Title
Effects of the gut-resident bacterial microbiome on host immune function during HIV/SIV disease

Author
Vujkovic-Cvijin, Ivan

Publication Date
2015

Peer reviewed|Thesis/dissertation

Effects of the gut-resident bacterial microbiome on host immune function during HIV/SIV disease

by

Ivan Vujkovic-Cvijin

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

ACKNOWLEDGEMENTS

First, I would like to thank my thesis advisors, J. "Mike" McCune and Susan V. Lynch. They have been truly superlative mentors: they have been patient, supportive, forthcoming with their wisdom, and courageous to be open-minded to my crazy ideas, something for which I will always be in their debt. I am immensely proud to have worked with them during my graduate work and hope to continue to do so in the future. I am also very appreciative to both the McCune and Lynch labs for maintaining an environment of fun and camaraderie that made my time spent in lab a pleasure. I thank my thesis committee: Anita Sil, Jason Cyster, J. "Mike" McCune, and Susan V. Lynch for their support, both moral and scientific. I also immensely thank the folks in the BMS administration, including Lisa, Monique, Nate, Caroline, Demian, and Ned, for all their hard work; it is very much noticed and very much appreciated. I also thank my friends for helping me to relax my tendency to take some things too seriously, for helping me recognize when I ought to take other things *more* seriously, and for laughing with me through most of all that by helping me laugh at myself. I also want to thank Mama, Tata, and Nora. You put me here; anything I accomplish is an extension of your accomplishments. Mnogo vas voli vaš sin i brat.

CONTRIBUTIONS OF CO-AUTHORS TO THE PRESENTED WORK

Chapter 2 of this dissertation is published at Science Translational Medicine{Vujkovic-Cvijin 2013} as "Dysbiosis of the Gut Microbiota Is Associated with HIV Disease Progression and Tryptophan Catabolism." The co-authors on this publication were Richard M. Dunham[1], Shoko Iwai[2], M. Cyrus Maher[3], Rebecca Albright[1], Mara J. Broadhurst[1], Ryan D. Hernandez[3], Michael M. Lederman[4], Yong Huang[5] , Ma Somsouk[1,2], Steven G. Deeks[6], Peter W. Hunt[6], and both Susan V. Lynch[2] and Joseph M. McCune[1] supervised the work.

[1]Division of Experimental Medicine, Department of Medicine, University of California, San Francisco (UCSF), San Francisco, CA 94110, USA.

[2]Division of Gastroenterology, Department of Medicine, UCSF, San Francisco, CA 94143, USA.

[3]Department of Epidemiology and Biostatistics, UCSF, San Francisco, CA 94107, USA.

[4]Division of Infectious Diseases, Department of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA.

[5]Department of Bioengineering and Therapeutic Sciences, UCSF, San Francisco, CA 94143, USA.

[6]Positive Health Program, Department of Medicine, UCSF, San Francisco, CA 94110, USA.

Chapter 3 of this dissertation is being prepared for submission as a manuscript, to which the following co-authors contributed substantially: Louise Swainson,[1] Simon Chu,[1] Clark Santee,[2] Annalise Petriello,[1] Richard M. Dunham,[1] Yong Huang,[3] Cristian Apetrei,[4] Ivona Pandrea,[4] Nichole Klatt,[5] Jason Brenchley,[6] Susan V. Lynch,[2] and Joseph M. McCune[1].

[1]Division of Experimental Medicine, Department of Medicine, University of California, San Francisco (UCSF), USA

[2]Division of Gastroenterology, Department of Medicine, UCSF, San Francisco, CA 94143, USA

[3]Department of Bioengineering and Therapeutic Sciences, UCSF, San Francisco, CA 94143, USA.

[4]Center for Vaccine Research, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA
[5]Department of Pharmaceutics, University of Washington, Seattle, WA 98121, USA.
[6]Program in Tissue Immunity and Repair and Immunopathogenesis Section, Lab of Molecular Microbiology, NIAID, NIH, Bethesda, MD 20892, USA.

Effects of the gut-resident bacterial microbiome on host immune function during HIV/SIV disease

Ivan Vujkovic-Cvijin

ABSTRACT

The mammalian gut bacterial microbiome is increasingly recognized as a vital immune organ. Numerous studies in murine models implicate the gut bacterial microbiome not only as being required for the induction and maintenance of various important cellular mediators of immunity, but also as being capable of initiating and/or exacerbating states of chronic inflammation in the gut and the periphery. HIV infection and subsequent disease progression is driven by chronic inflammation, and is concomitantly characterized by a disruption within many mucosal cellular compartments involved in regulating gut microbial community composition, including Paneth cells, macrophages, dendritic cells, and Th17 cells. Previous work from our laboratory has shown a link between Th17 cell depletion and activity of the immunomodulatory enzyme indoleamine 2,3-dioxygenase (IDO1), which produces molecules known as kynurenines that can directly inhibit differentiation of Th17 cells *in vitro* and *in vivo*. However, a putative interaction between the gut microbiome and the constellation of immune disruptions in HIV, including chronic inflammation and mucosal immune disruption (perhaps involving the IDO1 pathway), had not yet been described. We hypothesized that HIV infection is associated with a shift in gut microbiome composition, and that this shift could potentiate further mucosal immune disruptions and the chronic inflammation that drives HIV disease progression.

We profiled the gut mucosal-adherent microbiome of 32 human subjects at varying stages of HIV disease. We found that the microbiomes of untreated, infected individuals was

significantly different from those that were uninfected, and the extent to which this microbiome differed was proportionally related to markers of chronic inflammation and mucosal immune disruption (IDO1 activity). Furthermore, we found that some gut bacterial taxa that were enriched in HIV infection correlated strongly and positively with IDO1 activity, and representative members of the top such correlate could directly produce kynurenines *in vitro*. We further sought to identify gut-resident bacteria that may inhibit IDO1 activity, and longitudinally profiled 12 rhesus macaques before and after SIV infection toward that end. We found that fecal abundance of *Lactobacillus* was negatively correlated with IDO1 activity, and macaques given a *Lactobacillus*-containing probiotic exhibited diminished IDO1 activity as measured in peripheral blood.

TABLE OF CONTENTS

LIST OF TABLES

Chapter II:

Chapter III:

LIST OF FIGURES

Chapter II:

CHAPTER I: INTRODUCTION

HIV-associated inflammation and the gut microbiome

Though the advent of highly active antiretroviral therapy (HAART) has drastically improved long-term survival rates of HIV-infected subjects, life expectancy in HAART-suppressed HIV-infected individuals has not reached that of uninfected individuals{Lohse 2007} and is indeed shorter by 10-15 years. Though onset of AIDS is largely prevented in HAART-suppressed individuals, non-AIDS related morbidity and mortality are prevalent and the likelihood of their occurrence is tightly linked to levels of chronic inflammation that persist despite HAART. The etiology of this persistent immune activation has been proposed to stem from multifocal impairments observed in the gut immune barrier during progressive HIV and SIV infection. Depletion of gut mucosal Th17 cells is a salient feature of progressive SIV/HIV disease{Brenchley 2008; Favre 2009} and the loss of this cell type is thought to precipitate the loss of gut barrier integrity due to the critical role this T cell subset plays in stimulating epithelial cell proliferation, tight junction formation, and secretion of antimicrobial peptides via the action of the secreted cytokines IL-17 and IL-22{Ouyang 2008}. Concomitant with depletion of Th17 cells is induction of the antiviral type I interferon axis of inflammation, which induces catabolism of tryptophan into kynurenine compounds known to inhibit Th17 cell differentiation{Desvignes 2009}. Indeed, depletion of Th17 cells correlates with kynurenine pathway activity (as measured by the plasma ratio of kynurenine to tryptophan [Kyn:Trp]) in HIV-infected subjects, and this depletion further is associated with increased microbial translocation{Favre 2010; Raffatellu 2008}. These disruptions to gut barrier integrity among others are thought to be responsible for increases in circulating microbial products such as lipopolysaccharide (LPS) and the co-receptor for LPS, sCD14{Sandler 2011; Brenchley 2006;

Klatt 2012; Brenchley 2008}. Elevations in LPS and sCD14 have been linked to inflammation and poor CD4 T cell recovery after initiation of HAART{Brenchley 2006}, and indeed, up to a third of HIV-infected subjects on long-term HAART can be classified as "immunologic non-responders" (INR), exhibiting poor CD4 T cell recovery and heightened immune activation despite HAART-mediated viral suppression{Gazzola 2009}. Managing this persistent inflammatory state remains an unmet challenge, and there remains a need for novel and effective strategies for abrogating mucosal immune disruption and chronic inflammation particularly in HIV-infected INRs.

Emerging studies in the murine model have highlighted the importance of the gut bacterial microbiome in modulating gut mucosal immunity as well as the reciprocal effect of mucosal immunity on regulating the composition of the gut microbiome. Gut-resident commensal bacterial species in the clade *Clostridia* have been found to be necessary for robust induction of murine mucosal T cell subsets{Atarashi 2011; Geuking 2011} including Th17 cells in healthy mice as well as in models of Th17-driven disease{Gaboriau-Routhiau 2009; Wu 2010; Ivanov 2008}. Furthermore, impairments in intestinal immunity and microbial sensing give rise to characteristic communities of gut-resident microbes that are able to precipitate microbial translocation and exacerbate inflammation in models of chronic inflammatory diseases{Henao-Mejia 2012; Elinav 2011}. Indeed, intestinal inflammation itself – similar to that seen in pathogenic HIV/SIV infection – has been found to alter the composition of the gut microbial community of mice, and this altered community can reciprocally be sufficient towards sustaining chronic inflammation in the gut and periphery{Garrett 2007; Garrett 2010a}. Indeed, our group and others found distinct differences in the gut microbiomes of HIV-uninfected individuals and HIV-infected untreated individuals and that HIV-infected subjects harbor a greater abundance of

pathogenic *Proteobacteria* while exhibiting depletion of *Bacteroidetes* class bacteria{Ellis 2011; Vujkovic-Cvijin 2013}. Furthermore, a substantial proportion of HIV-infected subjects undergoing HAART also exhibit an apparently persistent dysbiosis{Vujkovic-Cvijin 2013; Lozupone 2013}, and the extent of this dysbiosis correlates with inflammatory markers of HIV disease progression{Vujkovic-Cvijin 2013} including plasma IL-6, a strong predictor of mortality among treated HIV-infected subjects{Kalayjian 2010}, and activity of the kynurenine pathway of tryptophan catabolism. We have further found that representative species of numerous bacterial taxa enriched in HIV-infected subjects encode enzymes homologous to those that produce kynurenines from tryptophan, suggesting that gut microbiome constituents of HIV-infected subjects may directly contribute to mucosal immune disruption via bacterial tryptophan catabolism.

The SIV-infected macaque as a model for host/microbiome interactions in HIV

SIV-infected macaques provide a well-described non-human primate model that recapitulates major features of HIV-associated chronic inflammation and mucosal immune disruption, including innate and adaptive immune activation, CD4 T cell depletion in the peripheral blood and gut-associated lymphoid tissue, and a loss of mucosal barrier-promoting Th17 cells{Favre 2009}. The etiology of these immune disruptions in the SIV-infected macaque model remain poorly understood, while distinct assemblages of the macaque gut microbiome have been shown to associate with abundance of Th17 cells{Ardeshir 2014}. Several recent macaque studies have shown that members of the bacterial genus *Lactobacillus* can modulate the Th17 axis, though mechanisms for this process remain poorly understood. Administration of VSL #3, a probiotic cocktail of which 4 of 7 constituent species are *Lactobacillus* spp., to chronically SIV-infected pigtail macaques resulted in greater Th17 polyfunctionality, as well as a

partial resolution of several inflammatory measures (e.g. lower fibrosis, reduced CD4 T cell activation, and increased CD4 T cell recovery){Klatt 2013; Ortiz 2014}. Ligated ileal loops of acutely SIV-infected rhesus macaques inoculated with *Lactobacillus plantarum* displayed increased IL-17 expression and epithelial tight junction formation{Hirao 2014}. Though studies of the macaque gut microbiome have shown few ecological differences in gut bacterial communities between SIV-uninfected and chronically SIV-infected macaques{Klatt 2013; Handley 2012; McKenna 2008}, which is in contrast to the pronounced microbiome perturbations observed in human cohorts{Lozupone 2013; Dillon 2014; Mutlu 2014; Dinh 2015}, host-microbiome dynamics during the acute phase of lentiviral infection when IDO1 is maximally induced by interferons, have thus far remained unexplored. We found that a transient perturbation of the fecal microbiome of SIV-infected macaques occurs during acute SIV infection, which coincides temporally with maximal IDO1 activity. This perturbation was characterized chiefly by a depletion of *Lactobacillus* genus members, and we provide evidence that gut-resident *Lactobacillus* spp. modulate primate IDO1 by inhibiting production of immunomodulatory kynurenine compounds that affect the Th17 cell arm of the mammalian mucosal immune system.

CHAPTER II: DYSBIOSIS OF THE COLONIC MUCOSAL-ADHERENT MICROBIOTA IS ASSOCIATED WITH HIV DISEASE PROGRESSION AND CHRONIC INFLAMMATION

Abstract

Progressive HIV infection is characterized by dysregulation of the intestinal immune barrier, translocation of immunostimulatory microbial products, and chronic systemic inflammation that is thought to drive progression of disease to AIDS. Elements of this pathologic process persist despite viral suppression during highly active antiretroviral therapy (HAART) and drivers of these phenomena remain poorly understood. Disrupted intestinal immunity can precipitate dysbiosis that induces chronic inflammation in the mucosa and periphery of mice. However, putative microbial drivers of HIV-associated immunopathology versus recovery have not been identified in humans. Using high-resolution bacterial community profiling, we identified a dysbiotic mucosal-adherent community enriched in *Proteobacteria* and depleted of *Bacteroidia* members that was associated with markers of mucosal immune disruption, T cell activation, and chronic inflammation in HIV-infected subjects. Furthermore, this dysbiosis was evident among HIV-infected subjects undergoing HAART, and the extent of dysbiosis correlated with activity of the kynurenine pathway of tryptophan metabolism and plasma concentrations of the inflammatory cytokine interleukin-6 (IL-6), two established markers of disease progression. Gut-resident bacteria with capacity to metabolize tryptophan through the kynurenine pathway were found to be enriched in HIV-infected subjects, strongly correlated with kynurenine levels in HIV-infected subjects, and capable of kynurenine production *in vitro*. These observations demonstrate a link between mucosal-adherent colonic bacteria and immunopathogenesis during

progressive HIV infection, which is apparent even in the setting of viral suppression during HAART. This link suggests that gut-resident microbial populations may influence intestinal homeostasis during HIV disease.

Introduction

Accumulating evidence from human and non-human primate studies supports the broad hypothesis that progression to AIDS during HIV infection is driven by chronically elevated T cell activation and systemic inflammation{Giorgi 1999; Sousa 2002; Deeks 2004; Estes 2008}. While the etiology of such persistent immune activation is incompletely understood, the gastrointestinal mucosal immune disruption that follows progressive HIV and SIV infection is postulated to play a role. Specifically, compromised mucosal barrier function and increased translocation of immunostimulatory microbial products from the gut lumen into systemic circulation have been implicated in this process{Estes 2010}. Indeed, the presence of microbial products in the peripheral blood of HIV-infected subjects has been linked to immune activation and increased morbidity and mortality{Brenchley 2006; Sandler 2011}. Among the reported defects in the mucosal barrier, a pronounced reduction in cell populations that are characterized by secretion of the enterocyte homeostasis-promoting factors, IL-17 and IL-22, has been associated with decreased gastrointestinal epithelial barrier integrity and accelerated disease{Raffatellu 2008; Favre 2010}. Accompanying and apparently precipitating these changes are reported increases in activation of the kynurenine pathway of tryptophan metabolism through the interferon-inducible enzyme, indoleamine 2,3-dioxygenase 1 (IDO1){Fuchs 1990; Huengsberg 1998; Favre 2010; Hunt 2012}, which produces tryptophan catabolites that can inhibit the differentiation of IL-17 secreting CD4+ T cells{Desvignes 2009; Romani 2008}. IDO1 activity is induced during the course of pathologic HIV and SIV infection, fails to

normalize during progressive disease, and is associated with impaired mucosal immunity and microbial translocation, the postulated drivers of chronic inflammation that precipitate AIDS{Sandler 2012}. Although highly active anti-retroviral treatment (HAART) results in a partial diminution in inflammatory markers, including IDO1 activity, residual markers of T cell activation and inflammation (e.g., interleukin-6 [IL-6] and interferon-inducible protein 10 [IP-10]) persist and the extent of their elevation correlates with incomplete immune recovery and an increased risk for non-AIDS related morbidities{Hunt 2003; Kalayjian 2010; Sauce 2011}. Nonetheless, the molecular and cellular mechanisms contributing to this inflammatory process during early infection and antiretroviral-treated infection remain poorly understood.

Gut-resident bacteria can modulate the mucosal immune system in ways that overlap with the salient features of HIV pathogenesis. For instance, the absence of a Clostridia class commensal bacterium, segmented filamentous bacteria (SFB), in the murine gut results in a decrease in IL-17-secreting CD4 T cell subpopulations{Ivanov 2009}, while other members of the class *Clostridia* (*Clostridium* clusters IV and XIVa) have been found to induce mucosal expression of IDO1{Atarashi 2011a}. Indeed, *Enterobacteriaceae* family members belonging to the genera *Klebsiella*{Garrett 2010}, *Citrobacter*{Lupp 2007}, and *Salmonella*{Stecher 2007; Winter 2010} can actively induce an inflammatory microenvironment in the gastrointestinal tract that supports their proliferation and persistence. Additionally, disruptions in the mucosal innate immune system can result in the outgrowth of a "dysbiotic" pro-inflammatory community that can be sufficient to sustain pathologic, chronic inflammation in the mucosa and the periphery {Garrett 2007; Garrett 2010; Elinav 2011; Henao-Mejia 2012}. These findings suggest not only that gut bacterial species may modulate mucosal immunity in humans, but also that inflammatory imbalance can promote outgrowth of pathobionts that further exacerbate mucosal immune

disequilibrium. Since HIV disease is characterized by pronounced disruptions to mucosal immunity as well as a sustained chronic inflammatory state, we sought to understand whether the gut microbiota is altered during HIV infection, and to assess relationships between these assemblages and characteristic immunologic states of HIV disease progression.

To capture a wide range of stages of HIV disease progression and treatment, a cohort of 32 male subjects was assembled to include six viremic, untreated (VU) HIV-infected subject, 18 virally-suppressed subjects on HAART with varying degrees of CD4+ T cell recovery (HAART), one HIV-infected long-term non-progressor, and nine uninfected risk-matched controls (HIV-); see Table 1 for defining characteristics of the cohort. Rectosigmoid biopsy specimens and peripheral blood were collected for immunophenotyping of immune cell subpopulations and for profiling of microbial communities. Biopsy tissue was chosen over stool sampling based on evidence that immunomodulatory bacterial species adhere closely to the gastrointestinal mucosal epithelium in the mouse{Ivanov 2009; Atarashi 2011a}, comprising a community that is unique from that within the gut lumen{Nava 2011}. Microbial profiling was performed using a standardized, high-density microarray containing oligonucleotide probes targeted against hypervariable regions of the bacterial 16S rRNA gene (G3 PhyloChip{Hazen 2010}). This approach is ideally suited for high-resolution microbiome comparative analyses and detection of lower-abundance species that may not be identified using sequencing approaches of relatively shallow read depth{DeSantis 2007}. This is particularly relevant given recent evidence demonstrating the presence of low-abundance "keystone" species that despite their low numbers are pivotal in modulating commensal community composition and host-microbe interactions in the oral cavity{Hajishengallis 2011}.

Materials and Methods

Subject recruitment and sample collection

HIV-infected subjects and controls were recruited from SCOPE{Deeks 2004}, a longitudinal cohort of mostly HIV-infected subjects enriched for unusual phenotypes such as long-term non-progressors. Uninfected, viremic untreated (VU), and subjects undergoing highly active antiretroviral therapy (HAART) were identified and enrolled. The HAART group included HIV-infected patients maintaining undetectable viral load on stable antiretroviral regimens for at least one year. Participants with suboptimal immune recovery (i.e., peripheral blood CD4 count < 350 cells/ mm^3) were enriched to capture a wide spectrum of HIV disease. Exclusion criteria included exposure to antibiotics within the 2 months preceding the mucosal sampling.

All participants gave written informed consent for research sigmoidoscopy and biopsies, a protocol approved by the Committee on Human Research, UCSF. Study participants underwent a blood draw and received a Fleet enema prior to the procedure, and rectosigmoid biopsies (each ~3 mm in diameter) were obtained between 10 and 20 cm from the anus using jumbo forceps. Four biopsies were formalin-fixed and paraffin-embedded for immunohistochemistry; 12 biopsies were placed immediately in 15 ml of RPMI 1640 with 15% fetal calf serum, and processed for flow cytometric analysis within 4 hours of collection; two biopsies were placed in RNAlater (Life Technologies, Carlsbad, CA) and frozen at -80°C for quantitation of inflammatory gene expression; finally, another two allocations of two biopsies each were snap-frozen in liquid N_2 and stored at -80°C for HIV viral nucleic acid quantitation and bacterial community profiling by PhyloChip microarray (Second Genome, Inc. San Bruno, CA).

