Systems analysis of gut microbiome influence on metabolic disease in HIV and high-risk populations

Abigail J.S. Armstrong¹,², Kevin Quinn³, Sam X. Li¹, Jennifer M. Schneider¹, Nichole M. Nusbacher¹, Katrina A. Doenges³, Suzanne Fiorillo¹, Tyson J. Marden⁴, Janine Higgins⁵, Nichole Reisdorph³, Thomas B. Campbell¹, Brent E. Palmer¹#, Catherine A. Lozupone¹#

Affiliations:
1. Department of Medicine, University of Colorado Denver, Anschutz Medical Campus, Aurora, CO, USA
2. Department of Immunology and Microbiology, University of Colorado Denver, Anschutz Medical Campus, Aurora, CO, USA
3. Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado Anschutz Medical Campus, 80045 CO, Aurora, USA
4. Colorado Clinical and Translational Sciences Institute; University of Colorado Anschutz Medical Campus, Aurora, Colorado, 80045. USA
5. Department of Pediatrics, Section of Endocrinology, University of Colorado Anschutz Medical Campus, Aurora, CO

# Corresponding authors
Brent E. Palmer: 303-724-7203; brent.palmer@cuanschutz.edu
Catherine A. Lozupone: 303-724-7942; catherine.lozupone@cuanschutz.edu
Abstract

Poor metabolic health, characterized by insulin resistance and dyslipidemia, is higher in people living with HIV (PLWH) and has been linked with inflammation, anti-retroviral therapy (ART) drugs, and ART-associated lipodystrophy (LD). Metabolic disease is associated with gut microbiome composition outside the context of HIV but has not been deeply explored in HIV infection nor in high-risk men who have sex with men (HR-MSM), who have a highly altered gut microbiome composition. Furthermore, the contribution of increased bacterial translocation and associated systemic inflammation that has been described in HIV-positive and HR-MSM individuals has not been explored. We used a multi-omic approach to explore relationships between gut microbes, immune phenotypes, diet, and metabolic health across ART-treated PLWH with and without LD; untreated PLWH; and HR-MSM. For PLWH on ART, we further explored associations with the plasma metabolome. Sixty-nine measures of diet, gut microbes, inflammation, and demographics were associated with impaired metabolic health defined using fasting blood markers including lipids, glucose and hormones. We found microbiome-associated metabolites associated with metabolic disease including the microbially produced metabolites, dehydroalanine and bacteriohopane-32,33,34,35-tetrol. Our central result was that elevated plasma lipopolysaccharide binding protein (LBP) was the most important predictor of metabolic disease in PLWH and HR-MSM, with network analysis of predictors showing that LBP formed a hub joining correlated microbial and immune predictors of metabolic disease. Our results suggest the role of inflammatory processes linked with bacterial translocation (measured by LBP) and interaction with dietary components and the gut microbiome in metabolic disease among PLWH and HR-MSM.
Importance Statement

The role of the gut microbiome in the health of HIV infected individuals is of interest because current therapies, while effective at controlling disease, still result in long term comorbidities. Metabolic disease is prevalent in HIV-infected individuals even in well-controlled infection. Metabolic disease has been linked with the gut microbiome in previous studies but little attention has been given to HIV infected populations. Furthermore, integrated analyses that consider gut microbiome composition together with data on diet, systemic immune activation, metabolites and demographic data have been lacking. By conducting a systems level analysis of predictors of metabolic disease in people living with HIV and men who are at high risk of acquiring HIV, we found that increased LBP, an inflammatory marker indicative of compromised intestinal barrier function, was associated with worse metabolic health. We also found this relationship to be associated with dietary, microbial, and metabolic factors suggesting a systemic gut microbiome influence on the presence of increased inflammatory markers which, in turn, influences the risk of metabolic disease. This work lays the framework for mechanistic studies aimed at targeting the microbiome and diet to prevent or treat metabolic endotoxemia in HIV-infected individuals.

Keywords

HIV; microbiome; men who have sex with men (MSM); metabolic disease

Background

The gut microbiome in people living with human immunodeficiency virus type 1 (HIV) (PLWH) is of interest as a potential contributor to infection, disease progression, and development of co-morbidities. Poor metabolic health characterized by insulin resistance and dyslipidemia is frequent in PLWH (1-3) and has been linked with chronic inflammation (4-7) and several anti-retroviral therapy (ART) drugs (8). Metabolic disease is particularly prevalent in HIV-positive individuals with lipodystrophy (LD), a disease linked with early ART drugs that is manifested by lipoatrophy.
in the face, extremities, and buttocks with or without visceral fat accumulation. LD can have a severe impact on the quality of life of PLWH and is associated with the development of diabetes and cardiovascular disease (9).

Metabolic disease has been linked with gut microbiome structure and function outside the context of HIV infection (10-14), but this relationship has not been explored deeply in PLWH. We and others have found an altered gut microbiome composition in both PLWH (15-17) and men who have sex with men at high-risk of contracting HIV (HR-MSM) (16, 18). Furthermore, we have demonstrated that the altered microbiome in HIV (15) and HR-MSM (15, 19) are pro-inflammatory both in vitro and/or in gnotobiotic mice (15, 19). This is of interest as peripheral inflammatory signals have been implicated in both cardiovascular disease risk (7, 20) and insulin sensitivity (4, 5, 21-23) in PLWH. Increased peripheral immune activation in HIV-positive individuals is driven in part by bacterial translocation (24, 25), as indicated by higher levels of the bacteria product lipopolysaccharide (LPS) or LPS-binding protein (LBP) in blood. Increased blood LPS levels have also been observed in MSM and linked with recent sexual behavior (26). An association between LBP and metabolic disease in other diseases (e.g. hemodialysis patients) has been described (27), however there are mixed data regarding a role in obesity associated metabolic disease (28-30). Additionally, a recent study of metabolic syndrome in PLWH found greater immune dysfunction and a more HIV-associated microbiome associated with risk of metabolic syndrome (31).

We hypothesized that PLWH and HR-MSM with poor metabolic health would harbor a distinct gut microbial signature that was in turn also associated with elevated peripheral immune activation. We evaluated this relationship while considering other factors known to influence the microbiome, immunity and metabolic health. This analysis included typical diet; HIV, ART, and LD status; and other demographic characteristics such as age and body mass index (BMI). For HIV-positive
individuals on ART with and without LD, we further explored associations with the plasma metabolome (Figure 1). Our results suggest a central role of inflammatory processes linked with bacterial translocation as measured by LBP, and co-correlated intestinal microbes, dietary and demographic attributes in metabolic disease risk.

Results

Study Population

This study examined a cohort of 113 men, including men who have sex with women (MSW; n=22, 19.5%) and men who have sex with men (MSM; n=91, 80.5%) (Table 1). Of the MSM, 32 were HIV-negative (35.2%), 14 were HIV-positive and not on ART (15.4%), and 45 were HIV-positive ART-treated (49.4%). The HIV-positive, treated group included those with lipodystrophy (LD; n=25, 55.6%) and those without (n=20, 44.4%). The HIV-negative MSM participated in activities that put them at high risk of contracting HIV as defined in a prior study of a candidate HIV vaccine: 1) a history of unprotected anal intercourse with one or more male or male-to-female transgender partners; 2) anal intercourse with two or more male or male-to-female transgender partners; or 3) being in a sexual relationship with a person who has been diagnosed with HIV (32). In order to focus on HIV-associated metabolic disease, obese individuals (BMI >30) were excluded. There was no significant difference in BMI between the cohorts (Kruskal-Wallis test, p = 0.085). HIV-positive, treated cohorts were significantly older than HIV-negative MSM and HIV-positive, untreated MSM (Kruskal-Wallis test, p < 0.001). Age matching across all cohorts was not feasible in part because LD is associated with early-generation ART drugs and thus most common in older HIV-positive individuals and HR-MSM behavior as well as new HIV infections are predominantly in younger individuals. However, age is carefully considered in downstream analyses. All treated, HIV-positive individuals were on successful ART with suppressed viral loads (Table 1).
Figure 1. Study design schematic. Measures were collected from four compartments: gut microbiome, peripheral immune, diet questionnaire, and plasma metabolome. These separate compartments can all influence each other, forming complex systems that together influence metabolic health. Furthermore, the compartments can be influenced by other clinical and demographic characteristics such as HIV and treatment status. In this study we examine all of these measures together in order to investigate metabolic health.
|                  | HIV-negative MSW | HIV-negative MSM | HIV-positive MSM, untreated | HIV-positive MSM, treated | HIV-positive MSM, treated, with LD |
|------------------|------------------|------------------|-----------------------------|---------------------------|----------------------------------|
| n                | 22               | 32               | 14                          | 20                        | 25                               |
| Age (years)      | 33 (27.3-38.5)   | 34 (29.8-44.5)   | 34 (26.5-40.3)              | 46 (42.8-50.5)            | 60 (54-64)                       |
| BMI (kg/m2)      | 25.2 (23.0-27.0) | 25.5 (20.2-28.0) | 21.4 (20.2-25.6)            | 23.9 (22.6-26.2)          | 25.8 (23.0-28.0)                 |
| CD4 cell count   | NA               | NA               | 538 (408.5-731.8)           | 586 (419.5-878.0)         | 659 (550.0-908.0)                |
| Viral load       | NA               | NA               | 101,400 (20,300-292,514)    | 20 (0-20)                 | 0 (0-20)                         |
| Cholesterol drugs/statins n (%) | 2 (9.1%) | 3 (9.4%) | 1 (7.1%) | 4 (20%) | 14 (56%) |

