nature publishing group

Altered metabolism of gut microbiota contributes to chronic immune activation in HIV-infected individuals

JF Vázquez-Castellanos[1,2,11], S Serrano-Villar[3,11], A Latorre[1,2], A Artacho[1], ML Ferrús[1], N Madrid[3], A Vallejo[3], T Sainz[4,5], J Martínez-Botas[6,7], S Ferrando-Martínez[4,5,8], M Vera[9], F Dronda[3], M Leal[8], J Del Romero[9], S Moreno[3], V Estrada[10], MJ Gosalbes[1,2] and A Moya[1,2]

Altered interplay between gut mucosa and microbiota during treated HIV infection may possibly contribute to increased bacterial translocation and chronic immune activation, both of which are predictors of morbidity and mortality. Although a dysbiotic gut microbiota has recently been reported in HIV + individuals, the metagenome gene pool associated with HIV infection remains unknown. The aim of this study is to characterize the functional gene content of gut microbiota in HIV + patients and to define the metabolic pathways of this bacterial community, which is potentially associated with immune dysfunction. We determined systemic markers of innate and adaptive immunity in a cohort of HIV-infected individuals on successful antiretroviral therapy without comorbidities and in healthy non-HIV-infected subjects. Metagenome sequencing revealed an altered functional profile, with enrichment of the genes involved in various pathogenic processes, lipopolysaccharide biosynthesis, bacterial translocation, and other inflammatory pathways. In contrast, we observed depletion of genes involved in amino acid metabolism and energy processes. Bayesian networks showed significant interactions between the bacterial community, their altered metabolic pathways, and systemic markers of immune dysfunction. This study reveals altered metabolic activity of microbiota and provides novel insight into the potential host–microbiota interactions driving the sustained inflammatory state in successfully treated HIV-infected patients.

INTRODUCTION

HIV infection is characterized by profound disruption of gut-associated lymphoid tissue and a chronic inflammatory state that persists even after restoration of circulating CD4 + T cell counts under successful antiretroviral therapy (ART).[1–3] Although HIV-infected adults (HIV +) with access to modern ART regimens will presumably be able to suppress HIV replication indefinitely, the profound CD4 + T cell depletion in gut-associated lymphoid tissue is incompletely reversed by long-term ART and microbial translocation continues long after peripheral CD4 + T cell restoration.[4–6] This residual activation of both the innate and the adaptive immune systems during treated HIV infection is also associated with markers of inflammation and coagulation, and decreased thymic output. Furthermore, it is an independent predictor of morbidity and mortality.[7–9]

The huge amount of metagenomic data generated in recent years has revolutionized the field of the human microbiome,

[1]Unidad Mixta de Investigación en Genómica y Salud de la Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO-Salud Pública) y el Instituto Cavanilles de Biodiversitad y Biología Evolutiva (Universitat de València), Valencia, Spain. [2]CIBER on Epidemiology and Public Health (CIBERESP), Madrid, Spain. [3]Department of Infectious Diseases, University Hospital Ramón y Cajal-IRYCIS, Madrid, Spain. [4]Laboratory of Molecular Immune Biology, University Hospital Gregorio Marañón and Gregorio Marañón Research Institute, Madrid, Spain. [5]CIBER on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain. [6]Department of Biochemistry, University Hospital Ramón y Cajal-IRYCIS, Madrid, Spain. [7]CIBER on Obesity and Nutrition Pathophysiology (CIBEROBN), Madrid, Spain. [8]Laboratory of Immunovirology, Department of Infectious Diseases, Biomedicine Institute of Seville (IBIS), University Hospital Virgen del Rocío, Sevilla, Spain. [9]Centro Sandoval, Madrid, Spain and [10]HIV Unit, Department of Internal Medicine, University Hospital Clínico San Carlos, Madrid, Spain. Correspondence: MJ Gosalbes (maria.jose.gosalbes@uv.es)
[11]These authors contributed equally to this work.

Received 30 April 2014; accepted 6 October 2014; published online 19 November 2014. doi:10.1038/mi.2014.107

and the gut microbiota has been shown to be a complex network with a critical impact on human biology and pathophysiology. It is now widely accepted that the gut microbiota is a key factor for immune homeostasis. For example, the lymphatic and colonic regulatory T cell repertoire is heavily influenced by microbiota composition.[10] However, the way the immune system shapes the microbiome and contributes to disease is poorly understood. Mounting evidence suggests that disruption of gut immunity in HIV infection also precipitates dysbiosis of the gut microbial community, which negatively affects critical pathways for healthy immune homeostasis.[11–17] In particular, the extent of dysbiosis correlates with the activity of the kynurenine pathway of tryptophan catabolism and with plasma concentrations of the inflammatory cytokine interleukin-6 (IL-6), both of which are established markers of disease progression.[11] Simultaneously, compositional and functional shifts in the gut microbiota have been described in most of the illnesses that drive excess of mortality during treated HIV infection.[11] These findings point to a novel link between the mucosal immunity, the gut bacterial community, chronic inflammation, and the risk of non-AIDS events. In other chronic inflammatory diseases, such as inflammatory bowel disease, metagenomic survey[12] and metaproteomic survey[13] have shown that the extent of dysbiosis is not limited to a shift in commensal organisms, but that it is also associated with up- or downregulation of pathways related to oxidative stress, virulence, and secretion, which are increased in the gut microbiota of patients with inflammatory bowel disease.[13] Although analysis of the more than 5,000,000 genes predicted in the human microbiome[14] will enable us to understand the metabolic influence of the dysbiotic bacterial community in the maintenance of persistent immune dysfunction during ART, the issue of whether the HIV-associated dysbiosis is also associated with a functional shift has not been addressed yet. Using *in silico* predictions based on the phylogenetic information of the 16S rRNA gene, McHardy *et al.*[15] recently described a trend of changes in different pathways such as amino acid biosynthesis and metabolism, CoA biosynthesis, folate biosynthesis, and glutathione and selenocompound metabolism in ART-naive HIV+ patients. However, data obtained from bioinformatics homology search against reference genomes might not be representative of the whole gut metagenome.

ART-treated HIV+ individuals range from otherwise healthy persons to patients with multiple chronic comorbidities. In fact, in most of HIV+ individuals on long-term ART, the risk of non-AIDS disorders,[9] which are in turn associated with a shift in microbiota, is higher than expected. This finding challenges the investigation of the persistence of intestinal dysbiosis during well-treated HIV infection, as gut microbiota might be an important target for novel nutritional interventions.

The present study is the first to determine from fecal samples the functional capacity profile of the intestinal microbiota of HIV+ patients without comorbidities during effective ART. We applied metagenome sequencing to identify a unique gene content that is characteristic of treated HIV-infected individuals. We also evaluated the relationship between functional capacity and variables involved in immune dysfunction and/or clinical progression. We used 16S rRNA gene analysis to determine the composition of fecal microbiota and examined its associations with bacterial translocation and immune activation. Finally, we built a Bayesian network that integrated the interactions of the main factors in HIV infection and thus enabled us to identify potential targets for intervention.

RESULTS

Differences in the clinical variables between ART-treated HIV+ individuals and healthy subjects

We included 15 chronically HIV-infected individuals on suppressive ART and 15 healthy controls. Patients had a median CD4 T cell nadir of 203 cells/μl, median cumulative ART exposure of 6 years, good CD4+ T cell reconstitution (584 cells/μl (466–794)), and had almost reached normal CD4/CD8 ratios (1,2 (0.9–1.3)). No meaningful differences were detected in age, body mass index, or glycemic and lipid plasma profiles, and there were non-statistically significant higher proportion of women in the control group. As for plasma biomarkers, only soluble CD14 (sCD14) levels were higher in patients than in controls ($P = 0.05$). As shown in **Supplementary Table S1** online, ART-treated HIV-infected individuals displayed lower CD4+ T cell counts and lower CD4/CD8 ratios than controls and significantly increased frequencies of CD4+ and CD8+ T cells expressing markers of T cell activation/senescence (HLA-DR/+CD38+, CD38+, CD25+) and senescence (CD57+).