Nucleic acid extraction and quantitative polymerase chain reaction (qPCR)

For quantitation of inflammatory and interferon-stimulated gene expression, RNA was extracted from snap-frozen biopsies frozen in RNAlater (Life Technologies, Carlsbad, CA) after homogenization in Trizol (Life Technologies, Carlsbad, CA) and purification using the RNeasy mini-kit (Qiagen, Valencia, CA). RNA was also prepared from whole blood, via collection using the PaxGene Blood RNA Kit (Qiagen, Valencia, CA). cDNA was prepared from RNA using Omniscript RT Kit (Qiagen, Valencia, CA) using random hexamer primers. qPCR was carried out using TaqMan Gene Expression MasterMix (Applied Biosystems) in a StepOne Plus Real Time PCR system (Applied Biosystems). IDO1 transcript was measured using custom primers and probe (F: 5'- GCCAAATCCACAGGAAAATCT, R: 5'- GCTGTGACTTGTGGTCTGTGA, P: 5'- AAACATCTGCCTGATCTCATAGAGT). Interferon stimulated genes (ISG) were measured using commercially available primer-probe sets from Applied Biosystems: GBP1 Hs00977005_m1, IFI27 HSs00271467_m1, MX1 Hs00895601_m1, OAS1 Hs00973637_m1. Relative expression was calculated using the ΔΔCT method, using HPRT as housekeeping gene. ISG expression was consolidated using the geometric mean of relative expression values for all four aforementioned genes. Expression of inflammatory genes (NADPH oxidase [CYBB subunit], arginase 1, nitric oxide synthase 2, type I interferons [IFN-α, IFN-β], type II interferon [IFN-γ], interferon-stimulated gene expression, tumor necrosis factor-alpha) was carried out using the Fluidigm (South San Francisco, CA, USA) platform, with gene-specific pre-amplification using proprietary computationally designed primers (DELTAgene system from Fluidigm, South San Francisco, CA, USA). Relative expression for these genes was also calculated using the ΔΔCT method, using HPRT as housekeeping gene.

Total bacterial burden quantification was carried out using DNA extracted from separate snap-frozen rectosigmoid biopsies that were collected at the same time as those above. Biopsies were resuspended in RNAlater ICE (Life Technologies, Carlsbad, CA), drained, and transferred to Lysis Matrix B (MP Biomedicals, Santa Ana, CA) tubes containing 600 ul RLT buffer and subjected to 30 seconds of bead beating at 5.5 m/s. Samples were briefly spun down,and DNA was extracted using the AllPrep Kit (Qiagen, Valencia, CA). TaqMan qPCR was performed, as above, using the following primers: P891F, 5'-TGGAGCATGTGGTTTAATTCGA; P1033R, 5'-TGCGGGACTTAACCCAACA, and UniProbe 5'-CACGAGCTGACGACARCCATGCA. A standard curve of known 16S rRNA copy numbers was included with each run and was used to calculate total bacterial burden normalized to total 16S rRNA genes. Enterobacteriaceae were quantitated using Brilliant II SYBR Green (Agilent Technologies, Santa Clara, CA) and the following primers: F: 5'- ATGGCTGTCGTCAGCTCGT; R: 5'-CCTACTTCTTTTGCAACCCACTC{Castillo 2006}.

T cell phenotyping analysis by flow cytometry

Mononuclear cells were prepared from peripheral blood by density gradient centrifugation and isolation using Ficoll-Paque (STEMCELL Technologies, BC, Canada). Rectosigmoid biopsy mononuclear cells (RMC) were isolated from biopsy specimens using a protocol optimized for lymphocyte viability and yield{Shacklett 2003}. Biopsy pieces were digested in three rounds of 0.5 mg/mL collagenase type II (Sigma-Aldrich, St. Louis, MO), after which tissue was disrupted with a syringe bearing a 16-gauge blunt-end needle and passaged through a 70mm cell strainer.

Cells were then stained directly or after phorbol 12-myristate 13-acetate (PMA)/ionomycin stimulation in the presence of brefeldin and monensin for surface marker expression, and then permeabilized and stained intracellularly for FoxP3 or cytokine production using the FoxP3 Fixation/Permeabilization Staining Set (eBiosciences, San Jose, CA) or Cytofix/Cytoperm (BD Biosciences, San Jose, CA), respectively. Acquisition was performed on a Becton Dickinson LSRII (BD Biosciences, San Jose, CA) and analyzed using FlowJo software, version 9 (Treestar, Portland, OR). Naive T cells were identified as CD27+ cells lacking CD45RO expression, and were excluded to yield total memory T cells. Proportions of activated CD4+ and CD8+ T cell were measured by the fraction of these cell populations staining positive for HLA-DR and CD38. Proportions of regulatory T cells (T_{reg}) from total CD4+ cells were determined by co-staining of CD25 and FoxP3.

Immunohistochemistry and viral load quantification in rectosigmoid tissues

Four rectosigmoid biopsies were collected into 4% paraformaldehyde, embedded in paraffin, cut into 5 μm sections and mounted on glass slides. Slides were deparaffinized in Clear-Rite 3 and rehydrated in an ethanol series, and antigenic epitopes were then retrieved by placing slides in a pressure cooker for 20 minutes at 121ºC in pH6 citrate buffer. Slides were blocked and stained with polyclonal rabbit anti-CD3 (DAKO, Denmark), and stain developed with DAB chromogen (anti-rabbit Polink-2 Plus HRP Detection Kit, GBI Labs, Mukilteo, WA). Whole slide composite images were acquired on a ScanScope (Aperio, Vista, CA) and analyzed with the accompanying software to quantify CD3 stained cells per area of tissue. To determine the number of CD4+ cells per area, the CD3+ cells/mm^2 value was multiplied by the fraction of CD3+ cells that stained with CD4 as determined by flow cytometry.

Two rectosigmoid biopsies were snap-frozen in liquid N_2 for quantification of HIV DNA and RNA as previously described{Yukl 2010}.

Plasma inflammatory marker measurements

Plasma was collected from peripheral blood by centrifugation and stored at -80ºC. The following soluble plasma proteins were measured by ELISA using commercially available kits and according to the manufacturer's instructions: IP-10, sTNF-RII, IL-6, FABP2, sCD14 (all from R&D Systems), zonulin (Immundiagnostik), D-dimer (Diagnostica Stago). Concentrations of kynurenine, tryptophan, and 3-hydroxyanthranilic acid were quantitated in plasma and in bacterial growth cultures by liquid chromatography-tandem mass spectrometry, as previously described{Favre 2010}.

16S rRNA amplicon preparation, Phylochip hybridization, scanning, and data normalization

Nucleic acid extraction was performed on snap-frozen rectosigmoid biopsies by re-suspending in RNAlater ICE (Life Technologies, Carlsbad, CA), according to the manufacturer's protocol and as previously described{Abreu 2012}. G3 PhyloChip microarrays were obtained from Second Genome, Inc. (San Bruno, CA). PCR amplification of the bacterial 16S rRNA gene, Phylochip hybridization, Phylochip scanning, and fluorescence intensity normalization and initial probeset filtering of Phylochip data were performed as previously described{Hazen 2010}, with the exception of performing 25 cycles of 16S PCR amplification (rather than 30) and the usage of customized quartile '*r* score' cutoff values. For each unique 16S rRNA gene

probeset, an *r* score was independently calculated that incorporated mismatch probe fluorescence intensity as well as distribution of total chip fluorescence to account for per-chip background fluorescence. Quartile *r* score cutoffs were chosen based on *r* scores retrieved for control probes spiked-in at known concentrations, as recorded from all Phylochip microarrays of the current study, and were as follows: $rQ_1 > .30$, $rQ_2 > .65$, $rQ_3 > .88$.

Phylochip analysis

Following probe intensity normalization to reduce the dataset for analyses, probesets belonging to phylogenetically related (based on species-level identity) 16S rRNA sequences and bearing similar probe set response patterns across all samples were represented by a single taxon, while others were eliminated from the dataset. This process was performed by calculating the variance in fluorescence intensity for each probeset and its two most phylogenetically related neighbors, and was repeated for each taxon across multiple different samples; the sum of all variances was used to define groups of taxa whose probe sets responded similarly. The resulting measures were ranked to select for removal the top ~15% of probesets exhibiting similar response trends across all samples, allowing reduction of redundant probesets for further analyses. Next, the 'vegan' package within the R programming environment was used to perform non-metric multidimensional scaling (NMDS) based on community dissimilarity using the Canberra dissimilarity metric. The 'Adonis' package was used to estimate strength of relationship between community dissimilarity and classification of subjects into either VU or HIV-uninfected. To identify bacterial taxa in differential abundance among viremic untreated and HIV- subjects, the software package Statistical Analysis of Microarrays (SAM{Tusher 2001}), which incorporates the Storey{Storey 2002} method of false discovery rate estimation in

the form of adjusted *q*-values, was used. Taxa were filtered down to a final list of 547 taxa by the following criteria: *q*-value < 0.05, raw MFI fold change > 2, and Mann-Whitney *U* test P value < 0.03. Genera to which belonged taxa within the final list were visualized using the Interactive Tree of Life{Letunic 2007; Letunic 2011} (iTOL, http://itol.embl.de/).

Comparisons of disease-associated microbial genera to immunologic parameters

The R software package 'multtest' was used to calculate Spearman rho correlation coefficient (R_S) values for comparisons of all 562 bacterial taxa that comprise the disease-associated microbial community (DMC) to each of the 24 immunologic parameters measured. Measures of correlations between immunologic variables and taxa were aggregated and condensed to genus-level classifications for visualization purposes in Fig. 3A, using the following methodology: for bacterial genera that were represented by more than one taxon in the DMC, R_S values representing correlations between taxa belonging to the same genus and an immunologic parameter were averaged to yield a R_{Save} that represents the correlation between relative abundance of that genus and each immune parameter. This was performed on the remaining immunologic parameters and all genera represented by more than one taxon, and was visualized using Cluster3 and Java TreeView (Version 1.1.6r2). To represent unadjusted significance values for each correlation shown in Fig. 3A, boxes were made to surround each correlation for which unadjusted P < 0.02. For genera that were represented by more than one taxon in the DMC, the median unadjusted P value for comparisons among immunologic parameters and taxa within that genus was used. Choice of mean for computing aggregated R_S and median for aggregated unadjusted P values for each genus was informed by assessment of global R_S and P value distributions (normal and non-normal, respectively, Fig. S9).

Principal component analysis

To quantify trends in co-variance among taxa that have similar patterns in relative abundance across all subject samples, a principal component analysis (PCA) was performed within the R programming environment (package 'prcomp'). This allowed the condensation of trends in co-varying microbial abundances into discrete variables, thereby yielding single measures that represented broad trends in abundance of multiple taxa. To reduce the number of relationships tested, Spearman rank correlations with immunologic parameters were calculated only for the principal component that represented the greatest portion of co-variance (PC1), and the resultant P values were adjusted for multiple comparisons by the Benjamini-Hochberg technique{Benjamini 1995}.

Multiple linear regression model selection analysis

Among treated subjects, associations were confirmed using linear regression, allowing for the inclusion of a variety of adjustment variables comprising age, race, antibiotic use, and duration of HAART, as well as the first five principal components representing the greatest proportion of microbial abundance co-variance (PC1, PC2, PC3, PC4, PC5). All possible models were fit and compared according to their predictive accuracy. Predictive accuracy was defined as the mean squared error (MSE) calculated from leave-one-out cross-validation (LOO-CV). Under this approach, the model is fit repeatedly, leaving out one observation each time in order to assess a given model's ability to predict new observations. For each outcome, the model with the

lowest LOO-CV MSE was chosen. Given the small number of observations available, p-values were then obtained by permutation of residuals{Freedman 1983}.

Analysis of kynurenine pathway enzyme annotations and associations with plasma Kyn:Trp ratios

The UniProt Consortium database{UniProt Consortium 2012} (www.uniprot.org) was queried for all annotated homologs of enzymes within the pathway to catabolize tryptophan to 3-hydroxyanthranilic acid (Enzyme Commission numbers: EC 1.13.11.11, EC 3.5.1.9, EC 3.7.1.3, EC 1.14.13.9). This list was pared to include only members of the domain Bacteria. Phylum classifications for each annotation were combined, counted, and represented as pie charts with colors representing phyla as denoted in legend of Fig. 5A. 'Totals' represent numbers of unique genera within the domain Bacteria that bear annotated homologs in the highly curated and comprehensive UniProt database for each respective enzyme at the time of publication.

Spearman correlation P values were calculated for each taxon in relation to subject plasma Kyn:Trp ratios. For Fig. 5C, each genus was ranked based on P values of the taxon with the strongest correlation (as defined by the taxon with the lowest P value within a given genus). Annotations for tryptophan catabolism enzymes for all sequenced species and strains within each genus were considered when counting numbers of tryptophan catabolism enzymes for each genus.

***Pseudomonas fluorescens* kynurenine production in bacterial cultures**

Pseudomonas fluorescens strain A506 (ATCC 31948) was grown in King's B Medium{KING 1954} within U-bottomed 96 well polystyrene culture plates at an initial

inoculum of 2.85 x 10^7 cfu per well. King's B Medium was prepared using Bacto Proteose Peptone No. 3 which, according to manufacturer-provided information, would contain 0.3 mM in the King's B Medium preparation. Thus, bacterial cultures contained the free tryptophan already present plus 0, 0.5, 1.0, and 5.0 mM tryptophan added exogenously. Cultures were grown at room temperature and, at 20 hours post inoculation, were spun at 12,000 rpm for 10 minutes in a tabletop microcentrifuge. Supernatants were filtered through a 0.22 μm filter and kanamycin was added to a final concentration of 50 μg/ml. Liquid chromatography-tandem mass spectrometry was used to assay concentrations of kynurenine and tryptophan in growth culture supernatants, as described above.

Results

Gut bacterial microbiota composition differs between HIV- and VU subjects

Analysis of rectosigmoid biopsies showed that total bacterial load was similar across all subjects, regardless of infection status (Fig. S1A), and that total bacterial load measures were concordant across different biopsies from a given individual as well as across different DNA extraction methods (Fig. S1B). Additionally, comparative analysis of gross microbial community metrics showed no significant differences between HIV-infected and uninfected subjects in community evenness or richness (Fig. S1C). By contrast, analysis of between-subject microbiota composition using a Canberra community dissimilarity matrix revealed distinct clustering of the subject groups representing extremes of health and disease: the HIV- and the VU group, respectively (Fig. 1). Permutational multivariate analysis of variance confirmed that the observed variability in community composition was significantly related to subjects belonging to the VU or HIV- group (Adonis; P = 0.002), indicating that untreated HIV infection is associated with a

distinct mucosal microbiota composition. Interestingly, HAART-treated patients, though uniformly exhibiting viral suppression (i.e., plasma viral load < 40 copies/ml), exhibited highly dissimilar community compositions: the microbiota of some patients clustered closely with VU subjects while others were more similar to HIV- subjects.

To identify differences in the adherent bacterial communities of HIV- and VU subjects, a permutational analysis algorithm with integrated false-discovery correction (SAM{Tusher 2001}) was used to detect specific bacterial taxa in differential fold abundance across these two subject groups. Of the 33,951 taxa (defined as a group of bacteria having at least 99% 16S rRNA sequence identity) that were detected in at least one of the samples studied, 625 exhibited significantly different relative abundance between VU and HIV- after stringent data filtering. This list comprised 579 taxa that were enriched and 45 taxa that were depleted in VU subjects compared with HIV- subject samples (Fig. 2). Among the taxa enriched in VU subjects, the most enriched was *Erysipelotrichaceae* in the class Mollicutes, a class that has been associated with obesity and heightened cardiovascular morbidity{Turnbaugh 2008; Zhang 2009; Koeth 2013}. A majority of taxa enriched in VU subjects included members of the phylum Proteobacteria, with a notably high representation of the Enterobacteriaceae family, a finding verified by quantitative polymerase chain reaction (PCR) (Fig. S2) and supported by a prior report{Ellis 2011}. Enriched genera from within this family included known pro-inflammatory pathobionts such as *Salmonella, Escherichia*, *Serratia, Shigella,* and *Klebsiella* species. Additionally, *Staphylococcus, Pseudomonas,* and *Campylobacter* spp., known opportunistic pathogens and sources of bacteremia in HIV-infected subjects{Tumbarello 1995; Bouza 1999; Manfredi 2002}, were highly enriched in the mucosae of VU subjects. Significant reductions in the relative abundance of specific members of Clostridia and Bacteroidia were observed in VU subjects, with

the greatest degree of depletion amongst members of the *Bacteroides* and *Alistipes* genera, which have been previously found to be depleted in inflammatory bowel disease{Frank 2007}. We refer to this community of bacterial taxa (identified as differing significantly in abundance between VU subjects and HIV-uninfected subjects) as the "disease-associated microbial community" (DMC).

Intestinal dysbiosis correlates with markers of disease progression among HIV-infected subjects

While levels of T cell activation (as measured by co-expression of CD38 and HLA-DR) and viral load are well-accepted predictors of HIV disease progression{Deeks 2004}, it is now also appreciated that other immune parameters predict morbidity and mortality, including plasma markers of inflammation, IDO1 activity, and markers of microbial translocation{Kalayjian 2010; Huengsberg 1998; Sandler 2011}. Of note, specific murine gut commensal species have been shown to adhere to mucosal epithelia and induce altered inflammatory cytokine secretion patterns{Ivanov 2009; Atarashi 2011a}, some in association with upregulated IDO1{Atarashi 2011a}. To determine whether associations exist between the mucosa-associated DMC and HIV disease progression, peripheral blood was drawn at the time of mucosal sampling and additional biopsies were obtained for immunophenotyping. Correlations between the relative abundance of taxa within the DMC and immunologic parameters among all HIV-infected subjects (HAART and VU) were examined (Fig. 3A). Although strong associations were not found between taxon abundances and peripheral blood $CD4^+$ T cell count or gut HIV RNA or DNA levels, taxa that were enriched in the DMC correlated with elevated levels of T cell activation, of catabolism of tryptophan through the kynurenine pathway (a marker of IDO1 activity measured by the ratio of plasma concentrations of the downstream product, kynurenine, to the parent compound,

tryptophan [Kyn:Trp]), and of soluble plasma markers of inflammation (such as IP-10, soluble TNF receptor II, and IL-6), and diminished levels of mucosal T cells secreting IL-17 and IL-22. Though abundances of many taxa were found to correlate with these immune parameters (Fig. 3A), Proteobacteria, particularly multiple phylogenetically distinct members of the Enterobacteriaceae members were among the strongest correlates of elevated gut CD4 and CD8 T cell activation (unadjusted P = 0.00004, R_S = 0.82; unadjusted P = 0.00006, R_S = 0.80; Table 3; R_S, Spearman rho correlation coefficient) while members of Staphylococcaceae were among the strongest correlates of higher Kyn:Trp and of elevated plasma IP-10 levels (unadjusted P = 0.00007, R_S = 0.76; unadjusted P = 0.0014, R_S = 0.64). Taxa within the families Micrococcaceae, Rhodobacteraceae, Halomonadaceae, and Pasteurellaceae also exhibited strong correlations with numerous markers of disease progression. Conversely, those taxa depleted in the DMC were found to correlate with lower levels of T cell activation and higher levels of mucosal T cells secreting IL-17, with the strongest correlates of lower T cell activation belonging to Bacteroidaceae (unadjusted P = 0.0017, R_S = -0.68) and the strongest correlates of higher gut IL-17 production being taxa within the Rikenellaceae family (unadjusted P = 0.0004, R_S = 0.72; see Fig. S3 and Table 3 for full list of all correlation coefficients). Collectively, these findings indicate higher abundance of multiple, diverse microbial taxa within the DMC is linked to immune activation and to poor gut IL-17 responses, and that this state is concomitantly characterized by an apparent depletion of taxa linked to immune homeostasis and conditions which promote gastrointestinal integrity, i.e., lower levels of T cell activation and higher levels of IL-17 secretion.

Gut-resident microbial communities function as discrete ecosystems bound together by interdependent metabolic exchange and collective protective mechanisms.

Accordingly, many taxa of the DMC exhibited similar patterns of correlation with immune parameters (Fig. 3A). Therefore, a holistic approach was taken to understand such trends in gross community abundance as they may relate to inflammatory states. Principal component analysis (PCA) was performed to consolidate patterns among co-varying microbial abundances into a handful of summary variables and to compare such variables to immune parameters. The first principal component (PC1) of this analysis comprised 52.8% of DMC variance (Fig. 3B and Fig. S4A), was able to differentiate HIV- from VU (Fig. 3C), and was chiefly characterized by opposing trends in the relative abundance of Proteobacteria (including Pasteurellaceae, Mycobacteriaceae, Pseudomonadaceae, and Enterobacteriaceae members) and Bacteroidaceae (Fig. S4B). We used PC1 as a measure of the degree to which a subject's gut microbiota was similar to that of VU or HIV- (and therefore a surrogate marker for dysbiosis), and related it to markers of immunopathology and disease progression among all HIV-infected subjects. PC1 was found to associate strongly with plasma IP-10 (adjusted P = 0.044), a marker of systemic inflammation shown to correlate with accelerated HIV disease progression{Sauce 2011; Kamat 2012}, with Kyn:Trp (adjusted P = 0.007), and with T cell activation in the gut and peripheral blood (respective significance values shown in figure) (Fig. 3D, Table 2).

