THE INTERACTIONS BETWEEN ENVIRONMENTAL EXPOSURES, THE GUT MICROBIOTA AND THE IMMUNE SYSTEM IN ALLOGENEIC STEM CELL TRANSPLANTATION

By

Chi Linh Nguyen

A Dissertation

Presented to the Faculty of the Louis V. Gerstner, Jr.

Graduate School of Biomedical Sciences,

Memorial Sloan Kettering Cancer Center

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

New York, NY

January 2024

_____ _____

Marcel R.M. van den Brink, MD, PhD
Dissertation Mentor

Date

DEDICATION

To my parents, Vu Thi Kim Thoa and Nguyen Dang Duc, and my sisters, Nguyen Ha Linh and Nguyen Quynh Mai, for their unconditional love and unwavering support. They have instilled in me a sense of curiosity about the world that fuels my pursuit of science and have taught me the importance of resilience. None of these achievements would have been possible without these cheerleaders by my side.

ABSTRACT

The intestinal microbiota, which refers to a community of bacteria, fungi, viruses, and protozoa that lives in the gastrointestinal tracts, plays an important role in human health and disease. Particularly in the context of allogeneic hematopoietic cell transplantation (allo-HCT), preclinical and clinical studies have demonstrated that the intestinal microbiota is an important predictor of clinical outcomes. Allo-HCT is a potentially curative treatment for a variety of hematological malignancies. However, transplant-associated complications such as acute graft-versus-host disease (GVHD) can compromise patient long-term survival after allo-HCT. Perturbations to the intestinal microbiota during allo-HCT may result in a decline of biomass diversity and an increase in domination by bacteria in the genus *Enterococcus*, which have been associated with an increased risk of GVHD-related mortality. These data highlight the potential clinical importance of preserving the intestinal microbiome to maximize patient outcomes post-allo-HCT. In addition, mechanistic understandings of the interactions between gut microbes and the immune system, specifically focusing on microbial modulation of antigen-presenting cells and T cells, are crucial to develop microbiome-based therapeutics to enhance the efficacy and alleviate the toxicity of allo-HCT.

In this thesis, we investigated the association between pharmacological exposures and the intestinal microbiome. We developed a novel computational method, termed PARADIGM (PARameters Associated with DynamIcs of Gut Microbiota), and applied it to a large data set of longitudinal fecal microbiome profiles and detailed medication-administration records from allo-HCT patients. We observed that several non-antibiotic

medications, including laxatives, antiemetics, and opioids, are associated with decreased intestinal bacterial diversity and increased *Enterococcus* relative abundance. We integrated drug-microbiome associations to predict clinical outcomes in two validation cohorts on the basis of drug exposures alone, suggesting that this approach can generate biologically and clinically relevant insights on how pharmacological exposures can perturb or preserve microbiota composition.

Next, we turned to mouse models of allo-HCT to understand the immunological mechanism by which *Enterococcus*, a genus of Gram-positive facultative bacteria, aggravates GVHD. We observed that endogenous *Enterococcus* relative abundance is associated with increased mortality and MHC-II expression by the intestinal epithelial cells (IECs) of GVHD mice. Monocolonization of gnotobiotic mice with *Enterococcus* is sufficient to induce MHC-II expression by colonic IECs at non-transplanted steady state. These observations suggest an enhanced capacity of IECs to present antigens and activate donor T cells, thereby initiating the inflammatory GVHD cascade. We also explored a potential interventional approach to improve GVHD survival using a strain of *Blautia producta* that confers colonization resistance against *Enterococcus* through lantibiotic production. Altogether, these findings highlight the regulation of gut epithelium antigen presentation by the intestinal microbiota as a potential therapeutic target to prevent and ameliorate GVHD pathophysiology.

BIOGRAPHICAL SKETCH

Chi Nguyen was born and raised in Hanoi, Vietnam before coming to the United States to attend Colgate University in 2014, where she studied Molecular Biology. During her time at Colgate University, Chi worked under the mentorship of Dr. Krista Ingram, investigating the genetic regulation and behavioral impacts of human circadian rhythm disruption and morning-night preferences. Concurrently, she also spent a semester training in the lab of Dr. Sushil Rane at the National Institutes of Health as part of the Colgate off-campus study program. There, Chi pursued a study examining the role of the TGF-β signaling pathway in brown adipocyte differentiation from mesenchymal stem cells. Upon receiving her Bachelor of Arts in Molecular Biology from Colgate University in 2018, Chi began her PhD training at the Louis V. Gerstner Jr. Graduate School of Biomedical Sciences at Memorial Sloan Kettering Cancer Center, where she worked under the mentorship of Dr. Marcel van den Brink.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my mentor, Dr. Marcel van den Brink, for all the effort and resources that he has put into mentoring me through my PhD. I am grateful for the trust he has shown in me and the independence he has given me, which has challenged me and allowed me to grow as a scientist-in-training. I will bring the lessons I have learned from him into my future career and personal life. Without his support and influence, I would not have accomplished the work presented in this thesis and definitely would not have run two marathons in the last two years of my PhD.

I would like to thank my thesis committee, Drs. Tobias Hohl and Joao Xavier, for their advice and guidance in driving my projects forward over the years. Their expertise in computational biology, microbiology and immunology is incredibly valuable to me, and their critiques and suggestions have allowed me to gain significant scientific understanding.

I am grateful to have been a part of the van den Brink lab, where I was surrounded by many passionate and hardworking individuals. I would like to thank all the members, past and present, of the lab for experimental support and insightful discussions. I especially would like to thank Drs. Antonio Gomes, Marina Burgos da Silva, Melissa Docampo, Jonathan Peled and Kate Markey for providing important hands-on training and invaluable feedbacks that have propelled forward not only my projects, but also my personal development as a scientist. They have been incredibly generous with their time and patience with me, teaching me from scruffing a mouse to thinking critically about data interpretation. To Drs. Oriana Miltiadous, Hana Andrlova, Sophia Chen, Adhithi

Rajagopalan and Jenny Paredes, for providing me much-needed emotional support through rough times. I would like to thank Gabrielle Armijo, Nicole Lee, Emmanuel Dwomoh, Brianna Gipson, Kristen Victor, Anqi Dai and Nicholas Waters, who have provided important technical support in my projects. To Romina Ghale, my partner in lab and in running, I wish you the best of luck on your own journey as a PhD student.

Lastly, I would like to thank the Gerstner Sloan Kettering Graduate School administration, Drs. Ken Marians, Michael Overholtzer, Linda Burnley and Thomas Magaldi, and the graduate school staff, David McDonough, Stacy De La Cruz and Julie Masen for ensuring a training environment that emphasizes scientific rigor and translational impact, and for taking care of the PhD students both inside and outside of the lab. Thank you to all the past and present GSK students for creating an extremely welcoming and supportive community. To my classmates and fellow immunologists, Sherry Gao and Celeste Dang, for all the fun we have had together. To Michael Waarts, for your love, support, and encouragement. To my family, for always believing in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABX	Antibiotics
AHR	Aryl hydrocarbon receptor
ALLO-HCT	Allogenic hematopoetic cell transplantation
AML	Acute myeloid leukemia
AMP	Anti-microbial peptide
APC	Antigen-presenting cell
ASV	Amplicon sequence variant
ATG	Antithymocyte globulin
AUC	Area under the curve
BHI	Brain heart infusion
BM	Bone marrow
BMDC	Bone marrow-derived dendritic cell
CD	Cluster of differentiation
CFSE	Carboxyfluoroscein succinimidyl ester
CFU	Colony forming unit
CI	Confidence interval
DC	Dendritic cell
DMM	Dirichlet Multinomial Mixture
FACS	Fluorescence-activated cell sorting
FMT	Fecal microbiota transplant
GVHD	Graft-versus-host disease

HDAC	Histone deacetylase
ICU	Intensive care unit
IEC	Intestinal epithelial cell
IFNγ	Interferon gamma
IFNγR	Interferon gamma receptor
IG	Immunoglobulin
IL	Interleukin
ILC	Innate lymphoid cell
IP3	Inositol 1,4,5-triphosphate
IQR	Interquartile range
IV	Intravenous
KO	Knockout
LPS	Lipopolysaccharide
LTA	Lipoteichoic acid
MAIT	Mucosal-associated invariant T
MFI	Mean fluorescent intensity
MHC-II	Major histocompatibility complex class II
NLRP	Nucleotide-binding oligomerization domain, Leucine rich Repeat and Pyrin domain containing
OD	Optical density
PARADIGM	PARameters Associated with DynamIcs of Gut Microbiota
PBSC	Peripheral blood stem cell
PCA	Principal component analysis

PCR	Polymerase chain reaction
SCFA	Short-chain fatty acid
SD	Standard deviation
SEM	Standard error of the mean
SPF	Specific pathogen-free
TB	Tributyrin
TBI	Total body irradiation
TH	T helper
TMP-SMX	Sulfamethoxazole/Trimethoprim
TLR	Toll-like receptor
TREG	Regulatory T cell
VAMN	Vancomycin, ampicillin, metronidazole, neomycin
VRE	Vancomycin-resistant *Enterococcus*
WT	Wildtype

CHAPTER ONE

INTRODUCTION*

The digestive tract harbors the largest microbial community of the mammalian body. These microorganisms have a significant capacity to interact with and influence their host through direct contact with both parenchymal and hematopoietic cell populations and production of circulating metabolites. Advances in sequencing technologies in recent years have facilitated the rapid progress in clinical associative studies. In turn, these studies have served as valuable hypothesis-generating tools for mechanistic studies in preclinical animal models.

* Nguyen, C.L., Docampo, M.D., van den Brink, M.R., Markey, K.A. (2021). The role of the intestinal microbiota in allogeneic HCT: clinical associations and preclinical mechanisms. Curr Opin Genet Dev *66*, 25-35.

Allo-HCT is a curative-intent treatment for patients with hematological malignancies. However, it is a high-risk therapy, as overall post-transplant mortality remains in the order of 50%, most commonly due to disease progression or relapse, infection and GVHD (D' Souza A, 2019). Acute GVHD (aGVHD) is a systemic, potentially fatal condition that occurs following allo-HCT, in which donor-derived T cells recognize host antigens as foreign, resulting in cytotoxic tissue damage to predominantly the skin, the liver and the gastrointestinal tract (Ferrara et al., 1999). This introductory section outlines our current understanding of the role of the intestinal microbiota in allo-HCT and GVHD from both clinical and preclinical studies (**Figure 1.1**).

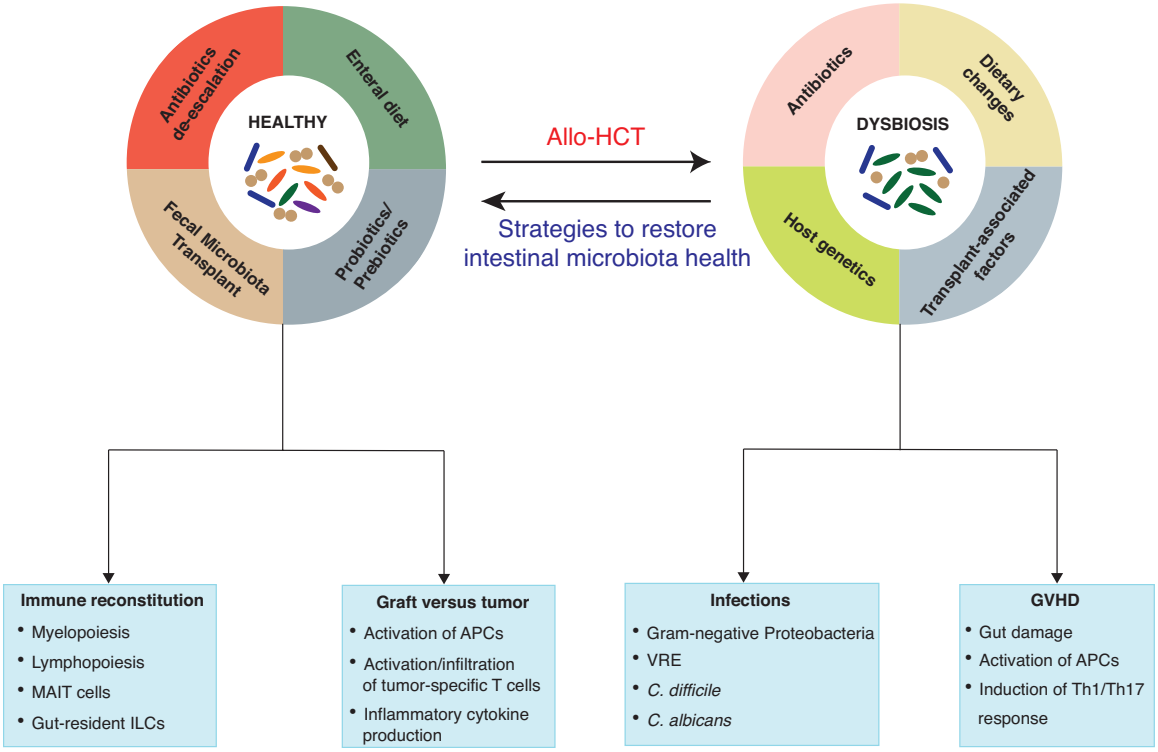

Figure 1.1. Intestinal microbiota dysbiosis and patient outcomes following allo-HCT. Allo-HCT patients are exposed to various environmental conditions including cytotoxic conditioning regimens, antibiotics, and dietary changes that might contribute to alterations in the intestinal microbiota. These injuries to the intestinal microbiota, in turn, are associated with transplant outcomes such as infections, immune reconstitution, GVHD and relapse through various different immunological mechanisms involving different

hematopoietic and non-hematopoietic cell populations. Strategies to restore the intestinal microbiota health include fecal microbiota transplantation, de-escalated antibiotic exposures and enteral diets that promote the growth of beneficial bacteria. VRE, vancomycin-resistant *Enterococcus*; MAIT, mucosal-associated invariant T cell; ILC, innate lymphoid cell; APC, antigen-presenting cell; Th, T helper.

Environmental exposures and intestinal microbiome dynamics

Microbial diversity, a summary measurement of the richness and evenness of unique bacterial taxa, is a useful measure of intestinal microbiota health. Several groups have observed a decline in intestinal microbial diversity manifesting early in the course of allo-HCT, and this injury to the intestinal microbiota is associated with lower overall survival after allo-HCT (Peled et al., 2020; Ying Taur, 2014). Exposures to broad-spectrum antibiotics are associated with a decrease in the intestinal microbiota biodiversity during allo-HCT, although other factors such as dietary changes, conditioning regimens and other medications are likely to influence the post-HCT intestinal microbiota dynamics (**Table 1.1**) (Stoma et al., 2020; Taur et al., 2014).

Table 1.1. Potential modulators of the intestinal microbiota during allo-HCT

Factors	Impacts	References
Antibiotics		
Broad-spectrum antibiotics (β-lactam, metronidazole, meropenem)	Broad-spectrum antibiotics target obligate anaerobes which are typically producers of short-chain fatty acids in the gut and allows the expansion of opportunistic pathogens	(Shono et al., 2016; Stoma *et al*., 2020)
Imipenem-cilastatin	Imipenem-cilastatin treatment in mice leads to expansion of the mucus-degrading bacteria *Akkermansia*	(Shono *et al*., 2016)

Fluoroquinolones	Prophylactic fluoroquinolones lead to a decreased risk of intestinal domination by Gram-negative bacteria	(Taur et al., 2012)
Transplant-associated factors		
Conditioning chemotherapy: carmustine, etoposide, aracytin and melphalan	After conditioning, the intestinal microbiota diversity declined, accompanied by a decrease in *Faecalibacterium* abundance and an increase in *Escherichia* abundance	(Montassier et al., 2014)
Conditioning intensity	Patients receiving non-myeloablative conditioning regimens have higher intestinal diversity at engraftment time compared to patients receiving myeloablative conditioning regimens	(Taur *et al*., 2014)
Total body radiation	Total body radiation in mice leads to reduction in the intestinal microbiota diversity and a shift in the overall microbial compositions compared to controls	(Zhao et al., 2019)
Immuno-suppressive medications (methotrexate, cyclosporine A)	In a high-throughput *in vitro* drug screen, some human-target drugs show antibacterial activities to some commensal bacterial strains often found in the human gut	(Maier et al., 2018)
Diet		
Total parenteral nutrition (TPN)	Use of long-term TPN (≥10 days) is associated with a loss of *Blautia*	(Jenq et al., 2015)
Western diet/Fiber-free diet	Mice fed on a prolonged fiber-free diet experience decreased intestinal microbiota diversity, a decrease in *Bifidobacteria* and an increase in mucus-degrading *Akkermansia*	(Desai et al., 2016; Schroeder et al., 2018)

Dietary fiber inulin supplementation	Inulin supplementation restores mucus functions in mice fed on a Western diet and restore microbiota loads	(Zou et al., 2018)
High-fiber (pectin, cellulose) diet	High-fiber diets containing cellulose and pectin lead to increased cecal and systemic levels of short-chain fatty acids such as acetate, propionate and butyrate	(Lewis et al., 2019)
Host genetics		
rs4988235 SNP conferring lactase expression and lactose intolerance in European populations	C/C SNP is associated with non-absorbers of lactose and increased risk of *Enterococcus* domination after allo-HCT	(Stein-Thoeringer et al., 2019)
Paneth cell α-defensin-5 gene SNPs	rs4415345G is associated with a higher abundance of a butyrogenic obligate anaerobe, *Odoribacter splanchnicus,* which is associated with a reduced risk of acute GHVD	(Rashidi et al., 2020)
NLRP6 deficiency	NLRP6 deficiency in the intestinal epithelial cells leads downregulation of IL-18 and alters the AMP production to a profile that favors intestinal microbiota dysbiosis	(Levy et al., 2015)

One of the first indications that the intestinal microbiota could play a role in transplant outcome came from the observation that transplanted germ-free mice had improved survival compared to specific-pathogen free (SPF) mice with intact microbiota (van Bekkum et al., 1974). Since then, clinical strategies aimed at "decontaminating" the gut have been an area of active investigation, although these clinical trials have produced conflicting results, suggesting the need for mechanistic studies using preclinical models to fully identify important taxonomic groups and their interactions with the host during allo-

HCT (Fredricks, 2019). Several clinical studies have reported associations between specific features of the intestinal microbiota and transplant outcomes, some favorable and some unfavorable (Golob et al., 2017; Holler et al., 2014; Jenq *et al*., 2015; Peled et al., 2017; Peled *et al*., 2020; Simms-Waldrip et al., 2017; Taur *et al*., 2012). Many currently active clinical trials employ therapeutic strategies that aim to eliminate microbiome features associated with adverse transplant outcomes, while preserving features associated with desirable outcomes (Taur et al., 2018).

Interactions between the intestinal microbiome and the mammalian host

The mechanisms by which the gut commensals influence local and systemic immunity include both direct interactions with immune cells, and indirect interactions via secreted microbial metabolites, such as short-chain fatty acids (SCFAs), secondary bile acids and other small molecules (Belkaid and Harrison, 2017) (**Figure 1.2**).

Short-chain fatty acids (SCFAs), such as butyrate, propionate and acetate, are produced as a result of the fermentation of dietary fiber by the intestinal microbiota. Butyrate (and to a lesser extent propionate) acts as a histone deacetylase inhibitor to promote the differentiation of colonic FoxP3^{+} regulatory T cells (Tregs) in mice (Arpaia et al., 2013; Furusawa et al., 2013). Propionate and acetate also induce anti-inflammatory responses and improve intestinal epithelial barrier integrity (Fukuda et al., 2011; Trompette et al., 2014). Secondary bile acids, which are produced as a result of microbial metabolism of primary bile acids excreted by the liver, can also induce differentiation of peripheral Tregs (Campbell et al., 2020; Hang et al., 2019). Microbe-derived indole derivatives from tryptophan metabolism confer a protective effect against epithelial barrier damage after

total body irradiation (TBI) and in a chemically induced colitis model (Shimada et al., 2013; Swimm et al., 2018). Moreover, the intestinal microbiota also produces many metabolites that act as ligands for the aryl hydrocarbon receptors (AhR), which are expressed by many cell types, and can play an important role in the induction and development of immune-mediated diseases (Lee et al., 2011; Manfredo Vieira et al., 2018; Zelante et al., 2013) (**Figure 1.2**).

The intestinal microbiota contributes to the maintenance and degradation of the mucus layer in the gastrointestinal tract, which forms a protective barrier between microbes and IECs. However, the outgrowth of bacteria with mucus-degrading activities can potentially lead to impaired intestinal barrier integrity in mice (Desai *et al*., 2016; Schroeder *et al*., 2018; Sovran et al., 2019). In addition, the intestinal microbiota also modulates antimicrobial peptide (AMP) production by IECs, such that disruption of the endogenous microbiota can affect AMP expression to favor the expansion of pathogenic bacteria in mice (Levy *et al*., 2015) (**Figure 1.2**). Alternatively, host genetics can also modulate the intestinal microbiota compositions: NLRP6 deficiency in mice leads to aberrant AMP production that favors intestinal microbiota dysbiosis (**Table 1.1**) (Levy *et al*., 2015). Gut commensals can also directly induce the expression of critical innate immune effectors and AMPs to maintain intestinal mucosal integrity and prevent colonization by opportunistic pathogens (Cash et al., 2006; Fan et al., 2015).

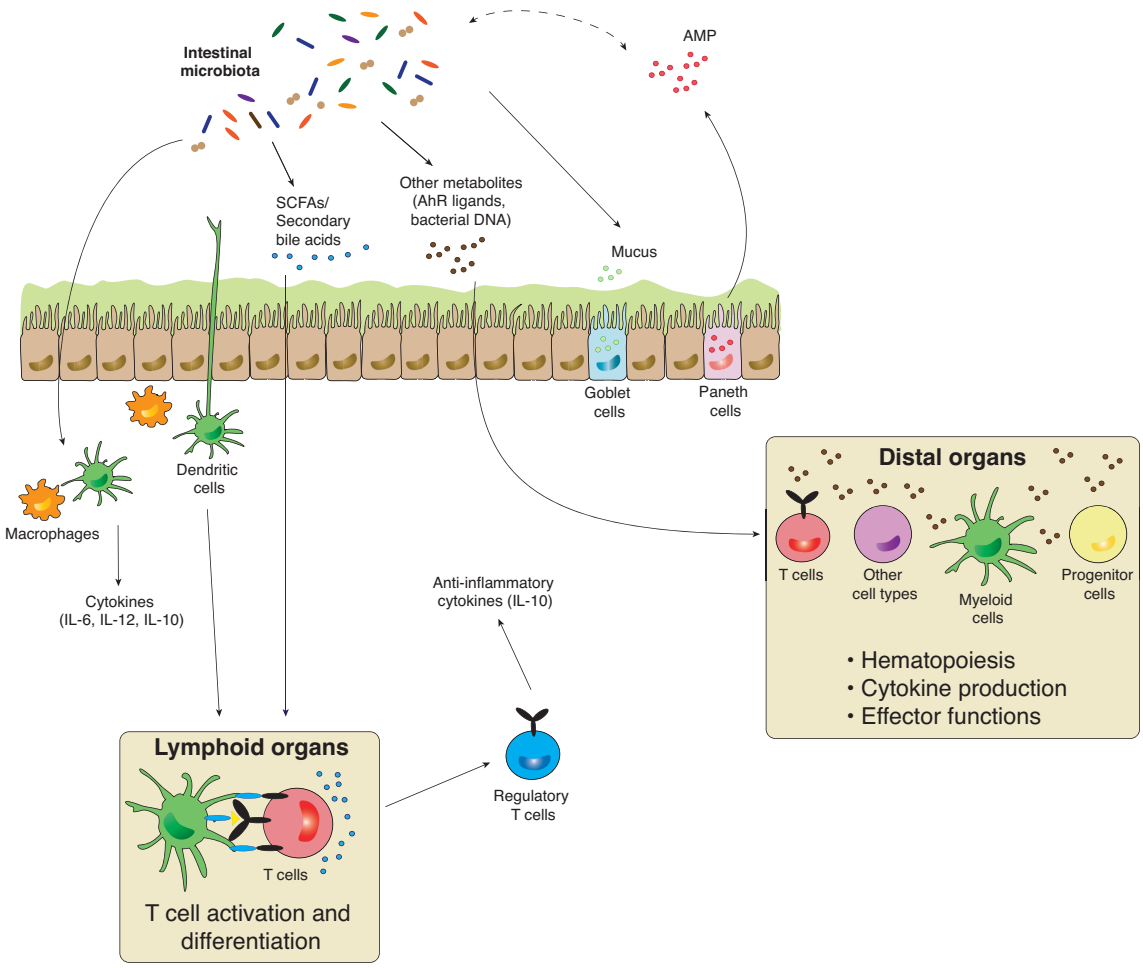

Figure 1.2. Interactions between the intestinal microbiota and the mammalian host. Within the gut lumen and intestinal epithelial environment, the intestinal microbiota promotes mucus production from goblet cells, as well as AMP production from Paneth cells to maintain gut homeostasis. AMPs, in turn, regulate the intestinal microbial compositions. Signals from the intestinal microbiota also activate APCs such as macrophages and dendritic cells, leading to downstream production of cytokines that promote context-specific inflammation or tolerance. Dendritic cells can also migrate to lymphoid organs to present antigens and activate T cells, and microbially derived metabolites such as SCFAs and secondary bile acids could mediate this process by influencing T cell activation and differentiation. Other microbially derived products such as AhR ligands, bacterial DNA and bacterial components can also translocate across the intestinal epithelial barrier into systemic circulation, reaching distal organs such as the liver, the lung and the bone marrow and modulate the generation, activation and differentiation of different cell subsets at these sites. SCFA, short-chain fatty acid; AhR, aryl-hydrocarbon receptor; AMP, anti-microbial peptide; IL, interleukin.

Intestinal microbiota and graft-versus-host disease (GVHD)

Several studies have reported that increased microbial diversity during the peri-neutrophil engraftment period is associated with a decreased risk of GVHD-related mortality (Golob *et al*., 2017; Holler *et al*., 2014; Peled *et al*., 2020; Simms-Waldrip *et al*., 2017; Taur *et al*., 2014). The intestinal microbial diversity during allo-HCT is most commonly determined through 16S ribosomal RNA sequencing of fecal samples that identify unique bacterial taxa and their relative abundance. Indoxyl sulfate, a product of the tryptophan metabolism by the intestinal microbiota, is a potential biomarker of intestinal biodiversity which can be measured in urine and has been studied in a pilot study of 31 allo-HCT patients (Holler *et al*., 2014). Decreased concentrations of these indole-derived metabolites were observed in allo-HCT patients with active GVHD compared to patients without GVHD (Michonneau et al., 2019).

Several specific members of the intestinal microbial community have been associated with either increased or decreased GVHD risks. SCFA-producing members of the families *Lachnospiraceae* and *Ruminococcaceae*, such as *Blautia, Clostridium* and *Lachnoclostridium*, are associated with protection against acute GVHD-related mortality and chronic GVHD development (Jenq *et al*., 2015; Markey et al., 2020). The absence of these beneficial bacteria, attributed to broad-spectrum antibiotic exposures, is correlated with an increased risk of GVHD development, specifically of the gastrointestinal tract (Han et al., 2018; Lee et al., 2019b). High relative abundance of *Lachnospiraceae* and *Ruminococcaceae* has also been correlated with an increased ratio of Treg/Th17 cells, potentially maintaining the balance of anti-inflammatory and pro-inflammatory responses during GVHD (Han *et al*., 2018; Han et al., 2019).

In addition, some bacterial taxa have been associated with increased GVHD severity. *Enterococcus* domination post-allo-HCT is associated with an increased risk of severe GVHD and GVHD-related mortality in allo-HCT patients and in mice (Holler *et al*., 2014; Stein-Thoeringer *et al*., 2019). Furthermore, *Enterococcus* requires lactose as a source of nutrition for maximal growth, and patients with the single-nucleotide polymorphism associated with lactose intolerance (thus retention of lactose in the gut lumen) also saw an increased risk of *Enterococcus* domination following allo-HCT (Stein-Thoeringer *et al*., 2019). *Akkermansia* expansion, which is associated with exposures to piperacillin/tazobactam and imipenem during transplant, is also correlated with an increased risk of GVHD-related mortality in both humans and mice (Desai *et al*., 2016; Shono *et al*., 2016). Moreover, this genus has been associated with an increased risk of chronic GVHD (Markey *et al*., 2020).

The intestinal microbiota potentially contributes to the development of GVHD through multiple mechanisms. Tissue damage caused by conditioning regimens allow translocation of pathogen-associated molecular patterns such as lipopolysaccharide (LPS) across the intestinal barrier, which can activate the innate immune system and is thought to be key for initiation of GVHD (Cooke et al., 2001; Zhao et al., 2013). Exposure to imipenem leads to decreased gut barrier integrity, resulting in more severe GVHD in mice, potentially due to the resultant increase in relative abundance of *Akkermansia* (Shono *et al*., 2016). Treatment with indole-3-carboxaldehyde, an indole derivative, limits gut epithelial damage and prevents the initial cascade of inflammatory responses triggered by TBI in a mouse model of GVHD (Swimm *et al*., 2018). Other indole derivatives can also activate aryl hydrocarbon receptor (AhR) on innate lymphoid cells (ILCs) and induce IL-

22 expression, which promotes epithelial regeneration after radiation and improves GVHD survival in mice (Lee *et al*., 2011; Lindemans et al., 2015; Zelante *et al*., 2013).

In addition, propionate, and to a lesser extent butyrate, has been shown to promote the activation of the NLPR3 inflammasome in non-hematopoietic cells to ameliorate GVHD (Fujiwara et al., 2018). Furthermore, oral dosing of mice with 17 butyrate-producing *Clostridium* strains after allo-HCT restores the intestinal butyrate levels and improves intestinal epithelial integrity, thereby reducing GVHD severity (Mathewson et al., 2016). Oral administration of butyrate in a mouse model of allo-HCT also mitigates GVHD through enhancing epithelial cell barrier functions, reducing apoptosis, and increasing expression of tight junction proteins (Mathewson *et al*., 2016).

Bacteria signal through toll-like receptors (TLRs) to recruit neutrophils into the ileum after allo-HCT, where they generate reactive oxygen species that further amplify radiation-induced damages in the gastrointestinal tract of recipient mice (Schwab et al., 2014). Neutrophil migration from the ileum to the mesenteric lymph nodes also depends on the intestinal microbiota, promoting alloantigen presentation to donor T cells (Hulsdunker et al., 2018). The intestinal microbiota also promotes MHC-II expression on IECs in a MyD88/TRIF-dependent manner, following the cues of inflammatory cytokines IFNγ and IL-12 (Koyama et al., 2019). In addition, professional donor-derived APCs such as CD103+ DCs migrate from the lamina propria to the mesenteric lymph nodes, where they promote the activation and proliferation of donor T cells through antigen presentation, and release inflammatory cytokines IL-6 and IL-12 in a MyD88/TRIF-dependent manner, suggesting the involvement of the intestinal microbiota in the activation and migration of APCs in GVHD pathophysiology (Koyama et al., 2015). Pre-transplant *Enterococcus*

colonization led to increased GVHD severity and mortality in a germ-free mouse model of GVHD. Feeding experimental mice a lactose-free diet prevented *Enterococcus* growth while preserving other members of the gut commensals, leading to improvement in GVHD-related survival in a mouse model of allo-HCT (Stein-Thoeringer *et al*., 2019). The specific mechanism by which *Enterococcus* modulates immune function during GVHD remains unclear and requires further investigation.

Not unsurprisingly, the influence of the complex intestinal bacterial community on GVHD outcome is complex, with some bacterial taxa appearing to be protective and others pathogenic. This highlights the need for novel therapeutic approaches to enhance the protective taxa while eliminating or decreasing the pathogenic taxa.

Introduction to the thesis

Clinical investigations have described relationships between the intestinal microbiota during allo-HCT and GVHD. Preclinical studies have complemented these clinical findings by elucidating some key mechanisms, specifically focusing on APC activation and antigen-presenting functions, effector functions of T cells and key mediating cytokines such as IL-6, IL-12 and IFNγ. In addition, studies into the potential modulators of the intestinal microbiota, including pharmacological exposures, dietary intakes, and host genetics, have provided additional insights regarding strategies to prevent and reverse microbiome perturbations. However, knowledge gaps persist that could potentially limit the clinical application of these findings.

In my thesis, I have combined computational analysis of human data with preclinical mechanistic studies using mouse models to further elucidate the interactions

between the environment, the intestinal microbiota, and the immune system during allo-HCT. Chapter two expands on our understanding of how non-antibiotic drug exposures are associated with the intestinal microbiome dynamics through the development of a novel computational method. Chapter three describes a mechanism by which a specific bacteria genus, *Enterococcus*, could aggravate GVHD through its influence on the intestinal epithelium antigen presentation. Altogether, these findings demonstrate a multi-pronged approach to preserve the health of the intestinal microbiota and its interaction with host immunity to maximize cancer treatment efficacy and safety.