Numbers reported are median (IQR)
Metabolic Disease Score as a Marker for Metabolic Health

We measured seven common clinical markers of metabolic health from fasting blood: triglycerides, glucose, insulin, LDL, HDL, leptin, and adiponectin. Since these markers are often correlated with each other, we used principal component analysis (PCA) to define a single continuous measure of overall metabolic health of study participants, as has been done previously (Figure 2A) (33, 34). The first principal component (PC1) explained 28.8% of the variability within the clinical marker data and separated individuals by multiple correlated markers of metabolic disease. Specifically, individuals with higher values of PC1 generally had high triglycerides and low HDL, indicating dyslipidemia, and higher levels of fasting blood glucose and insulin, indicating insulin resistance (Figure 2A, Supplemental Figure 1). We thus decided to use values of PC1 as an outcome. Values of PC1 were shifted to a minimum of one and log transformed to define the metabolic disease score, which ranged from 0 as healthy and 2.5 as impaired. In order to determine how this score related to metabolic health, we performed regressions between the metabolic disease score and individual measures to define a metabolic score threshold that corresponded with clinically defined cutoffs for normal levels (Supplemental Figure 1). For example, triglycerides positively correlated with metabolic disease score and almost all individuals with a score above 1.45 had triglyceride levels in the unhealthy range of greater than 200 mg/dL. Similar patterns and cutoffs were true for HDL, LDL, and glucose (Supplemental Figure 1). The intersect of the regression with these cutoffs were all averaged to a single number of 1.4. Individuals below the cutoff were categorized as metabolically normal and those above were categorized as metabolically impaired.

When comparing the metabolic disease score across cohorts, we found that ART-treated, HIV-positive individuals with LD trended higher in both the average metabolic disease score and the proportion of individuals with scores in the metabolically impaired group but intergroup
Figure 2. Metabolic disease score for marker of metabolic health. 

A. PCA of metabolic measures in fasting blood of 164 men and women: 113 participants described in this paper along with 51 individuals recruited at the same time and under the same exclusion criteria as study participants. Metabolic disease score is calculated as the PC1 coordinates shifted to a minimum of one and log transformed.

B. Metabolic disease scores broken up by cohort. The percentages noted above groups are the percent of individuals with a score above our metabolic impairment cutoff (Supplemental Figure 1). There is no significant difference between the proportions in each group (Fischer’s exact test, p = 0.11) or between mean ranks in each group (Kruskal-Wallis test p = 0.13).

C. Relationships between metabolic disease score and age stratified by cohort. Statistical significance of slopes are indicated and were calculated with the linear model: score ~ age + cohort + age*cohort. P-value annotations: ** < 0.01; * < 0.05.
significance was lost after multiple test corrections (Figure 2B). Furthermore, because our HIV-positive, treated cohorts were significantly older than our HIV-negative MSM and HIV-positive, untreated MSM, we used a linear model to explore differences in the metabolic disease score across cohorts while accounting for age (Figure 2C). This score was positively associated with age only in HIV-negative MSM and HIV-positive, untreated MSM (Figure 2C; linear model; p < 0.001 and p = 0.036 respectively) and only HIV-negative MSM had significantly higher metabolic disease score compared to HIV-negative MSW when accounting for age (linear model; p < 0.001).

**Selection of Features that Predict the Metabolic Disease Score and Interactions Between Selected Features**

To explore the complex relationships of the gut microbiome, peripheral immune activation, and diet to the metabolic disease score and to each other, we first selected features that were important predictors of the metabolic score using the tool VSURF (Variable Selection Using Random Forest) tool (35). The VSURF implementation of random forest is optimized for feature selection, returning all features that are highly predictive of the response variable, even when a smaller subset of highly predictive variables with redundant features removed could be just as accurate for prediction (35). We input the following features into the VSURF tool: 1) 130 microbial features (99% identity Operational Taxonomic Units (OTUs) with highly co-correlated OTUs binned into modules as described in the methods (detailed in Supplemental Table 1) and filtered to OTUs only in >20% of samples). 2) 21 immune features that were measured in plasma using multiplex ELISA (detailed in Supplemental Table 2). These immune measures were selected based on a literature search for those previously shown to be altered in HIV infection and/or metabolic disease. 3) 21 clinical/demographic features that were collected in questionnaires or measured in study participants such as age, BMI, HIV infection and treatment status, typical gastrointestinal symptoms including constipation, diarrhea and bloating, and sexual behavior.
(detailed in Supplemental Table 3). 4) 29 dietary features that were collected using a food frequency questionnaire of typical dietary intake over the prior year as detailed in the methods.

From the initial 201 measures, VSURF identified 69 important variables (four clinical data measures, six diet measures, 14 immune measures, 45 microbes) and a subset of ten highly predictive variables (Supplemental Table 4). These 69 features were sufficient to accurately predict metabolic disease score using traditional random forest (linear model: $r^2 = 31.05\%$, $p < 0.001$). Additionally, permutation testing revealed that VSURF performed better at selecting explanatory variables than a null model where the outcome was randomly permuted (permutation test; $p = 0.049$, Supplemental Figure 2). We found that 21 of the 69 selected variables were positively or negatively correlated with the metabolic disease score, indicating either increased or decreased risk respectively (Spearman rank correlation, FDR $p < 0.1$, Supplemental Table 4). Since random forest can detect non-linear relationships and/or features that are only important when also considering another feature, it is not surprising that all features were not correlated linearly with the metabolic disease score.

All VSURF selected clinical measures were positively correlated with metabolic disease score and included age, BMI, lipodystrophy, and bloating (Supplemental Table 4). None of the six selected diet measures correlated with metabolic disease score (Supplemental Table 4). VSURF selected several inflammatory immune measures that were positively correlated with metabolic disease score: LBP, intercellular adhesion molecule 1 (ICAM-1), interleukin (IL) 16, IL-12, and granulocyte-macrophage colony-stimulating factor (GM-CSF) (Supplemental Table 4). The most important feature as determined by random forest importance score was LBP.

Diet, the microbiome, and immune phenotypes can all influence each other (Figure 1). They can also relate to the measured clinical/demographic factors that we had identified as predictors of
the metabolic disease score such as BMI and age. For this reason, we also investigated the relationship between the 69 important factors using pairwise Spearman rank correlation and network visualization (Figure 3, Supplemental Figure 3, Supplemental Table 5). This approach revealed many within data type associations such as positive correlations within many selected dietary, microbiome, and immune features. It also identified correlations between data types such as a negative relationship between LBP and several Gram-negative bacteria or a positive correlation between age and immune measures such as LBP or IL-6 (Figure 3A, Supplemental Table 5).

The selected important microbes included many that positively or negatively correlated with the metabolic disease score (Supplemental Table 4) and that were highly co-correlated with each other and with dietary, clinical/demographic, and inflammatory phenotypes (Supplemental Figure 3A, Supplemental Table 5). For example, a module of bacteria identified within the *Prevotella* genus and the *Paraprevotellaceae* family, negatively associated with metabolic disease score and positively associated with dietary fiber (Supplemental Figure 3B). Because bacterial translocation is known to occur at increased levels in both HIV-positive individuals (36) and HR-MSM (26), we were specifically interested in looking at the selected microbes and other features that correlated with LBP, a marker of bacterial translocation. The network of LBP neighbors is shown in Figure 3B. All of the microbes correlating with LBP were classified to the order *Clostridiales*. More specifically, LBP is negatively correlated with several butyrate or putative butyrate producing bacteria/bacterial modules such as OTUs in the genera *Coprococcus* (37, 38). LBP was also positively correlated with *Dorea* species (Supplemental Table 5).