Differences in gut microbiota composition between HIV-infected individuals under ART and healthy subjects

We obtained an average of 5,392 16S rRNA gene sequences per sample by multiplex pyrosequencing from DNA extracted from fecal samples of a subset of our cohort (12 controls and 9 ART-treated HIV+ patients). The remaining nine participants initially consented to participate in the study but then refused to provide stool samples. Taxonomical assignation was performed at operational taxonomic units (OTUs) (97%) and at genus level, as they allowed a higher discriminatory power between samples. We used the weighted Unifrac distance and Bray–Curtis dissimilarity distance for the cluster analysis, because these metrics gave the optimal cluster configuration (see **Supplementary Methods**) at OTU (97%) and genus level, respectively (**Supplementary Figure S1**). The clustering showed that the samples formed two clear groups for both taxonomic levels (**Figure 1** and **Supplementary Figure S2**). These clusters and the group category (HIV− vs. HIV+) were validated by an ADONIS test at OTU (97%) and genus level ($P = 0.001$ and $P = 0.04$ for the cluster analysis and $P = 0.001$ and $P = 0.01$ for the group category, respectively). **Figure 1** showed three patients (H14, H18, and H24) located at the control cluster. The heat map showed in the **Supplementary Figure S3** also revealed two main clusters. One cluster (in red) was composed mainly of HIV+ individuals, with a high

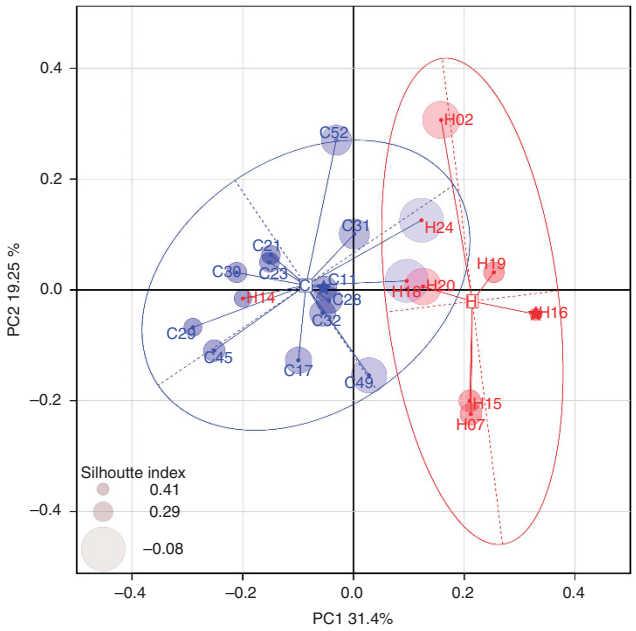

Figure 1 Comparison of microbiota between HIV + ART and uninfected subjects. Principal coordinates analysis of the bacterial composition in controls (blue dots) and cases (red dots) at operational taxonomic unit (OTU) (97%) level. The stars in blue and red correspond to the medoid retrieved from the PAM algorithm. The centroid is represented by a capital letter denoting the condition of the cluster (C for controls and H for HIV + subjects), while the blue and red ellipses represent the 95% of the samples belonging to each condition. Each point contains a halo proportional to its silhouette index value: as larger is the halo, more dissimilar is the element to its corresponding object.

abundance of *Prevotella* (44.1%) and *Succinivibrio* (14.6%). Subject C49, who harbored a high proportion of *Prevotella* genus, was also included in this cluster. The other cluster (in blue) was composed mainly of HIV − individuals in whom *Bacteroides* (27.5%) and *Faecalibacterium* (16.7%) were the most abundant genera (**Supplementary Figure S3**). Analysis of the clustering at the genus level showed that the subject H02 clustered with the control group (**Supplementary Figure S2**). This patient presented an unusual microbiota composition with a high abundance of *Bacteroides* and low level of *Faecalibacterium*. Intriguingly, the bacterial community of HIV + individuals had a much higher proportion of Gram-negative bacteria than HIV − individuals (ratio %Gram negative/%Gram positive: 71/18 vs. 45/44, respectively).

Comparison of richness estimators, ACE and Chao1, and the Shannon index revealed statistically significant differences only at OTU (97%) level (**Supplementary Table S2**), being the bacterial community of HIV − individuals more diverse. However, the overall bacterial load (expressed as number of 16S rRNA gene copies) was significantly higher in HIV + subjects than in HIV − subjects (**Supplementary Table S2**).

We used the linear discriminative analysis (LDA) effect size (LEfSe) biomarker discovery tool to elucidate which genera were driving divergence between the groups. We found 11 biomarkers for the HIV + cluster: 7 belonged to the Firmicutes phylum: *Acidaminococcus* ($P = 0.01$), *Butyrivibrio* ($P = 0.00$), *Eubacterium* ($P = 0.02$), *Mitsuokella* ($P = 0.00$), *Bulleidia* ($P = 0.00$), *Megasphaera* ($P = 0.01$), and *Catenibacterium* ($P = 0.00$); 3 belonged to the Proteobacteria phylum: *Succinivibrio* ($P = 0.00$), *Trabulsiella* ($P = 0.03$), and *Desulfovibrio* ($P = 0.02$); and finally, a single genus from the Bacteroidetes phylum: *Prevotella* ($P = 0.00$). In the control group, we observed nine biomarkers: six belonged to the Firmicutes phylum: *Faecalibacterium* ($P = 0.00$), *Roseburia* ($P = 0.00$), *Ruminococcus* ($P = 0.04$), *Blautia* ($P = 0.00$), *Coprococcus* ($P = 0.03$), and *Anaerostipes* ($P = 0.00$); two belonged to Bacteroidetes phylum: *Bacteroides* ($P = 0.00$) and *Parabacteroides* ($P = 0.00$); and a single genus representing Proteobacteria phylum: *Escherichia* ($P = 0.00$) (**Figure 2a**). These biomarkers presented high LDA scores (LDA > 3.5) and generated sample clustering similar to those obtained using all taxa (**Supplementary Figure S3** and **Figure 2b**). The *Prevotella* biomarker had the highest LDA score (5.2) and was 7.8 times more abundant in HIV + subjects than in HIV − subjects (1542.0 ± 641.9 vs. 196.9 ± 288.2), whereas *Succinivibrio* (LDA score = 4.8) was not present in HIV − subjects. In control group, *Bacteroides* and *Faecalibacterium* presented the highest discriminative power (LDA score = 5.1 and 4.8, respectively).

In order to elucidate whether specific clinical parameters might correlate with the extent of dysbiosis, we calculated the correlations (Spearman rank correlation coefficient (ρ)) between the first component of the principal coordinates analysis (PC1) and the CD4 + T cell counts, CD8 + T-cell counts, CD4/CD8 ratio, CD4 nadir, time from HIV diagnosis to ART initiation, and duration of ART. Only the CD4/CD8 ratio correlated with the PC1 ($\rho = -0.4$, $P = 0.045$), although statistical significance was lost after adjustment for multiple comparisons (adjusted $P = 0.117$).

Impact of total microbiota on immunological predictors of disease progression

Correlation analysis was performed between the first component of the PC1 at OTU level (97%) and markers of bacterial translocation (bactericidal-permeability increasing protein (BPI), sCD14), monocyte activation (T cell activation %HLA-DR + /CD38 + /CD25 +) and senescence (%CD57 +), thymic function (sj/β-TREC ratio), inflammation (high-sensitivity C-reactive protein (hs-CRP]) and IL6), and thrombosis (D-dimers). PC1 correlated positively with the inflammation marker hs-CRP and with markers of T cell activation, including %CD4 + HLA-DR + CD38 + T cells, %CD4 + CD25 + T cells, %CD8 + HLA-DR + CD38 + T cells, and %CD8 + CD38 + T cells (**Figure 3a–e**). No significant correlation was observed between PC1 and the markers of bacterial translocation (sCD14, $\rho = 0.2$, adjusted $P = 0.56$ and BPI, $\rho = 0.3$, adjusted $P = 0.45$), thrombosis (D-dimers, $\rho = 0.3$, adjusted $P = 0.41$), inflammation (IL6, $\rho = -0.2$, adjusted $P = 0.45$), and thymic function (sj/β-TREC ratio, $\rho = -0.25$, adjusted $P = 0.43$).

As lipopolysaccharide (LPS) is one of the principal antigens translocated from the gut to the bloodstream driving chronic

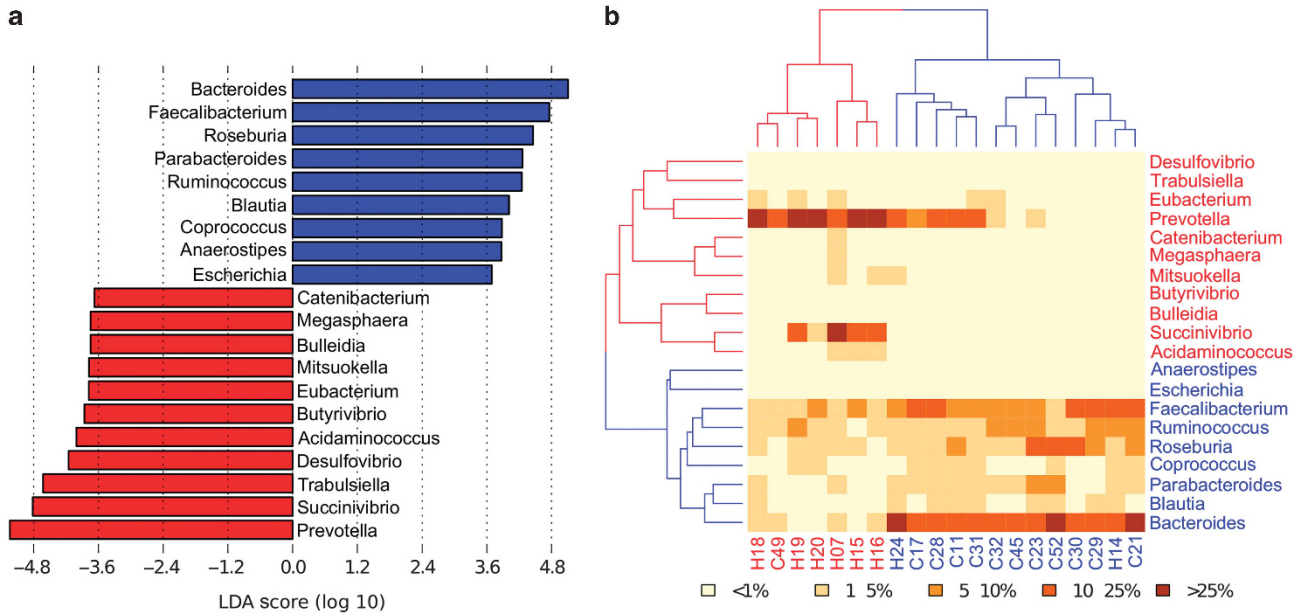

Figure 2 Taxonomic biomarkers. (**a**) linear discriminative analysis (LDA) effect size LEfSe) analysis between the case cluster (in red) and control cluster (in blue). LDA scores (log 10) for the most prevalent taxa in controls are represented on the positive scale, whereas LDA-negative scores indicate enriched taxa in cases. (**b**) Heat map of genus biomarkers. Biomarkers are represented in red and blue for cases and controls, respectively. In the heat map, the percentage range of sequences assigned to main taxa (abundance >1% in at least one sample) is represented by a color gradient.

immune activation, we assessed the correlation between PC1 and the number of sequences assigned to LPS biosynthesis pathway (see metagenome analyses below), and we found a significant positive correlation (**Figure 3f**).

