Microbiome

RESEARCH

Open Access

Alterations of the gut microbiome in HIV infection highlight human anelloviruses as potential predictors of immune recovery

Celia Boukadida[1], Amy Peralta-Prado[1], Monserrat Chávez-Torres[1], Karla Romero-Mora[1], Alma Rincon-Rubio[1], Santiago Ávila-Ríos[1], Daniela Garrido-Rodríguez[1], Gustavo Reyes-Terán[1,2] and Sandra Pinto-Cardoso[1*]

Abstract

Background HIV-1 infection is characterized by a massive depletion of mucosal CD4 T cells that triggers a cascade of events ultimately linking gut microbial dysbiosis to HIV-1 disease progression and pathogenesis. The association between HIV infection and the enteric virome composition is less characterized, although viruses are an essential component of the gut ecosystem. Here, we performed a cross-sectional analysis of the fecal viral (eukaryotic viruses and bacteriophages) and bacterial microbiome in people with HIV (PWH) and in HIV-negative individuals. To gain further insight into the association between the gut microbiome composition, HIV-associated immunodeficiency, and immune recovery, we carried out a longitudinal study including 14 PWH who initiated antiretroviral therapy (ART) and were followed for 24 months with samplings performed at baseline (before ART) and at 2, 6, 12, and 24 months post-ART initiation.

Results Our data revealed a striking expansion in the abundance and prevalence of several human virus genomic sequences (*Anelloviridae*, *Adenoviridae*, and *Papillomaviridae*) in stool samples of PWH with severe immunodeficiency (CD4 < 200). We also noted a decreased abundance of sequences belonging to two plant viruses from the *Tobamovirus* genus, a reduction in bacterial alpha diversity, and a decrease in *Inoviridae* bacteriophage sequences. Short-term ART (24 months) was linked to a significant decrease in human *Anelloviridae* sequences. Remarkably, the detection of Anellovirus sequences at baseline independently predicted poor immune recovery, as did low CD4 T cell counts. The bacterial and bacteriophage populations were unique to each PWH with individualized trajectories; we found no discernable pattern of clustering after 24 months on ART.

Conclusion Advanced HIV-1 infection was associated with marked alterations in the virome composition, in particular a remarkable expansion of human anelloviruses, with a gradual restoration after ART initiation. In addition to CD4 T cell counts, anellovirus sequence detection might be useful to predict and monitor immune recovery. This study confirms data on the bacteriome and expands our knowledge on the viral component of the gut microbiome in HIV-1 infection.

*Correspondence:
Sandra Pinto-Cardoso
sandra.pintocardoso.cieni@gmail.com
[1] Departamento del Centro de Investigación en Enfermedades Infecciosas, Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas, Ciudad de México, México
[2] Comisión Coordinadora de Institutos Nacionales de Salud y Hospitales de Alta Especialidad, Ciudad de México, México

Introduction

The gastrointestinal (GI) tract is a major site of HIV-1 replication, resulting in a massive depletion of mucosal CD4 T cells in the small and large intestines during both acute and chronic phases of HIV infection [1]. HIV-1 infection is also associated with enteropathy, mucosal inflammation, and disruption of the epithelial barrier leading to microbial translocation [2, 3]. Great strides have been made to understand the contribution of the gut microbiome to HIV-1 infection, primarily HIV-1 pathogenesis, and to leverage this knowledge to favorably impact the lives of people living with HIV (PWH). A large number of studies have indicated that microbial communities are implicated in fueling chronic activation [4] and inflammation and possibly causally involved in metabolic disorders and non-communicable diseases [5].

Bacterial dysbiosis is known to occur in HIV-1 infection [6, 7]. Data from cross-sectional (for the most part) and a handful of longitudinal studies indicate reduced bacterial diversity and enrichment in Pseudomonadota (formally Proteobacteria). The difficulty in defining an HIV-1-associated signature of bacterial dysbiosis is due to the ever-growing list of confounders and factors that also shape the gut microbiota, which are not always included or considered especially in earlier studies including sexual preference [8] and more recently geographical location [9, 10]. Furthermore, evidence from recent studies suggests that an HIV-1-associated bacterial dysbiosis signature is only apparent or detectable in later stages of HIV-1 infection (chronic infection) [6]. In an attempt to resolve these issues inherent to human studies, non-human primate (NHP) models have been used, where environmental and genetic factors are controlled for [11]. Studies have shown a lack of bacterial dysbiosis in NHP [12]. Furthermore, attempts to recreate an HIV-associated signature using specific antibiotics targeting Gram-positive bacteria failed to demonstrate a strong link between bacterial dysbiosis and HIV-1 disease progression in NHP [13].

In addition to bacteria, viruses are highly abundant and diverse in the human gut and exhibit important inter-individual variations [14]. The human gut virome is composed of eukaryotic and prokaryotic viruses infecting human cells, plants, or microorganisms such as bacteria, fungi, and archaea. The composition and potential contributions of viral communities to human diseases have been far less characterized than its bacterial counterparts, mainly due to technical limitations [15]. Studies have highlighted associations between the enteric virome composition and several diseases such as inflammatory bowel disease, diabetes, and graft versus host disease [16–19]. A seminal study of the human gut DNA virome in HIV-1 infection revealed an expansion of adenovirus sequences in highly immunosuppressed individuals, potentially contributing to HIV-associated enteropathy [20]. In another study, HIV-1 infection was not associated with changes in eukaryotic virus alpha- and beta-diversity, although a decrease in phage richness that reversed after antiretroviral therapy (ART) initiation was described [21]. In NHP models, expansion of adenoviruses during pathogenic simian immunodeficiency virus (SIV) infection was associated with damage to the epithelial barrier, gastrointestinal disease, and acquired immunodeficiency syndrome (AIDS)-related mortality [22, 23]. Interestingly, several mammalian viruses such as *Herpesviridae* and *Reoviridae* were also shown to be more abundant in gorillas naturally infected with SIV [24]. Altogether, these studies importantly demonstrate that HIV-1 and SIV infections can be associated with changes in the gut virome composition. Further studies should contribute to clarify the underlying mechanisms and potential role in HIV-1 disease progression.

Here, we explored the ecology of gut microbial communities including both the bacteriome and the RNA and DNA virome in HIV-1 infection. This study was conducted in Mexico where, despite universal availability of ART, PWH still frequently present with low CD4 T cell counts and acquired immunodeficiency syndrome (AIDS)-defining conditions. We first performed a cross-sectional study including 92 untreated PWH across different stages of HIV-1 infection and a control cohort of 53 HIV-seronegative individuals at high risk of HIV-1 infection. To gain further insight into the interplay between the gut microbiome composition and the immune status, we performed a longitudinal study following 14 highly immunocompromised PWH for 24 months after ART initiation. We used 16 S rRNA sequencing to characterize bacterial communities and conducted a comprehensive shotgun RNA and DNA virome analysis of fecal samples. Our data revealed significant alterations of the fecal virome in immunosuppressed PWH, and temporal evolution after ART initiation and suggested an association between the presence of *Anelloviridae* sequences and immune recovery.

Methods

Ethics

This study was approved by the Ethics, Research, and Safety Committees of the Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas (INER) under approval numbers B36-12, C59-15, and C15-16. All participants were informed of the study objectives and signed an informed consent. All participants were 18 years or older. The study was conducted according to the Declaration of Helsinki of 1964 and later amendments.

Metadata

The following metadata were included: age (numeric), sex assigned at birth (categorical: female, male), place of residence (categorical; CDMX Ciudad de México, EDOMEX State of Mexico, Others), gastrointestinal symptoms at study entry (categorical: yes, no), antibiotic use at study entry (categorical: yes, no), current smoking (categorical: yes, no), sexual preference (categorical: men who have sex with men, MSM, men who have sex with women, MSW, women who have sex with men, WSM), body mass index (BMI, kg/m^2, numeric), season of the year (categorical: summer, autumn, winter, spring), HIV status (categorical: negative, positive), absolute CD4 T cell count (cells/μL, numeric), CD4 percentage (%, numeric), CD4/CD8 ratio (numeric), and HIV plasma viral load (HIV-1 RNA copies/mL, numeric).

Study cohorts

A flow diagram of our study is summarized in Supplementary Figure S1.

Cross-sectional cohort

A total of 145 individuals were included: 53 were HIV-seronegative individuals (protocol C15-16), and 92 were PWH naïve to ART (protocols B36-12 and C59-15), and included both ambulatory and hospitalized individuals. HIV-1 diagnosis was confirmed by enzyme-linked immunosorbent assay and performed by the Virologic Diagnostic Laboratory, Departamento del Centro de Investigación en Enfermedades Infecciosas (CIENI), Instituto Nacional de Enfermedades Respiratorias (INER), in accordance with the manufacturer's instructions (VIDAS HIV DUO Ultra, VIDAS, BioMérieux, Marcy l'Etoile, France, and GENSCREEN ULTRA HIV Ag-Ab, Evolis, Bio-Rad, Hercules, CA, USA). HIV testing was also performed for all HIV-seronegative individuals. This control cohort included individuals at high-risk of HIV infection. To evaluate the risk of HIV acquisition, a risk behavior assessment questionnaire based on sexual practices was administered, as previously published [25]. Confirmatory negative HIV testing was performed prior to enrollment. Clinical and demographic data was obtained by performing standardized interviews. PWH naïve to ART were stratified into four groups according to their CD4 T cell counts: HIV CD4 $\geq$ 350 (n = 26), HIV CD4 [200:350) (n = 16), HIV CD4 [50:200) (n = 25), and HIV CD4 < 50 (n = 25).

Sub-cohort 1/longitudinal

Fourteen PWH naïve to ART were selected based on two criteria: (1) availability of paired stool samples at five-time points (baseline, previous to ART) and 2, 6, 12, and 24 months after ART initiation and (2) baseline CD4 T cell counts < 200 cells/mm^3. Demographic and clinical characteristics at baseline are summarized in Table S1. Data on the dynamics of HIV clinical parameters following ART initiation are summarized in Table S2. Following ART initiation, HIV plasma load decreased significantly at month 2 and thereafter ($p < 0.0001^{****}$, significance was determined by fitting a linear regression with plasma viral load as the outcome variable ($\log_{10}$), and time on ART as the independent variable). Complete virological suppression was achieved for 4 (28.57%) PWH at month 2, 11 (78.57%) at month 6, and 12 (85.71%) at month 12 (and at month 24). Some PWH changed the ART regimen at month 2 (n = 2) and month 6 (n = 1). There was a significant increase in both absolute CD4 T cell counts and CD4/CD8 ratio (significance was determined by pair-wise Wilcoxon test between time points).

Sub-cohort 2: immune recovery cohort

We explored associations between the gut microbiome and immune recovery. To do so, we collected follow-up clinical HIV data (pVL and CD4 T cell counts) for up to 60 months after ART initiation. PWH with baseline CD4 T cell count below 350 cells/μL and sustained plasma HIV viral load < 200 copies/ml after 6 months of ART were included (n = 42). Fifteen participants (35.7%) had at least one CD4 T cell count above 500 cells/mm^3 and were classified as adequate immune responders (AIR), whereas seventeen individuals (40.5%) had all CD4 T cell counts below 350 cells/μL and were classified as poor immunological responders (PIR).

Samples

Stool samples were collected, temporarily placed at + 4 °C, aliquoted (approximately 200–250 mg per tube), and stored at − 80°C within 2 h of reception at the CIENI.

HIV-1 RNA plasma viral load and lymphocyte populations

HIV-1 RNA plasma viral load was determined by automated real-time PCR using m2000 system (Abbot, Abbott Park, Illinois, USA). Lymphocyte populations were obtained by flow cytometry using Trucount in a FACSCanto II (BD Biosciences, San Jose, California, USA).

Bacterial 16S rRNA analysis

Stool DNA extraction

A total of 145 stool samples were available for DNA extraction. DNA was extracted using the PowerSoil DNA Extraction Kit (MO BIO Laboratories, Carlsbad, CA, USA) as instructed by the manufacturer. Briefly, stool samples were transferred from − 80 °C to the laboratory

on dry ice and then transferred in a Powerbead Tube. Next, 60 µl of solution C1 was added and the tube was inverted several times. Tubes were vortexed horizontally (bead beating) for 30 min on a vortex with a horizontal adapter. Next, tubes were centrifuged for 30 s min at 10,000×*g*. The supernatant was transferred to a new tube and 250 µl of the solution C2 was added, briefly vortexed, and incubated at 4 °C for 5 min. Tubes were centrifuged for 30 s at 10,000×*g*. Then, 600 µl of supernatant was transferred to a new tube and 200 µl of the solution C3 was added, briefly vortexed, and incubated at 4 °C for 5 min. Tubes were centrifuged for 30 s at 10,000×*g*. Next, 750 µl of the supernatant was transferred into a new tube and 1200 µl of solution C4 was added and vortexed for 5 s. The supernatant was loaded on to a spin column-tube assembly (three consecutive loadings of 675 µl) and centrifuged at 10,000×*g* for 1 min. Spin filters were washed by centrifugation with 500 µl of solution 5, and the total DNA was eluted in 100 µl of solution 6. DNA purity and quality were assessed by absorbance on a Nanodrop N1000 (ThermoFisher Scientific, Carlsbad, CA, USA) by measuring the A260/A280 ratio.