In addition to correlations, we evaluated interactions between variables using the tool iRF (iterative Random Forest)(39). These interactions represent variables that are in adjacent nodes
Figure 3. Networks of selected measures reveal several strong associations with metabolic disease score and between measures. Correlation sub-networks of A, all the non-microbe selected measures. B, the nearest neighbors of LBP. All Spearman rank correlations with an FDR p < 0.25 are shown. Subnetworks were pulled from a larger network of all VSURF selected measures (Supplemental Figure 3, Supplemental Table 2). C, Network of interactions between measures calculated using iRF. All edges represent an interaction (i.e. proximity in a decision tree) that occurred in 30% or more of the decision trees.
Interactions between variables in more than 30% of the trees were kept for further analysis (Figure 3C). This analysis identified a group of 5 interactive features: LBP, age, BMI, an OTU in the \textit{Lachnospiraceae} family, and microbiome module 24 (\textit{Coprococcus} sp. and \textit{Blautia} sp.). These features were also significantly correlated with LBP and suggest a subset of features that when taken together may be predictive of metabolic disease risk.

The Plasma Metabolome as a Potential Mechanism of Microbial Influence on Metabolic Disease Score in ART-treated HIV-positive Individuals with and without LD

To pursue a further mechanistic understanding of how the gut microbiome may influence the metabolic disease score in PLWH, we performed untargeted metabolomics (LC/MS) on plasma from our cohort of ART-treated, HIV-positive individuals with and without LD (n=44). Metabolite identities were then validated using untargeted MS/MS. We used two approaches to determine which plasma metabolites were either directly produced or indirectly influenced by the presence of a microbiome. First, metabolites found in the human plasma were run through the computational tool AMON (40), which uses the KEGG database (41) and inferred metagenomes (calculated using PICRUSt2 (42)) to determine which of the measured metabolites could have been produced by the microbiome. Second, plasma from both germ free (GF) and humanized mice was analyzed using metabolomics to determine metabolites that had significantly altered levels in mice with human microbiomes compared to GF mice, i.e. microbiome influenced metabolites. For this purpose, GF mice were gavaged using fecal samples from eight men from the study cohort (humanized mice) while two mice were gavaged using PBS as control (Supplemental Table 6). Plasma was collected before and after gavage. All mice were fed a high-fat western diet.
We found that 820 metabolites were different in abundance between GF and humanized mice after multiple test corrections (Student’s t-test, FDR p < 0.05), 493 of which were also present in the human plasma samples. However, only 376 of these 493 metabolites could be annotated (see Methods and Supplemental Table 7) while the remaining 148 were only assigned a mass. From the full set of 5,332 metabolites identified in the human plasma, 416 were able to be annotated with KEGG IDs. These were further analyzed using AMON. 146 microbiome-associated metabolites were identified that are putatively produced by the gut microbiome; however, many of these could also be produced by the host. Twenty-six of the 134 microbiome-associated metabolites identified by AMON were also identified in the gnotobiotic mouse analysis (Supplemental Table 7).

Of the 5,332 total measured metabolites in the human samples, 150 correlated with metabolic score (Spearman rank correlation, FDR p < 0.05; Supplemental Table 8). Of these 68 could not be annotated. The annotated correlated compounds were enriched in a number of different metabolic pathways with both the Phospholipid and the Glycerolipid pathways of the Small Molecule Pathway Database (SMDPH)(43) being highly enriched (Supplemental Table 9). Consistent with the metabolic score being defined in part by dyslipidemia, 17 of the significant compounds were annotated as triglycerides. Of the 150 correlated metabolites, seven were associated with the microbiome either because they were predicted microbial products of the gut microbiome (as determined with AMON (40)), because they were significantly different while comparing the metabolome of germ-free mice to that of mice colonized with the feces of study participants, or by literature search. We confirmed the identity of 5 of the 7 of these with MS/MS (Supplemental Table 8). Of these seven microbiome-associated metabolites, two could exclusively be explained by direct production by the microbiome. Specifically, dehydroalanine was identified as a microbial product with AMON and negatively correlated with metabolic disease score (Supplemental Table 8). Bacteriohopane-32,33,34,35-tetrol is a bacterial metabolite that
positively correlated with the metabolic disease score. Two additional microbiome-associated compounds were triglycerides (TG(54:6) and TG (16:0/18:2,20:4)) that were positively correlated with metabolic disease score and elevated in humanized mice compared to GF mice. Another of these metabolites, 1-Linoleyl-2-oleoyl-rac-glycerol is a 1,2-diglyceride in the triglyceride biosynthesis pathway. Finally, phosphatidylcholine (PC(17:0/18:2)) and phosphatidylethanolamine (PE(20:3/18:0)) compounds were identified as microbiome associated and positively and negatively correlated with the metabolic disease score respectively (44).

Discussion

In this study, we identified several bacterial, diet, and immune measures that predicted higher metabolic disease score in a cohort of MSM with and without HIV, ART, and LD. Notably, we identified a strong relationship between circulating LBP and higher metabolic disease score which in turn correlated with other markers of systemic inflammation, a loss of beneficial microbes such as Gram-positive, butyrate-producing bacteria, and higher BMI, indicating that diverse modifiable factors may influence LPS/inflammation driven metabolic disease in this population.

There was a positive association between metabolic disease score and age, as has been reported previously for non-HIV populations (45), but linear modeling suggested that this relationship was driven by an association in HIV-negative MSM and HIV-positive untreated MSM in our study, revealing a possibly larger effect size than in our other cohorts. Also, when controlling for age, HIV-negative and positive untreated MSM had the highest metabolic disease score, even compared to HIV positive individuals on ART with LD, a population that has previously been reported to have higher incidence (46). This result is intriguing given our results supporting a role for LBP driven inflammation in metabolic disease and prior research linking increased levels of LPS in blood with high-risk behavior in MSM (26). Larger cohorts and more detailed behavior
information are required, however, to make any definitive claims on impaired metabolic health in ageing in HR-MSM.

Consistent with prior studies that have associated high BMI with dyslipidemia, insulin resistance, and/or metabolic syndrome (9, 47, 48), BMI was a positive predictor of metabolic disease score in our cohort even though our study excluded obese individuals, but did include overweight. This suggests the importance of weight management even among overweight, non-obese individuals as a strategy for reducing metabolic health impairment in this population.

We did not find a positive association between ART treatment status and metabolic disease score, but this may be because individuals in our study were on a wide variety of drug combinations with the potential to have varied/contrasting effects. For instance, both integrase stand transfer inhibitors (ISTI) (49) and regimens including the nucleoside reverse transcriptase inhibitor (NRTI) tenofovir have been shown to increase risk of weight gain (50). Conversely, the CCR5 antagonist, maraviroc, may confer a benefit to cardiovascular function and body weight maintenance and evidence in mice suggests that this may be linked to differences in gut microbiome composition with treatment (21, 51). Thus, future studies more targeted to particular ART regimes will be required to look at factors important in particular drug contexts.

Several of the dietary components identified as important predictors of metabolic disease score in our cohort have been previously associated with metabolic health, including dietary carotenoid, lycopene, and fiber (52-56). Fiber’s benefit in glucose response has been linked with the activity of Prevotella copri. Individuals who had improved glucose response upon 3 days of a high-fiber diet consumption were characterized by a higher increase in P. copri (55) and beneficial effects of P. copri were confirmed in mice fed a high-fiber diet (55). Interactions between Prevotella, dietary fiber, and metabolic health were of particular interest in this cohort of HIV positive and
negative MSM since these individuals have much higher Prevotella, including *P. copri*, than in non-MSM (16, 18). However, other published studies suggested that high Prevotella might predict increased risk of metabolic disease. One group observed that *P. copri* in mice fed a western diet low in fiber could promote poor glucose response through the production of branched chain amino acids (BCAAs) (12). Additionally, our prior study using *in vitro* stimulations of human immune cells with fecal bacteria of HIV positive and negative MSM indicated that the Prevotella-rich microbiomes of MSM could drive systemic inflammation (15). Interestingly, in our data, a module of three OTUs, two of which are identified as within the genus *Prevotella* and the other within the family *Paraprevotellaceae*, negatively associated with metabolic disease score and positively associated with dietary fiber (Supplemental Figure 3B), supporting a relationship between particular *Prevotella* strains and dietary fiber towards improved metabolic health, and not supporting deleterious effects. Further work will be needed to decompose the complex relationship between dietary fiber, particularPrevotella strains, and metabolic health in HIV positive and negative MSM with unique Prevotella-dominated communities.