Differences in microbiota metabolic functions between HIV− and ART-treated HIV+ subjects

To explore functional hallmarks of the HIV-associated dysbiotic bacterial community, we analyzed the metagenome in both groups. Pyrosequencing of the samples yielded a total of 659 Mb with an average length of 655 ± 32 bp. All the high-quality reads (average 18.5 Mb per sample) were compared with the Kyoto Encyclopedia of Genes and Genomes orthology (KO) database at different hierarchical levels (level 1, top level; level 2, subcategories of the top level; pathway level, and KO, gene level),[16] giving a high functional assignment to 34% of the reads (49,517 open reading frames per sample).

At level 2, the functional profiles were fairly homogeneous for all individuals. The most abundant categories were "carbohydrate metabolism" (19.2%), "amino acid metabolism" (13.8%), "energy metabolism" (9.2%), and "nucleotide metabolism" (8.8%), thus highlighting the importance of the gut microbiota in these metabolic pathways. Similarly, major pathways in both groups were related to purine and pyrimidine metabolism, amino and nucleotide sugar metabolism, alanine, aspartate and glutamate metabolism, and transport systems.

We used the LEfSe package to identify significant variations in the functional profile of both groups at different hierarchical levels. At level 2, HIV+ individuals showed a unique differential category, i.e., "infectious diseases" ($P<0.001$), while "carbohydrate metabolism" ($P=0.02$) and "endocrine system" ($P=0.04$) were significantly more abundant in the HIV− group. However, we found 173 KOs that were significantly different between both groups. We then analyzed pathways that also had high discriminative power. **Figure 4** shows the biomarkers found at pathway level: 12 in the HIV+ group and 23 in the HIV− group. All biomarkers presented similar percentages of sequencing coverage in both groups (**Supplementary Table S3**). For HIV+ subjects, the pathways with the highest discriminative power were the "ribosome" and "LPS biosynthesis" pathways (LDA score = 3.2 and $P=0.01$ and LDA score = 3.2 and $P=0.00$, respectively), followed by the "phenylalanine tyrosine and tryptophan biosynthesis" pathway (ko00400) (LDA score = 2.8, $P=0.00$). In this group, functional biomarkers were mainly involved in biosynthetic pathways as "terpenoid backbone biosynthesis", "fatty acid biosynthesis", "ubiquinone and other terpenoid-quinone biosynthesis," and "zeatin biosynthesis". The KO-annotated genes belonging to pathogenesis processes as "Legionellosis" and "*Vibrio cholerae* pathogenic cycle" pathways were over-represented in the HIV+ group. In the HIV− subjects, nine biomarkers were related to metabolite degradation and eight to metabolism, with the "starch and sucrose metabolism" pathway presenting the highest LDA score (LDA score = 3.1, $P=0.03$). One of the biomarkers in this group, the peroxisome proliferator-activated receptor signaling pathway (LDA score = 2.6, $P=0.01$), contained phosphotransferases that have been related to anti-inflammatory responses.

The heat map obtained using the pathway biomarkers found for HIV+ patients and controls is shown in **Figure 5**. In contrast to that observed in the compositional clustering analysis, subject C49 (HIV−) showed a healthy functional

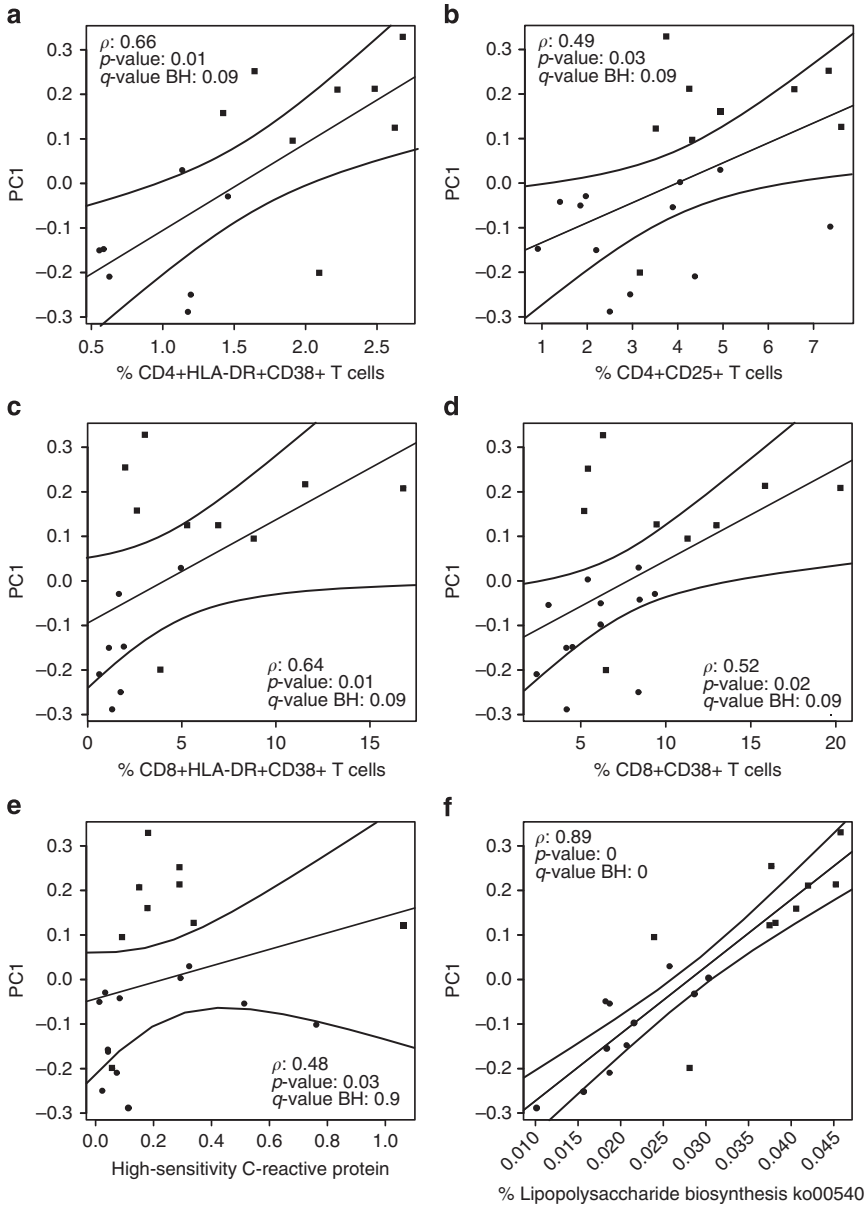

Figure 3 Associations between PC1 and markers of immune activation, inflammation, and bacterial translocation. (**a**–**d**) Correlation between PC1 and %CD4 + HLA-DR + CD38 +, %CD4 + CD25 +, %CD8 + HLA-DR + CD38 +, and %CD8 + CD38 + T cells. (**e**) Correlation between PC1 and high-sensitivity C-reactive protein (hs-CRP). (**f**) Correlation between PC1 and the lipopolysaccharide (LPS) biosynthesis pathway, as a maker of bacterial translocation. The best-fitted linear model that predicted T-cell markers and the LPS pathway is represented as a solid line. Dashed lines represent the 95% confidence interval for the linear regression coefficients. Dots represent controls and squares correspond to HIV + individuals. ρ represents the Spearman correlation coefficient and its corresponding *P*-value. *q*-value BH is the *P*-value adjusted using the Benjamini–Hochberg correction.

profile, whereas subject H24 (HIV +) clustered with the HIV + group and had a dysfunctional microbiota. Only subject H14 remained clustered with the HIV − group (sample H02 was not included in this analysis). In the HIV + group, the microbiota was depleted in genes belonging to main energetic processes as pyruvate metabolism, glycolysis and gluconeogenesis. In addition, a decrease in amino acid metabolism (glycine, serine, threonine, tryptophan, and histidine) was identified in this group (ko00260, ko00260, ko00260, ko00380, and ko00346, respectively). However, we detected enrichment for pathways involved in the metabolism of cofactors and vitamins (ubiquinone and other terpenoid–quinone biosynthesis, thiamine metabolism and nicotinate and nicotinamide metabolism) in the dysbiotic bacterial community.

