Title
Gut Microbial Taxa and Products that Regulate Immune Responses and Disease Features

Author
Chang, Yu-Ling

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Gut Microbial Taxa and Products that Regulate Immune Responses and Disease Features

A dissertation submitted in partial satisfaction of

the requirements for the degree Doctor of Philosophy

in Molecular Biology

by

Yu-Ling Chang

2017

ABSTRACT OF THE DISSERTATION

Gut Microbial Taxa and Products that Regulate Immune Responses and Disease Features

by

Yu-Ling Chang

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2017

Professor Matteo Pellegrini, Chair

Gut microbiome has a profound effect on human health and disease. The burgeoning body of research investigating the gut microbiome has illustrated its important roles on not only the local inflammatory diseases, such as inflammatory bowel disease, but also on systemic immunological disorders, such as autoimmune diseases. Many of these associations were determined at the level of metagenome analysis. This provides categorical information on both taxonomic composition and predicted functional capacities of the microbial community. This thesis tackles the problem of how gut microbiota affect human diseases using both these categories of assessment.

Although the accumulating evidence suggested the important roles of gut microbiota in the human immune system and Crohn's disease (CD), the molecular mechanisms of how microbiome directly modulates immune responses and leads to CD progression remain unclear. To advance the current understanding of microbial roles in CD, we developed a systems biology approach to assess the human immune function in response to disease-associated microbial products. Progress in computational omics analysis of the microbiome in CD enables the functional inference and prediction of microbial metabolites. Both genetic and clinical studies suggested the involvement of human CD4+ T cells in CD progression. We therefore tabulated a set of microbial metabolites predicted to be differentially abundant in CD, and screened these microbial metabolites for their bioactivity in human CD4+ T cell functions. Our screen revealed 15 bioactive microbial metabolites, 3 previously reported and 12 unprecedented, with selective action on CD4+ T cell cytokine production. Mechanistic assessment of one novel microbe-derived metabolite, ascorbate, revealed apoptosis of activated human CD4+ T cells associated with selective inhibition of glycolytic energy metabolism. These findings suggest a substantial rate (11%) of bioactive metabolites among the predicted CD-associated metabolite reservoirs and provide evidence for novel modes of microbial activity targeting T cell metabolism.

Emerging studies have characterized the association of taxonomic composition with systemic autoimmune diseases with various clinical phenotypes and modes of pathologic immune responses. In this thesis, we assessed systemic sclerosis (SSc), a previously unstudied autoimmune disease with respect to the microbiome, notable for both systemic manifestations and gastrointestinal involvement. The metagenomic study of mucosal-luminal interface (MLI) samples revealed a unique microbial composition associated with disease state. Patients with SSc had decreased abundances of *Faecalibacterium* and *Clostridium*, and increased abundances

of *Bifidobacterium*, *Lactobacillus, Fusobacterium* and γ-*Proteobacteria* compared with healthy controls. In addition, the increase of *Fusobacterium* species and the decrease of *Bacteroides fragilis* were associated with moderate/severe gastrointestinal symptoms. We conducted an independent study of faecal samples from two independent cohorts of patients, from United States and Norway, that revealed a similar group of microbial taxa associations, and a greater dysbiosis in patients from the United States. These studies enable the specific targets for intervention to avert or treat the gastrointestinal involvement in systemic sclerosis.

The study of gut microbiota is beginning to evolve from association toward causality. The functions of gut microbiota are not only restricted to local inflammatory responses but expanded to systemic immune regulations. We hypothesized that gut microbiome may regulate host immune cellular responses by producing metabolites. We investigated the bioactive microbial factors regulating Crohn's disease, and explored the potential microbial taxonomic biomarkers in a systemic autoimmune disorder, SSc. Those improved understandings of the microbiome activity and metabolism may inform therapeutic strategies and ultimately improve patient health.

This dissertation of Yu-Ling Chang is approved.

Jonathan Braun

Alexander Hoffmann

Elaine Reed

Lili Yang

Matteo Pellegrini, Committee Chair

University of California, Los Angeles

2017

DEDICATION

To my always supporting parents

Chin-Yang Chang and Hui-Chun Wang

and beloved life partner

Chun-Hao Chang

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AA	Ascorbic acid
ANOVA	Analysis of variance
AhR	Aryl hydrocarbon receptor
BGCs	Biosynthetic gene clusters
CA	Cholic acid
CD	Crohn's disease
CD28	Cluster of differentiation 28
CD3	Cluster of differentiation 3
CD4	Cluster of differentiation 4
CoA	Coenzyme A
DCA	Deoxycholic acid
DHA	Dehydroascorbic acid
GLP-1	Glucagon-like peptide-1
HDAC	histone-deacetylase
HMDB	Human Metabolome Database
HMG-CoA	3-hydroxy-3-methyl-glutaryl-coenzyme A
HUMAnN2	HMP Unified Metabolic Analysis Network2
I3A	indole-3-aldehyde
IBD	Inflammatory bowel disease
IFNγ	Interferon-gamma
IL	interleukin
LDA	linear discriminant analysis

MANOVA	multivariate analysis of variance
MIP-3α	macrophage inflammatory protein-3
NK	natural killer
OTUs	operational taxonomic units
PCoA	principal coordinate analysis
PICRUSt	phylogenetic investigation of communities by reconstruction of unobserved states
PVCA	principal variance component analysis
SCFA	short-chain fatty acid
PUFA	polyunsaturated fatty acid
MUFA	monounsaturated fatty acid
SEM	standard error of the mean
SSc	Systemic sclerosis
Th	T helper cells
TMA	Trimethyl-amine
TMAO	Trimethyl-amine N-oxide
TNFα	Tumor necrosis factor alpha

ACKNOWLEDGEMENT

I would like to first and foremost to express my sincere gratitude to my thesis mentor Prof. Jonathan Braun for the continuous and unconditional support of my Ph.D study. His patience and encouragement enabled me to overcome numerous obstacles through my research. The research resources and opportunities he provided created the knowledge-enriched environment where I experienced many innovative research projects. I would like to thank Dr. Elizabeth Volkmann for her support and guidance of systemic sclerosis studies and thank my thesis committee, Prof. Matteo Pellegrini, Alexander Hoffmann, Elaine Reed, and Lily Yang, for their insightful comments and discussions which incented me to deepen my research from many perspectives.

Chapter 2 is currently in submission. I particularly thank the Immune Assessment Core, Dr. Maura Rossetti and Dr. Elaine Reed, for their support in scaling up the assays, Dr. Hera Vlamakis from The Broad Institute of MIT and Harvard for predicting the microbial metabolites, Dr. Curtis Huttenhower, Dr. Ramnik Xavier, and Dr. Thaddeus Stappenbeck for their contribution to collaboration. I would also like to thank Dr. David Casero for the help of statistical analysis, Dr. Shelley Miller and Dr. Romney Humphries for microbial culture, Dr. Laurent Vergnes for the help with the Seahorse experiments, and the fund supported by the Crohn's and Colitis Foundation.

Chapter 3 has been published in Arthritis & Rheumatology in 2016 and Chapter 4 has been published in BMJ Open Gastroenterology in 2017. I would like to thank Dr. Elizabeth Volkmann for the opportunity to involve in this project and all the other authors to make the study possible.

I also want to thank the members in Braun/Gordon/Wadehra laboratory for their support throughout the 4-year journey of my graduate study.

VITA

EDUCATION

2005-2009 Bachelor of Science in Life Science, National Yang-Ming University (NYMU)

2009-2011 Master of Science in Dept. of Microbiology and Immunology, NYMU

2012-2013 Master of Science in Biomedical Imaging, University of California, San Francisco

PUBLICATION

Elizabeth R Volkmann, Anna-Maria Hoffmann-Vold, Yu-Ling Chang, Jonathan P Jacobs, Kirsten Tillisch, Emeran A Mayer, Philip J Clements, Johannes R Hov, Martin Kummen, Øyvind Midtvedt, Venu Lagishetty, Lin Chang, Jennifer S Labus, Øyvind Molberg, Jonathan Braun. "Systemic sclerosis is associated with specific alterations in gastrointestinal microbiota in two independent cohorts." **BMJ Open Gastro.** 2017, 4:e000134 doi:10.1136/bmjgast-2017-000134

Elizabeth R. Volkmann, Yu-Ling Chang, Nashla Barroso, Daniel E. Furst, Philip J. Clements, Alan H. Gorn, Bennett E. Roth, Jeffrey L. Conklin, Terri Getzug, James Borneman, Dermot P. B. McGovern, Maomeng Tong, Jonathan P. Jacobs, Jonathan Braun. "Association of Systemic Sclerosis With a Unique Colonic Microbial Consortium" **Arthritis Rheumatol.** 2016 Jun;68(6):1483-92.

Yu-Ling Chang, Tzu-Hui Chen, Yi-Hsiu Wu, Guann-An Chen, Tzu-Huei Weng, Ping-Hui Tseng, Shie-Liang Hsieh, Shu-Ling Fu, Chi-Hung Lin, Chun-Jen Chen, Ching-Liang Chu, Iok In Christine Chio, Tak Wah Mak and Nien-Jung Chen. "A novel TLR2-triggered signaling crosstalk synergistically intensifies TNF-mediated IL-6 induction." **J Cell Mol Med.** 2014 Jul;18(7):1344-57.

Kai-Wei Hsueh, Shu-Ling Fu, Chirn-Bin Chang, Yu-Ling Chang, Chao-Hsiung Lin. "A novel Aurora-A-mediated phosphorylation of p53 inhibits its interaction with MDM2" **Biochim Biophys Acta.** 2013 Feb;1834(2):508-15.

Iok In Christine Chio, Masato Sasaki, Danny Ghazarian, Juan Moreno, Susan Done, Takeshi Ueda, Satoshi Inoue, Yu-Ling Chang, Nien-Jung Chen and Tak Wah Mak. "TRADD Contributes to Tumor Suppression by Regulating ULF-dependent p19Arf Ubiquitination." **Nat Cell Biol.** 2012 May 6;14(6):625-33.

PUBLICATION-SUBMITTED

Yu-Ling Chang, Maura Rossetti, Hera Vlamakis, David Casero, Gemalene Sunga, Nicholas Harre, Shelley Miller, Romney Humphries, Thaddeus Stappenbeck, Kenneth Simpson, R. Balfour Sartor, Gary Wu, James Lewis, Frederic Bushman, Dermot P.B. McGovern, Nita Salzman, James Borneman, Ramnik Xavier, Curtis Huttenhower, Jonathan Braun. "A screen of Crohn's disease-associated microbial metabolites identifies ascorbate as a novel metabolic inhibitor of activated human T cells." *Under review (Mucosal Immunology)*

PRESENTATION

Su1873 Identification of IBD-Related Microbial Metabolites Affecting Human Th17 Differentiation. In: Digestive Disease Week, AGA Abstracts, Apr. 2016

Crohn's Disease-associated microbes produces ascorbate to regulate Human T cell responses. In: G. Oppenheimer Center for Neurobiology of Stress and Resilience Annual Symposium, Jan. 2017

AWARDS AND HONORS

Poster of Distinction Award at G. Oppenheimer Center for Neurobiology of Stress and Resilience Annual Symposium, 2017

Taiwanese Studying Abroad Scholarship, Taiwanese Government, 2014-2016

Fellowship, Molecular Biology Institute Graduate Program, 2013-2017

CHAPTER 1

Introduction

Gut microbiota associated with human health and disease has been appreciated in the past decade. It is not surprising that gut microbiota is linked to the local inflammatory conditions in the gut, such as irritable bowel syndrome[1] and inflammatory bowel diseases(IBD)[2]. There are also intriguing studies identifying the roles of gut microbiota in regulating systemic physiological conditions, such as obesity[3,4], cardiovascular disease[5], and rheumatoid arthritis[6]. Furthermore, the emerging science investigating the gut microbiome highlighted the link of gut-brain axis leading to major depression[7,8] and autism spectrum disorder[9,10]. Like human genome presenting individual identities, gut microbiome becomes the other critical factor to control human health.

Metagenomics analyses impute the potential microbial products

The human gut is colonized by a remarkably diverse microbial community. The advent of high-throughput DNA sequencing technology enables us to assay taxonomic composition and understands the whole genome contents for human microbiome[11]. An important insight is that the functional gene profiles predicted from metagenome are relative stable among healthy individuals despite highly divergent taxonomic composition[12,13]. This suggests that the functional activity of microbiota is the important indicator for physiological conditions. And, that microbial products, the net result of myriad metabolic interactions, could be leading mediators through which bacteria dictate disease.

Computational analyses of metagenome have suggested many ways in which microbiota can be associated with diseases; however, this raised an important issue of how to investigate the causality and mechanism of disease. Major studies of microbiome research in Crohn's diseases (CD) reported the altered metagenomes associated with CD states, which uncovered associations of overall metagenomic diversity and specific metagenes associated with CD,

together with the metabolic pathways and predicted microbial products altered in CD (Table. 1-1)[14-16]. The fundamental question is whether these observed microbial metabolic pathways are the cause of the immunological malfunction and whether there are key microbial products driving immune responses.

The prediction of microbial products from metagenome highly relies on the depth of sequencing and the algorithmic approaches to infer the functional and metabolic potentials of a microbial metagenome. One intriguing approach is on the concept and bioinformatics identification process for bacterial genomic clusters (BGCs), which are the clusters of chromosomally adjacent genes encode the enzymes and transporters that are necessary to produce and export a specialized metabolite. The prediction uses raw initial shotgun genomic sequencing acquisitions to assemble the reference genomic libraries; a PCR-based sequence tag approach can be further used to detect known gene clusters[17,18]. However, community metagenomes often comprise short reads from hundreds of different organisms, which are relatively challenging to assemble comparing to single-organism genomes[19].

In order to avoid the need for assembly of metagenomic reads, the pipeline of HMP Unified Metabolic Analysis Network2 (HUMAnN2) has been introduced to use the current databases of orthologous protein family categories[20]. This computational methodology enabled the efficient tool to infer presence/absence and abundance of microbial products in a microbial community from metagenome.

The reservoirs of microbial metabolites

Understanding the metabolic processes in microbiota enables the better prediction of microbial functional capacities from metagenome. The metabolomics measurement of microbial metabolites by mass spectrometry and nuclear magnetic resonance allowed the direct evidence

of the presence of secreted microbial products. Based on their metabolic processes and pathways, the metabolite repertoires of human gut microbiome can be categorized into three groups.

Primary metabolites are directly produced by the catabolic and anabolic reactions involved in cell growth, development, and reproduction. Large portions of these metabolites are either by-products or intermediates of energy metabolism, which can have a potential to influence host energy metabolism. Experimentally, short-chain fatty acids and tryptophan metabolites are the primary microbial metabolites reported to affect host biology[21,22].

The second group comprises secondary metabolites, which are not essential to the survival of an organism but have a broad spectrum of bioactivities, such as antibiotics, chemotherapeutics, or insecticides. The genes encoded enzymes that produce secondary metabolites are identified in bacterial genomic clusters (BGCs) which often acquired by horizontal gene transfer. Phenazine, for example, is a biologically well-characterized secondary metabolite that plays an important role in quorum sensing and biofilm formation in *Pseudomonas aeruginosa*[23,24]. Several studies developed computational platforms for gene cluster analysis to identify the microbial isolates produced phenazine and discover novel phenazine metabolic pathways in microbial genomes[18,25]. The analysis of gene clusters becomes an important bioinformatics tool to explore novel isolates produced secondary metabolites and predict the potential architectures for host-microbiota interactions[26].

The third group comprises the metabolites of non-microbial origin. These metabolites start as host-derived or diet-derived metabolites, which are modified by microbial enzymes in mucosal environments such as bile acids[27,28], or the small molecules introduced from the external environment (exposome) which are chemicals generated through exposure to food, medication, or environment[29].

As the knowledge in microbial metabolites expands, we may ultimately identify metabolic biomarkers for diagnostic and develop the novel therapeutic approach. The direct detection of metabolites from microbial samples showed that microbial metabolites are remarkably altered in CD disease[30,31]. Alteration of the metabolic profiling at multiple compartments in germ-free mice comparing with the wild-type mice suggested the role of gut microbiota influencing metabolism in host bodies[32]. Therefore, it is instructive to focus on microbial metabolites which represent the complex functional capacities of microbiota[33] and study their direct roles in regulating host physiology.

The bioactive roles of microbial metabolites in host immunity

Microbial metabolites are believed to be a key class of modulators through which microbiota directly interact with the host. Although a handful of bioactive metabolites were reported (Table 1-2), the large numbers of predicted metabolites suggest that myriad bioactive metabolites remain to be discovered[11].

One validated category of microbial metabolites influencing local cellular immune functions is short-chain fatty acids (SCFAs). SCFAs are produced by colonic microbiome when digesting fiber-containing food[34]. They are also produced in the process of bacterial fermentations or generated as byproducts from nitrogen recycling and amino acids metabolisms[35]. Selective molecules in SCFAs profoundly induced the anti-inflammation, by increasing regulatory T cell responses[21], inducing anti-inflammatory responses in macrophages via the inhibition of histone-deacetylase (HDAC)[36]. Additional studies demonstrated that SCFAs also regulated other non-immune cells, including the suppression of the intestinal epithelial cell proliferation[37]. These data suggested the role of SCFAs in local tissues.

As a second molecular class, indole and its metabolites affect both local and remote physiological responses. Indole and indole-3-aldehyde (I3A) are catabolized from tryptophan by commensal bacteria, *Lactobacillus* species or *Clostridium sporogenes*, and both metabolites are the agonists for the aryl hydrocarbon receptor (AhR). Through the AhR activity in either regulatory T cells or NKp46+ILCs at the mucosal site, indole derivatives can regulate the mucosal IL-22 production[38]. Dietary tryptophan metabolized by gut microbiota was also reported to remotely modulate astrocyte activity in CNS[39]. These studies demonstrated the potential roles of gut microbial metabolites in remote or systemic immunological regulations.

Conclusion

This introduction conveys how the assessment of host-microbiome physiology is evolving from microbial taxonomy to the products of the microbiome. This shift creates the opportunity to form mechanistic hypotheses on how the microbiome affects host physiology and disease biology. This thesis reflects this evolution.

In Chapter 2, we take advantage of a robust, established metagenomics dataset for Crohn's disease, to take up the challenge of assessing microbial function in regulating immune function in CD. We hypothesized that there are numbers of unidentified microbial metabolites as modulators to regulate host-microbiota interaction and to drive disease progression. The study established a novel pipeline for identifying candidate microbial products of the disease-associated microbiome, and subjecting them to validation of bioactivity in a central host trait of CD (CD4+ T cell activation and Th17 induction).

In Chapters 3 and 4, we studied the association of gut microbiota in a life-threatening and previously unstudied immunologic disease, SSc. We believed that gut microbiota may play an important role in affecting gastrointestinal tract symptoms in SSc. Understanding the association

of gut microbiome with symptoms may suggest the use of either antibiotics or probiotics in SSc and ultimately provide better quality of life in patients. The outcomes of these studies each have revealed new insights about the gut microbiome in SSc.

To conclude, this thesis has reinforced the roles of gut microbiota in the human diseases, and establishes a generalizable strategy of moving from taxonomic to functional microbiome analysis in disease biology.

Tables

Table 1-1. Alteration of microbial metabolic pathways in CD patients versus controls.

Study	Cohort	Samples	Increased in CD	Decreased in CD
Morgan et al. 2012[14]	PRISM	Biopsies	Nitrogen metabolism Sulfur metabolism Fructose and mannose metabolism(iCD) Riboflavin metabolism(iCD) Bacterial secretion system(iCD) Cysteine and methionine metabolism (iCD) Pentose phosphate pathway(iCD)	Butanoate metabolism (iCD) Lysine biosynthesis
Tong et al. 2013[15]	MLI	Lavage	Oxidative responses Glycan metabolism	Amino acid metabolism
Gevers et al. 2014[16]	RISK	Biopsies, Feces	Glycan biosynthesis and metabolism Glycerophospholipid and Lipopolysaccharide metabolism	Amino acid metabolism – Asparagine synthase – Glutamate synthase Lipid metabolism

Table 1-2. List of experimentally proved bioactive metabolites

Metabolites		Microbial genera/species[40]	Host	Cell type/Model	Molecular target/ Mechanism of action	Ref
SCFAs	Acetate, Propionate, Butyrate	*Most anaerobic gut bacteria studied produce acetate*	Human	Primary monocytes	Induce prostaglandin E2 (PGE-2), cytokine IL-10, and chemokine MCP-1 production	41
	Butyrate	*Bacteriodes, Ruminococcaceae, Lachnospiraceae*	Mouse	Treg cells	Induce the differentiation of Treg cells in the colonic lamina propria via inhibiting histone deacetylases	21,34
			Mouse	Macrophages	Inhibit the pro-inflammatory mediators in colon macrophages	36
			Human Mouse	Intestinal epithelial cells	Suppress the proliferation of epithelial cells	37
Indole derivatives (Trp)	Indole-3-aldehyde (I3A)	*Lactobacillus*	Mouse	*Ido1*$^{-/-}$ mice +dietary sup. with Trp	Induce aryl hydrocarbon receptor (AhR)-dependent Il-22 transcription	38
	Indole	*Lactobacillus, Bifidobacterium, Bacteroides fragilis, Parabacteroides, Clostridium*	Mouse	*Ido1*$^{-/-}$ mice	Activate AhR	38
				Colonic L cells	Modulate the secretion of glucagon-like peptide-1 (GLP-1)	42
	Indole sulfate		Human	THP-1 cells and HUVEC cells	Induce the leukocyte-endothelial interaction via E-selectin	43
Choline derivatives	Trimethyl-amine (TMA), Trimethyl-amine N-oxide (TMAO)	*Faecalibacteriu, Bifidobacterium, Clostridium, Proteus*	Mouse	Dietary sup. with choline or TMAO in germ-free mice	Upregulate the multiple macrophage scavenger receptors linked to atherosclerosis and cardiovascular disease	*44*
Bile acid	Cholic acid (CA), Deoxycholic acid (DCA)	*Lactobacillus, Bifidobacteria, Bacteroides, Clostridium*	Mouse	Germ-free mice	Microbial control the pool size of bile acid via FXR	28

CHAPTER 2

A screen of Crohn's disease-associated microbial metabolites identifies ascorbate as a novel metabolic inhibitor of activated human T cells

Abstract

Microbial metabolites are an emerging class of mediators influencing CD4+ T cell function. To advance the understanding of direct causal microbial factors contributing to Crohn's disease (CD), we screened 139 predicted CD-associated microbial metabolites for their bioactivity on human CD4+ T cell functions induced by disease-associated T helper 17 (Th17) polarizing conditions. We observed 15 metabolites with CD4+ T cell bioactivity, 3 previously reported and 12 unprecedented. Mechanistic assessment of one novel microbe-derived metabolite, ascorbate, revealed apoptosis of activated human CD4+ T cells associated with selective inhibition of glycolytic energy metabolism. These findings suggest a substantial rate of relevant T cell bioactivity among CD-associated microbial metabolites, and evidence for novel modes of bioactivity, including targeting of T cell energy metabolism.

Introduction

Abnormalities of intestinal microbial composition are strongly associated with inflammatory bowel diseases[45-47]. In the healthy state, much evidence suggests that small molecule products of the microbiota (microbial metabolites) are important modulators of normal host physiology[33,48,49]. Among these, the literature includes only a handful of bioactive microbial metabolites, and few metabolites directly affect cellular targets of CD relevance[21,37]. In CD, previous multi'omic studies suggest an association with alteration of microbial energy metabolism, enrichment of carbohydrate, lipid, cofactor and vitamin metabolism, and depletion of amino acid biosynthesis[16,50,51]. Certain of these microbial metabolites are known to influence important aspects of host metabolism[52], activity of intestinal epithelial stem cells[37], and regulation of diverse immune cellular responses[21,34,36]. For example, short chain fatty acids (SCFAs) such as butyrate are known for their mechanisms of action on immune responses through the inhibition of histone deacetylase[36]. Tryptophan catabolites induce T cell cytokine expression through a process involving ligand-activated aryl hydrocarbon receptor, AHR[53]. However, most inferred microbial metabolites remain untested for bioactivity on disease-relevant host responses, and unresolved for their mechanisms of action.

In genetically susceptible individuals, CD results from inappropriate immune responses to the unique intestinal microbiota[54,55]. Functional assessment of CD genetic loci revealed the involvement of CD4+ effector T helper signaling and the Th17-IL23 pathway[56,57]. Both clinical studies and pre-clinical colitis studies strongly link Th17 cell responses and the IL-17 pathway to the pathogenesis of CD[58-62]. Given their pathogenic role, Th17 cells became a therapeutic target[63], but clinical trials revealed unexpected inefficacy and disease exacerbation[64-68]. Subsequent work has uncovered dichotomous roles of the IL-17 pathway in CD progression, reflecting the

differential roles of Th17 cells in disease induction and resolution[69]. This affirms the importance of Th17 in CD, but suggests complexity in how the intestinal microbiome and its products may be targets and modifiers of mucosal effector T cell function.

To determine how CD-associated microbiome and its products might modulate CD4+ T cell responses, we tabulated a set of microbial metabolites predicted to be differentially abundant in CD. We functionally screened these metabolites by adding them to CD4+T cell cultures together with Th17-polarizing cytokines. This conceptually recapitulates the *in vivo* setting, in which lymphocytes simultaneously encounter bacterial products and inflammatory cytokines in the sub-mucosa or draining lymph nodes. Our screen revealed a substantial frequency (11%) of metabolites that modulated T cell activity, with different patterns of inhibition and augmentation. Detailed study of one microbe-derived metabolite, ascorbate, demonstrated potent inhibition of activated effector $CD4^+$ T cells associated with inactivation of glycolytic energy metabolism. These findings support the idea that disease-associated microbial metabolites may have a significant role in the pathogenesis of CD.

Materials and Methods

Human blood sample collection and T cell isolation

We recruited healthy donors at UCLA Medical Center, and collected blood samples according to an IRB protocol approved by the institutional review committee of University of California, Los Angeles. Human CD4+ T cells were isolated using Rosettesep Human CD4+ T cell enrichment cocktails (StemCell^{TM} technology) following manufacturer's instructions. Purified CD4+ T cells were confirmed to be > 95% pure by flow cytometry and the T cell subset composition was monitored (Figure S4-1). We cultured T cells with soluble anti-CD28 (0.2 ug/ml) antibodies on 96-well plates pre-coated with anti-CD3 antibodies (1ug/ml). Interleukin 2 (IL-2) was added at 4.8 U/ml (2 ng/ml). We induced Th17 differentiation with 50 ng/ml IL-6, 50 ng/ml IL-1β and 50 ng/ml IL-23 [70]. We refed T cells at day 7 with medium containing the same Th17 polarizing cytokines and further cultured for another 7 days. We collected supernatants for multiplex Luminex cytokine analysis, and analyzed T cells by intracellular flow cytometry.

Cell survival assay

The survival of metabolites-conditioned cells was examined by SYTOX™ Red dead cell stain (Molecular Probes, Cat.S34859), which is a nucleic acid stain that penetrates cells with compromised plasma membranes. We used the SYTOX™ Red dead cell stain after cell surface marker staining, and detected SYSTOX Red levels by flow cytometry.

Intracellular Flow Cytometry

At the end of the 14-day culture, T cells were stimulated with Leukocyte activation cocktail with BD GolgiPlug (BD Biosciences, Cat. 550583) for 4 hours and washed in FACS buffer. BV650-conjugated CD3 (Clone OKT3) and PE-CF594-conjugated CD4 (Clone OKT4) antibody

were used to detect these cell surface markers. We performed intracellular fixation and permeabilization with a commercial buffer set (eBiosciences, Cat. 88-8824). The following antibodies were used to detect intracellular cytokines: APC-conjugated IL-17A (Clone eBio64DEC17), FITC-conjugated IFNγ (Clone 4S.B3), PE-conjugated IL-10 (Clone JES3-9D7), and PE-Cy7-conjugated IL-4 (Clone 8D4-8) were purchased from eBioscience or Biolegend).

Cell respiration measurements

We activated isolated CD4+ T cells on plates coated with 10 ug/ml anti-CD3 and 10ug/ml anti-CD28, and stimulated in the presence of 10 ng/ml IL-2 as the activated condition, and 10 ng/ml IL-7 as the resting condition. After 48 hours, activated or resting cells were collected, washed with unbuffered DMEM (2 g/L glucose), and seeded into 0.2% gelatin-covered 24-well plates (XF24 from Seahorse) (1 million cells/well) for OCR and ECAR measurement. We measured cellular OCRs and ECARs as described previously[71]. Oxygen consumption was blocked by oligomycin (0.75 uM), an ATP synthase inhibitor; the ionophore FCCP (carbonyl cyanide-p-trifluoromethoxyphenylhydrazone; 1uM) assayed maximal respiratory capacity of mitochondria, whereas rotenone (3 uM), a mitochondrial inhibitor, was used to block mitochondrial respiration. Two independent experiments were performed, and 6 technical replicates per experiment were measured to calculate the mean ± SEM for each time point.

Bioinformatic prediction of CD-associated microbial metabolites.

Using a previously reported PRISM dataset[14] of microbial composition at the level of OTUs from 16S rRNA gene amplicon sequencing data, we tabulated microbial count data for healthy and CD subjects at the subject level, and converted them to imputed metagene data using PICRUSt[11]. We then determined the relative abundance of enzymatic pathways using HUMAnN2[20], with pathway end-products identified using UniRef and MetaCyc databases. Making the assumption

that metabolite abundance is the sum of abundances of the generating enzymes, we calculated the predicted abundance of each metabolite using the average of most abundant route through each pathway.

Disease-associated metabolites were determined using a multivariate model comparing predicted relative compound levels in health versus CD subjects. Specifically, we performed a generalized linear mixed model (glmmPQL, implemented in R), correcting for age and medications (antibiotics, immunosuppressants and anti-inflammatory agents). Disease-associated metabolites from this analysis were ranked by multi-testing corrected q value. From this raw list, the top molecules for which commercial sourcing could be identified were tabulated (Table S2-1), and procured for testing.

For prediction of ascorbate-producing microbiota, we used the reverse of this bioinformatics pipeline, with the following modifications. Microbial metabolic pathways and enzymes involved in ascorbate biosynthesis were defined using KEGG and literature reports. For those enzymes annotated in KEGG or MetaCyc databases, we performed a prediction of microbiota genomically bearing them using PICRUSt-precalculated gene content. We filtered these results for human fecal-associated OTUs, generating lists of OTUs bearing various numbers of ascorbate pathway genes. Finally, we filtered for OTUs validated for significant differential association in CD versus healthy subjects.

The computational resources for all analytic steps are available at https://bitbucket.org/biobakery/biobakery/wiki/Home.

Metabolite preparation

All chemical were of cell culture or higher grade, and purchased from Sigma Aldrich. The majority of chemicals were dissolved in aqueous solution; water-insoluble metabolites were dissolved in DMSO.