CHAPTER TWO

HIGH-RESOLUTION ANALYSES OF ASSOCIATIONS BETWEEN MEDICATIONS, MICROBIOME AND MORTALITY IN CANCER PATIENTS*

Introduction

Gut microbiota perturbations have been associated with various diseases and frequently linked to environmental exposures including antibiotic use and nutritional deficiencies (Johnson et al., 2019; Korpela et al., 2016). Non-antibiotic drugs can also contribute to intestinal microbiota changes (Maier *et al*., 2018), but their effects in humans are less well-understood and challenging to study due to a lack of reliable drug exposure data (e.g. recall-based surveys of habitual use of chronic medications) (Korpela *et al*., 2016; Vieira-Silva et al., 2020), and the absence of densely-collected longitudinal fecal samples (Falony et al., 2016; Vich Vila et al., 2020). Moreover, several pioneering studies of medication exposures and microbiome composition focused on volunteers at relatively healthy steady states (Falony *et al*., 2016; Vieira-Silva *et al*., 2020).

* Nguyen, C.L., Markey, K.A., Miltiadous, O., Dai, A., Waters, N., Sadeghi, K., et al. (2023). High-resolution analyses of associations between medications, microbiome, and mortality in cancer patients. Cell *186*(12): 2705-2718 e17.

Patients undergoing allo-HCT exhibit major perturbations of fecal microbiome composition that have been associated with increased risk of mortality (Holler *et al*., 2014; Peled *et al*., 2020; Stein-Thoeringer *et al*., 2019). These patients are exposed to a variety of drugs upon prolonged hospitalizations, during which a wealth of data is routinely gathered as part of their electronic health records. As such, this patient population, from which we have assembled a large bank of fecal specimens, presents a unique opportunity to investigate intestinal microbial responses to drug exposures *in vivo*.

Previous studies in these patients have largely focused on the effect of antibiotics on the intestinal microbiota (Morjaria et al., 2019; Peled *et al*., 2020; Taur *et al*., 2012), yet many non-antibiotic drugs routinely administered during allo-HCT have demonstrated anti-bacterial activities *in vitro* (Maier *et al*., 2018). Furthermore, the microbiome perturbations in allo-HCT patients are observed prior to broad-spectrum antibiotic administration, suggesting a potential influence of transplant-associated medications (Shouval et al., 2022). Here, we inferred relationships between medications, microbiome composition, and clinical outcomes by developing, applying, and validating a new computational method named PARADIGM (PARameters Associated with DynamIcs of Gut Microbiota) to a large dataset of 16S rRNA and shotgun metagenomic sequencing profiles of serially collected fecal samples from allo-HCT patients.

Results

Clustering captures the intestinal microbiome temporal dynamics during allo-HCT

The dataset consists of 9,167 fecal samples from 1,201 allo-HCT patients at Memorial Sloan Kettering Cancer Center (MSKCC; **Figure 2.1** and **Table 2.1**). We

divided the MSKCC cohort into discovery (7,454 samples; 778 patients) and validation (1,713 samples; 423 patients) cohorts (**Figure 2.1** and **Table 2.2**). We computed the compositional differences among discovery-cohort samples using Bray-Curtis β-diversity dissimilarity indices at the genus level for 16S rRNA-sequenced samples, or at the species level for shotgun metagenomic sequenced samples and visualized the high-dimensional stool composition data via t-stochastic neighbor embedding (tSNE; **Figures 2.2A-B**).

Table 2.1. Patient and sample cohorts from MSKCC and Duke cohorts.

	MSKCC **Discovery**	***MSKCC*** **Validation**	***Duke*** **Validation**
Number of patients	778	423	142
Number of patients with 16S-sequenced samples	778	405	138
Number of 16S-sequenced samples	7,454	1,713	473
Number of 16S-sequenced samples per patient, median (first-third quartile)	7 (3 – 14)	3 (2 – 5)	2 (1 – 5)
Number of patients with at least one pair of daily 16S-sequenced samples between day -14 and 14 relative to HCT	454	-	-
Number of pairs of daily 16S-sequenced samples between day -14 and 14 relative to HCT	2,039	-	-
Number of pairs of daily 16S-sequenced samples per patient, median (first-third quartile)	3 (1 – 7)	-	-
Number of patients with shotgun metagenomic samples	340	142	-
Number of shotgun metagenomic samples	980	200	-
Number of patients with drug exposures data	775	423	142

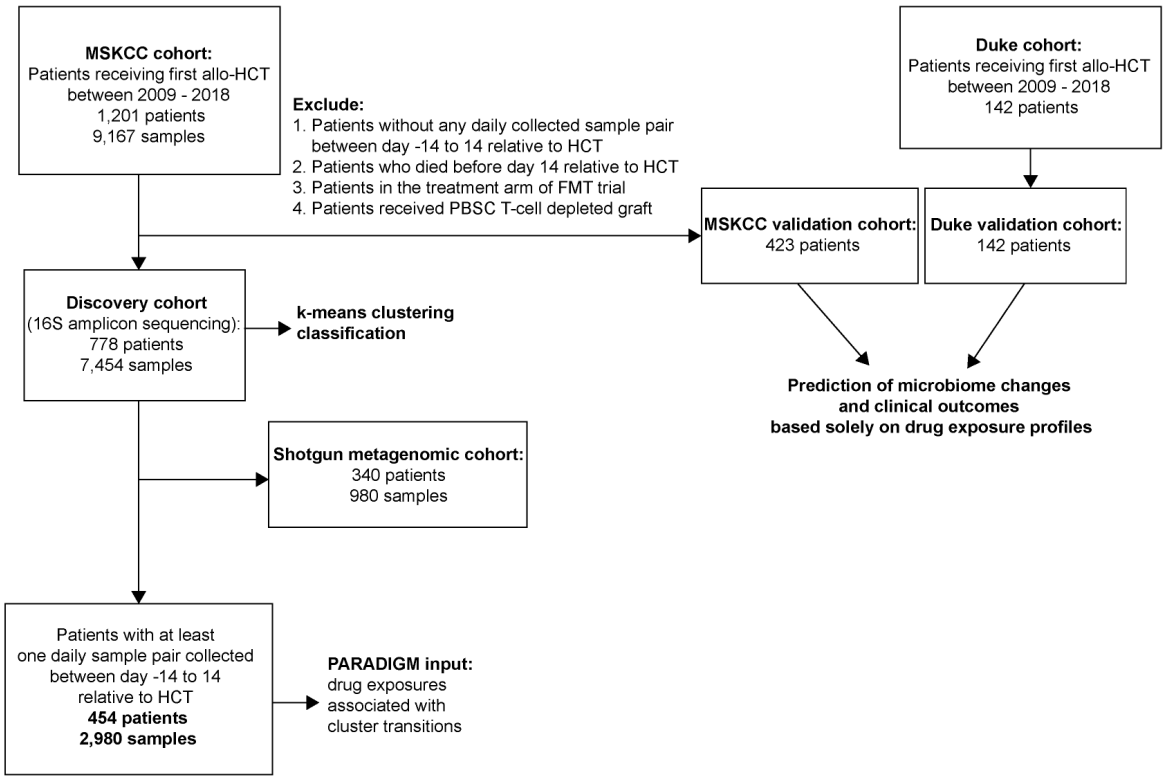

Figure 2.1. Patient selection criteria for the discovery and validation cohorts. The MSKCC discovery cohort was included in the clustering of sequencing data and PARADIGM algorithm training set. The validation cohorts were included in the analysis of clinical outcomes. FMT, fecal microbiota transplantation; PBSC, peripheral blood stem cell.

We observed patterns of microbiome injuries, including loss of alpha-diversity and enrichment of potentially pathogenic bacteria such as *Enterococcus* and Enterobacteriaceae (**Figures 2.2A-B**). As has been well-described in allo-HCT patients, these domination events can be profound, to the point of a single taxon comprising >90% of the relative abundance of a fecal sample (Peled *et al*., 2020; Taur *et al*., 2012), and are predictive of specific deleterious clinical outcomes such as bloodstream infections, GVHD and mortality (Stein-Thoeringer *et al*., 2019; Stoma et al., 2021; Taur *et al*., 2012). A subset of 980 samples with shotgun metagenomic profiling also showed similar patterns of microbiome injuries during allo-HCT (**Figure 2.2B**). Specifically, we observed a cluster

of samples whose most abundant organisms were various strict anaerobes such as *Ruminococcus gnavus* or *Erysipelatoclostridium ramosum*, as well as distinct clusters enriched for potentially pathogenic facultative species including *Enterococcus faecium*, *Klebsiella pneumoniae*, and *Escherichia coli*.

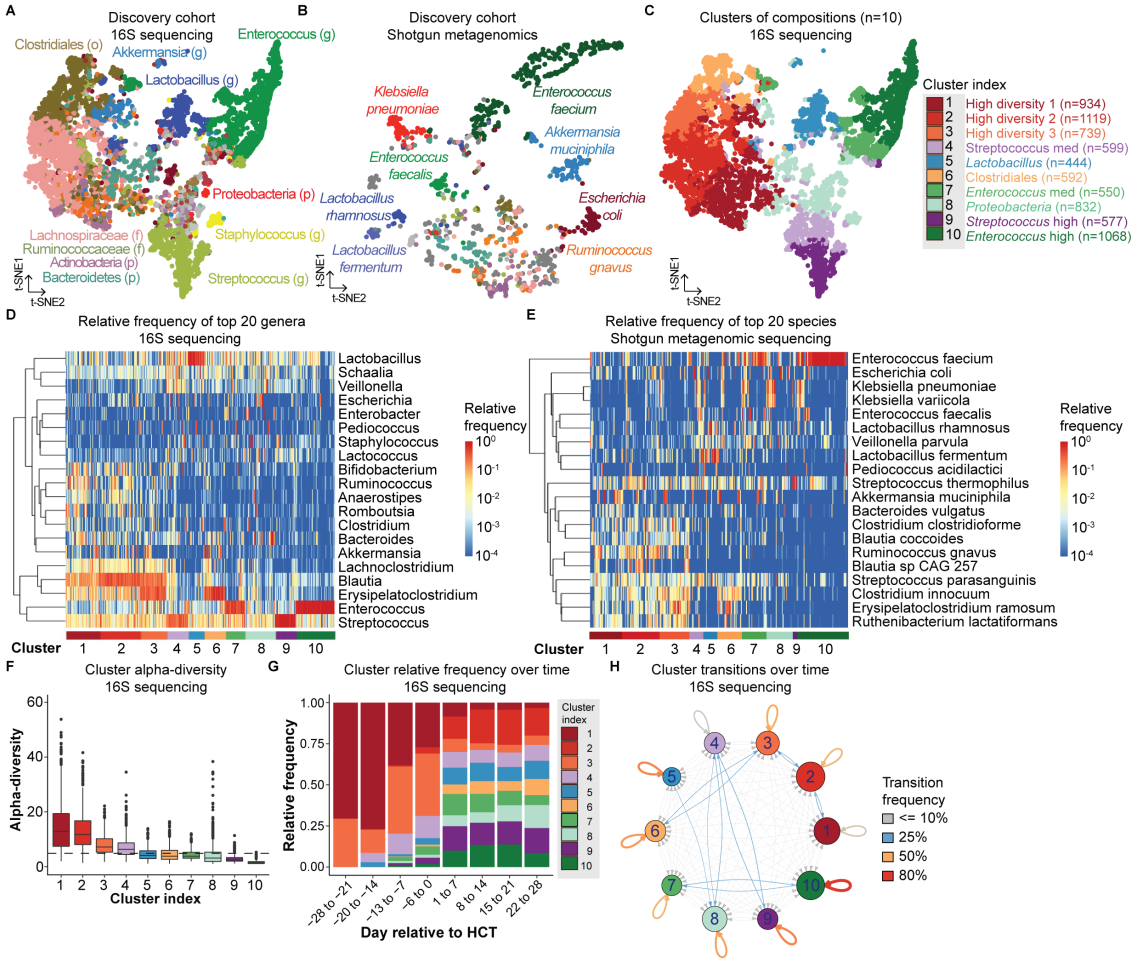

Figure 2.2. The intestinal microbiota of allo-HCT patients is highly dynamic. A, **B**, Compositional space of the intestinal microbiota in the MSKCC discovery cohort visualized by tSNE projection. Each point represents a sample, colored according to the taxon of highest relative abundance based on (**A**) 16S rRNA (7,454 samples; 778 patients) or (**B**) shotgun metagenomic sequencing profiles (980 samples; 340 patients) (p: phylum; f: family; o: order; g: genus). Samples were collected between day -30 and 2,205 relative to HCT. **c**, Ten clusters of intestinal microbiome compositions are assigned by k-means unsupervised clustering. **D**, **E**, Relative abundance of the top 20 most observed (**D**) genera in the 16S rRNA profiles and (**E**) species in the shotgun metagenomic profiles in the

MSKCC discovery cohort. Each column is one sample, each row is one genus or species. Rows are clustered by hierarchical clustering. **F**, Cluster alpha-diversity (reciprocal Simpson index). The horizontal dashed line represents the median alpha-diversity of the MSKCC discovery cohort. **G**, Cluster relative frequency over time relative to HCT. **H**, Network map depicting the transitions among the ten intestinal microbiota clusters over time (5,482 pairs of subsequent samples; 677 patients; collection between day -16 and 1,084 relative to HCT). The thickness of the line is proportional to transition frequency, while the node size is proportional to the number of samples per cluster.

The reproducibly observed microbiome perturbations in allo-HCT patients offer a unique opportunity to understand dynamics and evolution of relatively distinct perturbed microbiome compositions or states under environmental exposures, in contrast to the more fluid and non-discrete microbiota in healthy populations (Costea et al., 2018; DiGiulio et al., 2015; Human Microbiome Project, 2012; Munoz et al., 2021; Stewart et al., 2018). Given the mathematical challenge of reducing dimensionality complexity while preserving bacterial community structure, we performed unsupervised k-means clustering on the Bray-Curtis β-diversity matrix of samples in the discovery set and identified ten distinct microbiota clusters (**Figure 2.2C**). We also explored other clustering approaches including hierarchical clustering (**Figures 2.3A-C**) and Dirichlet Multinomial Mixtures (**Figures 2.3D-F**) (Holmes et al., 2012). Since k-means clustering partitioned samples more evenly (**Figure 2.3G**), we utilized k-means clusters for our subsequent analyses.

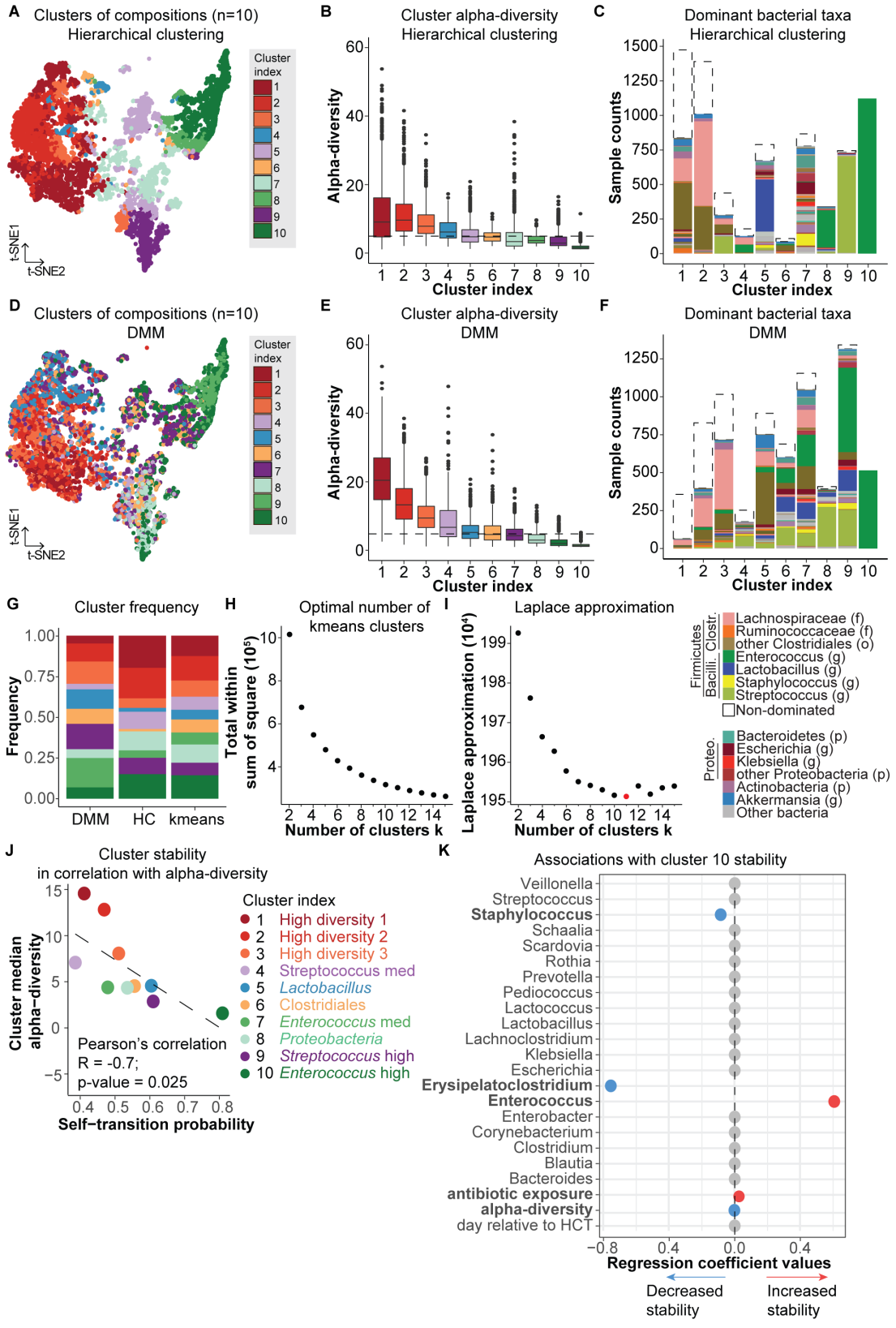
A
Clusters of compositions (n=10)
Hierarchical clustering
Cluster index
1
2
3
4
5
6
7
8
9
10
t-SNE1
t-SNE2
B
Cluster alpha-diversity
Hierarchical clustering
Alpha-diversity
Cluster index
C
Dominant bacterial taxa
Hierarchical clustering
Sample counts
Cluster index
D
Clusters of compositions (n=10)
DMM
E
Cluster alpha-diversity
DMM
F
Dominant bacterial taxa
DMM
G
Cluster frequency
Frequency
DMM
HC
kmeans
H
Optimal number of
kmeans clusters
Total within
sum of square (10⁵)
Number of clusters k
I
Laplace approximation
Laplace approximation (10⁴)
Number of clusters k
Firmicutes
Clostr.
Bacilli.
Lachnospiraceae (f)
Ruminococcaceae (f)
other Clostridiales (o)
Enterococcus (g)
Lactobacillus (g)
Staphylococcus (g)
Streptococcus (g)
Non-dominated
Proteo.
Bacteroidetes (p)
Escherichia (g)
Klebsiella (g)
other Proteobacteria (p)
Actinobacteria (p)
Akkermansia (g)
Other bacteria
J
Cluster stability
in correlation with alpha-diversity
Cluster median
alpha−diversity
Pearson's correlation
R = -0.7;
p-value = 0.025
Self−transition probability
Cluster index
1 High diversity 1
2 High diversity 2
3 High diversity 3
4 Streptococcus med
5 Lactobacillus
6 Clostridiales
7 Enterococcus med
8 Proteobacteria
9 Streptococcus high
10 Enterococcus high
K
Associations with cluster 10 stability
Veillonella
Streptococcus
Staphylococcus
Schaalia
Scardovia
Rothia
Prevotella
Pediococcus
Lactococcus
Lactobacillus
Lachnoclostridium
Klebsiella
Escherichia
Erysipelatoclostridium
Enterococcus
Enterobacter
Corynebacterium
Clostridium
Blautia
Bacteroides
antibiotic exposure
alpha-diversity
day relative to HCT
Regression coefficient values
Decreased
stability
Increased
stability

Figure 2.3. Clusters of intestinal microbiota during allo-HCT were identified by other unsupervised clustering methods. A, **D**, Compositional space of the intestinal microbiota visualized by tSNE projection in the MSKCC discovery cohort. Each dot represents a sample, colored according to clusters assigned by (**A)** hierarchical clustering or (**D**) Dirichlet Multinomial Mixture (DMM) model. **B**, **E**, Alpha-diversity per cluster identified by (**B)** hierarchical clustering or (**E**) DMM. The horizontal dashed line represents the median alpha-diversity of samples in the MSKCC discovery cohort. **C, F,** Compositional characteristics per cluster identified by (**C)** hierarchical clustering or (**F**) DMM. Samples in which the relative abundance of the most common taxon is ≥ 30% are color-coded by the most common taxon. Non-dominated samples are colored in white. (p: phylum; f: family; o: order; g: genus). **G**, Distribution of samples across 10 clusters identified by three unsupervised clustering methods. **H**, Optimal number of k-means clusters was estimated from the curve of within-cluster sum of square distances from each point to its cluster centroid. **I**, Optimal number of clusters identified by DMM was estimated by the smallest Laplace approximation metric. **J**, Correlation between cluster stability and alpha-diversity. Cluster stability was measured by self-transition probability. Alpha-diversity is defined as the cluster median reciprocal Simpson diversity. **K**, Co-exclusionary and inclusionary relationships associated with the stability of *Enterococcus*-high cluster 10.

Lachnospiraceae and *Clostridiales*, which constitute major commensal taxa present in the healthy human gut, were commonly observed in clusters 1-3, which were also characterized by high alpha-diversity (when compared to the median diversity of the discovery cohort) **(Figures 2.2D-F**) (Qin et al., 2010). Clusters 7-10 represented low-diversity "dysbiotic" states (**Figures 2.2D-F**). Intestinal domination by a single bacterial organism (≥ 30% relative abundance) is a hallmark of severe intestinal dysbiosis (Taur *et al*., 2012). *Lactobacillus*, Proteobacteria and *Streptococcus* were highly enriched in samples in clusters 5, 8 and 9. Clusters 7 and 10 consisted of *Enterococcus*-dominated samples, with cluster 10 specifically enriched for *E. faecium* (**Figure 2.2D-E)**. These compositional clusters also captured the temporal dynamics of the intestinal microbiota during allo-HCT: high-diversity clusters 1-3 were common pre-allo-HCT, while low-diversity states, particularly clusters 7-10, were more prevalent after allo-HCT (**Figure 2.2G**).

Table 2.2. Patient characteristics from MSKCC and Duke cohorts. IQR, interquartile range; sd, standard deviation; AML, acute myeloid leukemia; BM, bone marrow; PBSC, peripheral blood stem cell; PBSC T-cell depletion was performed by *ex vivo* CD34-selection of the graft.

	***MSKCC* Discovery**	***MSKCC* Validation**	***Duke* Validation**
Number of patients	778	423	142
Mean age at HCT, year (sd)	55 (13)	53 (13)	51 (13)
Year of HCT (%)			
2009 - 2015	374 (48)	331 (78)	33 (23)
2016 - 2019	404 (52)	92 (22)	109 (77)
Sex (%)			
Female	312 (40)	166 (39)	43 (30)
Male	466 (60)	257 (61)	99 (70)
Disease (%)			
AML	278 (36)	144 (34)	42 (30)
Others	500 (64)	279 (66)	100 (70)
Conditioning intensity (%)			
Nonmyeloablative	46 (6)	89 (21)	13 (9)
Reduced intensity	171 (22)	245 (58)	6 (4)
Ablative	561 (72)	89 (21)	123 (87)
Median follow-up, months	46	49	11
Graft type (%)			
Cord blood	62 (8)	131 (31)	19 (13)
BM unmodified	78 (10)	30 (7)	13 (9)
PBSC unmodified	176 (23)	262 (62)	110 (78)
PBSC T-cell depleted	462 (59)	-	-

This classification of samples into discrete microbiome states enabled us to model the complex changes in microbial communities as cluster transition probabilities. We observed the transition frequencies of consecutive samples collected at most 7 days apart between day -16 and day 1,084 relative to HCT. Patients remained in the same cluster over two consecutive samples in 2,987 (54.5%) pairs of samples (**Figure 2.2H**). Patients were less likely to remain in high-diversity clusters 1-3 among consecutive samples (mean frequency 46.4%; SD 5.0%), compared with dominated clusters such as 8-10 which are highly stable (mean frequency 65.1%; SD 14.1%). We observed a significant and negative

association between cluster alpha-diversity and self-transition probabilities, suggesting that high-diversity clusters are less stable compared to low-diversity clusters (**Figure 2.2J**). Transitions from one cluster to another were observed in 2,495 pairs of samples (45.5%).

We observed a particularly strong stability of the *Enterococcus*-high cluster 10, which is of interest due to the association of this genus with poor clinical outcomes following HCT (Holler *et al*., 2014; Stein-Thoeringer *et al*., 2019; Taur *et al*., 2012). To investigate potential drivers of *Enterococcus* domination stability, we developed a logistic regression model with lasso penalty analyzing cluster 10 stability as a function of parameters including antibiotic exposure, time of sample collection, alpha-diversity, and relative abundance of top 20 most abundant genera in cluster 10. We applied this model to a dataset of daily sample pairs collected between day -14 and 100 relative to HCT and found that higher relative abundances of *Staphylococcus* and *Erysipelatoclostridium* were associated with decreased cluster 10 stability (**Figure 2.3K)**. On the other hand, higher relative abundance of *Enterococcus* was associated with increased cluster 10 stability, indicating that *Enterococcus* domination leads to a positive feedback loop that support its own stability. As expected, antibiotic exposure was associated with increased cluster 10 stability. Here, using real-world data, we showed that both environmental factors such as medication exposures, and ecological relationships between bacteria, contribute to microbiota community stability, specifically regarding *Enterococcus* domination.

Non-antibiotic exposures during allo-HCT are associated with changes in the intestinal microbiome compositions

To investigate the associations between drug exposures and microbiome cluster transitions, we developed PARADIGM, a computational tool based on a logistic regression model integrated with first-order Markov-chain transitions. Markov-chain models have been utilized to investigate microbiome dynamics previously, but the associations between Markov transitions and environmental factors have not been extensively studied (Brooks et al., 2017; Cerdo et al., 2022; DiGiulio *et al*., 2015; Jin et al., 2017; Lee et al., 2019a). The model takes advantage of the high resolution daily 16S rRNA-sequenced fecal samples (2,039 sample pairs; 454 patients; **Table 2.3**) to infer associations between drug exposures and cluster transitions. For each cluster, we defined two transition types: self transitions (patients stay in the same cluster) and attractor transitions (patients move to a given cluster). The naming of "self transition" and "attractor transition" is motivated due to the intuitive connotation they convey about our cluster dynamics. Self transition describes the probability of a given cluster preserving its current state, and attractor transition describes the probability of a given cluster receiving transitions from any clusters other than itself, in a pair of daily collected samples (**Figure 2.4A**).

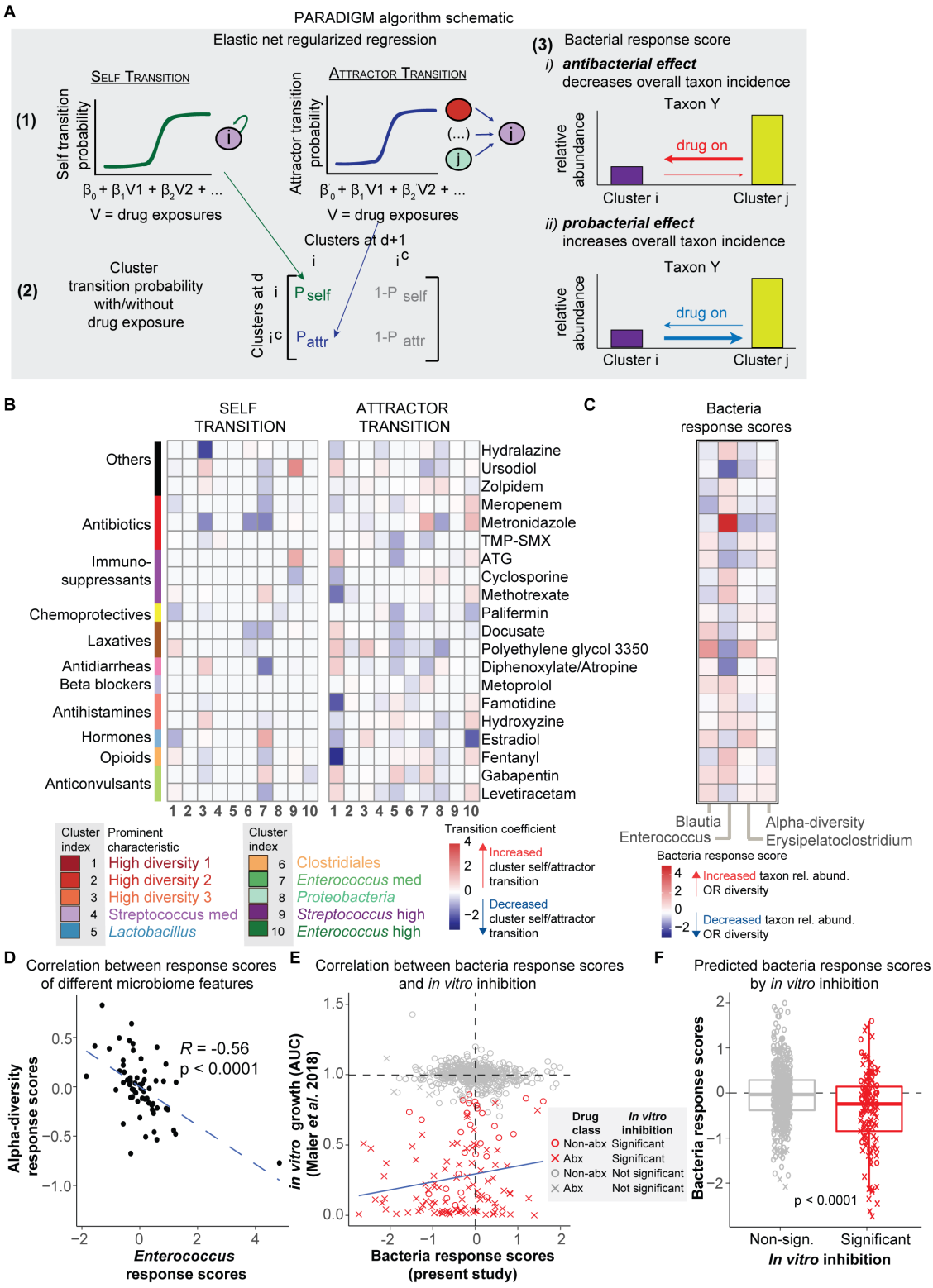

Figure 2.4. PARADIGM predicts changes in microbiome features such as genus relative abundance and alpha-diversity following drug exposures. A, Schematic representation of PARADIGM which takes advantage of daily sampling 16S rRNA-

sequenced samples and cluster transitions to infer how drug exposures are associated with microbial dynamics. Bacteria response scores translate drug-cluster associations into drug-genus associations. **B**, Pearson's correlation between *Enterococcus* response scores and alpha-diversity response scores. Each point represents an individual drug. **C**, Pearson's correlation between bacteria response scores and measurements of *in vitro* inhibition (Maier *et al*., 2018). Each point represents the association between a unique drug-species pair. **D**, Predicted bacteria response scores by *in vitro* inhibition. Two-sided Wilcoxon's rank-sum test.

Table 2.3. Patient characteristics of the MSKCC sub-cohort included in PARADIGM training set of daily collected fecal samples. IQR, interquartile range; sd, standard deviation; AML, acute myeloid leukemia; BM, bone marrow; PBSC, peripheral blood stem cell; PBSC T-cell depletion was performed by *ex vivo* CD34-selection of the graft.