Finally, as inflammation has been associated with oxidative stress, we investigated pathways that could be specifically related to this process. Even though the differences did not reach statistical significance, we observed higher relative abundance of genes belonging to glutathione metabolism and D-glutamine and D-glutamate metabolism in the HIV + group than in the HIV − group (0.66 vs. 0.28%; 0.26 vs. 0.21%, respectively).

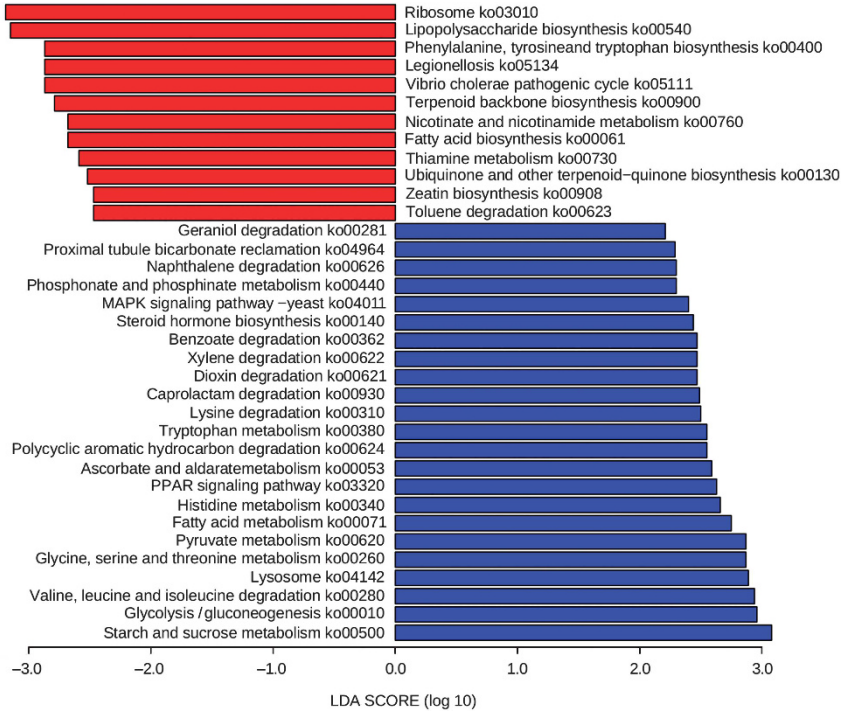

Figure 4 Linear discriminative analysis (LDA) effect size (LEfSe) analyses of statistically significant KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways. Negative LDA scores (red) are enriched in patients while positive LDA scores (blue) are enriched in controls.

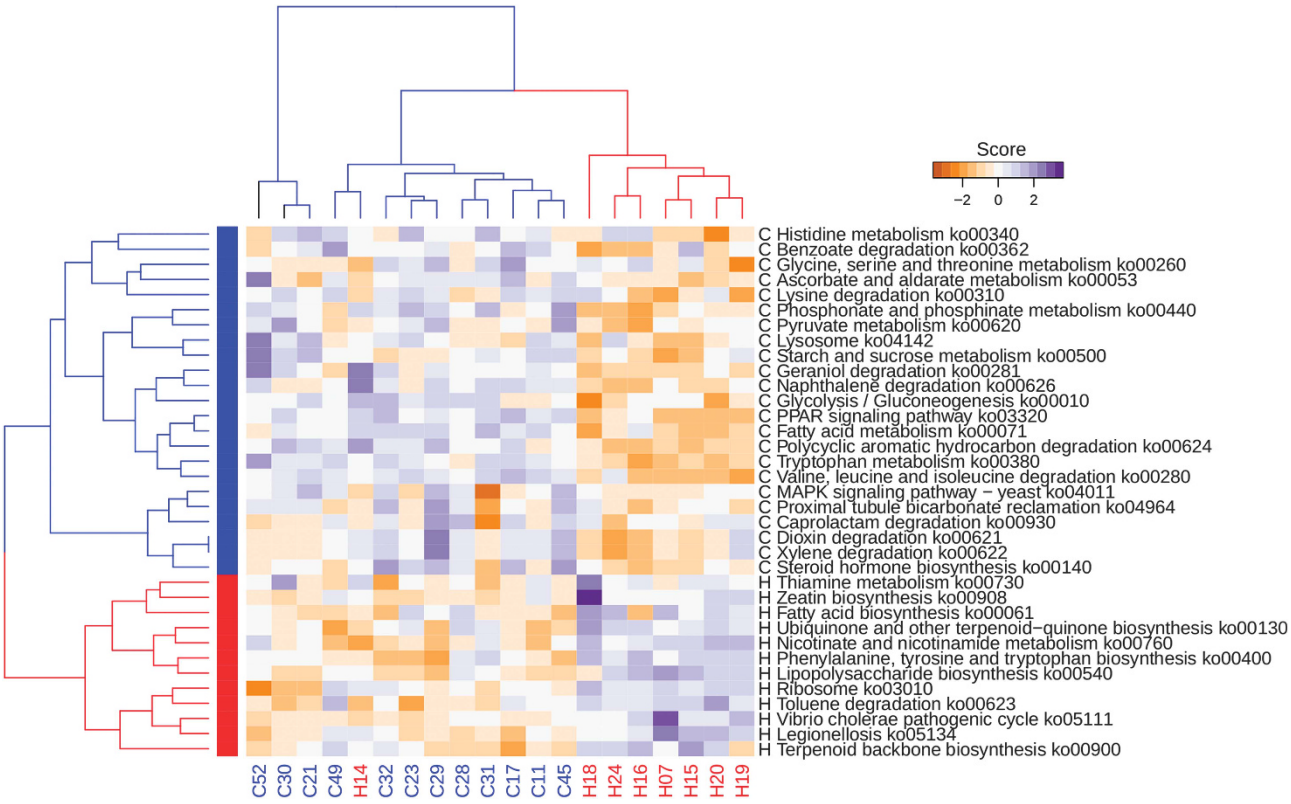

Figure 5 Heat map of the functional biomarkers for patients (in red) and controls (in blue). Over/underrepresentation is depicted by a color gradient. On the right, "H" identifies those pathways predicted as biomarkers for cases and "C" those pathways predicted as biomarkers for controls. The dendogram of the pathway abundances is divided in two main groups, a blue cluster (controls) and a red cluster (patients).

Correlations between markers of innate and adaptive immunity, and gut microbiota metabolic pathways in HIV + subjects on effective ART

We explored the correlations between metabolic functions of the intestinal microbiota and markers of T cell activation/senescence (% of CD4 + and CD8 + T cells expressing HLA-DR/ + CD38 +, CD38 +, CD25 +, or CD57 +), bacterial translocation (BPI and sCD14), inflammation (hs-CRP and IL6), thrombosis (D-dimers), and thymic function. Different associations were found (**Table 1**), although no relation remained statistically significant after adjusting for multiple comparisons. Positive correlations were detected between %CD4 + HLA-DR + CD38 +, and the "LPS biosynthesis" and "glutathione metabolism" pathways (**Table 1**). Interestingly, the"zeatin biosynthesis" pathway correlated negatively with bacterial translocation markers, sCD14 and BPI, as well as with the sj/β-TREC ratio. hs-CRP and the thrombosis D-dimers correlated negatively with the "thiamine metabolism pathway" and only the "glutathione metabolism" pathway correlated positively with hs-CRP (**Table 1**).

Bayesian networks and Markov Blankets estimation

Bayesian networks are probabilistic models in which the nodes correspond to random variables and the arcs represent causal relationships.[17] With cross-sectional data, the connecting arrows represent mutual associations rather than causality.[18]

Table 1 Significant correlation between pathways and clinical variables

	Spearman correlation index	*P*-value[a]	*q*-value[b]
T-cell markers			
Ribosome ko03010|CD4 T cells	−0.71	0.06	0.09
Terpenoid backbone biosynthesis ko00900|CD4 T cells	0.67	0.08	0.09
Glutathione metabolism ko00480|%CD4 + HLA-DR + CD38 + T cells	0.71	0.05	0.09
Lipopolysaccharide biosynthesis ko00540|%CD4 + HLA-DR + CD38 + T cells	0.71	0.05	0.09
D-Glutamine and D-glutamate metabolism ko00471|%CD4 + CD38 + T cells	−0.71	0.06	0.09
Nicotinate and nicotinamide metabolism ko00760|%CD4 + CD25 + T cells	0.95	0.00	0.00
Ribosome ko03010|%CD4 + CD25 + T cells	0.83	0.02	0.09
Glutathione metabolism ko00480|%CD8 + CD25 + T cells	0.79	0.03	0.09
Toluene degradation ko00623|%CD4 + CD25 + T cells	0.67	0.08	0.09
D-Glutamine and D-glutamate metabolism ko00471|%CD8 + CD57 + T cells	0.64	0.1	0.1
Bacterial translocation			
Legionellosis ko05134|BPI[c]	0.86	0.01	0.07
Glutathione metabolism ko00480|BPI[c]	0.79	0.03	0.09
Lipopolysaccharide biosynthesis ko00540|BPI[c]	0.71	0.06	0.09
Zeatin biosynthesis ko00908|BPI[c]	−0.71	0.06	0.09
Zeatin biosynthesis ko00908|sCD14[d]	−0.86	0.01	0.07
Ubiquinone and other terpenoid quinone biosynthesis ko00130|sCD14[d]	−0.64	0.10	0.10
Thymic function			
Zeatin biosynthesis ko00908|sj/β-TREC ratio	−0.75	0.03	0.09
Thiamine metabolism ko00730|sj/β-TREC ratio	−0.66	0.08	0.09
Inflammation			
Thiamine metabolism ko00730|hs-CRP[e]	−0.65	0.08	0.09
Glutathione metabolism ko00480|hs-CRP[e]	0.62	0.10	0.10
Thrombosis			
Thiamine metabolism ko00730|D-dimers	−0.67	0.08	0.09

[a]*P* is probability at $\alpha = 0.1$.
[b]*P*-value adjusted by the Benjamini–Hochberg method.
[c]Bactericidal-permeability increasing protein.
[d]Soluble CD14.
[e]High-sensitivity C reactive protein.