Luminex multiplex cytokine production

We measured cytokines from CD4+ T cell culture supernatants using Luminex. The following 11 cytokines were included in the Luminex panels and acquired on a Luminex® FLEXMAP 3D: Th17 cytokines, IL-17A, IL-17F, and IL-22; Th1 cytokines, TNFα and IFNγ; and Th2 cytokines, IL-5 and IL-13; Treg cytokine, IL-10; other related cytokines, IL-9, IL-21, and MIP-3a. We normalized the Luminex readouts for the cell count values.

Detecting ascorbate in bacterial culture

Pseudomonas aeruginosa included two strains, PA-01 and PA-14, obtained from the ATCC. *Burkholderia cepacia* included three clinical isolates (2-1, 3-1, 3-2; source strains ZK2853-ZK2862) from two cystic fibrosis patients[72], a gift from the collection of Robert Kolter (Harvard Medical School). Bacteria were cultured in 1 mL Luria-Bretani broth at 35 ± 2^{o}C under aerobic condition, and the bacterial cell numbers were monitored by McFarland standard [73]. After pelleting bacterial cells (5,000 rpm, 10 min), we filtered the supernatants using centrifugal filters per manufacturer instructions (Corning® Costar® Spin-X® centrifuge tube filters, 0.22μm). We measured ascorbate levels in supernatant filtrates by an enzyme-based assay (Abcam, ab65656).

Statistical analysis

To estimate contributors of experimental variability and control for potential technical effects, we performed principal variance component analysis (PVCA). The technical factors assessed included metabolite, experimental run, donor, and technical replicates in different plates. We used

PCA to reduce data dimensionality data from the 11 cytokines, and to maintain majority of the variability in the data. We used VCA to fit a mixed effect model using subject-specific factors as random effects to estimate and partition the total variability. Figure S2-3 shows one example. We ran PVCA for every experiment to ensure the stability of technical variation.

We used multivariate analysis of variance (MANOVA) to identify putative bioactive metabolites. In each condition treated with a specific metabolite concentration, we compared the levels of the 11 cytokines to those found in the vehicle control condition. We defined a metabolite as bioactive if p-values < 0.05; we confirmed bioactivity in a different healthy donor to confirm metabolite bioactivity.

To further validate biological effects and define specific cytokine responses to bioactive metabolites, we employed two-way analysis of variance (ANOVA) with metabolite and donor factors using the incorporated data from at least two independent experiments. We used the Holm-Sidak's post-test to correct for multiple comparisons. We defined statistical significance as adjusted p-values < 0.05.

We generated cytokine heatmaps from Luminex cytokine data by calculating the mean of log2 fold change between metabolite-untreated versus metabolite-treated samples (treated/untreated). We used the heatmap.2 function in gplot package to generate dendrogram and colorplot in R version 3.2.4.

We performed statistical analyses with Prism software (GraphPad) or using the R program (R version 3.1.1). We used the nonparametric Mann-Whitney test to compare two experimental conditions. We defined statistically significant results are indicated ($*p < 0.05$, $***p < 0.001$). We displayed error bars using mean ± SEM calculated from 3 or more experimental replicates.

Results

Screening of microbial metabolites identifies modulators Th17 responses

We identified a panel of fecal microbial metabolites predicted to be differentially produced by a CD-associated versus healthy microbiome using the previously reported PRISM cohort[14] (see Materials and Methods). From this panel, we selected a subset of 139 commercially available metabolites for further testing (Table S2-1). In Figure 2-1a, we summarize the workflow for the bioactivity screen of these metabolites on human Th17 polarization cultures. Briefly, purified human blood CD4+ T cells were cultured with different concentrations of each metabolite in the presence of standard 14 day Th17 polarization culture conditions (anti-CD3, anti-CD28, IL-6, IL-1β, and IL-23)[70]. We assessed Th17 polarization by levels of secreted cytokines, and validated by the percentage of Th17 cells by flow cytometry (Figure S2-2).

We first qualitatively and quantitatively screened for metabolite conditions that caused toxicity by the endpoints of cell death and reduced cell numbers. We tabulated metabolite conditions that significantly increased toxicity, and excluded them from downstream cytokine analyses. Notably, we did not observe any metabolites that significantly enhanced cell numbers (data not shown).

The primary screening revealed eighteen bioactive metabolites based on a relaxed cut-off in multivariate analysis of variance (Figure 2-1b). We performed a confirmatory screening in independent healthy individuals to validate the primary screen and refine dose-dependent responses (two-way ANOVA). Figure 2-1c summarizes results for all screened metabolites with respect to toxicity and secreted cytokine production. Principal variance component analysis (PVCA) demonstrated that the majority (57%) of variability was contributed by metabolite treatment, while only 10% of variability was attributed to a donor effect (Figure S2-3). Sixteen

metabolites (11%) were toxic based on cell death at all tested conditions in both donors. A total of fifteen metabolites (11%) displayed reproducible effects on T cell cytokine production, and were classified as *validated bioactive*. Three metabolites (2%) showed inconsistent cytokine changes between two individual experiments, and were classified as *putatively bioactive*.

Differential metabolite effects on T cell cytokine profiles

The fifteen validated bioactive metabolites varied in their pattern of effects on cytokine production (Table 2-1). Three of these metabolites were previously reported for analogous effects on T cells[34,74,75]. Our findings confirmed that glutamate enhances IFNγ production, butyrate inhibits IL-17A but increases IL-10 production, and arachidonic acid inhibits IFNγ production (Figure S2-4). Strikingly, the action of the other twelve metabolites are unprecedented, to our knowledge, for an effect on T cell cytokine production.

We observed distinct patterns of altered cytokine production. Some metabolites induced global changes in cytokine production. GDP-mannose decreased the production of eight cytokines (pan-inhibition) at the dose range of 0.1 to 1 mM (Figure 2-2a). Three other metabolites (HMG-CoA, CoA, and mevalonate) similarly decreased the production of all tested cytokines at high concentration (Figure S2-5a). Instead of pan-inhibition, L-homocysteine (1 mM) augmented production of most tested cytokines (Figure S2-5b).

Other metabolites showed selective effects on Th17 cytokines. For example, ascorbate decreased Th17-related cytokine production (Figure 2-2b), whereas there was no statistically significant change for IL-10, IL-5, and IL-13 production with ascorbate treatment. Three metabolites at indicated conditions (0.1 mM glutaryl-CoA, 1mM hexanoyl-CoA, and 1 mM isopentenyl diphosphate) enhanced the production of IL-17A and IL-17F (Figure S2-6a). Other

metabolites selectively altered production of certain non-Th17 cytokines. Oleic acid and linoleic acid increased IFNγ in a dose-dependent fashion (Figure 2-2c and 2-2d); In contrast, 0.1 mM pyridoxine phosphate decreased the production of TNFα, IFNγ, and IL-10 (Figure S2-6b).

Figure 2-2e summarizes the relative cytokine secretion for the twelve novel bioactive metabolites. Unsupervised hierarchical clustering classified the metabolites into patterns of inhibitory and augmented responses. Among them, six metabolites are fatty acid intermediates or end products based on the Human Metabolome Database (HMDB). Ascorbate is the only microbe-derived metabolite that humans cannot synthesize, which drew our attention to this metabolite for deeper validation and mechanistic assessment.

Ascorbate inhibits T cell cytokines in both polarized and non-polarized conditions

To further validate ascorbate bioactivity under Th17 polarizing conditions, we performed four additional biological replicates, and found that ascorbate consistently inhibited the production of Th17 cytokines. Flow cytometry analysis demonstrated that ascorbate induced a dose-dependent reduction in IL-17A-producing CD4+ T cells, including both IL-17A single positive T cells and IFNγ/IL-17A double positive cells (Figure 2-3a). Measurement of cytokine secretion by Luminex confirmed the ascorbate inhibitory effect. All the cardinal Th17 cytokines (IL-17A, IL-17F and IL-22) were reduced by more than 50% after 100 uM ascorbate treatment (Figure 2-3b). These data confirm that ascorbate inhibits the production of Th17 and Th17-related cytokines under Th17-polarizing conditions.

To investigate whether this inhibiting effect was specific to the Th17 subset, we cocultured CD4+ T cells with ascorbate under non-selective activating conditions (anti-CD3/CD28 without Th17-polaring cytokines (Figure 2-3c). In the presence of ascorbate, IFNγ, IL-4, and IL-17A-

producing CD4+ T cells were all reduced, whereas IL-10 producing cells were preserved. Both IL-17A-single positive cells and IFNγ/IL-17A-double positive cells were reduced in the presence of ascorbate. We concluded that ascorbate not only inhibited Th17 cells, but also reduced the activation of most other subsets of effector T cells (excluding IL-10-producing T cells).

Ascorbate selectively inhibits activated but not resting CD4+ T cells.

We next determined whether ascorbate regulated all CD4+ T cells, including resting T cells and activated T cells. We cultured CD4+ T cells in resting and activating conditions for 48 hours in the absence or presence of 100 uM ascorbate. Ascorbate selectively inhibited the upregulation of the activation marker CD154 in activating conditions (Figure 2-4a and 2-4b). AnnexinV and 7-AAD staining further demonstrated that activated but not resting T cells underwent programmed cell death after 48 hours of ascorbate treatment (Figure 2-4c). These data illustrated that ascorbate selectively inhibited activated versus resting T cells.

Ascorbate selectively inhibits glycolysis-dependent energy production in activated T cells

To study the molecular mechanisms involved in ascorbate-induced inhibition, we focused our analysis on the regulation of energy metabolic pathways. It is well known that TCR activation, specifically CD28 signaling, directly controls the metabolic switch to glycolysis by up-regulating the expression of glucose transporter 1, and remodeling activity of downstream biochemical pathways to favor glycolytic versus oxidative phosphorylation energy production [71,76,77]. In contrast, resting CD4+ T cells rely on fatty acid oxidation to generate energy through oxidative phosphorylation[76]. Since glucose transporter 1 (a transporter for the oxidized form of ascorbate) is up-regulated in activated T cells[71], we reasoned that an increased level of ascorbate is selectively transported in activated versus resting T cells. We therefore tested the possibility that ascorbate might selectively inhibit glycolysis-dependent activated T cells.

Consistent with previous reports, we observed that activated T cells augment their glycolysis compared to resting T cells, demonstrated by a significant increase in extracellular acidification rate (ECAR) (Figure 2-4d). Upon ascorbate treatment, we observed a significant reduction in ECAR by activated but not resting T cells, indicating that ascorbate decreased the glycolytic rate in activated T cells (Figure 2-4d and 2-4e). These data support the prediction that a selective inhibitory effect of ascorbate on activated T cells was related to its effect on glycolytic energy metabolism.

Ascorbate production by microbial taxa associated with CD

To investigate whether gut microbes can produce ascorbate, we performed a reverse metagenomic search for the bacterial genomes that contained ascorbate biosynthesis enzymes. The metabolic pathways utilized by microbes to produce ascorbate *de novo* are largely unknown, but biotechnology has suggested two microbial pathways to manufacture L-ascorbate from D-glucose[78]. We therefore investigated the chemical reactions in these processes, and tabulated the metabolic enzymes and corresponding metagenes involved in ascorbate synthesis reactions. Using this tabulation, we performed a reverse metagenomics search of human fecal microbiota bearing some or all of the genes in the archetypal D-sorbitol and 2,5-diketo-D-gluconic acid pathways of ascorbate biosynthesis (Table S2-2 and S2-3). We identified 5 Operational Taxonomic Units (OTUs) bearing the ascertainable metagenes of the D-sorbitol pathway, all of them in the CD-associated genus *Burkholderia* (Figure 2-5a). We also identified 120 OTUs for the 2,5-diketo-D-gluconic acid ascorbate pathway; of these, the most prevalent taxa belonged to the CD-associated genera *Pseudomonas* and *Erwinia* (Figure 2-5b). To directly assess ascorbate production by these microbes, we assembled a collection of three clinical isolates of *Burkholderia* and two strains of *Pseudomonas aeruginosa*. After aerobic culture to stationary phase, ascorbate production was

detectable in both *Pseudomonas* strains (Figure 2-5c), but not in the *Burkholderia* isolates (data not shown). These findings indicate that certain CD-associated microbiota are proficient at ascorbate biosynthesis, and point to the 2,5-diketo-gluconic acid pathway as a potential mechanism for this production.

Discussion

By screening microbial metabolites using primary human CD4+ T cells in Th17 polarizing culture conditions, we identified three known and twelve novel microbial metabolites with the potential to module T cell cytokine production. A detailed study of one novel microbial-derived metabolite, ascorbate, revealed inhibition of Th17 and activated T cell formation by a mechanism involving blockade of glycolytic energy metabolism. We metagenomically inferred and functionally validated examples of CD-associated microbiota proficient for ascorbate production. These findings suggest a substantial frequency of relevant T cell bioactivity among CD-associated microbial metabolites, and evidence for novel modes of bioactivity, including targeting of T cell energy metabolism.

This study emphasizes the importance of environmental metabolic cues to modulate CD4+ T cell functions, notably via differentiation and activation of specific subsets of CD4+ T cells with distinct cytokine expression patterns[79,80]. Three bioactive metabolites identified in our study confirmed previous observations of these metabolites as selective mediators of such differentiation and activation[34,74,75]. The bioactivity of twelve additional metabolites for human CD4+ T cells did not have literature precedents to our knowledge. The diversity of their effects on T cell cytokine function suggests that their mechanistic assessment may reveal new insights on the action of microbial metabolites in CD4+ T cell function, and the role of these actions in disease biology.

This study uncovered a previously unknown inhibitory action of ascorbate on activated CD4+ T cells. A recent report demonstrated that ascorbate selectively kills glycolysis-dependent cells (epithelial carcinoma) by its inactivation of glyceraldehyde 3-phosphate dehydrogenase, an apex enzyme in glycolytic energy metabolism[81]. The present study confirmed the selective effect of ascorbate on glycolytic energy production, and extended the observation to the CD4+ T cell

type. It is well known that the early phase of T cell activation induces a metabolic switch to and dependence upon aerobic glycolysis[71,76,77]. This provides a plausible mechanism by which ascorbate, via interruption of glycolytic energy production, may account for the selective functional inhibition and apoptosis of activated CD4+ T cells. However, a recent report observed that after completion of Th17 induction, these cells shift again to oxidative phosphorylation as their preferential energy source[82]. These findings suggest that the potential in vivo action of ascorbate on Th17 cells may be limited to the stages of Th17 induction and activation, and to their relevant anatomic sites (such as mesenteric nodes in the case of intestinal Th17 cells).

We also observed that ascorbate inhibits the production of IFNγ-, IL-4, and IL-17A, but not IL-10. Unlike other CD4+ T cell subsets, induced Tregs (which inhibit effector T cell responses in part through IL-10 release) do not substantially shift to glycolytic energy metabolism[76,77], and this may account for their resistance to ascorbate inhibition. Moreover, in mice ascorbate facilitates the demethylation of Foxp3 enhancer and promotes Treg cell function[83,84]. Taken together, these observations suggest that ascorbate as an environmental cue that may selectively inhibit effector CD4+ T cell responses.

An issue raised by these observations is the physiological role of dietary ascorbate versus local microbial ascorbate production. Unlike most eutherians, humans and primates require ascorbate intake, because they genetically lack the enzyme responsible for the last step in ascorbate synthesis, L-gulonolactone oxidase[85]. Previous work indicates that microbial ascorbate is not the major nutritional source for vitamin C in humans[86], suggesting the importance of dietary intake of ascorbate in human. However, dietary ascorbate is largely absorbed before transit to the colon, so local microbial production may be a significant contributor to luminal ascorbate in the distal intestine. The predicted ascorbate-producing genera are *Proteobacteria*, an uncommon intestinal

phylum most prevalent in the colon. *Pseudomonas aeruginosa* and related species are notable for their production of redox-active small molecules that modify the cellular redox state, acting as inter-microbial signals to control biofilm formation[23]. Ascorbate may represent an additional class of redox-active molecules involved in such inter-microbial interaction. Microbial genetics and ecologic studies are necessary to assess the control and bioavailability of ascorbate in the gut, and test its potential role in microbial ecology.

Clinically, ascorbate levels and the ratio of oxidized form (dehydroascorbic acid, DHA) versus reduced form (ascorbic acid, AA) of ascorbate are distinct in IBD patients versus healthy subjects [87]. Genetic variations in the sodium-dependent ascorbate transporter gene, *SLC23A1*, have been significantly associated with susceptibility to CD [88]. The disease variant of human *SLC23A1* results in lower transporter activity and decreased intracellular ascorbate. Decreased ascorbate transport limits the capacity to prevent oxidative tissue damage and impair recovery from mucosal injury [89]. However, the mechanistic basis of these disease-associated phenotypes has not yet been probed in experimental animal models, because except for primates, eutherians are proficient for ascorbate synthesis. Such studies, and testing the potential in vivo roles of microbial ascorbate on CD4+ T cells and CD, await genetic engineering of mice bearing the human-homologous L-gulonolactone oxidase deficiency.

Collectively, we screened a panel of predicted CD-associated microbial metabolites and identified 12 novel and 3 proven modulators of human CD4+ T cell function. We also demonstrated that a novel microbial-derived immune modulator, ascorbate, regulates host responses via inhibition of energy metabolism. Our results suggest that the balance of luminal metabolites could be important to maintain physiological CD4+ T cell functions of the intestinal mucosa.

Tables and Figures

Table 2-1. Validated bioactive metabolites.

Metabolite Name	HMDB_ID	Chemical Category (HMDB)
(S)-3-Hydroxy-3-methylglutaryl-CoA	HMDB01375	Fatty Acyls
Hexanoyl-CoA	HMDB02845	Fatty Acyls
GDP-mannose	HMDB01163	Purine nucleotides
Isopentenyl diphosphate	HMDB04196	Prenol lipids
CoA	HMDB01423	Purine nucleotides
L-Glutamate	HMDB00148	Amino acids, peptides, and analogues
L-Ascorbate	HMDB00044	Furanones
L-Homocysteine	HMDB00742	Carboxylic acids and derivatives
(R)-Mevalonate	HMDB59629	Fatty acids and conjugates
Glutaryl-CoA	HMDB01339	Fatty Acyls
Pyridoxine phosphate	HMDB01319	Pyridines and derivatives
Oleic acid	HMDB00207	Fatty Acid
Arachidonic acid	HMDB01043	Fatty Acid
Linoleic acid	HMDB00673	Fatty Acid
Butyric acid	HMDB00039	Fatty Acid

Figure 2-1

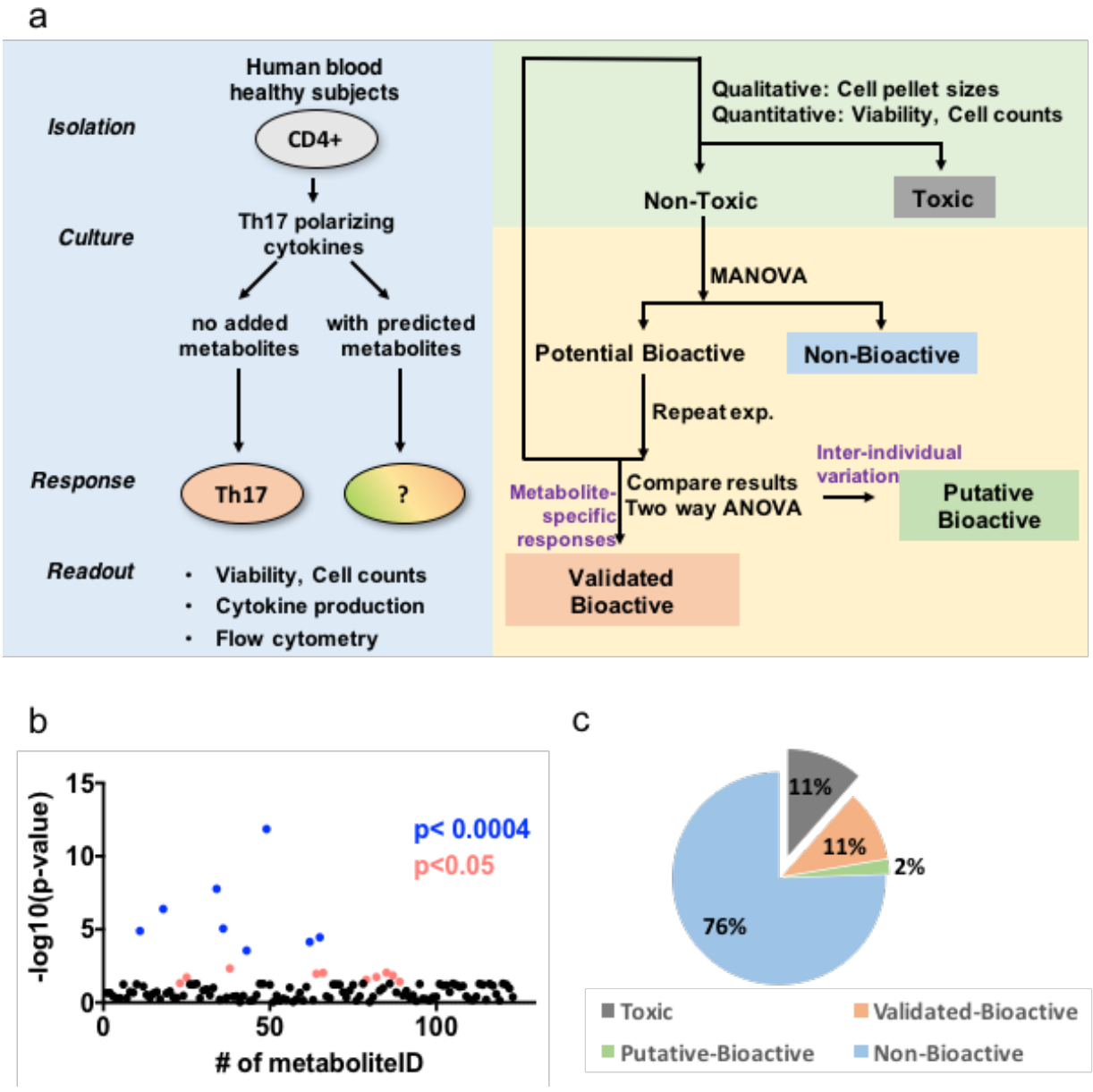
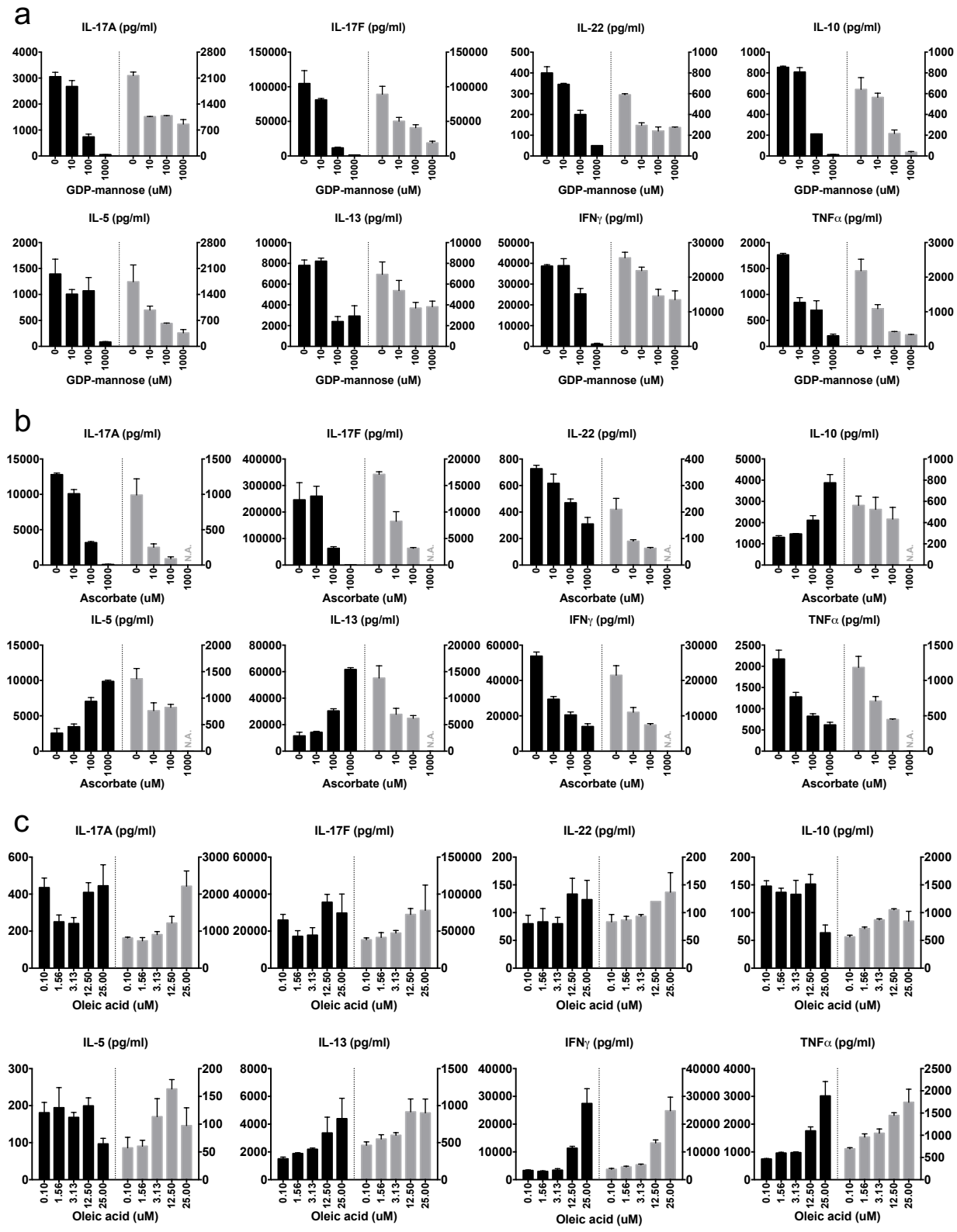


Figure 2-1. Microbial metabolites screen identifies bioactive metabolites modulating Th17 responses (a) Schematic of screening and analytic strategies. Toxic: cell death observed at all tested conditions; Potential Bioactive: achieved MANOVA statistical significance cut-off; Putative Bioactive: significant effects in first round but not second round of screening, potentially due to donor-specific variation; Validated Bioactive: reproducible effects in two individuals

(achieved statistical significance cut-off by two-way ANOVA test). (b) Aggregate cytokine response (negative logarithm of the association p-values calculated from multivariate analysis of variance) to candidate metabolites. $P<0.0004$ represents the significant threshold after Holm-Sidak correction for multiple comparisons (Blue dots). (c) Categories of response to metabolites (% of all tested metabolites).

Figure 2-2


a
IL-17A (pg/ml)
IL-17F (pg/ml)
IL-22 (pg/ml)
IL-10 (pg/ml)
IL-5 (pg/ml)
IL-13 (pg/ml)
IFNγ (pg/ml)
TNFα (pg/ml)
GDP-mannose (uM)
b
IL-17A (pg/ml)
IL-17F (pg/ml)
IL-22 (pg/ml)
IL-10 (pg/ml)
IL-5 (pg/ml)
IL-13 (pg/ml)
IFNγ (pg/ml)
TNFα (pg/ml)
Ascorbate (uM)
N.A.
c
IL-17A (pg/ml)
IL-17F (pg/ml)
IL-22 (pg/ml)
IL-10 (pg/ml)
IL-5 (pg/ml)
IL-13 (pg/ml)
IFNγ (pg/ml)
TNFα (pg/ml)
Oleic acid (uM)

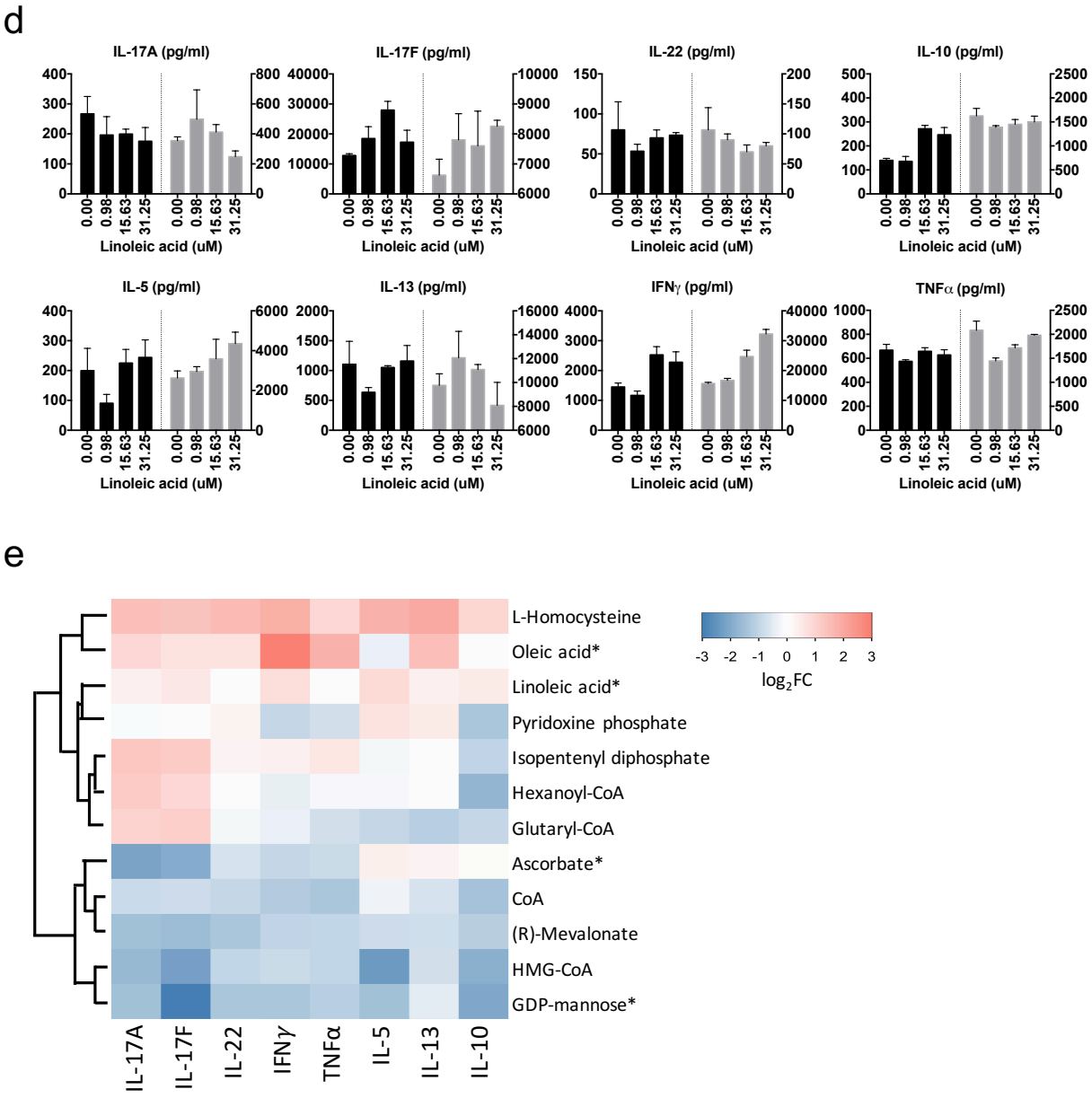


Figure 2-2. Representative bioactive metabolites altering CD4+ T cell cytokine secretion. (a) Pan-inhibition of cytokine secretion by GDP mannose. (b) Th17-specific cytokine inhibition by ascorbate. N.A.: condition that causes cell death and excluded. (c, d) Th1-specific cytokine augmentation by oleic acid (c) and linoleic acid (d). For each of the two donors (black and gray), we calculated mean ± SEM from three technical replicates. (e) Summary heatmap of multiplex cytokine expression data for the 12 novel validated-bioactive metabolites. We show metabolite-

treated samples (rows) at the optimal bioactive concentration. Asterisks (*) highlight metabolites with dose-dependent effects. CoA: Coenzyme A; HMG-CoA: 3-hydroxy-3-methyl-glutaryl-coenzyme A. Four metabolites with dose-dependent effects are marked with an asterisk (*).