16S rRNA gene amplification and sequencing

16S rRNA sequencing was performed according to the 16S metagenomic library preparation guide (Illumina® MiSeq System) using a two-stage PCR protocol. Primers for the V3-V4 16 s rRNA were previously published [26]. The forward primer was (341F), 5′-CCTACGGGNGGC WGCAG-3′, and the reverse primer was (805R), 5′-GAC TACHVGGGTATCTAATCC-3′. Primers were purchased with overhang adapter sequences, the forward overhang sequence was 5′-TCGTCGGCAGCGTCAGAT GTGTATAAGAGACAG-followed by the 341F primer, and the reverse overhang sequence was 5′-GTCTCG TGGGCTCGGAGATGTGTATAAG AGACAG-followed by the 805R primer (Invitrogen, Waltham, MA, USA). Briefly, each DNA sample was amplified in triplicate by PCR in a final volume of 25 µL. PCR reactions contained RNase/DNase-free water, 2.5 µL 10X high fidelity PCR buffer (Invitrogen), 0.5 µL 10 mM dNTPs, 1 µL 50 mM $MgSO_4$, 5 µL each of the forward and reverse primers (10 µM final concentration), and 0.1 µL platinum high-fidelity Taq (Invitrogen). Reactions were held at 95 °C for 3 min to denature the DNA, with amplification (25 cycles) at 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s; a final extension of 5 min at 72 °C was added to ensure complete amplification. PCR cycling was limited to 25 cycles. PCR amplification was verified on a 2% agarose gel (Sigma-Aldrich, MO, USA). Triplicate PCR reactions were pooled per sample and purified using AgenCourt AMPure XP beads (Beckman Dickson, Atlanta, GA, USA) as instructed by the manufacturer. Dual indices were attached by PCR to the purified PCR amplicon using the Nextera XT Index Kit (Illumina, San Diego, CA, USA). The index PCR clean-up step was repeated to guarantee complete purification of the indexed libraries using AgenCourt AMPure XP beads. Library quantification was performed using the Qubit dsDNA HS Buffer and Standards kit (Invitrogen) as instructed by the manufacturer. Next, each library was normalized to 4 nM by dilution and pooled together. Library size and molarity were checked on a 2100 Bioanalyzer instrument (Agilent Technologies, Santa Clara, CA, USA) and sequenced on the Illumina MiSeq™ platform using a final library molarity of 14 pM to obtain paired-end sequences (2×300 cycles). The internal PhiX control was used at 25% (also at 14 pM). Negative controls were included in parallel to account for contamination from reagents during DNA extraction and library preparation. Control samples did not amplify by PCR nor did they yield a quantifiable library, verified on the 2100 Bioanalyzer. Library preparation and sequencing were performed at the CIENI, INER. Raw 16S sequences were deposited at the National Center for Biotechnology Information-Sequence Read Archive (NCBI-SRA) under project PRJNA901491 (cross-sectional) and PRJNA996141 (longitudinal).

16S rRNA gene analysis

Raw demultiplex FastQ files (read 1 and read 2) were processed using QIIME2 (Quantitative Insights into Microbial Ecology, version 2022.2, [27]). First, sequences were imported into QIIME2, and primers were removed using qiime cutadapt trim-paired and checked for sequence quality and length. Next, DADA2 [28] was used to denoise, remove chimera and PhiX sequences, merge paired-end reads, and construct an amplicon variant sequences (ASVs) table. Negative controls yielded less than 1000 raw sequences (read 1 and read2), and no paired-sequences were obtained after DADA2. Taxonomy was assigned using the SILVA database (release 138, [29]) at 99% identity, pre-trained for the V3-V4 16S rRNA region [30]. Qiime2 artifacts were imported into R for further manipulation and graph visualization (R version 4.3.0). For the cross-sectional cohort, after DADA2, 32 samples were excluded due to a very low number of raw sequences (< 1000). A total of 113 samples were retained for studying the bacteriome, including 38 HIV-neg and 75 PWH. After excluding these samples, a total of 7,698,288 paired-sequences were available, with a median of 46,577 sequences (min, 6052-max, 343,650). Rarefaction was performed at a sampling depth of 5446 sequences per sample. For the longitudinal cohort, after DADA2, a median of 74,120.5 sequences were retained (min 12,086-max 260,599).

Virome analysis

Virus particle enrichment and nucleic acid extraction

Virus particles were purified and non-protected nucleic acids were digested prior to extraction. Stool samples were resuspended at 4% in phosphate-buffered saline (PBS) and centrifuged three times (resuspended sample centrifugation 5 min at 2500 g, supernatant centrifugation 15 min at 4800 g and supernatant centrifugation 30 min at 4800 g). Supernatants were filtered at 0.8 µm and 0.45 µm, concentrated using a 100 kDa Amicon molecular weight cutoff filter (Amicon), and treated with TURBO DNase (Ambion) and RNase I (Invitrogen) for 1 h 30 min at 37 °C. Nucleic acids were extracted using the QIAamp viral RNA mini kit (Qiagen).

Sample processing and whole metagenome sequencing

Amplified complementary DNA (cDNA) was prepared from extracted RNA using sequence-independent single primer amplification (SISPA) as previously described with some modifications [31]. Total cDNA was synthesized in triplicate reactions with three different sets of primers to reduce bias in sequence-independent nucleic acid amplification [16, 32]. Extracted RNA was converted to cDNA using rando nonamer primers attached to a conserved sequence using SuperScript III Reverse Transcriptase System (Thermo Fisher) and primer A1 (5′-GCCGGAGCTCTGCAGATATCNNNNNNNNN-3′), B1 (5′-GACCATCTAGCGACCTCCACNNNNNNNNN-3′), or C1 (5′-CCTTGAAGGCGGACTGTGAGNNNNNNNNN-3′). The second strand was synthesized using Klenow fragment polymerase (New England Biolabs) followed by 25 cycles of amplification using Expand High-Fidelity DNA polymerase (Roche) and primer A2 (5′-GCCGGAGCTCTGCAGATATC-3′), B2 (5′-GACCATCTAGCGACCTCCAC-3′), or C2 (5′-CCTTGAAGGCGGACTGTGAG-3′). For each sample, three independent cDNA reactions were performed and pooled, purified twice using AMPure XP beads (Beckman Coulter) and whole metagenome sequencing libraries were generated using the Nextera XT DNA library preparation kit (Illumina). Finally, samples were sequenced on the NextSeq 500 platform (Illumina) using 2×150-cycles mid-output kits (Illumina). Viral metagenomic raw sequencing reads were deposited at the National Center for Biotechnology Information-Sequence Read Archive (NCBI-SRA) under project PRJNA1068646.

Viral metagenomic sequence analysis

Sequences were analyzed using the SURPI bioinformatics pipeline [33]. Briefly, raw sequencing reads were preprocessed by removal of adapter, low-quality, and low-complexity sequences followed by subtraction of human reads. The remaining reads were then aligned to nucleotide and protein databases (NCBI NT and NR databases, July 2018) using SNAP and RAPSearch to identify viral sequences.

Multiple studies have demonstrated the presence of contaminant nucleic acids in various reagents and materials used in viral metagenomic studies highlighting the importance of using a proper control strategy [34–38]. To identify reads resulting from contaminants, one non-template control (NTC) was included in each batch of processed samples (25 NTC in total) and was sequenced in parallel. Reads corresponding to bacteriophages, large dsDNA viruses (*Phycodnaviridae*, *Mimiviridae*, *Pithoviridae*, *Iridoviridae*, and *Marseilleviridae*), parvoviruses, circular single-stranded DNA (ssDNA) genomes (*Circoviridae* and *Genomoviridae*), and Gammaretroviruses (Moloney Leukemia virus) were frequently identified in NTC and have been previously described as contaminants present in reagents and materials such as enzymes and extraction columns or resulting from bioinformatics misclassification in multiple studies [23, 34–39]. These viral sequences were not considered for downstream analyses. In addition, we occasionally observed a barcode hopping phenomenon resulting in false positive reads for a particular viral genus in samples sequenced on the same flow cell as a sample containing a large number of reads from this viral genus (over one million), as described previously [40]. Therefore, to minimize the risk from contaminant sequencing and barcode hopping, we applied a triple threshold strategy. A sample was considered positive for a particular genus and viral sequences were included for further analyses if the number of reads was higher than 10, higher than five times the mean NTC read number, and higher than 0.5% the number of reads found in the sample with the greater number of reads on the same flow cell.

Bacterial and viral alpha and beta-diversity analyses

Alpha-diversity was estimated using the Shannon index [how many ASVs and how evenly they are distributed] and observed ASVs (richness, how many ASVs there are), while beta-diversity was estimated using the Bray–Curtis dissimilarity index (vegan v.2.6–4, R). Beta-diversity was visualized using principal coordinate analysis (PCoA). The homogeneity of group dispersion was verified with the *betadisper* function (vegan, R). Permutational multivariate analysis of variance (adonis2, vegan) was used to test differences between groups. First, each categorical and numeric variable was tested for its impact on the fecal bacterial and viral structure using adonis2 (vegan, R, with 10,000 permutations) on the Bray–Curtis dissimilarity index. Missing values for each metadata variable were first removed. Categorical variables with three

or more levels were further tested for multilevel comparisons using pairwise adonis (adonis2, vegan, R), and *p* values were adjusted for multiple comparisons using Bonferroni.

Bacterial and viral differential abundance analyses

To identify differences in the abundances of individual taxa between groups, we used Microbiome Multivariate Association with Linear Models, MaAsLin2 [41]. MaAsLin2 relies on general linear models and a wide range of normalization and transformation methods. No collinearity was observed between the selected variables (pairwise Spearman correlation coefficients < 0.7) (Supplementary Figure S2). Prior to performing bacterial differential abundance analysis (DAA), ASVs collapsed at the genus level, and ASVs with less than 10% prevalence were removed as recommended [42].

Statistical analyses

Data were expressed as number (percentage), median (interquartile range), mean (plus or minus standard deviation), or number (percentage) as appropriate. Independent groups (cross-sectional study) were compared using the Mann–Whitney (two groups) or Kruskal–Wallis (three or more groups) test for continuous variables and Fisher's exact test or chi-square test (as required) for categorical values (GraphPad Prism v8, San Diego, California, USA). *P* values were adjusted for multiple comparisons when appropriate using Dunn's or false discovery rate (FRD). Paired data (longitudinal study) were compared using the Friedman test. Correlation coefficients and *p* values were computed with the nonparametric Spearman test using "cor.test" function in R, and *p* values were adjusted using the false discovery rate (FDR). For the bacteriome, all ASVs were collapsed at the genus level, and abundances were centered log ratio (CLR) transformed. Generalized linear regression analyses were used to assess relationships between alpha diversity metrics and CD4 T cell counts (numeric and categorical). Normal distribution was first assessed using the Shapiro–Wilk test. Alpha diversity indices were log10-transformed to reduce the skewness of the data. Data were visualized using "ggplot2", "ggcorrplot2", "ggboxplot," and "Heatmap".

Results

Microbiome study cohort

A total of 145 individuals were included, of those, 53 were HIV-seronegative individuals at high risk of HIV infection (HIV-negative) and 92 were PWH. The majority of participants were male (93.8%), resided in Mexico City (88.1%), self-declared as men who have sex with men (MSM, 85.8%), and their median age was 33 years. PWH presented a wide range of CD4 T cell counts (1–953 cells/mm^3, median 135.5 cells/mm^3), and 54.3% had progressed to acquired immunodeficiency syndrome (AIDS), the most severe stage of HIV infection defined by CD4 < 200. To reflect the degree of immunodeficiency, PWH were stratified into four groups according to their CD4 T cell counts: CD4 ≥ 350 (n = 26), 200 ≤ CD4 < 350 (n = 16), 50 ≤ CD4 < 200 (n = 25), and HIV CD4 < 50 (n = 25) (Table 1). Profound differences in age, sexual preference, body mass index (BMI), antibiotic use, hospitalization, and plasma HIV load were observed across our study groups. PWH with severe immunodeficiency (CD4 < 200) were older, with higher plasma HIV loads, lower BMI, higher frequencies of antibiotic use, and hospitalization, all indicative of advanced stages of HIV infection (Table 1 and Excel S1). Importantly, no differences across study groups were found for sex assigned at birth, place of residence, tobacco use, gastrointestinal symptoms, and fecal sample collection season. We used 16S rRNA amplicon sequencing and metagenomic viral sequencing to characterize the fecal bacteriome, and the fecal RNA and DNA virome, respectively. The overall study design is shown in Supplementary Figure S1.

