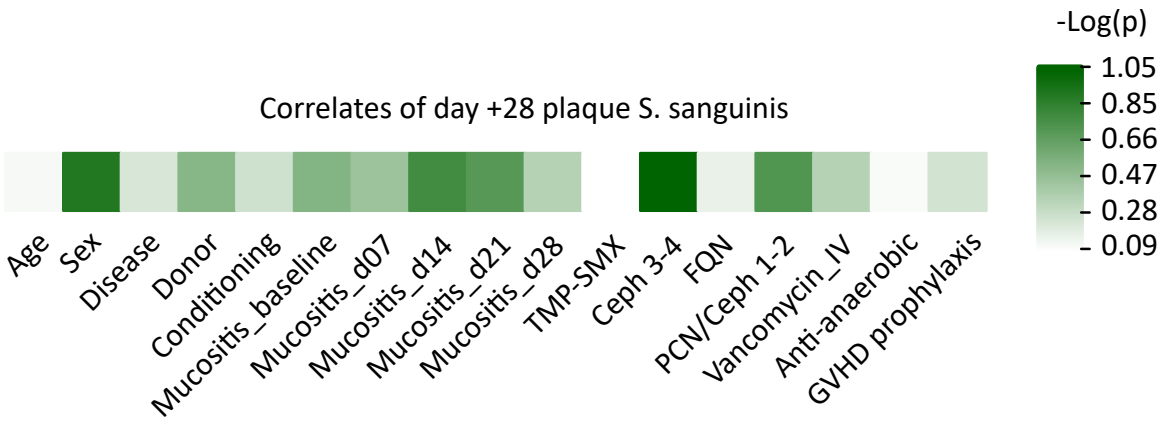

Correlates of day +28 plaque S. sanguinis
-Log(p)
1.05
0.85
0.66
0.47
0.28
0.09
Age
Sex
Disease
Donor
Conditioning
Mucositis_baseline
Mucositis_d07
Mucositis_d14
Mucositis_d21
Mucositis_d28
TMP-SMX
Ceph 3-4
FQN
PCN/Ceph 1-2
Vancomycin_IV
Anti-anaerobic
GVHD prophylaxis

Fig. S6: Baseline and early post-transplant clinical correlates of day +28 plaque *Streptococcus sanguinis*

Correlation plot showing log10-transformed P values along the color gradient for the association between baseline and early post-transplant clinical variables and *Streptococcus sanguinis* relative abundance in day +28 supragingival plaque. None of the variables reached statistical significance. A Wilcoxon's test was used for the categorical variables and Spearman's correlation test for continuous variables. Disease: acute leukemia vs. others. Donor type (multicollinear with graft source): HLA-matched donor vs. others. Conditioning: High-dose TBI-based vs. chemotherapy-based. Mucositis severity at each timepoint was derived from the quantitative OMAS score (total mucositis score). Antibiotic exposures are binary (0: not exposed; 1: exposed) and defined for the interval between day -7 and day +28. Ceph 1-2: first or second generation cephalosporins; Ceph 3+: third or higher generation cephalosporins; FQN: fluoroquinolones; PCN: penicillins; TMP-SMX: trimethoprim-sulfamethoxazole