Figure 2-3

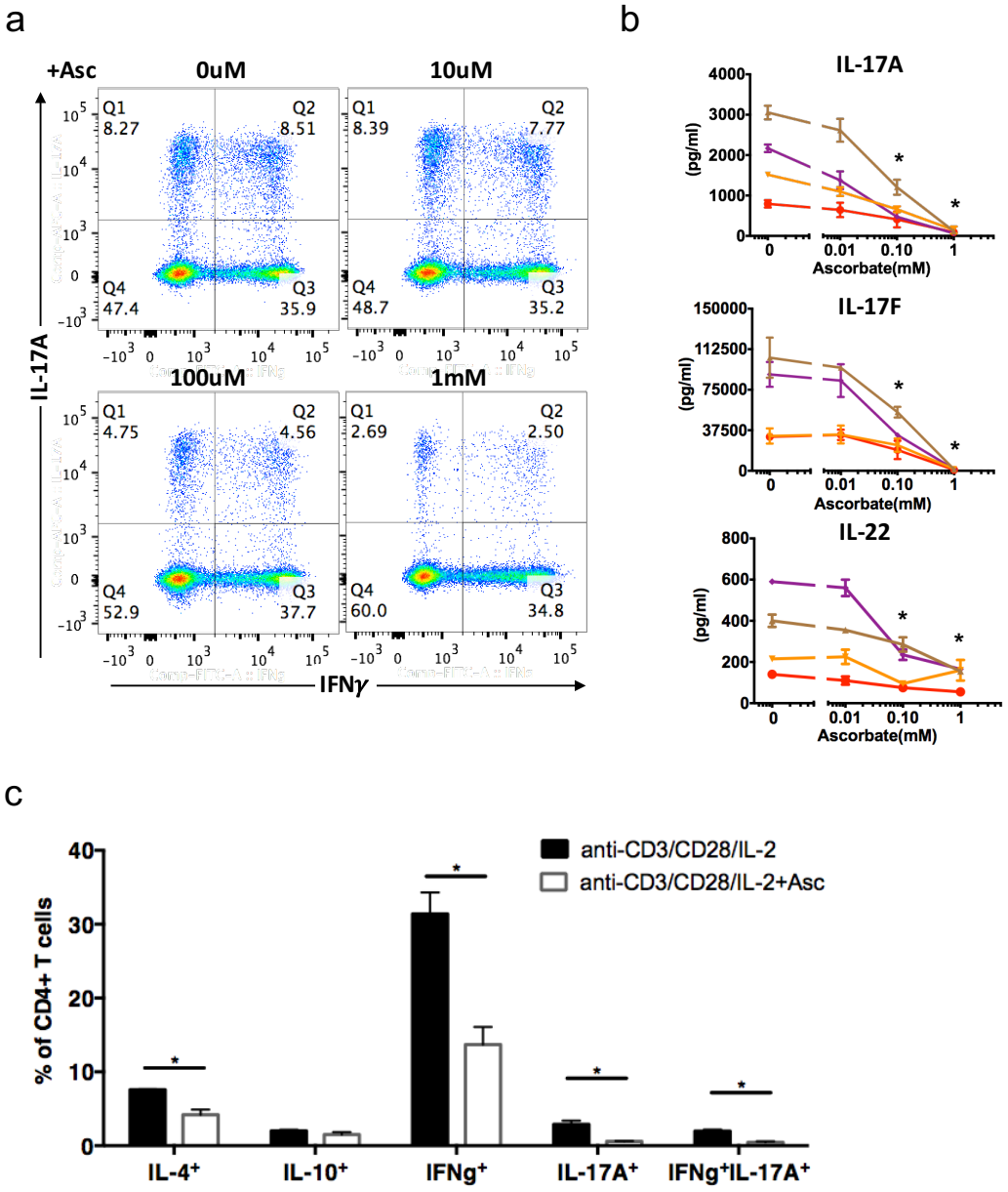


Figure 2-3. Ascorbate inhibits T cell cytokines in both polarized and non-polarized conditions. (a,b) CD4+ T cells were cultured with 0 to 1 mM of ascorbate as described in Figure 1. (a) Intracellular flow cytometry and (b) Luminex measurement of four biological replicates (four colors, mean ± SEM). (c) We cultured CD4+ T cells under T cell activation condition with 0 or 0.1 mM of ascorbate. We used intracellular flow cytometry to determine the percentages of indicated CD4+ T cells. Mean ± SEM calculated from three biological replicates.

Figure 2-4

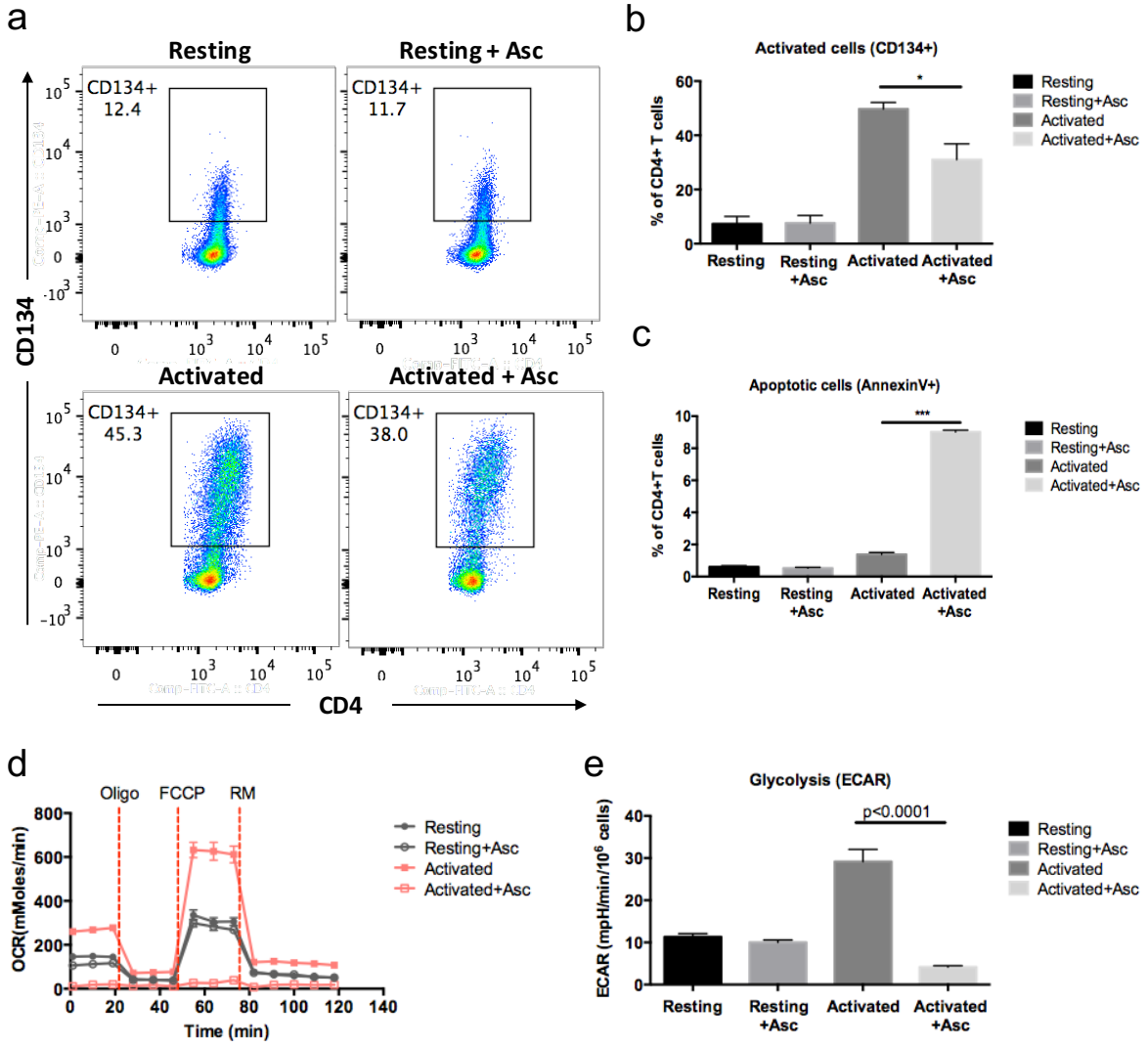


Figure 2-4. Ascorbate selectively inhibits activated but not resting CD4+ T cells. We cultured CD4+ T cells in 10 ng/ml of IL-7 (resting) or activated with 1ug/ml anti-CD3 and anti-CD28 antibody (activated) for 48 hours. We assessed T cell activation (CD134+) by flow cytometry. (a) Representative plot (b) Summary of the three independent experiments. (c) We measured cell death by Annexin V flow cytometry from three independent experiments. (d, e) Seahorse platform measurements of OCR and ECAR. (d) Representative plot; (e) Mean ± SEM calculated from 6 technical replicates.

Figure 2-5

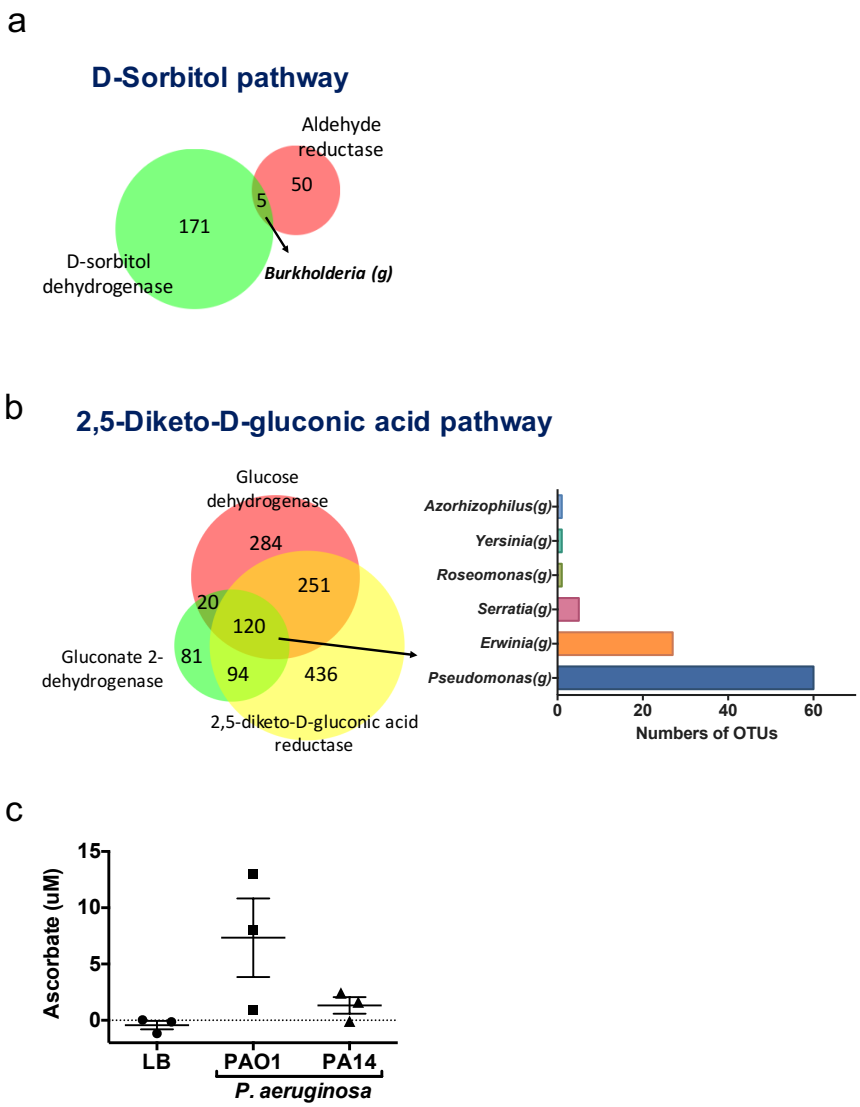


Figure 2-5. Gut microbes producing ascorbate. (a) Enzymes involved in the sorbitol pathway or (b) in the 2,5-diketo-D-gluconic acid pathway of ascorbate biosynthesis. OTUs: operational taxonomic unit. Left: Numbers of OTUs bearing metagenes for the indicated number of pathway enzymes. Right: Bar plots of the numbers of OTUs expressing indicated enzymes in bacterial genera. (c) We cultured bacteria for 24 hours and then collected supernatants to measure ascorbate levels. Since all cultures had reached equivalent McFarland readings at 24 hours, the ascorbate levels represent production from comparable cell numbers.

Supplementary Tables and Figures

Table S2-1. List of screened metabolites

Metabolite Name	HMDB_ID	KEGG_ID	Cat.# (Sigma-Aldrich)
L-Phenylalanine	HMDB00159	C00079	P5482-25G
gamma-L-Glutamyl-L-cysteine	HMDB01049	C00669	G0903-25MG
Sulfite	HMDB00240	C00094	S0505-250G
D-Fructose 1,6-bisphosphate	HMDB01058	C00354	F6803-1G
Thiamin monophosphate	HMDB02666	C01081	T8637-5G
(R)-5-Phosphomevalonate	HMDB01343	C01107	07841-10MG
O-Acetyl-L-serine	HMDB03011	C00979	A6262-10MG
L-Lysine	HMDB00182	C00047	L5501-10MG
2-Oxoglutarate	HMDB00208	C00026	K1875-1G
dTMP	HMDB01227	C00364	T7004-100MG
(S)-3-Hydroxy-3-methylglutaryl-CoA	HMDB01375	C00356	H6132-5MG
beta-Alanine	HMDB00056	C00099	A9920-100G
alpha-D-Glucose 6-phosphate	HMDB01401	C00668	G7250-10MG
Phosphoenolpyruvate	HMDB00263	C00074	P7127-100MG
Glutathione	HMDB00125	C00051	PHR1359-500MG
UTP	HMDB00285	C00075	U6875-100MG
dTTP	HMDB01342	C00459	GE27-1880-04
Hexanoyl-CoA	HMDB02845	C05270	H2012-5MG
D-Fructose 6-phosphate	HMDB00124	C00085	F3627-10MG
2-Dehydro-3-deoxy-D-gluconate	HMDB01353	C00204	12271-1MG
O-Phospho-L-homoserine	HMDB03484	C01102	04668-50MG
Succinate	HMDB00254	C00042	W327700-1KG-K
Glyoxylate	HMDB00119	C00048	G10601-25G
L-Ornithine	HMDB00214	C00077	O2375-10MG
GDP-mannose	HMDB01163	C00096	G5131-10MG
D-Glyceraldehyde 3-phosphate	HMDB01112	C00118	39705-1ML
IMP	HMDB00175	C00130	I4625-5G
L-Histidine	HMDB00177	C00135	H6034-10MG
Citrate	HMDB00094	C00158	W302600-1KG-K
Adenosine	HMDB00050	C00212	A9251-1G
D-Mannose 1-phosphate	HMDB06330	C00636	M1755-10MG
LL-2,6-Diaminoheptanedioate	HMDB01370	C00666	89469-10MG
Urocanate/4-Imidazoleacrylic acid	HMDB00301	C00785	859796-5G

Isopentenyl diphosphate	HMDB04196	C00129	00297-10MG
AMP	HMDB00045	C00020	01930-5G
CoA	HMDB01423	C00010	C4282-10MG
UDP	HMDB00295	C00015	94330-100MG
L-Glutamate	HMDB00148	C00025	G1501-100G
Glycine	HMDB00123	C00037	G7126-100G
L-Arginine	HMDB00517	C00062	A5006-100G
CTP	HMDB00082	C00063	C1506-25MG
Thiamin diphosphate	HMDB01372	C00068	C8754-1G
L-Ascorbate	HMDB00044	C00072	A7631-25G/A4034-100G
L-Methionine	HMDB00696	C00073	M9625-5G
Fumarate	HMDB00134	C00122	F1506-25G
L-Leucine	HMDB00687	C00123	L8000-25G
dATP	HMDB01532	C00131	D6500-10MG
myo-Inositol	HMDB00211	C00137	I5125-50G
L-Homocysteine	HMDB00742	C00155	69453-50MG
Acetoacetate	HMDB00060	C00164	A8509-10MG
Agmatine	HMDB01432	C00179	A7127-1G
D-Glucuronate	HMDB00127	C00191	G8645-5G
Acetyl phosphate	HMDB01494	C00227	A0262-500MG
L-Homoserine	HMDB00719	C00263	H6515-10MG
D-Mannose 6-phosphate	HMDB01078	C00275	M3655-100MG
Creatine	HMDB00064	C00300	C0780-10MG
Isocitrate	HMDB00193	C00311	58790-250MG
L-Citrulline	HMDB00904	C00327	C7629-1G
6-Phospho-D-gluconate	HMDB01316	C00345	55962-10MG
Phosphatidylethanolamine	HMDB60501	C00350	1535744-100MG
cis-Aconitate	HMDB00072	C00417	A3412-1G
(R)-Mevalonate	HMDB59629	C00418	50838-10MG
5,6-Dihydrouracil	HMDB00076	C00429	D7628-5G
dCTP	HMDB00998	C00458	D4635-25MG
Glutaryl-CoA	HMDB01339	C00527	G9510-5MG
Pyridoxine phosphate	HMDB01319	C00627	82890-10MG
(S)-Methylmalonyl-CoA	HMDB02310	C00683	M1762-1MG
Deamino-NAD (NAADP)	HMDB01179	C00857	N6506-25MG
L-Histidinol	HMDB03431	C00860	H6647-10MG
Crotonoyl-CoA	HMDB02009	C00877	28007-5MG
Dephospho-CoA	HMDB01373	C00882	D3385-5MG

L-Histidinol phosphate		C01100	41486-10MG
L-Galactono-1,4-lactone	HMDB02541	C01115	05313-100MG
D-Ribulose 1,5-bisphosphate	HMDB11688	C01182	83895-10MG
3-Methylbutanoyl-CoA	HMDB01113	C02939	I9381-10MG
cis-Homoaconitate		C04002	40487-10MG
1-(5'-Phosphoribosyl)-5-amino-4-imidazolecarboxamide (AICAR)	HMDB01517	C04677	A9978-25MG
Succinyl-CoA	HMDB01022	C00091	S1129-5MG
Glycerone phosphate	HMDB01473	C00111	37442-100MG-F
Proline betaine	HMDB04827	C10172	sc-296420(Santa Cruz)
Suberic acid	HMDB00893	C08278	S5200-5G
Oleic acid	HMDB00207	C00712	O1008-1G
Tyramine	HMDB00306	C00483	T90344-5G
Niacinamide	HMDB01406	C00153	N5535-100G
Arachidonic acid	HMDB01043	C00219	A3555-10MG
Caprylic acid	HMDB00482	C06423	PHR1202-1G
Linoleic acid	HMDB00673	C01595	L1376-1G
L-Tyrosine	HMDB00158	C00082	T8566-25G
Butyric acid	HMDB00039	C00246	B103500-100ML
cis-4,7,10,13,16,19-Docosahexaenoic acid	HMDB02183	C06429	D2534-25MG
L-Isoleucine	HMDB00172	C00407	I2752-5G
Sebacic acid	HMDB00792	C08277	283258-5G
4-Vinylphenol	HMDB04072	C05627	sc-267756; (Santa Cruz), 10% solution in propylene glycol
Oleamide	HMDB02117	C19670	O2136-100MG
Propionic acid	HMDB00237	C00163	P1386-500ML
Stigmasterol	HMDB00937	C05442	S2424-1G
Shikimic acid	HMDB03070	C00493	S5375-1G
N-Acetylputrescine	HMDB02064	C02714	A8784-25MG
Hexanoic acid	HMDB00535	C01585	153745-2.5G
Pyridoxamine dihydrochloride	HMDB01431	C00534	P9380-1G
N-a-Acetyl-L-arginine	HMDB04620		S451029-250MG
L-Carnitine	HMDB00062	C00318	C0158-1G
Azelaic acid (98%)	HMDB00784	C08261	246379-25G
Dodecanedioic acid	HMDB00623	C02678	D1009-100G
Nicotinic acid	HMDB01488	C00253	N4126-5G
1,2,4-Trimethylbenzene	HMDB13733	C14533	T73601-25ML

Isovaleric acid	HMDB00718	C08262	129542-100ML
Taurine	HMDB00251	C00245	T0625-10G
Phenylacetic acid	HMDB00209	C07086	P16621-5G
Putrescine	HMDB01414	C00134	51799-100MG
L-(+)-Lactic acid	HMDB00190	C00186	L1750-10G
Piperidine	HMDB34301	C01746	76046-100MG/411027
gamma-butyrobetaine/(3-Carboxypropyl)trimethylammonium chloride	HMDB01161	C01181	403245-1G
D-Glyceric acid	HMDB00139	C00258	61786-10MG
Cyclohexylamine	HMDB31404	C00571	G6503-1G
4-Guanidinobutyric acid	HMDB03464	C01035	G6503-1G
1-Oleoyl-sn-glycero-3-phosphocholine(synthetic, ≥99%)	HMDB02815	C04230	L1881-5MG
Cadaverine	HMDB02322	C01672	33211-10ML-F
L-(−)-Fucose	HMDB00174	C01019	F2252-10MG
2-Palmitoylglycerol	HMDB11533		75614-25MG
Cholesteryl oleate	HMDB00918	C14641	C9253-100MG
Histamine	HMDB00870	C00388	H7125-1G
D-Pantothenic acid	HMDB00210	C00864	P5155-100G
dGDP	HMDB00960	C00361	D9250-25MG
Tetradecanoyl-CoA	HMDB01521	C02593	M4414-5MG
NAD	HMDB00902	C00003	N7004-250MG
ADP	HMDB01341	C00008	A2754-100MG
FAD	HMDB01248	C00016	F6625-10MG
S-Adenosyl-L-methionine	HMDB01185	C00019	A7007-25MG
Acetyl-CoA	HMDB01206	C00024	A2056-10MG
GDP	HMDB01201	C00035	G7127-25MG
FMN	HMDB01520	C00061	F2253-25MG
GMP	HMDB01397	C00144	G8377-500MG
D-Erythrose 4-phosphate	HMDB01321	C00279	E0377-10MG
dGTP	HMDB01440	C00286	D4010-10MG
Acetoacetyl-CoA	HMDB01484	C00332	A1625-10MG
dUMP	HMDB01409	C00365	D3876-100MG
5-Aminolevulinate	HMDB01149	C00430	A3785-500MG
Homogentisate	HMDB00130	C00544	H0751-100MG

Table S2-2. Gut OTUs containing enzymes of the D-sorbitol pathway

OTU_IDs	Consensus Lineage
156580	k__Bacteria; p__Proteobacteria; c__Betaproteobacteria; o__Burkholderiales; f__Burkholderiaceae; g__Burkholderia; s__
4322414	k__Bacteria; p__Proteobacteria; c__Betaproteobacteria; o__Burkholderiales; f__Burkholderiaceae; g__Burkholderia; s__
49998	k__Bacteria; p__Proteobacteria; c__Betaproteobacteria; o__Burkholderiales; f__Burkholderiaceae; g__Burkholderia; s__
4311006	k__Bacteria; p__Proteobacteria; c__Betaproteobacteria; o__Burkholderiales; f__Burkholderiaceae; g__Burkholderia; s__
135993	k__Bacteria; p__Proteobacteria; c__Betaproteobacteria; o__Burkholderiales; f__Burkholderiaceae; g__Burkholderia; s__

Table S2-3. Gut OTUs containing enzymes of the 2,5-diketo-D-gluconic acid pathway

OTU_IDs	Consensus Lineage
4368451	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhizobiales; f__Phyllobacteriaceae; g__; s__
221365	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__; s__
4369186	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__; s__
4343884	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__; s__
648004	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__; s__
772368	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__; s__
88754	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__; s__
2228274	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__; s__
1044767	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__; s__
639311	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__; s__
163322	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__; s__
929312	k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Acetobacteraceae; g__Roseomonas; s__mucosa
236142	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
1119540	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
510870	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
819636	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
4456892	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
213522	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
540402	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
219248	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
3944484	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__

819387	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
1110763	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
238820	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
4432891	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__; s__
578058	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
239362	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
9939	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
138826	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
806143	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
813564	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
4430681	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
4318990	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
656889	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
4067355	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
835346	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
9826	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
9912	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
1022184	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
1134377	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
4417158	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
688934	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
167120	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
4250206	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__

585701	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
9918	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__
817254	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__dispersa
539107	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__dispersa
808540	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__dispersa
169538	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__dispersa
825033	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__dispersa
372473	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Erwinia; s__soli
3159130	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Serratia; s__
1726426	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Serratia; s__
4454068	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Serratia; s__
540793	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Serratia; s__
4360511	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Serratia; s__
759061	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Enterobacteriales; f__Enterobacteriaceae; g__Yersinia; s__
244248	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__; s__
4382169	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Azorhizophilus; s__
358042	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
133961	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
139321	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
813216	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
279948	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
171527	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
350105	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__

242070	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
104313	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
3290367	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
141206	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
218267	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
271238	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
553648	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
617271	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
4435983	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
144452	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
219457	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
99682	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
4475523	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
4320653	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
828623	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
217410	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
513808	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
28841	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
269901	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
1119175	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
4353093	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
780261	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
287032	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__

256834	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
170405	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
1118259	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
4451011	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
4408227	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
4372132	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
170417	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
3964538	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
109263	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
227761	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
144048	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
279231	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
4404312	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
544313	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
3913171	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
4389037	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
60329	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
728119	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__
61102	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__citronellolis
295031	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__fragi
327694	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__fragi
247615	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__nitroreducens

4339214	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__nitroreducens
226424	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__nitroreducens
138840	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__umsongensis
2589305	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__veronii
276867	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__veronii
556184	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__veronii
4432796	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__viridiflava
265715	k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Pseudomonadales; f__Pseudomonadaceae; g__Pseudomonas; s__viridiflava

Figure S2-1

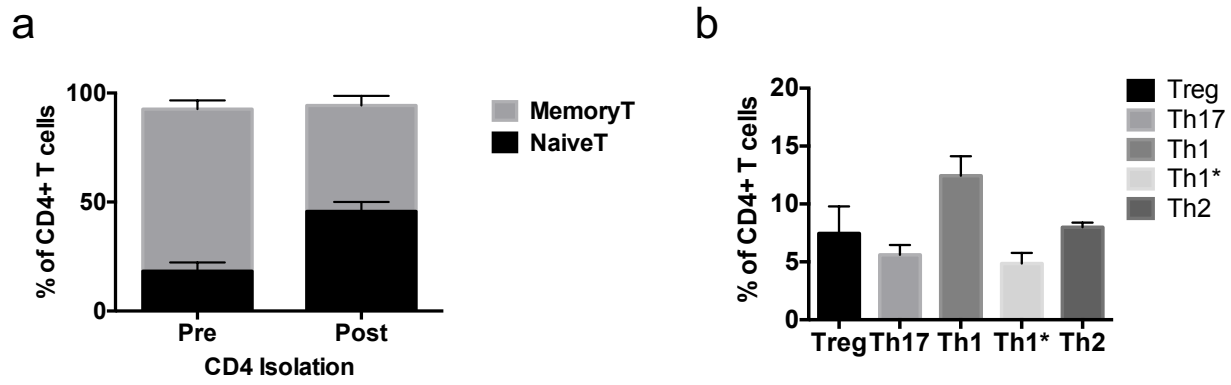


Figure S2-1. T cell subsets in the CD4+ T population before culture. (a) Pre- and post-isolated CD4+ T cells were stained with cell surface markers, and assessed for % naïve/memory population by flow cytometry. (b) Isolated CD4+ T cells stained with cell surface surrogate markers for T cell subsets, and assessed for % T cell subsets by flow cytometry.

Figure S2-2

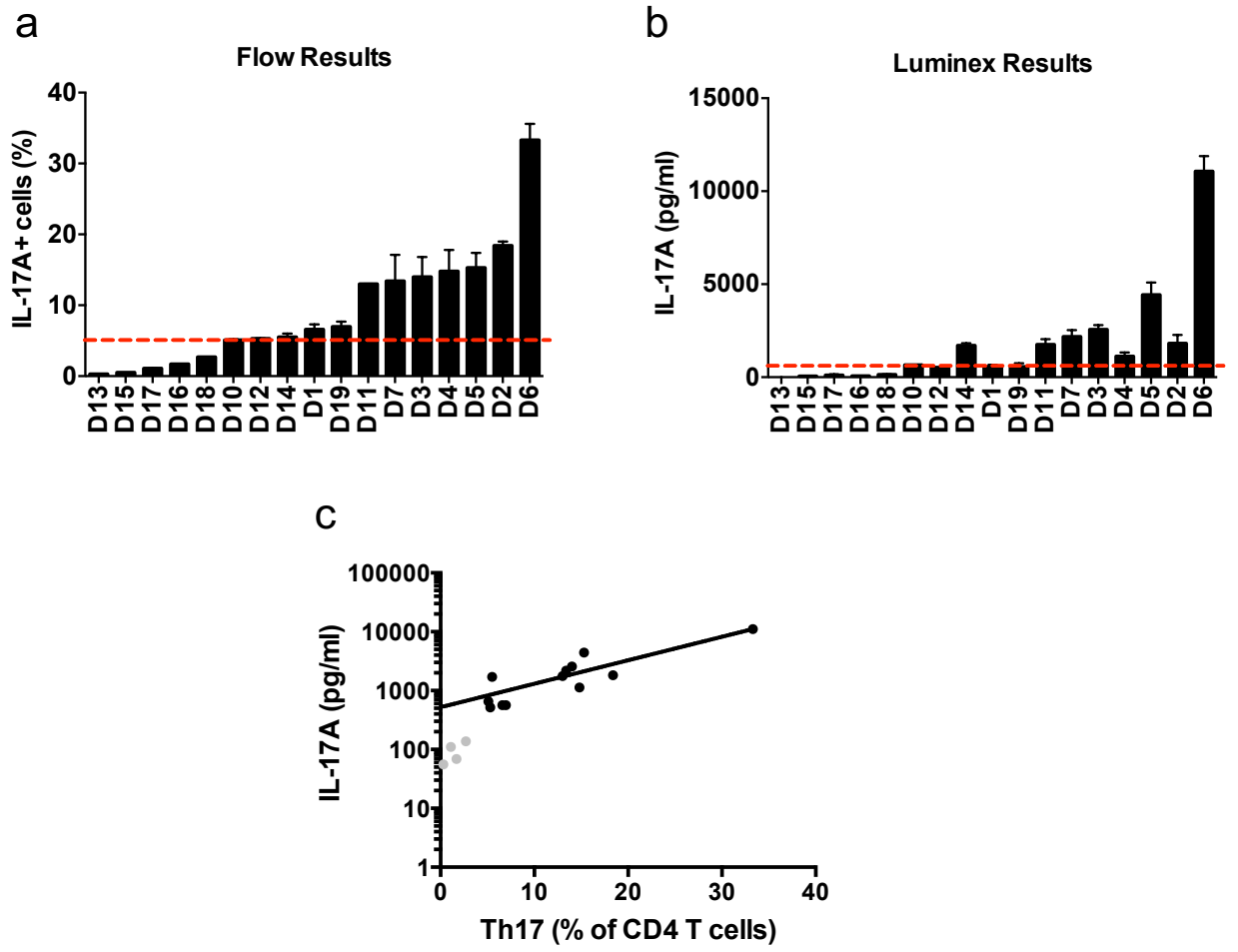


Figure S2-2. Levels of human Th17 induction among study donors. (a, b) We cultured isolated CD4+ T cells for 14 days with Th17 polarizing cytokines, and thereafter assessed cells for intracellular IL-17A expression by flow cytometry and supernatant IL-17A levels by multiplex Luminex assay. Each bar represents data from one donor. Mean ± SEM calculated from 3 technical replicates. (c) XY-plot shows the correlation between intracellular and secreted IL-17A. Each dot represents one donor. Gray dots indicate donor samples that did not meet assay quality control cut-off.