LBP was the most important feature in the random forest analysis and also a highly interactive measure in the iRF analysis. LBP binds to both microbial LPS and lipoteichoic acid (LTA) in the blood (57) and the presence of elevated LBP is indicative of increased intestinal barrier permeability (58). LBP levels were correlated with other inflammatory markers that have been linked with worse metabolic health in HIV-negative populations suggesting a role as a central mediator of metabolic-disease associated immune phenotypes. These included 1) I-CAM 1, whose expression in adipose tissue has been associated with diet-induced obesity in mice (59) and metabolic syndrome in humans (60) 2) IL-6, a pro-inflammatory cytokine that has been shown to play a direct role in insulin resistance (61), and 3) SAA, which is regulated in part by IL-6 and plays a role in cholesterol metabolism (60); SSA3 specifically has been shown to be produced in response to gut bacteria in obesity mice (62). We observed a positive association between

Armstrong et al. 18
metabolic disease score and frequency of abdominal bloating, further supporting a role of intestinal dysfunction in this population. Taken together these associations suggest that inflammation originating from an impaired intestinal barrier is promoting worse metabolic health.

Although prior studies have connected LBP-associated inflammation with worse metabolic health (27, 28, 63); the strength of this relationship is disease specific with less clear results in obesity-associated metabolic disease (29, 30). An importance of bacterial translocation in HIV-associated metabolic syndrome was demonstrated in a recent study of metabolic comorbidities in HIV-positive individuals which found that lower CD4 nadir and/or AIDS events, HIV-associated microbiota, and low alpha diversity was correlated with increases in sCD14 and LBP and increase risk of metabolic syndrome (31). Additionally, in our study LBP was correlated with age and BMI, a relationship that was previously observed in a cohort of HIV-negative men of African ancestry with this trio being further associated with adiposity and pre-diabetes (64). Lastly, the negative association of LBP with putative butyrate producing bacteria suggests that a lack of microbes that promote intestinal barrier integrity contributes to increased intestinal permeability and thus microbial components in circulating blood.

In our metabolomic analysis, we identified 150 metabolites in blood that correlated with metabolic disease score. In order to identify compounds whose prevalence may be related to the gut microbiome we used two complimentary approaches. First, we predicted which of these compounds could have been produced by the microbiome using information in KEGG and the bioinformatics tool AMON (40). Second, we measured which compounds changed in relative abundance in germ-free versus mice colonized with feces from our study cohort. The AMON analysis allows us to specifically evaluate which compounds could have been directly produced by the gut microbiome but is limited by a lack of KEGG annotations for many compounds (40). The gnotobiotic mouse experiments can identify microbial influence in unannotated compounds.
but cannot differentiate between direct production/consumption by microbes versus indirect influence. The results will also be influenced by physiological differences between mice and humans and the incomplete colonization of human microbes in humanized mice. Although these weaknesses may have led us to underestimate which of the 150 metabolic disease associated compounds may have been related to the microbiome, it still identified compounds that supported a mechanistic link between gut microbes, metabolites, and metabolic disease in HIV-infected individuals on ART.

Firstly, we found a negative correlation between the microbially-produced non-canonical amino acid, dehydroalanine, and metabolic disease score. Dehydroalanine is a component of lantibiotics that are active against Gram-positive bacteria. We observed Gram-positive Dorea to positively correlate with LBP, suggesting a role of lantibiotics in regulating our proposed LBP-centered metabolic disease in this population.

Secondly, we observed that bacteriohopane-32,33,34,35-tetrol positively correlated with the metabolic disease score. This compound has been found to be a lipoxygenase inhibitor that prevents the formation of hydroxyicosatetraenoic acid and various leukotrienes from arachidonic acid (65), which have been linked with the development of cardiovascular disease and metabolic syndrome (66). This association of a potentially protective metabolite increased in metabolic impairment seems counterintuitive; however, it may be indicative of larger systemic changes in arachidonic acid metabolism and is worthy of further exploration.

Thirdly, we identified a PC and a PE associated with both the microbiome and metabolic disease score. Changes in PCs and/or PEs have been previously implicated in atherosclerosis, insulin resistance and obesity (44). AMON analysis indicated that both PCs and PEs can be synthesized by intestinal bacteria; however, these compounds can also be synthesized in the host and may
be found in the diet. In our analysis, PE(18:1/20:1) levels were higher in colonized compared to
germin-free mice indicating that intestinal bacteria do influence overall levels despite diverse
potential sources.

Lastly, we observed increased levels of several plasma triglycerides in the humanized compared
to gern-free mice, including two plasma triglycerides that were significantly associated with
metabolic disease score. This confirms the influence of the gut microbiome on host plasma
triglycerides (67-69). However, we did not find any strong associations between these
triglycerides and specific microbes within our dataset, indicating a potential need for studies
conducted in larger cohorts or with shotgun metagenomics to look for functional correlates.

In conclusion, we observed a relationship between diet, gut microbiome, plasma metabolome,
and peripheral immune markers of inflammation and metabolic disease in MSM. These data pull
together several previously described relationships between pairs of these compartments
observed in other populations. However, in this work we demonstrate both a novel collection of
measures (microbiome, peripheral immune signaling, peripheral metabolites, demographic and
diet information) and a novel approach for integrating several host compartments in order to
examine a more complex system and applied it to the little studied population of HIV negative and
positive MSM with and without LD. Our results suggest that an overall gut environment driven by
low fiber, key vitamins, and microbes that promote intestinal barrier integrity and high in potentially
pathogenic bacteria may work in conjunction to increase levels of LBP and other inflammatory
cytokines to drive poor metabolic health. These results illuminate potential microbiome-targeted
therapies and personalized diet recommendation given an interacting set of gut microbes and
other host factors. Understanding these relationships further may provide novel treatments to
improve the metabolic disease and inflammatory outcomes of MSM living with HIV.
Methods

Subject Recruitment

Participants were residents of the Denver, Colorado metropolitan area and the study was conducted at the Clinical Translational Research Center of the University of Colorado Hospital. The study was reviewed and approved by the Colorado Multiple Institutional Review Board and informed consent was obtained from all participants. For detailed criteria on recruitment of our five cohorts (HIV negative MSW; HIV negative MSM; HIV positive, ART naïve MSM; HIV positive ART-treated MSM with LD; and HIV positive ART-treated MSM without LD) see supplemental methods.

Feces, a fasting blood sample, and clinical surveys were collected from participants in order to obtain analytes for the study design outlined in Figure 3 (Supplemental Table 2). To evaluate metabolic health, we measured seven common clinical markers from fasting blood: triglycerides, glucose, insulin, LDL, HDL, leptin, and adiponectin. Additional information about relevant clinical measures such as probiotic use were also collected via a questionnaire and study participants also filled out information on typical frequency of high-risk sexual practices and on typical levels of gastrointestinal issues such as bloating, constipation, nausea and diarrhea.

Diet Data FFQ Collection

Typical dietary consumption over the prior year was collected using Diet History Questionnaire II (70). Diet composition was processed using the Diet*Calc software and the dhq2.database.092914 database (71). All reported values are based on USDA nutrition guidelines. Reported dietary levels were normalized per 1000 kcal. To reduce the number of comparisons within the diet survey data, we binned highly co-correlating groups of measures within the data types into modules (Supplemental Table 3). These modules were defined using the tool, SCNIC (72).
Immune Data Collection

Whole blood was collected in sodium heparin vacutainers and centrifuged at 1700rpm for 10 minutes for plasma collection. Plasma was aliquoted into 1mL microcentrifuge tubes and stored at -80. For ELISA preparation, plasma was thawed, kept cold, and centrifuged at 2000xg for 20 minutes before ELISA plating. Markers for sCD14, sCD163, and FABP-2 were measured from plasma using standard ELISA kits from R&D Systems (DC140, DC1630 & DFBP20). Positive testing controls for each ELISA kit were also included (R&D Systems QC20, QC61, & QC213). LBP was measured by standard ELISA using Hycult Biotech kit HK315-02. Markers for IL-6, IL-10, TNF-α, MCP-1, and IL-22 were measured using Meso Scale Discovery’s U-PLEX Biomarker Group 1 multiplex kit K15067L-1. Markers for SAA, VCAM-1, ICAM-1, and CRP were measured using Meso Scale Discovery’s V-Plex Plus Vascular Injury Panel 2 multiplex kit K15198G-1. Vascular Injury Control Pack 1 C4198-1 was utilized as a positive control for this assay. Markers for GM-CSF, IL-7, IL-12/23p40, IL-15, IL-16, IL-17A, TNF-β, and VEGF were measured using Meso Scale Discovery’s V-Plex Plus Cytokine Panel 1 multiplex kit K151A0H-1. Cytokine Panel 1 Control Pack C4050-1 was utilized as a positive control for this assay. Plasma samples were diluted per manufacturer’s recommendation for all assays. Standard ELISA kit plates were measured using a Vmax® Kinetic Microplate Reader with Softmax® Pro Software from Molecular Devices LLC. Multiplex ELISA kits from Meso Scale Discovery were measured using the QuickPlex SQ 120 with Discovery Workbench 4.0 software.