Bacterial dysbiosis is associated with advanced stages of HIV infection

After quality control, sequences from 113 individuals were kept for 16S rRNA downstream analyses, including 38 HIV-negative and 75 PWH. First, we explored alpha diversity. Both richness and Shannon index were significantly lower in PWH with CD4 < 200 compared to HIV-negative individuals (Fig. 1B and Excel S1 for adjusted *p* values). Moderate positive correlations were found between CD4 T cell counts and both richness (rho = 0.47, unadjusted $p < 0.0001$) and Shannon index (rho = 0.49, unadjusted $p < 0.00001$, Excel S2). BMI was also weakly correlated with both richness and Shannon index (Excel S2). We further explored relationships between alpha diversity and CD4 T cell counts and other metadata by performing univariate linear regression analysis. CD4 T cell counts, BMI, antibiotic use, hospitalization, and sexual preference (tendency only) were associated with richness and Shannon index (Excel S3). After adjusting for BMI and sexual preference, associations between alpha diversity and CD4 T cell counts remained significant, though associations were lost when antibiotic use and hospitalization were included in the model (Excel S4). Next, we conducted an exploratory analysis to identify which metadata impacted the fecal bacteriome composition using PERmutational Multivariate ANalysis of VAriance (PERMANOVA). Genus-level bacterial communities were visualized using principal coordinate analysis (PCoA) based on the Bray–Curtis index of

Table 1 Cohort characteristics

Cohort characteristics	HIV-negative	HIV CD4 ≥ 350	HIV 200 ≤ CD4 < 350	HIV 50 ≤ CD4 < 200	HIV CD4 < 50	*Statistical significance*
Number of participants (***n***)	53	26	16	25	25	
Age (years)						***
Median	34	27.5	28.5	34	38	***p = 0.0005***
(IQR)	(29–38.5)	(24–33)	(22–40)	(27.5–39.5)	(32–45.5)	
Sex assigned at birth						*ns*
(male), ***n*** (%)	50 (94%)	26 (100%)	16 (100%)	22 (88%)	22 (88%)	*p = 0.2301*
Place of residence						*ns*
Mexico City, ***n*** (%)	49 (96%)	22 (85%)	13 (81%)	23 (82%)	19 (76%)	*p = 0.0632*
Other states of Mexico, ***n*** (%)	2 (4%)	4 (15%)	3 (19%)	2 (8%)	6 (24%)	
N/A, ***n***	2	0	0	0	0	
Sexual preference[a]						*
MSM, ***n*** (%)	48 (94%)	20 (95%)	12 (86%)	18 (78%)	17 (68%)	***p = 0.0151***
Heterosexual, ***n*** (%)	3 (6%)	1 (5%)	2 (14%)	5 (22%)	8 (32%)	
N/A, ***n***	2	5	2	2	0	
Tobacco use[a]						*ns*
n (%)	21 (41%)	8 (33%)	3 (21%)	8 (32%)	9 (36%)	*p = 0.7335*
N/A, ***n***	2	2	2	0	1	
Body mass index (BMI)						****
Median (IQR)	24.3 (22.4–26.7)	23.6 (21.2–25.4)	24.5 (20.5–26.3)	21.6 (19.1–22.9)	20.3 (18.5–22.8)	***p < 0.0001***
N/A, ***n***	1	3	3	1	0	
Gastrointestinal symptoms[a]						*ns*
(self-declared), ***n*** (%)	8 (16%)	2 (8%)	3 (20%)	5 (20%)	8 (33%)	*p = 0.2656*
N/A, ***n***	2	2	1	0	1	
Antibiotic use[b]						****
n (%)	7 (14%)	0 (0%)	6 (40%)	20 (80%)	21 (87%)	***p < 0.0001***
N/A, ***n***	2	2	1	0	1	
Hospitalization						****
n (%)	0 (0%)	0 (0%)	0 (0%)	5 (20%)	17 (68%)	***p < 0.0001***
CD4 T cell count						
(cells/mm^3), Median	908	488	280	81	21	****
(IQR)	(742–1124)	(426–675)	(231–323)	(64–120)	(12–38)	***p < 0.0001***
CD4 percent						
Median	40	23	12	8	4	****
(IQR)	(34–47)	(19–29)	(10–17)	(5–10)	(2–7)	***p < 0.0001***
CD4/CD8 ratio,						
Median	1.37	0.46	0.18	0.12	0.07	****
(IQR)	(1.03–1.83)	(0.32–0.65)	(0.15–0.31)	(0.07–0.15)	(0.04–0.12)	***p < 0.0001***
Plasma HIV load (HIV copies/mL)	-	53,344	91,933	225,253	380,428	**
Median		(10,942–	(11,999–	(80,378–	(104,556–	***p = 0.0011***
(IQR)		134,877)	292,737)	431,429)	966,670)	
Sample collection season						*Ns*
Winter, ***n*** (%)	14 (26%)	3 (12%)	2 (12%)	2 (8%)	4 (16%)	*p = 0.5073*
Spring, ***n*** (%)	12 (23%)	8 (31%)	6 (37%)	6 (24%)	7 (28%)	
Summer, ***n*** (%)	14 (26%)	7 (27%)	2 (12%)	6 (24%)	9 (36%)	
Autumn, ***n*** (%)	13 (25%)	8 (31%)	6 (37%)	11 (44%)	5 (20%)	

Statistical significance was assessed by two-sided Kruskal–Wallis and Fisher's exact tests for continuous and categorical variables, respectively

n Number of participants, *N/A* Not available, *IQR* Interquartile range, *ns* Not significant

[a] Self-declared

[b] Self-declared for non-hospitalized participants

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

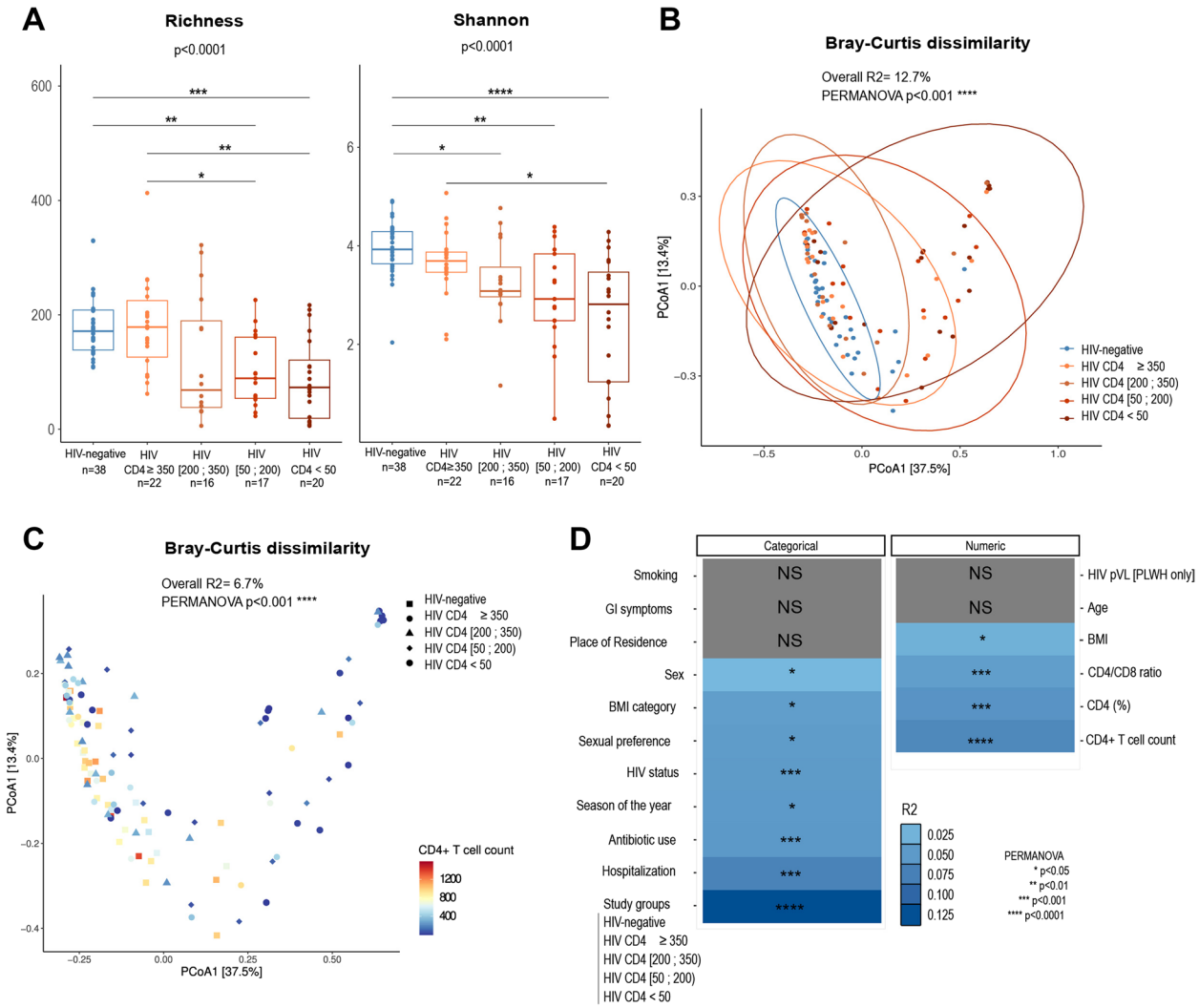

Fig. 1 Bacterial dysbiosis is associated with advanced stages of HIV infection. Alpha diversity is significantly lower in PWH with CD4 < 200 compared to HIV-negative individuals (**A**). Cross-sectional comparisons were performed on two alpha diversity metrics, richness (number of ASV observed), and the Shannon index across the five-study groups using a Kruskal–Wallis with a Dunn's post hoc test. The overall KW *p* value is shown above each graph. Only significant adjusted *p* values between pairs are shown. The fecal bacteriome is impacted by CD4 T cell counts. Principal coordinate analysis (PCoA) based on the Bray–Curtis dissimilarity index colored according to the five-group stratification (**B**) and CD4 T cell counts (**C**), respectively. First, ASV was agglomerated at genus level prior to calculating the Bray–Curtis dissimilarity index. The adonis2() function was used to calculate the effect size (R2) and *p* value. Ninety-five confidence ellipses are shown. Metadata that impacts the fecal bacteriome using univariate PERMANOVA (**D**). The impact of each metadata of interest is shown as a heatmap, where the intensity of the color blue indicates the strength of the effect size (R2), and is facetted by the type of variable (categorical or numeric). The season of the year refers to winter, spring, summer, and autumn (the time of year the stool sample was collected). ASVs were agglomerated at the genus level prior to calculating the Bray–Curtis dissimilarity index. The adonis2() function was used to calculate the effect size (R2) and *p* value. Abbreviations: % percent, ASV amplicon sequence variants, BMI body mass index, GI gastrointestinal, HIV human immunodeficiency virus, PERMANOVA permutational multivariate analysis of variance, pVL plasma viral load (HIV-1 copies/mL), *n* number of participants, NS not significant, **p* < 0.05, ***p* < 0.01, ****p* < 0.001, and *****p* < 0.0001

dissimilarity, and dots were colored according to study groups (Fig. 1B) and CD4 T cell counts (Fig. 1C). In univariate analysis, the variable that impacted most the fecal bacteriome was CD4 T cell counts (study groups) with an effect size (R^2) of 12.7%, and an unadjusted *p* value below 0.0001 (Fig. 1D and Excel S5). Differences were mainly centered between HIV-neg (reference) and HIV CD4 < 50 and HIV 50 ≤ CD4 < 200 as determined by pairwise adonis2 (Excel S6). The other variables that impacted the fecal bacterial composition were sex-assigned at birth,

sexual preference, BMI, antibiotic use, and hospitalization. Next, we conducted a multivariate analysis, where each term was tested sequentially (adonis2, by = "terms") and included sexual preference, BMI, age, HIV status, and CD4 + T cell counts. This analysis confirmed that CD4 T cell counts impacted the fecal bacterial ecology ($R^2 = 0.08$, $p = 0.0049^{**}$), independently of these factors. When including antibiotic use and hospitalization in the model, this impact was lost. Our results suggest that the individual contribution of CD4 T cell counts could not be established in relation to alpha and beta diversity, potentially due to moderate intercorrelations between the latter and antibiotic use and hospitalization (Supplementary Figure S2). Lastly, we explored differences in bacterial communities by differential abundance analysis (DAA) using Microbiome Multivariate Association with Linear Models (MaAsLin2). First, the top 3 phyla were Bacteroidota, Bacillota, and Pseudomonadota (formally Bacteroidetes, Firmicutes, and Proteobacteria, respectively), accounting for more than 97% of all phyla. At the genus level, *Prevotella*, *Bacteroides*, *Faecalibacterium*, *Alloprevotella*, and *UCG-002* accounted for more than 50% of all genera. For DAA, we ran two models: the first one included CD4 T cell counts and HIV status, while the second one considered the five-study group stratification, in addition to age, sexual preference, BMI, antibiotic use, and hospitalization. Both models demonstrated that hospitalization was positively associated with *Enterococcus* and *Streptococcus*, and negatively with alpha diversity (Shannon index), while HIV status was positively associated with *Coprococcus* (Excel S7). DAA results suggest that advanced stages of HIV infection are associated with bacterial dysbiosis as shown by others [6].