	MSKCC
Number of patients	454
Number of daily sample pairs	2,039
Number of pairs of samples per patient, median (first-third quartile)	3 (1 – 7)
Mean age at HCT, year (sd)	56 (13)
Sex (%)	
Female	183 (40)
Male	271 (60)
Disease (%)	
AML	159 (35)
Others	295 (65)
Conditioning intensity (%)	
Nonmyeloablative	40 (9)
Reduced intensity	162 (36)
Ablative	252 (55)
Graft type (%)	
Cord blood	56 (12)
BM unmodified	76 (17)
PBSC unmodified	166 (37)
PBSC T-cell depleted	156 (34)

We focused on 62 drugs to which patients in the discovery set were commonly exposed. To determine the contribution of each drug to the likelihood of self and attractor transitions, we utilized elastic-net logistic regression, where the resulting coefficients for each drug indicate both the direction and magnitude of the association between the drug and daily cluster transitions. Microbiome injury patterns in allo-HCT are strongly linked

with time relative to allo-HCT (Peled *et al*., 2020; Taur *et al*., 2018; Taur *et al*., 2012). Therefore, we included time as a co-variable to address the temporal patterns of drug exposures and attenuate time-dependent variability in microbiome cluster dynamics (**Figure 2.5**). To account for unequal patient contribution of data points to the training set (**Table 2.3**), we pre-specified the 10-fold cross-validation partitions in our training cohort such that samples of the same patient are always in the same partition.

We identified several associations between drug exposures and cluster self and attractor transitions (**Figure 2.6**). As expected, several antibiotics used to empirically treat neutropenic fever were associated with profound changes in the intestinal microbiota. Exposures to meropenem (2.5-fold increase), and metronidazole (3.4-fold increase) were associated with increased transitions to the *Enterococcus*-high cluster 10, consistent with previous reports (Soares et al., 2017; Taur *et al*., 2012). Several non-antibiotic medications were also associated with specific cluster dynamics. Exposure to aprepitant, a tachykinin receptor antagonist used for chemotherapy-induced nausea, was associated with a 2.8-fold increase in transition frequency to the *Enterococcus*-high cluster 10 (**Figure 2.6**). Similarly, exposure to the opioid analgesic fentanyl was associated with a 1.9-fold increase in transition frequency to the *Enterococcus*-high cluster 10. Other medications such as labetalol and insulin, which are not known to target intestinal bacteria, were associated with decreased stability of the *Enterococcus*-high cluster 10 (**Figure 2.6**). Altogether, PARADIGM identified the association between several non-antibiotic drugs and changes in the intestinal microbiome.

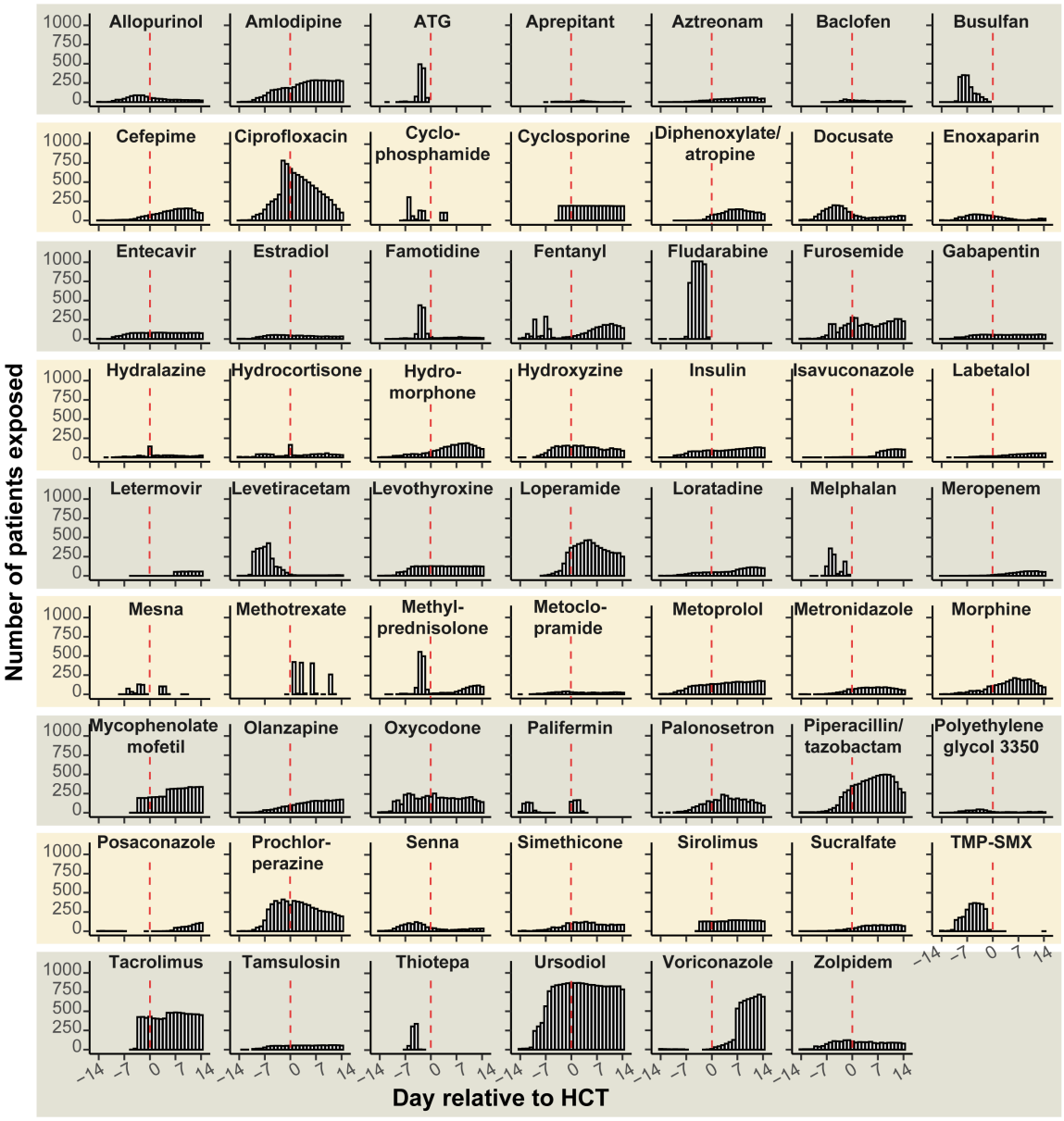

Figure 2.5. Time courses of drug exposures between day -14 to 14 relative to allo-HCT. Red dashed lines indicate day 0 which is the day of stem cell infusion.

Previous studies have identified specific bacteria that have either beneficial or deleterious associations with clinical outcomes following allo-HCT(Golob *et al*., 2017; Jenq *et al*., 2015; Simms-Waldrip *et al*., 2017). As such, we translated drug-cluster associations into drug-taxon associations by calculating bacteria response scores to identify

associations between drug exposure and changes in specific taxonomic groups of clinical interests (**Figure 2.7**). In our model, bacteria response scores estimate the association between a drug exposure and a microbiome feature, namely relative abundance of taxa or alpha diversity, where positive scores indicate association with higher relative abundances or diversity values (**Figure 2.4A**). We focused on four microbiome features previously associated with allo-HCT patient outcomes, namely relative abundance of *Enterococcus*, *Blautia*, *Erysipelatoclostridium*, and alpha-diversity (Golob *et al*., 2017; Holler *et al*., 2014; Jenq *et al*., 2015; Peled *et al*., 2020; Stein-Thoeringer *et al*., 2019).

Most antibiotics used in this cohort as empiric or pathogen-directed treatments (metronidazole, meropenem, aztreonam, and cefepime) were associated with increased relative abundance of *Enterococcus* as well as decreased alpha-diversity, consistent with previous studies (**Figures 2.7**) (Lee *et al*., 2019b; Shono *et al*., 2016; Taur *et al*., 2012). Our observation that cefepime exposure was associated with *Enterococcus* expansion is consistent with our previous report (Shono *et al*., 2016), and may be partly explained by the poor activity of cefepime and other cephalosporins against enterococci. Piperacillin-tazobactam exposures were associated with decreased *Enterococcus* relative abundance, as well as decreased relative abundance of intestinal commensals such as *Blautia* and *Erysipelatoclostridium* to a greater extent compared to other empiric antibiotics (**Figure 2.7**). We also observed that drugs most strongly associated with *Enterococcus* expansion were the non-antibiotic drugs including opioids such as fentanyl and hydromorphone, hormones such as levothyroxine, and anticonvulsants such as gabapentin (**Figure 2.7**).

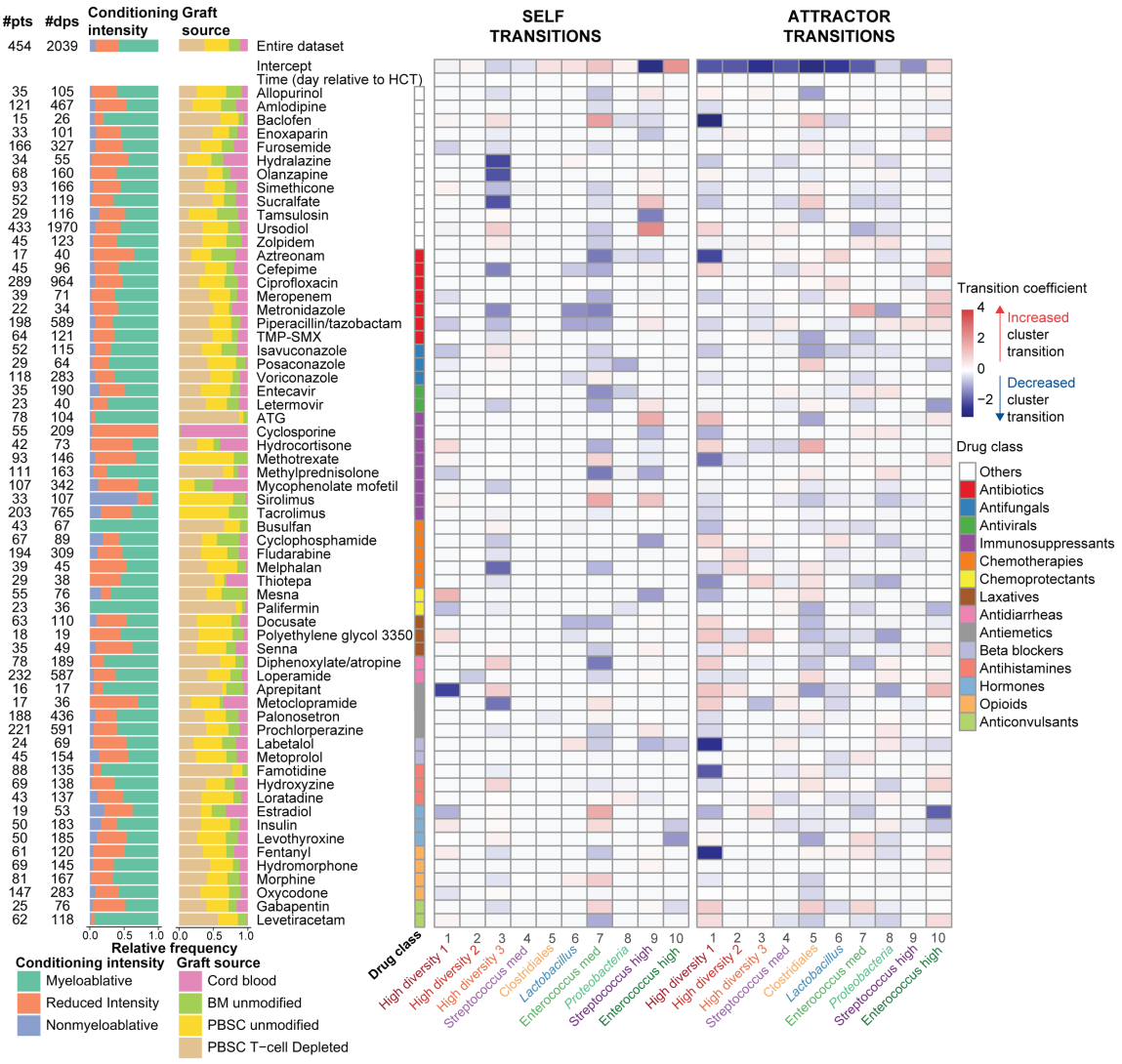

Figure 2.6. Associations between drug exposures and cluster self and attractor transitions. Self coefficients indicate whether drug exposure increases (positive coefficients, red shades) or decreases (negative coefficients, blue shades) the log-odds of cluster stability. Attractor coefficients indicate whether drug exposure increases (positive coefficients, red shades) or decreases (negative coefficients, blue shades) the log-odds of transition to a given cluster. #pts indicates the number of patients exposed to each drug, #dps indicates the number of sample pairs collected on the day of each drug exposure.

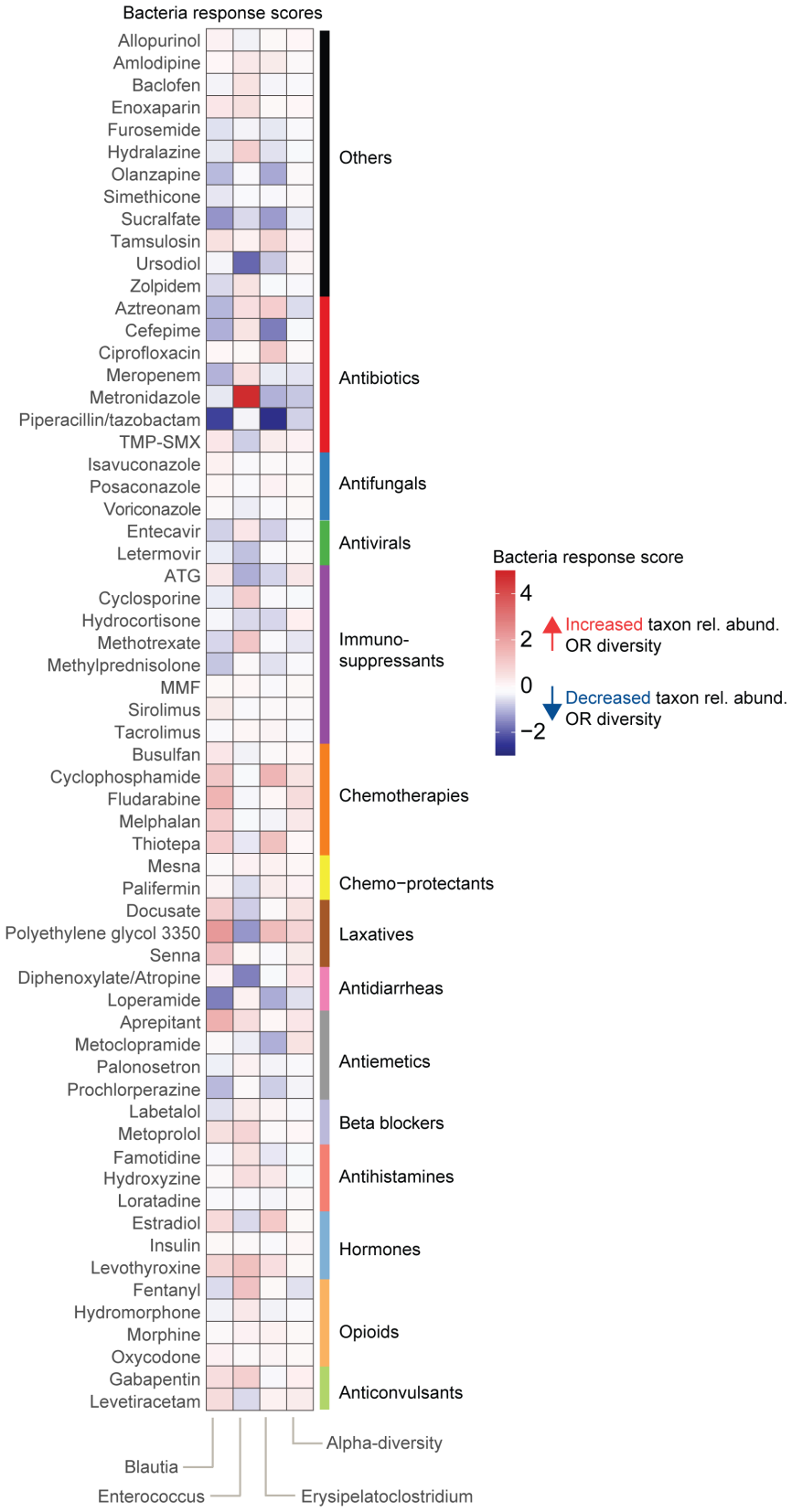
Bacteria response scores
Allopurinol
Amlodipine
Baclofen
Enoxaparin
Furosemide
Hydralazine
Olanzapine
Simethicone
Sucralfate
Tamsulosin
Ursodiol
Zolpidem
Aztreonam
Cefepime
Ciprofloxacin
Meropenem
Metronidazole
Piperacillin/tazobactam
TMP-SMX
Isavuconazole
Posaconazole
Voriconazole
Entecavir
Letermovir
ATG
Cyclosporine
Hydrocortisone
Methotrexate
Methylprednisolone
MMF
Sirolimus
Tacrolimus
Busulfan
Cyclophosphamide
Fludarabine
Melphalan
Thiotepa
Mesna
Palifermin
Docusate
Polyethylene glycol 3350
Senna
Diphenoxylate/Atropine
Loperamide
Aprepitant
Metoclopramide
Palonosetron
Prochlorperazine
Labetalol
Metoprolol
Famotidine
Hydroxyzine
Loratadine
Estradiol
Insulin
Levothyroxine
Fentanyl
Hydromorphone
Morphine
Oxycodone
Gabapentin
Levetiracetam
Others
Antibiotics
Antifungals
Antivirals
Immuno-
suppressants
Chemotherapies
Chemo−protectants
Laxatives
Antidiarrheas
Antiemetics
Beta blockers
Antihistamines
Hormones
Opioids
Anticonvulsants
Alpha-diversity
Blautia
Enterococcus
Erysipelatoclostridium
Bacteria response score
4
2
0
−2
Increased taxon rel. abund.
OR diversity
Decreased taxon rel. abund.
OR diversity

Figure 2.7. Bacteria response scores for four microbiome features of interest. Bacteria response scores predict the association between a given drug exposure and changes in genus relative abundance or alpha-diversity. Positive response scores (red shades) indicate that drug exposures are associated with increased genus relative abundance or alpha-diversity. Negative response scores (blue shades) indicate that drug exposures are associated with decreased genus relative abundance or alpha-diversity.

Opioid exposures have been previously associated with decreased relative abundance of *Blautia* in ICU patients who did not receive antibiotics (Pettigrew et al., 2019), an observation we also reported here for fentanyl and hydromorphone (**Figure 2.6**). In contrast, laxatives such as docusate and polyethylene glycol were strongly associated with decreased *Enterococcus* relative abundance (**Figure 2.7**). Previous experimental studies have demonstrated that polyethylene glycol induces global changes in bacterial compositions in mice, either through modulation of intestinal osmolality or through direct anti-bacterial inhibition (Nalawade et al., 2015; Tropini et al., 2018). Overall, drug exposures associated with alpha-diversity preservation were correlated with decreased *Enterococcus* expansion, and vice versa **(Figure 2.4B**).

Validation of in silico findings from real-world patient dataset against an independent in vitro dataset

We tested the predictive power of PARADIGM by comparing the *in silico* results described here using 'real-world data' from cancer patients with an independent published *in vitro* dataset (Maier *et al*., 2018) (**Figures 2.4C-D**). The *in vitro* screen and the present study share in common 19 bacterial species and 34 drugs. We observed an enrichment of drug-species pairs that showed *in vitro* inhibition (for both antibiotics and non-antibiotics) in the lower left quadrant of the plot, which corresponds to negative response scores *in*

silico (Fisher's Exact Test: odd ratio = 0.60, p-value = 0.01). Furthermore, drug-species pairs that showed *in vitro* inhibition had significantly lower response scores in the patient dataset when compared to those that did not show inhibition *in vitro* (**Figure 2.4D**). Altogether, these results suggest that PARADIGM can accurately predict *in vitro* anti-bacterial activity of both antibiotics and non-antibiotics and distinguish direct interactions of drugs with bacteria species from the potential confounding influence of the clinical symptoms prompting these drug exposures.

Antibiotic exposure is a strong predictor of subspecies dynamics

Several experimental studies have demonstrated that the bactericidal spectra of several drugs are species- and strain-specific even within the same genus (Maier et al., 2021; Maier *et al*., 2018). We therefore explored the associations between drug exposures and changes in relative abundances of species within genera of clinical importance in allo-HCT, using a subset of 980 specimens from 340 patients in the MSKCC discovery cohort with available shotgun metagenomic sequencing profiles (**Figure 2.8**). By applying a linear mixed-effects regression model, we identified associations between drug exposures and changes in bacterial species relative abundance. Again, we detected that several drug exposures spanning different drug classes (antibiotics, laxatives, anti-diarrhea, and opioids) were associated with changes in relative abundance of *Blautia coccoides, Blautia producta, Enterococcus faecalis, Enterococcus faecium* and *Erysipelatoclostridium ramosum,* although these associations were not statistically significant.

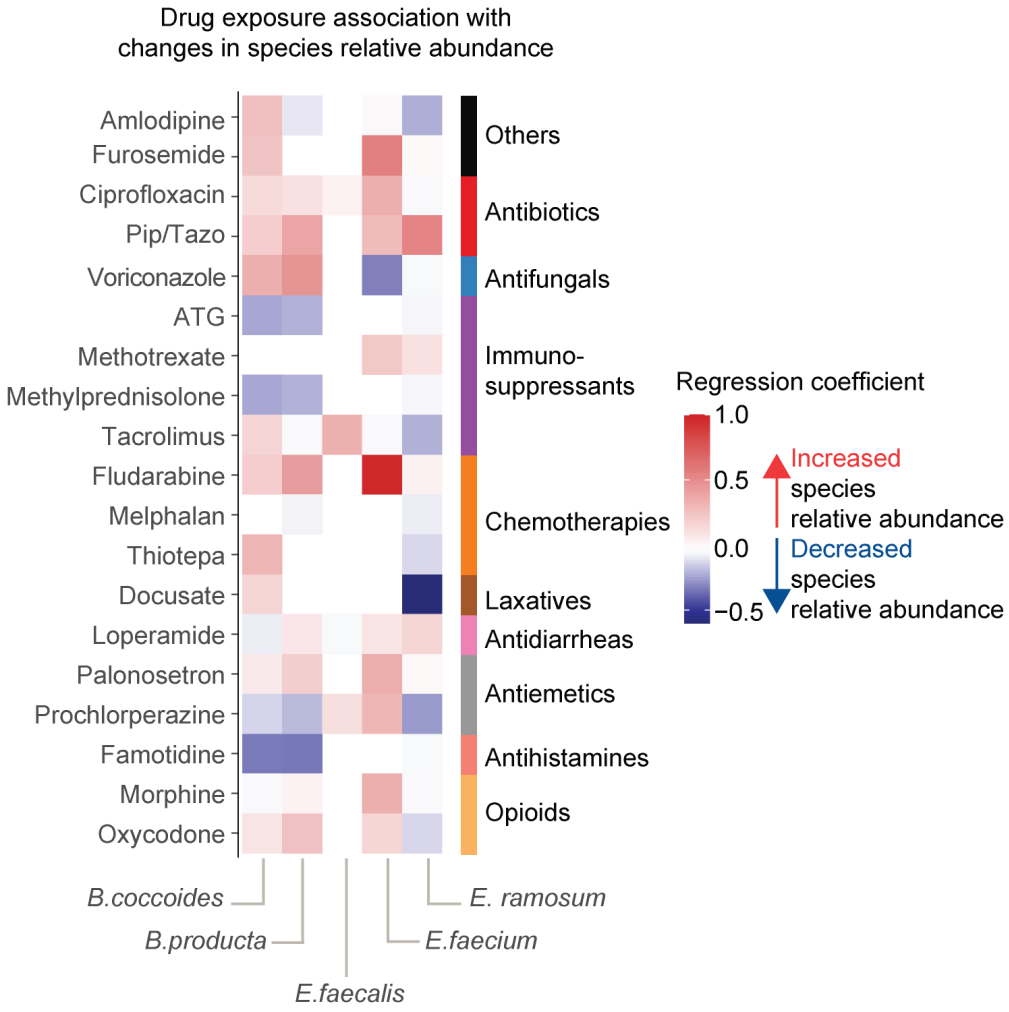

Figure 2.8. Associations between drug exposures and changes in species relative abundance from samples with shotgun metagenomic sequencing. Log changes in species relative abundance between subsequently collected samples were analyzed as a function of individual drug exposure in a linear mixed-effects regression model, with time of sample collection binned into weekly intervals as a random effect variable. Positive coefficient values (red shades) indicate that drug exposures are associated with increased species relative abundance. Negative coefficient values (blue shades) indicate that drug exposures are associated with decreased species relative abundance. A white box indicates that there is insufficient datapoint to fit the regression model for a given drug-species pair.

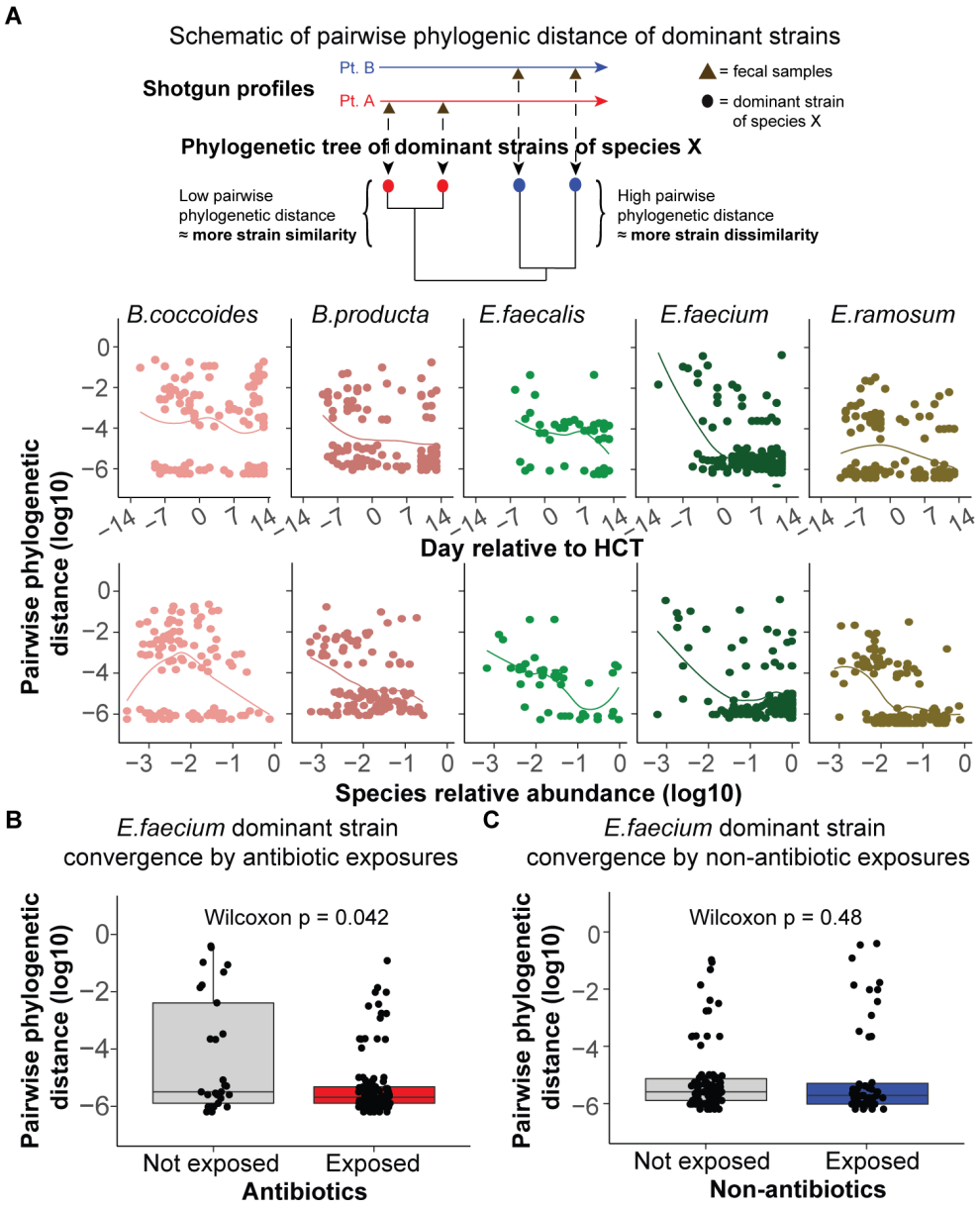

Figure 2.9. Antibiotics are strong predictors of strain genetic convergence during allo-HCT. A, Strain convergence over time relative to HCT (middle row), or by species relative abundance (bottom row). Each point represents the tree-based phylogenetic distance between the dominant strains of a given species in a pair of subsequently collected samples. Higher phylogenetic distance suggests genetic dissimilarity, while lower phylogenetic distance suggests strain genetic similarity. **B**, **C**, Antibiotic exposure (**B**), but not non-antibiotic exposure (**C**) is associated with increased *E. faecium* dominant strain convergence. Each point represents the phylogenetic distance between *E. faecium* dominant strains in a pair of subsequently collected samples, stratified by drug exposures during the time gap of sample pair collection. Two-sided Wilcoxon's rank-sum test.

Within human intestinal microbiome communities, most species are represented by a single dominant strain (Truong et al., 2017), although we have previously observed complex strain dynamics during episodes of domination by *E. faecium* in allo-HCT patients (Dubin et al., 2019). For five species within three genera of interest, we identified the sequence signatures of the dominant strains on the basis of marker gene polymorphisms using the StrainPhlAn algorithm (Truong *et al*., 2017) and calculated tree-based phylogenetic distances between dominant strains in consecutive patient samples (**Figure 2.9A**). Small phylogenetic distances between strains in consecutive samples indicate dominant-strain convergence, while large distances suggest dominant-strain divergence over time. For most species, and particularly *E. faecium*, the phylogenetic distance between consecutive samples declined over time (**Figure 2.9B**). This temporal pattern suggests reduction of within-species genetic variability across multiple species over time and the rise of a dominant subtype. Furthermore, subtype variability was inversely associated with species relative abundance (**Figure 2.9B**). This correlation might be a consequence of so-called "selective sweeps" (Bendall et al., 2016; Diaz Caballero et al., 2015) by comparatively better-fit strains or loss of variability due to a population bottleneck that may occur during allo-HCT (Ghalayini et al., 2018). Subspecies diversification following a parabolic fitness landscape has been observed in vancomycin-resistant *E. faecium* isolated from longitudinal stool sampling of allo-HCT patients (Dubin *et al*., 2019), which this method of strain classification focusing on dominant strains might fail to capture. We observed that antibiotic exposure was a significant predictor of dominant strain genetic convergence within species *E. faecium*, while non-antibiotic exposure was not (**Figures 2.9C-D**).