Gnotobiotic Mouse Protocols

Germ-free C57/BL6 mice were purchased from Taconic and bred and maintained in flexible film isolator bubbles, fed with standard mouse chow. Three days before they were gavaged, male mice between 5-7 weeks of age were switched over to a western high-fat diet and were fed this diet for the remainder of the experiment. Diets were all obtained from Envigo (Indiana): Standard
chow - Teklad global soy protein-free extruded (item 2920X - https://www.envigo.com/resources/data-sheets/2020x-datasheet-0915.pdf), Western Diet – New Total Western Diet (item TD.110919). See Supplemental Table 10 for detailed diet composition.

Mice were gavaged with 200 μL of fecal solutions prepared from 1.5 g of donor feces mixed in 3 mL of anaerobic PBS (19). Mice were housed individually following gavage for three weeks in a Tecniplast iso-positive caging system, with each cage having HEPA filters and positive pressurization for bioexclusion. Feces were collected from mice at day 21 for 16S rRNA gene sequencing. Mice were euthanized at 21 days post gavage using isoflurane overdose and all efforts were made to minimize suffering. Blood from euthanized animals was collected using cardiac puncture and cells were pelleted in K2-EDTA tubes; plasma was then aliquoted and stored at -80 °C.

Metabolomics Methods

Plasma Sample Preparation

A modified liquid-liquid extraction protocol was used to extract hydrophobic and hydrophilic compounds from the plasma samples (73). Briefly, 50 μL of plasma spiked with internal standards underwent a protein crash with 250 μL ice cold methanol. 750 μL methyl tert-butyl ether (MTBE) and 650 μL 25% methanol in water were added to extract the hydrophobic and hydrophilic compounds, respectively. 500 μL of the upper hydrophobic layer and 400 μL of the lower hydrophilic layer were transferred to separate autosampler vials and dried under nitrogen. The hydrophobic layer was reconstituted with 100 μL of methanol and the hydrophilic layer was reconstituted with 50 μL 5% acetonitrile in water. Both fractions were stored at -80 °C until LC/MS analysis.

Liquid Chromatography Mass Spectrometry
The hydrophobic fractions were analyzed using reverse phase chromatography on an Agilent Technologies (Santa Clara, CA) 1290 ultra-high precision liquid chromatography (UHPLC) system on an Agilent Zorbax Rapid Resolution HD SB-C18, 1.8um (2.1 x 100mm) analytical column as previously described (73, 74). The hydrophilic fractions were analyzed using hydrophilic interaction liquid chromatography (HILIC) on a 1290 UHPLC system using an Agilent InfinityLab Poroshell 120 HILIC-Z (2.1 x 100mm) analytical column with gradient conditions as previously described (75) with mass spectrometry modifications as follows: nebulizer pressure: 35psi, gas flow: 12L/min, sheath gas temperature: 275C, sheath gas flow: 12L/min, nozzle voltage: 250V, Fragmentor: 100V. The hydrophobic and hydrophilic fractions were run on Agilent Technologies (Santa Clara, CA) 6545 Quadrupole Time of Flight (QTOF) mass spectrometer. Both fractions were run in positive electrospray ionization (ESI) mode.

**Mass Spectrometry Data Processing**

Compound data was extracted using Agilent Technologies (Santa Clara, CA) MassHunter Profinder Version 10 software in combination with Agilent Technologies Mass Profiler Professional Version 14.9 (MPP) as described previously (40). Briefly, Batch Molecular Feature Extraction (BMFE) was used in Profinder to extract compound data from all samples and sample preparation blanks. The following BMFE parameters were used to group individual molecular features into compounds: charge state 1-2, with +H, +Na, +NH4 and/or +K charge carriers. To reduce the presence of missing values, a theoretical mass and retention time database was generated for compounds present in samples only from a compound exchange format (.cef) file. This .cef file was then used to re-mine the raw sample data in Profinder using Batch Targeted Feature Extraction.

An in-house database containing KEGG, METLIN, Lipid Maps, and HMDB spectral data was used to putatively annotate metabolites based on accurate mass (≤ 10 ppm), isotope ratios and isotopic...
distribution. This corresponds to a Metabolomics Standards Initiative metabolite identification level three (76). To improve compound identification, statistically significant compounds underwent tandem MS using 10, 20, and 40V. Fragmentation patterns of identified compounds were matched to either NIST14 and NIST17 MSMS libraries, or to the \textit{in silico} libraries, MetFrag (77) and Lipid Annotator 1.0 (Agilent) (78).

\textit{Microbiome-associated metabolites}

Microbiome-associated metabolites were defined using metabolites identified as significantly different in abundance between germ-free compared to humanized gnotobiotic mice and/or metabolites identified as microbially produced by the tool AMON (40).

For the gnotobiotic mouse analysis aqueous and lipid metabolites were analyzed separately (see mouse protocol above for details on experimental set-up). Metabolites that were present in <20\% of samples were filtered out before analysis. Significant difference was determined using a Student's t-test with FDR p-value correction. FDR-corrected p values < 0.05 were deemed significant. Significant metabolites also present in the human samples were retained for further analysis.

For the AMON-identified metabolites, the tool used an inferred metagenome, which was calculated using the PICRUST2 QILME2 plugin (42) and default parameters; a list of all identified KEGG IDs from the metabolite data (see metabolome methods); and KEGG flat files (downloaded 2019/06/10). AMON determined metabolites observed that could be produced by the given genome. These metabolites were kept for analysis in addition to the gnotobiotic mouse identified metabolites. Those without any putative classification were removed from analysis.

\textit{Microbiome Methods}
Sample Collection, Extraction, and Sequencing

Stool samples were collected by the patient within 24 hours prior to their clinic visit on sterile swabs dipped into a full fecal sample deposited into a commode specimen collector. Samples were kept cold or frozen at -20°C during transport prior to being stored at -80°C. DNA was extracted using the standard DNeasy PowerSoil Kit protocol (Qiagen). Extracted DNA was PCR amplified with barcoded primers targeting the V4 region of 16S rRNA gene according to the Earth Microbiome Project 16S Illumina Amplicon protocol with the 515F:806R primer constructs (79). Control sterile swab samples that had undergone the same DNA extraction and PCR amplification procedures were also processed. Each PCR product was quantified using PicoGreen (Invitrogen), and equal amounts (ng) of DNA from each sample were pooled and cleaned using the UltraClean PCR Clean-Up Kit (MoBio). Sequences were generated on six runs on a MiSeq sequencing platform (Illumina, San Diego, CA).