The fecal virome is composed of abundant and diverse communities of bacteriophages

We subsequently used metagenomic viral sequencing to characterize the fecal RNA and DNA virome. We obtained a median of 10.5 million sequences per sample (interquartile range (IQR) 9.6–11.6 million), with no significant differences between the five study groups (Kruskal–Wallis, $p = 0.20$). After quality control and subtraction of human, bacterial, and other non-viral, we obtained a median of 257,376 viral reads per sample (IQR: 76,151–717,790). All 145 individuals included in this study had viral sequences for downstream analysis. The most abundant viral sequences were from bacteriophages (42.5%), followed by plant viruses (37.5%), human viruses (17.4%), and, to a lesser extent, protist viruses (1.55%) (Fig. 2A, B). Protist virus sequences were found in 9% of samples. Interestingly, two samples had a high number of protist virus reads (> 10,000), and in both samples, these sequences belonged to *Cryptosporidium* virus, a double-stranded RNA virus harbored by the intracellular parasite *Cryptosporidium*. These samples belonged to two PWHs with severe immunodeficiency (CD4 = 11 and CD4 = 18), in agreement with the high prevalence of cryptosporidiosis in PWHs with low CD4 T cell counts [43]. Other identified viruses represented a minority of sequences and were attributed to dietary components. Bacteriophage sequences were identified in all 145 samples (100%) and represented over 10% of all reads in 71 samples (49.0%). Viral sequences belonging to the families *Microviridae* and *Inoviridae* and assigned to the morphology-based group Siphovirus were highly abundant (Fig. 2C, D). Phage communities were visualized using PCoA based on the Bray–Curtis index of dissimilarity and colored according to the study

(See figure on next page.)

Fig. 2 Fecal virome composition. **A** Host range of viral sequences identified in all samples ($n = 145$) based on mean relative abundance. **B** Abundance of sequences and proportion of positive samples across hosts. **C** Distribution of the most abundant bacteriophage classes and families/morphology-based groups in all samples ($n = 145$) based on mean relative abundance. **D** Abundance of virus sequences and proportion of positive samples across bacteriophage families/morphology-based groups detected in at least 5% of samples. *n*, number of participants. **E** Principal coordinate analysis based on the Bray–Curtis dissimilarity index colored according to the five-study groups for sequences belonging to bacteriophages only. When using the adonis2() function, the impact of CD4 T cell counts was minimal ($R^2 = 0.038$) with a PERMANOVA $p = 0.054$ (tendency). The only variable that impacted the fecal phageome composition was antibiotic use ($R.^2 = 0.023$, PERMANOVA $p = 0.0028$). **F** Richness is significantly lower in PWH with CD4 < 50 compared to HIV-negative individuals. No differences in the Shannon index were observed across the five-study groups. Cross-sectional comparisons were performed on two alpha diversity metrics, richness (number of ASV observed), and the Shannon index across the five-study groups using a Kruskal–Wallis with a Dunn's post hoc test. The overall KW *p* value is shown above each graph. Only significant adjusted *p* values between pairs are shown. **G** Inoviridae and Siphovirus are differentially abundant. Cross-sectional comparisons were performed across the five-study groups using a Kruskal–Wallis with a Dunn's post hoc test. The overall KW *p* value is shown above each graph. Only significant adjusted *p* values between pairs are shown. **H**, **I** Associations between bacteriophages and bacteria. Spearman correlations are shown as a correlogram where positive correlations are in blue and inverse correlations are in red. The intensity of the color represents the intensity of the spearman rho. Only significant correlations are shown. *P* values were corrected for multiple comparisons using the false discovery rate. Associations are shown overall (**H**) and stratified by HIV status (**I**). Abbreviations: % percent, ASV amplicon sequence variants, PERMANOVA permutational multivariate analysis of variance, PWH people living with HIV, *n* number of participants, KWp Kruskal–Wallis *p* value, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$

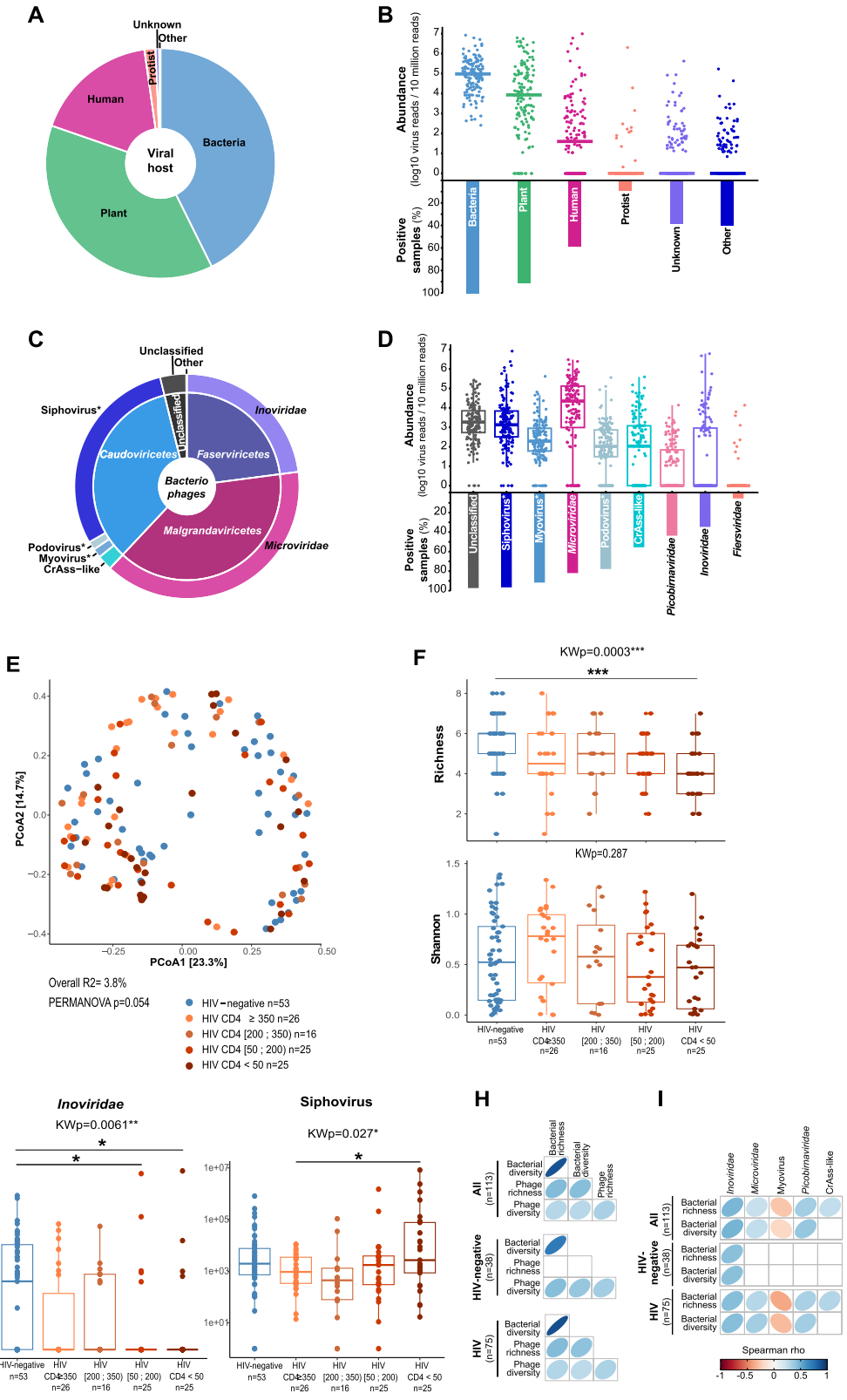

Fig. 2 (See legend on previous page.)

groups (Fig. 2E). Next, we assessed if CD4 T cell counts impacted the fecal phageome (bacteriophage communities) using univariate analysis (PERMANOVA) and found no impact (PERMANOVA $p=0.148$). When considering the five-study group stratification, the effect size was minimal ($R^2=0.038$) with a PERMANOVA $p=0.054$ (tendency). The only variable that impacted the fecal phageome composition, albeit minimally, was antibiotic use ($R^2=0.023$, $p=0.0028$). We found that phage richness was lower in HIV CD4 < 50 compared to HIV-negative (adjusted $p=0.00042$), while no differences in Shannon index (KWp = 0.2871) were observed across the five-study groups (Fig. 2F). We were unable to detect an expansion in *Caudoviricetes* nor did we find differences across our five-study groups for all other families or morphology-based groups except for *Inoviridae* and Siphovirus (Fig. 2G). *Inoviridae* sequences were less abundant in HIV CD4 < 50 and HIV 50 ≤ CD4 < 200 compared to HIV-negative (adjusted $p=0.0160$ and adjusted $p=0.0445$, respectively).

Relations between bacteriophages and bacteria

We then explored relationships between phages and bacteria. Phage diversity was moderately positively correlated with bacterial Shannon index and richness in both HIV-negative and PWH while phage richness was only moderately positively correlated with bacterial richness and diversity in PWH (Fig. 2H). Bacterial alpha diversity was positively correlated with *Inoviridae* in both HIV-negative and PWH. When stratifying by HIV status, we found that both the number of correlations between phages and bacteria and their strength increased in PWH. Myovirus were negatively correlated with both richness and Shannon index in PWH naïve to ART, while *Microviridae* and *Picobirnaviridae* were positively correlated with bacterial richness and Shannon index (Fig. 2I). CrAss-like phages were only positively correlated with bacterial richness in PWH.

Depletion of plant virus sequences in PWH with severe immunodeficiency

Plant virus sequences were identified in 132 samples (91.0%) and represented over 10% of the reads in 37 samples (25.5%). Twelve viral families were detected in at least 5% of samples (Fig. 3A, B) with a particularly high abundance of viruses from the *Virgaviridae* family and to a lesser extent from the *Tombusviridae, Betaflexiviridae*, and *Alphaflexiviridae* families. Notably, the viral species Pepper mild mottle virus (PMMoV) and topical soda apple mosaic virus (TSAMV), both belonging to the *Tobamovirus* genus within the *Virgaviridae* family, accounted for 27.8% and 35.7% of plant virus reads and were detected in 60.0% and 58.6% of samples, respectively. Surprisingly, plant virus sequences were significantly less abundant in PWH with CD4 < 50 compared to HIV-negative individuals and PWH with CD4 ≥ 350 (Fig. 3C). This effect was driven by the low abundance and prevalence of PMMoV and TSAMV in PWH with advanced immunodeficiency (Fig. 3E, F). No differences were detected for other plant viruses considered either individually or grouped (Fig. 3D and Excel S8).

Expansion of human viruses in PWH with severe immunodeficiency

Human viruses were detected in 57.9% of samples and 22 viral genera belonging to 10 viral families were identified (Fig. 4A). The viral families *Picornaviridae*, *Parvoviridae*, and *Caliciviridae* accounted for 64.0%, 25.2%, and 7.8% of human virus reads (Fig. 4A), respectively; due to the detection of very high number of sequences ($>10^6$) from the *Enterovirus* and *Salivirus*, *Sapovirus*, and *Bocaparvovirus* genera in a few samples (Fig. 4B, C). Remarkably, coxsackievirus A19 (*Enterovirus C* species, *Picornaviridae* family) sequences represented 96.8% of all reads in a sample from a PWH with severe immunodeficiency (CD4 = 28). Sequences from the viral families *Papillomaviridae*, *Anelloviridae*, and *Picornaviridae* were the most frequently detected in 33.1%, 31.7%, and 18.6%, respectively, of the samples. Some samples presented a higher diversity of human viruses, with 15 samples (10.3%) harboring at least four human virus genera (Fig. 4C, E). Notably, a sample from a PWH with AIDS (CD4 = 143) harbored sequences from 13 human virus genera belonging to seven viral families (Fig. 4C). Interestingly, PWH with CD4 < 350 presented an increased abundance, prevalence, and richness of human viruses compared to HIV-negative individuals (Fig. 4C, D, and E). The mean number of human virus genera detected was 0.4 in the HIV-negative group while it reached 2.4 and 2.6 in PWH with 50 ≤ CD4 < 200 and CD4 < 50, respectively.

(See figure on next page.)