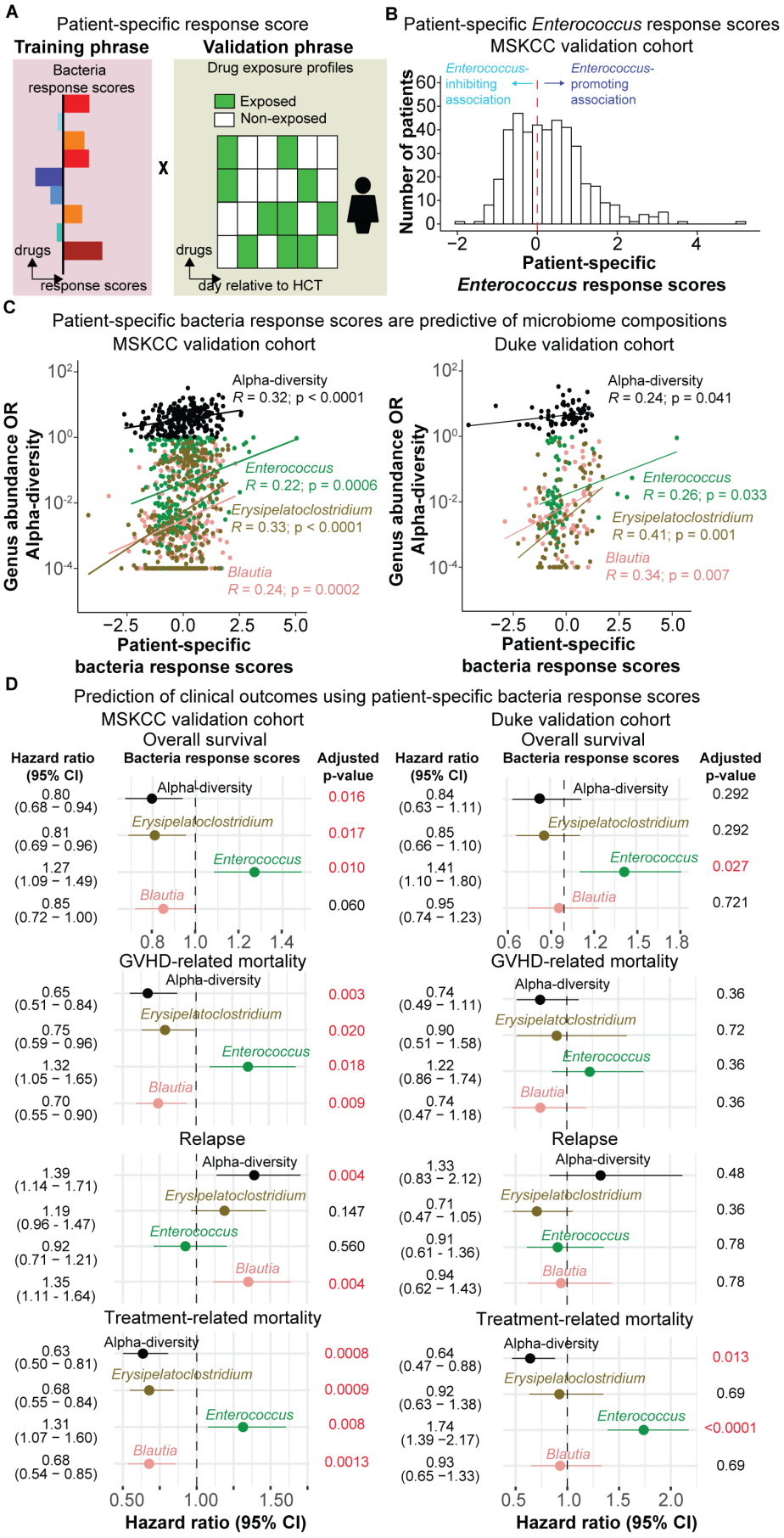
A
Patient-specific response score
Training phrase
Validation phrase
Bacteria response scores
drugs
response scores
X
Drug exposure profiles
Exposed
Non-exposed
drugs
day relative to HCT
B
Patient-specific Enterococcus response scores
MSKCC validation cohort
Enterococcus-inhibiting association
Enterococcus-promoting association
Number of patients
Patient-specific Enterococcus response scores
C
Patient-specific bacteria response scores are predictive of microbiome compositions
MSKCC validation cohort
Duke validation cohort
Genus abundance OR Alpha-diversity
Alpha-diversity R = 0.32; p < 0.0001
Enterococcus R = 0.22; p = 0.0006
Erysipelatoclostridium R = 0.33; p < 0.0001
Blautia R = 0.24; p = 0.0002
Alpha-diversity R = 0.24; p = 0.041
Enterococcus R = 0.26; p = 0.033
Erysipelatoclostridium R = 0.41; p = 0.001
Blautia R = 0.34; p = 0.007
Patient-specific bacteria response scores
D
Prediction of clinical outcomes using patient-specific bacteria response scores
MSKCC validation cohort
Duke validation cohort
Overall survival
GVHD-related mortality
Relapse
Treatment-related mortality
Hazard ratio (95% CI)
Bacteria response scores
Adjusted p-value
Alpha-diversity
Erysipelatoclostridium
Enterococcus
Blautia
0.80 (0.68 – 0.94) 0.016
0.81 (0.69 – 0.96) 0.017
1.27 (1.09 – 1.49) 0.010
0.85 (0.72 – 1.00) 0.060
0.84 (0.63 – 1.11) 0.292
0.85 (0.66 – 1.10) 0.292
1.41 (1.10 – 1.80) 0.027
0.95 (0.74 – 1.23) 0.721
0.65 (0.51 – 0.84) 0.003
0.75 (0.59 – 0.96) 0.020
1.32 (1.05 – 1.65) 0.018
0.70 (0.55 – 0.90) 0.009
0.74 (0.49 – 1.11) 0.36
0.90 (0.51 – 1.58) 0.72
1.22 (0.86 – 1.74) 0.36
0.74 (0.47 – 1.18) 0.36
1.39 (1.14 – 1.71) 0.004
1.19 (0.96 - 1.47) 0.147
0.92 (0.71 – 1.21) 0.560
1.35 (1.11 - 1.64) 0.004
1.33 (0.83 – 2.12) 0.48
0.71 (0.47 – 1.05) 0.36
0.91 (0.61 - 1.36) 0.78
0.94 (0.62 – 1.43) 0.78
0.63 (0.50 – 0.81) 0.0008
0.68 (0.55 – 0.84) 0.0009
1.31 (1.07 – 1.60) 0.008
0.68 (0.54 – 0.85) 0.0013
0.64 (0.47 – 0.88) 0.013
0.92 (0.63 – 1.38) 0.69
1.74 (1.39 –2.17) <0.0001
0.93 (0.65 –1.33) 0.69
Hazard ratio (95% CI)

Figure 2.10. Drug exposure profiles are predictive of future microbiome trajectories and allo-HCT patient outcomes in two distinct validation cohorts. A, Schematic of the patient-specific bacteria response score calculation. **B**, Patients-specific *Enterococcus* response scores in the validation cohort were derived based solely on drug exposure profiles (between day -14 to 14 relative to HCT) and bacteria response scores presented in **Figure 2.6C**. A negative score indicates that the drug exposure profile is associated with an *Enterococcus*-inhibiting effect, while a positive score indicates that the drug exposure profile is associated with an *Enterococcus*-promoting effect. **C**, Pearson's correlation between patient-specific bacteria response scores and observed genus relative abundance or alpha-diversity in samples collected between day 14 and 45 relative to HCT in the MSKCC validation cohort (423 patients) and Duke cohort (142 patients). Adjusted p-values by Benjamini-Hochberg's correction. **D**, Patient-specific bacteria response scores are predictive of overall and cause-specific mortality in the MSKCC and Duke validation cohorts, in each respective multivariate Cox proportional hazard or Fine-Gray model, controlled for age, sex, conditioning intensity, graft source and underlying disease. Adjusted p-values by Benjamini-Hochberg's correction.

Drug-microbiome associations are predictive of future microbiome trajectories and clinical outcomes following allo-HCT

Having demonstrated that drug exposures are associated with microbiota changes, and in light of previous reports associating fecal microbial composition with allo-HCT clinical outcomes, we next asked whether patterns of drug exposure alone could predict mortality independent of microbiome data. Using drug-exposure data from a separate MSKCC validation cohort, we defined patient-specific response scores as metrics that quantify the net response of microbiome features to drug-exposure profiles (**Figure 2.10A**). For example, the patient-specific *Enterococcus* response scores translated patient drug exposure profiles into relative risk of *Enterococcus* expansion (**Figure 2.10B**). We tested these patient-specific bacteria response scores against the outcomes of all-cause and specific-cause mortality in two independent validation cohorts: a subset of 423 MSKCC patients who were not included in the PARADIGM training cohort, as well as 142 patients from an independent cohort from Duke University Medical Center. All 62 drug exposures

were considered for the MSKCC validation cohort, while only antibiotic exposures were evaluated in the Duke validation cohort. Patient characteristics in these two cohorts are outlined in **Table 2.2**.

We observed that patient-specific bacteria response scores based on drug exposures between day -14 to 14 relative to HCT were significantly and positively correlated with observed taxa relative abundance and alpha-diversity, respectively, from samples collected between day 14 and 45 relative to HCT in both the MSKCC and Duke validation cohorts (**Figure 2.10C**). Furthermore, patients whose drug exposure profiles predicted higher *Enterococcus* expansion were at an increased risk of all-cause and transplant-related mortality following allo-HCT in the MSKCC and Duke validation cohorts. Specifically, in the MSKCC validation cohort, patient-specific *Enterococcus* response scores were also significantly associated with an increased risk of GVHD-related mortality (**Figure 2.10D**). Conversely, patients whose drug exposure profiles predicted *Eryispelatoclostridium, Blautia* or alpha-diversity preservation had a decreased risk of all-cause mortality in both the MSKCC discovery and validation cohorts (**Figure 2.10**). Overall, we demonstrated that drug-microbiome associations are predictive of subsequent changes in the intestinal microbiome compositions post-exposure, and of clinical outcomes in allo-HCT patients.

The framework for this analysis is a hypothesis that drug exposures affect the intestinal microbiota, which in turn shapes clinical outcomes. In some scenarios, however, patients at high risk for adverse outcomes (for a variety of reasons unrelated to the microbiome) might have received drugs that affect the intestinal bacteria. To explore these possibilities, we focused on *Enterococcus* and compared the hazard ratios of *Enterococcus* relative abundance with the corresponding patient-specific response scores in a Cox

proportional hazard ratio model. We reported a significant association between *Enterococcus* abundance and mortality risk, which was stronger than the association between patient-specific *Enterococcus* response scores and overall mortality, specifically in the MSKCC validation cohort (**Figure 2.11**). However, we also observed that patient-specific response score remained a statistically significant predictor of mortality risk when controlled for intestinal microbiome compositions (**Figure 2.11**). Altogether, these results suggest that association between drug exposures and clinical outcomes is partially dependent on drug interactions with the intestinal microbiota.

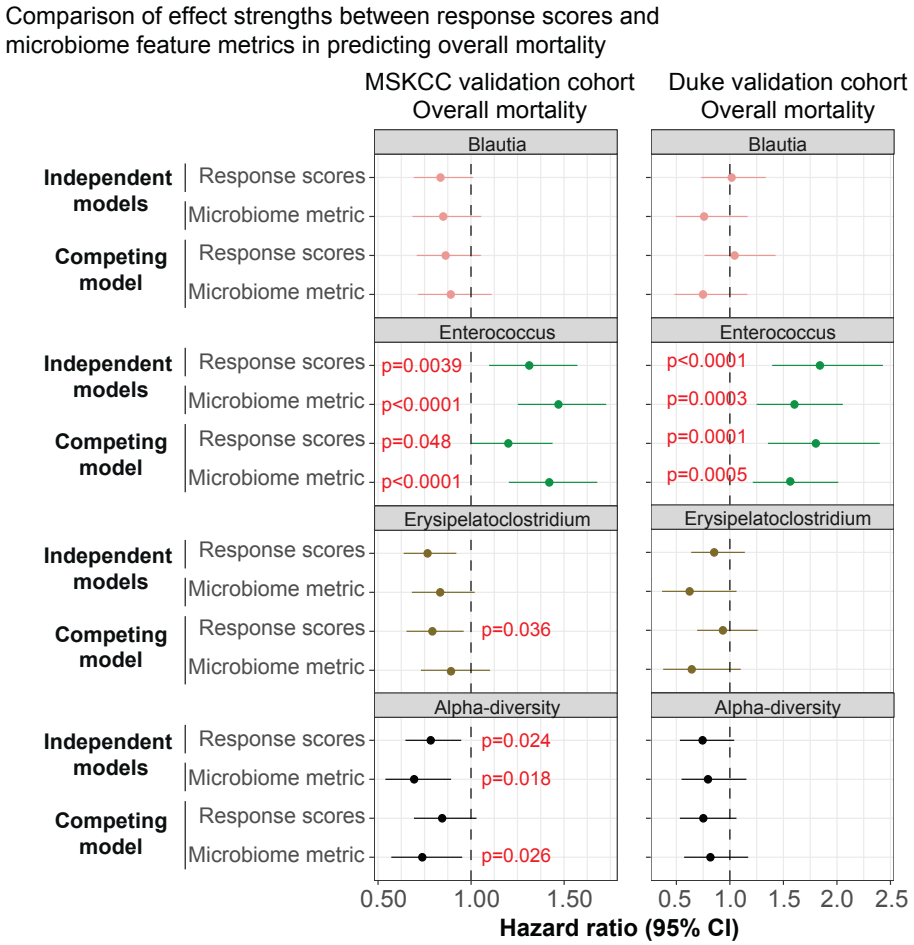

Figure 2.11. Investigation of causal relationship between drug exposures, microbiome and mortality. Microbiome feature metrics (taxa relative abundance or alpha-diversity) and corresponding response scores were compared in terms of their associations with

overall mortality risk. Measurements were fit into either an independent (which consider either microbiome metrics and response scores) or a competing (which considers microbiome metrics and response scores together) multivariate Cox proportional hazard model.

Discussion

We developed PARADIGM, a computational method that identifies the associations between drug exposures and intestinal microbial dynamics in humans. At its core, PARADIGM analyzes how discrete states of intestinal microbial compositions respond to both antibiotic and non-antibiotic drugs. While other computational methods have largely focused on antibiotics (Kanjilal et al., 2020; Vatanen et al., 2018), our approach further reveals the associations between many non-antibiotic drugs and microbial dynamics in high resolution by analyzing a large dataset of daily stool samples from allo-HCT patients.

Our method was able to infer meaningful associations between drug exposures and microbiome despite the various confounding parameters such as the clinical symptoms prompting these drug exposures, which is particularly important when analyzing medications used to treat gut toxicity. We validated our findings by comparing bacterial response scores derived from this real-world patient study with published *in vitro* observations (Maier *et al*., 2018) and found that our estimates were significantly correlated with the reported data. Furthermore, the bacterial response scores, calculated based solely on patient drug exposures, allowed us to predict future microbial changes and patient outcomes in two independent validation cohorts, particularly for *Enterococcus* responses. These results demonstrate that PARADIGM can generate hypotheses with both biological and clinical relevance.

In conclusion, we provide insights into the associations between pharmacological exposures and changes in intestinal microbial composition at the early stage of transplant. The algorithm we have developed, PARADIGM, identifies biologically meaningful and clinically relevant associations between drug exposures and intestinal microbial dynamics in humans. PARADIGM facilitates the integration of drug exposures from many classes of xenobiotic agents, microbiome dynamics, and clinical outcomes to understand the determinants of microbiome health. This computational framework is well-suited for longitudinal data and may be built upon in the future to investigate other environmental parameters of interest, such as dietary intake or other components of the "exposome", or applied to other disease settings in which drug-microbiome interactions are of clinical importance (Vermeulen et al., 2020).

Materials and Methods

Study population of human subjects

The patient and fecal sample cohort in this study has been described in previous studies(Liao et al., 2021; Peled *et al*., 2020; Stein-Thoeringer *et al*., 2019). Stool samples were collected at two different transplant centers, Memorial Sloan Kettering Cancer Center (MSKCC) from April 2009 to September 2019, and Duke University Medical Center from July 2012 to April 2018. Participants in the observational cohorts at both MSKCC and Duke provided written informed consent for the use of their stool samples and clinical data. The use and analysis of these specimens for this study was approved by Institutional Research Boards at both institutions (MSKCC: #16-834; Duke: PRO0006268 and Pro00050975). Stool samples collected after day -30 relative to a first allo-HCT, and before

day -10 relative to a second allogeneic transplant (if applicable) were included in the analysis. A subset of patients from MSKCC participated in a randomized clinical trial of fecal microbiota transplantation (FMT; NCT02269150) (Taur *et al*., 2018). Stool samples from patients in the control arm, as well as from patients in the FMT arm collected pre-FMT were included for analysis. Post-FMT stool samples were excluded.

Stool collection and storage

As DNA extraction procedures, sample-handling environment, sequencing and bioinformatics pipelines are important sources of variability in microbiome data, we minimized bias and institutional batch effects by collecting and freezing samples at each center following the same protocol. All stool samples were collected, aliquoted and frozen at their respective clinical centers; extraction, sequencing, and analyses were performed centrally at MSKCC.

Fecal samples were collected in both inpatient and outpatient settings. At MSKCC, inpatient samples were collected by nursing staff from toilet inverted "hats" into ~100 ml-sized containers at the bedside, promptly delivered to the laboratory via pneumatic tube, and refrigerated at 4°C until aliquoted for long term storage at -80°C. Outpatient stools were collected in the patients' homes using a commode specimen collection system, after which the entire collection bin was capped, placed inside a biohazard zip lock bag to prevent leakage, and deposited inside a 8 x 6 x 4.25" foam container along with pre-chilled freezer packs. Samples were either brought by the patient to a clinic appointment or shipped directly from the patient's home to the laboratory via courier.

Upon receipt by the Molecular Microbiology Facility Laboratory at MSKCC, each sample was given a unique ID and recorded into the fecal biobank database. Approximately 0.5 ml of whole stool was aliquoted without any preservative solution into four 2 ml cryovials using a clean disposable spatula and transferred into a -80°C freezer until thawed for DNA processing. Duke samples were aliquoted similarly at Duke, batch-shipped frozen to MSKCC, and extracted and sequenced in the same way as the MSK samples at the Molecular Microbiology Facility Laboratory.

DNA extraction

Bacterial DNA was extracted using an optimized phenol-chloroform protocol to recover nucleic acids from tough microbes commonly present in stool samples, as previously described(Peled *et al*., 2020; Rolling et al., 2021). Briefly, 200-300 mg of solid or 200-300 µl of liquid stool were aliquoted and the respective wet weight was recorded. Fecal samples were resuspended in 500 µl of extraction buffer (i.e., 0.2 M NaCl, 0.2 M Tris-HCl, pH 8.0, and 20 mM ethylenediaminetetraacetic acid, prepared fresh). The mixture was combined with 0.1 mm zirconia/silica beads (approximately 500 µl), 200 µl of sodium dodecyl sulfate, and 500 µl of phenol-chloroform-isoamyl alcohol (25:24:1 solution). The bacterial cells were lysed by mechanical disruption using a bead-beater for 2 minutes at over than 3000 rpm. Detritus was spun down at 16,000*g* at 4°C for 5 minutes. The upper aqueous layer was transferred to a clean 1.7 ml tube and mixed with 100 µl of extraction buffer as described above, and 500 µl of phenol-chloroform-isoamyl alcohol (25:24:1). The solutions were homogenized by inversion and spun down again at 16,000*g* for 5 minutes at 4°C. The upper aqueous layer was recovered, and this process was repeated

for a total of three times. After the final round of phenol-chloroform-isoamyl alcohol was completed, 400 µl of the top aqueous layer was combined with 40 µl of sodium acetate and 880 µl of cold 100% ethanol. The mixture was vortexed and then frozen for at least 20 minutes, preferably overnight. After the ethanol freeze incubation period, samples were spun down at 16,000*g* for 20 minutes at 4°C to pellet the DNA. The upper aqueous layer was aspirated, and the visible DNA pellet was resuspended in 200 µl of TE buffer containing 100 mg/mL of RNase solution, followed by an incubation at 50°C for 20 minutes. The genomic DNA was further purified using the QIAamp DNA mini kit (Qiagen) following the manufacturer instructions. The purified DNA was eluted in 100 µl of ultrapure water and stored at -80°C prior to quantification of DNA yield and subsequent PCR amplification.

Sequencing

16S rRNA V4-V5 barcoded amplification and multiplexing: The amplification, multiplexing and sequencing of 16S rRNA from extracted DNA has been previously reported(Taur *et al*., 2018). Briefly, genomic purified DNA was diluted if necessary and 50 ng was used as template during PCR amplification. The V4-V5 region of the 16S rRNA gene was amplified with the primers 563F (5′-nnnnnnnn-NNNNNNNNNNNN-AYTGGGYDTAAAGNG-3') and 926Rb (5′-nnnnnnnn-NNNNNNNNNNNN-CCGTCAATTYHTTTRAGT-3), where '*N*'s represent unique 12-base pair Golay barcodes and '*n*'s represent additional nucleotides to offset the sequencing of the primers(Caporaso et al., 2012). Duplicate PCR reactions were performed for each sample with 2.5 U of Platinum Taq DNA polymerase and 0.5 mM of forward and reverse primers

at 94°C for 3 minutes, followed by 27 cycles of 94°C for 50 seconds, 51°C for 30 seconds and 72°C for 1 minute and a final elongation step at 72°C for 5 minutes. Amplicons were purified using the Qiaquick PCR Purification Kit (Qiagen) after pooling the sample replicates. The purified PCR products were quantified using Agilent Technologies 4200 TapeStation and multiplexed at equimolar amounts. The obtained pools of barcoded 16S amplicons went to further processing for library preparation and sequencing on an Illumina MiSeq platform at paired-end 250 base pair (bp) at the MSKCC Integrated Genomics Operation sequencing core. Extraction blanks were included in each extraction batch as negative controls. These blanks were PCR-amplified but did not show PCR products and were subsequently removed from demultiplexing and sequencing.

Shotgun metagenomic sequencing: Shotgun metagenomic sequencing was conducted as previously described(Dubin *et al*., 2019). Extracted DNA was sheared to a target size of 650 bps using a Covaris ultrasonicator. DNA was then prepared for sequencing using the Illumina TruSeq DNA library preparation kit and sequenced using the Illumina platform targeting approximately 10–20 million reads per sample with 100-bp paired-end reads.

Sequencing bioinformatics pipeline

16S rRNA gene sequencing: Reads were quality-filtered, deduplicated, denoised and amplicon sequence variants (ASVs) were inferred following the DADA2 pipeline (Divisive Amplicon Denoising Algorithm) (Callahan et al., 2016). Reads were truncated at a length of 180 bp for forward and reverse reads to ensure sufficient quality and length to use overlapping paired end reads. Default values were used for filtering and trimming reads

prior to inferring sequence variants through function *filterAndtrim()* (maxN=0, maxEE=2, trunQ=2). Reads per sample were capped at 100,000 prior to sequence variance inference. Samples with more than 100,000 reads were sub-sampled. Each sequencing run was analyzed separately before merging the ASV counts across multiple runs.

Samples with poor sequencing quality were filtered out using the following criteria:

1. Less than 30% reads remaining after filtering and trimming
2. Less than 1,000 reads remaining after filtering and trimming
3. More than 5% adapter contamination

Taxonomic classification was annotated according to the NCBI 16S rRNA sequence database. Alpha-diversity was calculated using the reciprocal Simpson index at the ASV level, which is the metric used for all alpha-diversity measures in this study. Beta-diversity was computed according to the Bray-Curtis distances at the genus level using the *beta_diversity.py* script in the QIIME bioinformatic pipeline (Caporaso et al., 2010). We performed t-distributed stochastic neighbor embedding (tSNE) dimensionality reduction for visualization of the intestinal microbiota compositions using a Bray-Curtis β-diversity matrix at the genus level with R package Rtsne (max_iter = 10,000; perplexity = 75; theta = 0.2).

Shotgun metagenomic sequencing: The right and left side of a read in a pair was trimmed to Q10 using the Phred algorithm, using the bbduk.sh script in the BBMap package (https://www.sourceforge.net/projects/bbmap/). A pair of reads was dropped if either read had a length shorter than 51 nucleotides after trimming. The 3'-end adapters were trimmed using a kmer of length 31, and a shorter kmer of 9 at the other end of the read. One mismatch was allowed in this process, and adapter trimming was based on pair

overlap detection (which does not require known adapter sequences) using the 'tbo' parameter. The 'tpe' parameter was used to trim the pair of reads to the same length. Removal of human contamination was done using Kneaddata with paired end reads, employing BMTagger. The BMTagger database was built with human genome assembly GRCh38.

Samples with poor sequencing quality were filtered out using the following criteria:

1. Less than 50% reads remaining after trimming with BBMap
2. Less than 1 million reads remaining after trimming
3. More than 5% adapter contamination

After decontamination, the paired-end reads were concatenated to a single FASTQ file as the input for functional profiling with the HUMAnN 3.0 pipeline (Beghini et al., 2021). After aligning to the updated ChocoPhlAn and UniRef90 database with default settings, the samples were renormalized by library depth to copies per million. MetaPhlAn 3.0 was used to identify taxonomic compositions with the relative abundance of the species using parameter -t rel_ab (Beghini *et al*., 2021). Similar to the 16S rRNA-sequenced data, we performed tSNE analysis for visualization of shotgun metagenomic compositions using a Bray-Curtis β-diversity matrix at the species level with R package Rtsne (max_iter = 3,000; perplexity = 20; theta = 0.1).

Clustering of the intestinal microbiota compositions

Three unsupervised clustering methods were used to identify distinct clusters of the intestinal microbiota compositions within the MSKCC validation cohort. Unsupervised k-means clustering and hierarchical clustering methods were applied to the Bray-Curtis β-

diversity matrix. Dirichlet Multinomial Mixture (DMM) was applied to the raw count matrix at the genus level. For all three methods, cluster size parameter from 2 to 15 clusters was evaluated, ultimately choosing ten clusters for downstream analysis. Since k-means clustering partitioned samples more evenly, we utilized k-means clusters for our subsequent analyses. We determined ten clusters to be a good balance between reducing the complexity of the 16S rRNA sequencing data while still representing the variability of the intestinal microbiome compositions during allo-HCT.

Domination threshold

We defined sample domination by a single taxonomic units using the threshold of ≥ 0.3 of the 16S rRNA sequencing relative abundance (Peled *et al*., 2020). The taxonomic color scheme was adapted from the R package yingtool2 (https://github.com/ying14/yingtools2) and a previous publication (Peled *et al*., 2020).

Derivation of a mathematical model of microbial dynamics termed PARADIGM

A biologically motivated, simplified mathematical model termed PARADIGM was developed to model microbial dynamics while simultaneously considering the effects of drug exposures on the intestinal microbiota. A naïve model with n possible clusters in a time interval Δt for L distinct drugs would require $\Delta t \bullet n^2 2^L$ parameters, which would quickly saturate the amount of data available in most practical applications. The following characteristics of intestinal microbial dynamics during allo-HCT were considered to simplify our model: (i) The microbial composition is more likely to stay in same cluster within a day window; (ii) Drug exposures perturb clusters differently; (iii) Transitions are

easier to occur among close clusters compared to distant clusters; and (iv) Transition events vary over time. As such, we developed a model that include parameters for time, drug exposures and distance between clusters.

We defined microbiota dynamics in terms of two possibilities: the rate at which the microbiota composition stays in the same cluster (self transition) and the rate at which it attracts transitions from other clusters (attractor transition). Formally, attractor transition is a force, measured in terms of probability, that describes the magnitude in which a given cluster may receive transitions from any clusters other than itself, in a pair of daily collected samples. For example, the attractor transition to Cluster 1 is defined as the patient moves to cluster 1 at time t_{d+1}, from any clusters between 2 to 10 at time t_d. We note that, "attractor" has a particular definition in statistical physics, which does not directly translate into the definition outlined in this study.

Let $T_{i,j}$ represent the transition probability from cluster i to j at a single day resolution and w_{self} and $w_{attractor}$ represent the parameters associated with self and attractor transitions. The transition probabilities can be summarized as:

$$\mathrm{T}_{\copyright,\mathrm{j}} = \frac{\mathrm{e}^{\mathrm{w}_{\mathrm{self}}(\mathrm{j})\cdot\delta_{i,\mathrm{j}}+(1-\delta_{i,\mathrm{j}})\bullet(\mathrm{w}_{\mathrm{attractor}}(\mathrm{j}|\mathrm{i}))}}{\sum_{\mathrm{j}} \mathrm{e}^{\mathrm{w}_{\mathrm{self}}(\mathrm{j})\cdot\delta_{i,\mathrm{j}}+(1-\delta_{i,\mathrm{j}})\bullet(\mathrm{w}_{\mathrm{attractor}}(\mathrm{j}|\mathrm{i}))}} \quad (1)$$

Where $\delta_{\mathrm{i,j}}$ represents Kronecker delta.

The transition dynamics are simplified to a *2x2* transition matrix per cluster i that represents *self* (i) and *not-self* (i^c) states and the rates of transition between those states:

$$\tilde{T}(i) = \begin{pmatrix} p_{i,i} & p_{i,i^c} \\ p_{i^c,1} & p_{i^c,i^c} \end{pmatrix} \quad (2)$$

And the self and attractor parameters are formally represented as:

$$\begin{cases} \log\left(\frac{p_{i,i}}{p_{i,i^c}}\right) = \mathrm{w}_{self}(i) \\ \log\left(\frac{p_{i^c,i}}{p_{i^c,i^c}}\right) = w_{attractor}(i) \end{cases} \quad (3)$$

Inference of drug influence on cluster dynamics via elastic net regularized regression

To investigate high-resolution cluster dynamics in response to drug exposures, we included samples that were collected 1 day apart between day -14 and 14 relative to HCT from patients in the MSKCC discovery cohort. The criteria for drug inclusion in this analysis are as followed: (1) Drugs were administered via oral/IV routes; (2) Drug exposures occurred between day -14 to 14 relative to HCT; (3) Drugs were administered to at least 5% and no more than 90% patients in the discovery cohort; and (4) Drug exposures occurred in at least 15 patients with single-day resolution samples.

Equations 1-3 provide a formal definition of our microbial dynamics in terms of self and attractor transitions. In practice, we took advantage of a logistic regression fit to solve Equation 3 by assuming a binary value for each transition to a given cluster *i*. We defined the value of 1 for transitions towards cluster *i* and a value of 0 otherwise.

We used elastic net regularized regression for feature selection to estimate the influence of drug exposures on self and attractor transitions of each cluster, using the R package caret, e1071 and glmnet for model fitting and parameter tuning. Formally, the coefficients that impact transition probabilities were computed as following:

$$\begin{cases} \mathrm{w}_{self}(i,t,X) = \beta_{self,i,0} + \beta_{self,i,1}t + \sum_{k \in Drug\ set} \beta_{self,i,k} \bullet X_{k,t} \\ w_{attractor}(i,t,X) = \beta_{attr.,i,0} + \beta_{attr.,i,1}t + \beta_{attr.,i,2}d_{i,j} + \sum_{k \in Drug\ set} \beta_{attr.,i,k} \bullet X_{k,t} \end{cases} \quad (4)$$

The parameters $X_{k,t}$ have a value of *1* in days patients were exposed to drug *k* and *0* otherwise. The 10-fold cross-validation partitions were pre-specified such that samples from the same patient are always in the same partition using function *trainControl()* and *train()* in R package caret.

Calculation of bacterial response score of drugs

The dynamic model we proposed can estimate the association between a given drug exposure with quantitative microbiome features. The self and attractor parameters predicted in **Equation 4** can be used to construct transition probabilities according to **Equation 1**. Let the probability of switching from cluster *i* to *j* in the absence of any drug exposure be $\pi_{i\to j}(no\ drug)$, the bacterial response score of a drug *d* on taxon *y* is defined as:

$$C(y|d) = \sum_{i,j} (\pi_{i\to j}(d) - \pi_{i\to j}(no\ drug)) \bullet log(\frac{p_j(y)}{p_i(y)}) \quad (5)$$

where $p_i(y)$ represents the average relative abundance of taxon *y* from all samples in cluster *i*. A negative score indicates an antibacterial effect of drug *d* on taxon *y*. The transition matrix used to estimate antibacterial score was computed at time *t=0*.

The bacterial response scores estimated from **Equation 5** were compared to published *in vitro* measurements of anti-bacterial activities (Maier *et al*., 2018). The *in vitro* measurements evaluated the effect of specific drugs using the area under the curve (AUC) of bacterial growth. An AUC value significantly below 1 indicates that the drug inhibited bacterial growth *in vitro*. We restricted our analysis to 19 bacteria species that appeared in more than 10% of the MSKCC samples with relative abundance above 10^{-4}. We identified the relative abundance of a given species by summarizing the relative

abundance of all ASVs mapped to that species. Our prediction was considered consistent with *in vitro* measurements when bacterial response scores were negative for bacteria-drug pairs that showed inhibition in that study. Statistical significance was evaluated using two-sided Wilcoxon's rank-sum test.

Defining patient-specific bacteria response scores based on drug exposure profiles

A risk score that is associated with patient outcomes can be computed from how drug exposures are predicted to influence the microbial dynamics. We defined specific target values, $\square$, which are features of the intestinal microbiota associated with each cluster, and adapted **Equation 5** to compute a risk score, S, for each patient by averaging the contribution of all drug exposures that occurred within day -14 to 14 relative to HCT.

$$S = \sum_{t \in time} \sum_{d \in drugs} \sum_{i,j} (\pi_{i \to j}(d,t) - \pi_{i \to j}(no\ drug,t)) \bullet log(\frac{\tau_j(y)}{\tau_i(y)}) \quad (6)$$

For this work, we considered four target values: relative abundance of *Enterococcus, Erysipelatoclostridium, Blautia*, and alpha-diversity. Patient-specific bacteria response scores for each of the considered microbiome features were calculated based on drug exposure profiles between day -14 and 14 and compared with observed taxa relative abundance or alpha diversity in samples collected between day 14 to 45 relative to HCT using Pearson's correlation, with adjusted p-values by Benjamini-Hochberg's correction. The median values were taken for patients with multiple samples available.

Imputation of missing samples

The dynamic model defined in **Equation 1** can be used to simulate microbial dynamics and to impute cluster state in time points without sample collection. **Equation 1**

defines transition probabilities at single day resolution and forms the fundamental unit to predict microbial dynamics. Let $X(t)$ be the compositional state at time t, the value $T_{i,j}$ is equivalent to $P(X(t+1)|X(t))$. Assuming first-order Markov chain, the transition probability matrix between samples at longer time distance can be computed from its fundamental units as following:

$$\mathrm{T}(\mathrm{t}_0 \rightarrow \mathrm{t}_0 + \mathrm{d}) = \mathrm{T}(\mathrm{t}_0) * \mathrm{T}(\mathrm{t}_0 + 1) * \ldots * \mathrm{T}(\mathrm{t}_0 + d - 1) \quad (7)$$

Where $T(t_0)$ represents transition probabilities at single day resolution between time t_0 to t_0+1.