To unravel the complex interactions between microbiota and metabolic pathways contributing to T cell activation, thymic function and bacterial translocation, we fitted a Bayesian model. This analysis demonstrated a complex network in which most pathways and genera are interconnected. Most of these variables were associated with at least one genus or pathway, with the exception of BPI and the %CD8 + CD57 + T cells that exhibited a greater number of interactions with other network components (**Supplementary Figure S4**).

The set of nodes that predicts the behavior of another node in a Bayesian network is named the "Markov Blanket".[17] We estimated the Markov Blanket for the subset of nodes that showed a significant correlation with the immunological predictors. Three Markov Blankets were selected as follows: (i) one that included lipopolysaccharide biosynthesis and zeatin biosynthesis pathways, hereafter referred to as lipopolysaccharide–zeatin Markov Blanket (**Figure 6**); (ii) that included the %CD8 + CD57 + and %CD4 + CD38 + clinical variables, hereafter referred to as %CD8 + CD57 + and %CD4 + 38 + Markov Blanket (**Figure 7a**); and (iii) the Markov Blanket of the *Coprococcus* genus (**Figure 7b**). The lipopolysaccharide–zeatin Markov Blanket contained positive correlations between genera of Gram-negative bacteria and pathways related to inflammation as Legionellosis, *V. cholerae* pathogenic cycle or lipopolysaccharide biosynthesis. Together with glutathione metabolism, the lipopolysaccharide biosynthesis pathway correlated positively with BPI and %CD4 + HLA-DR + CD38 + T cells. The sCD14 marker correlated negatively with the zeatin biosynthesis pathway, and unexpectedly with *Prevotella* abundance. Finally, we observed no positive correlations with the sj/β-TREC ratio (**Figure 6**). Then, we explored the Markov Blanket for activated (%CD4 + CD38 +) and senescent (%CD8 + CD57 +) T cells (**Figure 7a**). The depletion of *Faecalibacterium* was associated with an

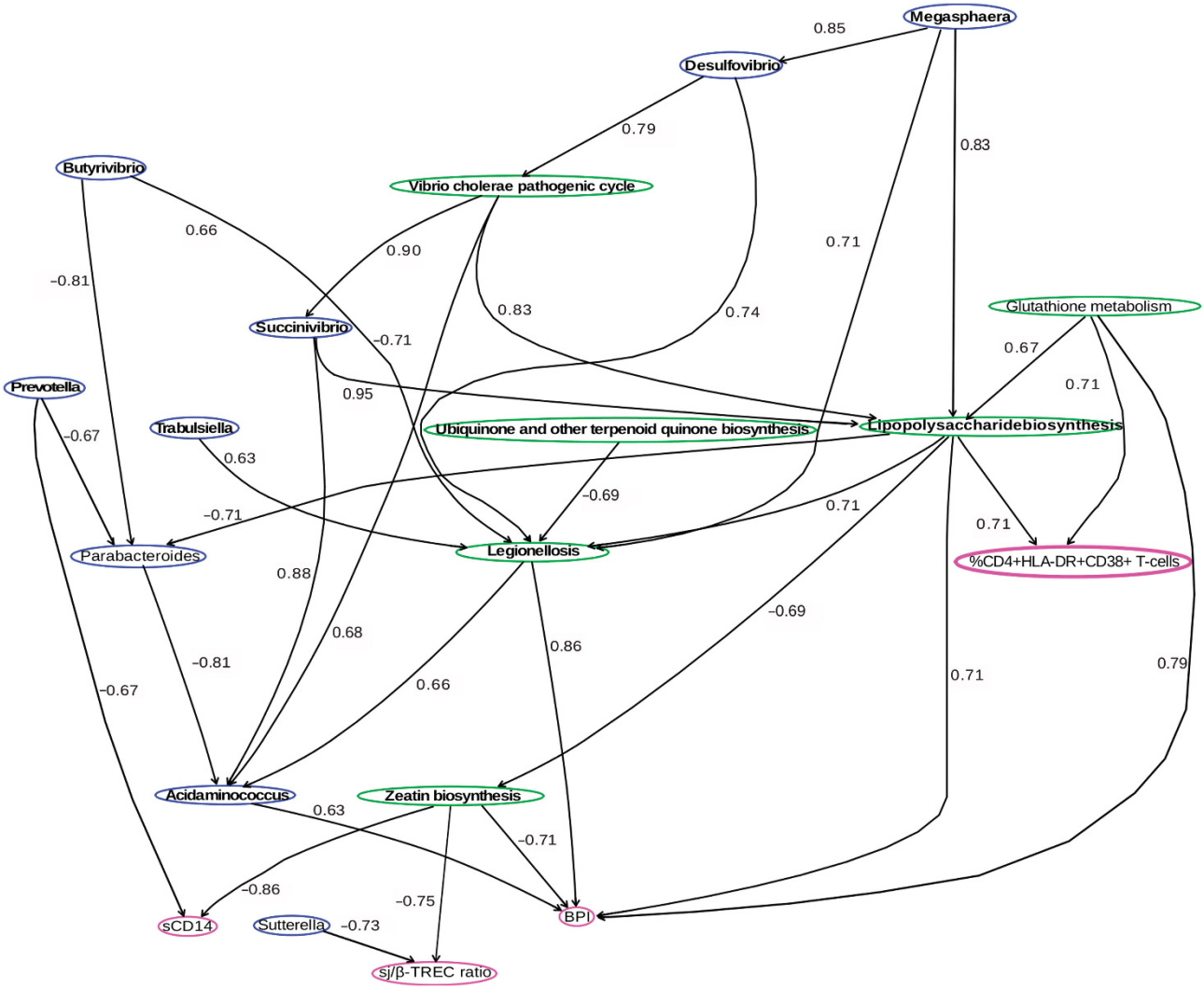

Figure 6 Lipopolysaccharide–zeatin Markov Blanket. Subgraph of the Bayesian networks represent the relationships between genera abundance (blue ellipses), pathway abundances (green ellipses) and markers of adaptive immunity, thymic function, and bacterial translocation (pink ellipses) related to the lipopolysaccharide and zeatin biosynthesis pathways. Taxa and pathway biomarkers of HIV + group are in bold. Arrows indicate conditional dependencies between variables. The Spearman correlation coefficient is indicated next to the lines.

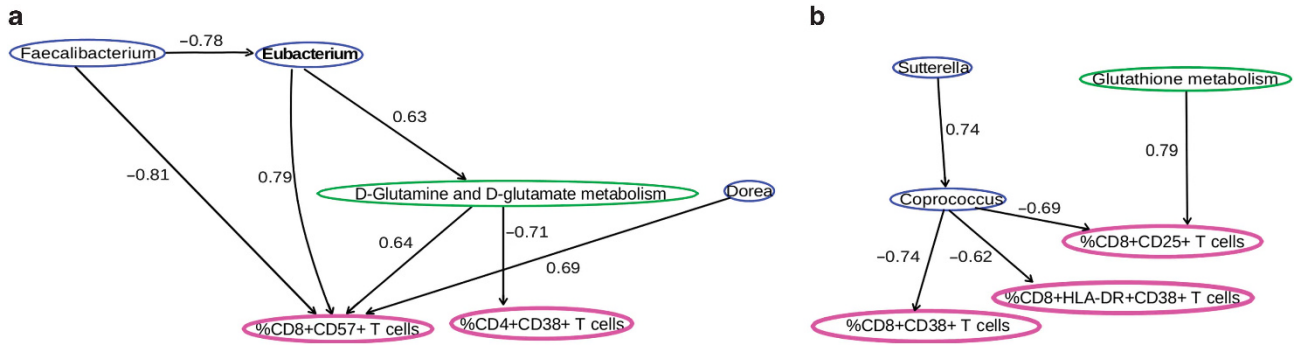

Figure 7 Network integrating genera, markers of adaptive immunity, and metabolic pathways. (**a**) Markov Blankets of the %CD8 + CD57 + T cells and %CD4 + CD38 + T cells (**b**) Markov Blanket of the *Coprococcus* genus. Bayesian networks show the relationships between genera abundance (blue ellipses), pathway abundance (green ellipses), and markers of adaptive immunity (pink ellipses). Taxa and pathway biomarkers of the HIV + group are indicated in bold. Arrows indicate conditional dependencies between variables. The Spearman correlation coefficient is indicated next to the lines.

overgrowth of *Eubacterium* and higher %CD8 + CD57 + T cells, which in turn correlated positively with the genus *Dorea* genus, and D-Glutamine and D-glutamate metabolism pathway. The %CD4 + CD38 + T cells showed a single negative correlation with the latter metabolic pathway. For *Coprococcus,* the Markov Blanket illustrated how a low abundance of this genus had an impact on different phenotypes of activated T cells (%CD8 + CD38 +, %CD8 + HLA-DR + CD38 +, and %CD8 + CD25 +). The glutathione metabolism pathway again correlated positively with markers of T cell activation (%CD8 + CD25 +) (**Figure 7b**).