Fig. 3 Composition and abundance of plant virus sequences. **A** Distribution of the most abundant plant virus families and species in all samples ($n=145$) based on mean relative abundance. **B** Abundance of virus sequences and proportion of positive samples for plant virus families detected in at least 5% of samples. Abundance of viral sequences and proportion of positive samples across study groups for all plant viruses (**C**), plant viruses except PMMoV and TSAMV (**D**), PMMoV (**E**), and TSAMV (**F**). Abbreviations: *n*, number of participants; pepper mild mottle mosaic virus (PMMoV), tropical soda apple mosaic virus (TSAMV); *$p<0.05$, **$p<0.01$, ***$p<0.001$, and ****$p<0.0001$

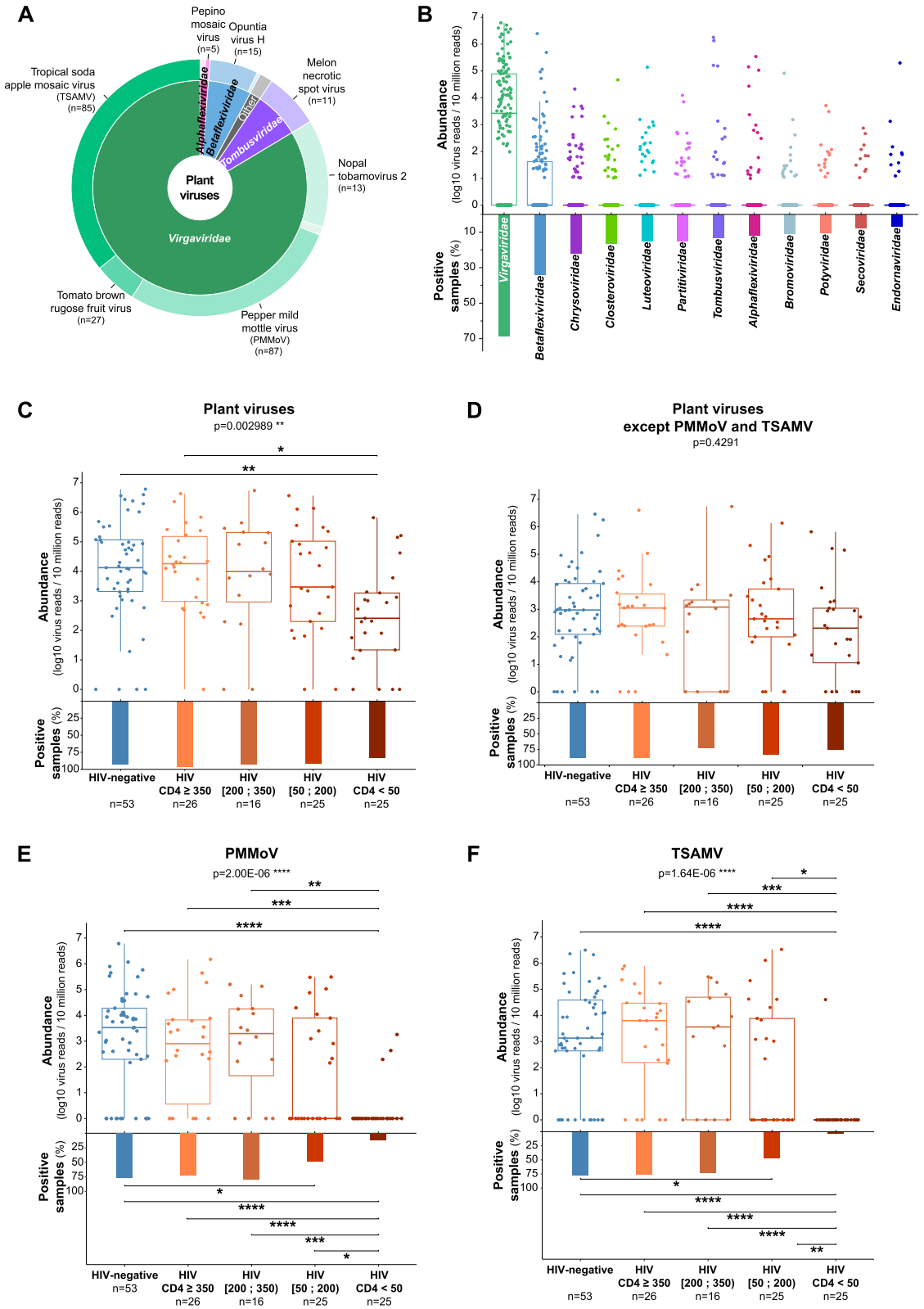

Fig. 3 (See legend on previous page.)

Differences in human virus communities between study groups were confirmed by principal coordinate analysis (PCoA) of beta-diversity based on the Bray–Curtis dissimilarity index (Fig. 4F). In univariate analysis, the five-study group stratification had the highest impact on the fecal human-host virome with an effect size of 13.3% and an unadjusted *p*-value below 0.0001. We observed a significant separation between HIV-negative individuals and PWH with $200 \leq CD4 < 350$, $50 \leq CD4 < 200$, and $CD4 < 50$ (PERMANOVA, unadjusted *p*-values of 0.029, 0.003, and 0.001, respectively).

Expansion of human viruses in PWH with advanced immunodeficiency

We next determined whether this expansion of human viruses in PWH with advanced immunodeficiency was driven by specific viral families. *Anelloviridae* sequences were markedly more abundant and more frequently detected in PWH with $CD4 < 200$ compared to HIV-negative individuals or PWH with $CD4 \geq 350$ (Fig. 5A and Excel S9). Strikingly, *Anelloviridae* sequences were not detected in any of the 53 samples from HIV-negative individuals while they were detected in 70% of the samples from PWH with $CD4 < 200$. It is typical of fecal viromes of healthy humans in which recovery of DNA eukaryotic viruses is rare [44]. Normalized number of *Anelloviridae* sequences showed a moderate negative correlation with CD4 counts (Fig. 5B). An expansion of *Papillomaviridae* and *Adenoviridae* sequences was also observed in PWH with advanced immunodeficiency (Fig. 5C, D). No differences were observed for other frequently detected viral families such as *Picornaviridae* and *Caliciviridae* (Excel S8). Next, we used MaAsLin2 to determine a multivariable association between metadata and the eukaryotic human virome. We ran two models, the first one included CD4 T cell counts and HIV status, while the second one considered the five-study group stratification, in addition to the metadata variables that were significantly different between study groups (age, sexual preference, BMI, antibiotic use, and hospitalization, Table 1). Both models consistently demonstrated that CD4 T cell counts were positively associated with *Anelloviridae*, *Adenoviridae*, and *Papillomaviridae* and negatively associated with PMMoV and TSAMV (Fig. 5E and Excel S9). No significant associations were detected for any of the other metadata variables. This multivariate analysis confirmed the expansion of human viruses belonging to the families *Anelloviridae*, *Adenoviridae*, and *Papillomaviridae* as well as the depletion of the plant viruses PMMoV and TSAMV in PWH with low CD4 T cell counts, independently of other factors.

Evolution of the bacterial and viral communities after antiretroviral therapy initiation

To address changes in the composition of the gut microbiome following ART initiation, we carried out a longitudinal characterization of the bacterial microbiome and virome in 14 PWH with AIDS ($CD4 < 200$) during the first 24 months of ART (see Supplementary Figure S1 for study design). At baseline, all PWH were naïve to ART and their median CD4 T cell count was 44 cells/µL (IQR 20.75–101) (Table S1). The predominant ART regimen was tenofovir disoproxil fumarate, emtricitabine, and efavirenz (TDF/FTC/EFV). Following ART initiation, pVL decreased significantly concurrent with CD4 T cell counts increase (Table S2). Importantly, all individuals had $pVL < 200$ after 2 to 12 months on ART, although an individual had low-level viremia after 24 months on ART. Despite a significant increase in CD4 T cell counts, nine individuals (64.3%) had $CD4 < 350$ after 24 months of ART, indicating that suboptimal immune recovery was frequent in this cohort of severely immunocompromised PWH. First, we explored how alpha diversity metrics were impacted by ART initiation. Richness increased (Friedman $p = 0.022$, Fig. 6A), while there was no significant increase in Shannon (Friedman $p = 0.66$, Fig. 6B). Phage richness and Shannon did not differ significantly across the five time-points (Friedman $p = 0.185$ and $p = 0.480$). As shown in Fig. 6C, D, highly individualized trajectories emerged, without a specific pattern of clustering of bacterial communities after 24 months on ART, as confirmed by PERMANOVA ($p > 0.05$). When exploring correlations between bacteria and phages at each time point, we found that bacterial alpha diversity metrics were strongly positively correlated with phage richness and Shannon, and *Inoviridae* and *Picobirnaviridae*

(See figure on next page.)

Fig. 4 Composition and abundance of human virus sequences. **A** Distribution of the most abundant human virus families and species in all samples ($n = 145$) based on mean relative abundance. **B** Abundance of virus sequences and proportion of positive samples for all human virus families identified. **C** Heatmap distribution of the $\log_{10}$ abundance of human virus genera in the 145 participants. Participants were ordered by study groups and by decreasing CD4 T cell counts (from left to right). [a]Mamastrovirus genus and MLB2 unclassified astrovirus. **D** Abundance of human virus sequences and proportion of positive samples across study groups. **E** Number of human virus genera identified across study groups. **F** Principal coordinate analysis based on the Bray–Curtis dissimilarity index colored according to the five-study groups. Abbreviations: *n*, number of participants; $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, and $****p < 0.0001$

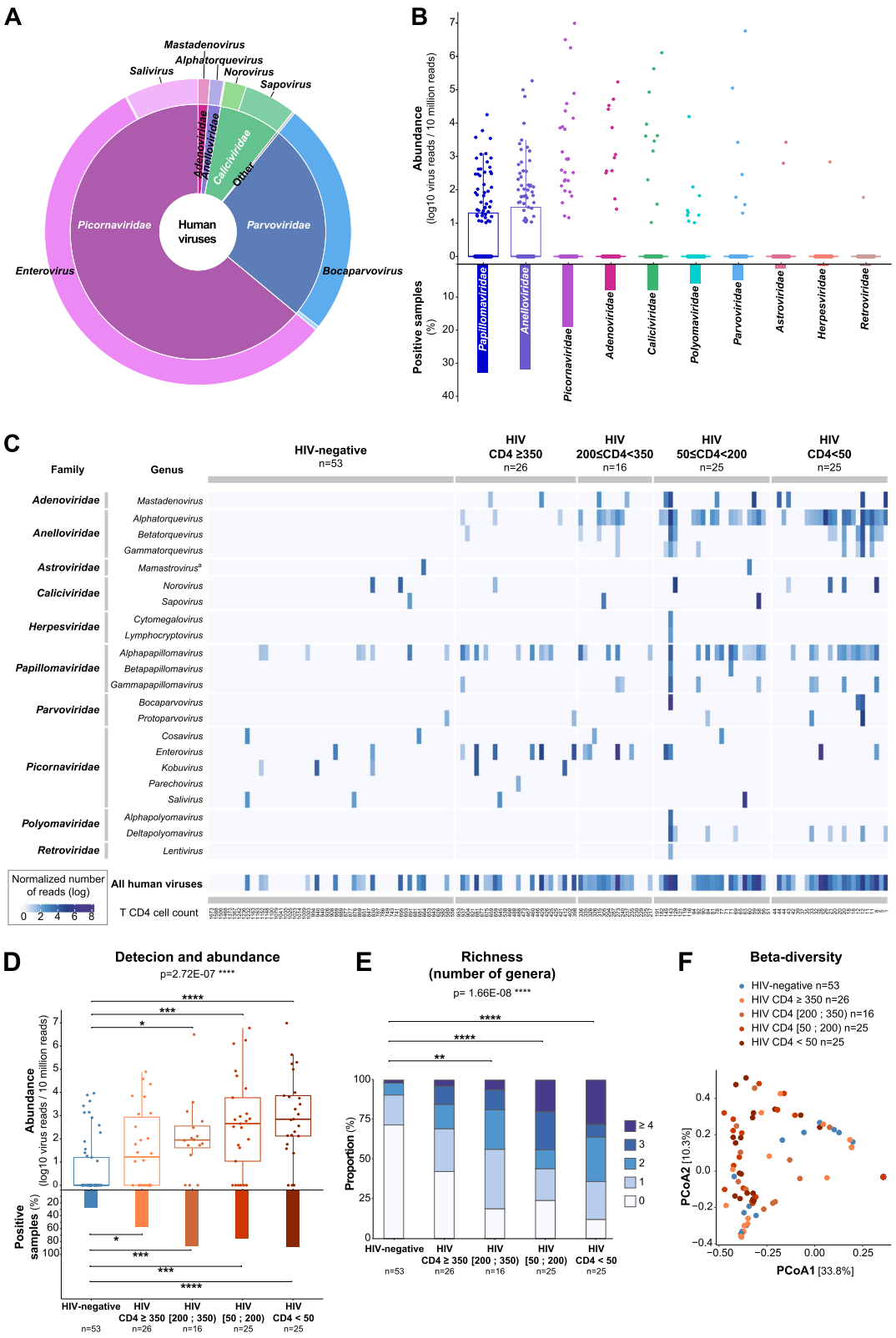

Fig. 4 (See legend on previous page.)