Heterogeneity in microbiota of subjects undergoing HAART correlates with markers of disease progression

Even though HAART is effective in diminishing plasma viremia to nearly undetectable levels, subjects receiving long-term treatment can exhibit persistently elevated levels of chronic immune activation{Hunt 2003}, of which the etiology is unknown. We hypothesized that persistence of even a subset of the DMC could continue to sustain pathologic

chronic immune activation in this population despite suppression of viral replication. To address this hypothesis, the relative abundance of DMC members was averaged across each of the VU, HAART, and HIV- subject groups and compared between these groups. The DMC in HAART subjects was found to represent an intermediate state between that observed in VU subjects and uninfected controls (Fig. 4A). Additionally, longitudinal sampling of two HIV-infected subjects prior to and following nine months of effective HAART demonstrated a relative shift in the DMC to more closely resemble that of the HIV- group (Fig. S5). However, individual subjects within the HAART group exhibited a spectrum of relative abundance for members of the DMC, such that the microbial community composition of some subjects was more similar to that of HIV- subjects while the composition of others was more similar to that of VU subjects (Fig. 4B).

We considered a number of possibilities to account for the presence of this spectrum in the DMC of subjects on HAART. Since prophylactic antibiotics are commonly used by subjects on HAART and may influence microbial community composition, relationships between PC1 and antibiotic administration history (for which data was available for 13 out of 18 subjects) were analyzed. No significant trends were observed, as was the case for comparisons to age and time on HAART (Fig. S6). Furthermore, gut inflammation itself has been found to stimulate the outgrowth of facultative anaerobes, including members of the Enterobacteriaceae family{Winter 2010; Winter 2013; Rivera-Chávez 2013}, by providing growth substrates in the form of anaerobic respiration terminal electron acceptors. Expression of genes involved in oxygen and nitrogen radical production (NADPH oxidase [CYBB subunit], arginase 1, nitric oxide synthase 2 [NOS2], type I interferons [IFN-α, IFN-β], type II interferon [IFN-γ], tumor necrosis factor-alpha [TNF-α]) was accordingly quantified in whole biopsies and compared to PC1, and no significant associations were found (Table 4). Rather, associations were found between PC1 and

numerous markers of disease progression, including plasma IL-6 levels, Kyn:Trp, soluble tumor necrosis factor receptor II levels in plasma, blood coagulation marker D-Dimer levels in plasma, gut IDO1 transcript, blood memory CD8 T cell activation, plasma IP-10 levels, and gut CD4 T cell activation ($P < 0.05$, $|R_s| > 0.5$; Table 3). After adjusting for multiple comparisons, significant relationships remained between PC1 and two disease markers: Kyn:Trp and plasma IL-6 concentration (Fig. 4C). To understand whether these associations were influenced by potentially confounding variables and to incorporate a greater proportion of microbial community abundance trends, multiple linear regression analysis was performed. This approach allows for simultaneous consideration of multiple adjustment variables (antibiotics usage history, days on HAART, age, race) and the five most representative principal components of DMC variance (PC1, PC2, PC3, PC4, and PC5) toward understanding which variables are most predictive of the immunologic parameters of interest (Kyn:Trp and plasma IL-6). Out of all possible models involving these variables, the model with the best leave-one-out cross-validated{Picard 1984} (LOO-CV) predictive accuracy was chosen and p-values were obtained by permutation. PC1 and ethnicity were significantly correlated with Kyn:Trp in the best scoring model ($P = 0.0006$ and $P = 0.0036$, respectively). In the case of plasma IL-6 levels, the variable lending greatest predictive power to the most predictive multivariate model was PC1 ($P = 0.018$), while four other variables contributed less significantly: PC4 ($P = 0.032$), ethnicity ($P = 0.048$), age ($P = 0.027$), and duration of HAART ($P = 0.020$). Comparisons of Kyn:Trp and plasma IL-6 to PC1 alone are shown in Fig. 4C.

Members of the DMC participate in immunoactive tryptophan catabolism

IDO1 is induced during inflammation and is thought to actively regulate immune activation by its numerous effects on T cells, including inhibition of IL-17 secretion and

induction of Foxp3-expressing regulatory T cells (T_{reg}){Mezrich 2010}. These effects are mediated by its production of kynurenine pathway compounds from its enzymatic substrate tryptophan, a reaction for which IDO1 is thought to govern the rate-limiting step{Fujigaki 2001}. Consistent with previous findings, we found that IDO1 is strongly associated with type II interferon expression in colonic mucosal tissue{Taylor 1991} and the ratio of plasma concentrations of kynurenine to tryptophan (Kyn:Trp) was indeed associated with reduced IL-17 secretion in CD4 T cells as well as in a higher abundance of T_{reg} in the colon. However, though there was a significant correlation between IDO1 expression and plasma Kyn:Trp in VU subjects, there was no such association in the HAART HIV-infected subject group (Fig. S7).

Given that a significant association was found between our marker of dysbiosis (PC1) and Kyn:Trp within the HAART subject group, we postulated that the DMC may directly contribute to immunomodulatory tryptophan catabolism through the kynurenine pathway. Comparing taxon abundances to plasma Kyn:Trp, we found that taxa within 140 genera positively correlated significantly with Kyn:Trp. Though bacteria are among the primary producers of the essential amino acid tryptophan, relatively few are known to encode the enzymatic machinery for tryptophan catabolism{Lima 2009; Matthijs 2004; Farrow 2007; Kurnasov 2003}. Indeed, analysis of genomic and proteomic annotations from the UniProt Consortium{UniProt Consortium 2012} revealed several bacterial genera within Proteobacteria that encode enzymes homologous to those necessary for catabolism of tryptophan to the most potent immunomodulatory product of the kynurenine pathway{Desvignes 2009; Favre 2010}, 3-hydroxyanthranilic acid (3-HAA; Fig. 5A). Comparing this list to that of genera enriched in the DMC, we found that genera that bore genetic homologs to tryptophan catabolism enzymes of the kynurenine pathway were preferentially enriched in the DMC (Fig. 5B), including *Pseudomonas,*

Xanthomonas, Burkholderia, Stenotrophomonas, Shewanella, and *Bacillus,* as well as members of the families Rhodobacteraceae, Micrococcaceae, and Halomonadaceae. Furthermore, genera whose members bore genetic homologs to 3-4 enzymes within the immunoactive kynurenine pathway correlated more strongly with Kyn:Trp ratios than genera that bore 1-2 tryptophan catabolism enzyme homologs, and even more so than genera that had no genetic evidence for kynurenine pathway metabolic machinery (Fig. 5C, Fig. S8).

To confirm the capacity for production of kynurenine pathway catabolites among members of the DMC, a representative species of the strongest bacterial correlate to plasma Kyn:Trp ratios (genus *Pseudomonas*) was assayed for kynurenine and 3-HAA production in pure culture. *Pseudomonas fluorescens* strain A506 was found to produce micromolar concentrations of kynurenine (Fig. 5D) *in vitro*, further supporting the hypothesis that gut mucosal-adherent bacteria associated with HIV infection might directly contribute to immunoactive tryptophan catabolism during HIV disease.

Discussion

Here, we report observations of differences in the composition of the colonic mucosal-adherent bacterial microbiota of untreated HIV-infected subjects as compared to uninfected risk-matched controls. This altered community prominently includes known pathobionts, such as *Staphylococcus spp., Pseudomonas spp.,* and Enterobacteriaceae family members with pro-inflammatory potential. While these bacteria are documented sources of bacteremia in HIV-infected subjects with advanced disease{Tumbarello 1995; Manfredi 2002}, it is possible that they assume an important but overlooked pathologic role outside the setting of

overt bacteremia. Within all HIV-infected subjects that we studied (n=24), the composition of this community correlated with markers of disease progression and correlates of morbidity and mortality in HIV-infected populations{Giorgi 1999, Sandler 2011; Huengsberg 1998; Hunt 2003; Kalayjian 2010; Sauce 2011}, namely, higher levels of T cell activation in the blood and gut as well as higher levels of tryptophan metabolism and elevated plasma markers of inflammation. Strikingly, when considering aviremic subjects on effective HAART (n=18), correlations remained evident between dysbiosis and two immunologic correlates of disease, production of kynurenine and plasma levels of the inflammatory cytokine IL-6, suggesting a relationship between the composition of the gut microbial community and predictors of disease outcomes even in treated subjects. Finally, numerous bacterial taxa of the DMC that were enriched in untreated HIV infection were found to encode the enzymatic machinery that performs the same catabolic function as human IDO1. Subsequent *in vitro* experiments confirmed the capacity of a model member of this community in the genus *Pseudomonas* to transform tryptophan to kynurenine.

While the mechanistic relationships defining the observed associations are not fully understood, a complex crosstalk between gut immune states and the microbiome in the context of HIV disease progression and treatment is tenable. Previous studies in the mouse support the idea that HIV-associated gut inflammation may contribute to the observed dysbiotic states{Lupp 2007; Nagalingam 2011; Stecher 2007; Winter 2010; Winter 2013}. Gut inflammation may induce changes in nutrient microenvironments via production of oxygen and nitrogen radicals that allow outgrowth of facultative anaerobes such as members of the Enterobacteriaceae family{Winter 2010; Winter 2013}. While our current investigation did not find correlations between enzymes responsible for reactive oxygen and nitrogen species and dysbiosis, we were

limited by measurements of enzyme expression in whole tissue as opposed to direct measurement of enzymatic endproducts or measurement of enzyme expression in specific cell lineages (e.g. tissue macrophages). Perturbations to gut immune barrier function (e.g., disruptions in microbial pattern recognition sensors) have also been found to result in the outgrowth of dysbiotic communities in mice{Elinav 2011; Henao-Mejia 2012}. Numerous cells that participate in microbial recognition and response are dysregulated during progressive HIV/SIV disease, including dendritic cells{Manches 2008}, macrophages{Müller 1990}, Th17 cells{Brenchley 2008}, enterocytes{Somsouk 2012}, and Paneth cells{Kelly 2004}. Such cells are critical regulators of gut microbial community composition and their disruption may allow for outgrowth of the dysbiotic community described herein.

While gut inflammation may well engender dysbiosis, murine studies also suggest that the dysbiotic communities that result from inflammation and immune barrier defects can reciprocally act to stimulate chronic inflammation in the gut and systemic circulation{Garrett 2007; Garrett 2010; Elinav 2011; Henao-Mejia 2012}. Although no bacterial species found in humans have been similarly described to directly alter or augment IDO1 activity, T cell differentiation, or gut inflammation, analogous immunomodulatory relationships are likely to exist between gut bacteria and human hosts. If so, the capacity for the mucosal immune system to regulate the abundance of such immunomodulatory species may be impaired during HIV disease. Indeed, we find that HIV-infected subjects are likely to harbor a community of enteropathogenic bacteria that can catabolize tryptophan into immunomodulatory kynurenine derivatives known to correlate with disease progression and thought to contribute to mucosal immune disruption{Favre 2010, Huengsberg 1990; Sandler 2012}. We show here that a model member of this community, a strain from within the genus bearing strongest correlation to

Kyn:Trp catabolism (*Pseudomonas*), is indeed able to catabolize tryptophan to kynurenine. Furthermore, members of this disease-associated community (including *Pseudomonas, Xanthomonas, Bacillus,* and *Burkholderia* spp.) have been shown to utilize kynurenine pathway catabolites as growth substrates{Palleroni 1964; Rosenfeld 1969; Farrow 2007; Bouknight 1975; Chang 2003; Davis 1951}. In aggregate, these observations suggest that the outgrowth of this community may be driven by the persistent host IDO1 activity found during chronic HIV infection and may in turn exacerbate IDO1-mediated immune disruption by accelerating tryptophan catabolism along the kynurenine pathway. If so, the potential exists for a self-amplifying cycle of dysbiosis and inflammation. A re-envisioning of the pathogenic mechanisms associated with HIV disease progression may thus include dysregulation of the balance between homeostasis-promoting versus gut barrier-disrupting microbes, contributing to weakened mucosal immune function, microbial translocation, and subsequent chronic inflammation. This process may further select for functionally distinct, enteropathogen-enriched communities in the gut mucosa that aggravate these immunologic phenomena. Intriguingly, the one HIV-infected long-term non-progressor assayed in this study (classified as such by a stable peripheral blood CD4 count despite over 21 years of untreated infection) revealed a gut community similar to that of uninfected subjects, presenting the interesting possibility that non-progressive disease may be linked to host-microbiome homeostasis.

Notably, a potentially detrimental host-microbiome interaction in the setting of HIV disease is not always reversed by effective antiretroviral treatment. Indeed, HAART subjects exhibited higher relative abundances of the identified disease-associated microbial community than did uninfected controls, and those with the greatest degrees of dysbiosis were characterized by elevated markers of chronic inflammation and disease progression, i.e., high levels of

kynurenine production and of the plasma inflammatory marker IL-6{Kalayjian 2010}. Previous findings support the possibility that the pathobiont community of Enterobacteriaceae enriched in untreated infection can indeed potentiate immunologic disruptions similar to those found in HIV disease{Garrett 2007; Garrett 2010; Stecher 2007; Winter 2010}. *Bacteroides* genus members, which can protect mice from colitis{Mazmanian 2008} and that are notably depleted in sufferers of inflammatory bowel disease{Frank 2007} and asthma{Björkstén 1999}, were similarly found in relative depletion among our cohort of HIV-infected untreated subjects.

Finally, we present data suggesting that a diverse community of gut-resident bacteria may augment IDO1-mediated mucosal disruption by direct catabolism of tryptophan through the kynurenine pathway. Targeting these microbial communities may present novel avenues for therapeutic interventions aimed at preventing or reversing HIV-associated immunopathology. Indeed, recent reports of trials using agents that modulate the composition of the microbial community in HIV-infected populations or SIV-infected macaques (e.g., wtih antibiotics{Walker 2010}, probiotics{Anukam 2008; Irvine 2010; Klatt 2013}, and prebiotics{Gori 2011}) have shown improvements in clinical outcomes as measured by $CD4^+$ T cell recovery, reduced mortality, and decreased markers of microbial translocation. However, microbial population dynamics during HIV disease progression remain poorly understood and efforts at developing therapeutics targeting the microbiome will certainly be augmented by a greater understanding of the host-microbiome relationship in HIV-infected subjects. Further efforts to determine the mechanistic processes underlying the complex relationships between the gut microbiome and its host will be critical toward developing the potential of gut microbial community modulation as a therapeutic strategy, and may lead to novel approaches for the management of HIV disease as well as of other chronic inflammatory conditions.

Acknowledgements & Footnotes

We would like to acknowledge S. Yukl for HIV viral load measurements, M. Krone for cohort characteristics data, M. Segal for statistical analysis consultation, and J. Robinson, B. Clagett, S. Sirdeshmukh and D. Dorazio for assistance in plasma inflammatory marker analyses. We would like to additionally thank P. Baum, W. Yonemoto, J. Mar, E. Krow-Lucal, M. Lowe, A. Chitre, C. Baker, B. Kanwar, T. Burt, N. Nagalingam, K. Fujimura, M. Rauch, and S. Galang for helpful discussion. This study was funded by NIH NRSA 5T32AI007334-23, NIH U19 AI96109, NSF 1144247, NIH UCSF-GIVI Center for AIDS Research P30-AI027763, and by the Harvey V. Berneking Living Trust. JMM is a recipient of the NIH Director's Pioneer Award Program, part of the NIH Roadmap for Medical Research, through grant number DPI OD00329. Additional support was provided by the American Foundation for AIDS Research (107854-48-RGRL to JMM and SGD), the Hurlbut-Johnson Fund administered by the AIDS Research Institute at UCSF (to RMD), R37 AI040312 (to JMM), K23 CA157929 (to MS), F32 AI091534 (to RMD), the California HIV/AIDS Research Program (ID09-SF-067 to PWH), and the UCSF Clinical and Translational Science Institute (RR024131-01). IVC, RMD, MJB, SVL, and JMM designed experiments. IVC and SI performed PhyloChip experiments. IVC, MCM, RDH, and SI performed PhyloChip analysis. RMD and RA performed flow cytometry and analysis. MS, PWH, and SGD recruited subjects and obtained human samples. YH and ML provided plasma inflammatory marker measurements. IVC, RMD, JMM, and SVL wrote the manuscript. PhyloChip data used in this study are available at the Gene Expression Omnibus web site (https://www.ncbi.nlm.nih.gov/geo/) under accession number: GSE43746.

Figures

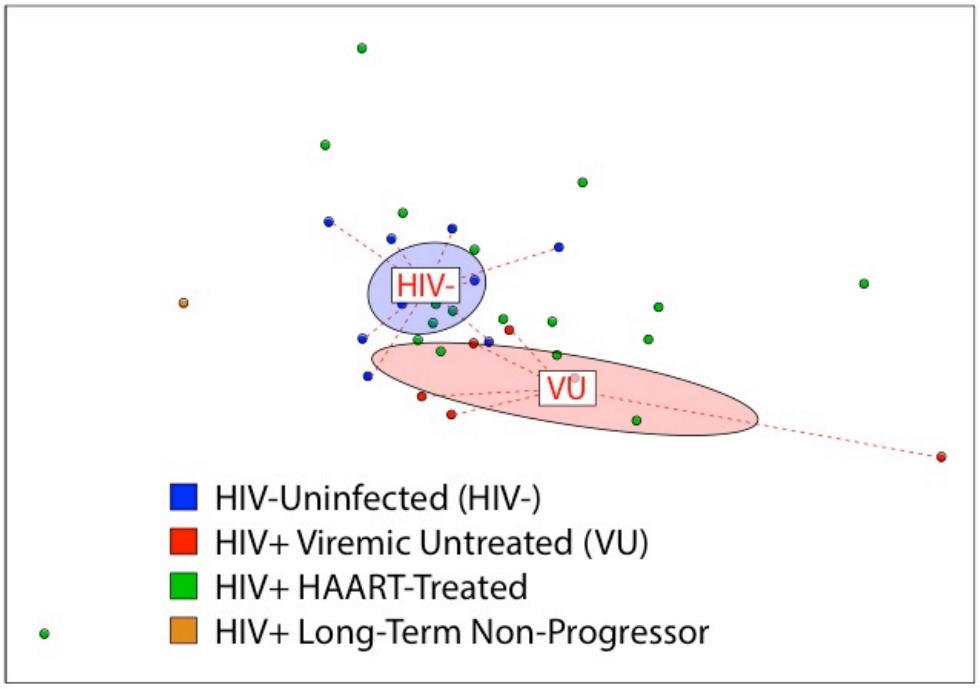


Figure 1: Gut bacterial microbiota composition in HIV-infected viremic untreated (VU) subjects differs from that of HIV-uninfected risk-matched controls.

Community composition dissimilarity was analyzed using microbiota profiles generated using PhyloChip. Non-metric dimensional scaling (NMDS) was used to plot each sample community based on a Canberra community dissimilarity matrix. Each dot represents the microbiome of a single subject, and color indicates subject group as denoted in legend. Ellipses represent 95% confidence intervals for standard error of weighted NMDS score means of respective groups, and are colored based on subject group (red denotes HIV+VU subjects, while blue denotes HIV-). Community differences were verified using Adonis, P = 0.002.

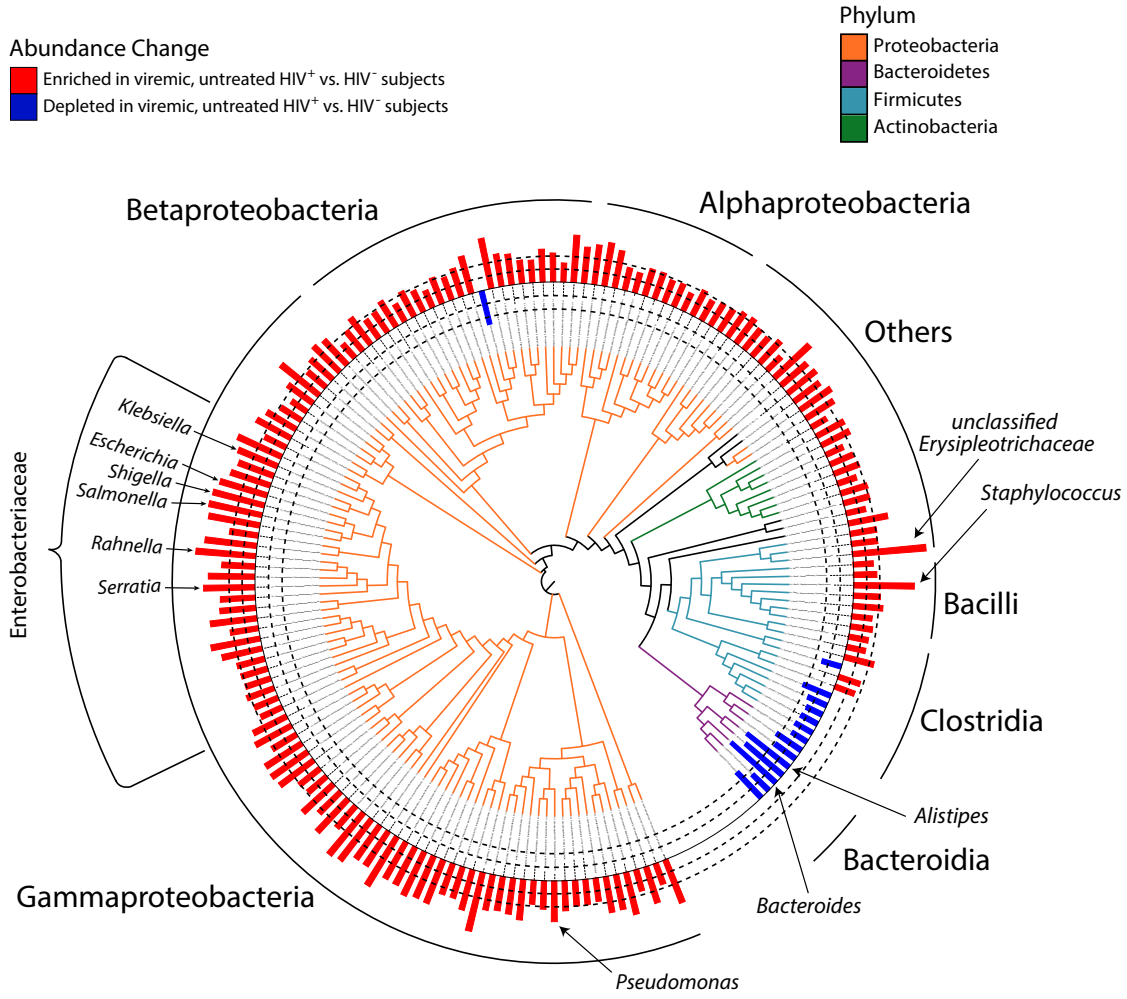


Figure 2: Phylogenetic distribution of the disease-associated microbial community (DMC), defined as significantly depleted or enriched in VU subjects compared with HIV- subjects.