DISCUSSION

Most studies show that despite modern ART, HIV + individuals have reduced life expectancy, mainly owing to the increase in morbidity and mortality associated with non-AIDS-related diseases, which are driven in part by persistent inflammation and immune activation.[3] Several recent studies[15,19–25] have explored the relationships between gut microbiota composition and HIV infection using different sampling techniques (stool, anal swab, biopsy, and sponge collection) and bacterial classification methods (16Sr RNA gene PCR, quantitative PCR and microarray). Even though stool microbiota might not be an exact reflection of the gut-resident bacterial community,[26] amplification of the16S rRNA gene in stools will probably be sufficiently informative, as it recovers bacteria from mucosal desquamation and is the easiest sampling technique to be considered standard for these investigations. Indeed, our results are congruent with those of previous studies based on colon biopsies that report similar dysbiotic microbiota depleted in *Bacteroides* and enriched in *Prevotella*.[20,25]

Given the broad range of clinical consequences of heightened immune activation in HIV + individuals, we focused our study on HIV + subjects without comorbidities receiving long-term effective ART. These patients are probably good candidates for strategies aimed at shaping the gut microbiome. In our study, HIV + individuals presented a distinctive microbiota composition characterized by high abundance of *Prevotella* and *Succinivibrio*, and depletion of *Bacteroides*, *Faecalibacterium*, and *Roseburia.* Some of these taxa were recently reported to be members of the gut-resident microbiota in HIV + individuals.[15,20,23,24] The *Enterobacteriaceae* family has been associated with bacterial translocation and immune activation,[27,28] although, in contrast with Vujkovic-Cvijn *et al.*[23] and consistent with Lozupone *et al.*,[24] we did not detect a high prevalence of this bacterial taxon in our HIV-infected cohort. This discrepancy is most likely due to the use of mucosal samples rather than luminal samples for these studies. Vujkovic-Cvijn and Dunham[23] examined colonic mucosal samples, which are widely colonized by Enterobacteriaceae, whereas Lozupone *et al.*[24] examined stool, as did we. However, our HIV-infected cohort showed a gut microbiota composition dominated by Gram-negative bacteria—representing 71% of the bacterial community—increased bacterial load and decreased diversity at the OTU (97%) level. Intriguingly, the two individuals (H14 and H24) with a "normal" microbiota composition showed the highest CD4/CD8 ratio in the cohort, and this ratio was the only variable available in clinical practice that correlated with the extent of dysbiosis, suggesting that normalization of this biomarker of immunological dysfunction during ART may also predict normalization of gut microbiota, at least in terms of composition. In ART-treated individuals, the fact that the CD4/CD8 ratio correlates with activity of the kynurenine pathway of tryptophan catabolism, an established marker of disease progression,[29] provides indirect support for the recent observation by Vujkovic-Cvijin and Dunham[23], who suggested that dysbiosis of the gut microbial community affects negatively this critical pathway in healthy immune system.[23]

In other studies, reduced diversity has been observed in treated HIV + individuals,[24,25] indicating that the combination of HIV infection and its treatment leads to decreased microbial diversity. This particular microbiota composition, characterized by *Prevotella* enrichment and *Bacteroides* depletion, has been described mainly in human populations with a carbohydrate-rich diet[30–33] and recently in HIV-infected subjects.[24] In addition, Lozupone *et al.* [24] showed that the microbiota of HIV + subjects is closer to those of people from agrarian

cultures than to that of healthy people from the United States. All subjects in our cohort followed a similar western diet, indicating that HIV infection is the likely factor driving changes in the gut microbiota. The observed depletion of *Faecalibacterium* and *Roseburia*, the major butyrate producers, might result in lower local production of short-chain fatty acids. In the healthy human gut, short-chain fatty acids such as butyrate, propionate, and acetate are an important energy source for the maintenance of homeostasis in the colonic mucosa and display anti-inflammatory properties.[34–36] Alterations in the short-chain fatty acid ratio have been related to an amplification of inflammatory responses in diseases that are typically associated with bacterial dysbiosis, such as ulcerative colitis and bacterial vaginosis.[12,37] In fact, a drastic decrease in the number of genes involved in glycolysis and pyruvate metabolism was observed in the metagenomes from HIV-infected individuals. Further short-chain fatty acid quantification should be performed to address its involvement in the inflammation process. On the other hand, depletion of the anti-inflammatory commensal genus *Faecalibacterium* has been reported in the anal microbiota of treated HIV + patients[25] and in patients with Crohn's disease.[38] Interestingly, the Markov Blankets of the %CD8 + CD57 + T cells (**Figure 7**) parallels previous observations in centenarians, in whom the depletion of *Faecalibacterium* genus correlates with an increase in *Eubacterium* species[39] and with an increase in %CD8 + CD57 + T cell, a hallmark of immunosenescence in HIV infection.[8,40] Furthermore, given that certain *Bacteroides* species are required for differentiation of the Th17 T cells, the significant loss of this bacterial genus in treated-HIV individuals might either aggravate the Th17 loss secondary to HIV infection or impair their reconstitution under ART.[41]

The high number of significantly abundant pathways and KOs in the bacterial community of HIV + individuals in our study revealed a functional dysbiosis that might explain, at least in part, the situation of chronic inflammation observed during treated HIV infection. Subject H24 (HIV +) had an altered metabolic profile despite clustering with individuals with a healthy bacterial composition. The fact that this subject was overweight (body mass index = 29 kg m^{-2}) suggests that other pro-inflammatory factors, such as obesity-related factors, might have a role. In addition, control C49 had an altered gut microbiota composition dominated by *Prevotella* genus '*Prevotella* enterotype'[42] but showed healthy functional capacity.

The significant correlation found between PC1 and the LPS biosynthesis pathway (P-value = 0, adjusted P-value BH = 0, $\rho = 0.89$) would indicate that Gram-negative-enriched microbiota was responsible for the high abundance of KOs belonging to the LPS biosynthesis pathway in HIV + individuals. These lipopolysaccharides are microbe-associated molecular patterns, which are potent immune activators that act via Toll-like receptor (TLR4) by promoting the inflammatory response.[43] In other gut microbiota-associated diseases and *in silico* predictions of the HIV microbiota, an increase in antioxidant pathways such as riboflavin, glutathione, and glutamine metabolism has been described and interpreted as the bacterial compensatory mechanism that attenuates the oxidative stress caused by epithelial damage.[12,15,36,44] Likewise, the microbiota of treated HIV + individuals was functionally enriched for ubiquinone and other terpenoid–quinone biosynthesis, nicotinate and nicotinamide metabolism, glutathione metabolism, and thiamine metabolic pathways. This latter pathway correlated negatively with hs-CRP and D-dimers. These results are congruent with the anti-inflammatory effect of thiamine that has been described in mammals[45,46] and with the fact that a depletion of this vitamin exists in other gastrointestinal illnesses, such as Crohn's disease.[47] Unlike the observations in a previous study,[23] we did not detect differences in genes involved in the kynurenine pathway of tryptophan catabolism. This inconsistency could be given by the different nature of the sample (colorectal biopsies vs. stool samples) or by the fact that the species capable to perform tryptophan catabolism are less frequent in the stool microbiota. In our view, the depletion of genes involved in amino acid metabolism and energy processes might be upregulating inflammatory pathways in HIV-infected individuals.

Altogether, we propose a complex network that integrates the different interactions between gut microbiota, metabolic functions, and host immunity. Although no relationship was found between microbial composition and sCD14, we found a correlation between bacterial translocation markers (sCD14 and BPI) and the zeatin biosynthesis pathway. As zeatins belong to a class of phytocytokine involved in the complex cell-signaling pathway, this observation suggests that the interaction between the microbiota and bacterial translocation could occur indirectly by cytokine signaling. Importantly, we detected a strong correlation between bacterial genera composition and the LPS pathway, suggesting that the abundance of Gram-negative bacteria in the dysbiotic microbiota of HIV-infected subjects might contribute to the burden of translocated bacterial antigens and, consequently, to chronic immune activation. This hypothesis is also supported by the significant association between the LPS biosynthesis pathway and both BPI and T-cell activation (%CD4 + HLA-DR + CD38 + T cells). Similarly, pathways related to bacterial antioxidant response, glutathione metabolism, and D-glutamine and D-glutamate metabolism correlated with different immune activation markers.