The imputation of cluster states at a given time point can be performed by considering the state at nearby time points. Interpolations can be made based on past samples (forward interpolations), future samples (backward interpolations) or both (forward-backward interpolations). Consider the estimation of a cluster state at time point t from samples available at time point $t-a$ and $t+b$. Forward interpolation imputes the data point based on the probability $P(X_t| X_{t-a})$, backward prediction from $P(X_t| X_{t+b})$ and backward-forward interpolation from $P(X_t| X_{t+a}, X_{t+b})$. Forward interpolations can be estimated from **Equation 7**. The backward interpolations were obtained by adapting model described in **Equation 1** to predict cluster state at time t based on cluster state at time $t+1$ and exposures at time t. The backward-forward interpolation takes advantage of both forward and backward equations and is computed as:

$$P\left(X_j \middle| X_{j-a}, X_{j+b}\right) \propto \frac{P\left(X_j \middle| X_{j+b} = x_b\right) \bullet P\left(X_j \middle| X_{j-a} = x_a\right)}{P(X_j)} \quad (8)$$

A modified forward-backward interpolation was also tested in which final predictions is obtained by either forward or backward probabilities:

$$P(X_j|X_{j-a}, X_{j+b}) \propto \begin{cases} P(X_j|X_{j+b} = x_b)\ if\ w(t) > c \\ P(X_j|X_{j-a} = x_a)\ if\ w(t) \leq c \end{cases} \quad (9)$$

Where *w(t)* indicates whether forward or backward interpolation should be used to predict state at time *t*, and c indicates the mean of the forward-backward prediction probabilities.

The imputation accuracy for each model described above was estimated using 10-fold cross validation, in which 90% of patients were selected as the training set and 10% as the test set. In the training set, the logistic regression model defined in **Equation 5** was used to estimate the parameters for *self* and *attractor* weights as well as the decision parameter for modified forward-backward prediction. In the test set, the regression coefficients were used to reconstruct transition matrices and compute forward, backward, forward-backward as well as modified forward-backward probabilities. The microbial state was classified by the cluster with the maximum probability.

Linear mixed-effects model of drug-species associations using shotgun metagenomic profiles

We defined a linear mixed-effects model to investigate the association between drug exposures and changes at a species-level resolution. Species relative abundance was predicted by MetaPhlAn 3.0 using shotgun metagenomic data. The model considers pairs of samples that were collected between day -14 and 14 relative to HCT and less than five days apart, and with a minimum species relative abundance of 10^{-4}. We assumed that changes in species relative abundance depend on time of sample collection and drug exposure. Formally, let a given sample pair be collected at time t_1 and t_2, respectively. The log difference in species relative abundance, ln (Sp_{t2}/ Sp_{t1}) depends on fixed effects of drug exposures and a random effect relative to time of sample collection, t_1, as following:

$$\ln\left(\frac{Sp_{t1}}{Sp_{t2}}\right)(t_1, t_2, \mathrm{X}) = 0 + \beta \bullet X_{t1,t2} + (1|t_1) \quad (10)$$

The parameter t_1 represents time of first sample collection in the sample pair, binned into weekly intervals relative to HCT, as a random effect variable. $X_{t1,t2}$ represents a given drug exposure and has the value of 1 if patients were exposed to the drug in the time interval $t_1$and t_2. P-values from the linear mixed-effects model were adjusted for multiple hypothesis testing using Benjamini-Hochberg's correction. Equation 10 was solved using R package lme4.

Dominant strain dynamics and its association with drug exposures

StrainPhlAn 3.0 was used to profile bacterial community at the strain-level resolution (Beghini *et al*., 2021; Truong *et al*., 2017). This algorithm identifies the most dominant strain per species per sample by reconstructing dominant consensus sequence variants across species-specific marker genes. We applied StrainPhlAn 3.0 to a dataset of 980 shotgun metagenomic shotgun samples in the MSKCC discovery cohort, which returned the multiple sequence alignment of the dominant strain for a given species and the RaxML (Randomized Axelerated Maximum Likelihood) phylogenetic tree across samples. The phylogenetic tree was then used to calculate the dominant strain phylogenetic distance across a pair of samples. Phylogenetic distance is defined as the branch length between two nodes in the StrainPhlAn phylogenetic tree, normalized over the total branch length of the tree. Branch length was calculated using R function *cophenetic()*. We focused on strains from five species of interest, including *Blautia coccoides*, *Blautia producta*, *Enterococcus faecalis*, *Enterococcus faecium* and *Erysipelatoclostridium ramosum*.

For analysis of dominant strain convergence in association with antibiotic and non-antibiotic exposures, we compared the phylogenetic distances of dominant strains within the species *E. feacium* across pairs of samples that were collected less than five days apart between day -14 and 14 relative to HCT. Exposures were considered true if patients received the drug during sample pair collection time. Antibiotic exposure includes the seven investigated antibiotics in this study. Non-antibiotic exposure comprises seven drugs with the highest absolute *Enterococcus*-response score values, indicating strongest associations with changes in *Enterococcus* relative abundance. These non-antibiotic drugs include diphenoxylate/atropine, polyethylene glycol, levothyroxine, fentanyl, methotrexate, anti-thymocyte globulin and cyclosporine. Statistical significance was evaluated using two-sided Wilcoxon's rank-sum test.

Co-inclusionary and exclusionary relationships between bacteria and cluster stability

To identify the potential microbiome influence on the stability of *Enterococcus* domination in allo-HCT patients, we developed a logistic regression model with lasso penalty. This model takes into account the relative abundance of different bacterial genera on the stability of *Enterococcus*-high cluster 10. We defined cluster 10 stability as the rate at which the microbiota compositions stay in the same cluster 10, assuming a value of 1 when patients stay in cluster 10, and a value of 0 when patients move from cluster 10 to any other cluster among a pair of daily collected samples. We included as parameters antibiotics exposure, day of sample collection relative to HCT, alpha-diversity and relative abundance of top 20 most abundant genera in cluster 10. Antibiotic exposure includes exposure to any of the seven antibiotics investigated in this study. The input dataset consists

of pairs of daily samples collected between day -14 and 100 relative to HCT. Formally, the coefficients of these potential drivers of cluster 10 stability were computed as following:

$$\mathrm{w}_{self}(i,t,abx,p) = \beta_{self,0} + \beta_{self,1}t + \beta_{self,2}abx_t + \beta_{self,3}simpson\_reciprocal_t + \sum_{i \in Top\ 20\ genera} \beta_{self,i} \bullet p_{i,t} \quad (11)$$

In which $p_{i,t}$ represents the relative abundance of each taxon in the top 20 most abundant genera in cluster 10, and abx indicates if patients received any of the seven investigated antibiotics in this study on day t. Coefficient values indicate the direction and magnitude of association between each parameter and the stability of *Enterococcus*-high cluster 10, with a negative value indicating a negative association with *Enterococcus* domination stability.

Statistical analyses of survival outcomes

We used landmark analyses of survival beyond day 14 relative to HCT using R package survival. Patients were censored at the time of last contact or at the time of second allo-HCT (when applicable). All survivors were censored at two years of follow-up. Patients randomized to the FMT arm of the trial were excluded from analysis of clinical outcomes. Patient-specific bacteria response scores corresponding to a given microbiome feature were considered as a continuous variable in a multivariate Cox proportional hazard model, controlled for age, sex, conditioning intensity, graft source and underlying disease as variables with R function *coxph()*. P-values were adjusted for multiple hypothesis testing using Benjamini-Hochberg's correction.

Competing risk analyses were performed to identify the association between patient-specific bacteria response scores and cause-specific mortality. We investigated three competing events, namely relapse (defined here as relapse or progression of disease), GVHD-related mortality (defined here as death due to GVHD or after GVHD onset, without relapse), and transplant-related mortality (encompassing deaths from GVHD, infections and organ toxicities). For each competing event, multivariate Fine-Gray subdistribution hazard models were fit by R function *crr()* from R package tidycmprks. Hazard ratios are presented with the 95% confidence interval indicated in parentheses.

To tease out the potential causal relationships between medication, microbiome and mortality, we compared the effect sizes and statistical strengths of patient-specific bacteria response scores and observed genus relative abundance or diversity values in predicting all-cause mortality. To assure proper magnitude comparison, microbiome measurements were rescaled by Z-score normalization and fit into either an independent (which consider either microbiome metrics or response scores) or a competing (which considers microbiome feature and response score together) multivariate Cox proportional hazard model, controlled for age, sex, conditioning intensity, graft source and underlying disease. The independent and competing models consider only patients with stool samples collected between day 0 and 45 relative to HCT (MSKCC validation cohort: 340 patients; Duke validation cohort: 108 patients). Genus abundance was calculated from stool samples collected between day 0 and 45 relative to HCT, response scores were calculated using drug exposure profiles between day -14 to 14 relative to HCT. The median values were taken for patients with multiple samples available. P-values were adjusted for multiple hypothesis testing by Benjamini-Hochberg's correction.

CHAPTER THREE

ENTEROCOCCUS INDUCES MHC-II EXPRESSION IN THE INTESTINAL EPITHELIUM TO AGGRAVATE GVHD

Introduction

The commensal intestinal microbiota has co-evolved with the mammalian hosts and played an important role in maintaining local and systemic homeostasis (Artis, 2008). Therefore, perturbations to the intestinal microbiota community could impair the interactions between the gut microorganisms and their hosts, which in turn are associated with various diseases including infection, autoimmunity, and cancer (Hou et al., 2022). We and others have previously reported the frequently observed intestinal microbiota injuries in allo-HCT patients (Golob *et al*., 2017; Peled *et al*., 2020; Taur *et al*., 2012). Specifically, intestinal domination by *Enterococcus*, a genus of Gram-positive facultative bacteria, is associated with an increased risk of GVHD-related mortality in allo-HCT patients and in preclinical models (Holler *et al*., 2014; Stein-Thoeringer *et al*., 2019). However, the immunological mechanisms underlying this association remain unclear.

Pathogenic *Enterococcus* strains are associated with hospital-acquired infections and autoimmunity such as inflammatory bowel disease (Levitus et al., 2023; Seishima et al., 2019; Zhou et al., 2016). Conversely, commensal Enterococcus strains have demonstrated clinical benefits as probiotics and in the context of immune checkpoint blockade (Griffin et al., 2021; Hanchi et al., 2018; Rashid et al., 2023). The dual role of this bacteria genus highlights the need to understand the biological context underlying the beneficial versus harmful effects of *Enterococcus* in human health and disease.

Previous literature has mostly focused on the effect of *Enterococcus* on innate immunity such as DCs and macrophages (Leendertse et al., 2008; Ocvirk et al., 2015; Park et al., 2013). Host- and donor-derived APCs play an important role in the development of GVHD through allo-antigen presentation and donor T cell activation (Koyama *et al*., 2015; Shlomchik et al., 1999). In addition, IECs could also present antigens and engage with alloreactive T cells to initiate GVHD through the expression of major histocompatibility complex class II (MHC-II) (Koyama et al., 2011; Koyama *et al*., 2019). While professional APCs constitutively express MHC-II, non-professional APCs do not typically present antigens through MHC-II. However, various factors such as diurnal oscillation, dietary changes, inflammation, and alterations to the intestinal microbiota could induce MHC-II expression in non-professional APCs such as IECs (Heuberger et al., 2023; Tuganbaev et al., 2020). Here, we investigate the regulation of intestinal epithelium MHC-II expression by *Enterococcus* as a mechanism that aggravates GVHD severity.

Results

Enterococcus abundance is associated with increased GVHD mortality

To assess the impact of *Enterococcus* on GVHD outcomes, we used an MHC-disparate model of GVHD (C57BL/6J donors into BALB/cJ recipients). Lethally irradiated BALB/cJ recipients were transplanted with C57BL/6J splenic T cells along with T-cell-depleted bone marrow (BM+T) to trigger GVHD, with the T-cell-depleted BM (BM only) group serving as controls (**Figure 3.1A**). We observed an increase in endogenous *Enterococcus* relative abundance in GVHD mice on day +7 post-BMT, compared to the BM only control group and pre-transplant naïve baseline, as quantified by 16S rRNA

sequencing of fecal samples (**Figure 3.1B**). This observation is consistent with previous findings demonstrating that *Enterococcus* bloom occurs both in mouse models of GVHD and in allo-HCT patients (Stein-Thoeringer *et al*., 2019).

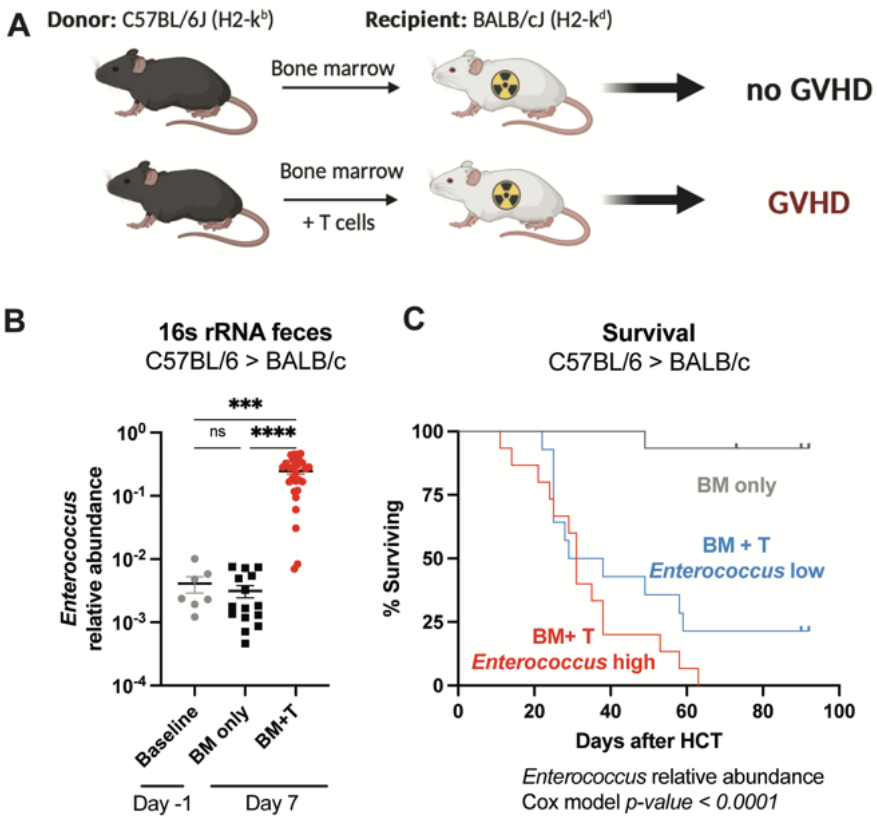

Figure 3.1. *Enterococcus* relative abundance is associated with increased mortality in am MHC-disparate mouse model of GVHD. A, Schematic of the MHC-disparate mouse model of GVHD in this study. Recipient BALB/cJ mice receiving bone marrow only from C57BL/6J donors do not develop GVHD, while recipient mice receiving bone marrow and allogeneic T cells from MHC-mismatched donor develop lethal GVHD. **B**, *Enterococcus* relative abundance quantified by 16S rRNA sequencing of fecal samples. **C**, Survival of GVHD mice, stratified by *Enterococcus* relative abundance on day +7 post-BMT. All data represent at least two independent experiments. Means ± SEM are plotted. ***p < 0.001, ****p < 0.0001 by Kruskal-Wallis test (**B**) or Cox-proportional hazard ratio (**C**).

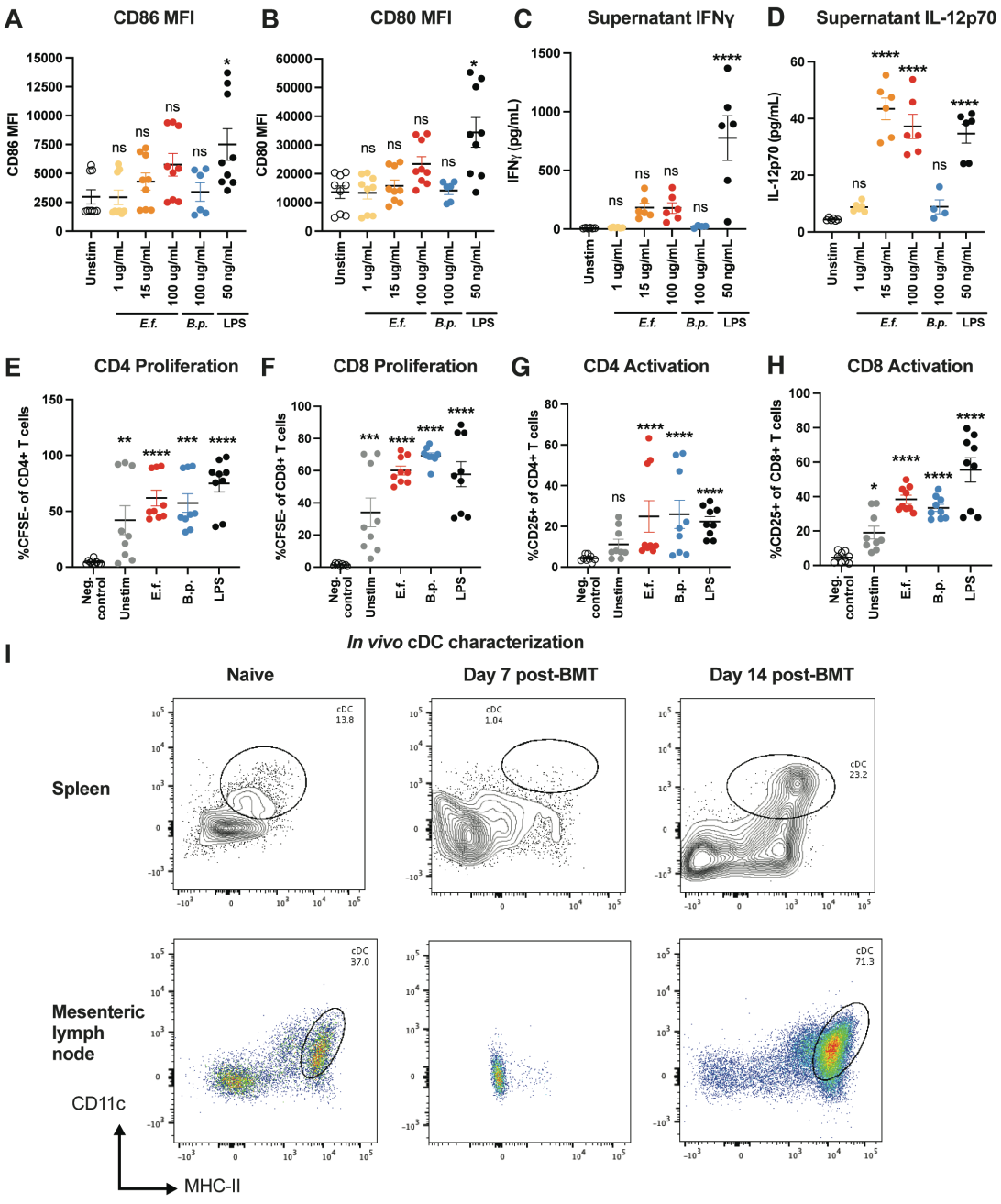

Figure 3.2. ***E. faecalis*** **crude protein lysates moderately activate DCs but did not affect allogeneic T cell activation and proliferation** ***in vitro*****. A-D**, BMDCs were cultured with various stimuli and collected 24-hour post-stimulation for flow cytometry of (**A**) CD86 and (**B**) CD80 expression on CD11c+ MHC-II+ cDCs, and ELISA of culture supernatant for (**C**) IFNγ and (**D**) IL-12p70. **E-H**, Mixed lymphocyte reaction between stimulus-activated BMDCs (BALB/cJ-derived) and allogeneic T cells (C56BL/6J-derived). On day +5 post-co-culture, allogeneic T cells were collected for flow cytometry. **I**, Reconstitution of donor cDCs in the spleen and mesenteric lymph nodes on day +7 and day +14 post-BMT. All data represent at least two independent experiments. Means ± SEM are plotted. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001 by one-way ANOVA test.

Given the variation in the level of endogenous *Enterococcus* bloom (**Figure 3.1B**), we asked whether *Enterococcus* relative abundance was correlated with GVHD severity. We stratified GVHD mice into two groups, *Enterococcus*-high and *Enterococcus*-low, based on the median *Enterococcus* relative abundance on day +7 post-BMT. We observed a significant and positive correlation between *Enterococcus* relative abundance and GVHD mortality, as the *Enterococcus*-high group had increased mortality compared to the *Enterococcus*-low group (**Figure 3.1C**). This observation in a mouse model of GVHD recapitulates previous studies in allo-HCT patients, where *Enterococcus* relative abundance in the peri-transplant period (day 0 to +21) is associated with increased GVHD-related mortality (Holler *et al*., 2014; Peled *et al*., 2020; Stein-Thoeringer *et al*., 2019).

Enterococcus abundance is positively correlated with MHC-II expression by colonic IECs during GVHD

We next investigated the mechanism underlying the association between *Enterococcus* and GVHD mortality. Professional APCs such as host- and donor-derived DCs are the primary cell populations to engage with and activate donor T cells (Koyama *et al*., 2015; Matte et al., 2004; Shlomchik *et al*., 1999). *Enterococcus* could activate DCs and induce inflammatory cytokine secretion (Molina et al., 2015; Ocvirk *et al*., 2015). To assess the effect of *Enterococcus* on DCs, we cultured bone marrow-derived DCs (BMDCs) in the presence of crude protein lysates from a murine-derived *E. faecalis* strain. We utilized *B. producta* lysates as a comparison group, given the association of this genus with improved GVHD outcomes (Jenq *et al*., 2015; Kim et al., 2019; Rashidi et al., 2022). *E. faecalis* lysates upregulated the expression of costimulatory molecules, CD80 and

CD86, on BMDCs in a dose-dependent manner, compared to unstimulated control, although not statistically significant (**Figures 3.2A-B**). *E. faecalis* lysates also led to a robust and significant increase in IL-12p70 secretion, but not IFNγ, from stimulated BMDCs (**Figures 3.2C-D**). On the other hand, *B. producta* lysates did not induce BMDC activation and inflammatory cytokine secretion.

We next asked whether BMDC activation by *E. faecalis* had an impact on alloreactive T cell in a mixed lymphocyte reaction. *E. faecalis*-stimulated BMDCs did not significantly activate alloreactive T cells, compared to *B. producta*-stimulated and LPS-stimulated BMDCs (**Figures 3.2E-H**). In addition to the *in vitro* experiments, we also examined DCs *in vivo* during GVHD. On day +7 post-BMT, which is the timepoint that *Enterococcus* bloom occurred in GVHD mice, DCs were mostly absent in the spleen and mesenteric lymph nodes (**Figure 3.2I**). Altogether, our data suggest that DCs are not the primary cell type mediating *E. faecalis* pathogenesis during GVHD.

Enterococcus upregulates MHC-II expression by IECs at steady state and during GVHD

Previous literature has demonstrated that non-professional APCs such as IECs could induce GVHD through allo-antigen presentation via MHC-II (Koyama *et al*., 2019). We observed that GVHD mice exhibited a significant increase in MHC-II expression by colonic IECs on day +7 post-BMT, compared to BM only and non-transplanted naïve controls (**Figure 3.3A**). Furthermore, MHC-II expression by colonic IECs was significantly and positively correlated with fecal *Enterococcus* relative abundance on day +7 post-BMT (**Figure 3.3B**).

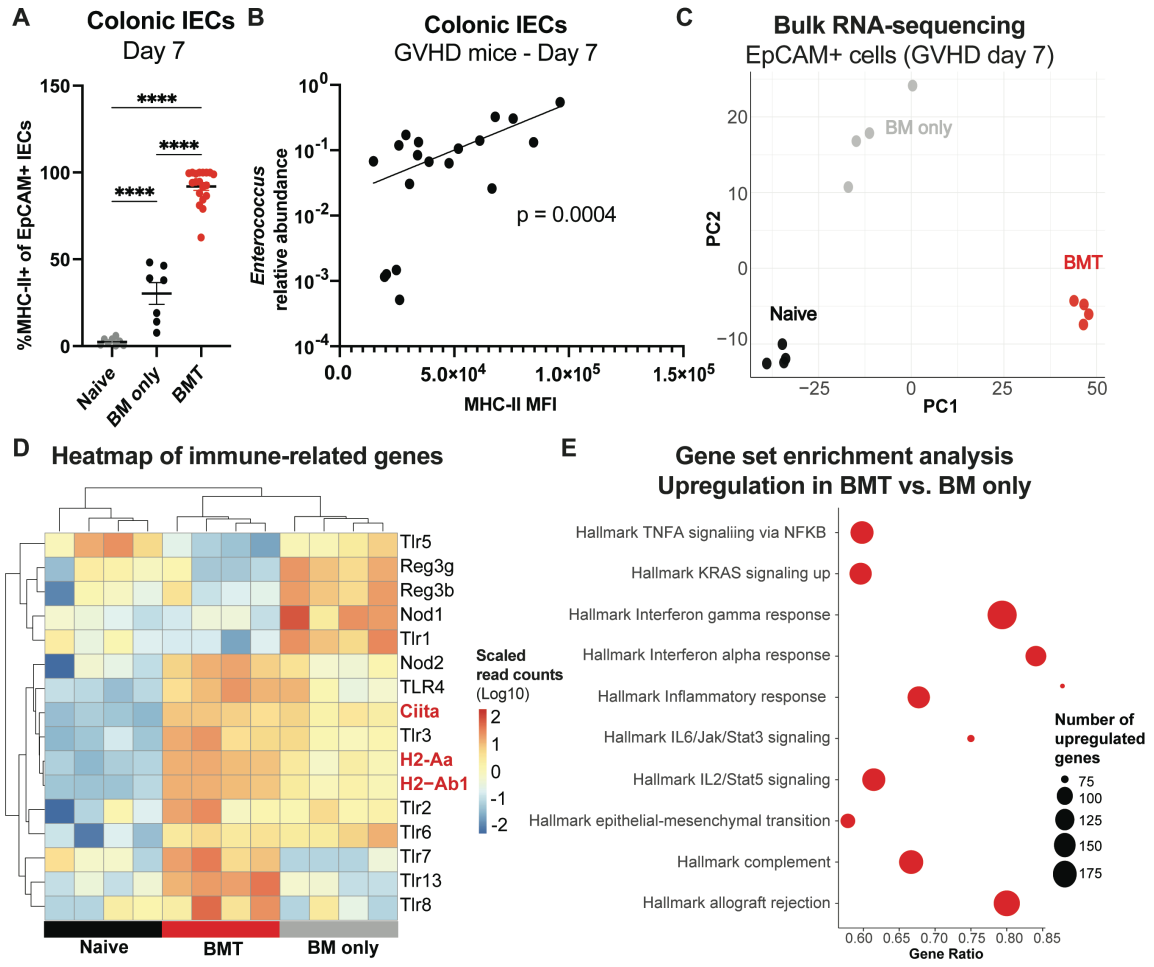

Figure 3.3. *Enterococcus* relative abundance is associated with increased MHC-II expression by colonic IECs in GVHD mice. A, Frequency of MHC-II+ EpCAM+ colonic IECs on day +7 post-BMT by flow cytometry. **B**, Simple linear regression between *Enterococcus* relative abundance and MHC-II expression by colonic IECs on day +7 post-BMT in GVHD mice. **C**, PCA of gene expression through bulk RNA-sequencing of sorted CD45- CD31- EPCAM+ colonic IECs isolated on day +7 post-BMT. **D**, Heatmap of selected genes related to innate sensing and antigen-presentation (in red) from bulk-RNA sequencing of CD45- CD31- EPCAM+ colonic IECs. **E**, Gene set enrichment analysis of upregulated pathways in BMT mice compared to BM only controls from bulk-RNA sequencing of CD45- CD31- EPCAM+ colonic IECs. All data represent at least two independent experiments. Means ± SEM are plotted. ****$p < 0.0001$ by one-way ANOVA test.

We explored other transcriptional changes in the colonic epithelium of GVHD mice through bulk RNA-sequencing of sorted EpCAM+ IECs (**Figures 3.3C-E**). The transcriptomic profiles of colonic epithelium during GVHD clustered distinctly from those

of BM only and non-transplanted naïve controls (**Figure 3.3C**). We confirmed the upregulation of genes related to anti-presentation at the transcriptional level (*H2-Aa, H2-Ab1, Ciita*; **Figure 3.3D**), along with genes associated with IFNα/γ inflammatory response in the colonic epithelium during GVHD (**Figure 3.3E**). Altogether, these data suggests that *Enterococcus* is associated with increased MHC-II expression by colonic IECs during GVHD.

Previous studies have demonstrated that irradiation-induced inflammation is sufficient to upregulate MHC-II expression by IECs at the early time points post-BMT (Koyama *et al*., 2019). To tease out the effect of *Enterococcus* on the colonic epithelium apart from irradiation-induced and GVHD-related inflammation, we monocolonized germfree BALB/cJ mice with *E. faecalis* and profiled the IEC compartment at non-transplanted steady state. As a control group, we utilized a cocktail of four bacteria, CBBP (*C. bolteae*, *B. sartorii*, *B. producta* and *P. distasonis*), which represents a minimal flora of potentially beneficial obligate anaerobes (Kim *et al*., 2019). *Enterococcus*-colonized mice exhibited increased MHC-II expression by colonic IECs compared to germfree and CBBP-colonized mice (**Figure 3.4A**). We did not observe a significant difference in the small IECs among the groups (**Figure 3.4B**).

To validate our observations in the gnotobiotic setting, we utilized a model of *Enterococcus* colonization in the SPF setting on the C57BL/6J background (**Figure 3.4C**). FMT from naïve Jackson Laboratory-sourced donors into antibiotic-treated recipients served as the control group with low *Enterococcus* abundance. *Enterococcus* colonization is sufficient to induce MHC-II expression by both colonic and small IECs compared to the

FMT and CBBP groups, consistent with our findings in the gnotobiotic setting (**Figure 3.4D-E**).

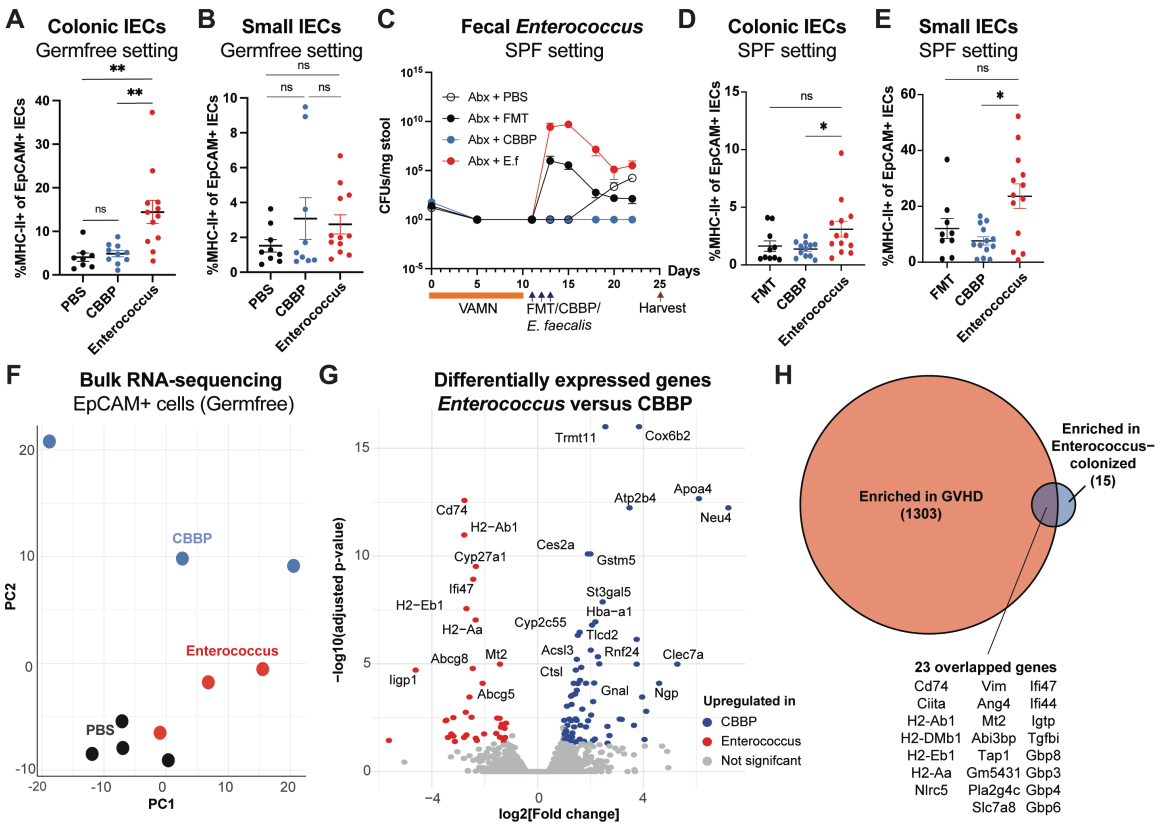

Figure 3.4. *Enterococcus* colonization is sufficient to induce MHC-II expression by colonic IECs in non-transplanted steady state. A-B, Frequency of MHC-II+ EpCAM+ (**A**) colonic and (**B**) small IECs of gnotobiotic mice on day +14 post-colonization by oral gavage. **C**, Fecal *Enterococcus* abundance in SPF mice treated with antibiotics (VAMN) prior to bacteria colonization by oral gavage. Abx: antibiotics. **D-E**, Frequency of MHC-II+ EpCAM+ (**D**) colonic and (**E**) small IECs of SPF mice on day +14 after the first oral gavage for bacterial colonization (day +25 after initiation of antibiotic treatment). **F**, PCA of gene expression through bulk RNA-sequencing of sorted CD45- CD31- EPCAM+ colonic IECs isolated from gnotobiotic mice on day +14 post-colonization by oral gavage. **G**, Volcano plot of differentially expressed genes from bulk RNA-sequencing of sorted CD45- CD31- EPCAM+ colonic IECs. **H**, A set of overlapped genes upregulated in both GVHD mice (SPF setting; day +7 post-BM) and *Enterococcus*-monocolonized mice (gnotobiotic setting; day +14 post-colonization. All data represent at least two independent experiments. Means ± SEM are plotted. ****p < 0.0001 by one-way ANOVA test.