Phylogenetic tree representing 16S rRNA gene sequence relatedness of taxa identified as significantly enriched or depleted in relative abundance between VU and HIV- subjects. The 625 taxa identified belonged to 172 unique genera, and representative taxa for each unique genus are depicted. Tree branches are colored-coded by phyla as indicated in legend. Red and blue bars depict bacterial genera in greater (red bars) and lesser (blue bars) abundance in VU subjects.

Dashed lines indicate magnitude of relative abundance change (inner dashed circle denotes 500 MFI, outer denotes 2000 MFI difference between comparative groups; 1,000 MFI change represents a $\log_2$ difference in relative abundance).

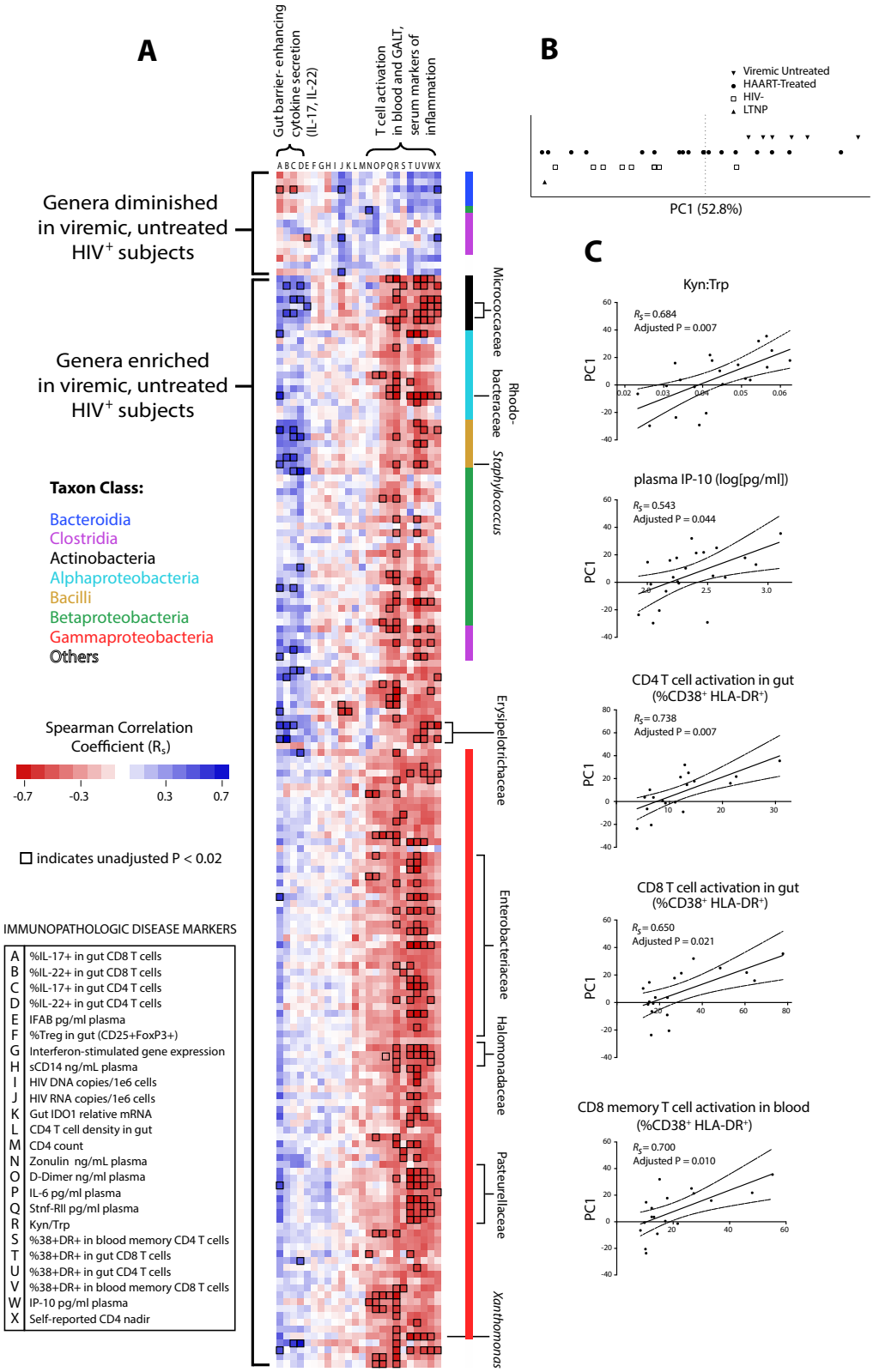

A
Gut barrier- enhancing cytokine secretion (IL-17, IL-22)
T cell activation in blood and GALT, serum markers of inflammation
A B C D E F G H I J K L M N O P Q R S T U V W X
Genera diminished in viremic, untreated HIV+ subjects
Genera enriched in viremic, untreated HIV+ subjects
Micrococcaceae
Rhodo-bacteraceae
Staphylococcus
Erysipelotrichaceae
Enterobacteriaceae
Halomonadaceae
Pasteurellaceae
Xanthomonas
Taxon Class:
Bacteroidia
Clostridia
Actinobacteria
Alphaproteobacteria
Bacilli
Betaproteobacteria
Gammaproteobacteria
Others
Spearman Correlation Coefficient (Rs)
-0.7 -0.3 0.3 0.7
indicates unadjusted P < 0.02
IMMUNOPATHOLOGIC DISEASE MARKERS
A %IL-17+ in gut CD8 T cells
B %IL-22+ in gut CD8 T cells
C %IL-17+ in gut CD4 T cells
D %IL-22+ in gut CD4 T cells
E IFAB pg/ml plasma
F %Treg in gut (CD25+FoxP3+)
G Interferon-stimulated gene expression
H sCD14 ng/mL plasma
I HIV DNA copies/1e6 cells
J HIV RNA copies/1e6 cells
K Gut IDO1 relative mRNA
L CD4 T cell density in gut
M CD4 count
N Zonulin ng/mL plasma
O D-Dimer ng/ml plasma
P IL-6 pg/ml plasma
Q Stnf-RII pg/ml plasma
R Kyn/Trp
S %38+DR+ in blood memory CD4 T cells
T %38+DR+ in gut CD8 T cells
U %38+DR+ in gut CD4 T cells
V %38+DR+ in blood memory CD8 T cells
W IP-10 pg/ml plasma
X Self-reported CD4 nadir
B
Viremic Untreated
HAART-Treated
HIV-
LTNP
PC1 (52.8%)
C
Kyn:Trp
Rs = 0.684
Adjusted P = 0.007
plasma IP-10 (log[pg/ml])
Rs = 0.543
Adjusted P = 0.044
CD4 T cell activation in gut (%CD38+ HLA-DR+)
Rs = 0.738
Adjusted P = 0.007
CD8 T cell activation in gut (%CD38+ HLA-DR+)
Rs = 0.650
Adjusted P = 0.021
CD8 memory T cell activation in blood (%CD38+ HLA-DR+)
Rs = 0.700
Adjusted P = 0.010
PC1

Figure 3: Bacterial community enriched in untreated infection associates with immunopathologic markers of HIV disease progression within HIV-infected subjects.

(A) Spearman correlations between individual taxa of the DMC and immunologic parameters were calculated among all HIV-infected subjects (treated and untreated), collated, and represented by color. Correlation coefficients for taxa with representative species belonging to the same genus and displaying similar FI trends were averaged for visualization purposes. Intersections of each genus within the DMC (rows) and immunologic parameters (columns) represent the strength of the positive (red) or negative (blue) correlation as determined by Spearman correlation coefficient (R_S). Bacteria depleted in VU subjects correlate positively with protective immune markers (low T cell activation in the blood and gut, higher gut IL-17 secretion), while bacteria enriched in VU subjects correlate with markers of disease progression (converse of above, as well as lower T cell-mediated IL-22 secretion in the gut, higher Kyn:Trp, and higher concentrations of plasma markers of inflammation [IP-10, IL-6, and soluble TNF receptor II]). Families and genera of taxa that exhibited numerous strong correlations with immunologic parameters are indicated. Intersections enclosed by a black box represent associations with unadjusted $P < 0.02$. (**B**) First principal component (PC1) representing trends in DMC bacterial community composition accounts for 52.8% of bacterial community variance, and separates VU subjects from HIV-, while subjects undergoing HAART have PC1 values distributed amongst those for VU and HIV- subjects (PC1, Principal Component 1; LTNP, Long Term Non-Progressor). **(C)** First principal component correlates with markers of disease progression: plasma concentration of inflammatory marker IP-10, IDO activity (Kyn:Trp), gut T cell activation, and peripheral blood CD8 T cell activation (R_S, Spearman rho correlation

coefficient; dotted lines represent 95% confidence interval bands for linear regression coefficients).

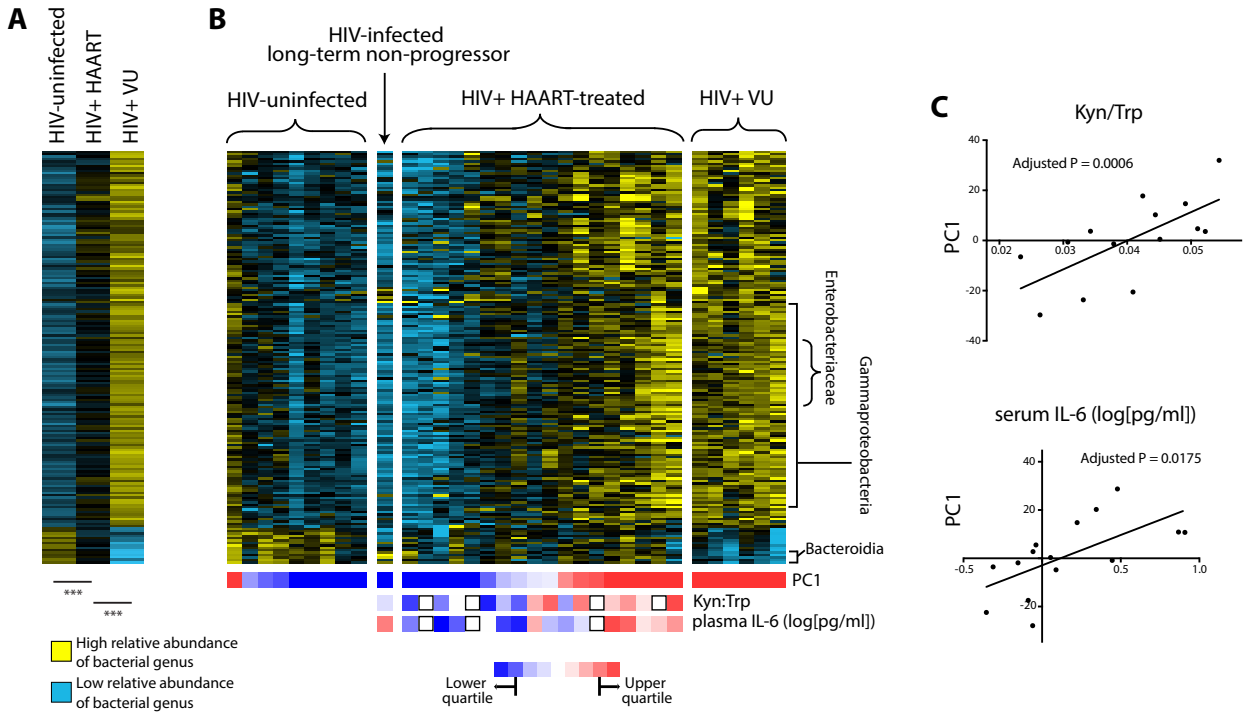


Figure 4: Relative abundance of DMC members is diminished in HAART subjects compared to VU and falls along a spectrum of VU-like or uninfected-like bacterial communities.

(A) Shown are relative abundances of genera within the DMC averaged within subject groups. Each row represents a unique genus. A higher relative abundance is denoted by yellow squares while blue denotes lower relative abundance. After assessment of data distribution, a two-tailed, unpaired T-test assuming unequal variance was used to test significance of differences (***$P < 5x10^{-10}$). **(B)** Heatmap of DMC member relative abundances, ordered first by subject grouping, then by subject PC1 values (relative values denoted by color below each subject column). Rows represent genera of the DMC while columns represent individual subject samples. In the treated

subject grouping, subjects were ranked for each variable (i.e. Kyn:Trp and IL-6 concentration), and these rankings are represented by colored squares below each subject column (red indicates high Kyn:Trp and high plasma IL-6 in the first and second rows, respectively). **(C)** Correlations between PC1 and Kyn:Trp or plasma IL-6 concentration. P values displayed represent the contribution of PC1 to the best-fit linear regression model that best predicts each immunologic variable, adjusted to account for multiple comparisons (see Materials & Methods for details).

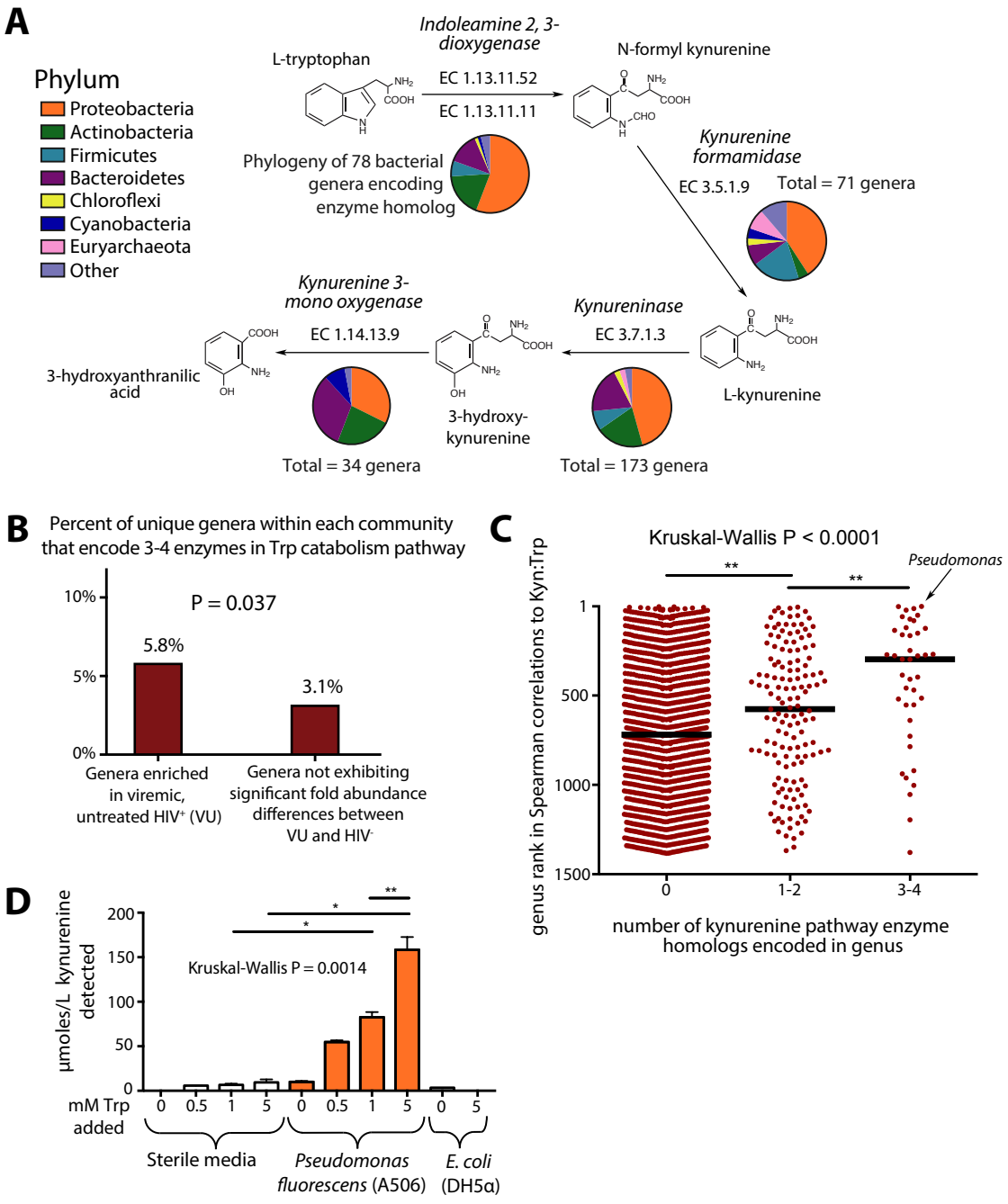


Figure 5: Bacterial tryptophan catabolism machinery is genetically and functionally homologous to IDO1 enzymatic activity and is enriched in the DMC.

(A) The tryptophan catabolism pathway to the most potent immunomodulatory kynurenine derivative, 3-hydroxyanthranilic acid (3-HAA), involves numerous enzymatic steps, for which

homologous enzymes are genetically encoded in numerous bacterial genera as denoted by 'Totals.' Phylogenetic distribution of annotated enzymes in the kynurenine pathway was collated and represented using pie charts as defined in legend. . **(B)** Distribution of genera with genetic homologs of kynurenine pathway enzymes was compared among genera enriched in the DMC and those not identified as having differential fold abundance between VU and HIV-. A Fisher's exact test for enrichment found that, compared to genera not exhibiting differential fold abundance between VU and HIV^-, a greater proportion of genera overrepresented in VU encoded most or all of the genes involved in the tryptophan catabolism pathway ($P = 0.037$). **(C)** All unique genera detected were ranked based on their strength of non-parametric Spearman correlation to Kyn:Trp ratios in HIV-infected subject plasma. Genera with 3 or 4 genetically encoded homologs to kynurenine pathway enzymes exhibited significantly stronger correlations to plasma Kyn:Trp ratios than those with only 1 or 2 homologs, and even more so than those with no annotated homologous enzymes. **(D)** *Pseudomonas fluorescens* strain A506 and a lab strain of *Escherichia coli* (DH5α) were grown in King's B Medium for 20 hours in the presence of varying concentrations of tryptophan. Supernatants from these cultures were assayed for the presence of kynurenine, with concentrations shown as μmoles/liter.

Figure S1: Gross bacterial community metrics do not differ significantly across subject groups or intra-subject samples.

(A) Total bacterial abundances compared between subject groups as measured by TaqMan qPCR assay. The Kruskal-Wallis test was used to test differences in subject groups ($P > 0.05$). **(B)** Total bacterial load is concordant across different biopsies from the same individual and irrespective of DNA extraction method. Y-axis represents total bacterial load measures on

biopsies with nucleic acids extracted using the AllPrep kit (Qiagen, Valencia, CA); X-axis represents separate biopsies from the same individual, taken at the same clinical visit and from the same relative anatomical location, with nucleic acids extracted using Trizol (Life Technologies, Carlsbad, CA). Correlation was tested by the Spearman method (R_S = 0.86, P = 0.0001). **(C)** Bacterial community evenness (Pielou's metric) and richness is not significantly different among subject groups. The Kruskal-Wallis test was used to detect significant differences across all subject groups (P > 0.05).

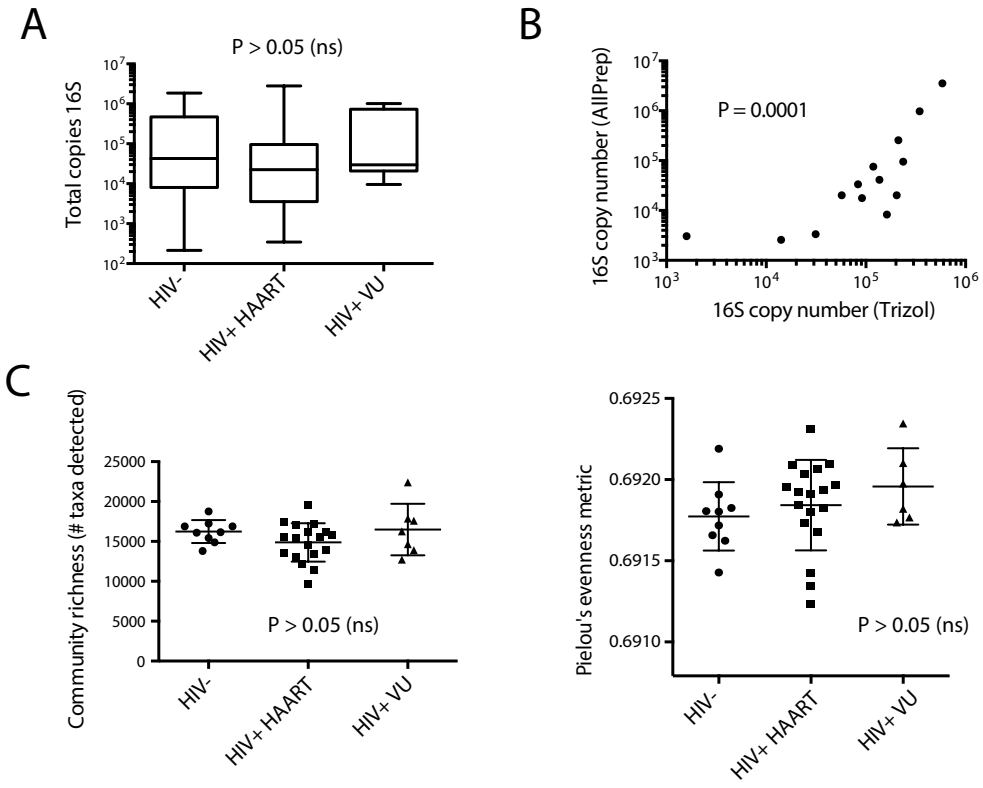


Figure S2: Comparisons of quantitative PCR (qPCR) measurements and PhyloChip relative abundance measurements [mean fluorescence intensity (MFI)] for selected taxa of interest.