Considering our findings and current knowledge in the field, we believe that the profound disruption of gut-associated lymphoid tissue secondary to HIV infection would generate a dysbiotic microbiota, both in terms of its composition—Gram-negative bacteria enrichment—and in terms of its altered metabolic profile with many genes involved in the LPS biosynthesis pathway, pathogenic pathways, and processes related with oxidative stress. This compositional and functional dysbiosis seems to fuel chronic innate and adaptive immune dysfunction and may be a viable target for interventions. Further longitudinal studies should be performed to provide evidence of causality in the correlations. A larger sample size and the inclusion of HIV + patients with a different

immunological response will be necessary to obtain a comprehensive understanding of the role of dysbiotic microbiota in HIV infection.

METHODS

Study design, participants, setting, and eligibility. We conducted a case–control study in 30 participants (15 cases and 15 controls). Cases were HIV+ patients attending the HIV clinics of two University hospitals in Madrid, Spain (University Hospital Clínico San Carlos and University Hospital Ramón y Cajal). The inclusion criteria were serologically documented HIV infection, age 18 years or older, at least 2 years under ART-mediated HIV RNA suppression with a regimen containing at least three antiretroviral drugs, and a CD4+ T cell count $\geq$350 cells per µl. The controls were healthy non-HIV-infected volunteers (mostly staff working in either institution), who were recruited to form an age-matched control group. The exclusion criteria were use of concomitant medications, use of systemic antibiotics during the previous 3 months, and any acute or chronic condition other than chronic HIV infection, including gastrointestinal symptoms (constipation, bloating, or diarrhea) or co-infections by hepatitis B or C viruses. We were only able to collect fecal samples from 12 controls and 9 HIV+ patients.

This study was conformed to the principles of the Declaration of Helsinki and the Good Clinical Practice Guidelines, and was approved by the Independent Ethics Committees at both recruiting institutions (University Hospital Clínico San Carlos ceic.hrc@salud.madrid.org and University Hospital Ramón y Cajal, ceic.hrc@salud.madrid.org). All patients and healthy donors provided written informed consent.

Nucleic acid purification. Fecal samples were stored in RNAlater solution (Life Technologies, Carlsbad, CA) at −80°C until use. Total DNA was extracted as described elsewhere.[48] Briefly, samples were diluted (dilution 1:2) in phosphate-buffered saline solution before being centrifuged at 2,000 r.p.m. at 4 °C for 2 min to remove fecal debris. The supernatant was centrifuged at 13,000 r.p.m. for 5 min to pellet the cells. Total DNA was extracted from pelleted cells using QIAamp DNA Stool Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

Amplification of the 16S rRNA gene. For each sample, a region of the 16S rRNA gene was amplified by PCR with the universal primers E8F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 530R (5′-CCGCGGC KGCTGGCAC-3′). The E8F primer included a sample-specific multiplex identifier to be multiplexed and sequenced simultaneously. The amplified region comprises hypervariable regions V1, V2, and V3. PCR products were purified using the NucleoFast 96 PCR Clean-Up Kit (Macherey-Nagel, Düren, Germany). The pooled PCR products were pyrosequenced directly following GS FLX+ emPCR amplification manual available at www.454.com/my454.

Pyrosequencing. Metagenomes and 16S rRNA gene amplicons were pyrosequenced using a Roche GS FLX sequencer and titanium chemistry at the Centre for Public Health Research (FISABIO-Salud Pública, Valencia, Spain). All sequences were deposited in the public European Nucleotide Archive server under accession number PRJEB5185.

Microbial quantification by quantitative PCR. Quantitative PCR conditions are detailed in **Supplementary Materials**.

16S RNA: Phylogenetic analysis, biodiversity and clustering. Amplicon data from the 16S rRNA gene were analyzed following the recommendations of the metagenomic state-of-the-art pipeline QIIME (v 1.6).[49] Detailed methods for the analysis on the amplicon data are described in **Supplementary Materials**.

Metagenome analysis. Sequence trimming, dereplication and removal of host sequences were performed using the MG-RAST pipeline (Release version 3.2) (default parameters) (http://metagenomics.anl.gov/).[50] Functional assignments were obtained from the MG-RAST pipe line using BLAT software (https://genome.ucsc.edu/FAQ/FAQblat.html) (*e*-value e−5, minimum identity 60%, and minimum alignment length 15 amino acids) against the KO database for each hierarchical level (Level 1, Level 2, Pathway level, and KO group).

Biomarker discovery. The LEfSe algorithm[51] was used to identify specific taxa and functions as biomarkers for cases and controls. We fixed an α-value <0.05. The bacterial taxa, or function, with significant differences between samples were used to build the LDA model and to estimate its effect as a discriminant feature between them. The threshold used to consider a discriminative feature for the logarithmic LDA score was set at >2. The biomarker discovery was performed at the genus level and for all hierarchical functional levels. We considered sample H02 as an outlier, owing to its bacterial and functional composition, and excluded it from these analyses.

Correlation analyses. We assessed correlations between markers of innate and adaptive immune activation, markers of innate immunity, and the first component of the principal coordinates analysis , as described elsewhere.[23] Linear regression coefficients (library "stats", function "lm") were calculated for all correlations with a significant *P*-value ($\alpha<0.05$) in the Spearman correlation index (library "stats", function "cor.test"). Correlation analyses were also performed between clinical variables and metabolic pathways. All *P*-values were adjusted using the Benjamini–Hochberg correction (library "stats", function "p.adjust").

Functional pathway abundances predicted by the shotgun 454 pyrosequencing sequences in the fecal samples from the controls and cases (metagenomes) were correlated using the Spearman correlation index (library "stats", function "cor.test") with a *P*-value cutoff of $\alpha<0.1$, and the markers of innate and adaptive immune activation and innate immunity.

Immunological predictors of disease progression. *Markers of innate immune activation and bacterial translocation.* A sample of fasting venous blood was obtained to determine concentrations of glucose, IL-6, total cholesterol, high-density lipoprotein cholesterol, and triglycerides using standard enzymatic methods. Plasma viral load was measured using the Cobas Taq-Man HIV-1 assay (Roche Diagnostics Systems, Branchburg, NJ). Plasma levels of hs-CRP were measured using nephelometry (VISTA System, Siemens Healthcare Diagnostics, Deerfield, IL). D-dimers were measured using turbidimetry (Beckman-Coulter, Munster, Germany). Cryopreserved plasma was assessed by immunoassay for BPI protein (Hycult Biotech, Uden, The Netherlands) and soluble CD14 (sCD14, BioVendor, Brno, Czech Republic).

Markers of adaptive immune activation. T-cell immunophenotyping: Fresh anticoagulated whole blood was used to analyze CD4+ and CD8+ T cells with the following antibody combination: CD3-eFluor450, CD4-peridinin chlorophyll protein complex, CD8-phycoerythrin-Cy7, CD38-Horizon V500, CD25-allophycocyanin (APC), HLA-DR-APC-Cy7, and CD57-fluorescein isothiocyanate. Antibodies were from Becton Dickinson (Becton Dickinson, Franklin Lakes, NJ) and unstained controls were performed for all samples. Cells were collected using a Gallios flow cytometer (Beckman-Coulter, Munster, Germany) and analyzed with Kaluza software (Beckman-Coulter) to quantify the percentage of CD4+ and CD8+ T cells expressing markers of activation (CD25+, CD38+, HLADR+, or CD38+/HLA-DR+) and senescence (CD57+).

sj/β-TREC ratio quantification: Thymic function was indirectly calculated in peripheral blood mononuclear cells DNA using the sj/β-TREC ratio, as previously described,[52] with minor modifications (see **Supplementary Materials**).

Bayesian network. The statistical R package "Bayesian network learning and inference" (bnlearn)[53] was used to estimate a probabilistic graph model between bacterial genera, functional composition and markers of innate and T-cell activation variables, thymic function, and bacterial translocation. The network topology was created using a hill-climbing score-based learning algorithm. The algorithm states the optimal network and in consequence the "father" to "child" node relationships as that which maximizes the Bayesian information criterion.

For the analyses in HIV-infected individuals, the input variables were as follows: markers of adaptive immunity, thymic function, and bacterial translocation (clinical variables); biomarker genera and taxa with relative abundance above 0.5% (bacterial genera); and pathways biomarkers, glutathione metabolism pathway (ko00480), and the D-glutamine and D-glutamate metabolism (ko00471) (metabolic functions). We did not include the sample H02 in the network estimation given that it was the most dissimilar in terms of genus composition.

The underlying graphical structure of the network and the conditional probability given the model parameters were estimated using the hill-climbing algorithm with Bayesian information criterion as the criteria of model selection (function "hc", package "bnlearn"). The option blacklist (R Package "bnlearn" function "hc") was used to define the set of arcs not included in the model, excluding those variables with a correlation *P*-value above 0.1 (R Package "stats", function "cor", method "Spearman"). Similarly, arcs with significant correlations that were not included in the final graph were incorporated by means of the whitelist option (function "hc", package "bnlearn"). Function mb (R Package "bnlearn" function "mb") was used to estimate the Markov Blanket from the lipopolysaccharide biosynthesis, zeatin biosynthesis, *Coprococcus*, and the markers of immune activation/senescence CD4 + CD38 + T cells and CD8 + CD57 + T cells.

Statistical analysis. Additional statistical methods are described in **Supplementary Materials**.