Figure S2-3

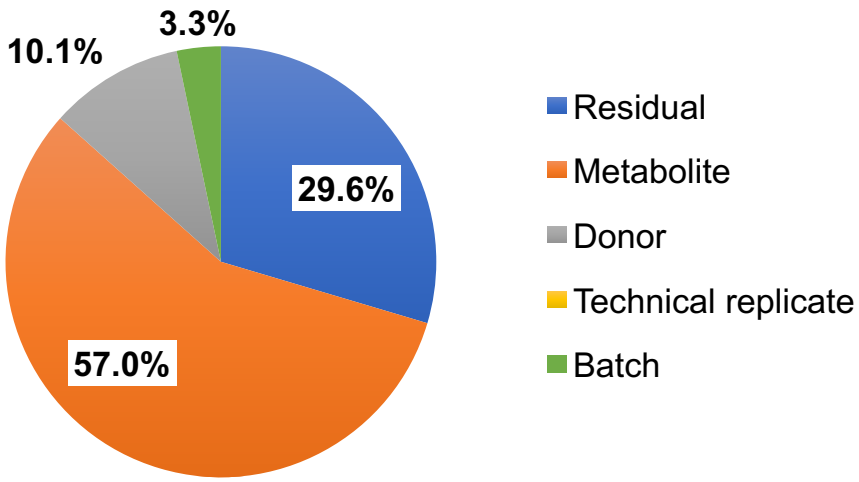


Figure S2-3. Principal variance component analysis (PVCA) of contributors to total variability. Secreted cytokine data from four donors with three technical replicates were analyzed in two independent batches. Metabolites for this analysis included GDP-mannose, ascorbate, and L-homocysteine. The variance attributed to technical replication was <1%.

Figure S2-4

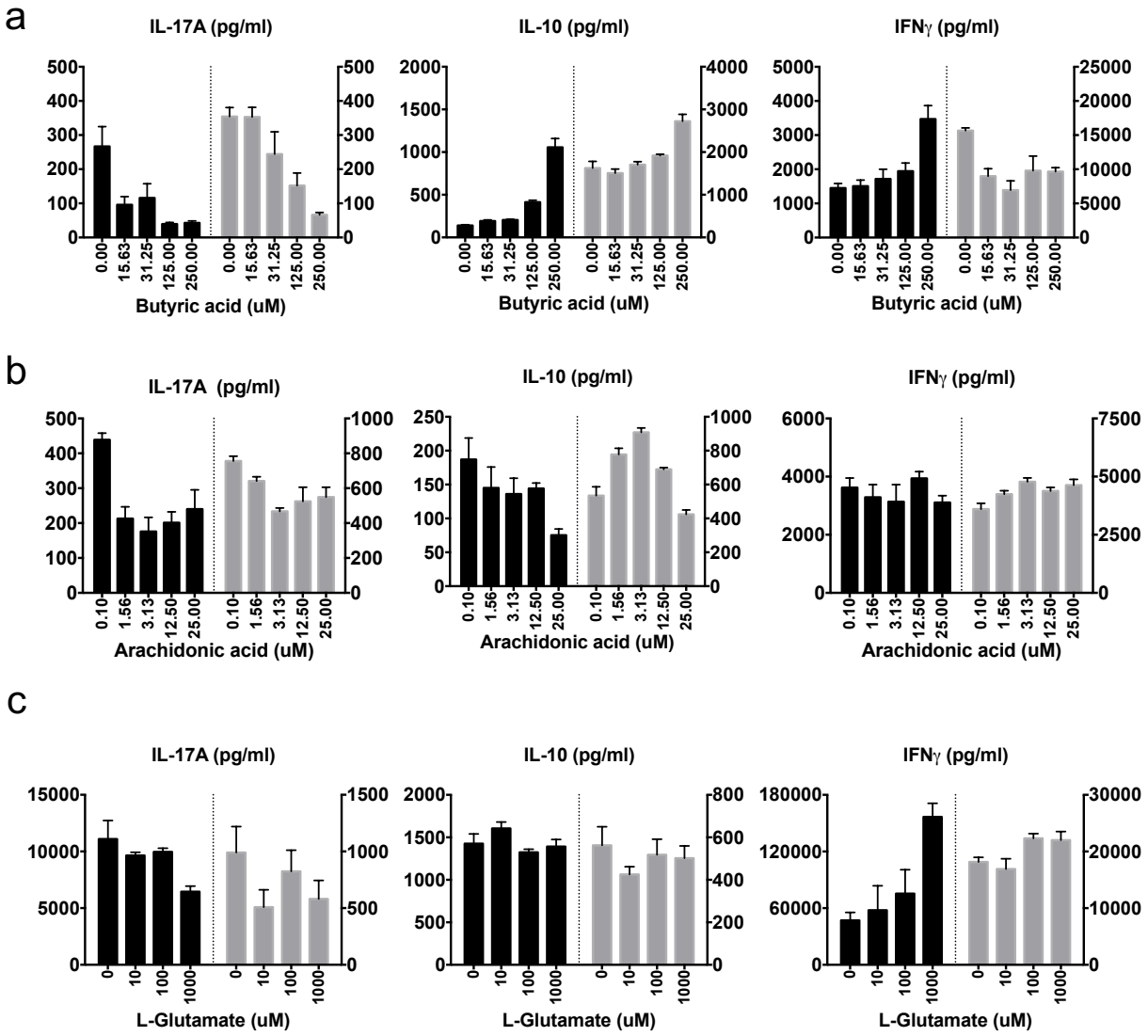


Figure S2-4. Bioactivity of metabolites with previously reported effects on CD4+ T cells. Secreted cytokines levels of CD4+ T cells cultured under polarizing Th17 conditions with the indicated metabolites. (a) Butyric acid, (b) arachidonic acid, and (c) L-glutamate. We show data from two biological replicates (black and gray). Mean ± SEM calculated from three technical replicates.

Figure S2-5

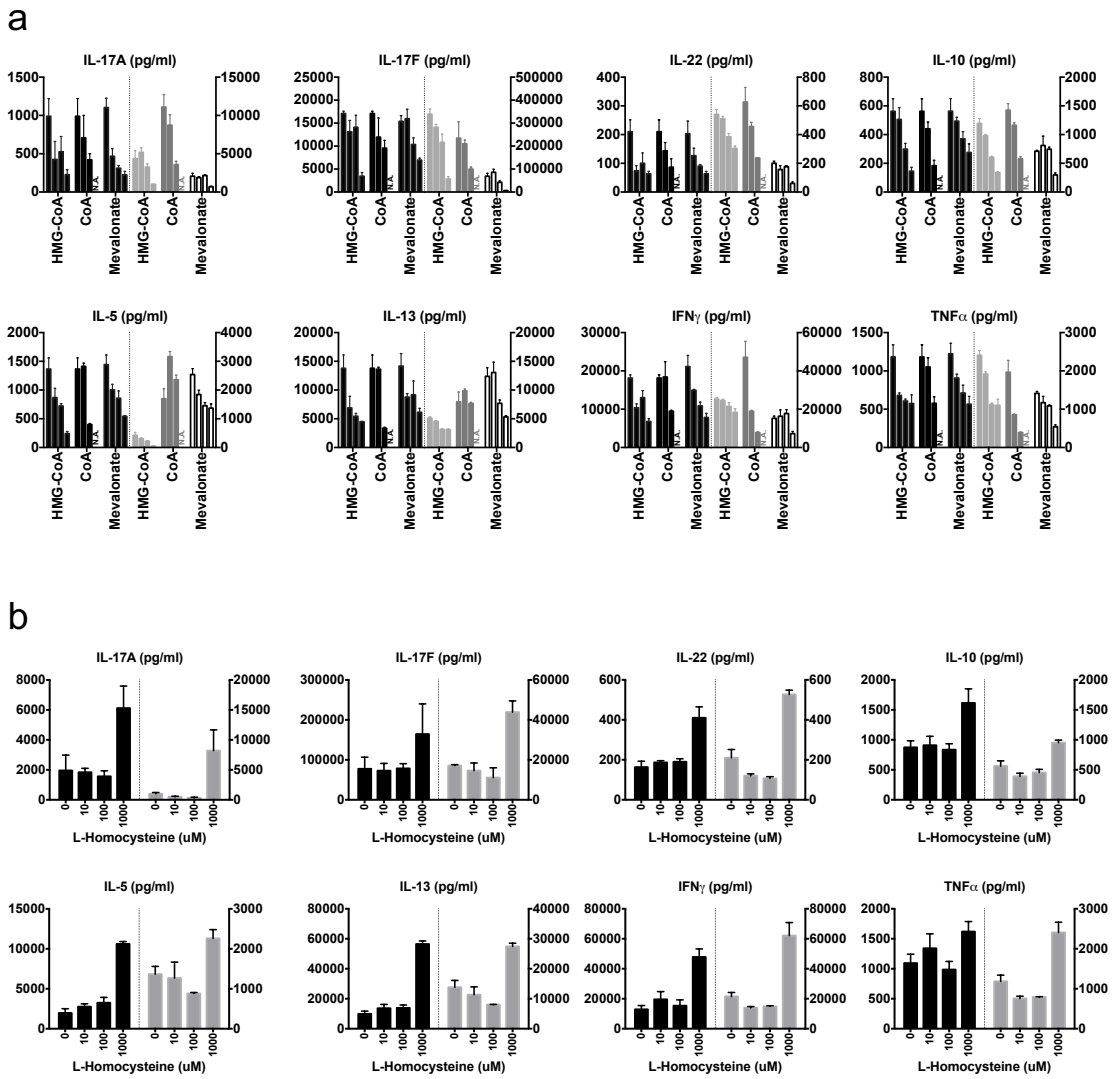


Figure S2-5. Metabolites with pan-effects on cytokine production. Secreted cytokines from CD4+ T cells cultured under polarizing Th17 conditions with the indicated metabolites. We show data from independent biological replicates for each metabolite (black, dark-gray, light-gray, and white). CoA: Coenzyme A; HMG-CoA: 3-hydroxy-3-methyl-glutaryl-coenzyme A. (a) Metabolites with apparent pan-inhibition. (b) Metabolites with apparent pan-enhancement. N.A., conditions causing cell death and hence excluded. Mean ± SEM calculated from 3 technical replicates.

Figure S2-6

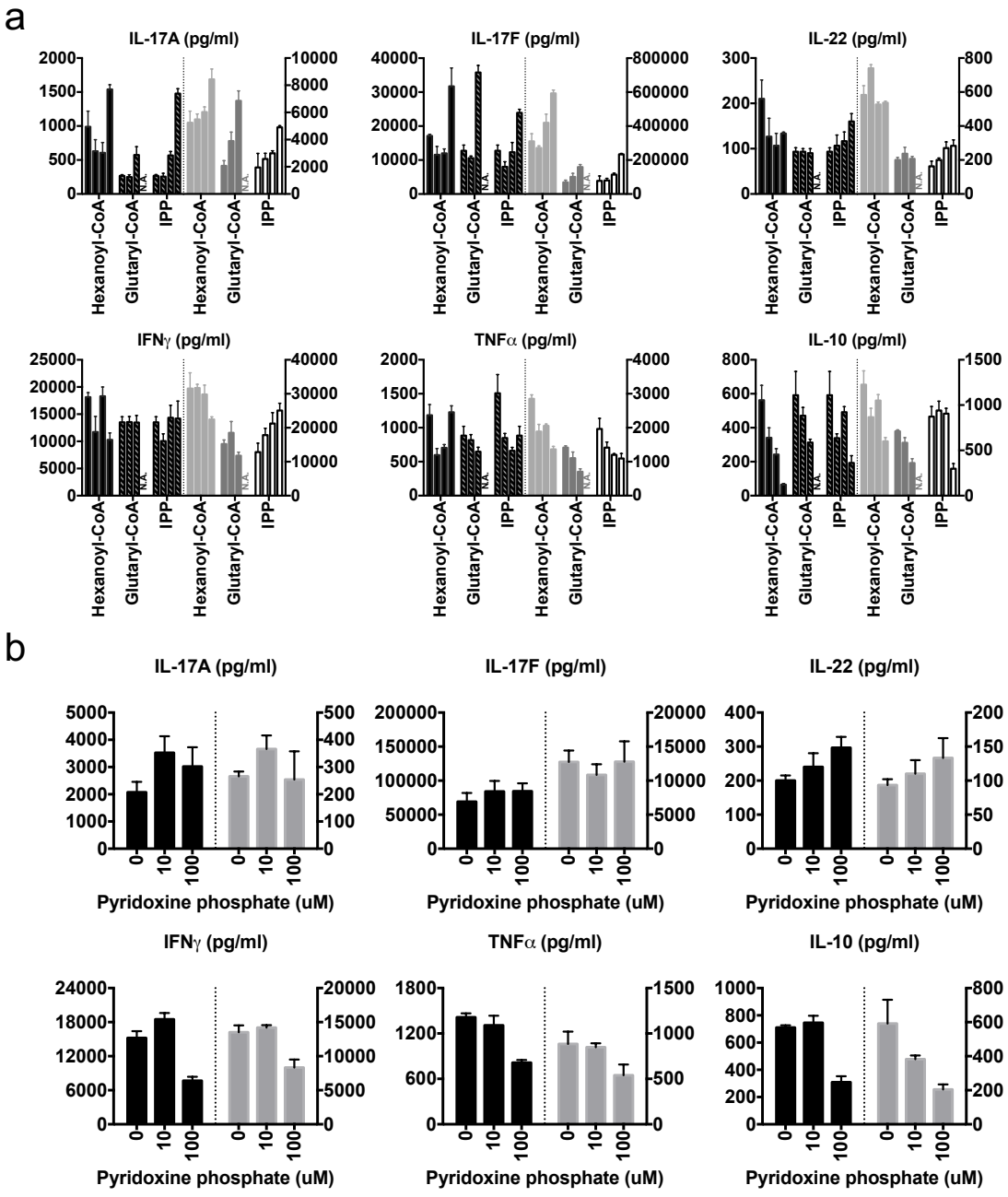


Figure S2-6. Metabolites with selective effects on cytokine production. Secreted cytokines from CD4+ T cells cultured under polarizing Th17 conditions with the indicated metabolites. (a) IPP: isopentenyl diphosphate. (b) Pyridoxine phosphate. N.A.: conditions that caused cell death and excluded. Mean ± SEM calculated from 3 technical replicates.

CHAPTER 3

Association of Systemic Sclerosis with a Unique Colonic Microbial Consortium

Abstract

Objective. To compare colonic microbial composition in systemic sclerosis (SSc) patients and healthy controls and to determine whether certain microbial genera are associated with gastrointestinal (GI) tract symptoms in patients with SSc.

Methods. Healthy controls were age and sex matched (1:1) with adult SSc patients. Cecum and sigmoid mucosal lavage samples were obtained during colonoscopy. The microbiota in these samples were determined by Illumina HiSeq 2000 16S sequencing, and operational taxonomic units were selected. Linear discriminant analysis effect size was used to identify the genera that showed differential expression in SSc patients versus controls. Differential expression analysis for sequence count data was used to identify specific genera associated with GI tract symptoms.

Results. Among 17 patients with SSc (88% female; median age 52.1 years), the mean 6 SD total GI Tract 2.0 score was 0.760.6. Principal coordinate analysis illustrated significant differences in microbial communities in the cecum and sigmoid regions in SSc patients versus healthy controls (both P=0.001). Similar to the findings in inflammatory disease states, SSc patients had decreased levels of commensal bacteria, such as *Faecalibacterium* and *Clostridium*, and increased levels of pathobiont bacteria, such as *Fusobacterium* and γ-*Proteobacteria*, compared with healthy controls. *Bifidobacterium* and *Lactobacillus*, which are typically reduced under conditions of inflammation, were also increased in abundance in patients with SSc. In SSc patients with moderate/severe GI tract symptoms, the abundance of *Bacteroides fragilis* was decreased, and that of *Fusobacterium* was increased, compared with patients who had no or mild symptoms.

Conclusion. This study demonstrates a distinct colonic microbial signature in SSc patients compared with healthy controls. This unique ecologic change may perpetuate immunologic aberrations and contribute to clinical manifestations of SSc.

Introduction

Gastrointestinal (GI) tract dysfunction is a leading cause of morbidity and mortality in patients with systemic sclerosis (SSc)[90,91]. Symptoms of lower GI tract involvement, such as constipation, abdominal pain, diarrhea, faecal incontinence, and weight loss[92], are among the most disruptive physical problems for SSc patients and compromise patient emotional well-being and quality of life[93,94].

The etiology of SSc-related lower GI tract dysfunction is largely unknown, and there are no effective treatment options[91]. The findings of some studies of upper GI tract dysfunction in patients with SSc have suggested that changes in the microvasculature, autonomic nervous system, and immune system may contribute to smooth muscle atrophy and gut wall fibrosis[95-97]. However, other evidence demonstrates that bacterial overgrowth may lead to malabsorption in these patients[98]. Marie et al [92] reported that in a study of 51 consecutively seen SSc patients, nearly half had positive results of an H_2/CH_4 breath test, and those with positive breath test results had more severe GI tract symptoms. Moreover, studies suggest that antibiotic therapy reduces lower GI tract symptoms[92,99]. Thus, alterations in microbial composition may potentially be a pathogenic factor in GI tract dysfunction in SSc.

While no reported studies have examined whether imbalances in colonic microbial composition are a feature of SSc, numerous studies have demonstrated that altered microbiota can induce or intensify inflammation in other autoimmune diseases, such as inflammatory bowel disease (IBD)[100]. There is substantial evidence suggesting that the abundance of beneficial human commensal genera known to produce key energy metabolites and antiinflammatory molecules for mucosal health (e.g., Bacteroides, Bifidobacterium, Clostridium types IV and XIV) is decreased, with a concurrent increase in potential pathobiont genera (e.g., sulfate- reducing d-

Proteobacterium, invasive γ-Proteobacterium, Fusobacterium, and Actinobacteria), in patients with IBD[101]. Moreover, metaproteome analyses at the mucosal–luminal interface in patients with IBD have shown that specific bacterial phylotypes were associated with increases in local host inflammatory products[102,103].

The present study was undertaken to investigate the hypothesis that the SSc disease state is associated with altered colonic microbial composition at the human mucosal–luminal interface. We also aimed to examine the relationship between microbial composition and self-reported symptoms of GI tract dysfunction in SSc patients, as an initial test of the hypothesis that certain microbial genera contribute to the GI tract phenotype in SSc. If the hypothesis is confirmed, such genera could provide specific targets for intervention to avert or treat this important clinical dimension of SSc.

Materials and Methods

Study Participants

Patient participants were consecutively enrolled from the outpatient rheumatology clinic at the University of California, Los Angeles (UCLA). Eligible participants included adult (age ≥18 years) patients with SSc. Exclusion criteria included comorbid IBD, contraindication to undergoing a colonoscopy, and inability to have antibiotic and probiotic treatment withheld for at least 3 weeks prior to colonoscopy. Patients were allowed to continue taking proton-pump inhibitors because these agents exert minimal-to-negligible effects on colonic microbiota[104].

Healthy control colonic lavage specimens were obtained from the UCLA Pathology Microbiome Repository, which includes specimens from 150 healthy adult subjects. Age- and sex-matched healthy controls were selected from this cohort by block randomization and matched with SSc patients at a 1:1 ratio.

The UCLA institutional review board (IRB) approved the study protocol (IRB #13-0011089), and written informed consent was obtained from each participant.

Specimen procurement and pre-processing

Participants underwent colonoscopy performed by a certified gastroenterologist (BER, JLC, or TG). We opted to collect mucosal instead of faecal samples, given the distinct and predictable compositional differences between colonic microbiota and stool microbiota demonstrated in prior studies[105,106]. In addition, by collecting lavage specimens we could ensure optimal preservation of the samples prior to preprocessing. Stool sample collection would have required patients to transport samples from home, and if sample thawing occurred, the integrity of the DNA and RNA would be compromised. We also planned to collect colonic biopsy samples if there was a clinical

indication for biopsy (ulceration, polyp, etc.). Complete details of specimen procurement and preprocessing are available upon request from the corresponding author.

Gene sequencing for 16S ribosomal RNA (rRNA) and microbial composition analysis.

The microbiota from the samples were profiled by multiplex sequencing for bacterial rRNA genes, using an Illumina HiSeq 2000. The exact details of this approach have been outlined in an earlier report by our group[103] and are available upon request from the corresponding author.

Assessment of GI tract symptoms.

Demographic and disease-related characteristics were recorded, and participants completed the GI Tract 2.0 instrument (GIT 2.0) on the morning of their colonoscopy. The GIT has been shown to be a valid measure of GI tract symptom severity in SSc patients[107]. The questionnaire consists of 7 scales (reflux, distension/bloating, diarrhea, faecal soilage, constipation, emotional well-being, and social functioning). The GIT 2.0 can furthermore discriminate between self-rated severity of GI tract involvement (none/very mild, mild, moderate, or severe/very severe), and results on this instrument correlate with objective measures of GI tract dysfunction, at least in the upper GI tract[108]. Participants also completed the disability index of the Scleroderma Health Assessment Questionnaire, a valid measure of self-reported function that is commonly used in SSc studies[109].

Bioinformatics and statistical analyses.

Operational taxonomic units (OTUs) were selected from the Greengenes database, and microbial OTUs were rarefied down to 15,000 reads per sample using Macqiime. To compare the microbial communities in SSc versus control samples, alpha diversity and beta diversity were analyzed. Alpha diversity, which represents the complexity of composition within members of a group, was calculated using the metrics of phylogenetic diversity, Chao1, observed species, and

Shannon index. Statistical comparison of alpha diversity between the 2 groups was performed using Student's 2-sided t-test at a depth of 15,000. Beta diversity, which represents the between-subject similarity of microbial composition and enables identification of differences between samples within a group, was computed in Macqiime, utilizing both unweighted and weighted UniFrac distances to estimate sample distributions. Analysis of variance using distance matrices (Adonis) significance analysis was performed for each pairwise comparison of sample groups, using the Adonis function from the R package. Principal coordinate analysis (PCoA) was performed to visualize the resulting UniFrac distance matrix.

Analyses were performed using R version 3.1.2. Means and standard deviations were used to describe continuous parametric data, and medians and interquartile ranges were used to describe continuous nonparametric data. All tests were 2-sided with an alpha level of 0.05. We used the false discovery rate (FDR) of Benjamini and Hochberg[110], and significant association was defined at the FDR q value threshold of <0.1. Cecum and sigmoid findings were treated as separate data sets because prior studies have demonstrated biogeographic differences in mucosal samples from these 2 regions[111].

Results

Participant characteristics.

Seventeen patients with SSc (15 women [88%] and 2 men [12%]) underwent colonoscopy and completed the questionnaire. Sigmoid and cecum lavage samples were obtained from all patients except 1, from whom only a sigmoid lavage sample was obtained. None of the patients had a clinical indication for colonic biopsy; therefore, these specimens were not collected. Clinical characteristics are shown in Table 3-1. The median age of the patients was 52.1 years, and the median disease duration was 6.6 years. The mean total score on the GIT 2.0 was 0.7, indicating moderate symptom severity. Scores on individual domains of the GIT 2.0 indicated moderate severity (distention/bloating, social functioning, emotional well-being, and constipation) or mild severity (faecal soilage and diarrhea). Six patients (35%) had taken antibiotics in the 3 months preceding the colonoscopy; the mean time between cessation of antibiotic treatment and colonoscopy was 6.5 weeks (range 4-12). None of the patients used tobacco products, and only 5 had consumed alcohol regularly within the previous month (mean ± SD 2.8 ± 1.5 servings of alcohol per week).

Altered colonic microbial diversity in patients with SSc.

After OTU selection, a total of 5,442 and 5,593 species-level OTUs were generated from the cecum and sigmoid data sets, respectively. The SSc and control participants exhibited similar alpha diversity (i.e., the complexity of microbial composition) based on the following metrics (cecum *P* value and sigmoid *P* value, respectively): phylogenetic diversity (*P*=0.2; *P*=0.1), observed species (*P*=0.4; *P*=0.2); and Shannon index (*P*=0.7; *P*=0.6). Compared with controls, SSc patients exhibited a trend toward increased alpha diversity based on the Chao 1 diversity metric, which estimates OTU richness in microbial communities[112] (*P*=0.07 and *P*=0.09 for the

cecum and sigmoid regions, respectively; data available upon request from the corresponding author); this suggests that the SSc disease state is associated with a potential increase in bacterial diversity.

Beta diversity was then computed to determine whether microbial composition differed between SSc patients and control subjects. The PCoA visualization of unweighted UniFrac distances and Adonis analysis of this comparison is shown in Figure 3-1. In both the cecum and the sigmoid regions, the microbial composition among SSc patients was significantly different from that in healthy controls (P=0.001 for both regions). Significant differences between SSc patients and healthy subjects were also observed by weighted UniFrac distances (P=0.03 for Adonis analysis of both regions).

Differential abundance of numerous colonic microbial genera in SSc patients.

To begin to define the compositional differences between SSc patients and healthy controls predicted by the beta diversity analysis, the relative abundances of microbial composition at different taxonomic levels were computed. The 3 predominant phyla in SSc samples were *Bacteroidetes* (cecum 47.2%; sigmoid 46.7%), *Firmicutes* (cecum 23.3%; sigmoid 25.2%), and *Proteobacteria* (cecum 23.1%; sigmoid 20.5%). Similar to findings in an earlier study of the microbiome in IBD[15], the relative abundances of *Verrucomicrobia*, *Fusobacterium*, and *Actinobacteria* were significantly increased in cecum samples from SSc patients compared with controls (q=0.03, q=0.02, and q=0.003, respectively). Additional data from this analysis are available upon request from the corresponding author.

The taxonomic differences between SSc patients and healthy controls at the genus level were assessed by linear discriminant analysis (LDA) effect size multivariate analysis (to correct for species abundance). The microbial enrichments and depletions in the cecum and the sigmoid

that reached significance after correction for multiple testing (q<0.1) are shown in Figure 3-2 and 2-3. In the cecum, commensal genera such as *Faecalibacterium*, *Clostridium*, and *Rikenella* were depleted in SSc patients, whereas *Fusobacterium*, *Prevotella*, and the uncommon γ-*Proteobacteria*, *Erwinia*, and *Trabsulsiella* were enriched in SSc patients. Similarly, in the sigmoid, *Faecalibacterium* and *Rikenella* were depleted in SSc patients, whereas numerous sigmoid genera, such as *Fusobacterium* and *Prevotella*, were enriched. Surprisingly, 2 commensal bacterial genera (*Lactobacillus* and *Bifidobacterium*) were found in greater abundance in SSc patients compared with controls in both the sigmoid and cecum regions (Figure 3-2 and 3-3). These 2 species are typically reduced in abundance in chronic inflammatory states, such as IBD[113].

Association of specific microbial genera and species with GI tract symptoms in patients with SSc.

Multivariate analysis with differential expression analysis for sequence count data was used to identify microbial taxa associated with GI tract symptoms as assessed based on the GIT 2.0 total score and individual domains (constipation, diarrhea, or distention/bloating). For the total score or domains, SSc patients were dichotomized into low (none-to-mild) or high (moderate-to-severe) disease severity groups. The fold difference in the abundance of specific organisms between disease severity groups was calculated (Figure 3-4), and for organisms with significant differences between groups, q values were tabulated (data available upon request from the corresponding author). The level of *Bacteroides fragilis* in both the cecum and the sigmoid was elevated in patients with mild disease (GIT 2.0 total score, diarrhea domain, and bloating/distension domain). *Candidatus arthromitus* had a similar association with mild disease, but only in the cecum. Conversely, genus-level members of the *Fusobacterium* genus and the

Actinobacillus genus were associated with severe disease, but more prominently in the sigmoid region and more heterogeneously with respect to symptom domains.

Differences in microbial composition by SSc subtype.

In an exploratory analysis, we compared beta diversity between SSc patients with limited cutaneous disease (lcSSc; n=11) and those with diffuse cutaneous disease (dcSSc; n=6). Differences in beta diversity were significant by Adonis in the cecum (R^2= 0.13, *P*=0.011) but not in the sigmoid (R^2=0.08, *P*=0.2), as calculated using unweighted UniFrac distances. However, only a small amount of the variation (13%, as reflected by the R^2 value) could be explained by this subtype grouping for the cecum. LDA effect size multivariate analysis was used to examine the taxonomic differences between patients with lcSSc and dcSSc at the genus level. After correction for multiple testing (q<0.1), the abundance of *Lactobacillus* from the *Firmicutes* phylum was higher in patients with lcSSc (LDA 2.7), while that of *Paludibacter* from the Bacteroidetes phylum was lower in patients with lcSSc (LDA 2.7), compared with patients with dcSSc.

Altered metabolic proficiencies of microbiota in SSc patients.