(A) PhyloChip MFI of Enterobacteriaceae members within the DMC were averaged and compared by Spearman correlation to qPCR measures of Enterobacteriaceae abundance using sample DNA extracted from mucosal biopsies, as determined by Ct (R_S = -0.66, $p<0.0001$), with linear regression line ($p<0.0001$). Fine dotted line represents qPCR limit of detection (Ct=37), as determined by Ct for which amplification in no-template controls was observed across replicates. Dotted lines represent 95% CI for linear regression coefficients. **(B)** qPCR measures of Enterobacteriaceae differ in VU subjects compared with HIV-uninfected risk-matched controls (P = 0.045). A Mann-Whitney non-parametric test was used to determine significance. Dotted line again represents limit of detection. **(C)** PhyloChip MFI of *Bacteroides fragilis* was compared to qPCR measurements for *B. fragilis* performed on pre-PhyloChip hybridization 16S amplicon samples. Spearman correlation was used to determine significance (R_S = -0.52, P = 0.0033), with linear regression line (P = 0.0002). Dotted lines represent 95% CI for linear regression coefficients.

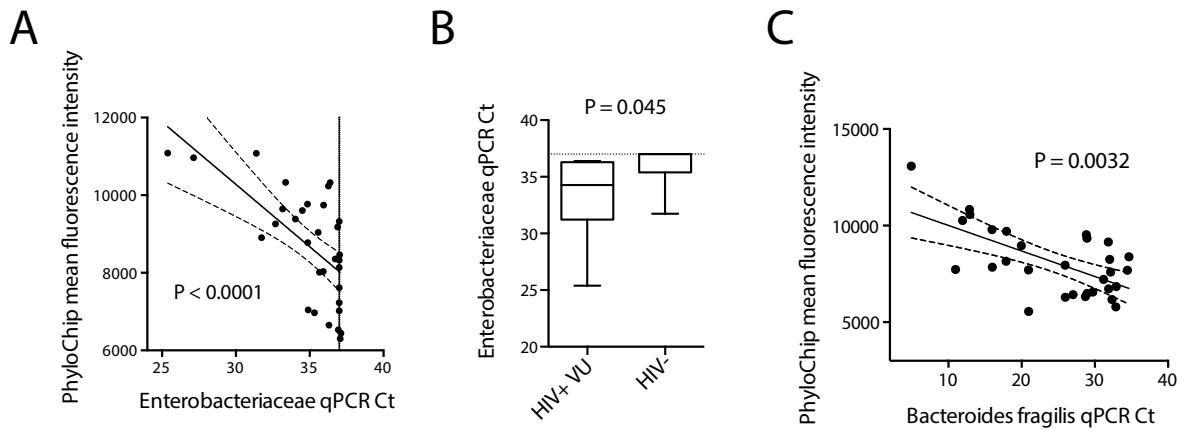


Figure S3: Genus and immunologic parameter identities for genera bearing one or more correlation with unadjusted P < 0.005, from analyses including all HIV-infected subjects studied.

As was done in Figure 3A, Spearman correlation coefficients for taxa with representative species belonging to the same genus and displaying similar MFI trends were averaged, while the median P value of the same Spearman correlations was used to define genera with unadjusted P < 0.005 (boxed intersections).

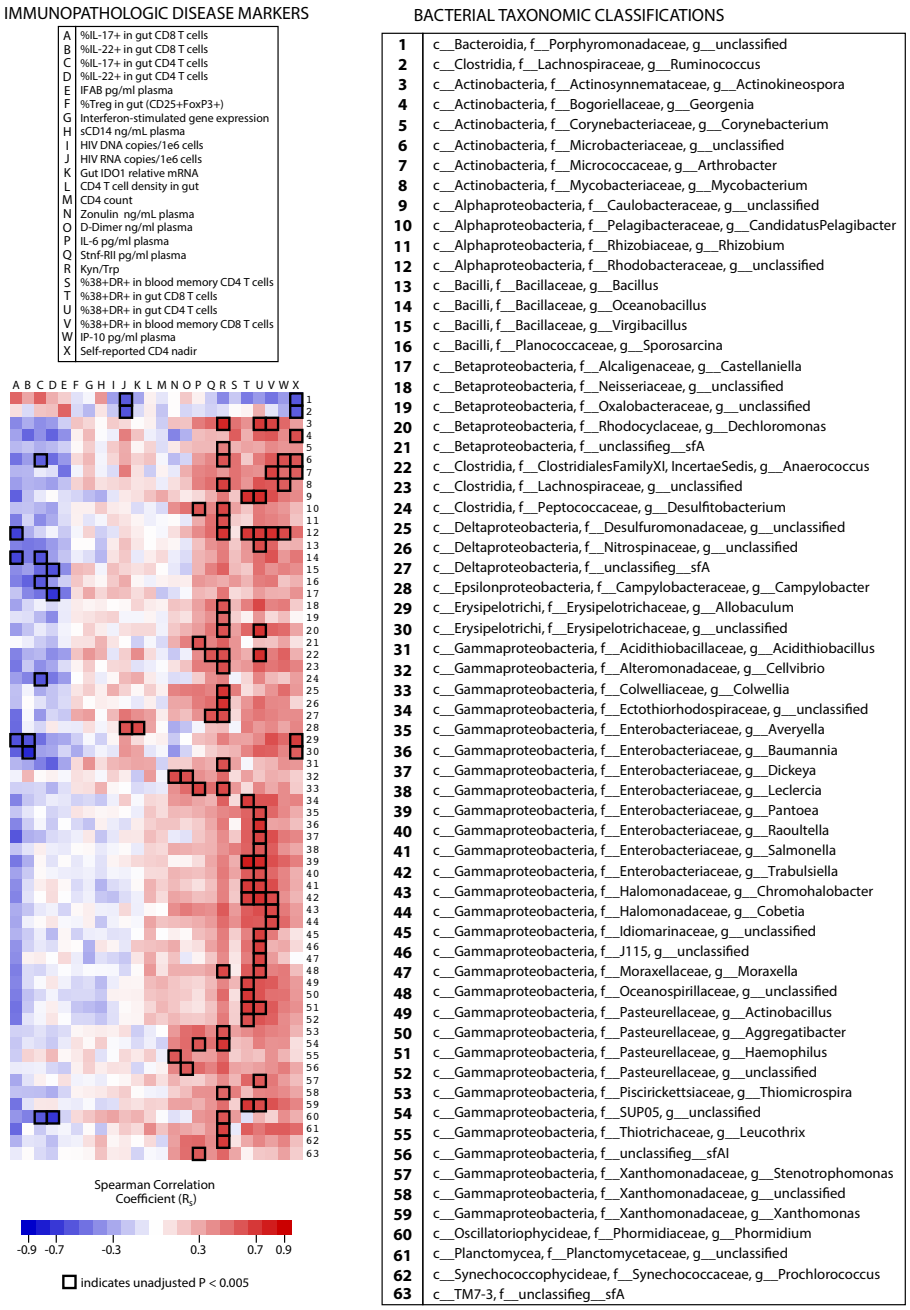


Figure S4: Principal component analysis characteristics.

(A) Proportions of covariance in DMC taxon abundance data explained by first thirty principal components. A principal component analysis was performed as detailed in Methods, the first thirty of which represented 99.45% of DMC community covariance, while the first five principal

components together represented 78.23% of covariance. **(B)** Loading scores for Principal Component 1 (PC1). Values of factor loading scores were ranked and maxima for the top genera loading most heavily toward the positive and negative extremes of PC1 are displayed.

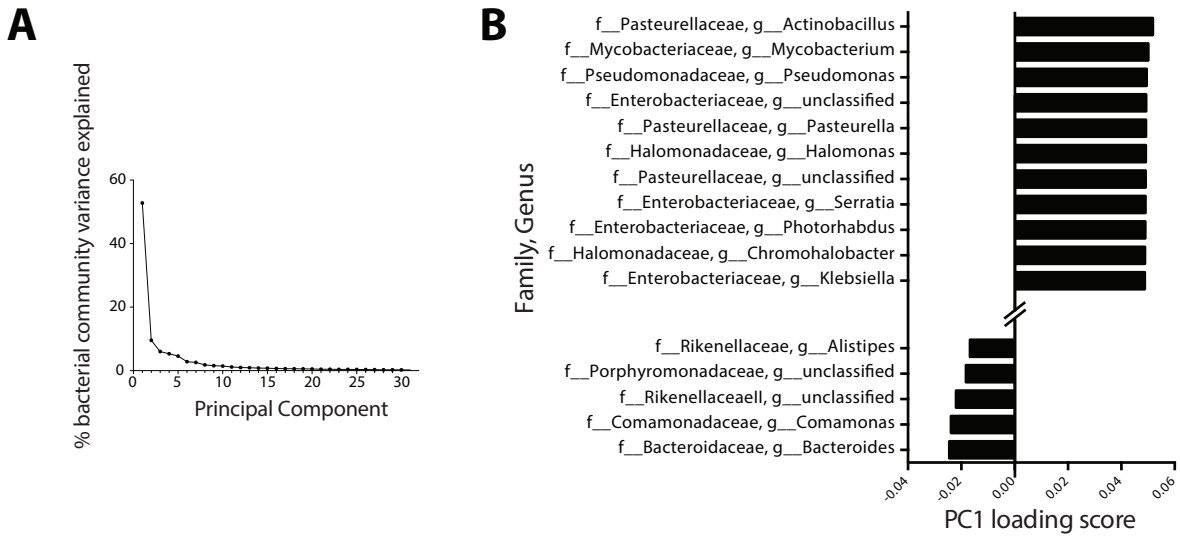


Figure S5: Highly active antiretroviral therapy causes a shift in DMC composition toward the healthy, uninfected state in two subjects.

(A) Two viremic, untreated (VU) subjects returned for follow-up after 9 months of successful HAART (as determined by sustained undetectable plasma viral load). The abundance of genera within the DMC diminished significantly in both subjects upon treatment, as determined by a two-tailed, paired T-test (***$P < 5x10^{-10}$). **(B)** PC1 values are shown for VU (top row), HIV-uninfected (bottom row), and two subjects who were sampled when viremic and untreated and then again after nine months of HAART (middle two rows).

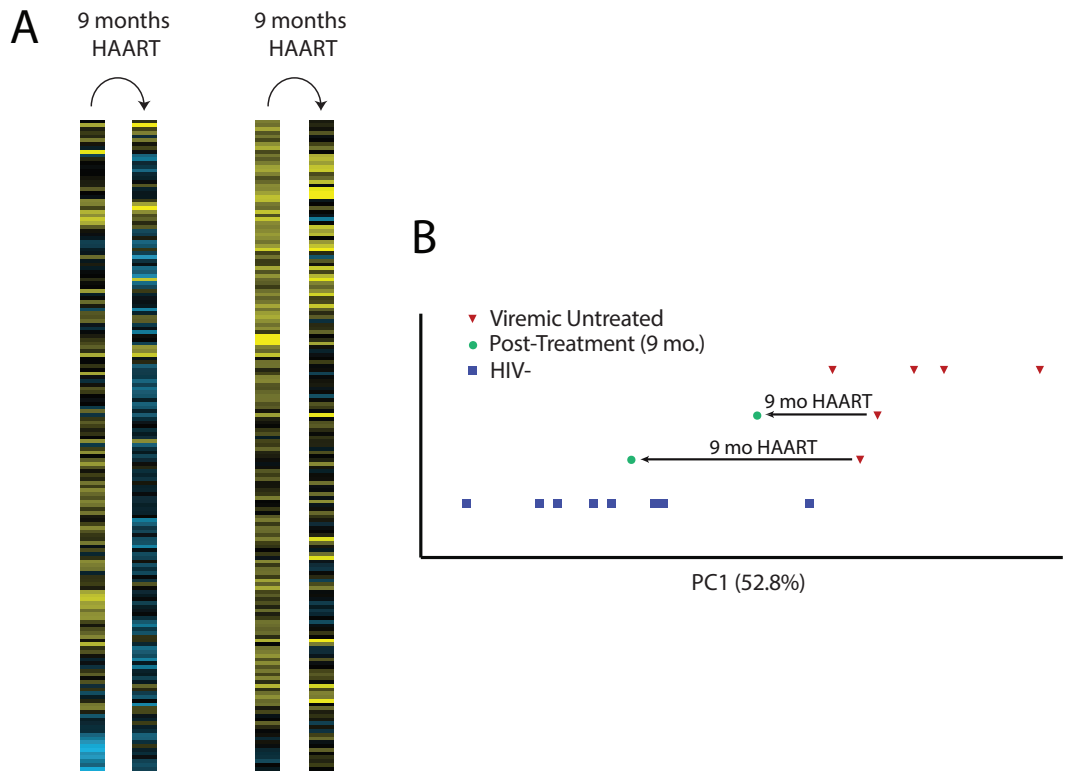


Figure S6: Comparison of adjustment variables (antibiotics usage, days on HAART, age) to PC1.

The Spearman correlation method was used to test trends between antibiotics usage history (data available for 13 out of 18 subjects; P = 0.38), days on HAART (data available for all subjects; P = 0.47), and age (data available for all subjects; P = 0.21).

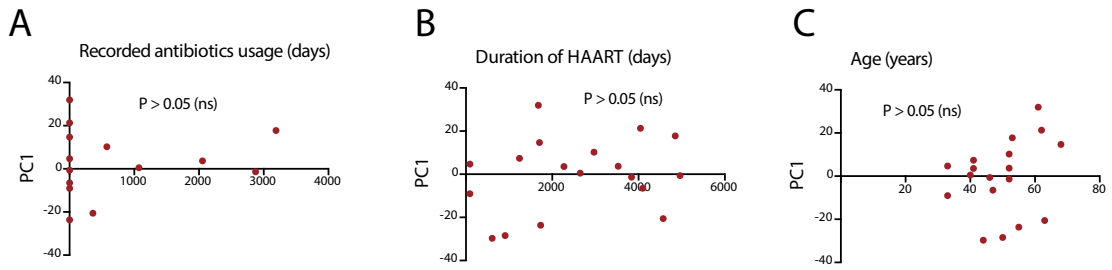


Figure S7: Gut IDO1 expression correlates with Kyn:Trp in VU subjects but not in HAART subjects.

Gut IDO1 expression was measured by quantitative PCR in DNA extracted from whole rectosigmoid biopsy specimens, and normalized to the housekeeping gene, HPRT.

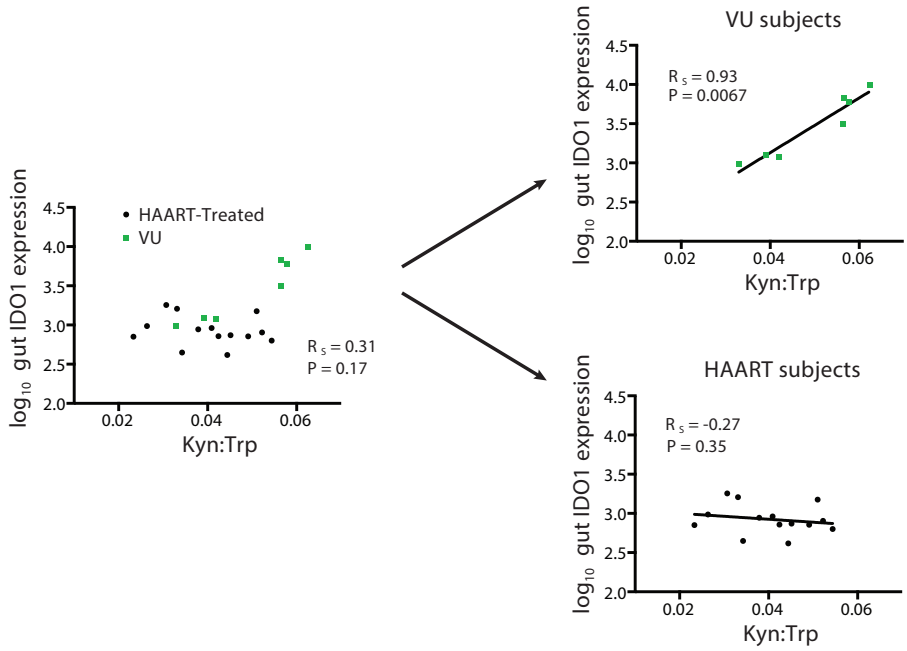


Figure S8: Bacterial taxa that predict plasma Kyn:Trp ratios in a multivariate machine learning analysis preferentially encode genetic homologs of tryptophan catabolism enzymes involved in the kynurenine pathway.

The R package 'randomForest' was used to incorporate analysis of the effects of complex dependencies in microbial abundance data, and allowing for high-resolution multivariate assessment of the predictive capacity of microbial abundances toward plasma Kyn:Trp ratios. Results were similar to those using the univariate Spearman correlation method, in that genera that genetically encoded greater numbers of tryptophan-catabolizing enzymes better predicted Kyn:Trp ratios in plasma of HIV-infected subjects than did those with fewer or no genetic homologs to tryptophan-catabolizing enzymes.

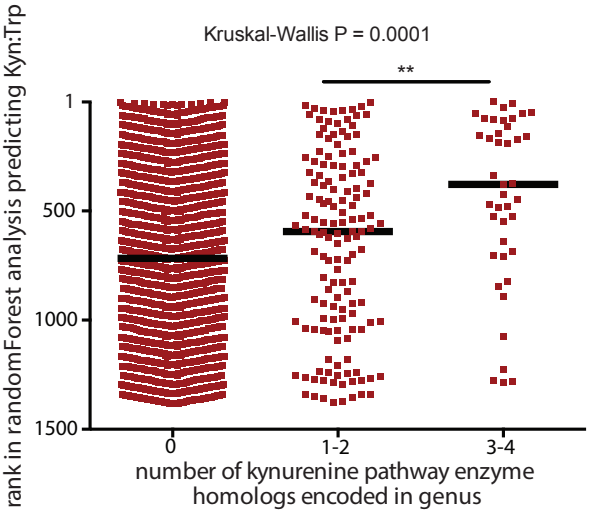
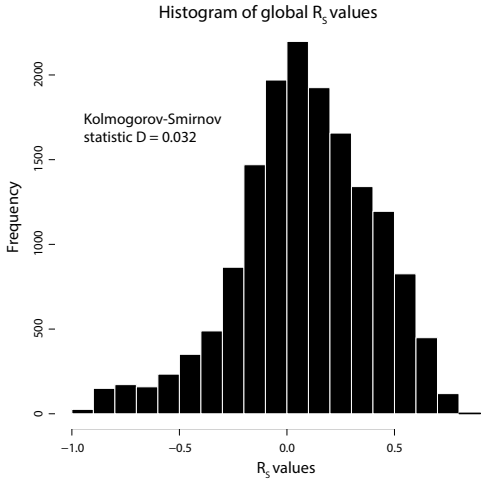


Figure S9: Data distribution of R_S and P values for all taxa as compared to all immunologic variables in Spearman correlation tests.

Histograms depict frequency (y-axis) of each value for R_S and P values (x-axes), and show a normal distribution for R_S values and a non-normal distribution for P values.


Histogram of global R_s values
Kolmogorov-Smirnov
statistic D = 0.032
Frequency
0
500
1000
1500
2000
−1.0
−0.5
0.0
0.5
R_s values

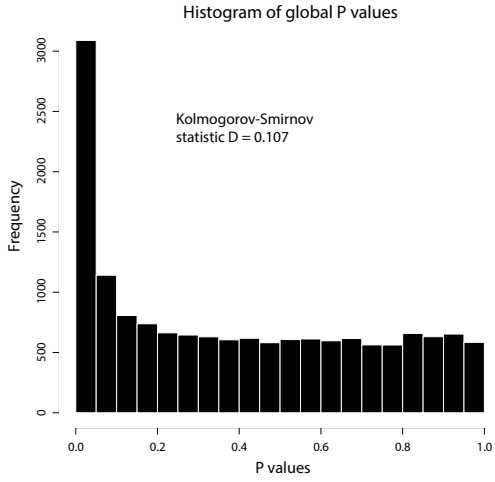

Histogram of global P values
Kolmogorov-Smirnov
statistic D = 0.107
Frequency
0
500
1000
1500
2000
2500
3000
0.0
0.2
0.4
0.6
0.8
1.0
P values

Table 1: Patient cohort data.