Having demonstrated that *Enterococcus* induced MHC-II expression by colonic IECs in gnotobiotic mice at steady state, we next profiled the overall transcriptomic profiles of sorted EpCAM+ colonic IECs upon *Enterococcus* monocolonization (**Figures 3.4F-H**). The transcriptomes of *Enterococcus*-colonized colonic epithelium clustered distinctly from those of CBBP-colonized group (**Figure 3.4F**). *Enterococcus*-colonized group exhibited upregulation in genes associated with antigen presentation (*H2-Eb1, H2-Aa, H2-Ab1, Cd74*), which confirmed the flow-cytometry-based protein expression results (**Figure 3.4A**), along with genes associated with IFNγ signaling response (*Ifi47, ligp1*; **Figure 3.4G**). Several genes were upregulated in both *Enterococcus*-colonized gnotobiotic mice at non-transplanted steady state and in *Enterococcus*-dominated SPF mice with GVHD (**Figure 3.4H**). These genes are associated with antigen presentation (*Cd74, H2-Ab1, H2-Eb1, H2-DMb1, Ciita, Nlrc5*) and with IFNγ signaling response (*Ifi44, Igtp, Gbp8, Gbp4, Gbp6, Tgfbi, Ifi47*). Thus, our results suggest that *Enterococcus* colonization is sufficient to induce MHC-II expression by IECs in the absence of systemic inflammation.

MHC-II upregulation by colonic IECs is regulated via IFNγ signaling

TLRs recognize the gut microbiota through various bacterial surface-associated and intracellular molecules (Abreu et al., 2005; Fang et al., 2022). TLR signaling results in a cascade of innate immune activation that maintains homeostasis or mediates responses to pathogens (Kubinak and Round, 2012; Rakoff-Nahoum et al., 2004). To identify which TLRs could sense *Enterococcus*, we screened the crude protein lysates of murine-derived *E. faecalis* and human-derived vancomycin-resistant *E. faecium* (VRE) against a panel of murine and human TLRs, respectively. Murine-derived *E. faecalis* activated murine TLR2

and TLR13, while human VRE only activated human TLR2-TLR6 (**Figures 3.5A-B**). This observation is consistent with previous studies reporting the role of TLR2 in host defense against pathogenic *Enterococcus* (Leendertse *et al*., 2008; Park *et al*., 2013).

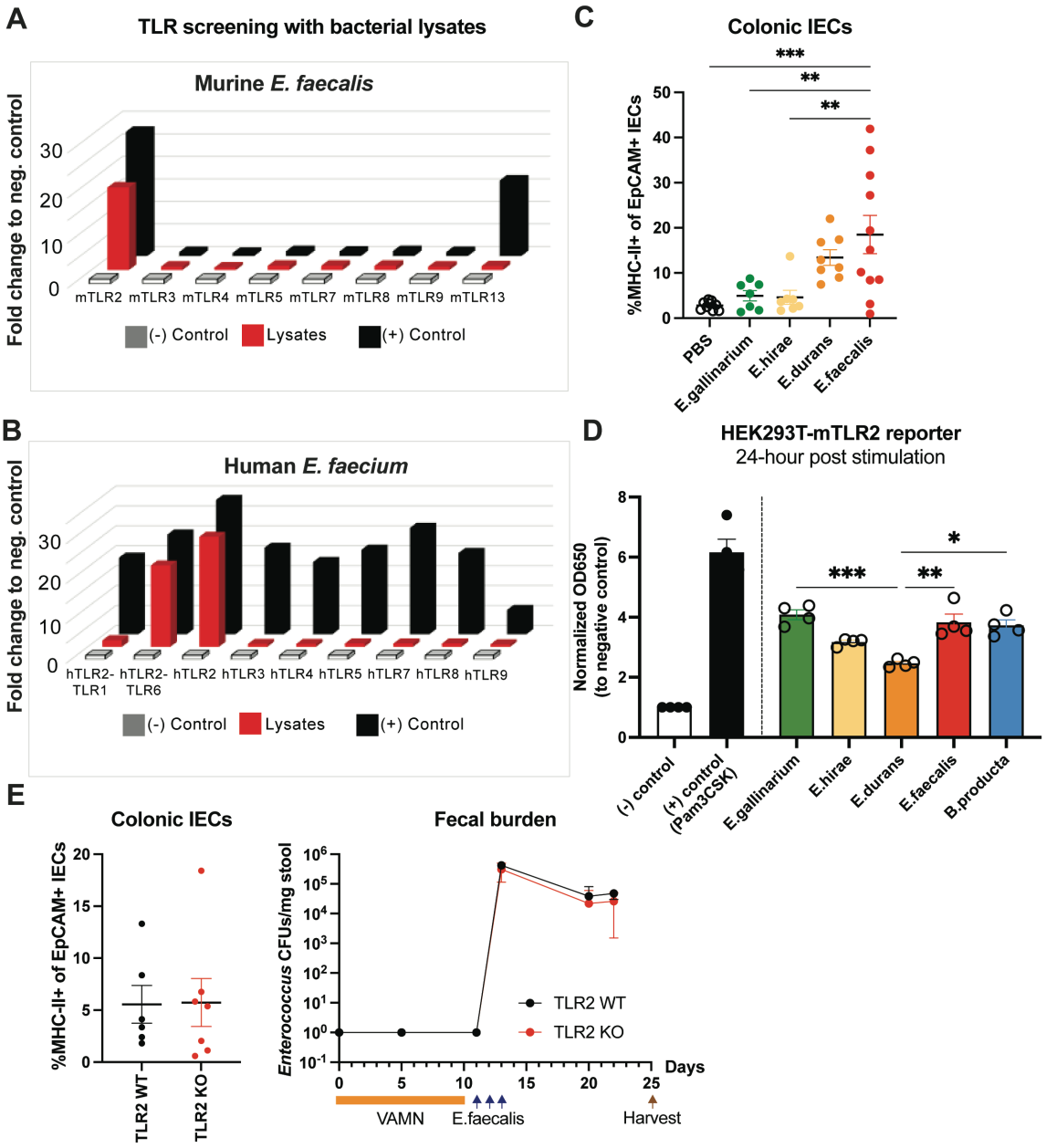

Figure 3.5. *Enterococcus*-induced MHC-II upregulation by colonic IECs is independent of TLR2 signaling. A, TLR activation screening using HEK293 cells expressing various mouse TLR co-cultured with crude protein lysates of murine-derived *E. faecalis*. **B**, TLR activation screening using HEK293 cells expressing various human

TLR co-cultured with crude protein lysates of human-derived *E. faecium*. **C**, Frequency of MHC-II+ EpCAM+ colonic IECs in gnotobiotic mice colonized with different *Enterococcus* strains by flow cytometry. **D**, TLR2 activation screen using HEK293 cells expressing mouse TLR2 co-cultured with crude protein lysates of different *Enterococcus* strains. **E**, Frequency of MHC-II+ EpCAM+ colonic IECs in TLR2 WT versus TLR2 KO mice colonized with *E. faecalis* post-antibiotic decontamination by flow cytometry. **F**, Absolute fecal *Enterococcus* abundance of TLR2 WT and KO mice colonized with *E. faecalis* post-antibiotic decontamination. VAMN: vancomycin, ampicillin, metronidazole, neomycin. All data represent at least two independent experiments. Means ± SEM are plotted. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ by one-way ANOVA test.

Lipoteichoic acid (LTA) is a cell-wall compenent of Gram-posive bacteria that serves as a ligand for TLR2 (Long et al., 2009). Given that *Enterococcus* species share similar LTA structures, we investigated whether *Enterococcus* species are similarly immunogenic in the context of the gut epithelium. We monocolonized germfree BALB/cJ mice with various different *Enterococcus* species and profiled MHC-II expression by colonic IECs at steady state. In this screen, *E. faecalis* was the only species that significanly upregulated MHC-II expression by colonic IECs, compared to germfree controls (**Figure 3.5C**). *E. durans* also induced MHC-II expression moderately, although not statistically signficantly. On the other hand, *E. gallinarium* and *E. hirae* did not affect MHC-II expression by the colonic epithelium. This species-specific phenotype *in vivo* did not correlate with TLR2 activation *in vitro*, given that MHC-II non-inducers such as *B. producta* and *E. gallinarium* could strongly induce TLR2 activation to a comparable level with *E. faecalis* (**Figure 3.5D**).

To further assess whether TLR2 signaling regulates *Enterococcus*-induced MHC-II expression in the colonic epithelium, we utilized TLR2-deficient mice in the SPF setting. While high *Enterococcus* colonization was achieved in both TLR2-competent and deficient mice, we observed no significant difference in MHC-II expression by colonic IECs in wild-

type versus TLR2-deficient mice (**Figure 3.5E**). Altogether, these data suggest that *Enterococcus* induces MHC-II expression in the colonic epithelium in a TLR2-independent manner.

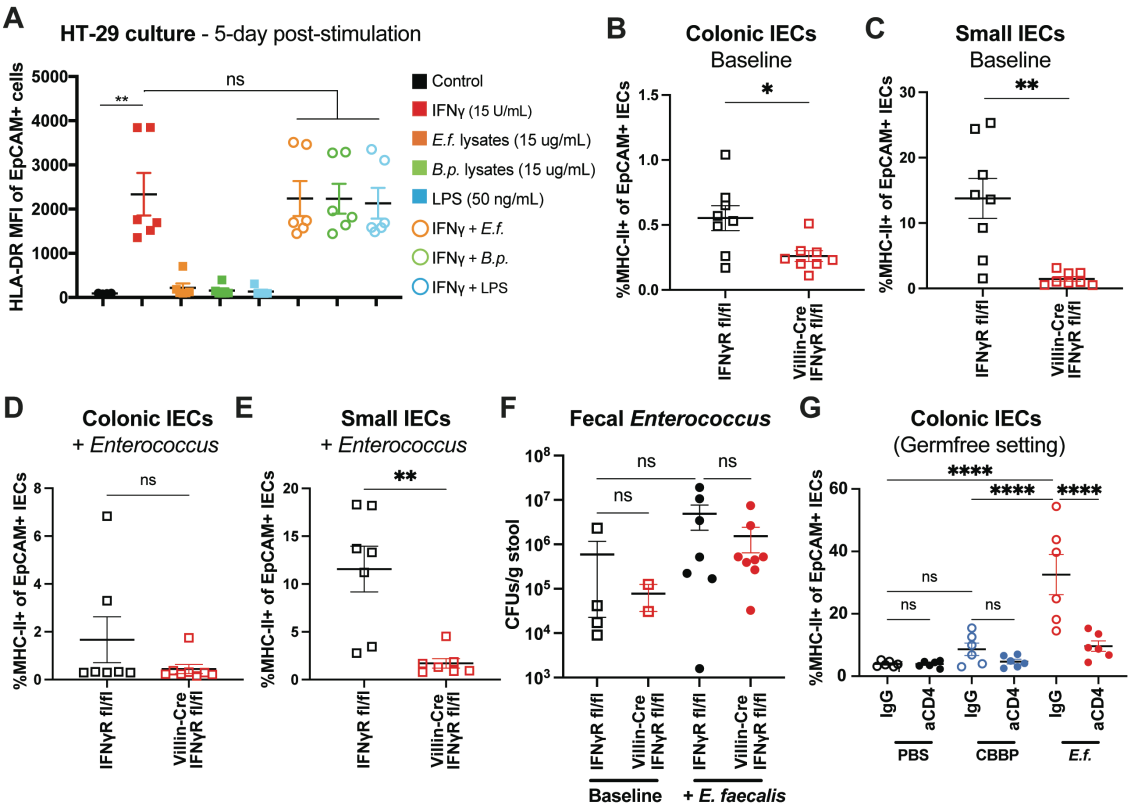

Figure 3.6. IFNγ signaling mediates *Enterococcus*-induced MHC-II upregulation by colonic IECs. A, HLA-DR expression of HT-29 cells after 5 days of stimulation with IFNγ and bacterial protein lysates. **B-C**, Frequency of MHC-II+ EpCAM+ in (**B**) colonic and (**C**) small IECs in Villin^{Cre+} IFNγR$^{fl/fl}$ mice at steady state. **D-E**, Frequency of MHC-II+ EpCAM+ in (**D**) colonic and (**E**) small IECs in mice with epithelial deficiency of IFNγR upon colonization with *E. faecalis* by oral gavage for 5 days. **F**, Fecal absolute *Enterococcus* abundance of mice with epithelial deficiency of IFNγR at steady state and upon colonization with *E. faecalis*. **G**, Frequency of MHC-II+ EpCAM+ colonic IECs in gnotobiotic mice 14-day post-colonization with *B. producta* or *E. faecalis* in addition to treatment with isotype control or anti-CD4 depletion antibody. All data represent at least two independent experiments. Means ± SEM are plotted. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001 by one-way ANOVA test (**A**, **F-G**) or Student's t-test (**B-E**).

Previous literature has established the role of IFNγ signaling in the regulation of MHC-II expression by non-professional APCs (Buttice et al., 2006; Koyama *et al.*, 2019;

Nikcevich et al., 1999). We also observed increased presence of IFNγ signaling upon *Enterococcus* colonization both at non-transplanted steady state and during GVHD. Thus, we next asked whether IFNγ signaling mediates the interaction between *Enterococcus* and the intestinal epithelium. We co-cultured HT-29 colorectal adenocarcinoma cells with either IFNγ alone or in combination with bacterial protein lysates. IFNγ treatment alone was sufficient to significantly upregulate MHC-II expression by HT-29 cells *in vitro* (**Figure 3.6A**). The addition of bacterial lysates from either *E. faecalis* or *B. producta* did not affect MHC-II expression in HT-29 cells compared to IFNγ treatment alone, suggesting an indirect interaction between bacteria and IECs potentially mediated by IFNγ signaling.

To further delineate the role of IFNγ signaling in regulating the interactions between *Enterococcus* and IECs, we utilized Villin^{Cre+} IFNγR$^{fl/fl}$ mice with specific deletion of IFNγR in Villin-expressing IECs. Villin^{Cre+} IFNγR$^{fl/fl}$ mice exhibited significantly decreased MHC-II expression by small IECs, compared to Villin^{Cre-} IFNγR$^{fl/fl}$ littermate controls, both at baseline microbiota and upon colonization with *Enterococcus* (**Figures 3.6B-E**). The intestinal microbiome profiles based on 16S rRNA sequencing of fecal pellets collected at the time of tissue isolation demonstrated comparable levels of *Enterococcus* relative abundance between Villin^{Cre+} and Villin^{Cre-} littermates (**Figure 3.6F**). Overall, these data suggest that IFNγ signaling is important for the induction of MHC-II expression by IECs both at steady state and during *Enterococcus* colonization.

While many immune cell types could secrete IFNγ, conventional CD4+ T cells are the most dominant source of this inflammatory cytokine in the gastrointestinal tract (Koyama *et al*., 2019). To test whether CD4+ T cells mediate the effect of *Enterococcus* on the intestinal epithelium, we administered antibodies depleting CD4+ T cells to

gnotobiotic mice monocolonized with *E. faecalis* or CBBP. Depletion of CD4+ T cells significantly abrogated the upregulation of MHC-II expression by colonic IECs in the presence of *E. faecalis* monocolonization (**Figure 3.6G**). These data support the hypothesis that IFNγ signaling from CD4+ T cells mediates the induction of antigen presentation in the intestinal epithelium by *Enterococcus*.

Butyrate does not regulate antigen presentation by the colonic epithelium in the context of Enterococcus domination

This study and others have demonstrated that only specific members of the intestinal microbiota could induce MHC-II expression by IECs (Koyama et al., 2023; Tuganbaev *et al*., 2020). Thus, we next investigated the microbial-derived factors driving this phenotype. Histone deacetylases (HDAC) such as HDAC3 play an important role in maintaining intestinal homeostasis and promoting barrier functions during GVHD (Alenghat et al., 2013; Mathewson *et al*., 2016; Reddy et al., 2004). Deletion of HDAC3 specifically in IECs significantly abrogates MHC-II expression in this cellular compartment, suggesting that role of HDAC3 in regulating MHC-II expression (Eshleman et al., 2023). The intestinal microbiota could modulate HDAC3 activity through the secretion of butyrate, a short-chain fatty acid that acts as a pan-HDAC inhibitor (Candido et al., 1978). Thus, we next assessed whether butyrate could suppress MHC-II expression by IECs. We induced MHC-II expression in HT-29 cells through IFNγ treatment in the presence of varying concentrations of sodium butyrate. Butyrate concentration as low as 5mM was sufficient to completely abrogate MHC-II expression by HT-29 cells, even when combined with a high dose of IFNγ (**Figure 3.7A**).

We next assessed the effect of butyrate *in vivo* using our gnotobiotic mouse model. Since orally supplemented butyrate is rapidly absorbed in the duodenum and therefore does not reach the colon, we administered tributyrin, a precursor of butyrate that is slowly converted to butyrate in the colon by pancreatic lipases, to gnotobiotic mice monocolonized with *E. faecalis* or CBBP. We observed a trend, although not statically significant, that tributyrin-treated mice exhibited decreased MHC-II expression by colonic IECs in the presence of *E. faecalis* monocolonization, compared to vehicle-treated group (**Figure 3.7B**). Thus, butyrate could suppress MHC-II expression *in vitro*, but its potential *in vivo* effects require further investigation.

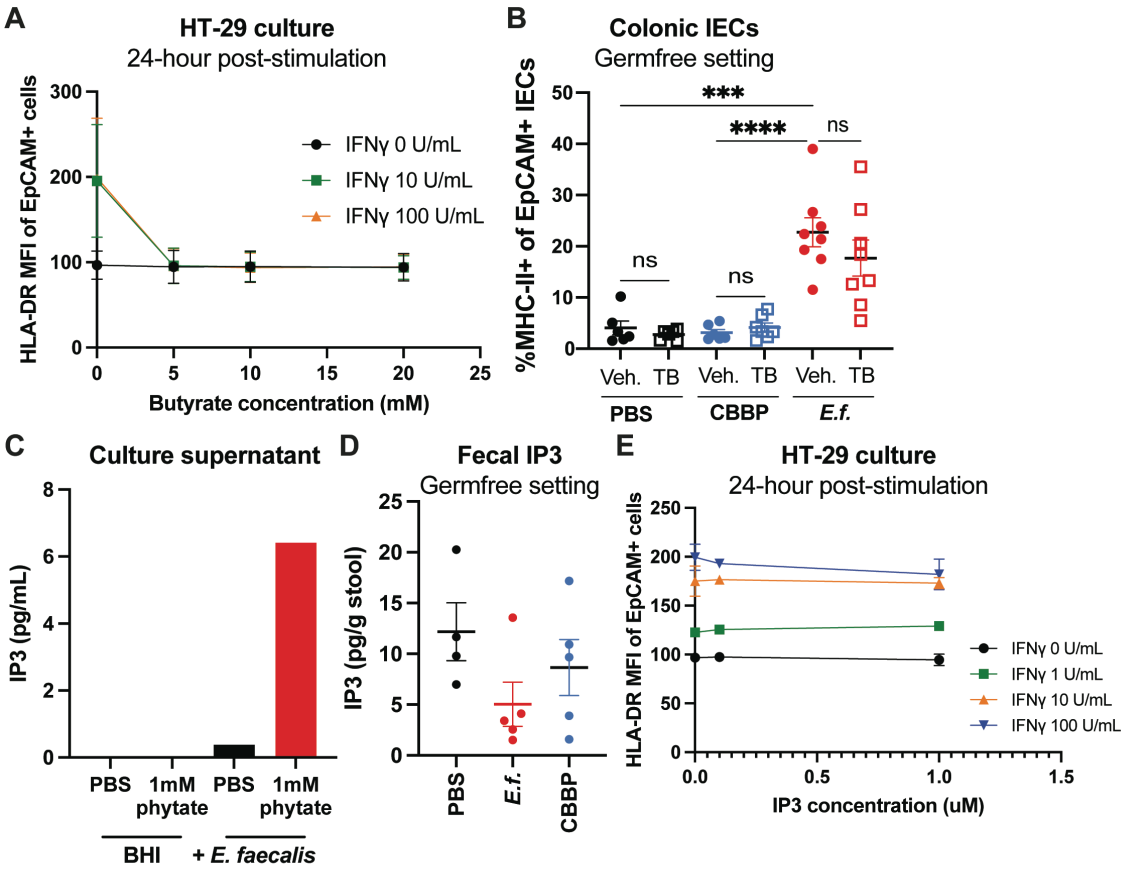

Figure 3.7. Butyrate suppresses MHC-II expression by colonic IECs *in vitro* but not *in vivo*. A, HLA-DR expression of HT-29 cells after 24 hours of stimulation with varying

concentrations of IFNγ and butyrate. **B**, Frequency of MHC-II+ EpCAM+ colonic IECs in gnotobiotic mice colonized with *B. producta* or *Enterococcus* strains upon treatment with vehicle control (glycerol) or tributyrin (TB). **C**, IP3 concentration in supernatant of *E. faecalis* culture in BHI with and without addition of phytate. **D**, Fecal IP3 concentration of gnotobiotic mice colonized with *B. producta* or *E. faecalis*. **E**, HLA-DR expression of HT-29 cells after 24 hours of stimulation with varying concentrations of IFNγ and IP3. All data represent at least two independent experiments. Means ± SEM are plotted. ***p < 0.001, ****p < 0.0001 by one-way ANOVA test (**B**, **D**) or two-way ANOVA test (**A**, **E**).

The intestinal microbiota could also promote HDAC activity in the gut through the microbial-derived metabolite inositol-1,4,5-triphosphate (IP3), a byproduct of dietary phytate metabolism (Wu et al., 2020). We hypothesized that *E. faecalis* stimulates MHC-II expression by IECs through IP3 production. We confirmed the capability of *E. faecalis* to metabolize dietary phytate, as demonstrated by the detection of IP3 in the culture supernatant upon addition of phytate to *E. faecalis* culture (**Figure 3.7C**). However, we did not detect a significant difference in IP3 production by *E. faecalis in vivo*, compared to MHC-II non-inducers such as CBBP (**Figure 3.7D**). In addition, treatment of HT-29 cells with IP3, with or without IFNγ, did not affect MHC-II expression (**Figure 3.7E**). Altogether, these data do not support the hypothesis that *E. faecalis* induces MHC-II expression by IECs through modulation of HDAC activity.

Lantibiotic-producing B. producta improves GHVD survival through suppression of Enterococcus bloom

Having demonstrated CBBP as MHC-II non-inducers, we next assessed whether the CBBP cocktail has a potential therapeutic effect in the setting of GVHD. We colonized GVHD mice with either *E. faecalis* and CBBP on day +4 and +5 post-transplant and observed that CBBP colonization significantly improved survival compared to *E. faecalis* colonization (**Figure 3.8A**). Colonization with CBBP also led to a reduction in fecal

Enterococcus burden on day +7 post-BMT, compared to PBS-treated and *E. faecalis*-colonized group (**Figure 3.8B**).

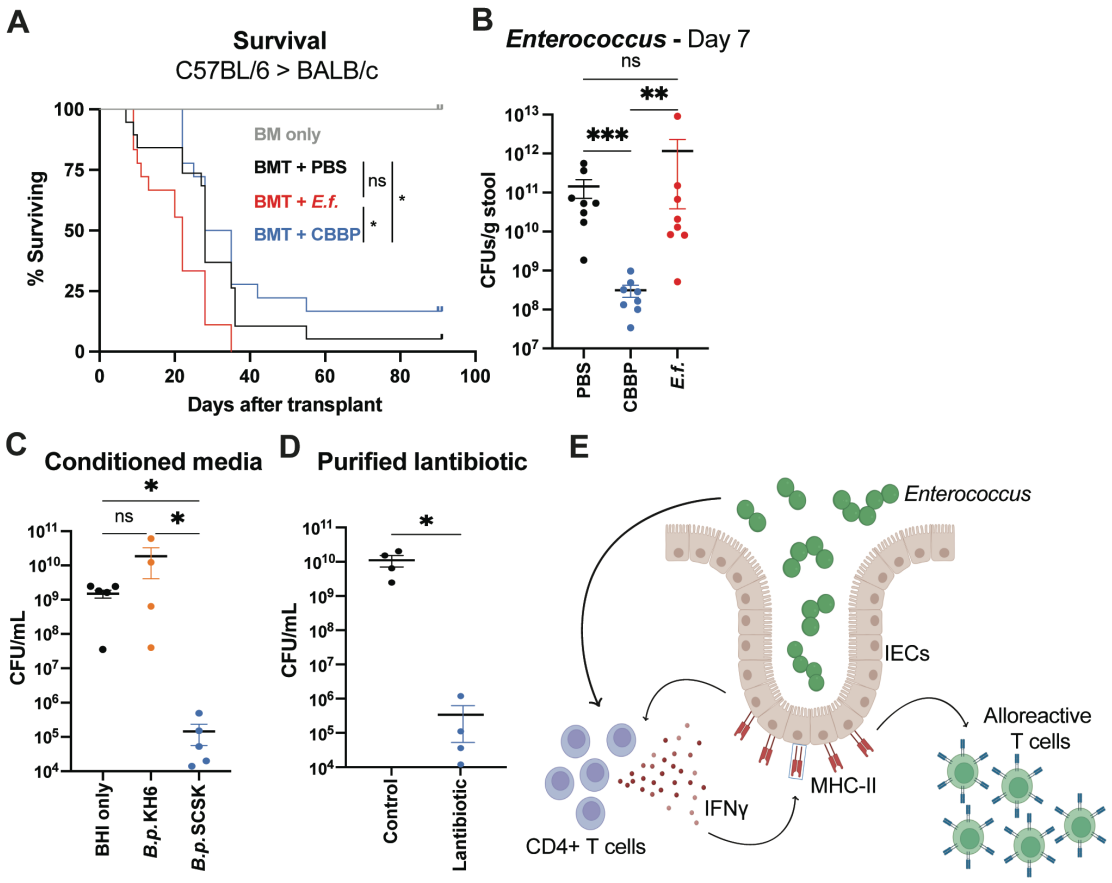

Figure 3.8. Colonization resistance against *Enterococcus* by a *B. producta* strain improves GVHD survival. A, Survival of GVHD mice colonized with lantibiotic-producing *B. producta* or *E. faecalis*. **B**, Fecal *Enterococcus* absolute abundance on day +7 post-BMT. **C**, *Enterococcus* absolute abundance in co-culture with conditioned media from the lantibiotic-producing *B. producta* strain (*B.p.* SCSK) and a non-producer control strain (*B.p.* KH6). **D**, *Enterococcus* absolute abundance in co-culture with purified lantibiotic from the culture supernant of *B.p.* SCSK compared to vehicle control (PBS). **E**, Schematic of the interaction between *Enterococcus* and the intestinal epithelium through CD4+ T cells and IFNγ signaling pathway, leading to increased allo-antigen presentation by IECs that induces activation of alloreactive T cells and subsequent development of GVHD. All data represent at least two independent experiments. Means ± SEM are plotted. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001 by Cox-proportional hazard ratio (**A**), Kruskal-Wallis test (**B-C**) or Mann-Whitney U-test (**D**).

The CBBP cocktail contains a strain of *B. producta*, termed *B.p.* SCSK, which could suppress VRE growth through lantibiotic production (Kim *et al.*, 2019). We next assessed whether *B.p.* SCSK strain could also suppress commensal *E. faecalis* growth. Treatment with conditioned media from *B.p.* SCSK culture strongly suppressed *E. faecalis* growth *in vitro*, compared to untreated control and treatment with conditioned media from the control strain *B.p.* KH6 that does not produce lantibiotic (**Figure 3.8C**). In addition, treatment with purified lantibiotic also led to a strong reduction in *E. faecalis* growth *in vitro* (**Figure 3.8D**). Altogether, these data support the potential therapeutic effect of CBBP in ameliorating GVHD severity through suppression of *Enterococcus* growth.

Discussion

In this study, we described the role of the intestinal microbiota in the regulation of MHC-II expression by IECs (Koyama *et al.*, 2023; Tuganbaev *et al.*, 2020). Specifically, we demonstrated that *E. faecalis* could strongly upregulate MHC-II expression by IECs both at steady state and during GVHD. Conversely, other strains of *Enterococcus* and of specific obligate anaerobes such as *B. producta* did not induce MHC-II expression. Importantly, *Enterococcus* domination was also strongly correlated with increased MHC-II expression by colonic IECs during GVHD, further highlighting the importance of antigen presentation by non-professional APCs to initiate the GVHD inflammatory cascade (Koyama *et al.*, 2011; Koyama *et al.*, 2019). We also reported that *Enterococcus* induction of MHC-II expression by IECs was mediated by IFNγ signaling. Antigen presentation by non-professional APCs, as regulated by the intestinal microbiota, also play a crucial role in other disease settings (Beyaz et al., 2021; Kreisel et al., 2010). Thus, targeting the

microbiome-gut epithelium MHC class II- IFNγ axis presents a window of opportunity across many therapeutic areas (**Figure 3.8E**).

In addition, we identified the potentially therapeutic effect of *B. producta* SCSK strain to ameliorate GVHD severity and improve survival. *B. producta* SCSK strain produces lantibiotic that suppresses both commensal and pathogenic *Enterococcus* growth *in vitro* and *in vivo* (Kim *et al*., 2019). During allo-HCT, disruptions to the gut microbial community could lead to pathogen outgrowth and subsequent enteric and systemic infections, which could be prevented or reversed through restoring colonization resistance by beneficial gut commensals (Buffie and Pamer, 2013; Khan et al., 2021; Taur *et al*., 2012). Here, we demonstrated that colonization resistance against *Enterococcus* could also ameliorate GVHD severity. We envision the use of *B. producta* SCSK strain as probiotics prior to allo-HCT to prevent *Enterococcus* bloom or as a treatment in response to intestinal *Enteroocccus* colonization to prevent the risk of GVHD-related mortality.

In conclusion, we reported a mechanism by which *Enterococcus* aggravates GVHD through the upregulation of MHC-II expression by IECs. Future preclinical and clinical studies may further delineate the relevance of these observations in allo-HCT patients. Despite its pathogenic impact in the context of GVHD, enhanced antigen presentation by the gut epithelium could be beneficial in other disease settings such as enteric infections and gastrointestinal cancer (Alenghat *et al*., 2013; Beyaz *et al*., 2021; Heuberger *et al*., 2023; Kreisel *et al*., 2010; Tuganbaev *et al*., 2020). *Enterococcus* could also confer benefits in boosting immune responses during infection and cancer immunotherapy (Griffin *et al*., 2021; Rashid *et al*., 2023). Additional mechanistic insights into the microbial-derived

factors driving the interactions between *Enterococcus* with the gut epithelium, therefore, may allow more precise utilization of this bacteria in a clinical setting.