To compare the functional proficiencies of the microbial communities in SSc patients and healthy controls, the imputed metagenomic composition in each patient was determined using PICRUSt (phylogenetic investigation of communities by reconstruction of unobserved states), and compared by calculating Bray-Curtis distance matrices and visualizing by PCoA. Differences in metagenomic content between SSc patients and healthy controls reached significance by Adonis in both the cecum region (R^2=0.9, *P*=0.016) and the sigmoid region (R^2=0.9, *P*=0.007).

Among all of the imputed genes, 66 in the cecum region and 89 in the sigmoid region showed significant differences (q<0.1) in SSc patients compared with healthy controls. Of these significantly dysregulated genes, 41 in the cecum region and 47 in the sigmoid region were mapped

to biologic pathways through the Kyoto Encyclopedia of Genes and Genomes pathway database. A complete list of gene differences and their associated pathways is presented in Supplementary Table 3-1, available on the Arthritis & Rheumatology website at http://onlinelibrary.wiley.com/doi/10.1002/art. 39572/abstract. SSc patients had a significantly decreased abundance of genes involved in amino sugar and nucleotide sugar metabolism, such as N-acylneuraminate cytidylyltransferase (K00983; log LDA effect size in the cecum and sigmoid, respectively 2.0 [q=0.001] and 2.1 [q=0.001]) and β-hexosaminidase (K12373; log LDA effect size in the cecum and sigmoid, respectively 2.4 [q=0.07] and 2.5 [q=0.04]), as well as genes involved in sphingolipid metabolism, such as β-galactosidase (K01190; log LDA effect size in the cecum and sigmoid, respectively 2.7 [q=0.02] and 2.8 [q=0.007]) (Figure 3-5).

Discussion

The present study is, to our knowledge, the first to define colonic microbial composition in adult patients with SSc. We report a unique colonic microbial consortium in SSc patients, characterized by significant increases in *Fusobacterium*, *Prevotella*, and uncommon γ-*Proteobacteria* (i.e., *Erwinia* and *Trabsulsiella*) genera and significant decreases in *Faecalibacterium* and *Clostridium* genera, compared with age- and sex-matched healthy controls. Interestingly, the SSc microbial consortium was also enriched with *Lactobacillus* and *Bifidobacterium*, 2 commensal genera typically found in lower abundance in chronic inflammatory states.

Fusobacterium species represent a group of gram-negative anaerobes that principally colonize the oral cavity. When present in the colon, these species are considered pathobionts given their invasive nature and ability to translocate into the blood and contribute to systemic processes, including bacteremia, organ abscesses, and possibly coronary artery disease[114,115]. In patients with Crohn's disease, *Fusobacterium* species are increased compared with controls[116], and human isolates of *Fusobacterium varium* induce colonic mucosal erosions in mice[117]. Moreover, *Fusobacterium* isolates recovered from IBD patients demonstrate enhanced invasive and pro-inflammatory properties in cultured epithelial cell assays than those strains isolated from healthy individuals [118], further implicating these species in the pathogenesis of IBD.

Prevotella species were also enriched in SSc patients compared with healthy controls. These genera are increased in patients with Crohn's disease[119], as well as in patients with new-onset rheumatoid arthritis[6]. A recent study demonstrated that inflammation was increased in mice with dextran sulfate sodium-induced colitis that were colonized with *Prevotella copri*, in particular[6].

The observation that *Fusobacterium* and *Prevotella* are increased in SSc patients compared with controls suggests that these species may play a role in the development of the SSc GI tract phenotype. Additional studies are needed to elucidate their pathogenic potential in SSc. Given the established link between diet and *Prevotella*[120], there may be a future role for dietary modification in averting GI tract symptoms in SSc.

The genera that were depleted in SSc, most notably *Faecalibacterium* and *Clostridium* species, are commensal organisms, which may protect against mucosal inflammation and colonization of pathogenic species[121]. For example, *Faecalibacterium prausnitzii* is decreased in IBD[122]. Low levels of mucosa-associated *F. prausnitzii* are associated with a higher risk of recurrent Crohn's disease after surgery[123], whereas recovery of *F. prausnitzii* after relapse is associated with maintenance of clinical remission in ulcerative colitis[124]. Moreover, *Clostridium* species have been found to induce the expansion of regulatory T cells, thereby reducing intestinal inflammation[125].

A surprising observation in this study was the significant increase in *Bifidobacterium* and *Lactobacillus* among SSc patients compared with healthy controls. *Lactobacillus* was also more abundant in patients with lcSSc compared with those with dcSSc, although the latter finding should be interpreted with caution given the exploratory nature of this subgroup analysis. While *Bifidobacterium* and *Lactobacillus* are typically depleted in chronic inflammatory states, a recent study of patients with enthesitis-related arthritis also demonstrated an increase in *Bifidobacterium* in stool samples from children with this disorder, compared with controls[126].

Taken together, these findings suggest that our traditional views of "protective" commensal species need to be considered in the context of the underlying disease state. Given the myriad of ways in which the phenotypic expression of SSc differs from that of other autoimmune diseases,

it is not surprising that the microbial consortium in SSc is unique. Moreover, our results may indicate that therapeutic efforts to increase commensal organisms in SSc (i.e., currently available commercial probiotics, which commonly include Lacto- bacillus or *Bifidobacter* species) should be targeted only toward select organisms.

In addition to identifying distinct features of the SSc colonic microbial consortium, we have elucidated relationships between specific microbial genera/species and the SSc GI tract phenotype. Increased abundance of *Fusobacterium* and decreased abundance of *B. fragilis* and *C. arthromitus* were associated with increased GI tract symptom severity, as measured by the GIT 2.0 total score as well as many of the individual GIT 2.0 domains reflecting lower GI tract dysfunction. If *B. fragilis* and *C. arthromitus* are validated in future studies to attenuate mucosal inflammation in SSc, therapy directed at increasing the growth or activity of these species to ameliorate GI tract symptoms in SSc would merit testing.

Furthermore, the finding of differences in predicted metagenomic pathways between SSc patients and healthy controls implies that the compositional shifts observed in this study likely correspond to altered functional capacity of the SSc-associated microbiome relative to a healthy microbiome. These pathways may mediate the relationship between microbiota composition and GI tract symptoms and warrant further investigation.

The nature and findings of the present study should be placed in the context of certain limitations. First, the study was cross-sectional, and it is unclear whether the observed relationships between specific genera and GI tract symptoms persist with time. To address this question, a 12-month longitudinal study of this cohort is currently under way. Second, there may be a small batch effect as the sequencing analysis was per- formed for the control samples prior to the SSc samples. Given that the sequencing protocol was identical for both groups and that the observed differences

between SSc patients and controls were highly significant, however, a possible batch effect is unlikely to fully explain the differences. Third, the sample size was small, and the results would need to be confirmed in a validation cohort. Despite the small sample size, however, we observed several significant associations, suggesting that it is unlikely the present findings are due to chance alone. Fourth, because of the small sample size and concerns regarding multiple hypothesis testing (Type I error), we were unable to perform meaningful subgroup analyses to compare the microbiota in patients with different disease characteristics (e.g., negative versus positive for Scl-70 antibody). Fifth, colonic motility was not assessed in this study since there is presently no valid measure of colonic motility in SSc. However, for future studies investigators may consider using potential surrogate measures of motility, such as anal endosonography and/or manometry. This may help elucidate the effects (if any) of intestinal dysmotility on microbial composition in SSc, or vice versa. Future studies assessing dietary intake patterns to ascertain whether certain dietary features (e.g., gluten- free, dairy-free, high animal protein, etc.) affect colonic microbiota may also provide useful insights.

An additional unanswered question of our study is whether the observed shifts in the colonic microbial consortium in SSc are present prior to the development of GI tract symptoms. A study of patients with very early SSc, such as the Very Early Diagnosis of SSc cohort[127], may help to discern whether these shifts contribute to SSc GI tract pathogenesis.

The present study also has several strengths. First, we obtained specimens during endoscopy at 2 colonic regions, which, in contrast to stool collection, facilitates in-depth investigation of the microbiota at the mucosal–luminal surface. Second, we ensured that all antibiotic and probiotic medications were withheld for at least 3 weeks prior to colonoscopy, by verifying medication lists 3 times in the month preceding the colonoscopy. Third, we performed

sophisticated statistical analyses to correct for multiple hypothesis testing, using a relatively conservative FDR *q* value. Fourth, the multivariate method used for comparing metagenomic data (LDA effect size) allowed us to control for relative species abundance and also provided an estimate of the magnitude of the observed difference. Finally, by including a clinical outcome measure, we have endeavored to discern how changes in the microbiome in SSc contribute to GI tract symptoms.

In conclusion, utilizing an innovative experimental strategy to identify bacterial genera associated with SSc we have identified, for the first time, a colonic microbial consortium unique to the SSc disease state. Using an integrative bioinformatics approach, we also demonstrated relationships between specific genera and clinical manifestations of SSc, which merit further investigation. Interestingly, many of the increased and decreased organisms identified in SSc are similar to those found in Crohn's disease, which, like SSc, has both inflammatory and fibrosing pathologic features. Further studies are needed to validate and expand on the present findings.

Tables and Figures

Table 3-1. Study patient characteristics*

	SSc participants (N=17)	Healthy controls (N=17)
Age, median (IQR) years	52.1 (46.6-63.0)	55.0 (51.0-62.0)
Female	15 (88.2)	15 (88.2)
Race		
White	9 (52.9)	15 (88.2)
Asian	2 (11.8)	0
More than one race	4 (23.5)	0
Other	2 (11.8)	1 (5.9)
Hispanic Ethnicity	6 (35.3)	1 (5.9)
Diffuse cutaneous disease	6 (35.3)	NA
SSc duration, median (IQR) years	6.6 (2.5-16.4)	NA
Antinuclear antibody positive	15 (93.8)	NA
Anti-Scl-70 positive	3 (27.3)	NA
Anticentromere antibody positive	5 (45.5)	NA
HRCT-defined ILD	12 (70.6)	NA
Current prednisone use†	3 (17.6)	NA
Current use of other immunosuppressive agent‡	3 (17.6)	NA
Current use of probiotic oral supplement§	3 (17.6)	NA
Current use of proton-pump inhibitor	10 (58.8)	NA
Scleroderma HAQ DI (023), mean ± SD	1.1±0.6	NA

GIT 2.0, mean ± SD		
Total score	0.7±0.6¶	NA
Distension/bloating score	1.5±0.9¶	NA
Diarrhea score	0.4±0.6#	NA
Faecal soilage score	0.5±0.9#	NA
Constipation score	0.7±0.7¶	NA
Emotional well-being score	0.5±0.7¶	NA
Social functioning score	0.5±0.5¶	NA

* Information on antibody status was missing in some cases (data on antinuclear antibody positivity, anti-Scl-70 positivity, and anticentromere antibody positivity are from 16 patients, 11 patients, and 11 patients, respectively). Except where indicated otherwise, values are the number (%). SSc = systemic sclerosis; IQR = interquartile range; NA = not applicable; HRCT = high-resolution computed tomography; ILD = interstitial lung disease; HAQ DI = Health Assessment Questionnaire disability index; GIT 2.0 = Gastrointestinal Tract 2.0 questionnaire.

† Dosages of ≤10 mg daily in all cases.

‡ Mycophenolate mofetil in 1 patient; azathioprine in 2 patients.

§ Culturelle (*Lactobacillus rhamnosus GG*) in 1 patient, Align (*Bifidobacterium infantis*) in 1 patient, and Florify in 1 patient. Probiotics were not consumed within 3 weeks of the colonoscopy.

¶ Score indicates moderate symptom severity[107]

Score indicates mild symptom severity[107]

Figure 3-1

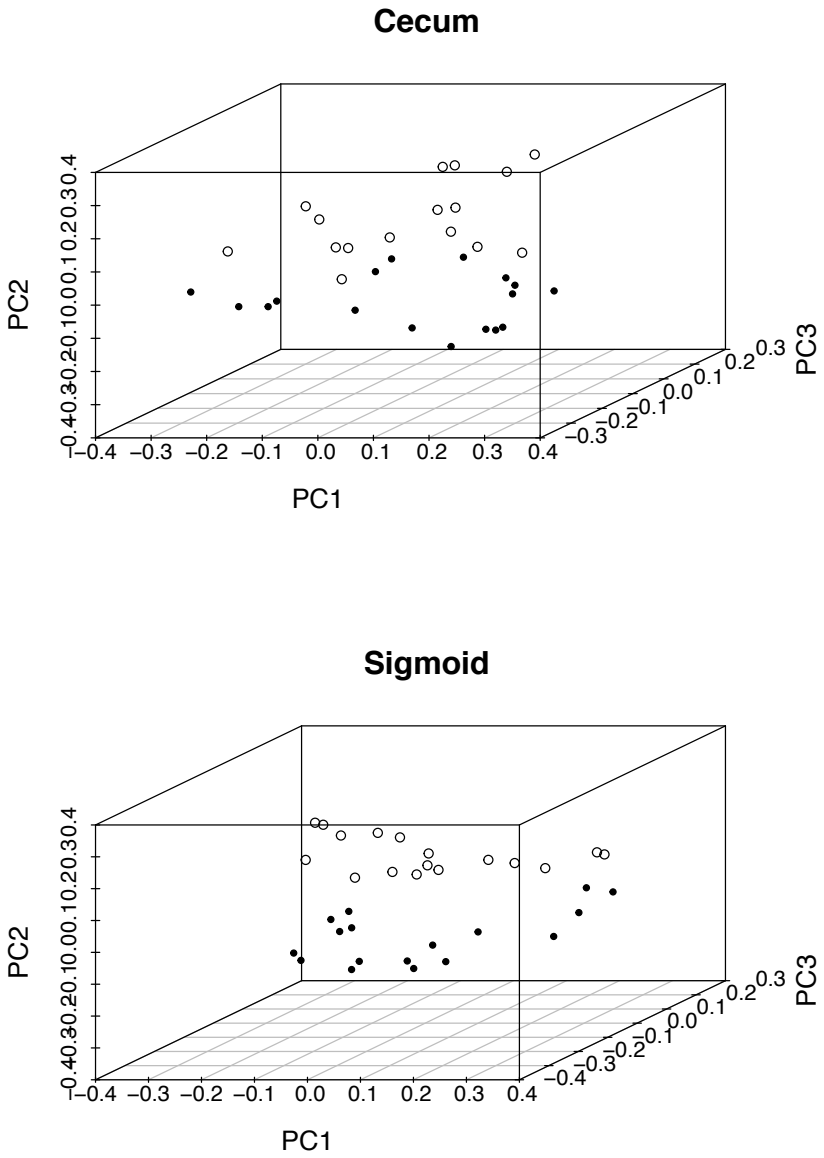


Figure 3-1. Significant differences in the beta diversity of samples from the cecum and sigmoid regions of systemic sclerosis patients (open circles) and healthy controls (solid circles), as demonstrated in principal coordinate (PC) analysis plots of the unweighted UniFrac distance. For both the cecum region and the sigmoid region, R^2=0.9, *P*=0.001 (by analysis of variance using distance matrices) between patients and controls.

Figure 3-2

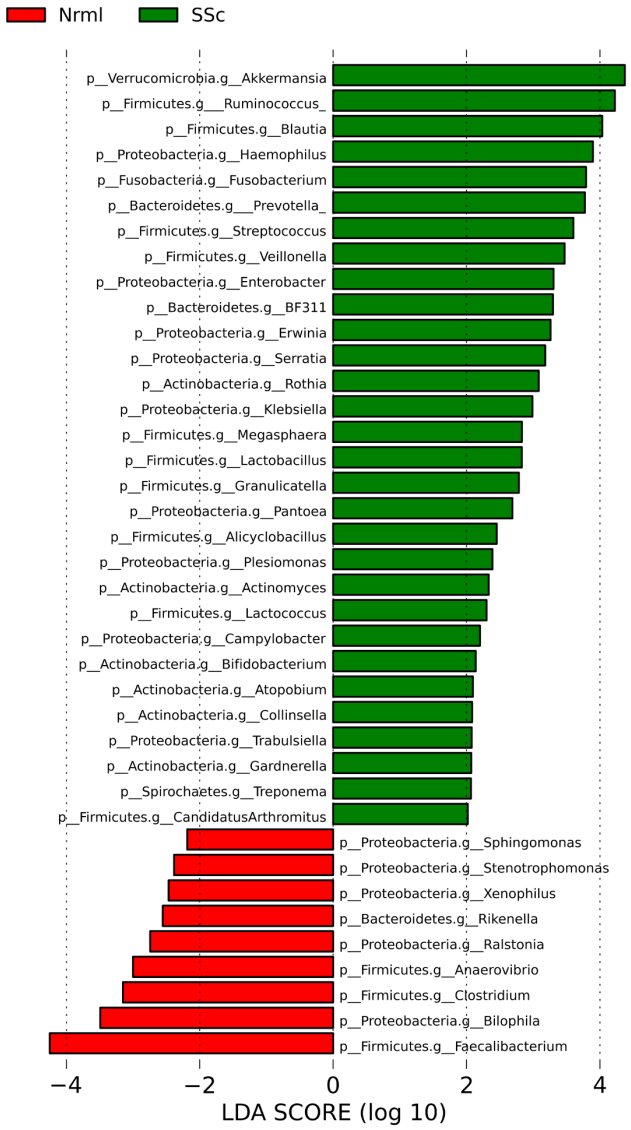


Figure 3-2. Genus-level taxa in the cecum that were associated with systemic sclerosis (SSc) (as determined by comparing with healthy subjects). Linear discriminant analysis (LDA) effect size multivariate analysis was used to identify significant associations (q<0.1), and LDA was used to calculate the effect size for these associations. Negative (red) and positive (green) effect sizes denote genera that were decreased and increased, respectively, in patients with SSc. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/journal/doi/10.1002/art.39572/abstract.

Figure 3-3

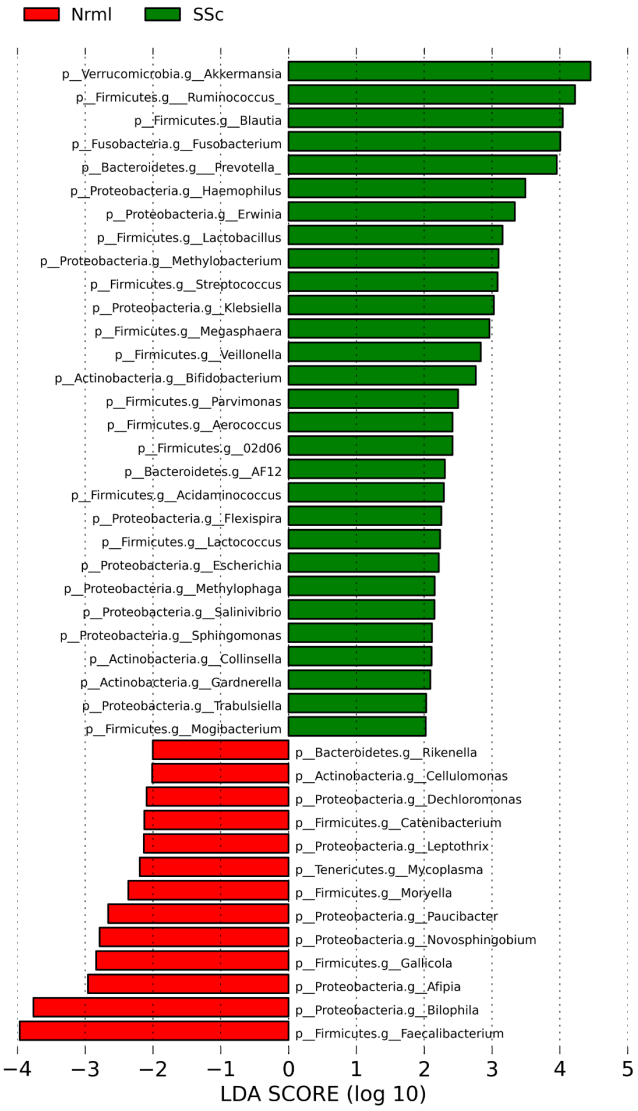


Figure 3-3. Genus-level taxa in the sigmoid that were associated with systemic sclerosis (SSc) (as determined by comparing with healthy subjects). Linear discriminant analysis (LDA) effect size multivariate analysis was used to identify significant associations (q<0.1), and LDA was used to calculate the effect size for these associations. Negative (red) and positive (green) effect sizes denote genera that were decreased and increased, respectively, in patients with SSc. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/journal/doi/10.1002/art.39572/abstract.

Figure 3-4

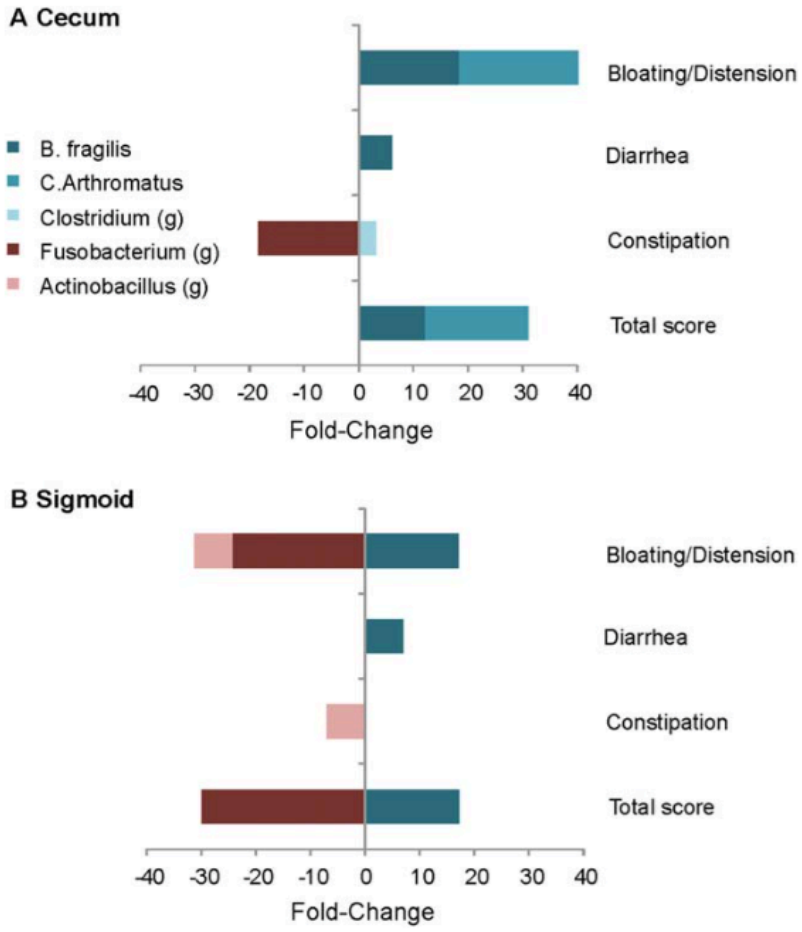


Figure 3-4. Bacterial taxa associated with the Gastrointestinal Tract 2.0 (GIT 2.0) total score and individual domains. Patients were dichotomized into low (none-to-mild) or high (moderate-to-severe) disease severity groups for the GIT 2.0 total score and the individual domains bloating/distension, diarrhea, and constipation. Multivariate analysis with differential expression analysis for sequence count data was used to identify microbial taxa that were significantly associated with low versus high disease severity and to calculate the fold change between groups. Positive and negative change scores denote organisms whose abundance was increased and decreased, respectively, in low disease severity groups. *B. fragilis* =*Bacteroides fragilis*; *C. arthromitus* =*Candidatus arthromitus*; g=genus-level taxa.

Figure 3-5

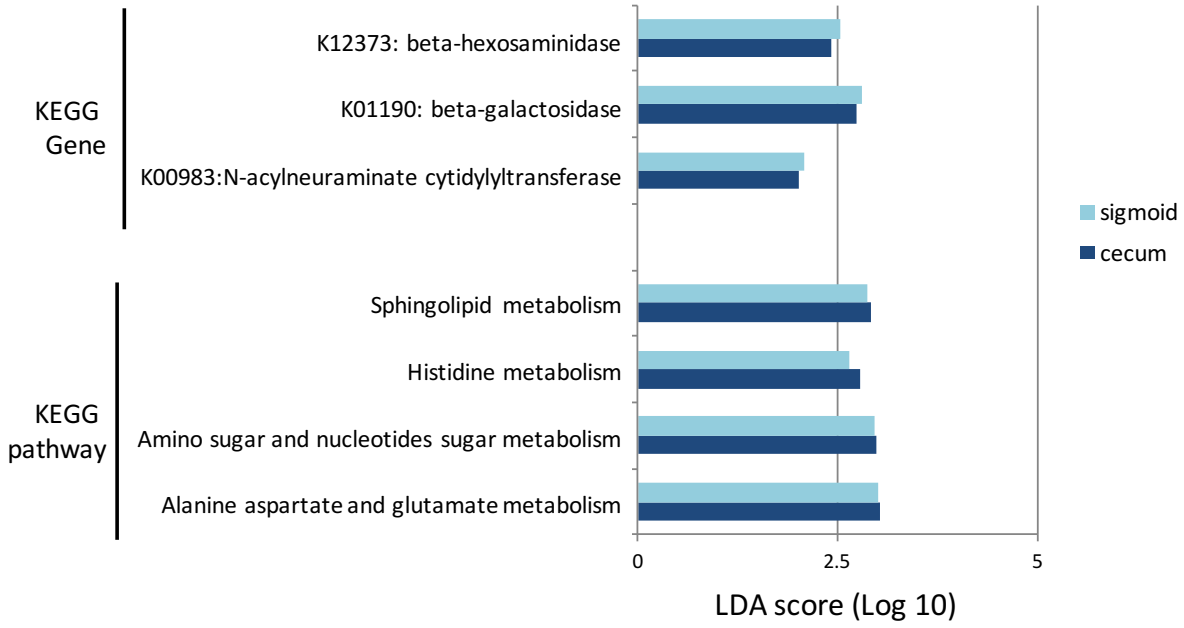


Figure 3-5. Significant differences in metagenomic content in the sigmoid and cecum regions between systemic sclerosis (SSc) patients and healthy controls. Linear discriminant analysis (LDA) effect size multivariate analysis was used to identify significant associations (q<0.1), and LDA was used to calculate the effect size for these associations (see Supplementary Table 3-1, available on the Arthritis & Rheumatology website at http://onlinelibrary.wiley.com/doi/10.1002/art.39572/abstract). A positive LDA score indicates genes and pathways that were decreased in SSc patients compared with healthy controls. KEGG = Kyoto Encyclopedia of Genes and Genomes. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/journal/doi/10.1002/art.39572/abstract.

CHAPTER 4

Systemic sclerosis is associated with specific alterations in gastrointestinal microbiota in two independent cohorts

Abstract

Objective. To compare faecal microbial composition in patients with systemic sclerosis (SSc) from 2 independent cohorts with controls and to determine whether certain genera are associated with SSc gastrointestinal tract (GIT) symptoms.

Design. Adult patients with SSc from the University of California, Los Angeles (UCLA) and Oslo University Hospital (OUH) and healthy controls participated in this study (1:1:1). All participants provided stool specimens for 16S rRNA sequencing. Linear discriminant analysis effect size demonstrated genera with differential expression in SSc. Differential expression analysis for sequence count data identified specific genera associated with GIT symptoms as assessed by the GIT 2.0 questionnaire.

Results. The UCLA-SSc and OUH-SSc cohorts were similar in age (52.1 and 60.5 years, respectively), disease duration (median (IQR): 6.6 (2.5–16.4) and 7.0 (1.0–19.2) years, respectively), gender distribution (88% and 71%, respectively), and GIT symptoms (mean (SD) total GIT 2.0 scores of 0.7 (0.6) and 0.6 (0.5), respectively). Principal coordinate analysis illustrated significant microbial community differences between SSc and controls (UCLA: p=0.001; OUH: p=0.002). Patients with SSc had significantly lower levels of commensal genera deemed to protect against inflammation, such as *Bacteroides* (UCLA and OUH), *Faecalibacterium* (UCLA), *Clostridium* (OUH); and significantly higher levels of pathobiont genera, such as *Fusobacterium* (UCLA), compared with controls. Increased abundance of *Clostridium* was associated with less severe GIT symptoms in both cohorts.

Conclusions. The present analysis detected specific aberrations in the lower GIT microbiota of patients with SSc from 2 geographically and ethnically distinct cohorts. These findings suggest that GIT dysbiosis may be a pathological feature of the SSc disease state.

Summary box

What is already known about this subject?

- Gastrointestinal tract dysfunction affects over 90% of patients with systemic sclerosis.
- The pathogenesis of lower gastrointestinal tract dysfunction in systemic sclerosis is largely unknown.
- Emerging evidence suggest that gastrointestinal tract dysbiosis may be a feature of the systemic sclerosis disease state.

What are the new findings?

- The study found specific alterations in the gastrointestinal microbiota in two independent systemic sclerosis cohorts.
- The extent of dysbiosis (i.e. alterations in the intestinal microbiota) appeared greatest in patients from the American SSc cohort compared with those from the Norwegian SSc cohort.
- Specific genera were associated with severity of gastrointestinal tract symptoms.

How might it impact on clinical practice in the foreseeable future?

- Currently, few effective treatment options exist for managing lower gastrointestinal tract symptoms in patients with systemic sclerosis. If specific genera are found to contribute the gastrointestinal tract phenotype in systemic sclerosis, such genera could provide specific targets for intervention to avert or treat this important clinical dimension of systemic sclerosis.

Introduction

The majority of patients with systemic sclerosis (SSc) experience gastrointestinal tract (GIT) dysfunction.[90,91] Symptoms of lower GIT involvement[92] adversely affect quality of life and social functioning.[93,94] Unfortunately, no effective treatment options exist for eliminating these disruptive symptoms, largely because the pathogenesis of this dimension of SSc is poorly understood.

GIT dysbiosis occurs in a number of chronic inflammatory conditions, including inflammatory bowel disease (IBD),[100-103] and recent studies suggest that alterations in GIT microbiota may be a feature of the SSc disease state.[92,128] Our prior study demonstrated that patients with SSc had decreased abundance of beneficial human commensal genera known to produce key energy metabolites and anti-inflammatory molecules for mucosal health (eg, *Faecalibacterium*; *Clostridium*) with a concurrent increase in potential pathobiont genera (eg, invasive γ-*Proteobacteria*; *Fusobacterium*), compared with controls.[128]

This study also demonstrated that increased abundance of Clostridium and decreased abundance of Fusobacterium were independently associated with decreased GIT symptoms.[128] To further explore whether alterations in microbial composition may serve as a pathogenic factor in SSc-GIT dysfunction, the present study examined the lower GIT microbiota from two independent SSc cohorts. Whereas our prior study examined colonic mucosal microbes via colonoscopy, the present study examined stool specimens from the same patients with SSc.[128] Assessment of stool specimens may reflect other aspects of the gut microbiota (eg, metabolic features) and can be obtained by a less invasive approach.

The objectives of this study, not previously explored in the SSc literature, were to: (1) establish whether analysis of stool specimens can distinguish the GIT microbiota of patients with

SSc versus healthy controls; (2) compare the GIT microbiota of patients with SSc from two geographically independent cohorts; and (3) evaluate the hypothesis that specific microbial genera are associated with SSc-GIT symptom severity.