Paired blood and rectosigmoid biopsy specimens were collected from a cohort of male HIV-infected and uninfected subjects receiving care at San Francisco General Hospital with characteristics described below:

Median (Range)	Age	Viral Load	Duration of HAART	CD4 counts, cells/ul blood	Years since first seropositive test
Viremic Untreated (n=6)	33 (28-55)	11716 copies/ml (1077-83800)	0 years	356.3 (313-819)	2.20 (0.37-5.81)
HAART-Treated (n=16)*	52 (40-68)	<40 copies/ml	> 2 years	374.5 (251-1110)	21.22 (7.81-32.17)
Uninfected (n=9)	46 (36-59)	<40 copies/ml	0 years	803 (460-1487)	0
Long Term Non-Progressor (n=1)	34	820 copies/ml	0 years	505	21.12

Ethnicity (n)	African American	Asian	Caucasian	Hispanic/ Latino	Multiracial	Native American	Pacific Islander
Viremic Untreated (n=6)	1	1	3	0	1	0	0
HAART-Treated (n=16)*	3	0	10	2	0	0	1
Uninfected (n=9)	4	1	3	1	0	0	0
Long Term Non-Progressor (n=1)	0	0	1	0	0	0	0

***Note:** Two viremic untreated subjects returned after nine months of effective HAART (as determined by lack of detectable plasma viremia)

Table 2: Spearman correlations between PC1 and immunologic variables among all HIV-infected subjects.

Immune Variable	Spearman R_S value	Unadjusted P value	Benjamini -Hochberg adjusted P value
%CD38+DR+ amongst gut CD4+ T cells	0.738	0.00047	0.00746
Kyn:Trp ratio	0.684	0.00062	0.00746
%CD38+DR+ amongst blood memory CD8+ T cells (all CD8+ after exclusion of CD27+CD45RO-)	0.7	0.00121	0.00966
%CD38+DR+ amongst gut CD8+ T cells	0.649	0.00354	0.02121
IP-10 pg/ml	0.543	0.00907	0.04355
Soluble TNF-Receptor II pg/ml plasma	0.493	0.01975	0.07902
%IL-17+ in gut CD8+ T cells	-0.423	0.06344	0.21691
Self-reported CD4+ T cell nadir	0.366	0.0723	0.21691
IL-6 pg/ml plasma	0.379	0.08205	0.21881
%CD38+DR+ in blood memory CD4+ T cells (all CD4+ after exclusion of CD27+CD45RO-)	0.389	0.11054	0.2653
D-Dimer ng/ml plasma	0.223	0.31842	0.63683
Zonulin ng/ml plasma	0.223	0.31842	0.63683
CD4+ T cell density in gut	0.222	0.37623	0.65517
%IL-22+ amongst gut CD8+ T cells	-0.196	0.38218	0.65517
%IL-17+ amongst gut CD4+ T cells	-0.162	0.45994	0.69009
HIV RNA copies/1e6 cells	0.162	0.46006	0.69009
IFAB pg/ml plasma	-0.15	0.50629	0.69746
%IL-22+ amongst gut CD4+ T cells	-0.14	0.5231	0.69746
Cumulative measure of interferon-stimulated gene (ISG) expression*	0.104	0.65403	0.80936
sCD14 ng/ml plasma	-0.097	0.67447	0.80936
%Treg in gut (CD25+FoxP3+) of total CD4+ T cells	0.078	0.71192	0.81363
Gut IDO1 relative mRNA	0.069	0.74976	0.81792
HIV DNA copies/1e6 cells	0.057	0.80098	0.8358
CD4+ T cell count	0.031	0.88452	0.88452

*A cumulative measure of ISG expression was obtained by quantitative PCR of four selected ISGs (GBP1, IFI27, MX1, and OAS1). Cts were subtracted from the housekeeping gene HPRT, and a geometric mean was calculated for all four genes on a per subject basis, yielding the final measure used for comparison to PC1.

Table 3: Spearman correlations between PC1 and immunologic variables among HIV+ subjects on HAART.

Immune Variable	Spearman R_S value	Unadjusted P value	Benjamini-Hochberg adjusted P value
IL-6 pg/ml plasma	0.743	0.0015	0.036
Kyn:Trp ratio	0.727	0.0032	0.038
Soluble TNF-Receptor II pg/ml plasma	0.654	0.0082	0.066
D-Dimer ng/ml plasma	0.607	0.0164	0.095
Gut IDO1 relative mRNA	-0.559	0.0197	0.095
IP-10 pg/ml	0.561	0.0297	0.119
%CD38+DR+ amongst gut CD4+ T cells	0.556	0.0389	0.134
HIV RNA copies/1e6 cells	-0.491	0.0534	0.16
Zonulin ng/ml plasma	0.486	0.0664	0.177
%CD38+DR+ in blood memory CD8+ T cells (all CD8+ after exclusion of CD27+CD45RO-)	0.489	0.076	0.182
%IL-17+ amongst gut CD8+ T cells	-0.382	0.1598	0.349
%CD38+DR+ amongst gut CD8+ T cells	0.337	0.2392	0.478
sCD14 ng/ml plasma	0.213	0.4643	0.812
%CD38+DR+ amongst blood memory CD4+ T cells (all CD4+ after exclusion of CD27+CD45RO-)	0.209	0.4738	0.812
HIV DNA copies/1e6 cells	-0.186	0.5075	0.812
%Treg in gut (CD25+FoxP3+) of total CD4+ T cells	0.123	0.6273	0.889
IFAB pg/ml plasma	-0.136	0.6296	0.889
Self-reported CD4+ T cell nadir	0.088	0.7293	0.966
CD4 T+ cell density in gut	0.077	0.8122	0.966
Cumulative measure of interferon-stimulated gene (ISG) expression*	-0.053	0.8456	0.966
%IL-17+ amongst gut CD4+ T cells	0.041	0.8797	0.966
%IL-22+ amongst gut CD8+ T cells	-0.039	0.8894	0.966
%IL-22+ amongst gut CD4+ T cells	-0.021	0.9397	0.966
CD4+ T cell count	0.011	0.9665	0.966

*****A cumulative measure of ISG expression was obtained by quantitative PCR of four selected ISGs (GBP1, IFI27, MX1, and OAS1). Cts were subtracted from the housekeeping gene HPRT,

and a geometric mean was calculated for all four genes on a per subject basis, yielding the final measure used for comparison to PC1.

Table 4: Comparisons of PC1 marker of dysbiosis to gut inflammatory gene expression in whole rectosigmoid biopsy specimens.

Variable	Spearman R_S value	Unadjusted P value
TNFa	0.35	0.27
CYBB	-0.28	0.29
IFNa2	-0.2	0.45
IFNb	0.15	0.57
IFNg	-0.15	0.59
Arginase 1	-0.14	0.62
NOS2	0.08	0.76

CHAPTER III: GUT-RESIDENT *LACTOBACILLUS* ABUNDANCE ASSOCIATES WITH IDO1 INHIBITION AND TH17 LOSS IN SIV-INFECTED RHESUS MACAQUES

Abstract

Gut microbes have been shown to modulate mucosal barrier-promoting Th17 cells in mammals. A salient feature of HIV/SIV immunopathogenesis is the loss of Th17 cells, which has been linked to upregulation of activity of the immunomodulatory enzyme indoleamine 2,3-dioxygenase (IDO1). The role of gut microbes in this system remains unknown, while the SIV-infected rhesus macaque provides a well-described model for HIV-associated inflammation and mucosal immune disruption. We observed a specific depletion of gut-resident *Lactobacillus* in rhesus macaques during acute and chronic SIV infection. This depletion correlated with increased IDO1 activity and Th17 cell loss. While products of IDO1 were found not to inhibit *Lactobacillus* growth in vitro, macaques supplemented with a *Lactobacillus*-containing probiotic exhibited decreased IDO1 activity during chronic SIV infection. We propose that *Lactobacillus* species inhibit mammalian IDO1 and thus may help preserve Th17 cells during pathogenic SIV infection, providing support for *Lactobacillus* species as modulators of mucosal immune homeostasis.

Introduction

Untreated HIV infection is characterized by chronic immune activation, which is postulated to be spurred by decreased mucosal barrier function and subsequent microbial translocation. The gut mucosa of HIV-infected individuals exhibits a relative depletion of Th17 cells, a cell type that participates in containment of gut luminal bacteria and supports barrier

integrity via secretion of IL-17 and IL-22, cytokines which act on epithelial cells to promote production of antimicrobial peptides and mucins. Concomitant with Th17 cell depletion in HIV, which also occurs in non-human primate models using the closely related simian immunodeficiency virus (SIV), has been observed a simultaneous upregulation of the immunomodulatory enzyme indoleamine 2,3-dioxygenase (IDO1){Favre 2010}. IDO1 is an interferon-stimulated gene induced primarily in macrophages and plasmacytoid dendritic cells by type I and II interferons, cytokines that are highly abundant in circulation during untreated HIV/SIV infection{Hosmalin 2006}. IDO1 catalyzes the conversion of tryptophan to kynurenine and other downstream catabolites, which bind the aryl hydrocarbon receptor and have been shown to inhibit Th17 differentiation{Desvignes 2009; Romani 2008}. IDO1 activity, as measured by the ratio of kynurenine to its parent compound tryptophan (Kyn:Trp) in blood plasma, accordingly correlates strongly with the depletion of mucosal-barrier promoting Th17 cells in the gut mucosa of HIV-infected subjects{Favre 2010; Vujkovic-Cvijin 2013} and it is postulated that this disruption of mucosal immunity accounts for associations between Kyn:Trp and microbial translocation{Chen 2014} as well as with peripheral immune activation{Jenabian 2013; Jenabian 2015} and mortality{Byakwaga 2014} in HIV-infected cohorts. These observations support a role for IDO1 activity in the pathophysiology of HIV/AIDS, and highlight the need for a greater understanding of the precise factors regulating the IDO1 pathway, be they extrinsic or host-mediated, and the depletion of Th17 cells in HIV/SIV infection.

Human-associated bacteria of the gastrointestinal tract vastly outnumber human cells within each individual, and evidence suggests that complex bidirectional relationships have evolved between gut-resident microbes and the mammalian mucosal immune system, likely to ensure maintenance of homeostasis wherein the health of both host and microbe is enabled.

Indeed, Th17 cells and Treg are poised at mucosal surfaces where they interact with gut-resident microbes{Lathrop 2011} and regulate their containment in the gut lumen{Raffatellu 2008}, while conversely these cells themselves are dynamically modulated by the presence of specific gut microbes{Ivanov 2009; Atarashi 2011a}. Uncovering pathways by which gut-resident bacteria modulate mucosal T cell differentiation is an active area of research. Putative microbial drivers of mucosal immune homeostasis in HIV/SIV, specifically those that may inhibit the immunomodulatory IDO1 axis or otherwise preserve the Th17 cell pool, are yet to be identified.

SIV-infected macaques provide a well-described non-human primate model that recapitulates major features of HIV-associated chronic inflammation and mucosal immune disruption, including innate and adaptive immune activation, CD4 T cell depletion in the peripheral blood and gut-associated lymphoid tissue, and a loss of mucosal barrier-promoting Th17 cells{Favre 2009}. Several recent macaque studies have shown that members of the bacterial genus *Lactobacillus* can modulate the Th17 axis, though mechanisms for this process remain poorly understood. Administration of VSL #3, a probiotic cocktail of which 4 of 7 species are *Lactobacillus* spp., to chronically SIV-infected pigtail macaques resulted in greater Th17 polyfunctionality, as well as a partial resolution of several inflammatory measures (e.g. lower fibrosis, reduced CD4 T cell activation, and increased CD4 T cell recovery){Klatt 2013; Ortiz 2014}. Ligated ileal loops of acutely SIV-infected rhesus macaques inoculated with *Lactobacillus plantarum* displayed increased IL-17 expression and epithelial tight junction formation{Hirao 2014}. Though studies of the macaque gut microbiome have shown few ecological differences in gut bacterial communities between SIV-uninfected and chronically SIV-infected macaques{Klatt 2013; Handley 2012; McKenna 2008}, which is in contrast to the pronounced microbiome perturbations observed in human cohorts{Lozupone 2013; Dillon 2014;

Mutlu 2014; Dinh 2015}, host-microbiome dynamics during the acute phase of lentiviral, infection when IDO1 is maximally induced by interferons, remain unexplored. We found that a transient perturbation of the fecal microbiome of SIV-infected macaques occurs during acute SIV infection, which coincides with maximal IDO1 activity. This perturbation was characterized chiefly by a depletion of *Lactobacillus* genus members, and we provide evidence that gut-resident *Lactobacillus* spp. modulate primate IDO1 by inhibiting production of immunomodulatory kynurenine compounds that affect the Th17 cell arm of the mammalian mucosal immune system.

Materials and Methods

Study cohorts and sample collection. Six female rhesus macaques (RM, *Macaca mulatta*) were infected with SIVmac251-2010 on April 23, 2013 (Cohort A) and eight more females were infected on July 16, 2013 (Cohort B) at Advanced BioScience Laboratories, Inc (IACUC Animal Use Protocol #552). Viral inocula were diluted 1:10 in RPMI and then 1:20 in RPMI with 1ml administered i.v. Animals were fed High Protein Monkey Diet 5045 (LabDiet), with occasional enrichment using PRIMA-Treats or Monkey Dough (BioServ). Animals were singly housed on stool collection days, and stool was collected either from the pan (if fresh) or digitally from the rectum prior to freezing at -80C. Stool from Cohort A was preserved in RNAlater (Life Technologies) prior to cryopreservation while stool from Cohort B was immediately cryopreserved. Peripheral blood plasma and stool were collected and stored at -80C at the following days relative to the dates of infection for all animals: -7, 3, 7, 10, 14, and 56.

Peripheral blood mononuclear cells were obtained and cryopreserved at -80C at days -7, 14, and 56 post-infection.

Fecal sample processing, 16S rRNA amplification, and 16S rRNA sequencing. Fecal sample DNA was extracted using the QIAamp DNA Stool Mini Kit (QIAGEN) as per manufacturer instructions, in conjunction with Tallprep Lysing Matrix E tissue disruption tubes and the FastPrep-24 Tissue and Cell Homogenizer (MP Biomedical). DNA was amplified using the 515F/806R primer pair targeting the V4 region of the 16S rRNA gene, with barcodes affixed to the 806R primer and Illumina adapter oligonucleotides on both primers. Amplification was performed using Hot Start ExTaq (Takara) for 30 cycles (98 °C for 20 seconds, 50 °C for 30 seconds, 72 °C for 45 seconds). Amplicon concentrations were quantified by comparison to Low Mass DNA Ladder (Life Technologies) on an agarose gel, pooled at equimolar concentrations, and purified by agarose gel extraction using the QIAquick Gel Extraction Kit (QIAGEN). Pooled amplicon libraries were quantified using the Qubit 2.0 Fluorometer (Life Technologies) and sequenced on a MiSeq Genome Sequencer (Illumina) using the V3 Reagent Kit (Illumina).

16S rRNA sequence analysis and statistical methods. Sequencing reads were split by barcodes using a Phred q score threshold cutoff of 30 and collapsed into operational taxonomic units (OTUs) via an open-reference OTU picking scheme as implemented in the software package QIIME{Caporaso 2010} using a 97% sequence identity threshold cutoff. OTUs with read abundance below 0.001% of total reads were removed. Beta diversity and alpha diversity calculations were performed in QIIME, while the R package ‘adonis’ was used to calculate permANOVA statistics. For identifying discriminating OTUs after SIV infection, ten rarefaction

tables were generated at 34,000 reads each, and custom scripts were designed to perform serial paired Wilcoxon signed-rank tests (R package 'exactRankTests') on all OTUs at each time point post-infection compared to each respective animal's pre-infection time point sample. Statistics were collated between the resultant ten output tables by calculating median P values and mean $\log_{10}$ fold changes for each OTU. A linear mixed effects modeling implementation in R (package 'lme4') fitted by maximum likelihood was used to compare abundances of all bacterial families to peripheral blood IDO1 activity and Th17 cell abundances. Visualizations of data were generated using Prism 6 (GraphPad Software, Inc.) and the R package 'ggplot2'.

VSL#3 trial characteristics and analysis. Two separate studies using pigtail macaques (*Macaca nemestrina*) infected with 3,000 TCID of SIV_{mac239} i.v. were combined for analysis of the effects of VSL#3 treatment on IDO1 activity. One cohort, initiated on 6/25/2010{Klatt 2013}, consisted of 4 controls animals and 7 VSL#3-treated animals, with VSL#3-treated animals beginning daily administration of probiotic beginning at 100 days post-infection and both groups beginning antiretroviral therapy at day 160 post-infection (0–30 mg/kg PMPA, 30 mg/kg FTC once daily, s.c., and 120 mg L812, 50 mg L564 twice daily, oral), as described previously{Klatt 2013}. The second cohort, initiated on 10/9/12, consisted of 5 control animals and 6 VSL#3/IL-21-treated animals, with antiretroviral therapy commencing on the day 98 p.i. for both groups and the VSL#3/IL-21 group receiving daily VSL#3 beginning on day 98 p.i. as well as two courses of rMamu IL-21-IgFc (50 ug/kg at 1 mg/mL in H_2O s.c., 5 once-weekly doses) initiated at days 98 and 333 p.i. as described{Ortiz 2014}. One animal in the control group for the second cohort was sacrificed before the study end due to severe diarrhea. Linear mixed effects modeling fitted by maximum likelihood (R package 'lme4') was used to test

significance of differences between VSL#3-treated and untreated animals, incorporating stratification by study cohort.

***Lactobacillus* isolation and *in vitro* experimentation.** *Lactobacillus* primary isolates were isolated by plating frozen stool from animals in Cohort B onto MRS selection agar (Beckton Dickinson). 16S sequencing to verify taxonomic identity of resultant isolates was performed using the 27f/1492r 16S rRNA primer pair for full-length 16S PCR amplification. Taxonomy was assigned to full-length 16S sequence reads using NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Growth assessment with addition of exogenous kynurenine was performed in MRS liquid media using a spectrophotometer machine and taking OD_{600} readings at regular intervals.

Results

Transient microbiome shift occurs during acute SIV infection of rhesus macaques. Two temporally independent cohorts of six rhesus macaques (RM) each were experimentally infected with SIV_{mac251} and followed longitudinally via collections of stool and peripheral blood before SIV infection, as well five time points through acute and chronic infection. We profiled fecal bacterial microbiomes by sequencing of the V4 region of the 16S rRNA gene and collapsing reads with 97% sequence homology into operational taxonomic units (OTUs). We found a transient shift in the fecal microbiome at 2 weeks post-SIV infection that subsides by 8 weeks post-infection (**Figure 1**). This shift was characterized by an increase in beta diversity (defined as the compositional difference between two ecological communities) within each animal as compared to its own pre-infection time point, using both the Canberra beta diversity metric,

which takes into account relative abundances of bacterial taxa among both communities when assessing their compositional similarity, and the unweighted UniFrac beta diversity metric, which incorporates phylogenetic relationships of taxa present in common between communities (Figure 1A and Figure S1). This shift was characterized also by a decrease in alpha diversity (Figure 1B), a measure for community richness (the total number of unique taxa) and evenness (the relative dispersion of abundances of all taxa). A concerted shift in community composition from pre-infection to 2 weeks post-infection (p.i.) was discernible in a principal coordinates ordination (permANOVA P=0.039, Figure 1C), though a concerted shift was not evident when comparing pre-infection communities to 8 weeks p.i. (permANOVA P = 0.194). Intriguingly, when comparing RM fecal microbiomes at each time point to fecal microbiomes of chronically HIV-infected humans, there was no increase in similarity of the RM fecal microbiome to that of chronically HIV-infected humans after SIV infection (Figure S1B), indicating that SIV infection in this non-human primate model does not cause a microbiome shift that increases in resemblance to the HIV-associated gut microbiome.

Acute SIV-associated microbiome shift is characterized by depletion of *Lactobacillus spp.*

To investigate shifts in microbiome composition as a function of SIV infection over the acute and chronic stages, with particular interest in the 2 week p.i. community, we compared OTU abundances within each animal between its pre-infection time point and each time point after infection. Notable differences were observed at 2 weeks p.i., wherein a depletion of *Lactobacillus* spp. occurred (**Figure 2A**). This depletion was observed in both temporally independent cohorts of animals and persisted into the chronic phase of infection 8 weeks p.i. (**Figure 2B**). Similarly to the timing of the peak of *Lactobacillus* depletion at 2 weeks p.i., IDO1

activity was heightened at 2 weeks p.i. and did not resolve to baseline, pre-infection levels during the chronic phase of infection (**Figure 2C**), a pattern similar in kinetics and relative scale to the acute depletion of *Lactobacillus* with a persistent but intermediate depletion extending into the chronic phase of SIV infection. Interestingly, in the non-pathogenic SIV infection model of the African green monkey - a species of non-human primate which, when experimentally infected with SIV, does not exhibit SIV-associated chronic inflammation and does not develop simian AIDS{Jacquelin 2009} - IDO1 activity does resolve to pre-infection levels during chronic infection and coincides with an increased abundance of fecal *Lactobacillus* at that time (Figure S2), rather than a decrease as was observed in pathogenic SIV infection of RM.