Microbiome Sequence Processing and Analysis

Microbiome processing was performed using QIIME2 version 2018.8.0 (80). Data was sequenced across five sequencing runs. Each run was demultiplex and denoised separately using the DADA2 q2 plugin (81). Individual runs were then merged together and 99% de novo OTUs were defined using vSEARCH (82). Features were classified using the skLearn classifier in QIIME2 with a classifier that was pre-trained on GreenGenes13_8 (83). The phylogenetic tree was building using the SEPP plugin (84). Features that did not classify at the phylum level or were classified as mitochondria or chloroplast were filtered from the analysis. Samples were rarefied at 19,986 reads. To reduce the number of comparisons within the microbiome, we binned highly co-correlating groups of measures within the data types into modules (Supplemental Table 1). These modules were defined using the tool, SCNIC (72). For statistical analysis features present in <20% of samples were filtered out.
Module definition

Modules were called on microbiome and diet data. Modules were defined using the tool SCNIC (72). The q2-SCNIC plugin was used with default parameters for the microbiome data and standalone SCNIC was used for the diet data (https://github.com/shafferm/SCNIC). Specifically, for each data type SCNIC was used to first identify pairwise correlations between all features. Pearson correlation was used for diet and SparCC (85), which takes into account compositionality, was used for microbiome data. Modules were then selected with a shared minimum distance (SMD) algorithm. The SMD method defines modules by first applying complete linkage hierarchical clustering to correlation coefficients to make a tree of features. Modules are defined as subtrees where all pairwise correlations between all pairs of tips have an R value greater than defined minimum. The diet modules were defined using a Pearson r^2 cutoff of 0.75. The microbiome modules were defined using a SparCC minimum r cutoff of 0.35. To summarize modules SCNIC uses a simple summation of count data from all features in a module. Application of SCNIC reduced the number of evaluated features from 6,913 to 6,818 for microbiome and 59 to 29 for diet data.

Statistical Analysis

All statistics were performed in R. For non-parametric tests Spearman rank correlation and Kruskal-Wallis test were used. For parametric tests linear models and Student's t-test were used.

Data analysis tools

Metabolic disease score was calculated using PCA in R with prcomp. Data was scaled using default method within the prcomp library. All random forest analysis tools were used in R. Standard random forest was performed using randomForest. Variable selection was performed in R using the tool VSURF (35). Interaction analysis was performed in R using the tool iRF (39).
Data Availability

All data will be publicly available upon publishing. Microbiome data in QIITA (https://qiita.ucsd.edu) Study ID 13338 and available upon request and will be publicly available in EBI/ENA (https://www.ebi.ac.uk/ena) upon publishing. Immune and diet data are available along with the microbiome data as associated metadata. Metabolomics data will be available on Metabolomic Workbench (https://www.metabolomicsworkbench.org) upon publishing. Until publicly available it is available upon request.

Statements

Acknowledgements

We would like to thank our study participants for contributing their samples and time to this study. We would also like to thank Christine Griesmer for her contributions to subject recruitment and Brandi Wagner and Maggie Stanislawski for insights into statistical analysis methods.

Funding sources

This work was funded by R01 DK104047 and R01 DK108366 with additional support from NIH/NCATS Colorado CTSA Grant Number UL1 TR002535. High performance computing was supported by a cluster at the University of Colorado Boulder funded by National Institutes of Health 1S10OD012300. AJS Armstrong was supported by T32 AI052066. Contents are the authors’ sole responsibility and do not necessarily represent official NIH views.

Declarations of interests

The authors declare that they have no conflicts of interest.

Author contributions

Armstrong et al. 29
AJSA analyzed and interpreted all data. AJSA and CAL wrote the manuscript. NR guided generation and interpretation of metabolomics data. KQ and KAD prepared, ran, and processed metabolomics. KQ ran metabolic pathway analysis. SXL prepared and conducted mouse experiments. JMS ran immunological assays. NMN prepared and ran sequencing and coordinated fecal sample and metadata collection from study subjects. SF recruited subjects, collected samples, and maintained regulatory compliance. TJM and JH collected and aided in interpretation and processing of diet data. CAL, BEP, and TC conceptualized and led the study. TC guided all clinical data collection and subject recruitment and provided clinical insight into study populations. BEP guided generation and interpretation of immune data. CAL guided microbiome data generation and multi’omic data analysis. All authors read and approved the final manuscript.
Supplemental Tables

Supplemental Table 1. Microbiome modules calculated by SCNIC

Supplemental Table 2. Study measures by datatype

Supplemental Table 3. Diet survey data modules calculated by SCNIC

Supplemental Table 4. VSURF-selected features and correlation with metabolic disease score

Supplemental Table 5. Edge table for VSURF-selected inter-variable correlations

Supplemental Table 6. Gnotobiotic mouse experiment set-up

Supplemental Table 7. Microbiome-associated metabolites list and source of identification

Supplemental Table 8. Metabolites correlating with metabolic score

Supplemental Table 9. mBrole pathway analysis results

Supplemental Table 10. Western diet for gnotobiotic mice
Supplemental Figure 1. Metabolic Score Cutoff Calculations using regressions between metabolic disease score and the metrics used in the PCA analysis. Triglycerides, fasting glucose, HDL and LDL have well-defined clinical cut-offs for high values and were used to calculate the healthy-unhealthy cutoff. Linear model was calculated modeling metabolic disease score by each marker. The high value intercept of the regression line is marked with a dotted line and value annotated on the plot. The solid line is the defined healthy-unhealthy cutoff calculated as the mean of the four cutoff values. P values are from the linear model.
Supplemental Figure 2. Histogram of percent variation explained in permuted VSURF.

Metabolic disease score was permuted 1,000 times and passed through VSURF. The resulting variables were run through a standard random forest and the percent variation explained was calculated. The blue line represents the percent variation explained for the true VSURF. P value was calculated using a one tailed test.
Supplemental Figure 3. Correlation network of VSURF-selected variables. Correlation network of A) all VSURF-selected variables and B) neighboring nodes of dietary fiber. All Spearman rank correlations with an FDR p < 0.25 are shown. See Supplemental Table 5 for the edge table and Supplemental Table 4 for the node table.
1. Non LR, Escota GV, Powderly WG. 2016. HIV and its relationship to insulin resistance and lipid abnormalities. Translational Research 183:41-56.

2. Nix LM, Tien PC. 2014. Metabolic syndrome, diabetes, and cardiovascular risk in HIV. Curr HIV/AIDS Rep 11:271-8.

3. Lake JE, Li X, Palella FJ, Jr., Erlandson KM, Wiley D, Kingsley L, Jacobson LP, Brown TT. 2018. Metabolic health across the BMI spectrum in HIV-infected and HIV-uninfected men. AIDS 32:49-57.

4. Willig AL, Overton ET. 2016. Metabolic Complications and Glucose Metabolism in HIV Infection: A Review of the Evidence. Curr HIV/AIDS Rep 13:289-96.

5. Monczor AN, Li X, Palella FJ, Jr., Erlandson KM, Wiley D, Kingsley LA, Post WS, Jacobson LP, Brown TT, Lake JE. 2018. Systemic Inflammation Characterizes Lack of Metabolic Health in Nonobese HIV-Infected Men. Mediators Inflamm 2018:5327361.

6. Funderburg NT, Mehta NN. 2016. Lipid Abnormalities and Inflammation in HIV Infection. Curr HIV/AIDS Rep 13:218-25.

7. Beltran LM, Rubio-Navarro A, Amaro-Villalobos JM, Egido J, Garcia-Puig J, Moreno JA. 2015. Influence of immune activation and inflammatory response on cardiovascular risk associated with the human immunodeficiency virus. Vasc Health Risk Manag 11:35-48.

8. Tsiodras S, Mantzoros C, Hammer S, Samore M. 2000. Effects of Protease Inhibitors on Hyperglycemia, Hyperlipidemia, and Lipodystrophy A 5-Year Cohort Study. Arch Intern Med 160:2050-2056.

9. Galescu O, Bhangoo A, Ten S. 2013. Insulin resistance, lipodystrophy and cardiometabolic syndrome in HIV/AIDS. Reviews in Endocrine and Metabolic Disorders 14:133-140.

10. Tilg H, Kaser A. 2011. Gut microbiome, obesity, and metabolic dysfunction. The Journal of Clinical Investigation 121:2126-2132.

11. Sonnenburg JL, Bäckhed F. 2016. Diet–microbiota interactions as moderators of human metabolism. Nature 535:56-64.

12. Pedersen HK, Gudmundsdottir V, Nielsen HB, Hyotylainen T, Nielsen T, Jensen BA, Forslund K, Hildebrand F, Prifti E, Falony G, Le Chatelier E, Levenez F, Dore J, Mattila I, Plichta DR, Pohoh P, Hellgren LI, Arumugam M, Sunagawa S, Vieira-Silva S, Jorgensen T, Holm JB, Trost K, Meta HITC, Kristiansen K, Brix S, Raes J, Wang J, Hansen T, Bork P, Brunak S, Oresic M, Ehrlich SD, Pedersen O. 2016. Human gut microbes impact host serum metabolome and insulin sensitivity. Nature 535:376-81.