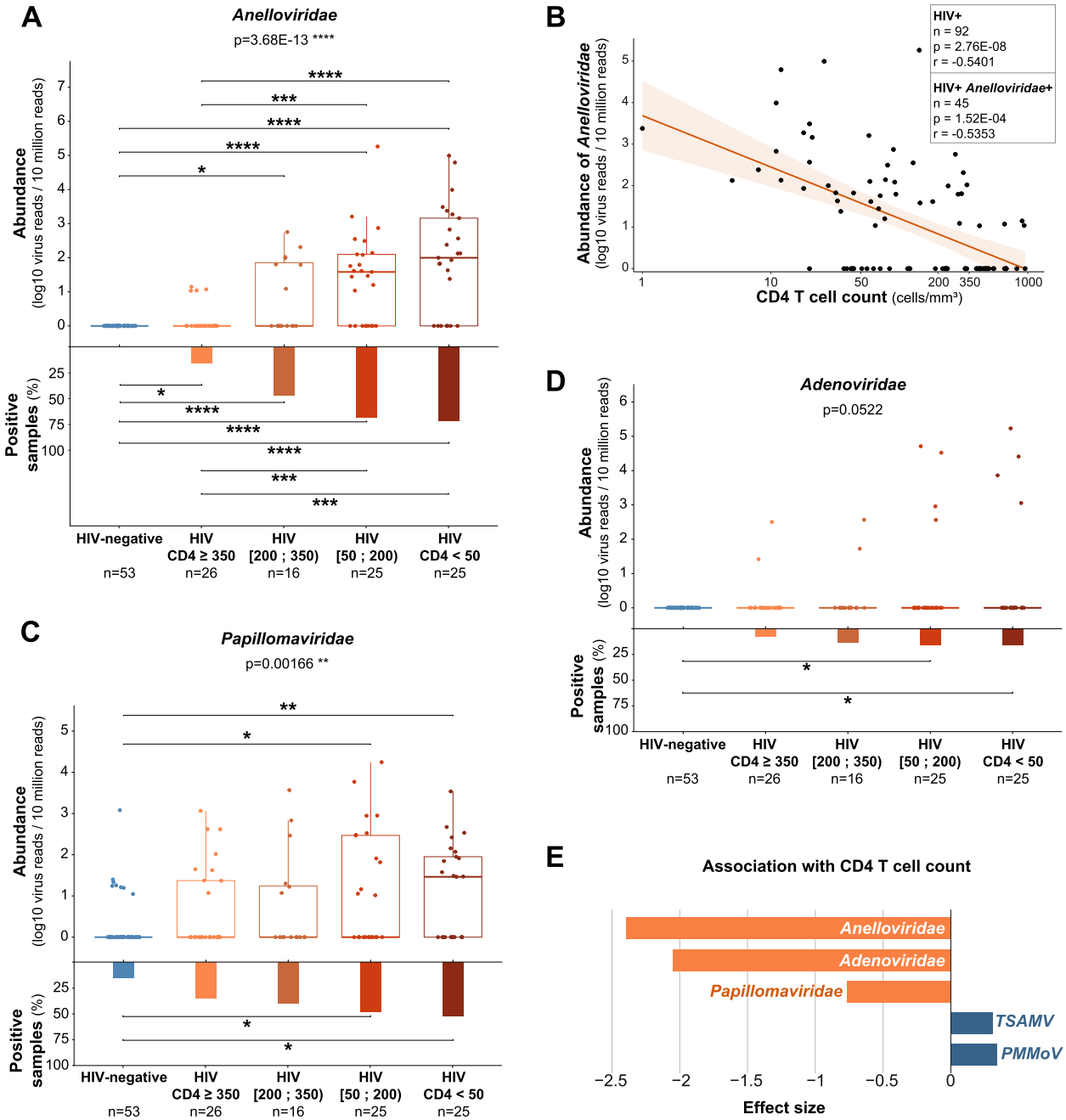

Fig. 5 Association between HIV-associated immunodeficiency, human virus sequence detection, and immune recovery. The abundance of viral sequences and proportion of positive samples across study groups for *Anelloviridae* (**A**), *Adenoviridae* (**C**), and *Papillomaviridae* (**D**) families. **B** Correlation between *Anelloviridae* sequence abundance and CD4 T cell counts. **E** Significant associations between CD4 T cell counts and the abundance of viral sequences using Microbiome Multivariate Association with Linear Models (MaAsLin2). Abbreviations: *n*, number of participants; **p* < 0.05, ***p* < 0.01, ****p* < 0.001, and *****p* < 0.0001

at baseline (Excel S10). After 2 months on ART, we found strong negative correlations between bacterial richness and Shannon and phage richness and Shannon. Bacterial richness and Shannon diversity were also negatively correlated with Myovirus, and bacterial Shannon diversity was negatively correlated with Podovirus. At month 12, a negative correlation was found between bacterial diversity and *Fiersviridae* (formally *Leviviridae*). At months 6 and 24 after ART initiation, no correlations were found between bacteria and phages.

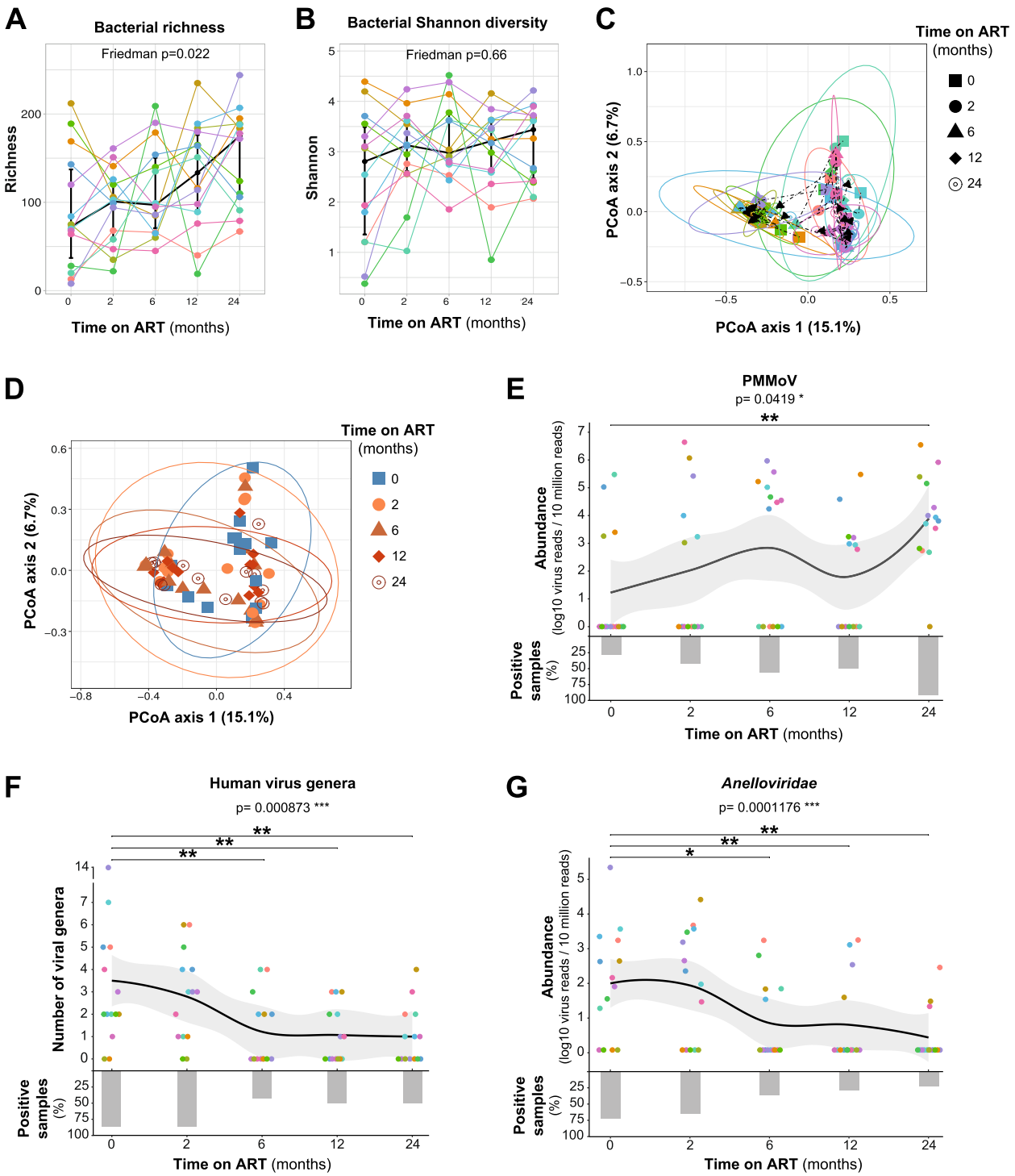

Fig. 6 Evolution of the bacterial and viral communities after ART initiation. Fourteen PWH with AIDS (CD4 < 200) were followed during the first 24 months of ART. **A**, **B** Richness increased after ART initiation while there was no significant increase in Shannon. Longitudinal comparisons were performed using the Friedman test. Individual colors were used for each participant. The median and interquartile range is also shown in black. **C**, **D** Principal coordinate analysis (PCoA) based on the Bray–Curtis dissimilarity index colored according to each individual where arrows indicate the trajectory of each individual (**C**) and according to each time point on ART (**D**). Ninety-five confidence ellipses are shown. No specific pattern of clustering of bacterial and phage communities after 24 months on ART was found, as confirmed by PERMANOVA ($p > 0.05$). **E** The abundance of viral sequences and proportion of positive samples during the first 24 months of ART for PMMoV (**E**), human viruses (**F**), and Anelloviridae (**G**). Abbreviations: % percent, AIDS acquired immunodeficiency syndrome, ART antiretroviral therapy, PCoA principal coordinate analysis, Pepper Mild Mottle Mosaic Virus (PMMoV), *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$

Next, we performed a longitudinal analysis of the fecal eukaryotic virome after ART initiation. We observed a significant increase in the abundance of the plant viruses PMMoV and TSAMV after 24 months of ART compared to baseline ($p=0.00518$ and $p=0.0310$, respectively) (Fig. 6E and Excel S11). Conversely, the number of human virus genera as well as their abundance and prevalence progressively decreased during the 24 months following ART initiation (Fig. 6F and Excel S11). Notably, the abundance of *Anelloviridae* sequences was significantly lower after 6, 12, and 24 months of ART compared to the baseline ($p=0.0415$, $p=0.00805$, and $p=0.00592$, respectively) (Fig. 6G and Excel S11). No significant differences were observed for other plant and human viruses (Excel S11). Pre-ART samples harbored significantly less PMMoV and TSAMV sequences, and more *Anelloviridae* sequences compared to HIV-negative individuals ($p=0.00635$, $p=0.0440$, and $p=4.04E-09$) (Excel S12), in agreement with our results described above for the complete cross-sectional cohort. After 24 months on ART, the abundance of PMMoV and TSAMV was similar when comparing treated PWH and HIV-negative individuals. However, *Anelloviridae* sequences were still detected in three individuals (21.4%) with CD4 < 300 after 24 months on ART and their abundance and prevalence were significantly higher compared to HIV-negative individuals ($p=0.00760$) (Excel S12), potentially reflecting the suboptimal immune recovery in some PWH. Altogether, these results revealed a gradual restoration of the virome in PWH after ART initiation and a potential association between immune reconstitution and *Anelloviridae* sequence detection.

Baseline CD4 T cell counts and human Anelloviridae detection can predict immune recovery during ART

Our results demonstrated a remarkable expansion of *Anelloviridae* sequences in PWH with severe immunodeficiency, suggesting that the presence of human Anelloviruses could be a marker of the immune status. We next assessed whether the detection of human *Anelloviridae* sequences in fecal samples from untreated PWH could predict suboptimal immune recovery after ART initiation. We first performed univariable logistic regression analyses in a subset of PWH who had follow-up clinical HIV data (pVL and CD4 T cell counts) up to 60 months after ART initiation (see Supplementary Figure S1). Individuals with baseline CD4 T cell count below 350 cells/mm^3 and sustained plasma HIV viral load < 200 copies/ml after 6 months of ART were included ($n=42$). Fifteen participants (35.7%) had at least one CD4 T cell count above 500 cells/mm^3 and were classified as adequate immune responders (AIR) whereas seventeen individuals (40.5%) had all CD4 T cell counts below 350 cells/mm^3 and were classified as poor immunological responders (PIR). Low baseline CD4 T cell counts and *Anelloviridae* sequence detection before ART initiation were significantly associated with poor immune recovery ($p=0.0173$ and $p=0.0078$, respectively) while no association was identified for other variables (p-values > 0.20 for age, sex assigned at birth, BMI, and plasma HIV load) (Fig. 7 and Excel S13). These two potential predictors were then included in a multivariable logistic regression model. Importantly, both low baseline CD4 T cell counts and detection of human *Anelloviridae* sequences in stool samples before ART initiation independently predicted poor immune recovery ($p=0.0279$ and $p=0.0205$, Excel S13).

Discussion

The overall objective of this study was to gain a better understanding of the viral component of the gut microbiome and its association with HIV infection and immunodeficiency. We demonstrate a striking expansion in *Anelloviridae* intimately linked to CD4 T cell counts and severe immunodeficiency, bacterial dysbiosis linked to advanced HIV disease, and no major change in bacteriophage populations. Short-term ART (24 months) had an impact on the eukaryotic virome, with a gradual restoration of plant viruses and a decrease in human anellovirus sequences. Furthermore, human *Anelloviridae* sequences predicted poor immune recovery after ART initiation.