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/mi

ACKNOWLEDGMENTS

We thank Dr Nuria Jiménez for her help with pyrosequencing (Sequencing Service of FISABIO-Salud Pública, Valencia, Spain). We also thank Dr MP Francino for helpful comments. This work was supported by "Generalitat Valenciana" (Prometeo/2009/092) and by the Spanish Ministry of Economy and Competitiveness (SAF2012-31187), Redes Temáticas de Investigación en SIDA (ISCIII RETIC RD12/0017/0029 and RD12/0017/0037), and Junta de Andalucia (Proyecto de Excelencia CTS-6313). J.F.V.C. was supported by a fellowship "Ayudas Predoctorales de Formación en Investigación en Salud" from the Instituto de Salud Carlos III, Spain. S.S.V. and T.S. were supported by a grant from the Spanish Ministry of Science and Innovation (Ayudas para Contratos de Formación en Investigación Río Hortega). S.F.M. holds a grant from the Fondo de Investigaciones Sanitarias (CD10/00382).

DISCLOSURE

The author declared no conflict of interest.

REFERENCES

1. Sandler, N.G. & Douek, D.C. Microbial translocation in HIV infection: causes, consequences and treatment opportunities. *Nat. Rev. Microbiol.* **10**, 655–666 (2012).
2. Brenchley, J.M. *et al.* Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat. Med.* **12**, 1365–1371 (2006).
3. Deeks, S.G., Tracy, R. & Douek, D.C. Systemic effects of inflammation on health during chronic HIV infection. *Immunity* **39**, 633–645 (2013).
4. Jiang, W. *et al.* Plasma levels of bacterial DNA correlate with immune activation and the magnitude of immune restoration in persons with antiretroviral-treated HIV infection. *J. Infect. Dis.* **199**, 1177–1185 (2009).
5. Mavigner, M. *et al.* Altered CD4 + T cell homing to the gut impairs mucosal immune reconstitution in treated HIV-infected individuals. *J. Clin. Invest.* **122**, 62–69 (2012).
6. Neuhaus, J. *et al.* Markers of inflammation, coagulation, and renal function are elevated in adults with HIV infection. *J. Infect. Dis.* **201**, 1788–1795 (2010).
7. Sandler, N.G. *et al.* Plasma levels of soluble CD14 independently predict mortality in HIV infection. *J. Infect. Dis.* **203**, 780–790 (2011).
8. Appay, V. & Sauce, D. Immune activation and inflammation in HIV-1 infection: causes and consequences. *J. Pathol.* **214**, 231–241 (2008).
9. Deeks, S., Lewin, S. & Havlir, D. The end of AIDS: HIV infection as a chronic disease. *Lancet* **382**, 1525–1533 (2013).
10. Cebula, A. *et al.* Thymus-derived regulatory T cells contribute to tolerance to commensal microbiota. *Nature* **497**, 258–262 (2013).
11. Cho, I. & Blaser, M.J. The human microbiome: at the interface of health and disease. *Nat. Rev. Genet.* **13**, 260–270 (2012).
12. Morgan, X.C. *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol.* **13**, R79 (2012).
13. Presley, L.L. *et al.* Host-microbe relationships in inflammatory bowel disease detected by bacterial and metaproteomic analysis of the mucosal-luminal interface. *Inflamm. Bowel Dis.* **18**, 409–417 (2012).
14. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
15. McHardy, I.H. *et al.* HIV Infection is associated with compositional and functional shifts in the rectal mucosal microbiota. *Microbiome* **1**, 26 (2013).
16. Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. & Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**, 277–280 (2004).
17. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, 2nd edn (Morgan Kaufmann Publishers Inc, San Francisco California USA, 1988).
18. Su, C., Andrew, A., Karagas, M.R. & Borsuk, M.E. Using Bayesian networks to discover relations between genes, environment, and disease. *BioData Min.* **6**, 6 (2013).
19. Yu, G., Fadrosh, D., Ma, B., Ravel, J. & Goedert, J.J. Anal microbiota profiles in HIV-positive and HIV-negative MSM. *AIDS* **28**, 753–760 (2013).
20. Dillon, S., Lee, E., Kotter, C. & Austin, G. An altered intestinal mucosal microbiome in HIV-1 infection is associated with mucosal and systemic immune activation and endotoxemia. *Mucosal Immunol.* **7**, 983–994 (2014).
21. Ellis, C., Ma, Z., Mann, S. & Li, C. Molecular characterization of stool microbiota in HIV-infected subjects by panbacterial and order-Level 16S Ribosomal DNA (rDNA) quantification and correlations with immune activation. *J. Acquir. Immune Defic. Syndr.* **57**, 363–370 (2011).
22. Pérez-Santiago, J., Gianella, S. & Massanella, M. Gut lactobacillales are associated with higher CD4 and less microbial translocation during HIV infection. *AIDS* **27**, 1921–1931 (2013).
23. Vujkovic-Cvijin, I. *et al.* Dysbiosis of the gut microbiota is associated with HIV disease progression and tryptophan catabolism. *Sci. Transl. Med.* **5**, 193ra91 (2013).
24. Lozupone, C.A. *et al.* Alterations in the gut microbiota associated with HIV-1 infection. *Cell Host Microbe* **14**, 329–339 (2013).
25. Mutlu, E.A. *et al.* A compositional look at the human gastrointestinal microbiome and immune activation parameters in HIV infected subjects. *PLoS Pathog.* **10**, e1003829 (2014).
26. Durbán, A. *et al.* Assessing gut microbial diversity from feces and rectal mucosa. *Microb. Ecol.* **61**, 123–133 (2011).
27. Winter, S.E. *et al.* Host-derived nitrate boosts growth of E. coli in the inflamed gut. *Science* **339**, 708–711 (2013).
28. Mukhopadhya, I., Hansen, R., El-Omar, E.M. & Hold, G.L. IBD-what role do Proteobacteria play?. *Nat. Rev. Gastroenterol. Hepatol.* **9**, 219–230 (2012).
29. Serrano-Villar, S. *et al.* HIV-infected individuals with low CD4/CD8 ratio despite effective antiretroviral therapy exhibit altered T cell subsets, heightened CD8 + T cell activation, and increased risk of non-AIDS morbidity and mortality. *PLoS Pathog.* **10**, e1004078 (2014).

30. De Filippo, C. *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl Acad. Sci. USA* **107**, 14691–14696 (2010).
31. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227 (2012).
32. Wu, G.D. *et al.* Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108 (2011).
33. Ou, J. & Carbonero, F. Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. *Am. J. Clin. Nutr.* **98**, 111–120 (2013).
34. Louis, P., Scott, K.P., Duncan, S.H. & Flint, H.J. Understanding the effects of diet on bacterial metabolism in the large intestine. *J. Appl. Microbiol.* **102**, 1197–1208 (2007).
35. Vinolo, M.R., Rodrigues, H.G., Nachbar, R.T. & Curi, R. Regulation of inflammation by short chain fatty acids. *Nutrients* **3**, 858–876 (2011).
36. Carvalho, B.M. & Saad, M.J.A. Influence of gut microbiota on subclinical inflammation and insulin resistance. *Mediators Inflamm.* **2013**, 13 (2013).
37. Rook, G. Regulation of the immune system by biodiversity from the natural environment: An ecosystem service essential to health. *Proc. Natl Acad. Sci. USA* **110**, 1–8 (2013).
38. Sokol, H. *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl Acad. Sci. USA* **105**, 16731–16736 (2008).
39. Biagi, E. *et al.* Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. *PLoS ONE* **5**, e10667 (2010).
40. Ivanov, I.I. *et al.* Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine. *Cell Host Microbe* **4**, 337–349 (2008).
41. Deeks, S.G., Verdin, E. & McCune, J.M. Immunosenescence and HIV. *Curr. Opin. Immunol.* **24**, 501–506 (2012).
42. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473**, 174–180 (2011).
43. Verhasselt, V. *et al.* Bacterial lipopolysaccharide stimulates the production of cytokines and the expression of costimulatory molecules by human peripheral blood dendritic cells. *J. Immunol.* **158**, 2919–2925 (1997).
44. Karlsson, F.H. *et al.* Symptomatic atherosclerosis is associated with an altered gut metagenome. *Nat. Commun.* **3**, 1245 (2012).
45. Naik, S.R., Rupawalla, E.N. & Sheth, U.K. Anti-inflammatory activity of thiamine and nicotinic acid. *Biochem. Pharmacol.* **19**, 2867–2873 (1970).
46. Moallem, S., Hosseinzadeh, H. & Farahi, S. A study of acute and chronic anti-nociceptive and anti-inflammatory effects of thiamine in mice. *Iran Biomed. J.* **12**, 173–178 (2008).
47. Kuroki, F. *et al.* Multiple vitamin status in Crohn's disease. *Dig. Dis. Sci.* **38**, 1614–1618 (1993).
48. Pérez-Cobas, A.E. *et al.* Gut microbiota disturbance during antibiotic therapy: a multi-omic approach. *Gut* **62**, 1591–1601 (2013).
49. Caporaso, J. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
50. Meyer, F. *et al.* The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**, 386 (2008).
51. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60 (2011).
52. Ferrando-Martínez, S. *et al.* A reliable and simplified sj/beta-TREC ratio quantification method for human thymic output measurement. *J. Immunol. Methods* **352**, 111–117 (2010).
53. Scutari, M. Learning Bayesian Networks with the bnlearn R Package. *J. Stat. Softw.* **35**, 1–22 (2010).