Materials and Methods

Study Participants

Patient participants were consecutively enrolled from the outpatient rheumatology clinics at University of California, Los Angeles (UCLA) and Oslo University Hospital (OUH). Eligible participants included adult (≥18years) patients with SSc, according to the 2013 American College of Rheumatology/European League Against Rheumatism Classification Criteria for SSc.[129] Exclusion criteria included IBD, inability to withstand from taking an antibiotic and a probiotic for at least 3weeks prior to the stool collection. Patients were allowed to continue taking a proton pump inhibitor medication because this agent exerts negligible effects on colonic microbiota[104].

Healthy controls stool specimens were obtained from the UCLA Specialized Center of Research in Neurovisceral Sciences and Women's Health Repository, which consists of stool specimens from healthy adult participants (12 men and 9 women) who do not have GIT symptoms. These control participants were different than the control participants used in our prior study.[128] We attempted to match 17 of these healthy participants by age and sex with the UCLA-SSc cohort patients (1:1); however, the healthy control cohort had more men than women so gender matching was not always possible. These healthy controls specimen also served as the controls for the OUH-SSc cohort.

The UCLA Institutional Review Board and the Regional Committee of Health and Medical Research Ethics in Norway approved the study protocol and written informed consent was obtained from each participant.

Specimen procurement and processing

Participants at both sites collected stool specimens using our previously published, standard collection method[130] and immediately froze the specimen. Frozen specimens were

subsequently transferred on ice to the participant's study centre and stored at −80°C. OUH specimens were shipped overnight on dry ice to UCLA for further processing. Please see the online supplementary appendix for complete details of specimen procurement and processing.

16S rRNA gene sequencing and microbial composition analysis

The microbiota from the stool specimens were profiled by multiplex sequencing for bacterial rRNA genes using an Illumina HiSeq 2500 (Illumina, San Diego, California, USA) sequence technique. All samples (UCLA-SSc, OUH-SSc, controls) were analyzed simultaneously at UCLA to avoid any batch effects. The exact details of this approach have been outlined in our prior publication,[103,128] and are summarized in the online supplementary appendix.

Faecal calprotectin

Faecal calprotectin concentrations were measured in duplicate from the stool specimens of the UCLA-SSc and healthy control participants using a commercial kit (KTR-849, Epitope Diagnostics, San Diego, California, USA) according to the manufacturer's instructions.

Assessment of GIT symptoms

On the day of their stool collection, the UCLA and OUH-SSc participants completed the GIT 2.0, a valid measure of GIT symptom severity in SSc patients[107]. The questionnaire consists of seven domains and has been translated and validated in several languages, including Norwegian. The GIT 2.0 can furthermore discriminate between self-rated severity (i.e. none/mild versus moderate versus severe/very severe disease) of GIT involvement[107].

Statistical and bioinformatics analyses

Statistical approach

Analyses were performed using R V.3.1.2. Mean and SDs were used to describe continuous parametric data; median and IQRs were used to describe continuous non-parametric data. All tests

were two-sided with a 0.05 α level. To compare the microbial communities of SSc versus control samples, α and β diversity were analysed. The α diversity represents the complexity of composition within members of a group, and β diversity represents the between-participant similarity of microbial composition and enables the identification of differences between samples within a group. We also examined taxo- nomic differences at specific levels (ie, phylum, genus) using linear discriminant analysis effect size (LefSe). We used the false discovery rate (FDR) of Benjamini and Hochberg[110], and a significant association was defined at the FDR q-value ≤0.1. Please see the online supplementary appendix for further details.

Results

Participant characteristics

Seventeen UCLA-SSc cohort patients (88% women; median age 52.1 years) and 17 OUH-SSc cohort patients (71% women; median age 60.5years) were enrolled (Table 4-1). The healthy control cohort (N=17) was younger than the SSc cohorts (median age 29.0 years), and had a predominance of women (60% women). The OUH-SSc cohort was mostly Caucasian (94%); whereas the UCLA-SSc and healthy control cohorts had more Hispanics (35% and 24%, respectively, vs 6% in the OUH-SSc cohort). The mean body mass index (BMI) was similar in the two SSc cohorts (~24 for both cohorts) and slightly higher in the controls (28.1), although the difference in BMI between SSc participants and healthy controls was not significant.

UCLA-SSc and OUH-SSc cohort patients had similar SSc disease durations, proportion of patients with diffuse cutaneous sclerosis, and autoantibody profiles (Table 4-1). However, more patients in the UCLA-SSc cohort (71%) had clinically significant interstitial lung disease (ILD) by high-resolution CT[131], compared with the OUH-SSc cohort (47%).

The UCLA-SSc and OUH-SSc cohorts also had similar GIT symptom profiles (Table 4-1). The mean GIT 2.0 scores indicated moderate symptom severity[107] for the total score, as well as for the following domains: distension/bloating, emotional well-being and constipation (Table 4-1). The mean GIT 2.0 scores for faecal soilage and diarrhea indicated mild symptom severity[107] (Table 4-1). Six UCLA-SSc participants (35%) had taken antibiotics in the 3 months preceding stool collection; the mean time between cessation of antibiotics and stool collection was 6.5 weeks (range 4–12 weeks). None of the OUH-SSc cohort patients had taken antibiotics 3 months prior to stool sampling.

Only three of the UCLA-SSc cohort patients had ele- vated levels of faecal calprotectin, which was defined as >50 μg/g. None of the control participants had elevated levels of faecal calprotectin.

Colonic microbial diversity in SSc

After the operational taxonomic unit (OTU) selection process, a total of 231, 250 and 184 species-level OTUs were generated from the UCLA-SSc, OUH-SSc and healthy control cohorts, respectively. The α diversity (ie, the complexity of microbial composition) was similar between the UCLA-SSc and control participants (Figure S4-1) and between the OUH-SSc and control participants (Figure S4-2). There were also no significant differences in α diversity between the UCLA-SSc and OUH-SSc participants (Figure S4-3).

The β diversity was then computed to determine whether SSc and control participants differed in their microbial composition. The principle coordinate analysis visualization of the weighted Unifrac distances and analysis of variance using distance matrices (Adonis) of this comparison are shown in Figure 4-1. In the UCLA-SSc (R^2 0.355; p=0.001) and the OUH-SSc (R^2 0.126; p=0.002) cohorts, the microbial composition among patients with SSc was significantly different than healthy controls. The magnitude of this difference appeared greater between the UCLA-SSc cohort and healthy con- trols than between the OUH-SSc cohort and healthy controls based on the R^2 values and as illustrated in Figure 4-1. The β diversity also differed between the UCLA-SSc and OUH-SSc cohort patients (R^2 0.145; p=0.002).

Colonic microbial genera are differentially abundant in patients with SSc

To begin to define the compositional differences between patients with SSc and healthy controls predicted by the β diversity analysis, the relative abundances of microbial composition at different taxonomic levels were computed (Figure 4-2). The two predominant phyla in SSc

samples were *Bacteroidetes* (UCLA: 21.3%; OUH: 45.0%; controls: 63.2%), and *Firmicutes* (UCLA: 63.5%; OUH: 42.8%; controls: 33.0%). The relative abundance of *Bacteroidetes* was significantly decreased in the UCLA-SSc cohort compared with the healthy control cohort ($p<0.0001$) and in the OUH-SSc cohort compared with the healthy control cohort ($p=0.009$).

LefSe multivariate analysis demonstrated significant taxonomic differences between SSc and healthy controls at the genus level for the UCLA-SSc (Figure 4-3) and OUH-SSc (Figure 4-4) cohorts. The LefSe analysis comparing the UCLA-SSc and healthy control cohorts yielded more significant genus-level differences than the analysis comparing the OUH-SSc and healthy control cohorts.

Commensal genera such as *Faecalibacterium* (UCLA), *Clostridium* (OUH) and *Bacteroides* (UCLA, OUH) were depleted in patients with SSc; whereas, *Fusobacterium* (UCLA), Ruminococcus (UCLA) and the uncommon γ-Proteobacteria, Erwinia (UCLA), were enriched in patients with SSc, compared with health controls. Consistent with our prior study[128], *Lactobacillus* (considered to be a commensal genus[113]) was found in greater abundance in patients with SSc compared with controls in the UCLA and OUH cohorts (Figure 4-3 and 4-4).

When comparing the UCLA-SSc and OUH-SSc cohorts, the LefSe analysis revealed that specific commensal genera, including *Faecalibacterium* and *Bacteroides*, were significantly more abundant in the OUH-SSc cohort patients compared with the UCLA-SSc cohort patients (fold change scores of 4.85 and 3.75, respectively).

Specific microbial genera and species are associated with SSc-GIT symptoms

Differential expression analysis for sequence count data multivariate analysis identified specific microbial genera associated with SSc-GIT symptom severity based on the scores for the total GIT 2.0 and individual domains (eg, constipation, diarrhea or distension/bloating). For the

total score and each domain, patients were dichotomised into low (none-to-mild) or high (moderate-to-severe) GIT symptom severity groups, and the fold change and q-values for organisms reaching significance were tabulated (Table S4-1). The frequency of distribution of patients in each disease severity category were fairly well balanced (N for low vs high groups): total score (18 vs 16); constipation (16 vs 18); distention/bloating (8 vs 26) and diarrhea (18 vs 15).

In both SSc cohorts, *Clostridium* was more abundant in patients with low GIT symptom severity (for the total GIT score and the bloating/distension domain) compared with patients with high GIT symptom severity (Figure 4-5 and see Table S4-1). *Lactobacillus* was more abundant in patients with none-to-mild constipation compared with patients with moderate-to-severe constipation. In contrast, *Prevotella* was more abundant in patients with moderate-to-severe GIT symptom severity (bloating/distention domain, diarrhea domain) compared with patients with none-to-mild GIT symptom severity.

Specific microbial genera are associated with SSc-ILD phenotype

To ask if gut microbiome is associated the ILD phenotype in SSc, we included 34 adult SSc patients from UCLA and OUH cohorts. They were grouped into two clinical phenotypes based on their ILD records: Yes, the patients have developed ILD; No, the patients have not yet developed ILD. The method of linear discriminant analysis (LDA) effect size were used to compute the features most likely to explain the differences between clinical phenotypes (i.e. ILD record) by coupling with the features of sample sources (i.e. UCLA v.s. OUH). Our result showed five genera are positively associated with ILD phenotype with the effect sizes all greater than 3 (Table 4-2). Genera *Blautia* and *Dorea* in *Lachnospiraceae* family have significant p-value below

0.001. Genera *Rothia* and *Collinsella* in *Actinobacteria* phylum have p-value below 0.05. We have found several genera that are associated with SSc-ILD phenotype.

Discussion

Consistent with our prior findings[128], the present study demonstrated that the SSc disease state is associated with alterations in the GIT microbial consortium. Using data from two independent SSc cohorts, we found that specific faecal microbial taxa were enriched or depleted among patients with SSc compared with healthy controls. We also demonstrated that specific taxa were associated with severity of GIT symptoms in both SSc cohorts.

In terms of phylum-level differences, the relative abundance of *Bacteroidetes* was significantly decreased in the UCLA and OUH-SSc cohorts compared with the healthy control cohort. This finding is of interest in light of the recent evidence suggesting that the ratio of *Firmicutes* to *Bacteroidetes* may have an important effect on human health[3,132]. Patients with SSc in the UCLA cohort had the lowest proportion of *Bacteroidetes* (21.3%), followed by patients with SSc in the OUH cohort (45.0%) and healthy controls (63.2%).

Consistent with the observed decrease in *Bacteroidetes*, lower level taxonomic analysis revealed that both SSc cohorts had significantly lower levels of the commensal genera *Bacteroides*, which is thought to protect the host from mucosal inflammation and against colonization of pathogenic species. Low relative abundance of *Bacteroides* is associated with increased disease activity in other chronic inflammatory diseases, such as Crohn's disease (CD)[133].The fold change scores for the UCLA-SSc and OUH-SSc cohorts relative to controls was nearly 5 signifying a substantial shift in this genus in the SSc disease state.

Concordant with our prior study using colonic lavage specimens[128],we found that *Faecalibacterium* (UCLA) and *Clostridium* (OUH) were also depleted in SSc. Low abundance of these genera is associated with increased disease activity in IBD[134-136].As *Clostridium* species have been found to induce the expansion of regulatory T cells[137], it is tempting to speculate that

depletion of these taxa may reduce the number of regulatory T cells, which may play a role in SSc pathogenesis[138]. It is important to note, however, that species within this genus vary in terms of their metabolic properties. Future studies may explore the relationship between the abundance of specific species within the *Clostridium* genus and levels of circulating regulatory T cells.

In line with the potential immunoregulatory properties of *Clostridium*, we found that increased abundance of this genus was associated with decreased GIT symptoms (for the total GIT symptom severity score and the bloating/distension domain score). Our prior study also reported a link between Clostridium and GIT symptoms[128]. If future studies demonstrate that *Clostridium* attenuates local or systemic inflammation in SSc (through effects on T regulatory cells), therapy aimed at increasing the abundance or activity of species in this genus may improve SSc disease activity.

In contrast, *Prevotella* was more abundant in patients with more severe GIT symptoms (bloating/distention domain, diarrhea domain scores). This genus appears to have effects on intestinal T helper 17 cells[139], and it has been found to be increased in faecal samples from patients with CD[102,140], and new-onset rheumatoid arthritis[6].

Also consistent with our prior study, the abundance of the pathobiont genera, *Fusobacterium*, and uncommon *γ-Proteobacteria* (eg, *Erwinia*) were higher in the UCLA-SSc faecal samples compared with controls. *Fusobacterium* species are increased in patients with CD compared with healthy controls and have been mechanistically proven to enhance proinflammatory responses[16,141].

We are intrigued by the increase of *Lactobacillus* in the UCLA-SSc and OUH-SSc cohorts compared with healthy controls. This finding has been observed now in three different SSc cohorts (UCLA (colonic lavage specimens, faecal specimens), OUH (faecal specimens), Swedish cohort

(faecal specimens))[128,142]. We do not attribute this expansion to probiotic use. First, due to the exceptionally larger scale of the intestinal microbiome, ingestion of lactobacillus-bearing probiotics does not measurably alter faecal microbial composition[143]. Second, very few patients in the present cohorts were taking probiotics on a regular basis, and none were taking probiotics within 3 weeks of the stool collection. Deemed a beneficial commensal genus by several host inflammatory and physiological end points in animal models, its abundance is typically reduced in chronic inflammatory human diseases[144]. We further note that members of this genus and their products alter the broader intestinal microbial ecosystem gene expression and function, which may be an intermediary of their action on systemic host phenotypes[143,145]. Since many patients with SSc consume commercial probiotic supplements that commonly include *Lactobacillus* species, the utility of *Lactobacillus* as a biomarker or potential intervention in patients deficient for this genus warrants further study.

The overall extent of dysbiosis appeared greatest in the patients with SSc from UCLA compared with patients with SSc from the OUH cohort. Possible explanations for this disparity include different SSc phenotypes between the cohorts. For instance, more patients with SSc in the UCLA cohort had ILD than in the OUH cohort. Accumulating evidence suggests that GIT microbiota plays a central role in regulating immune responses in pulmonary disease, such as asthma[146]. It is possible that patients with SSc-ILD have distinct microbiota features compared with patients with SSc without ILD. Indeed, a recent analysis by Andréaasen et al[142] found that among patients with SSc from a single-centre cohort, the extent of GIT dysbiosis was more severe in patients with ILD compared with patients without ILD. Furthermore, although there was a similar ratio of patients with diffuse versus limited cutaneous disease in each SSc cohort, there

may have been subtle differences in the peak severity of cutaneous sclerosis between the two cohorts.

In addition to SSc disease phenotype, genetic variation may account for some of the observed differences in severity of dysbiosis between the two SSc cohorts. The majority of the OUH-SSc cohort was Caucasian with a genetic Norwegian background; whereas the UCLA cohort was more ethnically diverse. A study for ulcerative colitis reported that ethnicity stratifies a small proportion (7.5%) of observed variation in β diversity[147].Dietary variation may also contribute to the differences observed between the two SSc cohorts.

In contrast with a prior research[142,148], faecal calprotectin concentrations were not elevated in the majority of patients with SSc (from the UCLA cohort), although different kits/protocols may have been used in prior studies. While the accumulating evidence suggests that faecal calprotectin is seen in inflammatory intestinal diseases (eg, IBD, colorectal cancer), future studies are needed to determine whether this protein is a marker of disease activity or the extent of dysbiosis in SSc.

The findings of the present study should be considered within the context of certain limitations. First, the present study is cross-sectional. It is unclear whether the relationships observed between specific genera and GIT symptoms are causational and/or persist with time. Second, the sample size for both SSc cohorts is small. Despite the small sample size, we observed several significant associations consistent with the associations reported in our prior study[128], suggesting that the present findings are unlikely to be due to chance alone. Third, the same healthy control cohort was used for the UCLA and OUH-SSc cohorts. It may have been more ideal to use a Norwegian healthy control cohort for the OUH-SSc analyses to control for genetic and dietary variation.

In addition, it may be prudent to include a measure of GIT transit in future SSc-GIT microbiome studies. This study did not assess GIT transit as there is no valid measure of GIT transit in SSc; however, diarrhea and constipation are intrinsically related to GIT transit. Future studies of this nature may consider using potential surrogate measures of GIT transit, such as wireless capsule endoscopy. Finally, although it is a challenge to quantify dietary patterns due to patient recall bias, future studies could also consider the impact of diet on microbial composition in SSc. Accordingly, a distinct study design will be required to assess the differential microbiota associated with disease state or phenotypic progression.

The present study also has important strengths. By studying two independent SSc cohorts, we have minimized the risk of type 1 error that often ensues in discovery cohort analyses of this nature. In addition, by performing all of the sequencing analyses simultaneously, we have eliminated the possibility of a batch effect. We also took substantial caution to ensure that all patients withheld medications, such as antibiotics and probiotics, at least 3 weeks prior to the stool collection, by verifying medication lists three times in the month preceding the collection.

To conclude, this study identified bacterial genera associated with SSc using sophisticated sequencing analyses of faecal specimens. Many of the genus-level gains and losses appreciated in the faecal specimens examined in the present study were also observed in our prior study using colonic-mucosal specimens from the same patients with SSc (but different control patients)[128]. Using two independent SSc cohorts, this study also uncovered relationships between specific genera and severity of SSc-GIT symptoms, which merit further investigation. While larger studies are needed to validate and expand on these findings, the present study is the first to characterize the lower GIT microbiota in two geographically and ethnically distinct SSc cohorts.

Tables and Figures

Table 4-1. Systemic sclerosis (SSc) participant characteristics

	UCLA-SSc participants (N=17)	**OUH-SSc participants (N=17)**
Age (years)	Median 52.1 (IR 46.6, 63.0)	60.5 (IR 46.0, 71.0)
Female	15 (88.2%)	12 (70.6%)
Race		
White	9 (52.9%)	16 (94.1%)
Asian	2 (11.8%)	0
More than one race	4 (23.5%)	0
Other	2 (11.8%)	0
Hispanic	6 (35.3%)	1 (5.9%)
Diffuse cutaneous disease	6 (35.3%)	7 (41.2%)
SSc disease duration (years)	Median 6.6 (IR 2.5, 16.4)	Median 7.0 (IR 1.0, 19.2)
ANA positive	15/16 (93.8%)	17/17 (100%)
Scl-70 positive	3/11 (27.3%)	4/17 (23.5%)
Anti-centromere positive	5/11 (45.5%)	9 (52.9%)
HRCT-Defined Interstitial Lung Disease	12/17 (70.6%)	8/17 (47.1%)
Current prednisone use*	3 (17.6%)	2 (11.8%)
Current other immunosuppressant use †	6 (35.3%)	2 (11.8%)

Current use of probiotic oral supplement ‡	3 (17.6%)	3 (17.6%)
Current use of proton pump inhibitor	10 (58.8%)	5 (29.4%)
Gastrointestinal Tract (GIT) 2.0 Total Score	Mean 0.7 (0.6) §	Mean 0.6 (0.5) §
Distension/Bloating	Mean 1.5 (0.9) §	Mean 1.2 (0.8) §
Diarrhea	Mean 0.4 (0.6) ¶	Mean 0.3 (0.3) ¶
Faecal Soilage	Mean 0.5 (0.9) ¶	Mean 0.3 (0.6) ¶
Constipation	Mean 0.7 (0.7) §	Mean 0.6 (1.1) §
Emotional Well-Being	Mean 0.5 (0.7) §	Mean 0.5 (0.7) §
Social Functioning	Mean 0.5 (0.5) §	Mean 0.4 (0.6) ¶

Values are n (%), except where otherwise noted.

***** Dosages of prednisone were ≤10 mg daily.

† Immunosuppressant medications used included mycophenolate (UCLA: N=1; OUH: N=2) and azathioprine (UCLA: N=2).

‡ Probiotic used in the UCLA-SSc cohort included Culturelle (N=1), Florify (N=1), and Align (N=1). For the OUH-SSc cohort, patients consumed probiotic enriched sour milk products (N=3). Probiotics were not consumed within 3 weeks of the stool collection.

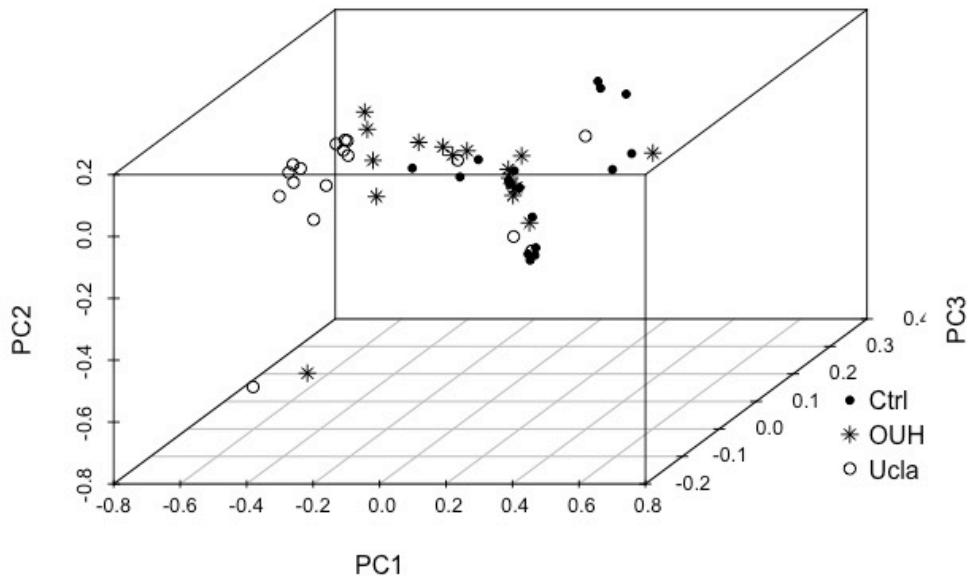
§ Score indicates moderate symptom severity.

¶ Score indicates mild symptom severity.

Table 4-2. List of bacteria genera that associated with ILD

Genera	the logarithm value of the highest mean among all the classes	the feature is discriminative	the logarithmic LDA score	P-value
g__Blautia	4.560723361	Y	4.200746601	0.000635892
g__Anaerofustis	1.207918363	Y	3.756217898	0.05804888
g__Rothia	1.32995036	Y	3.651263102	0.015062241
g__Dorea	3.836995687	Y	3.530337979	0.003004453
g__Collinsella	3.456704308	Y	3.181864475	0.024915323

Figure 4-1



Control vs OUH:
P=0.002 (R2=0.126)

Control vs UCLA:
P=0.001 (R2=0.355)

Figure 4-1. Significant differences in the β diversity of the SSc and healthy samples as demonstrated by principal coordinate analysis plots of the weighted UniFrac distance. Each dot represents a sample from a UCLA-SSc cohort patient (open circle) or a healthy control (closed circle). Each star represents a sample from a OUH-SSc cohort patient. The p values provided were calculated by analysis of variance using distance matrices. OUH, Oslo University Hospital; PCoA, principle coordinate analysis; SSc, systemic sclerosis; UCLA, University of California, Los Angeles.

Figure 4-2

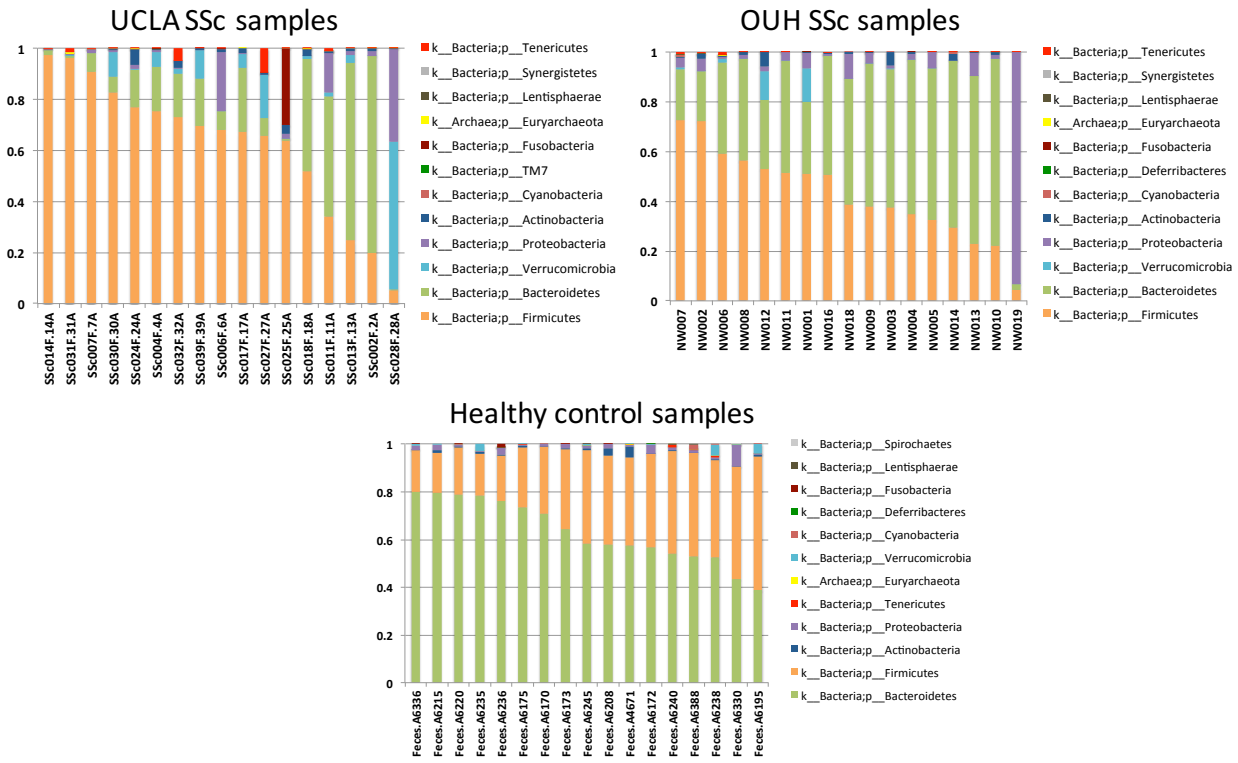


Figure 4-2. Microbial composition at the phylum level in UCLA-SSc samples (top left), OUH-SSc samples (top right) and healthy samples (bottom left). Legend provides colour coding specific to each phylum. OUH, Oslo University Hospital; SSc, systemic sclerosis; UCLA, University of California, Los Angeles.

Figure 4-3

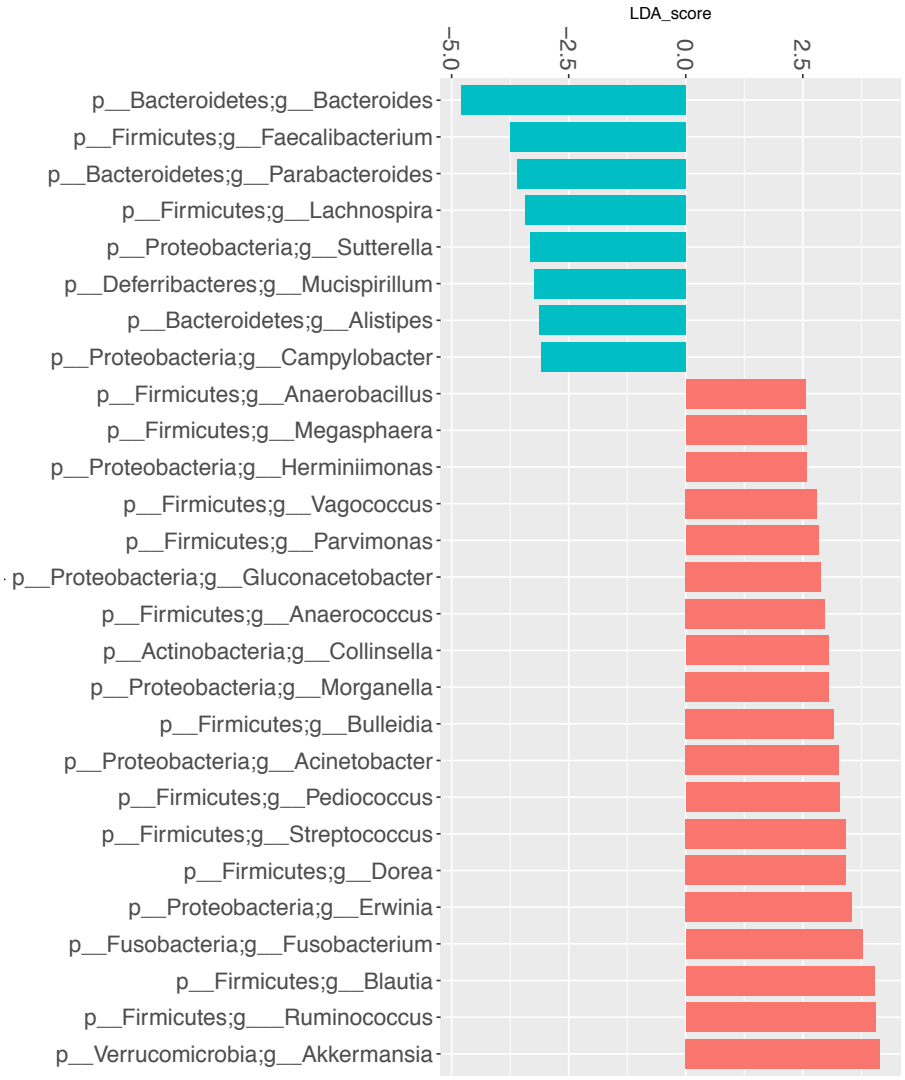


Figure 4-3. Genus-level taxa associated with UCLA-SSc cohort patients versus healthy participants. LefSe multivariate analysis was used to identify significant associations ($q<0.1$), and LDA was used to calculate the effect size for these associations. Negative and positive effect sizes denote genera decreased (blue) or increased (red) in patients with SSc, respectively. All genera with an absolute LDA score >2.5 were included in this figure. LDA, linear discriminant analysis; LefSe, linear discriminant analysis effect size; SSc, systemic sclerosis; UCLA, University of California, Los Angeles.