***Lactobacillus* depletion associates with increased IDO1 activity and loss of Th17 cell abundance.** To identify candidate bacterial taxa that may modulate IDO1 activity, we compared abundances of all bacterial genera detected in all animals at all time points to respective plasma Kyn:Trp ratios using a linear mixed effects models approach, which accounts for inter-individual variation in longitudinal time-series data. We found that fecal abundance of the *Lactobacillus* genus correlated significantly and inversely with IDO1 activity (P=0.00151, Figure 3A), and was the top inverse correlate of all genera detected in our dataset (total=86, Table 1). Furthermore, IDO1 activity was correlated with peripheral blood Th17 cell loss across acute and chronic SIV infection (P=0.0000052, Figure 3B). Accordingly, abundance of the *Lactobacillus* genus exhibited a significant positive correlation with Th17 cell abundances (P=0.00187, Figure 3C), and was among the top correlates of bacterial genera with Th17 cell abundances in our dataset (Table 2). To address the hypothesis that other fecal bacteria that correlated positively with IDO1 activity may themselves produce kynurenine compounds, we utilized the UniProt consortium of

genome sequencing data to identify bacterial families that encode enzymes that participate in the kynurenine pathway of tryptophan catabolism. We found families that encode such enzymes do not preferentially correlate with Kyn:Trp ratios (Figure S3), suggesting that direct bacterial tryptophan catabolism may not participate in correlations between bacterial abundances and Kyn:Trp ratios in the SIV-infected RM.

Supplementation of SIV-infected macaques with *Lactobacillus*-containing probiotic reduces peripheral blood IDO1 activity. To further investigate cause and effect relationships responsible for the correlation between *Lactobacillus* genus abundance and the immunomodulatory IDO1 pathway, we asked whether supplementation of SIV-infected macaques with *Lactobacillus spp.* may modulate IDO1 activity *in vivo*. Plasma was analyzed from two cohorts of SIV-infected pigtail macaques that were treated with VSL#3, a probiotic mixture composed of 7 bacterial strains of which 4 are *Lactobacillus spp.*[VSL ref] after twelve weeks of SIV infection{Klatt 2013; Ortiz 2014}. While some animals were given VSL#3, all animals received antiretroviral therapy (ART) as detailed in Experimental Procedures. At days 325-373 p.i., VSL#3-treated animals exhibited significantly decreased plasma concentrations of kynurenine, the effector molecule produced by IDO1, as compared to untreated control animals (Figure 5A) using linear mixed effects models to control for between-cohort differences. A trend was additionally observed for lower ratios of Kyn:Trp (Figure 5B), a marker of flux through the IDO1 catabolic pathway, and trends were additionally observed for differences in both kynurenine concentration and Kyn:Trp ratios at the 325-372 days p.i. time point as compared to the same measurements in pre-infection plasma. While one of the two cohorts treated with VSL#3 received an additional agent (the Th17-promoting IL-21 cytokine, which has no reported

relationships to IDO1 activity), there were no differences between Kyn:Trp ratios or kynurenine concentrations of the VSL#3-treated arms of these cohorts (Figure S4A). Furthermore, to understand whether the inverse correlation between IDO1 activity and *Lactobacillus* abundance found in our rhesus cohorts was due to an antagonistic effects of kynurenine on *Lactobacillus* growth, we isolated *Lactobacillus spp*. from the stool samples obtained from the experimental rhesus study animals. Resulting primary isolates were dominated by *Lactobacillus animalis* and *Lactobacillus reuteri* (Figure S4B). Physiologic and supra-physiologic concentrations of kynurenine were added to growth cultures of *L. animalis* and *L. reuteri* isolates and growth was measured by light absorption. No effect of kynurenine on growth of these primary isolates over time was observed (Figure S4C).

Discussion

We observed a specific depletion of gut-resident *Lactobacillius* in rhesus macaques, seen most dramatically during acute SIV infection and persisting into the chronic phase. This depletion correlated with increased IDO1 activity and Th17 cell loss across time during SIV infection. Products of the IDO1 pathway were found to have no effect on *Lactobacillus* growth *in vitro*, while pigtail macaques supplemented with a *Lactobacillus*-containing probiotic exhibited decreased IDO1 activity during chronic SIV infection. We propose that gut-resident *Lactobacillus* species inhibit primate IDO1 and thus may help preserve the Th17 cell compartment during pathogenic SIV infection, providing support for *Lactobacillus* species as modulators of mucosal immune homeostasis.

Furthermore, we profiled the fecal microbiomes of several African green monkeys at different stages of SIV infection. African green monkeys exhibit a non-pathogenic course of SIV

infection characterized by a resolution of IDO1 activity during the chronic phase and a preservation of the Th17 cell pool, which is in contrast to SIV-infected macaques. These animals did not exhibit a reduction in fecal *Lactobacillus* abundance but rather an increase during chronic infection, consistent with a role for *Lactobacillus* in inhibiting the IDO1 pathway. However, these observations are from a small, cross-sectional sample size (n=4-6 for pre-infection, acute, and chronic SIV-infected animals), and so should be cautiously interpreted. Additionally, power to detect differences between VSL#3-treated and untreated SIV-infected animal groups was enhanced by the combination of two temporally separated VSL#3 supplementation experiments. One of these trials included administration of the cytokine IL-21 in addition to VSL#3. Though to our knowledge no precedence exists for IL-21 directly affecting IDO1 activity, this possibility cannot be ruled out. Despite this, plasma kynurenine concentrations and Kyn:Trp ratios did not differ significantly between the cohort that received IL-21 and that which didn't, suggesting that the interpretation of the effect being largely due to the VSL#3 probiotic may be appropriate. Furthermore, while *Lactobacillus* has been previously associated with lower peripheral blood kynurenine in rats{Valladares 2013}, mono-administration of a *Bifidobacterium sp.* contained in VSL#3 did not affect peripheral blood kynurenine levels{Bercik 2010}, while products of the other genus composing VSL#3, *Streptococcus*, have been shown to induce rather than inhibit IDO1 activity{Murr 2001}.

While shifts in the gut bacterial microbiome of chronically HIV-infected humans have been well-described and with relative consistency in findings, several studies including that presented herein suggest that significant shifts in the gut microbiome do not occur in macaques in the chronic phase of SIV infection as compared to uninfected macaques{Klatt 2013; Handley 2012; McKenna 2008}. Though we observed a depletion of *Lactobacillus* during acute and

chronic SIV infection, numerous studies of HIV-associated microbiome shifts have reported no changes in the abundance of gut *Lactobacillus* during chronic HIV infection in humans. As HIV-associated microbiome shifts may be relevant to disease pathogenesis, knowledge of this as a potential caveat to the macaque model of SIV infection may inform interpretation of results from this model. Regardless, however, of its capacity to recapitulate HIV-associated microbiome shifts, the macaque model offers various notable advantages toward probing host-microbiome immunologic relationships, including genomic proximity to humans and a higher capacity for tissue sampling across time.

Lactobacillus-supplemented yogurt has been investigated as a possible therapeutic agent in African cohorts of treated and untreated HIV infected adults with results showing improvements in CD4 count, though mechanisms of this improvement remain poorly understood{Anukam 2008; Irvine 2010}. *Lactobacilli* may inhibit IDO1 of humans as well as non-human primates as suggested in this study, and this effect may have the capacity to mitigate loss of gut barrier-promoting Th17 cells. This could allow for maintenance of mucosal integrity, reduction in microbial translocation, and an overall diminution of HIV-associated inflammation. Clinical trials using over-the-counter probiotics such as VSL#3 and others containing *Lactobacillus spp.* may have efficacy in HIV-infected humans, and perhaps attention should be placed on IDO1 activity as an outcome.

Acknowledgements & Footnotes

Research reported in this publication was supported by CHRP D13-SF-388 (to IVC), NSF 1144247 (to IVC), and NIH U19 AI96109 (to JMM). IVC, CS, and SC performed DNA extraction on fecal samples. IVC performed 16S library preparation, sequencing, and analysis.

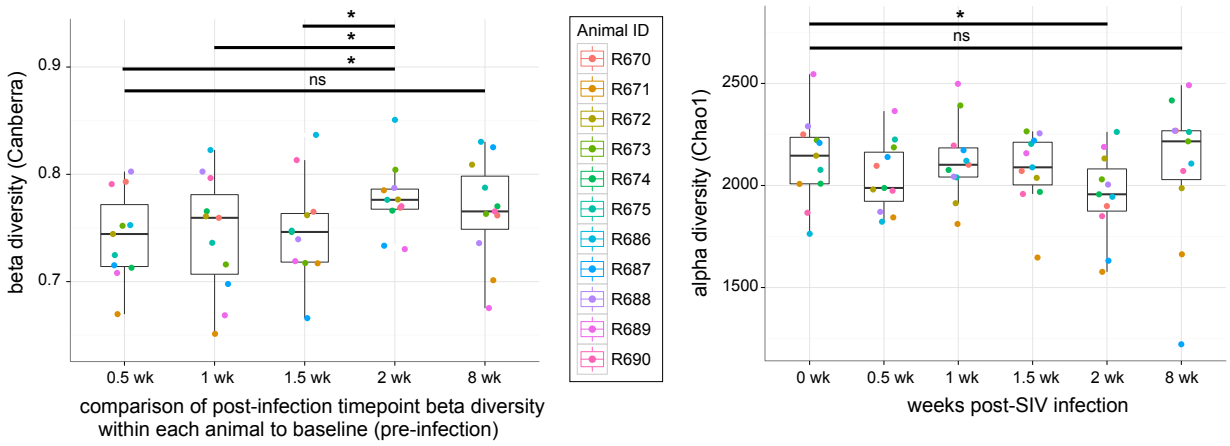
LS performed flow cytometry measurements and analysis. IVC, RMD, and LS designed the rhesus macaque studies. IVC and SC isolated *Lactobacillus* primary isolates. AP performed *in vitro* isolate growth experiments. YH measured kynurenine and tryptophan in macaque plasma. CA and YP designed African green monkey experiments and provided samples. NK and JB designed the VSL#3 trial in pigtail macaques and provided samples. IVC, SVL, and JMM wrote the manuscript. SVL and JMM oversaw all experiments and analyses.

Figures

Figure 1: Ecological dynamics of microbiome composition across SIV infection in rhesus macaques.

A) Difference between baseline (pre-infection) and post-SIV infection (p.i.) fecal microbiomes significantly increases at 2 weeks post-infection (p.i.) and returns by 8 weeks p.i. (Wilcoxon matched-pairs signed rank test *P < 0.05).

B) Alpha diversity (richness and evenness) of the rhesus macaque fecal microbiome drops at 2 weeks p.i. (Wilcoxon matched-pairs signed rank test *P < 0.05).



C) Principal Coordinates Analysis (PCoA) plot of fecal microbiome shift from pre-infection to 2 weeks p.i. (P = 0.039 by permANOVA, stratified by individual, Canberra beta diversity metric). No such shift is evident at 8 weeks p.i. (P = 0.194 by permANOVA, Canberra)

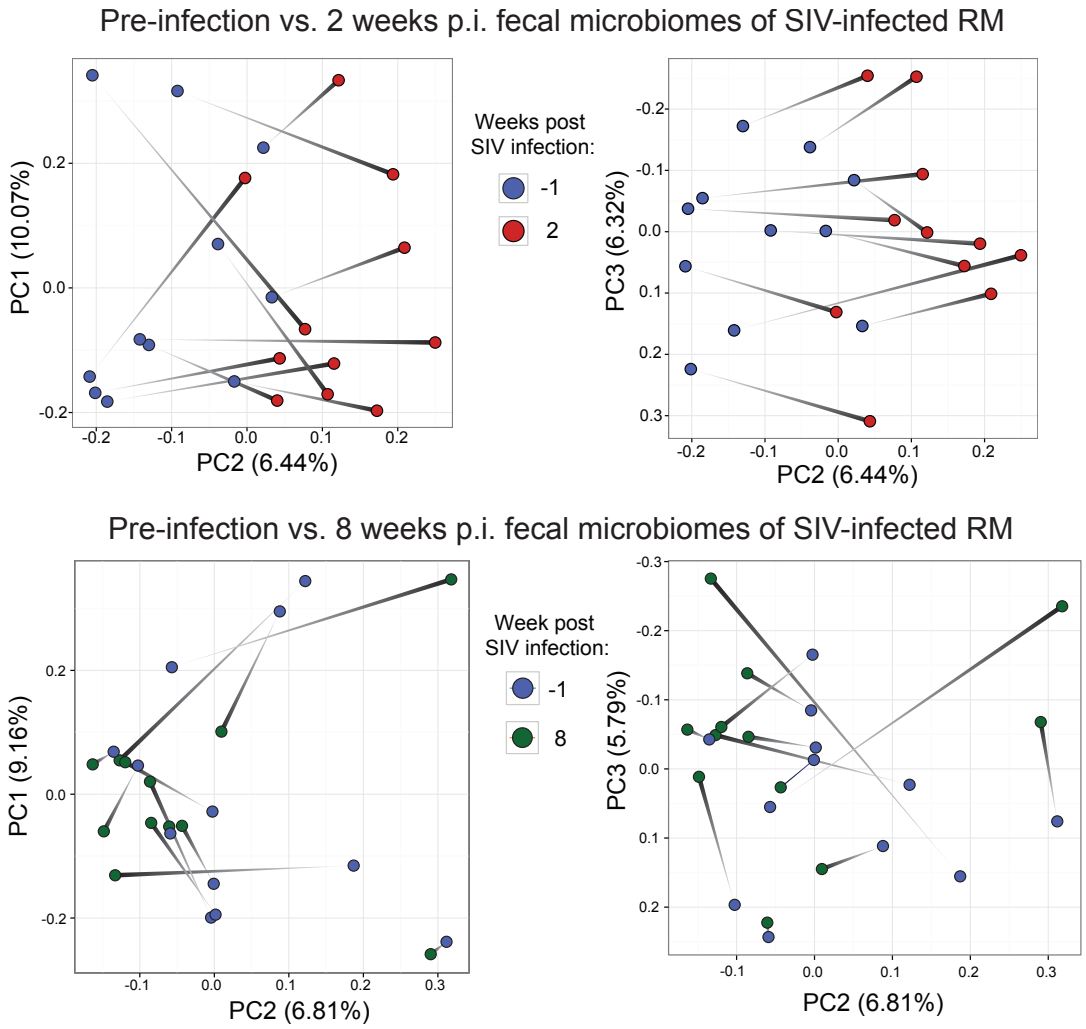
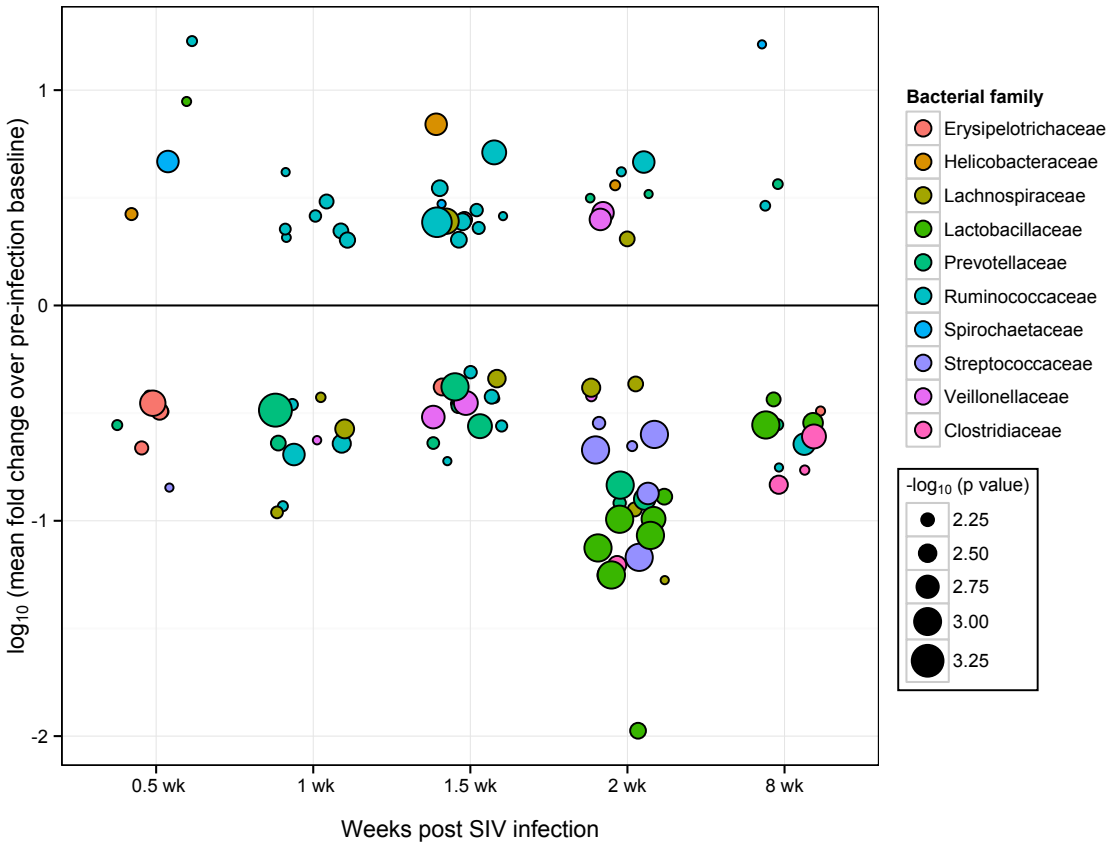
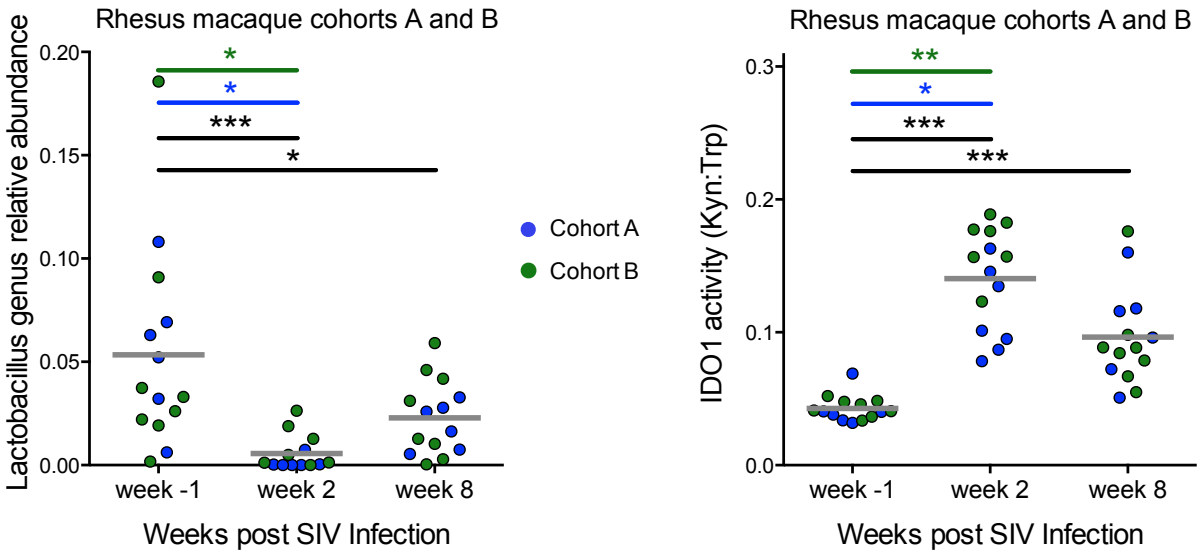


Figure 2: Taxonomic profile of microbiome shifts across time in SIV-infected rhesus macaques.

A) Shown are OTUs that differ significantly between pre-infection time point and each post-infection time point within each animal, with Wilcoxon matched-pairs signed rank test $P < 0.01$ and mean read abundance fold change > 2. Point colors indicate bacterial family to which each OTU belongs, and relative sizes of points are proportional to the inverse log of Wilcoxon P values.



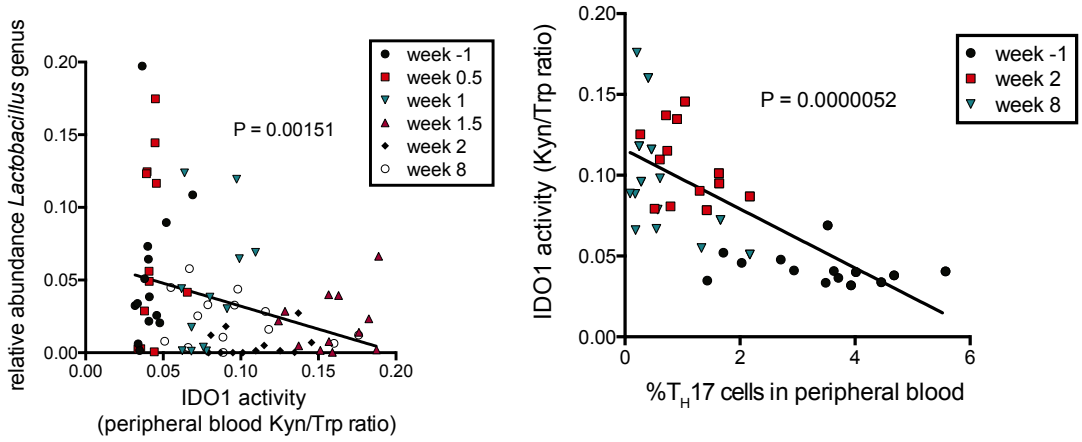
B) *Lactobacillus* genus is depleted at 2 weeks p.i. and this depletion persists into chronic phase of SIV infection at 8 weeks p.i. as compared to pre-infection. Wilcoxon matched-pairs signed rank test was used for each comparison.