Materials and Methods

Mice

Female C57BL/6J and BALB/cJ mice were purchased from the Jackson Laboratory and maintained in our SPF facility. Female germfree BALB/cJ mice were born and maintained in flexible isolators at the Weill Cornell Gnotobiotic Mouse Facility, fed autoclaved feed and water, and routinely monitored for sterility. Villin-Cre (B6.Cg-Tg(Vil1-cre)1000Gum/J), TLR2 KO (B6.129-Tlr2tm1Kir/J) and IFNγR fl/fl (C57BL/6N-Ifngr1tm1.1Rds/J) mice were purchased from the Jackson Laboratory. In all experiments, sex- and age-matched littermate controls were used. Mice were 6-12 weeks of age at the beginning of experiments. All experiments procedures were conducted in compliance with the institutional guidelines at MSKCC.

Bone marrow transplantation and assessment of GVHD

Bone marrow transplantation were performed as previously described (Shono et al., 2014; Tsai et al., 2018). Briefly, BALB/cJ recipients received 900-cGy split-dosed lethal irradiation and were transplanted with T-cell-depleted bone marrow (10 x 10^6 cells) and enriched T cells (1 x 10^6 cells) from donor C57BL/6J mice. T cells were depleted from the bone marrow with anti-Thy-1.2 and Low-Tox-M rabbit complement (CEDARLANE Laboratories). Donor T cells were prepared from splenocytes, enriched for CD5+ T cells

by Miltenyi MACS purification kit (routinely >90% purity). Mice were monitored daily for survival and weekly for GVHD clinical scores (Cooke et al., 1996).

Bacterial species

The *E. faecalis* strain used in this study was isolated from a BALB/cJ mouse with GVHD on day +7 post-transplant by plating homogenized fecal pellets in sterile PBS on Enterococcosel agar medium (Becton Dickinson). The following commercial bacteria species and strains were sourced as follows: *E. durans* 23C2 (ATCC 6056), *E. hirae* R (ATCC 8043) and *E. gallinarum* NCDO 2313 (ATCC 51559). All *Enterococcus* strains and species were grown at 37°C under ambient atmosphere in autoclaved Gibco Bacto Brain Heart Infusion medium (Fisher Scientific).

C. boltea, B. producta SCSK, *B.producta* KH6, *B. sartorii* and *P. distasonis* strains were kindly provided by Eric Pamer (University of Chicago). These strains were cultured in pre-reduced and autoclaved Brain Heart Infusion medium supplemented with yeast extract (5 g/L) and L-cysteine (1 g/L) at 37°C under anaerobic atmosphere.

For colony-forming unit (CFU) assay, fecal samples were sterilely collected, weighed, resuspended in sterile PBS, homogenized by douncing with sterile pestles, serially diluted in sterile PBS, and then plated by drip assay onto selective Enterococcosel agar plates.

Bacterial lysates were prepared by growing bacteria until log phase overnight. Pelleted bacteria were washed twice with sterile PBS, resuspended in sterile PBS supplemented with 250 μL/mL of MD solution. MD solution contained 0.1M $MgCl_2$ and 100 μg/mL DNase I in sterile water. Autoclaved zirconium beads (0.1mm size, 400

mg/tube) were added to the bacterial suspension, and bacteria was disrupted using a Bead beater (Fast-Prep24; program: 6.5 M/S; 30 seconds) for three times, with rest on ice in between each repeat for one minute. After the last bead beating cycle, samples were rested on ice for 5 min, centrifuged at 7500g for 5 min. Supernatant was collected, sterile filtered and stored in -20°C. Protein concentration was quantified by Bradford assay (Thermo Scientific).

Cell culture

The following cell lines were used and sourced as follows: HT-29 (ATCC HTB-38) and HEK293-mTLR2 (Invivogen hkb-mtlr2). HT-29 cells were cultured at 37°C and 5% CO_2 in complete McCoy's 5A (ATCC 30-2007) supplemented with 10% fetal bovine serum, 100 U/mL penicillin and 100 μg/mL streptomycin. HEK293-mTLR2 cells were cultured at 37°C and 5% CO_2 in complete DMEM (ThermoFisher) supplemented with 10% fetal bovine serum, 100 U/mL penicillin, 100 μg/mL streptomycin and 100 μg/mL Normocin (Invivogen). Cells were maintained at no greater than 80% confluency and subcultured accordingly using Trypsin. Cells were routinely tested for mycoplasma.

For experiments using bacterial protein lysates, HT-29 cells were cultured in complete media supplemented with recombinant human IFNγ (15 U/mL; Peprotech), along with bacterial crude protein lysates (15 μg/mL) or LPS (50 ng/mL; ThermoFisher) for 5 days before collecting the cells for flow cytometry analysis. For experiments using butyrate and IP3, HT-29 cells were cultured in complete media supplemented with varying concentrations of recombinant human IFNγ and sodium butyrate (Sigma) or D-myo-

Inositol 1, 4, 5-triphosphate trisodium salt (Sigma) for 24 hours before collecting the cells for flow cytometry analysis.

Bone marrow-derived dendritic cell culture and mixed lymphocyte reaction

Bone marrow cells were isolated from wildtype BALB/cJ mice, cultured in complete RPMI supplemented with 10% fetal bovine serum, 100 U/mL penicillin and 100 μg/mL streptomycin, along with recombinant GM-CSF (20ng/mL; Peprotech) and IL-4 (5 ng/mL; Preprotech) for 6 days. On day 6, BMDCs were matured with various concentrations of bacterial lysates or with LPS (50 ng/mL; Sigma) for 24 hours. BMDCs were collected by scraping and used for flow cytometry analysis of activation markers or subsequent MLR. Culture supernatant was collected for cytokine analysis with ProcartaPlex Multiplex Immunoassay per the manufacturer's instructions (Affymetrix). Cytokine concentration results were acquired with a Luminex 200 instrument and analyzed with xPONENT software (Luminex Corporation).

For subsequent MLR, activated BMDCs were used as stimulators. Splenic T cells from wildtype C57BL/6J mice were isolated using Miltenyi MACS purification of Pan T cells (routinely >95% purity). Isolated T cells were labeled with carboxyfluorescein succinimidyl ester (CFSE) to measure proliferation, as indicated by the percentage of $CFSE^{low/medium}$ cells by flow cytometry analysis. 100,000 T cells were mixed with 10,000 stimulated BMDCs and co-cultured in a 96-well U-bottom plate for 96 hours.

Bacterial administration

In the germfree setting, bacteria were administered through a single dose of oral gavage. On the day of administration, frozen bacterial stocks were thawed, pelleted by centrifugation at 4,000g for 10 min and resuspended in sterile PBS (3 x 10^8 CFUs/mouse). Mice were euthanized on day +14 relative to colonization day for profiling of IECs by flow cytometry, unless otherwise indicated.

For GVHD experiments in the SPF setting, BALB/cJ recipients were provided with sterile-filtered drinking water supplemented with ampicillin (0.5 g/L; Sigma) and enrofloxacin (0.25 g/L; Sigma) as previously described (Staffas et al., 2018). Antibiotic treatment started at day -2 pre-BMT and continued until day 3 post-BMT until antibiotic solutions were replaced with regular drinking water. On day 4 and 5, bacteria were administered via oral gavage (3 x 10^8 CFUs/mouse).

For steady state experiments in the SPF setting, mice were provided with sterile-filtered drinking water supplemented with ampicillin (1 g/L; Sigma), vancomycin (0.5 g/L; Sigma), metronidazole (1 g/L; Sigma), neomycin (1 g/L; Sigma) and glucose (1 g/L; Sigma) for 10 days. Bacteria were administered via oral gavage for the subsequent 3 consecutive days. On the day prior to administration, bacteria were inoculated into growth medium and grown overnight to late logarithmic phase. The following day, overnight cultures were pelleted by centrifugation and resuspended in sterile PBS (8 x 10^{10} CFUs/day/mouse). FMT was carried out as fast as possible to avoid prolonged aerobic exposure. Microbiome samples were collected from the cecal content of a healthy, naïve C57BL/6 into a Falcon tube containing 1 mL sterile PBS. Samples were homogenized,

filtered through a 100 μM strainer and immediately administered to recipient mice pre-treated with antibiotics.

Cell isolation and flow cytometry

Single cell suspensions from intestinal tissues were prepared as previously described (Stein-Thoeringer *et al*., 2019). Briefly, intestinal tissue was excised from mice, thoroughly rinsed with ice-cold PBS to clean out luminal content and opened longitudinally. The tissue was then cut into pieces of 1cm length and incubated in PBS containing 1mM EDTA and 1mM DTT (Teknova) at 37°C for 30 min while shaking at 250 rpm. Epithelial cells were collected, filtered, centrifuged, and subsequently incubated with staining buffer (PBS, 0.5% BSA, 2mM EDTA) containing anti-mouse CD16/32 blocking agent (1:400; BD Pharmingen 553142) for 15 min on ice. Samples were protected from light from this point forward. Samples were washed with staining buffer and then incubated with antibody cocktail for 30 min on ice. Antibodies were procured and used as follows: anti-CD45 (1:800; BV711, 30-F11, BioLegend 103147), anti-CD31 (1:200; BV605, 390, BioLegend 102427), anti-CD326 (1:200; PE-Cy7, G8.8, eBioscience 25-5791-80) and anti-I/A-I/E (1:200; APC, M5/114.15.2, eBioscience 17-5321-82). After staining, cells were washed with staining buffer and resuspended in staining buffer containing DAPI (1:200). Multiparameter analysis was performed on a FACS Symphony A5 (BD Biosciences) flow cytometer and analyzed with FlowJo software (Tree Star). When applicable, gating was determined using isotype control.

FACS sorting experiment followed a similar protocol as described above. Antibodies were procured and used as follows: anti-CD45 (1:1600; FITC), anti-CD31

(1:200;) and anti-CD326 (1:200; PE-Cy7, G8.8, eBioscience 25-5791-80). Samples were sorted on SH800S Cell Sorter (SONY) to >98% purity.

DNA extraction and 16S rRNA sequencing

DNA was extracted from mouse fecal samples using a phenol-chloroform bead beating protocol and genomic 16S ribosomal RNA gene V4-V5 variable region was amplified and sequenced on the Illumina MiSeq platform as previously described (Jenq *et al*., 2015). Amplicon sequence variants (ASVs) were inferred following the DADA2 pipeline (Callahan *et al*., 2016) and classified to the genus level against the NCBI database.

RNA sequencing and analysis

CD45- CD31- EpCAM+ intestinal epithelial cells from the colon of GVHD mice on day +7 post-BMT and from gnotobiotic mice on day +14 post-colonization were isolated and FACS-sorted. RNA was isolated using TRIzol (Invitrogen) and phase separation as induced with chloroform. RNA was extraction from the aqueous phase using the miRNeasy Micro or Mini Kit (Qiagen) on the QIAcube Connect (Qiagen) according to the maufacturer's instructions. Smaple was eluted in RNase-free water.

After RiboGreen quantification and quality control by Agilent BioAnalyzer, total RNA was amplified using TruSeq Stranded mRNA LT kit (Illumna) according to the manufacturer's instructions, with 8 cycles of PCR. Samples were barcoded and run on the NovaSeq 6000 in a PE100 run using the NovaSeq 6000 S4 Reagent Kit (200 Cycles; Illumina).

TLR stimulation test

TLR stimulation is tested by assessing NF-kB activation in the TLR expressing cell lines (InvivoGen). The activity of the test articles is tested on nine different human TLRs (TLR1/2, 2/6, 2, 3, 4, 5, 7, 8 and 9) and eight mouse TLRs (mTLR2, 3, 4, 5, 7, 8, 9 and 13) as potential agonists. HEK-Blue hTLR2 cells have been stably transfected with human TLR2 and CD14. In HEK-Blue hTLR2-TLR1 cells endogenous TLR1 and TLR6 genes have been neutralized and human TLR1, TLR2 and CD14 have been stably transfected. In HEK-Blue hTLR2-TLR6 cells endogenous TLR1 and TLR6 genes have been neutralized and human TLR2, TLR6 and CD14 have been stably transfected. The activity of the test articles is tested at the concentration of 10 μg/mL and compared to control ligands.

The secreted embryonic alkaline phosphatase (SEAP) reporter is under the control of a promoter inducible by the transcription factor NF-κB. This reporter gene allows the monitoring of signaling through the TLR, based on the activation of NF-κB. In a 96-well plate (200 μL total volume) containing the appropriate cells (50,000 – 75,000 cells/well), 20 μL of the test article or the positive control ligand is added to the wells. The media added to the wells is designed for the detection of NF-κB induced SEAP expression. After a 24-hour incubation the optical density (OD) is read at 650 nm on a Molecular Devices SpectraMax 340PC absorbance detector.

In vivo CD4 depletion

Monocolonized mice were treated with anti-CD4-depletion antibody (250 μg/dose/mouse; BioXCell) or isotype control (rat IgG2b; BioXCell) by intraperitoneal injection on day -3, -1, +4 and +11 relative to colonization day.

Tributyrin administration in gnotobiotic setting

Monocolonized mice were treated with tributyrin (150 mM; Sigma) or equimolar glycerol vehicle control by oral gavage three times a week starting on the day of bacterial administration.

IP3 ELISA

Fecal pellets were homogenized in 250 µL of sterile PBS then centrifuged at 1,000 *g* for 10 min. *E. faecalis* was grown in BHI overnight supplemented with phytic acid sodium salt hydrate (1 mM; Sigma). Overnight culture was centrifuged at 4,000 RPM for 10 min Supernatants were collected and IP3 levels were determined using mouse inositol 1, 4, 5, -triphosphate ELISA kit (MyBioSource) following the manufacturer's instructions.

Lantibiotic purification and co-culture

Co-culture of *E. faecalis* with conditioned supernatant from *B. producta* SCSK or KH6 strains and purified lantibiotic was performed as previously described (Kim *et al.*, 2019). Briefly, a frozen aliquot of *B. producta* was inoculated and cultured for 24 hours. Culture supernatant was collected by centrifugation at 4,000 RPM for 10 min and subsequent filtration (0.22 µm). Supernatants were diluted 1:2 with culture broth. *E. faecalis* was subsequent inoculated (10^3 CFUs/mL), cultured for 24 hours and enumerated at the end of the experiment.

To purify lantibiotic from the culture supernatant of *B. producta*, ammonium sulfate was added to 45% saturation and equilibrated overnight stirring at 4°C. The precipitate was collected by centrifugation at 3000 *g* for 30 min at 4°C, resuspended with 5 mL sterile

dPBS and dialyzed through the Amicon Ultra 3 KDa filter with dPBS to wash out the residual ammonium sulfate. Total protein concentrations were quantified with the BCA protein assay kit (Thermo Fisher), normalized (2 mg/mL), sterile filtered and diluted in culture broth (20 μg/mL). *E. faecalis* was inoculated (10^3 CFUs/mL), cultured for 24 hours and enumerated at the end of the experiment.

Statistical analysis

Results are shown as means ± SEM, unless otherwise indicated. Tests between two groups used a two-tailed Student's t-test or a Mann-Whitney U-test. Tests between more than two groups used a one-way ANOVA with Tukey or a Kruskal-Wallis test with Dunn for multiple comparisons. Linear regression was used to assess correlations between two data sets. Survival curves were evaluated using a log-rank Mantel-Cox test with p-values adjusted for multiple comparisons. Results were considered significant at $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$. Statistical significance was calculated using Prism version 9.0 (GraphPad Software).

CHAPTER FOUR

CONCLUDING REMARKS

This thesis focuses on the interactions between pharmacological exposures, the intestinal microbiota and the immune system in the context of allo-HCT. Several studies have demonstrated the association between the intestinal microbiota and clinical outcomes following allo-HCT (Holler *et al*., 2014; Peled *et al*., 2020; Stein-Thoeringer *et al*., 2019). Therefore, interventions to minimize disruptions to the intestinal microbiota are of significant clinical importance. One such intervention involves the stewardship of pharmacological exposures, minimizing the collateral damage of medications on the intestinal microbiota (Weersma et al., 2020; Zimmermann et al., 2021). While preclinical and clinical studies have demonstrated the detrimental impact of broad-spectrum antibiotics on the intestinal microbiota (Fishbein et al., 2023), the influence of non-antibiotic medications on the gut microbes is less well-understood, especially in the human setting. We investigated the associations between various pharmacological exposures (encompassing both antibiotic and non-antibiotic drug classes) on the intestinal microbiota.

Our study enabled the identification of medication-microbiome associations using a high-throughput approach and showed the power of those associations in predicting clinical outcomes. However, important limitations exist, and the interpretation of the results should be performed with care. (i) This study analyzes a retrospective cohort study. Controlled experiments are required to confirm whether our predictions are causal or not. (ii) Drug-microbiome interactions are potentially dependent on drug dosing, which we did not account for. (iii) While synergistic and antagonistic drug interactions have been

reported (Yeh et al., 2009), our model assumes drugs to act independently from each other. (iv) Transplant-specific effects, such as conditioning intensity and graft type, are not fully captured by our model. We assume those effects are partially attenuated by a time parameter and the medication protocol depending on transplant type and conditioning regimen. (v) Medication exposure is only one component among the various perturbations a patient is subjected to. Other environmental factors have also been associated with changes in the intestinal microbiome. For example, dietary intake has been shown to play a major role in shaping the intestinal microbiota in both healthy individuals and cancer patients (Gacesa et al., 2022; Spencer et al., 2021; Wu et al., 2011). Future efforts to collect dietary intake data could help elucidate further the association between diet, changes in the intestinal microbiota and clinical outcomes of allo-HCT patients. Finally, we envision PARADIGM to be a valuable tool to investigate microbiome dynamics *in vivo* and await its application to other comparable datasets, encompassing drug exposures and other environmental factors, to replicate and extend the results in this study.

In the events that microbiome disruption is unavoidable, understanding of the immunological mechanism underlying microbe-host interactions could provide additional therapeutic approaches to maximize clinical outcomes following allo-HCT. Several preclinical and clinical studies have reported that the genus *Enterococcus* is strongly associated with an increased risk of GVHD-related mortality following allo-HCT (Holler *et al*., 2014; Stein-Thoeringer *et al*., 2019). Our study demonstrated that intestinal *Enterococcus* colonization could upregulate MHC-II expression by IECs at non-transplanted steady state and during GVHD. Modulation of MHC-II expression by IECs is dependent on IFNγ signaling from CD4+ T cell subsets. Several microbial taxa have been

identified as inducers of MHC-II expression by IECs (Koyama *et al*., 2023; Tuganbaev *et al*., 2020). Although previous studies have demonstrated the involvement of the innate immune signaling such as TLR (Beyaz *et al*., 2021; Koyama *et al*., 2019), we observed that the regulation of gut epithelium antigen presentation by *Enterococcus* occurred independently of TLR2 signaling and in the absence of DCs. Future studies are important to delineate further the involvement of other immune and non-immune cell types in the regulation of antigen presentation by IECs. In addition, identification of microbial derived molecules driving IFNγ signaling from CD4+ T cells and subsequent MHC-II expression by IECs will further provide important mechanistic insights into gut epithelium-microbe interactions.

Collectively, this thesis supports the development of microbiome-based therapeutics to prevent and ameliorate GVHD pathophysiology and improve patient outcomes following allo-HCT. We envision various potential approaches, including (i) the optimization of pharmacological exposures to limit perturbations to the intestinal microbiota, and (ii) the targeted elimination of pro-inflammatory microbial taxa by rationally designed probiotics to restore intestinal epithelium homeostasis. Importantly, the data from our studies could be extended beyond allo-HCT, given the involvement of the gut microbiota in several other diseases, including inflammatory bowel disease and gastrointestinal cancer.

BIBLIOGRAPHY

Abreu, M.T., Fukata, M., and Arditi, M. (2005). TLR signaling in the gut in health and disease. J Immunol *174*, 4453-4460. 10.4049/jimmunol.174.8.4453.

Alenghat, T., Osborne, L.C., Saenz, S.A., Kobuley, D., Ziegler, C.G., Mullican, S.E., Choi, I., Grunberg, S., Sinha, R., Wynosky-Dolfi, M., et al. (2013). Histone deacetylase 3 coordinates commensal-bacteria-dependent intestinal homeostasis. Nature *504*, 153-157. 10.1038/nature12687.

Arpaia, N., Campbell, C., Fan, X., Dikiy, S., van der Veeken, J., deRoos, P., Liu, H., Cross, J.R., Pfeffer, K., Coffer, P.J., and Rudensky, A.Y. (2013). Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. Nature *504*, 451-455. 10.1038/nature12726.

Artis, D. (2008). Epithelial-cell recognition of commensal bacteria and maintenance of immune homeostasis in the gut. Nat Rev Immunol *8*, 411-420. 10.1038/nri2316.

Beghini, F., McIver, L.J., Blanco-Miguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., Thomas, A.M., et al. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife *10*. 10.7554/eLife.65088.

Belkaid, Y., and Harrison, O.J. (2017). Homeostatic Immunity and the Microbiota. Immunity *46*, 562-576. 10.1016/j.immuni.2017.04.008.

Bendall, M.L., Stevens, S.L., Chan, L.K., Malfatti, S., Schwientek, P., Tremblay, J., Schackwitz, W., Martin, J., Pati, A., Bushnell, B., et al. (2016). Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. ISME J *10*, 1589-1601. 10.1038/ismej.2015.241.

Beyaz, S., Chung, C., Mou, H., Bauer-Rowe, K.E., Xifaras, M.E., Ergin, I., Dohnalova, L., Biton, M., Shekhar, K., Eskiocak, O., et al. (2021). Dietary suppression of MHC class II expression in intestinal epithelial cells enhances intestinal tumorigenesis. Cell Stem Cell *28*, 1922-1935 e1925. 10.1016/j.stem.2021.08.007.

Brooks, J.P., Buck, G.A., Chen, G., Diao, L., Edwards, D.J., Fettweis, J.M., Huzurbazar, S., Rakitin, A., Satten, G.A., Smirnova, E., et al. (2017). Changes in vaginal community state types reflect major shifts in the microbiome. Microb Ecol Health Dis *28*, 1303265. 10.1080/16512235.2017.1303265.

Buffie, C.G., and Pamer, E.G. (2013). Microbiota-mediated colonization resistance against intestinal pathogens. Nat Rev Immunol *13*, 790-801. 10.1038/nri3535.

Buttice, G., Miller, J., Wang, L., and Smith, B.D. (2006). Interferon-gamma induces major histocompatibility class II transactivator (CIITA), which mediates collagen repression and major histocompatibility class II activation by human aortic smooth muscle cells. Circ Res *98*, 472-479. 10.1161/01.RES.0000204725.46332.97.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods *13*, 581-583. 10.1038/nmeth.3869.

Campbell, C., McKenney, P.T., Konstantinovsky, D., Isaeva, O.I., Schizas, M., Vetrter, J., Mai, C., Jin, W., Guo, C., Violante, S., et al. (2020). Bacterial metabolism of bile acids promotes generation of peripheral regulatory T cells. Nature *https://doi.org/10.1038/s41586-020-2193-0*.

Candido, E.P., Reeves, R., and Davie, J.R. (1978). Sodium butyrate inhibits histone deacetylation in cultured cells. Cell *14*, 105-113. 10.1016/0092-8674(78)90305-7.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat Methods *7*, 335-336. 10.1038/nmeth.f.303.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J *6*, 1621-1624. 10.1038/ismej.2012.8.

Cash, H.L., Whitham, C.V., Behrendt, C.L., and Hooper, L.V. (2006). Symbiotic bacteria direct expression of an intestinal bactericidal lectin. Science *313*, 1126-1130. 10.1126/science.1127119.

Cerdo, T., Ruiz, A., Acuna, I., Nieto-Ruiz, A., Dieguez, E., Sepulveda-Valbuena, N., Escudero-Marin, M., Garcia-Santos, J.A., Garcia-Ricobaraza, M., Herrmann, F., et al. (2022). A synbiotics, long chain polyunsaturated fatty acids, and milk fat globule membranes supplemented formula modulates microbiota maturation and neurodevelopment. Clinical Nutrition *41*, 1697-1711. 10.1016/j.clnu.2022.05.013.

Cooke, K.R., Kobzik, L., Martin, T.R., Brewer, J., Delmonte, J., Jr., Crawford, J.M., and Ferrara, J.L. (1996). An experimental model of idiopathic pneumonia syndrome after bone marrow transplantation: I. The roles of minor H antigens and endotoxin. Blood *88*, 3230-3239.

Cooke, K.R., Gerbitz, A., Crawford, J.M., Teshima, T., Hill, G.R., Tesolin, A., Rossignol, D.P., and Ferrara, J.L. (2001). LPS antagonism reduces graft-versus-host disease and preserves graft-versus-leukemia activity after experimental bone marrow transplantation. J Clin Invest *107*, 1581-1589. 10.1172/JCI12156.

Costea, P.I., Hildebrand, F., Arumugam, M., Backhed, F., Blaser, M.J., Bushman, F.D., de Vos, W.M., Ehrlich, S.D., Fraser, C.M., Hattori, M., et al. (2018). Enterotypes in the landscape of gut microbial community composition. Nat Microbiol *3*, 8-16. 10.1038/s41564-017-0072-8.

D' Souza A, F.C. (2019). Current uses and outcomes of hematopoietic cell transplantation (HCT): CIBMTR Summary Slides.

Desai, M.S., Seekatz, A.M., Koropatkin, N.M., Kamada, N., Hickey, C.A., Wolter, M., Pudlo, N.A., Kitamoto, S., Terrapon, N., Muller, A., et al. (2016). A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. Cell *167*, 1339-1353 e1321. 10.1016/j.cell.2016.10.043.

Diaz Caballero, J., Clark, S.T., Coburn, B., Zhang, Y., Wang, P.W., Donaldson, S.L., Tullis, D.E., Yau, Y.C., Waters, V.J., Hwang, D.M., and Guttman, D.S. (2015). Selective Sweeps and Parallel Pathoadaptation Drive Pseudomonas aeruginosa Evolution in the Cystic Fibrosis Lung. mBio *6*, e00981-00915. 10.1128/mBio.00981-15.

DiGiulio, D.B., Callahan, B.J., McMurdie, P.J., Costello, E.K., Lyell, D.J., Robaczewska, A., Sun, C.L., Goltsman, D.S., Wong, R.J., Shaw, G., et al. (2015). Temporal and spatial variation of the human microbiota during pregnancy. Proc Natl Acad Sci U S A *112*, 11060-11065. 10.1073/pnas.1502875112.

Dubin, K.A., Mathur, D., McKenney, P.T., Taylor, B.P., Littmann, E.R., Peled, J.U., van den Brink, M.R.M., Taur, Y., Pamer, E.G., and Xavier, J.B. (2019). Diversification and Evolution of Vancomycin-Resistant Enterococcus faecium during Intestinal Domination. Infect Immun *87*. 10.1128/IAI.00102-19.

Eshleman, E.M., Shao, T.Y., Woo, V., Rice, T., Engleman, L., Didriksen, B.J., Whitt, J., Haslam, D.B., Way, S.S., and Alenghat, T. (2023). Intestinal epithelial HDAC3 and MHC class II coordinate microbiota-specific immunity. J Clin Invest *133*. 10.1172/JCI162190.

Falony, G., Joossens, M., Vieira-Silva, S., Wang, J., Darzi, Y., Faust, K., Kurilshikov, A., Bonder, M.J., Valles-Colomer, M., Vandeputte, D., et al. (2016). Population-level analysis of gut microbiome variation. Science *352*, 560-564. 10.1126/science.aad3503.

Fan, D., Coughlin, L.A., Neubauer, M.M., Kim, J., Kim, M.S., Zhan, X., Simms-Waldrip, T.R., Xie, Y., Hooper, L.V., and Koh, A.Y. (2015). Activation of HIF-1alpha and LL-37 by commensal bacteria inhibits Candida albicans colonization. Nat Med *21*, 808-814. 10.1038/nm.3871.

Fang, Y., Yan, C., Zhao, Q., Zhao, B., Liao, Y., Chen, Y., Wang, D., and Tang, D. (2022). The Association Between Gut Microbiota, Toll-Like Receptors, and Colorectal

Cancer. Clin Med Insights Oncol *16*, 11795549221130549. 10.1177/11795549221130549.

Ferrara, J.L., Levy, R., and Chao, N.J. (1999). Pathophysiologic mechanisms of acute graft-vs.-host disease. Biol Blood Marrow Transplant *5*, 347-356. 10.1016/s1083-8791(99)70011-x.

Fishbein, S.R.S., Mahmud, B., and Dantas, G. (2023). Antibiotic perturbations to the gut microbiome. Nat Rev Microbiol. 10.1038/s41579-023-00933-y.

Fredricks, D.N. (2019). The gut microbiota and graft-versus-host disease. J Clin Invest *129*, 1808-1817. 10.1172/JCI125797.

Fujiwara, H., Docampo, M.D., Riwes, M., Peltier, D., Toubai, T., Henig, I., Wu, S.J., Kim, S., Taylor, A., Brabbs, S., et al. (2018). Microbial metabolite sensor GPR43 controls severity of experimental GVHD. Nat Commun *9*, 3674. 10.1038/s41467-018-06048-w.

Fukuda, S., Toh, H., Hase, K., Oshima, K., Nakanishi, Y., Yoshimura, K., Tobe, T., Clarke, J.M., Topping, D.L., Suzuki, T., et al. (2011). Bifidobacteria can protect from enteropathogenic infection through production of acetate. Nature *469*, 543-547. 10.1038/nature09646.

Furusawa, Y., Obata, Y., Fukuda, S., Endo, T.A., Nakato, G., Takahashi, D., Nakanishi, Y., Uetake, C., Kato, K., Kato, T., et al. (2013). Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature *504*, 446-450. 10.1038/nature12721.

Gacesa, R., Kurilshikov, A., Vich Vila, A., Sinha, T., Klaassen, M.A.Y., Bolte, L.A., Andreu-Sanchez, S., Chen, L., Collij, V., Hu, S., et al. (2022). Environmental factors shaping the gut microbiome in a Dutch population. Nature *604*, 732-739. 10.1038/s41586-022-04567-7.

Ghalayini, M., Launay, A., Bridier-Nahmias, A., Clermont, O., Denamur, E., Lescat, M., and Tenaillon, O. (2018). Evolution of a Dominant Natural Isolate of Escherichia coli in the Human Gut over the Course of a Year Suggests a Neutral Evolution with Reduced Effective Population Size. Appl Environ Microbiol *84*. 10.1128/AEM.02377-17.

Golob, J.L., Pergam, S.A., Srinivasan, S., Fiedler, T.L., Liu, C., Garcia, K., Mielcarek, M., Ko, D., Aker, S., Marquis, S., et al. (2017). Stool Microbiota at Neutrophil Recovery Is Predictive for Severe Acute Graft vs Host Disease After Hematopoietic Cell Transplantation. Clin Infect Dis *65*, 1984-1991. 10.1093/cid/cix699.

Griffin, M.E., Espinosa, J., Becker, J.L., Luo, J.D., Carroll, T.S., Jha, J.K., Fanger, G.R., and Hang, H.C. (2021). Enterococcus peptidoglycan remodeling promotes checkpoint inhibitor cancer immunotherapy. Science *373*, 1040-1046. 10.1126/science.abc9113.

Han, L., Jin, H., Zhou, L., Zhang, X., Fan, Z., Dai, M., Lin, Q., Huang, F., Xuan, L., Zhang, H., and Liu, Q. (2018). Intestinal Microbiota at Engraftment Influence Acute Graft-Versus-Host Disease via the Treg/Th17 Balance in Allo-HSCT Recipients. Front Immunol *9*, 669. 10.3389/fimmu.2018.00669.

Han, L., Zhang, H., Chen, S., Zhou, L., Li, Y., Zhao, K., Huang, F., Fan, Z., Xuan, L., Zhang, X., et al. (2019). Intestinal Microbiota Can Predict Acute Graft-versus-Host Disease Following Allogeneic Hematopoietic Stem Cell Transplantation. Biol Blood Marrow Transplant *25*, 1944-1955. 10.1016/j.bbmt.2019.07.006.

Hanchi, H., Mottawea, W., Sebei, K., and Hammami, R. (2018). The Genus Enterococcus: Between Probiotic Potential and Safety Concerns-An Update. Front Microbiol *9*, 1791. 10.3389/fmicb.2018.01791.

Hang, S., Paik, D., Yao, L., Kim, E., Trinath, J., Lu, J., Ha, S., Nelson, B.N., Kelly, S.P., Wu, L., et al. (2019). Bile acid metabolites control TH17 and Treg cell differentiation. Nature *576*, 143-148. 10.1038/s41586-019-1785-z.