Figure 4-4

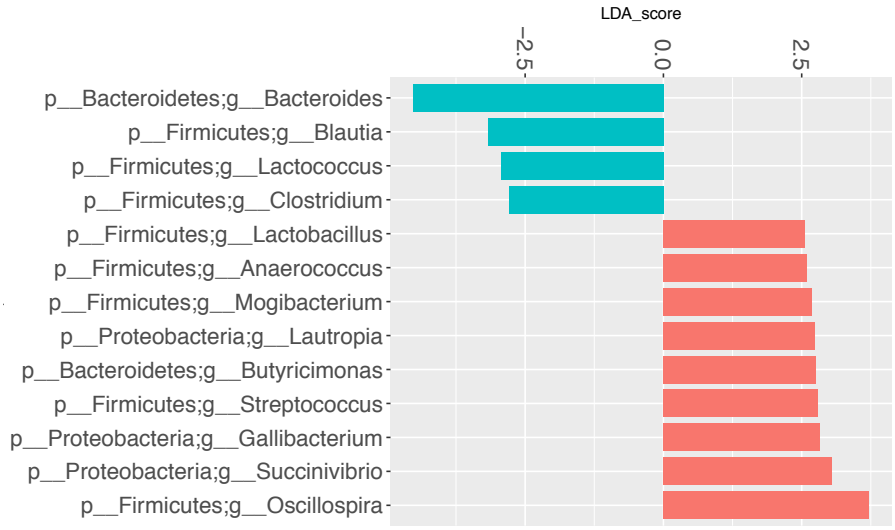


Figure 4-4. Genus-level taxa associated with OUH-SSc cohort patients versus healthy participants. LefSe multivariate analysis was used to identify significant associations ($q<0.1$), and LDA was used to calculate the effect size for these associations. Negative and positive effect sizes denote genera decreased (blue) or increased (red) in patients with SSc, respectively. All genera with an absolute LDA score >2.5 were included in this figure. LDA, linear discriminant analysis; LefSe, linear discriminant analysis effect size; OUH, Oslo University Hospital; SSc, systemic sclerosis.

Figure 4-5

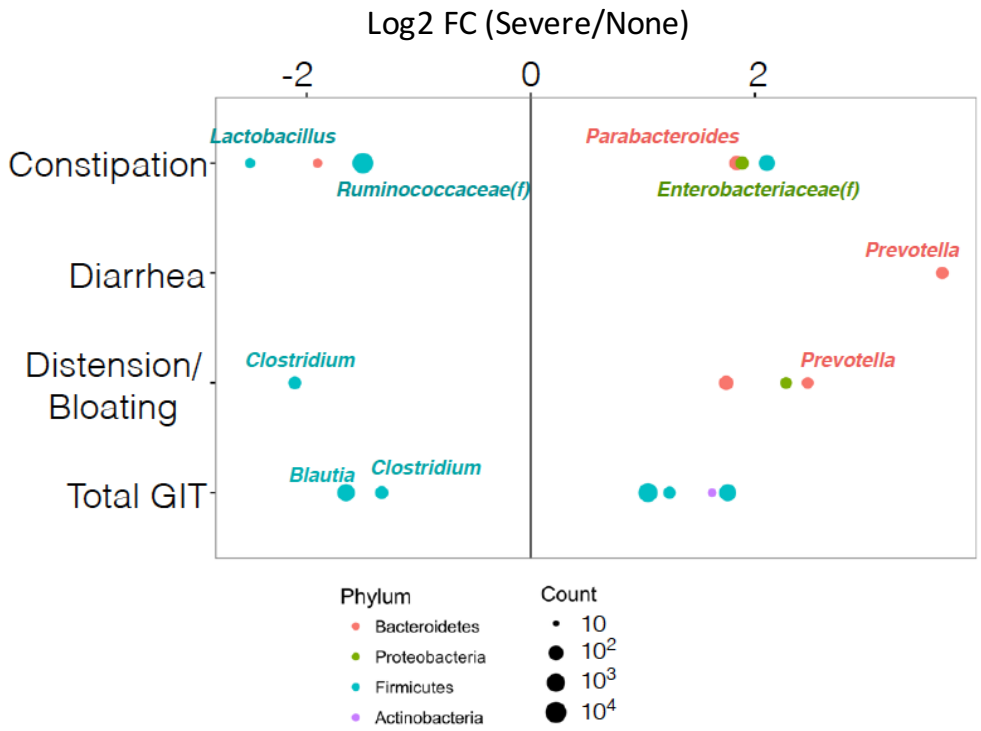


Figure 4-5. Bacterial taxa associated with GIT disease score and domains. Patients were dichotomised into low (none-to-mild) or high (moderate to severe) disease severity groups for the total GIT 2.0 score and its individual domains (constipation, diarrhoea or distension/bloating). DESeq2 multivariate analysis was used to identify microbial taxa significantly associated with low versus high groups (Table S4-1), and calculate the fold change between groups. Negative fold change scores (log2) denote organisms decreased in high disease severity groups; whereas, positive fold change scores denote organisms increased in high disease severity groups. Legend provides colour code of bacterial taxa at the phylum level; 'f" denotes family-level taxa. The size of the coloured dots represents the square root of the absolute mean counts of the OTUs at the genus level. DESeq2, differential expression analysis for sequence count data; GIT, gastrointestinal tract; OUT, operational taxonomic unit.

Supplementary Tables and Figures

Table S4-1. Genera associated with GIT 2.0 total score and individual domains

Organism	Group	Fold change	q value
Actinomyces	Total GIT score, High	-1.618	0.025
Ruminococcus	Total GIT score, High	-1.758	0.023
	Constipation, High	-2.107	0.003
Dorea	Total GIT score, High	-1.239	0.092
Parabacteroides	Constipation, High	-1.838	0.021
	Distention/bloating, High	-1.744	0.057
Undefined genus from *Enterobacteriaceae* family	Constipation, High	-1.887	0.036
Prevotella	Diarrhea, High	-3.671	<0.0001
	Distention/bloating, High	-2.472	0.025
Sutterella	Distention/bloating, High	-2.278	0.024
SMB53 from the *Clostridiaceae* family	Distention/bloating, Low	2.101	0.025
Blautia	Total GIT score, Low	1.648	0.025
Clostridium	Total GIT score, Low	1.329	0.092
	Distension/bloating, Low	2.104	0.024
Lactobacillus	Constipation, Low	2.505	0.003

Figure S4-1

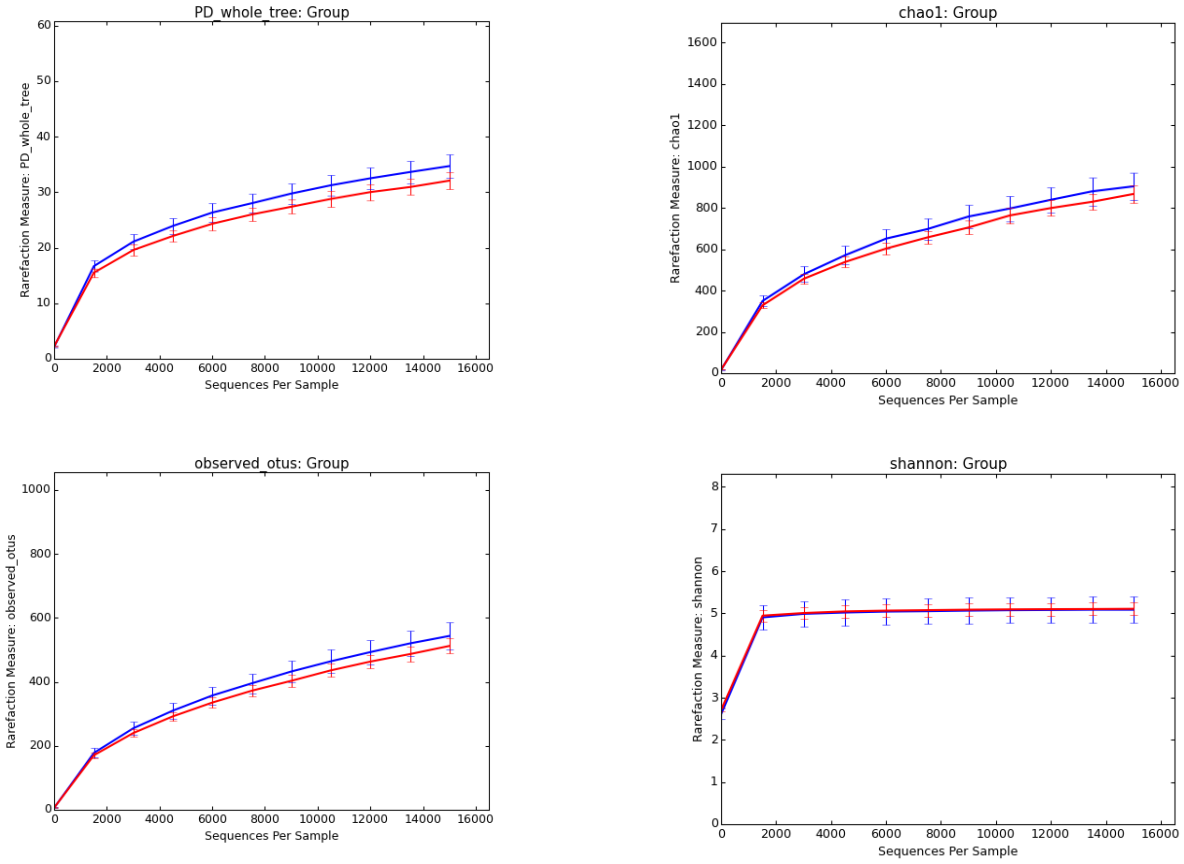


Figure S4-1. Alpha diversity of UCLA-SSc (blue line) and healthy (red line) samples. There were no significant differences in any of the alpha diversity metrics between the UCLA-SSc and healthy control samples based on the following metrics: phylogenetic diversity (p=0.330); observed species (p=0.528); Shannon index (p=0.958); and Chao 1 diversity metric (p=0.657).

Figure S4-2

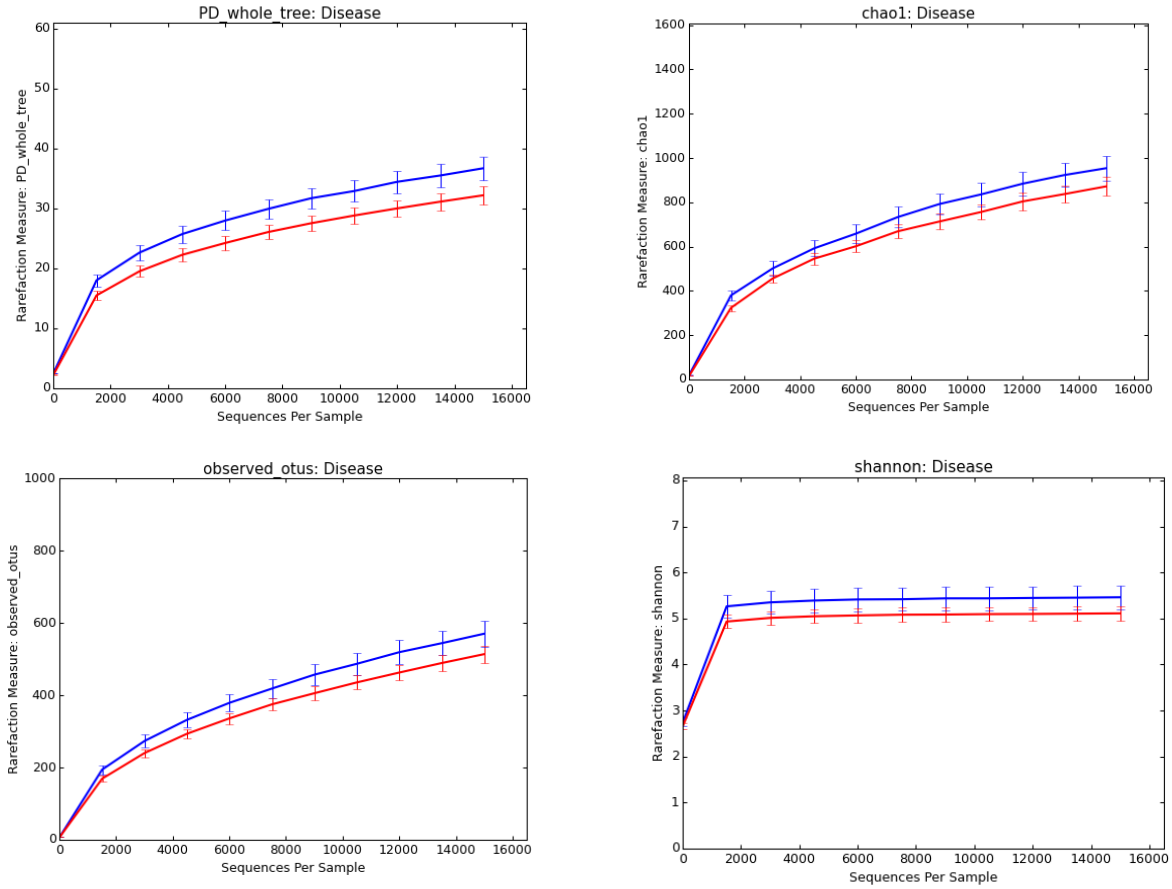


Figure S4-2. Alpha diversity of OUH-SSc (blue line) and healthy samples (red line). There were no significant differences in any of the alpha diversity metrics between the OUH-SSc and healthy control samples by the following metrics: phylogenetic diversity (p=0.078); observed species (p=0.206); Shannon index (p=0.265); and Chao 1 diversity metric (p=0.245)

Figure S4-3

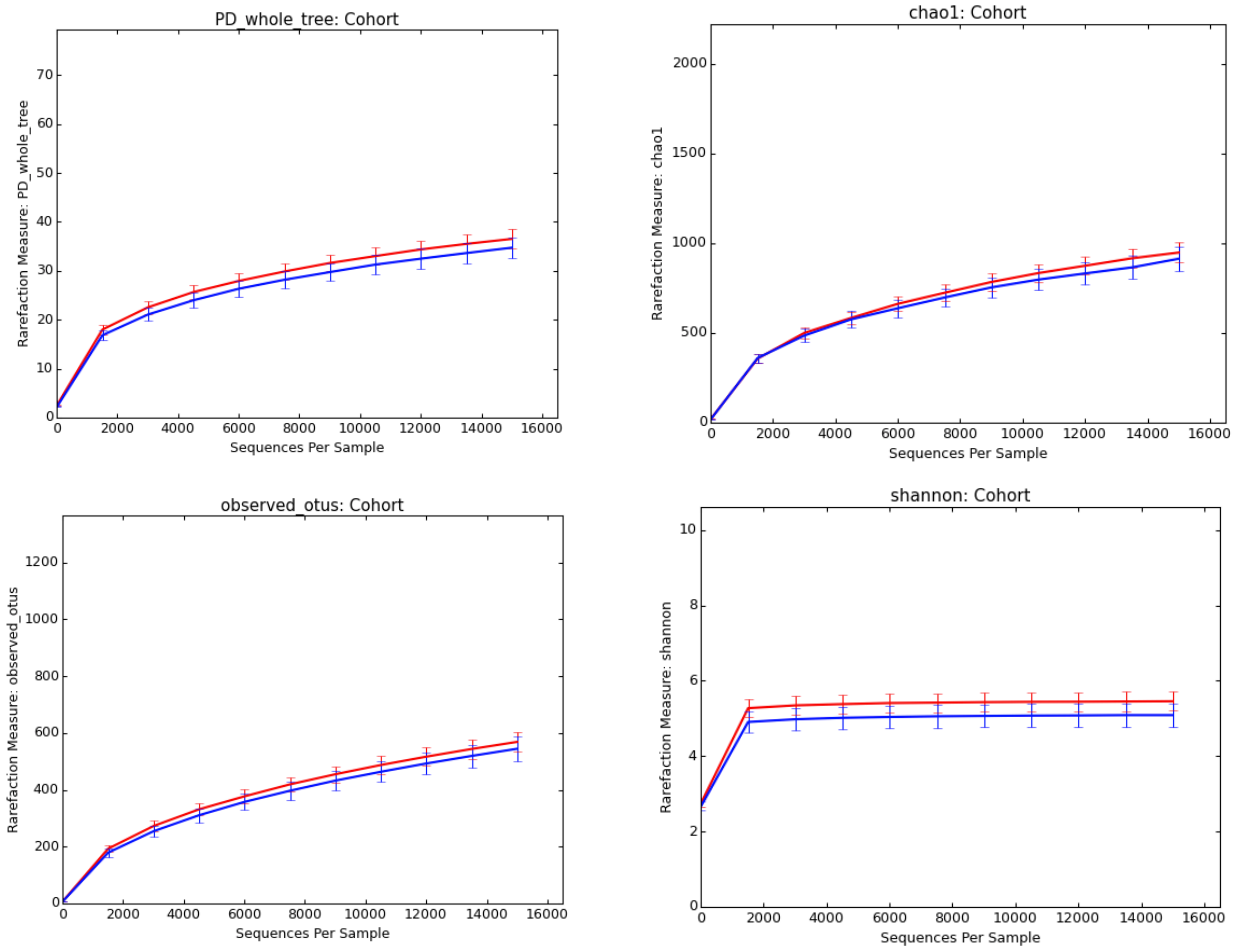


Figure S4-3. Alpha diversity of UCLA-SSc (blue line) and OUH-SSc (red line) samples. There were no significant differences in any of the alpha diversity metrics between the UCLA or OUH-SSc samples by the following metrics: phylogenetic diversity (p=0.542); observed species (p=0.680); Shannon index (p=0.381); and Chao 1 diversity metric (p=0.674).

CHAPTER 5

Conclusion

The study for host-microbiome interaction is beginning to move from association towards causality, that enables the prediction and testing of mechanisms by which the microbiome affects disease biology. In this thesis, we established a system biology screening to investigate the causal microbial factors regulating local chronic inflammatory disorder, CD, and explored the association of microbial taxa with the disease features in a systemic autoimmune disorder, SSc. In this conclusion, we reiterate the important observations, and consider the potential follow-up studies and implications.

T cell energy metabolisms affected by metabolites

In Chapter 2, we emphasized the importance of environmental metabolic cues to modulate the differentiation and activation of specific subsets of CD4+ T cells. This observation can be a critical mechanism to explain the recently discovered phenomenon that the environmental and microbial factors influence the cytokine production capacity[149-151]. We believed the other identified bioactive metabolites in this study could potentially regulate T cell intracellular metabolisms (Table 5-1). Additional investigation is required to identify the mechanisms.

Three identified bioactive metabolites, HMG-CoA, mevalonate, and isopentenyl diphosphate are the essential intermediates in an overlapping metabolic pathway, mevalonate metabolism, which suggested the potential role of this pathway in T cell activation or Th17 differentiation[152]. However, we found the divergent biological effects among the three molecules: HMG-CoA(pan-inhibition), mevalonate(pan-inhibition), and isopentenyl diphosphate (selective enhancement on IL-17A and IL-17F), indicating the potential involvement of multiple regulations. Interestingly, TCR stimulation is reported to induce expression of genes encoding HMG-CoA synthase and HMG-CoA reductase[153], the enzymes to facilitate the production of HMG-CoA and mevalonate, respectively. How the activities of these enzymes be regulated by

the levels of HMG-CoA and mevalonate in cell cultures can be the important follow-up question. Future experimentation involving the blockade of these enzymes using chemical inhibitors or genetic manipulations can be used to assess T cells or Th17 functions. In addition, isopentenyl diphosphate is the first isoprenoids, the precursors of steroid, in mevalonate-cholesterol biosynthesis pathway. It is required to understand how the levels of isopentenyl diphosphate in T cell culture affects the intracellular sterol availability, which may be mediated by liver X receptors (LXRs)[154] or the transcription factors, the sterol regulatory element-binding proteins (SREBPs)[155] to further elucidate the mechanism of action on how isopentenyl diphosphate regulates CD4+ T cell function. Furthermore, isopentenyl diphosphate is also believed to be the agonist of human γδ T cells. The approach to investigate the level and activity of the primate-specific Vγ9Vδ2 T cells, is required to test the hypothesis whether isopentenyl diphosphate enhances the production of IL-17A and IL-17F via the activation of Vγ9Vδ2 T cells.

The production of microbial metabolites

Our metagenomics analyses yielded a substantial successful prediction rate for CD-associated microbial metabolites with relevant T cell bioactivity. Based on our ascorbate study, the design was to search the metabolite producing genera based on their orthologue genomic contents, and direct culture of candidate isolates to validate production of metabolites. The detecting approach afforded an *a priori* selection for the targeted metabolomics approach which requires a customized optimization of MS parameters for each metabolite of interest. However, it should be emphasized that it remains challenging for the metabolomics detection from the bacteria-cultured samples. The first difficulty is the sample preparation to remove matrix effects from culture media (i.e. salts or cell debris) which could dramatically impact downstream liquid chromatography mass spectrometry (LC-MS) results[156]. Even if some extraction methods have

been developed (for example solid-phase or liquid-liquid extraction[157]), the time-consuming steps of extraction can potentially suppress the signals, especially for the unstable, reactive metabolites. The second challenge is the unknown regulation of metabolite production in microbes. Many genes are expressed only under specific circumstances. The optimal production of microbial metabolites can only achieve when the culture condition fulfills the requirements. A precise and refined genomic sequencing for the bacterial gene cluster in candidate organisms may permit the prediction of operon features for gene regulation and that could be targeted for culture condition evaluation.

To prove the presence of metabolites at the local gut region, several untargeted metabolomics approaches have been developed[158,159]. However, the major challenge to the field is to determine whether a gut metabolite is the product of host, food, or microbiota. It is required to develop the culture-independent high-throughput shotgun sequencing approaches for measuring levels of mRNA (metatranscriptomics[160]) or proteins (metaproteomics[161,162]), which can uncover the production of metabolites generated from microbes and enable the functional assessment of disease-associated features that metagenomics studies overlook.

The alterations of gut microbiota in SSc

To explore the potential roles of gut microbiota in systemic autoimmune disorders, we studied the association of gut microbiome with systemic sclerosis. The alteration of gut microbial composition in SSc has been characterized in our group[128,163] and other studies[142,164] (Table 5-2). We observed the decreased abundances of genera *Faecalibacterium* (UCLA cohort), *Clostridium* (OUH cohort), and *Bacteroides* (both cohorts) in SSc patients, and the augmented abundances of genera *Lactobacillus* (both cohorts), *Fusobacterium* (UCLA cohort), and *Ruminococcus* (UCLA cohort). Interestingly, *Ruminococcus* in both studies was found to be

largely enriched (LDA score > 3) in SSc patients. This genus was recently associated with a stricturing (fibrotic) complications in a multicenter inception cohort of pediatric CD study[165], suggesting that the genus *Ruminococcus* could potentially perpetuate the generation of fibrotic tissues. Additional study is required to address causality.

The finding of the increased abundance of the genus *Lactobacillus* in SSc patients was remarkable. Both the colonic lavage samples from the cecum and sigmoid regions and the faecal samples in UCLA cohort demonstrated the increase of *Lactobacillus* in SSc patients. This observation has been appreciated in three other SSc cohorts in which faecal samples were analyzed: Norwegian cohort[163]; Swedish cohort[142]; and Italian cohort[164]. This is contradicting to current knowledge that the abundance of *Lactobacillus* genus is typically reduced in chronic inflammatory states[144] and several *Lactobacillus spp.* are beneficial bacteria with anti-inflammatory effects in several animal models[166,167]. Studies in probiotics suggested that *Lactobacillus reuteri* reduced the colon motility via affecting enteric neuron[168,169]. The reduction of colonic motility can lead to diminished peristalsis, which is a common feature of SSc. Future studies are required to determine whether members of the *Lactobacillus spp.* are concurrent or causative in SSc.

The aforementioned studies are essential to establish the knowledge foundation for the transition of association to causality in the field of the microbiome, and the additional mechanistic validations (i.e. the *in vivo* model) should be pursued with the combined effort in basic and translational scientific community.

Table

Table 5-1. Information for bioactive microbial metabolites

Metabolite Name	Category (HMDB)	Pathway	HMDB (conc. in bloodstream)
(S)-3-Hydroxy-3-methylglutaryl-CoA	Fatty Acyls	Mevalonate metabolism	N.A.
Hexanoyl-CoA	Fatty Acyls	Fatty acids biosynthesis and oxidation; ceramide formation	N.A.
GDP-mannose	Purine nucleotides	Production of fucosylated oligosaccharides	N.A.
Isopentenyl diphosphate	Prenol lipids	Mevalonate metabolism	N.A.
CoA	Purine nucleotides	Fatty acids biosynthesis and oxidation	N.A.
L-Glutamate	Amino acids, peptides and analogues	Amino acid metabolism	8-97.4 uM
L-Ascorbate	Furanones	Vitamin C metabolism	14.0-125uM
L-Homocysteine	Carboxylic acids and derivatives	Methionine metabolism	2.74-10.4 uM
(R)-Mevalonate	Fatty acids and conjugates	Mevalonate metabolism	N.A.
Glutaryl-CoA	Fatty Acyls	Fatty acids biosynthesis and oxidation	N.A.
Pyridoxine phosphate	Pyridines and derivatives	Vitamin B6 metabolism	0.0-0.01 uM
Oleic acid	Fatty Acid (MUFA)	Fatty acids biosynthesis and oxidation	11-122 uM
Arachidonic acid	Fatty Acid (PUFA)	Fatty acids biosynthesis and oxidation	5-55.4 uM
Linoleic acid	Fatty Acid (PUFA)	Fatty acids biosynthesis and oxidation	14.7-207.9uM
Butyric acid	Fatty Acid (SCFA)	Fatty acids biosynthesis and oxidation	1.0 (0.3- 1.5) uM

Table 5-2. Increased and decreased microbial taxa in SSc patients versus controls.

Study	Region	Design/Sample	N	Increased in SSc*	Decreased in SSc*
Andréasson et al. 2016	Lund, Sweden	Cross-sectional/ Fecal sample	98	*Lactobacillus*	*Faecalibacterium prausnitzii; Clostridiaceae(f)*
Volkmann et al. 2016	Los Angeles, USA	Cross-sectional/ Colonic lavage sample	17†	*Lactobacillus; Bifidobacterium; Fusobacterium; Erwinia; Ruminococcus Prevotella*	*Faecalibacterium; Clostridium; Rikenella*
Volkmann et al. 2017	Oslo, Norway	Cross-sectional/ Fecal sample	17	*Lactobacillus*	*Clostridium; Bacteroides*
	Los Angeles, USA	Cross-sectional/ Fecal sample	17†	*Lactobacillus; Fusobacterium; Erwinia; Ruminococcus*	*Faecalibacterium; Bacteroides*
Bosello et al. 2016 (abstract)	Rome, Italy	Cross-sectional/ Fecal sample	66	*Lactobacillus; Ruminococcus; Roseburia; Faecalibacterium*	*Clostridium; Odoribacter; Veillonella; Prevotella*

* Relative to healthy controls.

† These are the same subjects. In Volkmann et al. 2016, we collected lavage specimens from the cecum and sigmoid colon during colonoscopy; in Volkmann et al. 2017, we collected fecal specimens. The healthy control groups were comprised of different individuals in the two studies.

REFERENCES

1 Distrutti, E., Monaldi, L., Ricci, P. & Fiorucci, S. Gut microbiota role in irritable bowel syndrome: New therapeutic strategies. *World journal of gastroenterology : WJG* **22**, 2219-2241, doi:10.3748/wjg.v22.i7.2219 (2016).

2 Frank, D. N. *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A* **104**, 13780-13785, doi:10.1073/pnas.0706625104 (2007).

3 Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022-1023, doi:10.1038/4441022a (2006).

4 Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027-1031, doi:10.1038/nature05414 (2006).

5 Koeth, R. A. *et al.* Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med* **19**, 576-585, doi:10.1038/nm.3145 (2013).

6 Scher, J. U. *et al.* Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. *Elife* **2**, e01202, doi:10.7554/eLife.01202 (2013).

7 Jiang, H. *et al.* Altered fecal microbiota composition in patients with major depressive disorder. *Brain Behav Immun* **48**, 186-194, doi:10.1016/j.bbi.2015.03.016 (2015).

8 Aizawa, E. *et al.* Possible association of Bifidobacterium and Lactobacillus in the gut microbiota of patients with major depressive disorder. *J Affect Disord* **202**, 254-257, doi:10.1016/j.jad.2016.05.038 (2016).

9 Wang, L. *et al.* Increased abundance of Sutterella spp. and Ruminococcus torques in feces of children with autism spectrum disorder. *Mol Autism* **4**, 42, doi:10.1186/2040-2392-4-42 (2013).

10 Tomova, A. *et al.* Gastrointestinal microbiota in children with autism in Slovakia. *Physiol Behav* **138**, 179-187, doi:10.1016/j.physbeh.2014.10.033 (2015).

11 Langille, M. G. *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* **31**, 814-821, doi:10.1038/nbt.2676 (2013).

12 Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220-230, doi:10.1038/nature11550 (2012).

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

14 Morgan, X. C. *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* **13**, R79, doi:10.1186/gb-2012-13-9-r79 (2012).

15 Tong, M. *et al.* A modular organization of the human intestinal mucosal microbiota and its association with inflammatory bowel disease. *PloS one* **8**, e80702, doi:10.1371/journal.pone.0080702 (2013).

16 Gevers, D. *et al.* The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**, 382-392, doi:10.1016/j.chom.2014.02.005 (2014).

17 Medema, M. H. & Fischbach, M. A. Computational approaches to natural product discovery. *Nat Chem Biol* **11**, 639-648, doi:10.1038/nchembio.1884 (2015).

18 Hadjithomas, M. *et al.* IMG-ABC: A Knowledge Base To Fuel Discovery of Biosynthetic Gene Clusters and Novel Secondary Metabolites. *MBio* **6**, e00932, doi:10.1128/mBio.00932-15 (2015).

19 Charlop-Powers, Z., Milshteyn, A. & Brady, S. F. Metagenomic small molecule discovery methods. *Curr Opin Microbiol* **19**, 70-75, doi:10.1016/j.mib.2014.05.021 (2014).

20 Abubucker, S. *et al.* Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol* **8**, e1002358, doi:10.1371/journal.pcbi.1002358 (2012).

21 Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446-450, doi:10.1038/nature12721 (2013).

22 Monteleone, I., MacDonald, T. T., Pallone, F. & Monteleone, G. The aryl hydrocarbon receptor in inflammatory bowel disease: linking the environment to disease pathogenesis. *Curr Opin Gastroenterol* **28**, 310-313, doi:10.1097/MOG.0b013e328352ad69 (2012).