13. Matey-Hernandez ML, Williams FMK, Potter T, Valdes AM, Spector TD, Menni C. 2018. Genetic and microbiome influence on lipid metabolism and dyslipidemia. Physiol Genomics 50:117-126.

14. Karlsson F, Tremaroli V, Nielsen J, Bäckhed F. 2013. Assessing the human gut microbiota in metabolic diseases. Diabetes 62:3341-9.

15. Neff CP, Krueger O, Xiong K, Arif S, Nusbacher N, Schneider JM, Cunningham AW, Armstrong A, Li S, McCarter MD, Campbell TB, Lozupone CA, Palmer BE. 2018. Fecal Microbiota Composition Drives Immune Activation in HIV-infected Individuals. EBioMedicine 30:192-202.

16. Armstrong AJS, Shaffer M, Nusbacher NM, Griesmer C, Fiorillo S, Schneider JM, Preston Neff C, Li SX, Fontenot AP, Campbell T, Palmer BE, Lozupone CA. 2018. An exploration of Prevotella-rich microbiomes in HIV and men who have sex with men. Microbiome 6:198.

17. Vujkovic-Cvijin I, Somsouk M. 2019. HIV and the Gut Microbiota: Composition, Consequences, and Avenues for Amelioration. Curr HIV/AIDS Rep doi:10.1007/s11904-019-00441-w.
18. Noguera-Julian M, Rocafort M, Guillén Y, Rivera J, Casadellà M, Nowak P, Hildebrand F, Zeller G, Parera M, Bellido R, Rodríguez C, Carrillo J, Mothe B, Coll J, Bravo I, Estany C, Herrero C, Szaj J, Sirera G, Torrela A, Navarro J, Crespo M, Brander C, Negredo E, Blanco J, Guarner F, Calle ML, Bork P, Sönnerborg A, Clotet B, Paredes R. 2016. Gut Microbiota Linked to Sexual Preference and HIV Infection. EBioMedicine 5:135-46.

19. Li SX, Sen S, Schneider JM, Xiong KN, Nusbacher NM, Moreno-Huizar N, Shaffer M, Armstrong AJS, Severs E, Kuhn K, Neff CP, McCarter M, Campbell T, Lozupone CA, Palmer BE. 2019. Gut microbiota from high-risk men who have sex with men drive immune activation in gnotobiotic mice and in vitro HIV infection. PLoS Pathog 15:e1007611.

20. Cui HL, Ditiatkovski M, Kesani R, Bobryshev YV, Liu Y, Geyer M, Mukhamedova N, Bukrinsky M, Sviridov D. 2014. HIV protein Nef causes dyslipidemia and formation of foam cells in mouse models of atherosclerosis. FASEB Journal 28:2828-2839.

21. Auclair M, Guenantin AC, Fellahi S, García M, Capeau J. 2020. HIV antiretroviral drugs, dolutegravir, maraviroc and ritonavir-boosted atazanavir use different pathways to affect inflammation, senescence and insulin sensitivity in human coronary endothelial cells. PLoS One 15:e0226924.

22. Ahmed D, Roy D, Cassol E. 2018. Examining Relationships between Metabolism and Persistent Inflammation in HIV Patients on Antiretroviral Therapy. Mediators Inflamm 2018:6238978.

23. Reid M, Ma Y, Scherzer R, Price JC, French AL, Plankey MW, Grunfeld C, Tien PC. 2017. Higher CD163 levels are associated with insulin resistance in hepatitis C virus-infected and HIV-infected adults. AIDS 31:385-393.

24. Somsouk M, Estes JD, Deleage C, Dunham RM, Albright R, Inadomi JM, Martin JN, Deeks SG, McCune JM, Hunt PW. 2015. Gut epithelial barrier and systemic inflammation during chronic HIV infection. AIDS (London, England) 29:43-51.

25. Klase Z, Ortiz A, Deleage C, Mudd JC, Quiñones M, Schwartzman E, Klatt NR, Canary L, Estes JD, Brenchley JM. 2015. Dysbiotic bacteria translocate in progressive SIV infection. Mucosal Immunology 8:1009-1020.

26. Palmer CD, Tomassilli J, Sirignano M, Romero-Tejeda M, Arnold KB, Che D, Lauffenburger DA, Jost S, Allen T, Mayer KH, Altfeld M. 2014. Enhanced immune activation linked to endotoxemia in HIV-1 seronegative MSM. AIDS 28:2162-6.

27. Lim PS, Chang YK, Wu TK. 2019. Serum Lipopolysaccharide-Binding Protein is Associated with Chronic Inflammation and Metabolic Syndrome in Hemodialysis Patients. Blood Purif 47:28-36.

28. Hersoug LG, Moller P, Loft S. 2018. Role of microbiota-derived lipopolysaccharide in adipose tissue inflammation, adipocyte size and pyroptosis during obesity. Nutr Res Rev 31:153-163.

29. Awoyemi A, Troseid M, Arnesen H, Solheim S, Seljeffot I. 2018. Markers of metabolic endotoxemia as related to metabolic syndrome in an elderly male population at high cardiovascular risk: a cross-sectional study. Diabetol Metab Syndr 10:59.

30. Liu X, Lu L, Yao P, Ma Y, Wang F, Jin Q, Ye X, Li H, Hu FB, Sun L, Lin X. 2014. Lipopolysaccharide binding protein, obesity status and incidence of metabolic syndrome: a prospective study among middle-aged and older Chinese. Diabetologia 57:1834-41.

31. Gelpi M, Vestad B, Hansen SH, Holm K, Drivsholm N, Goetz A, Kirkby NS, Lindegaard B, Lebech AM, Hoel H, Michelsen AE, Ueland T, Gerstoft T, Lundgren J, Hov JR, Nielsen SD, Troseid M. 2020. Impact of HIV-Related Gut Microbiota Alterations on Metabolic Comorbidities. Clin Infect Dis doi:10.1093/cid/ciz1235.

32. Hammer SM, Sobieszczyk ME, Janes H, Karuna ST, Mulligan MJ, Grove D, Koblin BA, Buchbinder SP, Keefer MC, Tomaras GD, Frahm N, Hural J, Anude C, Graham BS,
Enama ME, Adams E, DeJesus E, Novak RM, Frank I, Bentley C, Ramirez S, Fu R, Koup RA, Mascola JR, Nabel GJ, Montefiori DC, Kublin J, McElrath MJ, Corey L, Gilbert PB, Team HS. 2013. Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine. N Engl J Med 369:2083-92.

Wiley JF, Carrington MJ. 2016. A metabolic syndrome severity score: A tool to quantify cardio-metabolic risk factors. Prev Med 88:189-95.

Hillier TA, Rousseau A, Lange C, Lepinay P, Cailleau M, Novak M, Calliez E, Ducimetiere P, Balkau B, Cohort DESIR. 2006. Practical way to assess metabolic syndrome using a continuous score obtained from principal components analysis. Diabetesologia 49:1528-35.

Genuer R, Poggi J-M, Tuleau-Malot C. 2015. VSURF: An R Package for Variable Selection Using Random Forests. The R Journal, R Foundation for Statistical Computing 7:12-33.

Marchetti G, Tincati C, Silvestri G. 2013. Microbial translocation in the pathogenesis of HIV infection and AIDS. Clinical microbiology reviews 26:2-18.

Duncan SH, Holtrop G, Lobley GE, Calder AG, Stewart CS, Flint HJ. 2004. Contribution of acetate to butyrate formation by human faecal bacteria. Br J Nutr 91:915-23.

Duncan SH, Barcenilla A, Stewart CS, Pryde SE, Flint HJ. 2002. Acetate utilization and butyryl coenzyme A (CoA):acetate-CoA transferase in butyrate-producing bacteria from the human large intestine. Appl Environ Microbiol 68:5186-90.

Basu S, Kumbier K, Brown JB, Yu B. 2018. Iterative random forests to discover predictive and stable high-order interactions. Proc Natl Acad Sci U S A 115:1943-1948.

Shaffer M, Thurimella K, Quinn K, Doenges K, Zhang X, Bokatzian S, Reisdorph N, Lozupone CA. 2019. AMON: annotation of metabolite origins via networks to integrate microbiome and metabolome data. BMC Bioinformatics 20:614.

Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Research 45:D353-D361.

Douglas GM, Maffei VJ, Zaneveld J, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. 2020. PICRUSt2: An improved and extensible approach for metagenome inference. doi:10.1101/672295.