Heuberger, C.E., Janney, A., Ilott, N., Bertocchi, A., Pott, S., Gu, Y., Pohin, M., Friedrich, M., Mann, E.H., Pearson, C., et al. (2023). MHC class II antigen presentation by intestinal epithelial cells fine-tunes bacteria-reactive CD4 T cell responses. Mucosal Immunol. 10.1016/j.mucimm.2023.05.001.

Holler, E., Butzhammer, P., Schmid, K., Hundsrucker, C., Koestler, J., Peter, K., Zhu, W., Sporrer, D., Hehlgans, T., Kreutz, M., et al. (2014). 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. 10.1016/j.bbmt.2014.01.030.

Holmes, I., Harris, K., and Quince, C. (2012). Dirichlet multinomial mixtures: generative models for microbial metagenomics. PLoS One *7*, e30126. 10.1371/journal.pone.0030126.

Hou, K., Wu, Z.X., Chen, X.Y., Wang, J.Q., Zhang, D., Xiao, C., Zhu, D., Koya, J.B., Wei, L., Li, J., and Chen, Z.S. (2022). Microbiota in health and diseases. Signal Transduct Target Ther *7*, 135. 10.1038/s41392-022-00974-4.

Hulsdunker, J., Ottmuller, K.J., Neeff, H.P., Koyama, M., Gao, Z., Thomas, O.S., Follo, M., Al-Ahmad, A., Prinz, G., Duquesne, S., et al. (2018). Neutrophils provide cellular communication between ileum and mesenteric lymph nodes at graft-versus-host disease onset. Blood *131*, 1858-1869. 10.1182/blood-2017-10-812891.

Human Microbiome Project, C. (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207-214. 10.1038/nature11234.

Jenq, R.R., Taur, Y., Devlin, S.M., Ponce, D.M., Goldberg, J.D., Ahr, K.F., Littmann, E.R., Ling, L., Gobourne, A.C., Miller, L.C., et al. (2015). Intestinal Blautia Is Associated with Reduced Death from Graft-versus-Host Disease. Biol Blood Marrow Transplant *21*, 1373-1383. 10.1016/j.bbmt.2015.04.016.

Jin, B., Liu, R., Hao, S., Li, Z., Zhu, C., Zhou, X., Chen, P., Fu, T., Hu, Z., Wu, Q., et al. (2017). Defining and characterizing the critical transition state prior to the type 2 diabetes disease. PLoS One *12*, e0180937. 10.1371/journal.pone.0180937.

Johnson, A.J., Vangay, P., Al-Ghalith, G.A., Hillmann, B.M., Ward, T.L., Shields-Cutler, R.R., Kim, A.D., Shmagel, A.K., Syed, A.N., Personalized Microbiome Class, S., et al. (2019). Daily Sampling Reveals Personalized Diet-Microbiome Associations in Humans. Cell Host Microbe *25*, 789-802 e785. 10.1016/j.chom.2019.05.005.

Kanjilal, S., Oberst, M., Boominathan, S., Zhou, H., Hooper, D.C., and Sontag, D. (2020). A decision algorithm to promote outpatient antimicrobial stewardship for uncomplicated urinary tract infection. Sci Transl Med *12*. 10.1126/scitranslmed.aay5067.

Khan, I., Bai, Y., Zha, L., Ullah, N., Ullah, H., Shah, S.R.H., Sun, H., and Zhang, C. (2021). Mechanism of the Gut Microbiota Colonization Resistance and Enteric Pathogen Infection. Front Cell Infect Microbiol *11*, 716299. 10.3389/fcimb.2021.716299.

Kim, S.G., Becattini, S., Moody, T.U., Shliaha, P.V., Littmann, E.R., Seok, R., Gjonbalaj, M., Eaton, V., Fontana, E., Amoretti, L., et al. (2019). Microbiota-derived lantibiotic restores resistance against vancomycin-resistant Enterococcus. Nature *572*, 665-669. 10.1038/s41586-019-1501-z.

Korpela, K., Salonen, A., Virta, L.J., Kekkonen, R.A., Forslund, K., Bork, P., and de Vos, W.M. (2016). Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children. Nat Commun *7*, 10410. 10.1038/ncomms10410.

Koyama, M., Kuns, R.D., Olver, S.D., Raffelt, N.C., Wilson, Y.A., Don, A.L., Lineburg, K.E., Cheong, M., Robb, R.J., Markey, K.A., et al. (2011). Recipient nonhematopoietic antigen-presenting cells are sufficient to induce lethal acute graft-versus-host disease. Nat Med *18*, 135-142. 10.1038/nm.2597.

Koyama, M., Cheong, M., Markey, K.A., Gartlan, K.H., Kuns, R.D., Locke, K.R., Lineburg, K.E., Teal, B.E., Leveque-El Mouttie, L., Bunting, M.D., et al. (2015). Donor colonic CD103+ dendritic cells determine the severity of acute graft-versus-host disease. J Exp Med *212*, 1303-1321. 10.1084/jem.20150329.

Koyama, M., Mukhopadhyay, P., Schuster, I.S., Henden, A.S., Hulsdunker, J., Varelias, A., Vetizou, M., Kuns, R.D., Robb, R.J., Zhang, P., et al. (2019). MHC Class II Antigen Presentation by the Intestinal Epithelium Initiates Graft-versus-Host Disease and Is

Influenced by the Microbiota. Immunity *51*, 885-898 e887. 10.1016/j.immuni.2019.08.011.

Koyama, M., Hippe, D.S., Srinivasan, S., Proll, S.C., Miltiadous, O., Li, N., Zhang, P., Ensbey, K.S., Hoffman, N.G., Schmidt, C.R., et al. (2023). Intestinal microbiota controls graft-versus-host disease independent of donor-host genetic disparity. Immunity *56*, 1876-1893 e1878. 10.1016/j.immuni.2023.06.024.

Kreisel, D., Richardson, S.B., Li, W., Lin, X., Kornfeld, C.G., Sugimoto, S., Hsieh, C.S., Gelman, A.E., and Krupnick, A.S. (2010). Cutting edge: MHC class II expression by pulmonary nonhematopoietic cells plays a critical role in controlling local inflammatory responses. J Immunol *185*, 3809-3813. 10.4049/jimmunol.1000971.

Kubinak, J.L., and Round, J.L. (2012). Toll-like receptors promote mutually beneficial commensal-host interactions. PLoS Pathog *8*, e1002785. 10.1371/journal.ppat.1002785.

Lee, J.S., Cella, M., McDonald, K.G., Garlanda, C., Kennedy, G.D., Nukaya, M., Mantovani, A., Kopan, R., Bradfield, C.A., Newberry, R.D., and Colonna, M. (2011). AHR drives the development of gut ILC22 cells and postnatal lymphoid tissues via pathways dependent on and independent of Notch. Nat Immunol *13*, 144-151. 10.1038/ni.2187.

Lee, K.H., Gordon, A., Shedden, K., Kuan, G., Ng, S., Balmaseda, A., and Foxman, B. (2019a). The respiratory microbiome and susceptibility to influenza virus infection. PLoS One *14*, e0207898. 10.1371/journal.pone.0207898.

Lee, S.E., Lim, J.Y., Ryu, D.B., Kim, T.W., Park, S.S., Jeon, Y.W., Yoon, J.H., Cho, B.S., Eom, K.S., Kim, Y.J., et al. (2019b). Alteration of the Intestinal Microbiota by Broad-Spectrum Antibiotic Use Correlates with the Occurrence of Intestinal Graft-versus-Host Disease. Biol Blood Marrow Transplant *25*, 1933-1943. 10.1016/j.bbmt.2019.06.001.

Leendertse, M., Willems, R.J., Giebelen, I.A., van den Pangaart, P.S., Wiersinga, W.J., de Vos, A.F., Florquin, S., Bonten, M.J., and van der Poll, T. (2008). TLR2-dependent MyD88 signaling contributes to early host defense in murine Enterococcus faecium peritonitis. J Immunol *180*, 4865-4874. 10.4049/jimmunol.180.7.4865.

Levitus, M., Rewane, A., and Perera, T.B. (2023). Vancomycin-Resistant Enterococci. In StatPearls.

Levy, M., Thaiss, C.A., Zeevi, D., Dohnalova, L., Zilberman-Schapira, G., Mahdi, J.A., David, E., Savidor, A., Korem, T., Herzig, Y., et al. (2015). Microbiota-Modulated Metabolites Shape the Intestinal Microenvironment by Regulating NLRP6 Inflammasome Signaling. Cell *163*, 1428-1443. 10.1016/j.cell.2015.10.048.

Lewis, G., Wang, B., Shafiei Jahani, P., Hurrell, B.P., Banie, H., Aleman Muench, G.R., Maazi, H., Helou, D.G., Howard, E., Galle-Treger, L., et al. (2019). Dietary Fiber-Induced Microbial Short Chain Fatty Acids Suppress ILC2-Dependent Airway Inflammation. Front Immunol *10*, 2051. 10.3389/fimmu.2019.02051.

Liao, C., Taylor, B.P., Ceccarani, C., Fontana, E., Amoretti, L.A., Wright, R.J., Gomes, A.L.C., Peled, J.U., Taur, Y., Perales, M.A., et al. (2021). Compilation of longitudinal microbiota data and hospitalome from hematopoietic cell transplantation patients. Sci Data *8*, 71. 10.1038/s41597-021-00860-8.

Lindemans, C.A., Calafiore, M., Mertelsmann, A.M., O'Connor, M.H., Dudakov, J.A., Jenq, R.R., Velardi, E., Young, L.F., Smith, O.M., Lawrence, G., et al. (2015). Interleukin-22 promotes intestinal-stem-cell-mediated epithelial regeneration. Nature *528*, 560-564. 10.1038/nature16460.

Long, E.M., Millen, B., Kubes, P., and Robbins, S.M. (2009). Lipoteichoic acid induces unique inflammatory responses when compared to other toll-like receptor 2 ligands. PLoS One *4*, e5601. 10.1371/journal.pone.0005601.

Maier, L., Pruteanu, M., Kuhn, M., Zeller, G., Telzerow, A., Anderson, E.E., Brochado, A.R., Fernandez, K.C., Dose, H., Mori, H., et al. (2018). Extensive impact of non-antibiotic drugs on human gut bacteria. Nature *555*, 623-628. 10.1038/nature25979.

Maier, L., Goemans, C.V., Wirbel, J., Kuhn, M., Eberl, C., Pruteanu, M., Muller, P., Garcia-Santamarina, S., Cacace, E., Zhang, B., et al. (2021). Unravelling the collateral damage of antibiotics on gut bacteria. Nature *599*, 120-124. 10.1038/s41586-021-03986-2.

Manfredo Vieira, S., Hiltensperger, M., Kumar, V., Zegarra-Ruiz, D., Dehner, C., Khan, N., Costa, F.R.C., Tiniakou, E., Greiling, T., Ruff, W., et al. (2018). Translocation of a gut pathobiont drives autoimmunity in mice and humans. Science *359*, 1156-1161. 10.1126/science.aar7201.

Markey, K.A., Schluter, J., Gomes, A.L., Littmann, E., Pickard, A., Taylor, B.P., Giardina, P.A., Weber, D., Dai, A., Docampo, M., et al. (2020). Microbe-derived short chain fatty acids butyrate and propionate are associated with protection from chronic GVHD. Blood. 10.1182/blood.2019003369.

Mathewson, N.D., Jenq, R., Mathew, A.V., Koenigsknecht, M., Hanash, A., Toubai, T., Oravecz-Wilson, K., Wu, S.R., Sun, Y., Rossi, C., et al. (2016). Gut microbiome-derived metabolites modulate intestinal epithelial cell damage and mitigate graft-versus-host disease. Nat Immunol *17*, 505-513. 10.1038/ni.3400.

Matte, C.C., Liu, J., Cormier, J., Anderson, B.E., Athanasiadis, I., Jain, D., McNiff, J., and Shlomchik, W.D. (2004). Donor APCs are required for maximal GVHD but not for GVL. Nat Med *10*, 987-992. 10.1038/nm1089.

Michonneau, D., Latis, E., Curis, E., Dubouchet, L., Ramamoorthy, S., Ingram, B., de Latour, R.P., Robin, M., de Fontbrune, F.S., Chevret, S., et al. (2019). Metabolomics analysis of human acute graft-versus-host disease reveals changes in host and microbiota-derived metabolites. Nat Commun *10*, 5695. 10.1038/s41467-019-13498-3.

Molina, M.A., Diaz, A.M., Hesse, C., Ginter, W., Gentilini, M.V., Nunez, G.G., Canellada, A.M., Sparwasser, T., Berod, L., Castro, M.S., and Manghi, M.A. (2015). Immunostimulatory Effects Triggered by Enterococcus faecalis CECT7121 Probiotic Strain Involve Activation of Dendritic Cells and Interferon-Gamma Production. PLoS One *10*, e0127262. 10.1371/journal.pone.0127262.

Montassier, E., Batard, E., Massart, S., Gastinne, T., Carton, T., Caillon, J., Le Fresne, S., Caroff, N., Hardouin, J.B., Moreau, P., et al. (2014). 16S rRNA gene pyrosequencing reveals shift in patient faecal microbiota during high-dose chemotherapy as conditioning regimen for bone marrow transplantation. Microb Ecol *67*, 690-699. 10.1007/s00248-013-0355-4.

Morjaria, S., Schluter, J., Taylor, B.P., Littmann, E.R., Carter, R.A., Fontana, E., Peled, J.U., van den Brink, M.R.M., Xavier, J.B., and Taur, Y. (2019). Antibiotic-Induced Shifts in Fecal Microbiota Density and Composition during Hematopoietic Stem Cell Transplantation. Infect Immun *87*. 10.1128/IAI.00206-19.

Munoz, A., Hayward, M.R., Bloom, S.M., Rocafort, M., Ngcapu, S., Mafunda, N.A., Xu, J., Xulu, N., Dong, M., Dong, K.L., et al. (2021). Modeling the temporal dynamics of cervicovaginal microbiota identifies targets that may promote reproductive health. Microbiome *9*, 163. 10.1186/s40168-021-01096-9.

Nalawade, T.M., Bhat, K., and Sogi, S.H. (2015). Bactericidal activity of propylene glycol, glycerine, polyethylene glycol 400, and polyethylene glycol 1000 against selected microorganisms. J Int Soc Prev Community Dent *5*, 114-119. 10.4103/2231-0762.155736.

Nikcevich, K.M., Piskurich, J.F., Hellendall, R.P., Wang, Y., and Ting, J.P. (1999). Differential selectivity of CIITA promoter activation by IFN-gamma and IRF-1 in astrocytes and macrophages: CIITA promoter activation is not affected by TNF-alpha. J Neuroimmunol *99*, 195-204. 10.1016/s0165-5728(99)00117-4.

Ocvirk, S., Sava, I.G., Lengfelder, I., Lagkouvardos, I., Steck, N., Roh, J.H., Tchaptchet, S., Bao, Y., Hansen, J.J., Huebner, J., et al. (2015). Surface-Associated Lipoproteins Link Enterococcus faecalis Virulence to Colitogenic Activity in IL-10-Deficient Mice Independent of Their Expression Levels. PLoS Pathog *11*, e1004911. 10.1371/journal.ppat.1004911.

Park, O.J., Han, J.Y., Baik, J.E., Jeon, J.H., Kang, S.S., Yun, C.H., Oh, J.W., Seo, H.S., and Han, S.H. (2013). Lipoteichoic acid of Enterococcus faecalis induces the expression

of chemokines via TLR2 and PAFR signaling pathways. J Leukoc Biol *94*, 1275-1284. 10.1189/jlb.1012522.

Peled, J.U., Devlin, S.M., Staffas, A., Lumish, M., Khanin, R., Littmann, E.R., Ling, L., Kosuri, S., Maloy, M., Slingerland, J.B., et al. (2017). Intestinal Microbiota and Relapse After Hematopoietic-Cell Transplantation. J Clin Oncol *35*, 1650-1659. 10.1200/JCO.2016.70.3348.

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., et al. (2020). Microbiota as Predictor of Mortality in Allogeneic Hematopoietic-Cell Transplantation. N Engl J Med *382*, 822-834. 10.1056/NEJMoa1900623.

Pettigrew, M.M., Gent, J.F., Kong, Y., Halpin, A.L., Pineles, L., Harris, A.D., and Johnson, J.K. (2019). Gastrointestinal Microbiota Disruption and Risk of Colonization With Carbapenem-resistant Pseudomonas aeruginosa in Intensive Care Unit Patients. Clinical Infectious Diseases *69*, 604-613. 10.1093/cid/ciy936.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature *464*, 59-65. 10.1038/nature08821.

Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., and Medzhitov, R. (2004). Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. Cell *118*, 229-241. 10.1016/j.cell.2004.07.002.

Rashid, M., Narang, A., Thakur, S., Jain, S.K., and Kaur, S. (2023). Therapeutic and prophylactic effects of oral administration of probiotic Enterococcus faecium Smr18 in Salmonella enterica-infected mice. Gut Pathog *15*, 23. 10.1186/s13099-023-00548-x.

Rashidi, A., Herman, A., Gomes, A.L.C., Peled, J.U., Jenq, R.R., Brereton, D.G., Staley, C., Blazar, B.R., and Weisdorf, D.J. (2020). An alpha-defensin gene single nucleotide polymorphism modulates the gut microbiota and may alter the risk of acute graft-versus-host disease. Br J Haematol *189*, 926-930. 10.1111/bjh.16458.

Rashidi, A., Peled, J.U., Ebadi, M., Rehman, T.U., Elhusseini, H., Marcello, L.T., Halaweish, H., Kaiser, T., Holtan, S.G., Khoruts, A., et al. (2022). Protective Effect of Intestinal Blautia Against Neutropenic Fever in Allogeneic Transplant Recipients. Clin Infect Dis *75*, 1912-1920. 10.1093/cid/ciac299.

Reddy, P., Maeda, Y., Hotary, K., Liu, C., Reznikov, L.L., Dinarello, C.A., and Ferrara, J.L. (2004). Histone deacetylase inhibitor suberoylanilide hydroxamic acid reduces acute graft-versus-host disease and preserves graft-versus-leukemia effect. Proc Natl Acad Sci U S A *101*, 3921-3926. 10.1073/pnas.0400380101.

Rolling, T., Zhai, B., Gjonbalaj, M., Tosini, N., Yasuma-Mitobe, K., Fontana, E., Amoretti, L.A., Wright, R.J., Ponce, D.M., Perales, M.A., et al. (2021). Haematopoietic cell transplantation outcomes are linked to intestinal mycobiota dynamics and an expansion of Candida parapsilosis complex species. Nat Microbiol *6*, 1505-1515. 10.1038/s41564-021-00989-7.

Schroeder, B.O., Birchenough, G.M.H., Stahlman, M., Arike, L., Johansson, M.E.V., Hansson, G.C., and Backhed, F. (2018). Bifidobacteria or Fiber Protects against Diet-Induced Microbiota-Mediated Colonic Mucus Deterioration. Cell Host Microbe *23*, 27-40 e27. 10.1016/j.chom.2017.11.004.

Schwab, L., Goroncy, L., Palaniyandi, S., Gautam, S., Triantafyllopoulou, A., Mocsai, A., Reichardt, W., Karlsson, F.J., Radhakrishnan, S.V., Hanke, K., et al. (2014). Neutrophil granulocytes recruited upon translocation of intestinal bacteria enhance graft-versus-host disease via tissue damage. Nat Med *20*, 648-654. 10.1038/nm.3517.

Seishima, J., Iida, N., Kitamura, K., Yutani, M., Wang, Z., Seki, A., Yamashita, T., Sakai, Y., Honda, M., Yamashita, T., et al. (2019). Gut-derived Enterococcus faecium from ulcerative colitis patients promotes colitis in a genetically susceptible mouse host. Genome Biol *20*, 252. 10.1186/s13059-019-1879-9.

Shimada, Y., Kinoshita, M., Harada, K., Mizutani, M., Masahata, K., Kayama, H., and Takeda, K. (2013). Commensal bacteria-dependent indole production enhances epithelial barrier function in the colon. PLoS One *8*, e80604. 10.1371/journal.pone.0080604.

Shlomchik, W.D., Couzens, M.S., Tang, C.B., McNiff, J., Robert, M.E., Liu, J., Shlomchik, M.J., and Emerson, S.G. (1999). Prevention of graft versus host disease by inactivation of host antigen-presenting cells. Science *285*, 412-415. 10.1126/science.285.5426.412.

Shono, Y., Tuckett, A.Z., Ouk, S., Liou, H.C., Altan-Bonnet, G., Tsai, J.J., Oyler, J.E., Smith, O.M., West, M.L., Singer, N.V., et al. (2014). A small-molecule c-Rel inhibitor reduces alloactivation of T cells without compromising antitumor activity. Cancer Discov *4*, 578-591. 10.1158/2159-8290.CD-13-0585.

Shono, Y., Docampo, M.D., Peled, J.U., Perobelli, S.M., Velardi, E., Tsai, J.J., Slingerland, A.E., Smith, O.M., Young, L.F., Gupta, J., et al. (2016). Increased GVHD-related mortality with broad-spectrum antibiotic use after allogeneic hematopoietic stem cell transplantation in human patients and mice. Sci Transl Med *8*, 339ra371. 10.1126/scitranslmed.aaf2311.

Shouval, R., Waters, N.R., Gomes, A.L.C., Zuanelli Brambilla, C., Fei, T., Devlin, S.M., Nguyen, C.L., Markey, K.A., Dai, A., Slingerland, J.B., et al. (2022). Conditioning regimens are associated with distinct patterns of microbiota injury in allogeneic hematopoietic cell transplantation. Clin Cancer Res. 10.1158/1078-0432.CCR-22-1254.

Simms-Waldrip, T.R., Sunkersett, G., Coughlin, L.A., Savani, M.R., Arana, C., Kim, J., Kim, M., Zhan, X., Greenberg, D.E., Xie, Y., et al. (2017). Antibiotic-Induced Depletion of Anti-inflammatory Clostridia Is Associated with the Development of Graft-versus-Host Disease in Pediatric Stem Cell Transplantation Patients. Biol Blood Marrow Transplant *23*, 820-829. 10.1016/j.bbmt.2017.02.004.

Soares, F.S., Amaral, F.C., Silva, N.L.C., Valente, M.R., Santos, L.K.R., Yamashiro, L.H., Scheffer, M.C., Castanheira, F., Ferreira, R.G., Gehrke, L., et al. (2017). Antibiotic-Induced Pathobiont Dissemination Accelerates Mortality in Severe Experimental Pancreatitis. Front Immunol *8*, 1890. 10.3389/fimmu.2017.01890.

Sovran, B., Hugenholtz, F., Elderman, M., Van Beek, A.A., Graversen, K., Huijskes, M., Boekschoten, M.V., Savelkoul, H.F.J., De Vos, P., Dekker, J., and Wells, J.M. (2019). Age-associated Impairment of the Mucus Barrier Function is Associated with Profound Changes in Microbiota and Immunity. Sci Rep *9*, 1437. 10.1038/s41598-018-35228-3.

Spencer, C.N., McQuade, J.L., Gopalakrishnan, V., McCulloch, J.A., Vetizou, M., Cogdill, A.P., Khan, M.A.W., Zhang, X.T., White, M.G., Peterson, C.B., et al. (2021). Dietary fiber and probiotics influence the gut microbiome and melanoma immunotherapy response. Science *374*, 1632-+. 10.1126/science.aaz7015.

Staffas, A., Burgos da Silva, M., Slingerland, A.E., Lazrak, A., Bare, C.J., Holman, C.D., Docampo, M.D., Shono, Y., Durham, B., Pickard, A.J., et al. (2018). Nutritional Support from the Intestinal Microbiota Improves Hematopoietic Reconstitution after Bone Marrow Transplantation in Mice. Cell Host Microbe *23*, 447-457 e444. 10.1016/j.chom.2018.03.002.

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., et al. (2019). Lactose drives Enterococcus expansion to promote graft-versus-host disease. Science *366*, 1143-1149. 10.1126/science.aax3760.

Stewart, C.J., Ajami, N.J., O'Brien, J.L., Hutchinson, D.S., Smith, D.P., Wong, M.C., Ross, M.C., Lloyd, R.E., Doddapaneni, H., Metcalf, G.A., et al. (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. Nature *562*, 583-588. 10.1038/s41586-018-0617-x.

Stoma, I., Littmann, E.R., Peled, J.U., Giralt, S., van den Brink, M.R.M., Pamer, E.G., and Taur, Y. (2020). Compositional flux within the intestinal microbiota and risk for bloodstream infection with gram-negative bacteria. Clin Infect Dis. 10.1093/cid/ciaa068.

Stoma, I., Littmann, E.R., Peled, J.U., Giralt, S., van den Brink, M.R.M., Pamer, E.G., and Taur, Y. (2021). Compositional Flux Within the Intestinal Microbiota and Risk for Bloodstream Infection With Gram-negative Bacteria. Clin Infect Dis *73*, e4627-e4635. 10.1093/cid/ciaa068.

Swimm, A., Giver, C.R., DeFilipp, Z., Rangaraju, S., Sharma, A., Ulezko Antonova, A., Sonowal, R., Capaldo, C., Powell, D., Qayed, M., et al. (2018). Indoles derived from intestinal microbiota act via type I interferon signaling to limit graft-versus-host disease. Blood *132*, 2506-2519. 10.1182/blood-2018-03-838193.

Taur, Y., Xavier, J.B., Lipuma, L., Ubeda, C., Goldberg, J., Gobourne, A., Lee, Y.J., Dubin, K.A., Socci, N.D., Viale, A., et al. (2012). Intestinal domination and the risk of bacteremia in patients undergoing allogeneic hematopoietic stem cell transplantation. Clin Infect Dis *55*, 905-914. 10.1093/cid/cis580.

Taur, Y., Jenq, R.R., Perales, M.A., Littmann, E.R., Morjaria, S., Ling, L., No, D., Gobourne, A., Viale, A., Dahi, P.B., et al. (2014). The effects of intestinal tract bacterial diversity on mortality following allogeneic hematopoietic stem cell transplantation. Blood *124*, 1174-1182. 10.1182/blood-2014-02-554725.

Taur, Y., Coyte, K., Schluter, J., Robilotti, E., Figueroa, C., Gjonbalaj, M., Littmann, E.R., Ling, L., Miller, L., Gyaltshen, Y., et al. (2018). Reconstitution of the gut microbiota of antibiotic-treated patients by autologous fecal microbiota transplant. Sci Transl Med *10*. 10.1126/scitranslmed.aap9489.

Trompette, A., Gollwitzer, E.S., Yadava, K., Sichelstiel, A.K., Sprenger, N., Ngom-Bru, C., Blanchard, C., Junt, T., Nicod, L.P., Harris, N.L., and Marsland, B.J. (2014). Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. Nat Med *20*, 159-166. 10.1038/nm.3444.

Tropini, C., Moss, E.L., Merrill, B.D., Ng, K.M., Higginbottom, S.K., Casavant, E.P., Gonzalez, C.G., Fremin, B., Bouley, D.M., Elias, J.E., et al. (2018). Transient Osmotic Perturbation Causes Long-Term Alteration to the Gut Microbiota. Cell *173*, 1742-1754 e1717. 10.1016/j.cell.2018.05.008.

Truong, D.T., Tett, A., Pasolli, E., Huttenhower, C., and Segata, N. (2017). Microbial strain-level population structure and genetic diversity from metagenomes. Genome Res *27*, 626-638. 10.1101/gr.216242.116.

Tsai, J.J., Velardi, E., Shono, Y., Argyropoulos, K.V., Holland, A.M., Smith, O.M., Yim, N.L., Rao, U.K., Kreines, F.M., Lieberman, S.R., et al. (2018). Nrf2 regulates CD4(+) T cell-induced acute graft-versus-host disease in mice. Blood *132*, 2763-2774. 10.1182/blood-2017-10-812941.

Tuganbaev, T., Mor, U., Bashiardes, S., Liwinski, T., Nobs, S.P., Leshem, A., Dori-Bachash, M., Thaiss, C.A., Pinker, E.Y., Ratiner, K., et al. (2020). Diet Diurnally Regulates Small Intestinal Microbiome-Epithelial-Immune Homeostasis and Enteritis. Cell *182*, 1441-1459 e1421. 10.1016/j.cell.2020.08.027.

van Bekkum, D.W., Roodenburg, J., Heidt, P.J., and van der Waaij, D. (1974). Mitigation of secondary disease of allogeneic mouse radiation chimeras by modification of the intestinal microflora. J Natl Cancer Inst *52*, 401-404. 10.1093/jnci/52.2.401.

Vatanen, T., Franzosa, E.A., Schwager, R., Tripathi, S., Arthur, T.D., Vehik, K., Lernmark, A., Hagopian, W.A., Rewers, M.J., She, J.X., et al. (2018). The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. Nature *562*, 589-594. 10.1038/s41586-018-0620-2.

Vermeulen, R., Schymanski, E.L., Barabasi, A.L., and Miller, G.W. (2020). The exposome and health: Where chemistry meets biology. Science *367*, 392-396. 10.1126/science.aay3164.

Vich Vila, A., Collij, V., Sanna, S., Sinha, T., Imhann, F., Bourgonje, A.R., Mujagic, Z., Jonkers, D., Masclee, A.A.M., Fu, J., et al. (2020). Impact of commonly used drugs on the composition and metabolic function of the gut microbiota. Nat Commun *11*, 362. 10.1038/s41467-019-14177-z.

Vieira-Silva, S., Falony, G., Belda, E., Nielsen, T., Aron-Wisnewsky, J., Chakaroun, R., Forslund, S.K., Assmann, K., Valles-Colomer, M., Nguyen, T.T.D., et al. (2020). Statin therapy is associated with lower prevalence of gut microbiota dysbiosis. Nature *581*, 310-315. 10.1038/s41586-020-2269-x.

Weersma, R.K., Zhernakova, A., and Fu, J. (2020). Interaction between drugs and the gut microbiome. Gut *69*, 1510-1519. 10.1136/gutjnl-2019-320204.

Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.Y., Keilbaugh, S.A., Bewtra, M., Knights, D., Walters, W.A., Knight, R., et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. Science *334*, 105-108. 10.1126/science.1208344.

Wu, S.E., Hashimoto-Hill, S., Woo, V., Eshleman, E.M., Whitt, J., Engleman, L., Karns, R., Denson, L.A., Haslam, D.B., and Alenghat, T. (2020). Microbiota-derived metabolite promotes HDAC3 activity in the gut. Nature *586*, 108-112. 10.1038/s41586-020-2604-2.

Yeh, P.J., Hegreness, M.J., Aiden, A.P., and Kishony, R. (2009). Drug interactions and the evolution of antibiotic resistance. Nat Rev Microbiol *7*, 460-466. 10.1038/nrmicro2133.

Ying Taur, R.R.J., Miguel-Angel Perales, et al. (2014). The effects of intestinal tract bacterial diversity on mortality following allogeneic hematopoietic stem cell transplantation. Blood *124*, 1174-1182.

Zelante, T., Iannitti, R.G., Cunha, C., De Luca, A., Giovannini, G., Pieraccini, G., Zecchi, R., D'Angelo, C., Massi-Benedetti, C., Fallarino, F., et al. (2013). Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. Immunity *39*, 372-385. 10.1016/j.immuni.2013.08.003.

Zhao, Y., Liu, Q., Yang, L., He, D., Wang, L., Tian, J., Li, Y., Zi, F., Bao, H., Yang, Y., et al. (2013). TLR4 inactivation protects from graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. Cell Mol Immunol *10*, 165-175. 10.1038/cmi.2012.58.

Zhao, Y., Zhang, J., Han, X., and Fan, S. (2019). Total body irradiation induced mouse small intestine senescence as a late effect. J Radiat Res *60*, 442-450. 10.1093/jrr/rrz026.

Zhou, Y., Chen, H., He, H., Du, Y., Hu, J., Li, Y., Li, Y., Zhou, Y., Wang, H., Chen, Y., and Nie, Y. (2016). Increased Enterococcus faecalis infection is associated with clinically active Crohn disease. Medicine (Baltimore) *95*, e5019. 10.1097/MD.0000000000005019.

Zimmermann, M., Patil, K.R., Typas, A., and Maier, L. (2021). Towards a mechanistic understanding of reciprocal drug-microbiome interactions. Mol Syst Biol *17*, e10116. 10.15252/msb.202010116.

Zou, J., Chassaing, B., Singh, V., Pellizzon, M., Ricci, M., Fythe, M.D., Kumar, M.V., and Gewirtz, A.T. (2018). Fiber-Mediated Nourishment of Gut Microbiota Protects against Diet-Induced Obesity by Restoring IL-22-Mediated Colonic Health. Cell Host Microbe *23*, 41-53 e44. 10.1016/j.chom.2017.11.003.