When analyzing the bacterial component of the gut microbiome, we found that advanced HIV infection and immunodeficiency were linked to bacterial dysbiosis as shown by others [6, 45]. We report a decrease in alpha diversity metrics and a bacterial signature of increased *Enterococcus*, *Streptococcus*, and *Coprococcus*, linked to advanced HIV disease. The lack of a more pronounced HIV-associated bacterial signature was a surprise to us, given the wide range of CD4 T cell counts and the inclusion of PWH with AIDS-defining conditions, as shown by others [8]. After ART initiation, we found large interpersonal differences in the gut bacterial composition without a specific pattern of clustering of bacterial communities after 24 months on ART and increased bacterial richness. A recent study demonstrated differences in cohorts from different geographical locations (USA and two sub-Saharan countries: Uganda and Botswana), suggesting that the gut microbiome is highly contextualized to each population and their characteristics [9]. Our results were likely affected by the inclusion of PWH with severe immunodeficiency and suboptimal immune recovery despite short-term ART. Also, we included 14 PWH in the longitudinal study and most likely did not have enough power to observe changes associated with ART initiation. High intra- and interindividual variability

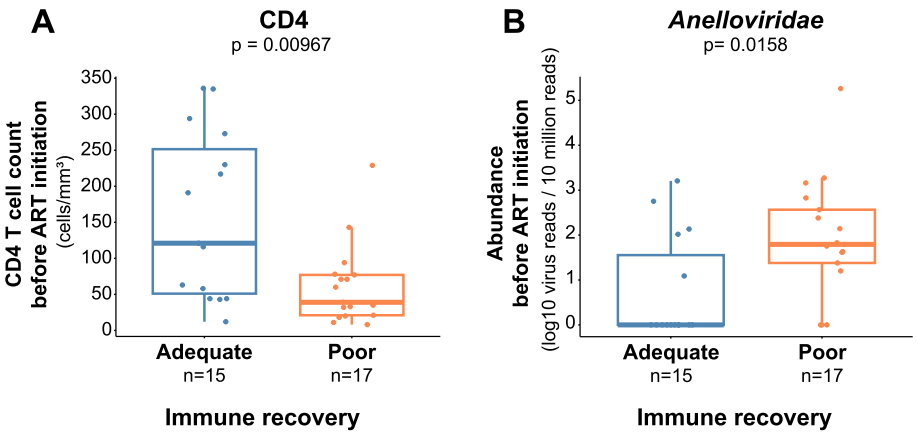

Fig. 7 Potential predictors of immune recovery on ART. Baseline CD4 T cell counts (**A**) and *Anelloviridae* sequences (**B**) in adequate and poor immune responders. Abbreviations: *n*, number of participants

could have masked the impact of ART on the gut microbiome [46].

The phageome of untreated PWH was dominated by temperate-tailed phages mainly *Caudoviricetes* and within this order, *Siphovirus* and lytic phages of the *Microviridae* family. Surprisingly, we did not detect an expansion in *Caudoviricetes* as shown in Crohn's disease, which could be indicative of gut virome dysbiosis [16, 47, 48]. Either *Caudoviricetes* populations remain unaffected by HIV-1 infection and profound immunodeficiency or we were unable to detect it. We did find a decrease in *Inoviridae*, a group of filamentous phages, in PWH with severe immunodeficiency. *Inoviridae* are known to infect Gram-negative bacteria without killing their host, establishing a productive chronic infection [49, 50]. As shown in previous studies, external factors, such as antibiotics induce the switch from lysogenic to lytic [51].

Our study showed drastic changes in the gut virome composition in immunocompromised PWH, characterized by a depletion in plant viruses, an expansion in Adenovirus, *Papillomaviridae*, and Adenoviridae sequences. First, plant virus sequences were less frequently and less abundantly detected in PWH with severe immunodeficiency compared to HIV-negative individuals and PWH with CD4 > 200. Notably, this depletion reverted after 24 months of ART and was specific to pepper mild mottle virus (PMMoV) and tropical apple soda mosaic virus (TSAMV). These two plant viruses share high-nucleotide sequence identity (80.4%) and are non-enveloped, positive-sense single-stranded RNA viruses belonging to the *Tobamovirus* genus (*Virgaviridae* family) [52, 53]. They mainly infect *Capsicum* spp (chili, bell, and ornamental pepper plants) and can cause mild to severe symptoms in infected plants [52, 54]. PMMoV has been found in pepper-containing food products and in drinking water sources [55–57]. Due to its high abundance in human fecal samples [56, 58] and its long persistence in environmental samples, PMMoV has been used as an indicator of human fecal contamination in wastewater [52, 59, 60]. TSAMV was identified more recently and has been far less studied than PMMoV [52]. Interestingly, it was abundantly and frequently detected in stool samples from Mexican children [61]. Since PMMoV (and likely TSAMV) RNA in fecal samples may result from the consumption of pepper-containing food products and potentially contaminated water, variability in their fecal shedding could be due to differences in diet. Gastrointestinal disorders and higher hospitalization rates in PWH with severe immunosuppression might be associated with a lower consumption of pepper-containing food products (chili pepper, dry spices, and sauces) and/or consumption of higher quality drinking-water, possibly resulting in a lower abundance of PMMoV and TSAMV in stool samples. Dietary information was unavailable in our study and these hypotheses would require further specific investigations.

Next, we showed a moderate expansion of human *Adenoviridae* sequences in PWH with severe immunodeficiency, in agreement with previous studies carried out in Uganda and Mozambique [6, 20]. However, this expansion was less marked in our study, with a lower prevalence of *Adenoviridae* sequences in Mexico compared to these East African cohorts. Similarly, a study analyzing the geographic variation of the eukaryotic gut virome revealed a higher prevalence of adenoviruses in diarrheic children from Gambia compared to Australia and Seattle [62]. These results highlight the impact of the geographic region on the composition of the gut virome, possibly due to environmental, socioeconomic, and host factors such as climate, water quality, diet, lifestyle, health status,

and age [63, 64]. Expansion of enteric viruses in severely immunocompromised might contribute to AIDS pathogenesis by damaging the intestinal barrier and potentially promoting chronic immune activation. In non-human primate models, pathogenic simian immunodeficiency virus (SIV) infection was associated with an expansion of adenoviruses [22, 23]. Importantly, these studies demonstrated a significant association between gastrointestinal disease, AIDS-related mortality, and *Adenoviridae* sequence detection in SIV-infected rhesus macaques. Additional studies are needed to determine whether virome alterations may contribute to HIV-associated enteropathy and disease progression in PWH.

We also showed an expansion of *Papillomaviridae* sequences in stool samples from PWH with AIDS compared to HIV-negative individuals. Human papillomavirus (HPV) sequences detected in fecal samples could originate from the anal canal and potentially from the oral cavity [65]. Interestingly, high anal HPV prevalence was reported among young HIV-positive and HIV-negative MSM, with amplifying effects of HIV infection and HIV-related immunosuppression [66]. Alpha-, beta-, and gamma HPV were found to be significantly more prevalent in the anal canal of PWH, with genus- and species-specific results depending on the study [61, 67, 68]. In the oral cavity, HPV prevalence was found to be higher in PWH and was strongly associated with low CD4 T cell counts [69, 70]. HPV expansion may be explained by the higher incidence, persistence, and recurrence of HPV infections, resulting in a higher risk of HPV-associated malignancies in PWH with persistently low CD4 T cell counts [71, 72].

A striking association between HIV-associated immunodeficiency and *Anelloviridae* sequence detection was found. *Anelloviridae* are small non-enveloped DNA viruses with circular, negative-sense single-stranded genomes. They are extremely diverse, highly prevalent, and present in a wide range of human biological samples [73–76]. Anelloviruses cause persistent infections in humans that have not been associated with the disease. Several studies have demonstrated a marked expansion of anellovirus sequences in plasma samples from immunocompromised hosts, in particular transplant recipients receiving immunosuppressive therapy [75, 77] and PWH with low CD4 T cell counts [78–80]. Interestingly, high anellovirus loads have been associated with an increased risk of infection while low anellovirus loads were associated with a higher risk of graft rejection in solid-organ transplant recipients [81]. Due to the extremely high prevalence of anelloviruses and their association with immunosuppression, plasma anellovirus load has been proposed as a marker of the overall state of immunosuppression [80, 82–84]. Our longitudinal study demonstrated that short-term ART was associated with a gradual decrease in the abundance of human *Anelloviridae* sequences in severely immunocompromised PWH, suggesting that *Anelloviridae* detection could be a useful marker for immune reconstitution after ART initiation as already suggested previously [78]. We further showed that baseline human anellovirus detection and CD4 T cell count were significantly associated with immune recovery after ART initiation. While low nadir CD4 T cell count has been consistently associated with poor immune recovery [85–88], the predictive value of anellovirus load on immune reconstitution led to contrasting results in two recent studies [89, 90]. Our results suggest that anellovirus abundance in fecal samples could be used as a predictive marker for immune recovery in PWH, extending previous results obtained in plasma samples [90]. Although both peripheral blood and gastrointestinal tract are relevant components to characterize HIV disease progression and immune recovery [1, 91], there are notable differences in the kinetics of CD4 T cell restoration after ART initiation between these anatomical sites [92–94]. Comparative studies of anellovirus prevalence and abundance at different anatomical sites in PWH would be needed to assess the value of enteric versus plasma anellovirus load as an immune marker and as a predictor of CD4 restoration after ART initiation.

Some limitations of this study should be acknowledged. Dietary information was not collected although diet has been previously shown to have significant effects on the microbiome composition. Associations between diet and the abundance and prevalence of plant viruses in the gut could notably be an interesting area for future investigations. Other limitations include the inclusion of severely immunocompromised PWH with antibiotic use and the need for hospitalization, which albeit truly informative for the virome, might have obscured bacteriome alterations. Also the characteristics of our cohort could affect the generalizability of these results. Notably, our cohort was predominantly composed of males and men who self-declared as MSM contrasting with the small number of females and MSW included in this study. Although this is representative of the HIV-1 epidemic in Mexico, we certainly lacked statistical power to accurately control this covariate (sexual preference) and confounders. We leveraged these limitations by conducting both a cross-sectional with a large number of PWH ($n=92$) and a control cohort of HIV-seronegative individuals at high risk of HIV-1 infection ($n=53$). Our longitudinal study included complete data

at five-time points. We included both 16rRNA bacterial sequencing and comprehensive shotgun RNA and DNA virome analysis of fecal samples to characterize the bacterial and viral communities.

Conclusion

Overall, our data revealed a striking expansion in the abundance and prevalence of *Anelloviridae*, *Adenoviridae*, and *Papillomaviridae* and a decreased abundance of sequences belonging to two plant viruses from the *Tobamovirus* genus, a reduction in bacterial alpha diversity, and a decrease in *Inoviridae* bacteriophage sequences in stool samples of PWH with severe immunodeficiency (CD4 < 200). Indeed, advanced HIV-1 infection was associated with marked alterations in the virome composition, in particular a remarkable expansion of human anelloviruses. Short-term ART (24 months) was linked to a significant decrease in human *Anelloviridae* sequences and the detection of anellovirus sequences at baseline independently predicted poor immune recovery, as did low CD4 T cell counts. In addition to CD4 T cell counts, anellovirus sequence detection might be useful to predict and monitor immune recovery. This study confirms data on the bacteriome and expands our knowledge on the viral component of the gut microbiome in HIV-1 infection.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40168-024-01925-7.

Supplementary Material 1.

Acknowledgements
The authors would like to thank all the participants for their valued contribution to HIV research. We would also like to thank the Laboratorio de Diagnóstico Viral (LDV, CIENI-INER) head by QC. Ramón Hernández Juan, blood collection team (CIENI-INER) headed by QC. Edna Rodríguez.

Authors' contributions
CB, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Writing – original draft, Writing – review and editing AMP, Conceptualization, Data curation, Resources, Writing – editing,, Resources, Writing – review and editing MCT, Investigation, Methodology, Writing – review and editing KRM, Conceptualization, Data curation, Resources, Writing – editing, AR, Investigation SAR, Conceptualization, Funding acquisition, Investigation, Resources, Writing – review and editing DGR, Investigation, Methodology, Writing – review and editing GRT, Conceptualization, Funding acquisition, Investigation, Resources, Writing – review and editing SPC, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Writing – original draft, Writing – review and editing.

Funding
This work was supported by grants from the Programa Presupuestal P016, Anexo 13 del Decreto del Presupuesto de Egresos de la Federación.

Data availability
Raw sequencing reads were deposited at the National Center for Biotechnology Information-Sequence Read Archive (NCBI-SRA) under projects PRJNA901491 (cross-sectional 16S rRNA sequences) and PRJNA996141 (longitudinal 16S rRNA sequences) and PRJNA1068646 (metagenomic sequences).