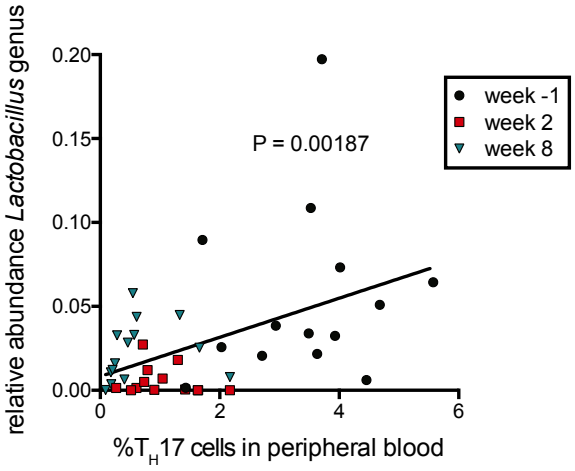
C) Rhesus macaques exhibit elevated IDO1 activity that persists into the chronic phase of infection. Wilcoxon matched-pairs signed rank test, * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$ for RM cohorts A and B considered together (black error bars).

Figure 3: IDO1 activity and Th17 cell dynamics associate with loss of gut-resident *Lactobacillus*.

A) IDO1 activity (Kyn:Trp) correlates strongly and significantly with Th17 cell abundances in peripheral blood of RM. Linear mixed effects (LME) modeling was used to allow for inclusion of longitudinal data from each animal and test for correlations between all bacterial families detected and IDO1 activity (peripheral blood Kyn:Trp ratios). Linear regression lines shown in black.



B) Abundance of *Lactobacillus* genus correlates strongly with IDO1 activity as measured by Kyn:Trp ratios in peripheral blood, using LME modeling. (top inverse correlate of all taxa deteted, Table 2)



C) *Lactobacillus* abundance correlates with Th17 cell abundance in peripheral blood. As with B), LME modeling was used to test for correlations between all bacterial families detected and peripheral blood Th17 cell abundance as assessed by flow cytometry gating %IL-17+ of total CD4+ cells.

Figure 4: Effects of supplementation of SIV-infected pigtail macaques with *Lactobacillus*-containing probiotic on markers of IDO1 activity.

A) Two cohorts of pigtail macaques (PTMs) given daily VSL#3 exhibit a decrease in peripheral blood kynurenine concentrations compared to untreated control PTMs at days 325-372 p.i., P = 0.046, by LME.

B) Trends toward signifiance in Kyn:Trp ratios are seen in PTMs treated with VSL#3 at at days 325-372 p.i. (P= 0.091).

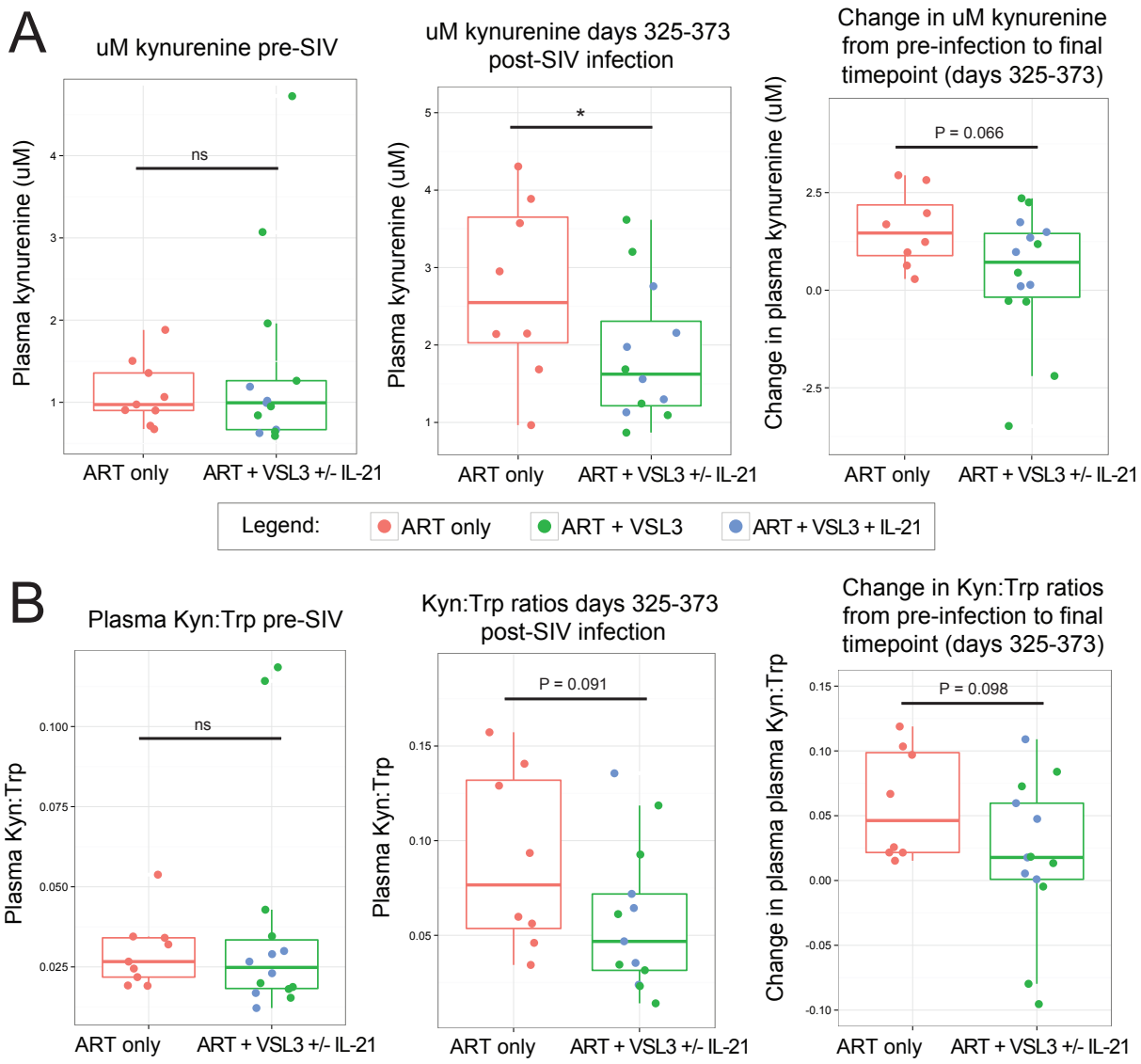


Table 1: Correlations between IDO1 activity (Kyn:Trp ratios) and fecal bacterial genus abundances.

Rank in strength of correlation (out of 86 total)	Taxonomic information for OTU	LME T value	LME P value
1	k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f___g__	3.797278	0.00028989
2	k__Archaea_p__Euryarchaeota_c__Methanobacteria_o__Methanobacteriales_f__Methanobacteriaceae_g__Methanobrevibacter	3.635978	0.00049473
3	k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Dehalobacteriaceae_g__	3.381543	0.00112993
4	k__Bacteria_p__Firmicutes_c__Bacilli_o__Bacillales_f__Planococcaceae_g__	3.317599	0.00134191
5	k__Bacteria_p__Firmicutes_c__Bacilli_o__Lactobacillales_f__Lactobacillaceae_g__Lactobacillus	-3.282863	0.00151547

Table 2: Correlations between peripheral blood Th17 cell abundance and fecal bacterial genus abundances.

Rank in strength of correlation (out of 86 total)	Taxonomic information for OTU	LME T value	LME P value
1	k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Clostridiaceae_g__	3.787026	0.00051473
2	k__Bacteria_p__Firmicutes_c__Bacilli_o__Lactobacillales_f__Lactobacillaceae_g__Lactobacillus	3.365279	0.00187481
3	k__Bacteria_p__Firmicutes_c__Clostridia_o__Clostridiales_f__Clostridiaceae_g__Sarcina	2.687198	0.01118711
4	k__Bacteria_p__Actinobacteria_c__Coriobacteriia_o__Coriobacteriales_f__Coriobacteriaceae_g__Slackia	2.529289	0.01662025

Figure S1A: Beta diversity dynamics using Unweighted Unifrac mirrors Canberra metric dynamics.

Wilcoxon matched-pairs signed rank test *P < 0.05, by Unweighted Unifrac distance metric.

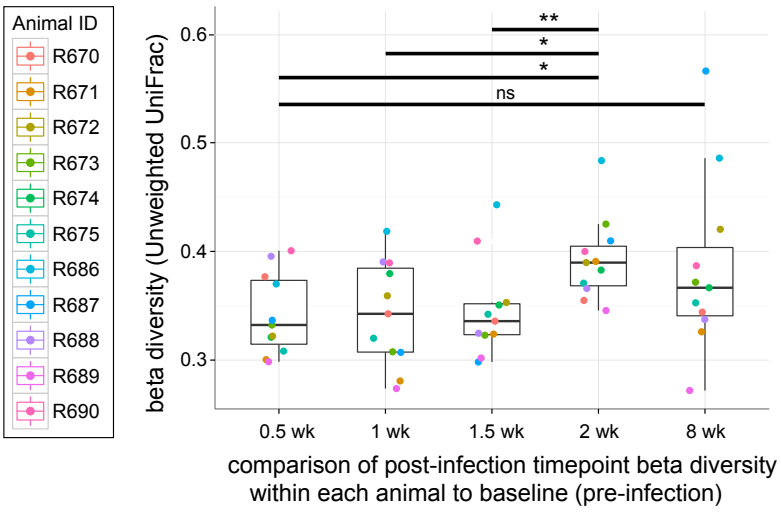


Figure S1B: Comparison of RM microbiomes to HIV-infected humans.

There is no significant change in similarity of the macaque fecal microbiome to HIV-infected human fecal microbiomes after SIV infection as compared to before infection (unweighted UniFrac metric; P=ns by one-way repeated measures ANOVA within SIV-infected cohorts)

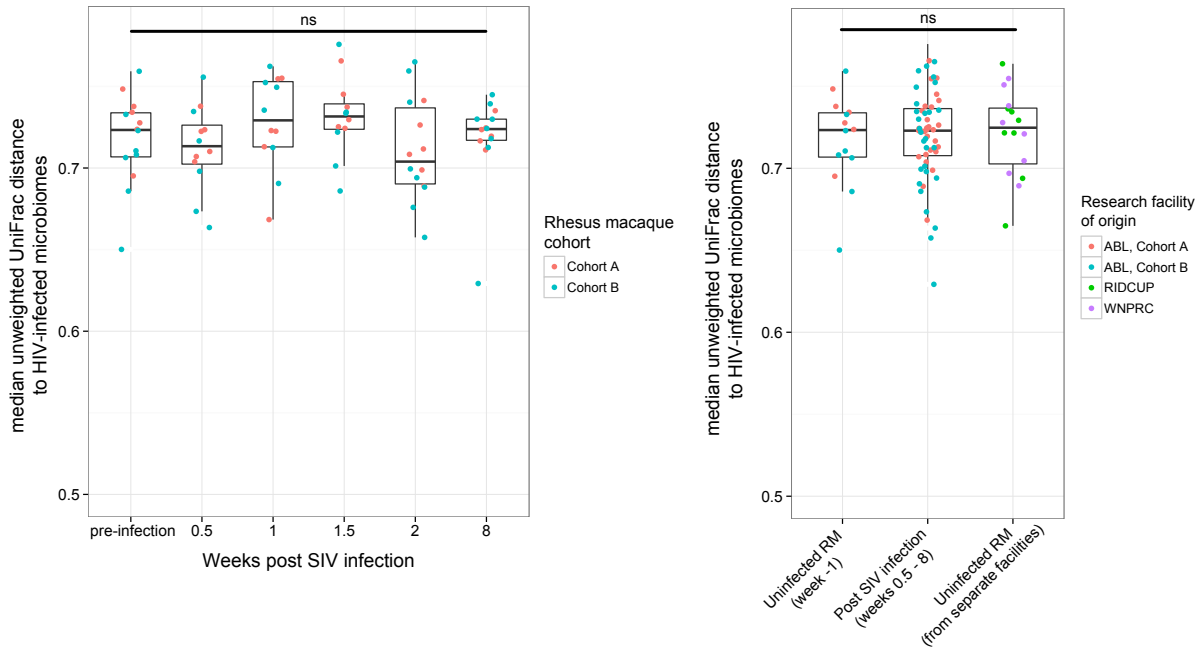


Figure S2: African green monkey fecal Lactobacillus abundance increases during chronic SIV infection, while IDO activity reverts during chronic infection back to pre-infection levels.

Cross-sectional analysis of separate animals at each time point was employed and an unpaired Student's T test was used, * $P < 0.05$.

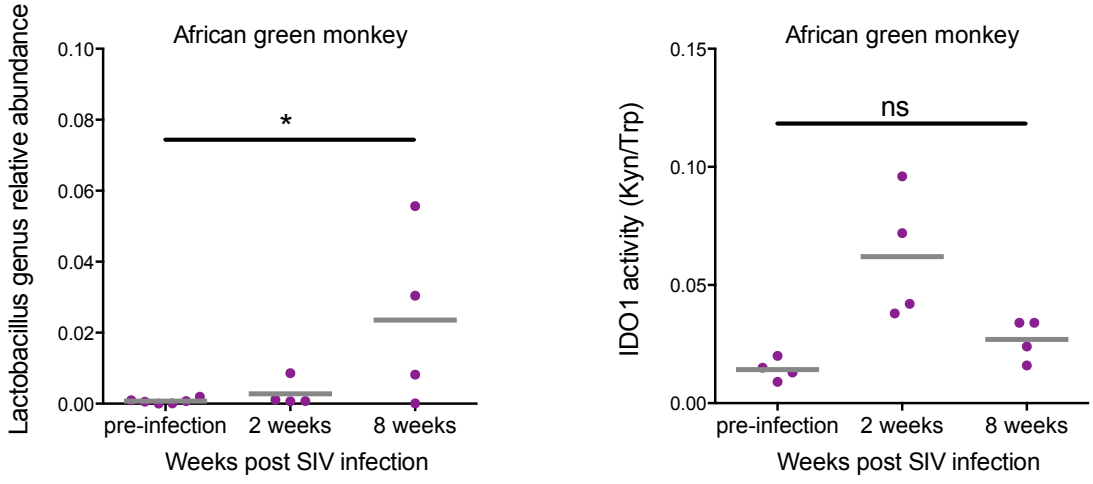


Figure S3: Bacterial families that encode kynurenine pathway enzymatic machinery do not preferentially correlate with peripheral blood K:T ratios of SIV-infected rhesus macaques.

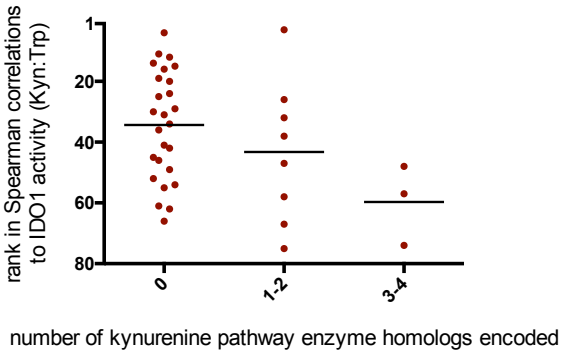


Figure S4A: No significant differences in plasma Kyn or Kyn:Trp observed between studies (with or without IL-21 supplement).

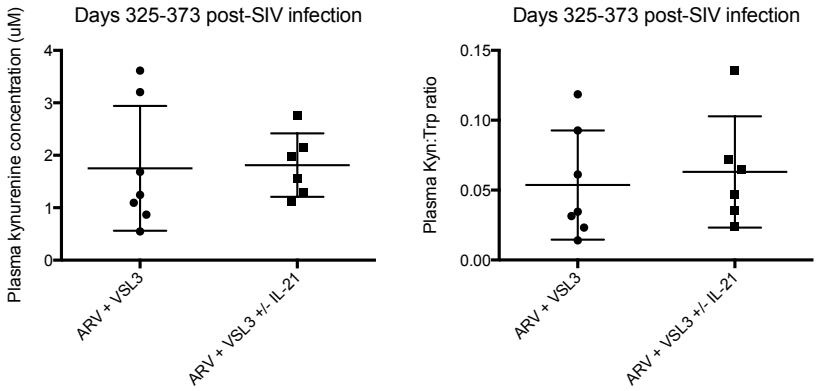


Figure S4B: Composition of Lactobacillus spp. isolated from rhesus macaque stool specimens.

Isolations were performed anaerobically on MRS selective agar.

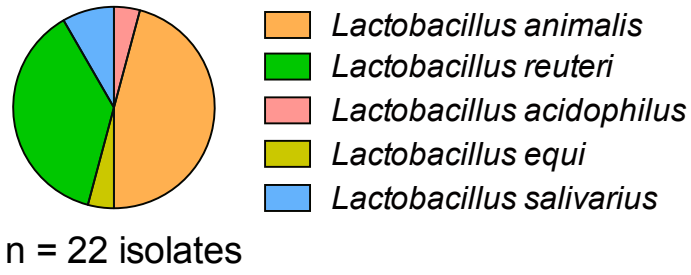
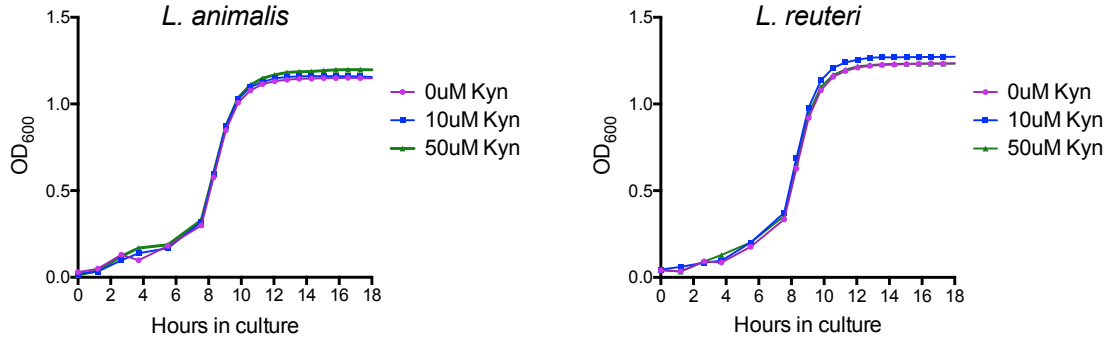


Figure S4C: Kynurenine does not impact *Lactobacillus* primary isolate growth.

Exogenous kynurenine was added to growth cultures in MRS medium and OD_{600} was observed at regular time intervals to determine growth kinetics. The same result was observed for an isolate of the next most dominant Lactobacillus species retrieved (L. reuteri). Results shown are from representative primary isolates collated from sixplicate cultures.



CHAPTER IV. CONCLUSIONS AND FUTURE DIRECTIONS

Summary

Our work focused on investigating the possible contribution of the gut microbiome to HIV-associated chronic inflammation and mucosal immune disruptions. We utilized microarray-based as well as high throughput sequencing-based microbiome profiling techniques, *in vitro* microbiological culture, flow cytometry on a variety of human and non-human primate tissues, liquid chromatography and tandem mass spectrometry, quantitative PCR, immunohistochemistry, ELISAs, and a suite of bioinformatic and biostatistical methods to demonstrate that the HIV-associated microbiome differs from that of uninfected subjects, and the extent to which it is different correlates with markers of chronic inflammation and mucosal immune disruption. We propose a mechanistic link in HIV between gut microbes and mucosal immune disruption by the direct production of immumodulatory kynurenine compounds by gut bacteria enriched in HIV infection. We also showed that the rhesus macaque fecal microbiome does not change significantly as a result of chronic SIV infection, consistent with previous reports{Klatt 2013; Handley 2012; McKenna 2008}, and differing from observations in HIV-infected humans, which do exhibit significant microbiome perturbations during chronic HIV infection. We did, however, observe an acute microbiome disruption at 2 weeks post-SIV infection characterized by a loss of *Lactobacillus* and coinciding with a rise in IDO1 activity. We found abundance of fecal *Lactobacillus* correlated strongly and inversely with IDO1 activity in these animals across time. An *in vivo* trial using a *Lactobacillus*-containing probiotic in pigtail macaques revealed lower IDO1 activity in the treated animals, consistent with the hypothesis that gut-resident *Lactobacillus* may inhibit IDO1.

Future directions

Numerous questions remain outstanding, and specifically those regarding mechanism, with regards to how gut-resident bacteria may impact host immune states during HIV infection, merit further investigation. While we demonstrated that bacteria related closely to those that correlated most strongly with Kyn:Trp ratios in our human subjects cohort could indeed produce kynurenine *in vitro*, whether they can and do so to an appreciable extent *in vivo*, where environmental stimul and selection pressures are dramatically different than in bacterial culture media, should be investigated. It is likely that experiments such as this and others geared toward understanding the effects of other members, or perhaps all members, of the HIV-associated gut microbiome may be enabled by further study in reductionist models such as the mouse.

The observation that a *Lactobacillus*-containing probiotic may impact IDO1 activity in untreated SIV infection also may merit further investigation as to its potential therapeutic applications in humans. Currently, our collaborator and co-author Jason Brenchley, is engaged in organizing a trial through the AIDS Clinical Trials Group (ACTG) to test whether such an effect may occur in HIV-infected humans supplemented with the same probiotic. We are excited to test plasma from these subjects before and after treatment in order to measure Kyn:Trp ratios and understand whether our observation in the pigtail macaque model will translate to humans. Further forays into understanding the mechanism by which *Lactobacillus spp*. might inhibit IDO1 activity, be it directly or indirectly, are absolutely merited. Again, utilization of reductionist models, such as *in vitro* culture systems using primary bacterial isolates and human IDO1-producing cells or *in vivo* models such as the mouse, will likely be invaluable toward this end. It is my hope that expansion of the tools to understand host-microbiome relationships will

enable such studies in the future, and for such studies investigating microbial control of host immunity to translate into novel avenues toward alleviating human disease.

REFERENCES

{Bibliography}

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature

6/5/15

Date