Frolkis A, Knox C, Lim E, Jewison T, Law V, Hau DD, Liu P, Gautam B, Ly S, Guo AC, Xia J, Liang Y, Shrivastava S, Wishart DS. 2010. SMPDB: The Small Molecule Pathway Database. Nucleic Acids Res 38:D480-7.

van der Veen JN, Kennelly JP, Wan S, Vance JE, Vance DE, Jacobs RL. 2017. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. Biochim Biophys Acta Biomembr 1859:1558-1572.

Verhaegen P, Borchelt M, Smith J. 2003. Relation between cardiovascular and metabolic disease and cognition in very old age: cross-sectional and longitudinal findings from the berlin aging study. Health Psychol 22:559-69.

Montessori V, Press N, Harris M, Akagi L, Montaner JSG. 2004. Adverse effects of antiretroviral therapy for HIV infection. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne 170:229-38.

Alexander CM, Landsman PB, Teutsch SM, Haffner SM, Third National H, Nutrition Examination S, National Cholesterol Education P. 2003. NCEP-defined metabolic syndrome, diabetes, and prevalence of coronary heart disease among NHANES III participants age 50 years and older. Diabetes 52:1210-4.

Meigs JB. 2003. Epidemiology of the insulin resistance syndrome. Curr Diab Rep 3:73-9.
Weight Gain Associated with Integrase Stand Transfer Inhibitor Use in Women. Clin Infect Dis doi:10.1093/cid/ciz853.

Venter WDF, Moorhouse M, Sokhela S, Fairlie L, Mashabane N, Masenya M, Serenata C, Akpomiemie G, Qavi A, Chandiwana N, Norris S, Chersich M, Clayden P, Abrams E, Arulappan N, Vos A, McCann K, Simmons B, Hill A. 2019. Dolutegravir plus Two Different Prodrugs of Tenofovir to Treat HIV. N Engl J Med 381:803-815.

Pérez-Matute P, Pérez-Martínez L, Aguilera-Lizarraga J, Blanco JR, Oteo JA. 2015. Maraviroc modifies gut microbiota composition in a mouse model of obesity: a plausible therapeutic option to prevent metabolic disorders in HIV-infected patients. Revista española de quimioterapia : publicación oficial de la Sociedad Española de Quimioterapia 28:200-6.

Han GM, Meza JL, Soliman GA, Islam KM, Watanabe-Galloway S. 2016. Higher levels of serum lycopene are associated with reduced mortality in individuals with metabolic syndrome. Nutr Res 36:402-7.

Zeng YC, Peng LS, Zou L, Huang SF, Xie Y, Mu GP, Zeng XH, Zhou XL, Zeng YC. 2017. Protective effect and mechanism of lycopene on endothelial progenitor cells (EPCs) from type 2 diabetes mellitus rats. Biomed Pharmacother 92:86-94.

O'Keefe SJD, Li JV, Lahti L, Ou J, Carbonero F, Mohammed K, Posma JM, Kinross J, Wahl E, Ruder E, Vipperla K, Naidoo V, Mtshali L, Tims S, Puylaert PGB, DeLany J, Krasinskas A, Benefiel AC, Kaseb HO, Newton K, Nicholson JK, de Vos WM, Gaskins HR, Zoetendal EG. 2015. Fat, fibre and cancer risk in African Americans and rural Africans. Nature communications 6:6342.

Hsu LA, Chang CJ, Wu S, Teng MS, Chou HH, Chang HH, Chang PY, Ko YL. 2010. Association between functional variants of the ICAM1 and CRP genes and metabolic syndrome in Taiwanese subjects. Metabolism 59:1710-6.

Senn JJ, Klover PJ, Nowak IA, Mooney RA. 2002. Interleukin-6 induces cellular insulin resistance in hepatocytes. Diabetes 51:3391-9.

Tilves CM, Zmuda JM, Kuipers AL, Nestlerode CS, Evans RW, Bunker CH, Patrick AL, Miljkovic I. 2016. Association of Lipopolysaccharide-Binding Protein With Aging-Related Adiposity Change and Prediabetes Among African Ancestry Men. Diabetes Care 39:385-91.
65. Moreau RA, Agnew J, Hicks KB, Powell MJ. 1997. Modulation of lipoxygenase activity by bacterial hopanoids. J Nat Prod 60:397-8.

66. Tsai IJ, Croft KD, Mori TA, Falck JR, Beilin LJ, Puddey IB, Barden AE. 2009. 20-HETE and F2-isoprostanes in the metabolic syndrome: the effect of weight reduction. Free Radic Biol Med 46:263-70.

67. Raza GS, Putaala H, Hibberd AA, Alhoniemi E, Tiihonen K, Makela KA, Herzig KH. 2017. Polydextrose changes the gut microbiome and attenuates fasting triglyceride and cholesterol levels in Western diet fed mice. Sci Rep 7:5294.

68. Rabot S, Membrez M, Bruneau A, Gérard P, Harach T, Moser M, Raymond F, Mansourian R, Chou CJ. 2010. Germ-free C57BL/6J mice are resistant to high-fat-diet-induced insulin resistance and have altered cholesterol metabolism. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 24:4948-59.

69. Velagapudi VR, Hezaveh R, Reigstad CS, Gopalacharyulu P, Yetukuri L, Islam S, Felin J, Perkins R, Boren J, Oresic M, Backhed F. 2010. The gut microbiota modulates host energy and lipid metabolism in mice. J Lipid Res 51:1101-12.

70. National Institutes of Health EaGRP, National Cancer Institute. 2010. Diet History Questionnaire, Version 2.0.

71. National Cancer Institute EaGRP. 2012. Diet*Calc Analysis Program, Version 1.5.0. .

72. Shaffer M. SCNIC: Sparse Cooccurrence Network Investigation for Compositional data. https://github.com/shafferm/SCNIC. Accessed

73. Yang Y, Cruickshank C, Armstrong M, Mahaffey S, Reisdorph R, Reisdorph N. 2013. New sample preparation approach for mass spectrometry-based profiling of plasma results in improved coverage of metabolome. J Chromatogr A 1300:217-26.

74. Halper-Stromberg E, Gillenwater L, Cruickshank-Quinn C, O’Neal WK, Reisdorph N, Petrache I, Zhuang Y, Labaki WW, Curtis JL, Wells J, Rennard S, Pratte KA, Woodruff P, Stringer KA, Kechris K, Bowler RP. 2019. Bronchoalveolar Lavage Fluid from COPD Patients Reveals More Compounds Associated with Disease than Matched Plasma. Metabolites 9.

75. Kennedy A, Bivens A. 2017. Methods for the Analysis of Underivatized Amino Acids by LC/MS: For Food, Life Science, and Metabolomics Applications, Agilent Application Note.

76. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW, Fiehn O, Goodacre R, Grillin JL, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane AN, Lindon JC, Marriott P, Nicholls AW, Reilly MD, Thaden JJ, Viant MR. 2007. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 3:211-221.

77. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S. 2016. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. J Cheminform 8:3.

78. Agilent. 2020. Lipidomics Analysis with Lipid Annotator and Mass Profiler Professional, vol 5994-1111EN., Technical Overview.

79. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G, Navas-Molina JA, Janssen S, Kopylova E, Vazquez-Baeza Y, Gonzalez A, Morton JT, Mirarab S, Zech Xu Z, Jiang L, Haroon MF, Kanbar J, Zhu Q, Jin Song S, Kosciolok T, Bokulich NA, Lefler J, Brislawn CJ, Humphrey G, Owens SM, Hampton-Marcell J, Berg-Lyons D, McKenzie V, Fierer N, Fuhrman JA, Clauset A, Stevens RL, Shade A, Pollard KS, Goodwin KD, Jansson JK, Gilbert JA, Knight R, Earth Microbiome Project C. 2017. A communal catalogue reveals Earth’s multiscale microbial diversity. Nature 551:457-463.
Brown CT, Callahan BJ, Caraballo-Rodriguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvall C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852-857.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13:581-3.

Rognes T, Flouri T, Nichols B, Quince C, Mahe F. 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4:e2584.

McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P. 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. The ISME journal 6:610-8.

Janssen S, McDonald D, Gonzalez A, Navas-Molina JA, Jiang L, Xu ZZ, Winker K, Kado DM, Orwoll E, Manary M, Mirarab S, Knight R. 2018. Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. mSystems 3.

Friedman J, Alm EJ. 2012. Inferring correlation networks from genomic survey data. PLoS Comput Biol 8:e1002687.