23 Costa, K. C., Glasser, N. R., Conway, S. J. & Newman, D. K. Pyocyanin degradation by a tautomerizing demethylase inhibits Pseudomonas aeruginosa biofilms. *Science* **355**, 170-173, doi:10.1126/science.aag3180 (2017).

24 Costa, K. C., Bergkessel, M., Saunders, S., Korlach, J. & Newman, D. K. Enzymatic Degradation of Phenazines Can Generate Energy and Protect Sensitive Organisms from Toxicity. *MBio* **6**, e01520-01515, doi:10.1128/mBio.01520-15 (2015).

25 Weber, T. *et al.* antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* **43**, W237-243, doi:10.1093/nar/gkv437 (2015).

26 Guo, C. J. *et al.* Discovery of Reactive Microbiota-Derived Metabolites that Inhibit Host Proteases. *Cell* **168**, 517-526 e518, doi:10.1016/j.cell.2016.12.021 (2017).

27 Jones, M. L., Tomaro-Duchesneau, C. & Prakash, S. The gut microbiome, probiotics, bile acids axis, and human health. *Trends Microbiol* **22**, 306-308, doi:10.1016/j.tim.2014.04.010 (2014).

28 Ridlon, J. M., Kang, D. J., Hylemon, P. B. & Bajaj, J. S. Bile acids and the gut microbiome. *Curr Opin Gastroenterol* **30**, 332-338, doi:10.1097/MOG.0000000000000057 (2014).

29 Clayton, T. A., Baker, D., Lindon, J. C., Everett, J. R. & Nicholson, J. K. Pharmacometabonomic identification of a significant host-microbiome metabolic interaction affecting human drug metabolism. *Proc Natl Acad Sci U S A* **106**, 14728-14733, doi:10.1073/pnas.0904489106 (2009).

30 Jacobs, J. P. *et al.* A Disease-Associated Microbial and Metabolomics State in Relatives of Pediatric Inflammatory Bowel Disease Patients. *Cell Mol Gastroenterol Hepatol* **2**, 750-766, doi:10.1016/j.jcmgh.2016.06.004 (2016).

31 Marchesi, J. R. *et al.* Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *Journal of proteome research* **6**, 546-551, doi:10.1021/pr060470d (2007).

32 Claus, S. P. *et al.* Systemic multicompartmental effects of the gut microbiome on mouse metabolic phenotypes. *Mol Syst Biol* **4**, 219, doi:10.1038/msb.2008.56 (2008).

33 Dorrestein, P. C., Mazmanian, S. K. & Knight, R. Finding the missing links among metabolites, microbes, and the host. *Immunity* **40**, 824-832, doi:10.1016/j.immuni.2014.05.015 (2014).

34 Smith, P. M. *et al.* The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341**, 569-573, doi:10.1126/science.1241165 (2013).

35 Bergen, W. G. & Wu, G. Intestinal nitrogen recycling and utilization in health and disease. *The Journal of nutrition* **139**, 821-825, doi:10.3945/jn.109.104497 (2009).

36 Chang, P. V., Hao, L., Offermanns, S. & Medzhitov, R. The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. *Proc Natl Acad Sci U S A* **111**, 2247-2252, doi:10.1073/pnas.1322269111 (2014).

37 Kaiko, G. E. *et al.* The Colonic Crypt Protects Stem Cells from Microbiota-Derived Metabolites. *Cell* **167**, 1137, doi:10.1016/j.cell.2016.10.034 (2016).

38 Zelante, T. *et al.* Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity* **39**, 372-385, doi:10.1016/j.immuni.2013.08.003 (2013).

39 Rothhammer, V. *et al.* Type I interferons and microbial metabolites of tryptophan modulate astrocyte activity and central nervous system inflammation via the aryl hydrocarbon receptor. *Nat Med* **22**, 586-597, doi:10.1038/nm.4106 (2016).

40 Zhang, L. S. & Davies, S. S. Microbial metabolism of dietary components to bioactive metabolites: opportunities for new therapeutic interventions. *Genome medicine* **8**, 46, doi:10.1186/s13073-016-0296-x (2016).

41 Cox, M. A. *et al.* Short-chain fatty acids act as antiinflammatory mediators by regulating prostaglandin E(2) and cytokines. *World journal of gastroenterology : WJG* **15**, 5549-5557 (2009).

42 Chimerel, C. *et al.* Bacterial metabolite indole modulates incretin secretion from intestinal enteroendocrine L cells. *Cell Rep* **9**, 1202-1208, doi:10.1016/j.celrep.2014.10.032 (2014).

43 Ito, S. *et al.* Indoxyl sulfate induces leukocyte-endothelial interactions through up-regulation of E-selectin. *The Journal of biological chemistry* **285**, 38869-38875, doi:10.1074/jbc.M110.166686 (2010).

44 Wang, Z. *et al.* Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57-63, doi:10.1038/nature09922 (2011).

45 Integrative, H. M. P. R. N. C. The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. *Cell Host Microbe* **16**, 276-289, doi:10.1016/j.chom.2014.08.014 (2014).

46 Huttenhower, C., Kostic, A. D. & Xavier, R. J. Inflammatory bowel disease as a model for translating the microbiome. *Immunity* **40**, 843-854, doi:10.1016/j.immuni.2014.05.013 (2014).

47 Sartor, R. B. & Wu, G. D. Roles for Intestinal Bacteria, Viruses, and Fungi in Pathogenesis of Inflammatory Bowel Diseases and Therapeutic Approaches. *Gastroenterology* **152**, 327-339.e324, doi:http://doi.org/10.1053/j.gastro.2016.10.012 (2017).

48 Brestoff, J. R. & Artis, D. Commensal bacteria at the interface of host metabolism and the immune system. *Nature immunology* **14**, 676-684, doi:10.1038/ni.2640 (2013).

49 Thorburn, A. N., Macia, L. & Mackay, C. R. Diet, metabolites, and "western-lifestyle" inflammatory diseases. *Immunity* **40**, 833-842, doi:10.1016/j.immuni.2014.05.014 (2014).

50 Tong, M. *et al.* Reprograming of gut microbiome energy metabolism by the FUT2 Crohn's disease risk polymorphism. *ISME J* **8**, 2193-2206, doi:10.1038/ismej.2014.64 (2014).

51 Lewis, J. D. *et al.* Inflammation, Antibiotics, and Diet as Environmental Stressors of the Gut Microbiome in Pediatric Crohn's Disease. *Cell Host Microbe* **18**, 489-500, doi:10.1016/j.chom.2015.09.008 (2015).

52 Gao, Z. *et al.* Butyrate improves insulin sensitivity and increases energy expenditure in mice. *Diabetes* **58**, 1509-1517, doi:10.2337/db08-1637 (2009).

53 Mascanfroni, I. D. *et al.* Metabolic control of type 1 regulatory T cell differentiation by AHR and HIF1-alpha. *Nat Med* **21**, 638-646, doi:10.1038/nm.3868 (2015).

54 Hooper, L. V., Littman, D. R. & Macpherson, A. J. Interactions between the microbiota and the immune system. *Science* **336**, 1268-1273, doi:10.1126/science.1223490 (2012).

55 Knights, D. *et al.* Complex host genetics influence the microbiome in inflammatory bowel disease. *Genome medicine* **6**, 107, doi:10.1186/s13073-014-0107-1 (2014).

56 Khor, B., Gardet, A. & Xavier, R. J. Genetics and pathogenesis of inflammatory bowel disease. *Nature* **474**, 307-317, doi:10.1038/nature10209 (2011).

57 Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119-124, doi:10.1038/nature11582 (2012).

58 Yen, D. *et al.* IL-23 is essential for T cell-mediated colitis and promotes inflammation via IL-17 and IL-6. *The Journal of clinical investigation* **116**, 1310-1316, doi:10.1172/JCI21404 (2006).

59 Fujino, S. *et al.* Increased expression of interleukin 17 in inflammatory bowel disease. *Gut* **52**, 65-70 (2003).

60 Kleinschek, M. A. *et al.* Circulating and gut-resident human Th17 cells express CD161 and promote intestinal inflammation. *The Journal of experimental medicine* **206**, 525-534, doi:10.1084/jem.20081712 (2009).

61 Bogaert, S. *et al.* Differential mucosal expression of Th17-related genes between the inflamed colon and ileum of patients with inflammatory bowel disease. *BMC immunology* **11**, 61, doi:10.1186/1471-2172-11-61 (2010).

62 Kobayashi, T. *et al.* IL23 differentially regulates the Th1/Th17 balance in ulcerative colitis and Crohn's disease. *Gut* **57**, 1682-1689, doi:10.1136/gut.2007.135053 (2008).

63 Neurath, M. F. New targets for mucosal healing and therapy in inflammatory bowel diseases. *Mucosal Immunol* **7**, 6-19, doi:10.1038/mi.2013.73
mi201373 [pii] (2014).

64 Hueber, W. *et al.* Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. *Gut* **61**, 1693-1700, doi:10.1136/gutjnl-2011-301668 (2012).

65 Sandborn, W. J. *et al.* Ustekinumab induction and maintenance therapy in refractory Crohn's disease. *N Engl J Med* **367**, 1519-1528, doi:10.1056/NEJMoa1203572 (2012).

66 Sands, B. E., Chen, J., Penney, M., Newbold, P., Faggioni, R., van der Merwe, R., Patra, K., Klekotka, P., Pulkstenis, E., Drappa, J., Gasser, R.A. OP025: A randomized, double-blind placebo-controlled phase 2a induction study of MEDI2070 (anti-p19 antibody) in patients with active Crohn's disease who have failed anti-TNF antibody therapy. *J. Crohn's Colitis* **9**, S15–S16 (2015).

67 Targan, S. R. *et al.* Mo2083 A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Safety, Tolerability, and Efficacy of AMG 827 in Subjects With Moderate to Severe Crohn's Disease. *Gastroenterology* **143**, e26, doi:http://dx.doi.org/10.1053/j.gastro.2012.07.084 (2012).

68 Colombel, J. F., Sendid, B., Jouault, T. & Poulain, D. Secukinumab failure in Crohn's disease: the yeast connection? *Gut* **62**, 800-801, doi:10.1136/gutjnl-2012-304154 (2013).

69 Maxwell, J. R. *et al.* Differential Roles for Interleukin-23 and Interleukin-17 in Intestinal Immunoregulation. *Immunity* **43**, 739-750, doi:10.1016/j.immuni.2015.08.019 (2015).

70 Wilson, N. J. *et al.* Development, cytokine profile and function of human interleukin 17-producing helper T cells. *Nature immunology* **8**, 950-957, doi:10.1038/ni1497 (2007).

71 Macintyre, A. N. *et al.* The glucose transporter Glut1 is selectively essential for CD4 T cell activation and effector function. *Cell metabolism* **20**, 61-72, doi:10.1016/j.cmet.2014.05.004 (2014).

72 Sokol, P. A., Darling, P., Woods, D. E., Mahenthiralingam, E. & Kooi, C. Role of ornibactin biosynthesis in the virulence of Burkholderia cepacia: characterization of pvdA, the gene encoding L-ornithine N(5)-oxygenase. *Infect Immun* **67**, 4443-4455 (1999).

73 Mc, F. J. The nephelometer:aN instrument for estimating the number of bacteria in suspensions used for calculating the opsonic index and for vaccines. *Journal of the American Medical Association* **XLIX**, 1176-1178, doi:10.1001/jama.1907.25320140022001f (1907).

74 Wang, C. *et al.* CD5L/AIM Regulates Lipid Biosynthesis and Restrains Th17 Cell Pathogenicity. *Cell* **163**, 1413-1427, doi:10.1016/j.cell.2015.10.068 S0092-8674(15)01424-5 [pii] (2015).

75 Pacheco, R. *et al.* Glutamate released by dendritic cells as a novel modulator of T cell activation. *Journal of immunology* **177**, 6695-6704 (2006).

76 Buck, M. D., O'Sullivan, D. & Pearce, E. L. T cell metabolism drives immunity. *The Journal of experimental medicine* **212**, 1345-1360, doi:10.1084/jem.20151159 (2015).

77 Slack, M., Wang, T. & Wang, R. T cell metabolic reprogramming and plasticity. *Mol Immunol* **68**, 507-512, doi:10.1016/j.molimm.2015.07.036 (2015).

78 Hancock, R. D. & Viola, R. Biotechnological approaches for L-ascorbic acid production. *Trends Biotechnol* **20**, 299-305 (2002).

79 DuPage, M. & Bluestone, J. A. Harnessing the plasticity of CD4(+) T cells to treat immune-mediated disease. *Nature reviews. Immunology* **16**, 149-163, doi:10.1038/nri.2015.18 (2016).

80 Sallusto, F. Heterogeneity of Human CD4(+) T Cells Against Microbes. *Annual review of immunology* **34**, 317-334, doi:10.1146/annurev-immunol-032414-112056 (2016).

81 Yun, J. *et al.* Vitamin C selectively kills KRAS and BRAF mutant colorectal cancer cells by targeting GAPDH. *Science*, doi:10.1126/science.aaa5004 (2015).

82 Franchi, L. *et al.* Inhibiting Oxidative Phosphorylation In Vivo Restrains Th17 Effector Responses and Ameliorates Murine Colitis. *Journal of immunology* **198**, 2735-2746, doi:10.4049/jimmunol.1600810 (2017).

83 Sasidharan Nair, V., Song, M. H. & Oh, K. I. Vitamin C Facilitates Demethylation of the Foxp3 Enhancer in a Tet-Dependent Manner. *Journal of immunology* **196**, 2119-2131, doi:10.4049/jimmunol.1502352 (2016).

84 Yue, X. *et al.* Control of Foxp3 stability through modulation of TET activity. *The Journal of experimental medicine* **213**, 377-397, doi:10.1084/jem.20151438 (2016).

85 Nishikimi, M., Fukuyama, R., Minoshima, S., Shimizu, N. & Yagi, K. Cloning and chromosomal mapping of the human nonfunctional gene for L-gulono-gamma-lactone oxidase, the enzyme for L-ascorbic acid biosynthesis missing in man. *J Biol Chem* **269**, 13685-13688 (1994).

86 Naidu, K. A. Vitamin C in human health and disease is still a mystery? An overview. *Nutr J* **2**, 7, doi:10.1186/1475-2891-2-7 (2003).

87 Vera, J. C. *et al.* Resolution of the facilitated transport of dehydroascorbic acid from its intracellular accumulation as ascorbic acid. *The Journal of biological chemistry* **270**, 23706-23712 (1995).

88 Amir Shaghaghi, M., Bernstein, C. N., Serrano Leon, A., El-Gabalawy, H. & Eck, P. Polymorphisms in the sodium-dependent ascorbate transporter gene SLC23A1 are associated with susceptibility to Crohn disease. *The American journal of clinical nutrition* **99**, 378-383, doi:10.3945/ajcn.113.068015 (2014).

89 Hengstermann, S. *et al.* Altered status of antioxidant vitamins and fatty acids in patients with inactive inflammatory bowel disease. *Clin Nutr* **27**, 571-578, doi:10.1016/j.clnu.2008.01.007 (2008).

90 Lock, G., Holstege, A., Lang, B. & Schölmerich, J. Gastrointestinal manifestations of progressive systemic sclerosis. *The American journal of gastroenterology* **92**, 763-771 (1997).

91 Sallam, H., McNearney, T. A. & Chen, J. Z. Systematic review: pathophysiology and management of gastrointestinal dysmotility in systemic sclerosis (scleroderma). *Alimentary pharmacology & therapeutics* **23**, 691-712 (2006).

92 Marie, I., Ducrotté, P., Denis, P., Menard, J.-F. & Levesque, H. Small intestinal bacterial overgrowth in systemic sclerosis. *Rheumatology* **48**, 1314-1319 (2009).

93 Franck-Larsson, K., Graf, W. & Rönnblom, A. Lower gastrointestinal symptoms and quality of life in patients with systemic sclerosis: a population-based study. *European journal of gastroenterology & hepatology* **21**, 176-182 (2009).

94 Bodukam, V. *et al.* Association of gastrointestinal involvement and depressive symptoms in patients with systemic sclerosis. *Rheumatology (Oxford)* **50**, 330-334, doi:10.1093/rheumatology/keq296 (2011).

95 Manetti, M. *et al.* Severe fibrosis and increased expression of fibrogenic cytokines in the gastric wall of systemic sclerosis patients. *Arthritis Rheum* **56**, 3442-3447, doi:10.1002/art.22940 (2007).

96 Kawaguchi, Y. *et al.* Muscarinic-3 acetylcholine receptor autoantibody in patients with systemic sclerosis: contribution to severe gastrointestinal tract dysmotility. *Ann Rheum Dis* **68**, 710-714, doi:10.1136/ard.2008.096545 (2009).

97 Roberts, C. G., Hummers, L. K., Ravich, W. J., Wigley, F. M. & Hutchins, G. M. A case-control study of the pathology of oesophageal disease in systemic sclerosis (scleroderma). *Gut* **55**, 1697-1703, doi:10.1136/gut.2005.086074 (2006).

98 Kaye, S. A. *et al.* Small bowel bacterial overgrowth in systemic sclerosis: detection using direct and indirect methods and treatment outcome. *Rheumatology* **34**, 265-269 (1995).

99 Parodi, A. *et al.* Small intestinal bacterial overgrowth in patients suffering from scleroderma: clinical effectiveness of its eradication. *The American journal of gastroenterology* **103**, 1257-1262 (2008).

100 Frank, D. N. *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proceedings of the National Academy of Sciences* **104**, 13780-13785 (2007).

101 Fava, F. & Danese, S. Intestinal microbiota in inflammatory bowel disease: friend of foe? *World journal of gastroenterology: WJG* **17**, 557 (2011).

102 Presley, L. L. *et al.* Host–microbe relationships in inflammatory bowel disease detected by bacterial and metaproteomic analysis of the mucosal–luminal interface. *Inflammatory bowel diseases* **18**, 409-417 (2012).

103 McHardy, I. H. *et al.* Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. *Microbiome* **1**, 17 (2013).

104 Tsuda, A. *et al.* Influence of Proton-Pump Inhibitors on the Luminal Microbiota in the Gastrointestinal Tract. *Clin Transl Gastroenterol* **6**, e89, doi:10.1038/ctg.2015.20 (2015).

105 Stearns, J. C. *et al.* Bacterial biogeography of the human digestive tract. *Scientific reports* **1** (2011).

106 Costello, E. K. *et al.* Bacterial community variation in human body habitats across space and time. *Science* **326**, 1694-1697 (2009).

107 Khanna, D. *et al.* Reliability and validity of the University of California, Los Angeles scleroderma clinical trial consortium gastrointestinal tract instrument. *Arthritis Care & Research* **61**, 1257-1263 (2009).

108 Bae, S. *et al.* Associations between a scleroderma-specific gastrointestinal instrument and objective tests of upper gastrointestinal involvements in systemic sclerosis. *Clin Exp Rheumatol* **31**, 57-63 (2013).

109 Steen, V. D. & Medsger, T. A. The value of the health assessment questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. *Arthritis & Rheumatism* **40**, 1984-1991 (1997).

110 Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300 (1995).

111 Morgan, X. C. *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* **13**, R79 (2012).

112 Hill, T. C., Walsh, K. A., Harris, J. A. & Moffett, B. F. Using ecological diversity measures with bacterial communities. *FEMS Microbiology Ecology* **43**, 1-11 (2003).

113 Kostic, A. D., Xavier, R. J. & Gevers, D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* **146**, 1489-1499 (2014).

114 Roberts, G. Fusobacterial infections: an underestimated threat. *British journal of biomedical science* **57**, 156-162 (1999).

115 Koren, O. *et al.* Human oral, gut, and plaque microbiota in patients with atherosclerosis. *Proceedings of the National Academy of Sciences* **108**, 4592-4598 (2011).

116 Ohkusa, T. *et al.* Fusobacterium varium localized in the colonic mucosa of patients with ulcerative colitis stimulates species-specific antibody. *Journal of gastroenterology and hepatology* **17**, 849-853 (2002).

117 Ohkusa, T. *et al.* Induction of experimental ulcerative colitis by Fusobacterium varium isolated from colonic mucosa of patients with ulcerative colitis. *Gut* **52**, 79-83 (2003).

118 Strauss, J. *et al.* Invasive potential of gut mucosa-derived fusobacterium nucleatum positively correlates with IBD status of the host. *Inflammatory bowel diseases* **17**, 1971-1978 (2011).

119 Benjamin, J. L. *et al.* Smokers with active Crohn's disease have a clinically relevant dysbiosis of the gastrointestinal microbiota. *Inflamm Bowel Dis* **18**, 1092-1100, doi:10.1002/ibd.21864 (2012).

120 Wu, G. D. *et al.* Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105-108, doi:10.1126/science.1208344 (2011).

121 Kamada, N., Chen, G. & Núñez, G. A complex microworld in the gut: harnessing pathogen-commensal relations. *Nature medicine* **18**, 1190-1191 (2012).

122 Sokol, H. *et al.* Low counts of Faecalibacterium prausnitzii in colitis microbiota. *Inflammatory bowel diseases* **15**, 1183-1189 (2009).

123 Sokol, H. *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proceedings of the National Academy of Sciences* **105**, 16731-16736 (2008).

124 Varela, E. *et al.* Colonisation by Faecalibacterium prausnitzii and maintenance of clinical remission in patients with ulcerative colitis. *Aliment Pharmacol Ther* **38**, 151-161, doi:10.1111/apt.12365 (2013).

125 Atarashi, K. *et al.* Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* **500**, 232-236, doi:10.1038/nature12331 (2013).

126 Stoll, M. L. *et al.* Altered microbiota associated with abnormal humoral immune responses to commensal organisms in enthesitis-related arthritis. *Arthritis research & therapy* **16**, 486 (2014).

127 Lepri, G. *et al.* Evidence for oesophageal and anorectal involvement in very early systemic sclerosis (VEDOSS): report from a single VEDOSS/EUSTAR centre. *Ann Rheum Dis* **74**, 124-128, doi:10.1136/annrheumdis-2013-203889 (2015).

128 Volkmann, E. R. *et al.* Association of Systemic Sclerosis With a Unique Colonic Microbial Consortium. *Arthritis Rheumatol* **68**, 1483-1492, doi:10.1002/art.39572 (2016).

129 van den Hoogen, F. *et al.* 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. *Arthritis Rheum* **65**, 2737-2747, doi:10.1002/art.38098 (2013).

130 Tong, M., Jacobs, J. P., McHardy, I. H. & Braun, J. Sampling of Intestinal Microbiota and Targeted Amplification of Bacterial 16S rRNA Genes for Microbial Ecologic Analysis. *Current protocols in immunology / edited by John E. Coligan ... [et al.]* **107**, 7 41 41-47 41 11, doi:10.1002/0471142735.im0741s107 (2014).

131 Goh, N. S. *et al.* Interstitial lung disease in systemic sclerosis: a simple staging system. *Am J Respir Crit Care Med* **177**, 1248-1254, doi:10.1164/rccm.200706-877OC (2008).

132 Mariat, D. *et al.* The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC Microbiol* **9**, 123, doi:10.1186/1471-2180-9-123 (2009).

133 Schaffler, H. *et al.* Mucosa-attached bacterial community in Crohn's Disease coheres with the Clinical Disease Activity Index. *Environ Microbiol Rep*, doi:10.1111/1758-2229.12411 (2016).

134 Sokol, H. *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci U S A* **105**, 16731-16736, doi:10.1073/pnas.0804812105 (2008).

135 Varela, E. *et al.* Colonisation by Faecalibacterium prausnitzii and maintenance of clinical remission in patients with ulcerative colitis. *Alimentary pharmacology & therapeutics* **38**, 151-161 (2013).

136 Prosberg, M., Bendtsen, F., Vind, I., Petersen, A. M. & Gluud, L. L. The association between the gut microbiota and the inflammatory bowel disease activity: a systematic review and meta-analysis. *Scand J Gastroenterol* **51**, 1407-1415, doi:10.1080/00365521.2016.1216587 (2016).

137 Atarashi, K. *et al.* Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* **500**, 232-236 (2013).

138 Slobodin, G. & Rimar, D. Regulatory T Cells in Systemic Sclerosis: a Comprehensive Review. *Clin Rev Allergy Immunol* **52**, 194-201, doi:10.1007/s12016-016-8563-6 (2017).

139 Maeda, Y. *et al.* Dysbiosis Contributes to Arthritis Development via Activation of Autoreactive T Cells in the Intestine. *Arthritis Rheumatol* **68**, 2646-2661, doi:10.1002/art.39783 (2016).

140 De Cruz, P. *et al.* Association between specific mucosa-associated microbiota in Crohn's disease at the time of resection and subsequent disease recurrence: a pilot study. *J Gastroenterol Hepatol* **30**, 268-278, doi:10.1111/jgh.12694 (2015).

141 Strauss, J. *et al.* Invasive potential of gut mucosa-derived Fusobacterium nucleatum positively correlates with IBD status of the host. *Inflamm Bowel Dis* **17**, 1971-1978, doi:10.1002/ibd.21606 (2011).

142 Andreasson, K., Alrawi, Z., Persson, A., Jonsson, G. & Marsal, J. Intestinal dysbiosis is common in systemic sclerosis and associated with gastrointestinal and extraintestinal features of disease. *Arthritis Res Ther* **18**, 278, doi:10.1186/s13075-016-1182-z (2016).

143 McNulty, N. P. *et al.* The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. *Sci Transl Med* **3**, 106ra106, doi:10.1126/scitranslmed.3002701 (2011).

144 Walter, J. Ecological role of lactobacilli in the gastrointestinal tract: implications for fundamental and biomedical research. *Applied and environmental microbiology* **74**, 4985-4996, doi:10.1128/AEM.00753-08 (2008).

145 Tillisch, K. *et al.* Consumption of fermented milk product with probiotic modulates brain activity. *Gastroenterology* **144**, 1394-1401, 1401 e1391-1394, doi:10.1053/j.gastro.2013.02.043 (2013).

146 Samuelson, D. R., Welsh, D. A. & Shellito, J. E. Regulation of lung immunity and host defense by the intestinal microbiota. *Front Microbiol* **6**, 1085, doi:10.3389/fmicb.2015.01085 (2015).

147 Mar, J. S. *et al.* Disease Severity and Immune Activity Relate to Distinct Interkingdom Gut Microbiome States in Ethnically Distinct Ulcerative Colitis Patients. *MBio* **7**, doi:10.1128/mBio.01072-16 (2016).

148 Marie, I. *et al.* Fecal calprotectin in systemic sclerosis and review of the literature. *Autoimmunity reviews* **14**, 547-554, doi:10.1016/j.autrev.2015.01.018 (2015).

149 Li, Y. *et al.* A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans. *Cell* **167**, 1099-1110 e1014, doi:10.1016/j.cell.2016.10.017 (2016).

150 Schirmer, M. *et al.* Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. *Cell* **167**, 1125-1136 e1128, doi:10.1016/j.cell.2016.10.020 (2016).

151 Ter Horst, R. *et al.* Host and Environmental Factors Influencing Individual Human Cytokine Responses. *Cell* **167**, 1111-1124 e1113, doi:10.1016/j.cell.2016.10.018 (2016).

152 Thurnher, M. & Gruenbacher, G. T lymphocyte regulation by mevalonate metabolism. *Science signaling* **8**, re4, doi:10.1126/scisignal.2005970 (2015).

153 Kidani, Y. *et al.* Sterol regulatory element-binding proteins are essential for the metabolic programming of effector T cells and adaptive immunity. *Nature immunology* **14**, 489-499, doi:10.1038/ni.2570 (2013).

154 Kidani, Y. & Bensinger, S. J. Liver X receptor and peroxisome proliferator-activated receptor as integrators of lipid homeostasis and immunity. *Immunological reviews* **249**, 72-83, doi:10.1111/j.1600-065X.2012.01153.x (2012).

155 Brown, M. S. & Goldstein, J. L. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* **89**, 331-340 (1997).

156 Huang, L. *et al.* Determination of meropenem in bacterial media by LC-MS/MS. *J Chromatogr B Analyt Technol Biomed Life Sci* **961**, 71-76, doi:10.1016/j.jchromb.2014.05.002 (2014).

157 Faijes, M., Mars, A. E. & Smid, E. J. Comparison of quenching and extraction methodologies for metabolome analysis of Lactobacillus plantarum. *Microb Cell Fact* **6**, 27, doi:10.1186/1475-2859-6-27 (2007).

158 Matsumoto, M. *et al.* Impact of intestinal microbiota on intestinal luminal metabolome. *Sci Rep* **2**, 233, doi:10.1038/srep00233 (2012).

159 Yen, S. *et al.* Metabolomic analysis of human fecal microbiota: a comparison of feces-derived communities and defined mixed communities. *Journal of proteome research* **14**, 1472-1482, doi:10.1021/pr5011247 (2015).

160 Franzosa, E. A. *et al.* Relating the metatranscriptome and metagenome of the human gut. *Proc Natl Acad Sci U S A* **111**, E2329-2338, doi:10.1073/pnas.1319284111 (2014).

161 Mottawea, W. *et al.* Altered intestinal microbiota-host mitochondria crosstalk in new onset Crohn's disease. *Nat Commun* **7**, 13419, doi:10.1038/ncomms13419 (2016).

162 Li, X. *et al.* A metaproteomic approach to study human-microbial ecosystems at the mucosal luminal interface. *PloS one* **6**, e26542, doi:10.1371/journal.pone.0026542 (2011).

163 Volkmann, E. R. *et al.* Systemic sclerosis is associated with specific alterations in gastrointestinal microbiota in two independent cohorts. *BMJ Open Gastroenterology* **0**, e000134, doi:doi:10.1136/bmjgast-2017-000134 (2017).

164 Bosello S *et al.* The Intestinal Involvement in Systemic Sclerosis Is Characterized By a Peculiar Gut Microbiota [abstract]. *Arthritis Rheumatol.* **68 (suppl 10).** (2016).

165 Kugathasan, S. *et al.* Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. *Lancet* **389**, 1710-1718, doi:10.1016/S0140-6736(17)30317-3 (2017).

166 Aktas, B. *et al.* The Effect of Lactobacillus casei 32G on the Mouse Cecum Microbiota and Innate Immune Response Is Dose and Time Dependent. *PloS one* **10**, e0145784, doi:10.1371/journal.pone.0145784 (2015).

167 Lee, J. *et al.* Characterization of the anti-inflammatory Lactobacillus reuteri BM36301 and its probiotic benefits on aged mice. *BMC Microbiol* **16**, 69, doi:10.1186/s12866-016-0686-7 (2016).

168 Kunze, W. A. *et al.* Lactobacillus reuteri enhances excitability of colonic AH neurons by inhibiting calcium-dependent potassium channel opening. *J Cell Mol Med* **13**, 2261-2270, doi:10.1111/j.1582-4934.2009.00686.x (2009).

169 Wang, B. *et al.* Lactobacillus reuteri ingestion and IK(Ca) channel blockade have similar effects on rat colon motility and myenteric neurones. *Neurogastroenterol Motil* **22**, 98-107, e133, doi:10.1111/j.1365-2982.2009.01384.x (2010).