Declarations

Ethics approval and consent to participate
This study was approved by the Ethics, Research and Safety Committees of the Instituto Nacional de Enfermedades Respiratorias Ismael Cosío Villegas (INER) under approval numbers B36-12, C59-15, and C15-16. All participants were informed of the study objectives and signed an informed consent. All participants were 18 years or older. The study was conducted according to the Declaration of Helsinki of 1964 and later amendments.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 17 July 2024 Accepted: 3 September 2024
Published online: 17 October 2024

References

1. Brenchley JM, Douek DC. HIV infection and the gastrointestinal immune system. Mucosal Immunol. 2008;1(1):23–30.
2. Brenchley JM, Price DA, Schacker TW, Asher TE, Silvestri G, Rao S, et al. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. Nat Med. 2006;12(12):1365–71.
3. Brenchley JM, Douek DC. Microbial translocation across the GI tract. Annu Rev Immunol. 2012;30:149–73.
4. Nganou-Makamdop K, Douek DC. The gut and the translocated microbiomes in HIV infection: current concepts and future avenues. Pathog Immun. 2024;9(1):168–94.
5. Vujkovic-Cvijin I, Sortino O, Verheij E, Sklar J, Wit FW, Kootstra NA, Sellers B, Brenchley JM, Ananworanich J, Loeff MSV, Belkaid Y, Reiss P, Sereti I. HIV-associated gut dysbiosis is independent of sexual practice and correlates with noncommunicable diseases. Nat Commun. 2020;11(1):2448. https://doi.org/10.1038/s41467-020-16222-8.
6. Rocafort M, Noguera-Julian M, Rivera J, Pastor L, Guillén Y, Langhorst J, et al. Evolution of the gut microbiome following acute HIV-1 infection. Microbiome. 2019;7(1):73.
7. Borgognone A, Noguera-Julian M, Oriol B, Noël-Romas L, Ruiz-Riol M, Guillén Y, et al. Gut microbiome signatures linked to HIV-1 reservoir size and viremia control. Microbiome. 2022;10(1):59.
8. Noguera-Julian M, Rocafort M, Guillén Y, Rivera J, Casadellà M, Nowak P, et al. Gut microbiota linked to sexual preference and HIV infection. EBioMedicine. 2016;5:135–46.
9. Burkhart Colorado AS, Lazzaro A, Neff CP, Nusbacher N, Boyd K, Fiorillo S, et al. Differential effects of antiretroviral treatment on immunity and gut microbiome composition in people living with HIV in rural versus urban Zimbabwe. Microbiome. 2024;12:18.
10. Rocafort M, Gootenberg DB, Luévano JM, Paer JM, Hayward MR, Bramante JT, et al. HIV-associated gut microbial alterations are dependent on host and geographic context. Nat Commun. 2024;15(1):1055.
11. Brenchley JM, Ortiz AM. Microbiome studies in non-human primates. Curr HIV/AIDS Rep. 2021;18(6):527–37.
12. Ortiz AM, Flynn JK, DiNapoli SR, Sortino O, Vujkovic-Cvijin I, Belkaid Y, et al. Antiretroviral therapy administration in healthy rhesus macaques is associated with transient shifts in intestinal bacterial diversity and modest immunological perturbations. J Virol. 2019;93(18):e00472-e519.
13. Ortiz AM, Flynn JK, DiNapoli SR, Vujkovic-Cvijin I, Starke CE, Lai SH, et al. Experimental microbial dysbiosis does not promote disease progression in SIV-infected macaques. Nat Med. 2018;24(9):1313–6.
14. Liang G, Bushman FD. The human virome: assembly, composition and host interactions. Nat Rev Microbiol. 2021;19(8):514–27.
15. Khan Mirzaei M, Xue J, Costa R, Ru J, Schulz S, Taranu ZE, et al. Challenges of studying the human virome - relevant emerging technologies. Trends Microbiol. 2021;29(2):171–81.
16. Clooney AG, Sutton TDS, Shkoporov AN, Holohan RK, Daly KM, O'Regan O, et al. Whole-virome analysis sheds light on viral dark matter in inflammatory bowel disease. Cell Host Microbe. 2019;26(6):764-778.e5.

17. Cao Z, Fan D, Sun Y, Huang Z, Li Y, Su R, et al. The gut ileal mucosal virome is disturbed in patients with Crohn's disease and exacerbates intestinal inflammation in mice. Nat Commun. 2024;15(1):1638.
18. Zhao G, Vatanen T, Droit L, Park A, Kostic AD, Poon TW, et al. Intestinal virome changes precede autoimmunity in type I diabetes-susceptible children. Proc Natl Acad Sci U S A. 2017;114(30):E6166–75.
19. Legoff J, Resche-Rigon M, Bouquet J, Robin M, Naccache SN, Mercier-Delarue S, et al. The eukaryotic gut virome in hematopoietic stem cell transplantation: new clues in enteric graft-versus-host disease. Nat Med. 2017;23(9):1080–5.
20. Monaco CL, Gootenberg DB, Zhao G, Handley SA, Ghebremichael MS, Lim ES, et al. Altered virome and bacterial microbiome in human immunodeficiency virus-associated acquired immunodeficiency syndrome. Cell Host Microbe. 2016;19(3):311–22.
21. Villoslada-Blanco P, Pérez-Matute P, Íñiguez M, Recio-Fernández E, Blanco-Navarrete P, Metola L, et al. Integrase inhibitors partially restore bacterial translocation, inflammation and gut permeability induced by HIV infection: impact on gut microbiota. Infect Dis Ther. 2022;11(4):1541–57.
22. Handley SA, Thackray LB, Zhao G, Presti R, Miller AD, Droit L, et al. Pathogenic simian immunodeficiency virus infection is associated with expansion of the enteric virome. Cell. 2012;151(2):253–66.
23. Handley SA, Desai C, Zhao G, Droit L, Monaco CL, Schroeder AC, et al. SIV infection-mediated changes in gastrointestinal bacterial microbiome and virome are associated with immunodeficiency and prevented by vaccination. Cell Host Microbe. 2016;19(3):323–35.
24. D'arc M, Furtado C, Siqueira JD, Seuánez HN, Ayouba A, Peeters M, et al. Assessment of the gorilla gut virome in association with natural simian immunodeficiency virus infection. Retrovirology. 2018;15(1):19.
25. Briceño O, Peralta-Prado A, Garrido-Rodríguez D, Romero-Mora K, Chávez-Torres M, Pinto Cardoso S, et al. Double-negative T cell number and phenotype alterations before and after effective antiretroviral treatment in persons living with HIV. AIDS Res Hum Retroviruses. 2023;39(3):104–13.
26. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. 2013;41(1):e1.
27. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(8):852–7.
28. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581–3.
29 Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41(Database issue):D590-6.
30. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome. 2018;6(1):90.
31. Victoria JG, Kapoor A, Li L, Blinkova O, Slikas B, Wang C, et al. Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. J Virol. 2009;83(9):4642–51.
32. Temmam S, Monteil-Bouchard S, Robert C, Pascalis H, Michelle C, Jardot P, et al. Host-associated metagenomics: a guide to generating infectious RNA viromes. PLoS One. 2015;10(10):e0139810.
33. Naccache SN, Federman S, Veeraraghavan N, Zaharia M, Lee D, Samayoa E, et al. A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. Genome Res. 2014;24(7):1180–92.
34. Zheng H, Jia H, Shankar A, Heneine W, Switzer WM. Detection of murine leukemia virus or mouse DNA in commercial RT-PCR reagents and human DNAs. PLoS One. 2011;6(12):e29050.
35. Naccache SN, Greninger AL, Lee D, Coffey LL, Phan T, Rein-Weston A, et al. The perils of pathogen discovery: origin of a novel parvovirus-like hybrid genome traced to nucleic acid extraction spin columns. J Virol. 2013;87(22):11966–77.
36. Smuts H, Kew M, Khan A, Korsman S. Novel hybrid parvovirus-like virus, NIH-CQV/PHV, contaminants in silica column-based nucleic acid extraction kits. J Virol. 2014;88(2):1398.
37. Asplund M, Kjartansdóttir KR, Mollerup S, Vinner L, Fridholm H, Herrera JAR, et al. Contaminating viral sequences in high-throughput sequencing viromics: a linkage study of 700 sequencing libraries. Clin Microbiol Infect. 2019;25(10):1277–85.
38. Rosseel T, Pardon B, De Clercq K, Ozhelvaci O, Van Borm S. False-positive results in metagenomic virus discovery: a strong case for follow-up diagnosis. Transbound Emerg Dis. 2014;61(4):293–9.
39. Greninger AL, Naccache SN, Messacar K, Clayton A, Yu G, Somasekar S, et al. A novel outbreak enterovirus D68 strain associated with acute flaccid myelitis cases in the USA (2012–14): a retrospective cohort study. Lancet Infect Dis. 2015;15(6):671–82.
40. Siqueira JD, Dominguez-Bello MG, Contreras M, Lander O, Caballero-Arias H, Xutao D, et al. Complex virome in feces from Amerindian children in isolated Amazonian villages. Nat Commun. 2018;9(1):4270.
41. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, et al. Multivariable association discovery in population-scale meta-omics studies. PLoS Comput Biol. 2021;17(11):e1009442.
42. Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al. Microbiome differential abundance methods produce different results across 38 datasets. Nat Commun. 2022;13(1):342.
43. Ahmadpour E, Safarpour H, Xiao L, Zarean M, Hatam-Nahavandi K, Barac A, Picot S, Rahimi MT, Rubino S, Mahami-Oskouei M, Spotin A, Nami S, Baghi HB. Cryptosporidiosis in HIV-positive patients and related risk factors: A systematic review and meta-analysis. Parasite. 2020;27:27. https://doi.org/10.1051/parasite/2020025.
44. Carding SR, Davis N, Hoyles L. Review article: the human intestinal virome in health and disease. Aliment Pharmacol Ther. 2017;46(9):800–15.
45. Guillén Y, Noguera-Julian M, Rivera J, Casadellà M, Zevin AS, Rocafort M, et al. Low nadir CD4+ T-cell counts predict gut dysbiosis in HIV-1 infection. Mucosal Immunol. 2019;12(1):232–46.
46. Guthrie L, Spencer SP, Perelman D, Van Treuren W, Han S, Yu FB, et al. Impact of a 7-day homogeneous diet on interpersonal variation in human gut microbiomes and metabolomes. Cell Host Microbe. 2022;30(6):863-874.e4.
47. Norman JM, Handley SA, Baldridge MT, Droit L, Liu CY, Keller BC, et al. Disease-specific alterations in the enteric virome in inflammatory bowel disease. Cell. 2015;160(3):447–60.
48. Tun HM, Peng Y, Massimino L, Sin ZY, Parigi TL, Facoetti A, et al. Gut virome in inflammatory bowel disease and beyond. Gut. 2024;73(2):350–60.
49. Dion MB, Oechslin F, Moineau S. Phage diversity, genomics and phylogeny. Nat Rev Microbiol. 2020;18(3):125–38.
50. Shkoporov AN, Hill C. Bacteriophages of the human gut: the "known unknown" of the microbiome. Cell Host Microbe. 2019;25(2):195–209.
51. Townsend EM, Kelly L, Muscatt G, Box JD, Hargraves N, Lilley D, et al. The human gut phageome: origins and roles in the human gut microbiome. Front Cell Infect Microbiol. 2021;11:643214.
52. Adkins S, Kamenova I, Rosskopf EN, Lewandowski DJ. Identification and characterization of a novel tobamovirus from tropical soda apple in Florida. Plant Dis. 2007;91(3):287–93.
53. Fillmer K, Adkins S, Pongam P, D'Elia T. The complete nucleotide sequence and genomic characterization of tropical soda apple mosaic virus. Arch Virol. 2016;161(8):2317–20.
54. Kumari N, Sharma V, Patel P, Sharma PN. Pepper mild mottle virus: a formidable foe of capsicum production—a review. Front Virol. 2023;3. Available from: https://doi.org/10.3389/fviro.2023.1208853. Cited 2024 Jun 3.
55. Colson P, Richet H, Desnues C, Balique F, Moal V, Grob JJ, et al. Pepper mild mottle virus, a plant virus associated with specific immune responses, Fever, abdominal pains, and pruritus in humans. PLoS One. 2010;5(4):e10041.
56. Zhang T, Breitbart M, Lee WH, Run JQ, Wei CL, Soh SWL, et al. RNA viral community in human feces: prevalence of plant pathogenic viruses. PLoS Biol. 2006;4(1):e3.
57. Haramoto E, Kitajima M, Kishida N, Konno Y, Katayama H, Asami M, et al. Occurrence of pepper mild mottle virus in drinking water sources in Japan. Appl Environ Microbiol. 2013;79(23):7413–8.
58 Arts PJ, Kelly JD, Midgley CM, Anglin K, Lu S, Abedi GR, et al. Longitudinal and quantitative fecal shedding dynamics of SARS-CoV-2, pepper mild mottle virus, and crAssphage. mSphere. 2023;8(4):e0013223.
59. Symonds EM, Nguyen KH, Harwood VJ, Breitbart M. Pepper mild mottle virus: a plant pathogen with a greater purpose in (waste)water treatment development and public health management. Water Res. 2018;144:1–12.
60. Chettleburgh C, Ma SX, Swinwood-Sky M, McDougall H, Kireina D, Taggar G, et al. Evaluation of four human-associated fecal biomarkers in wastewater in Southern Ontario. Sci Total Environ. 2023;904:166542.
61. Aguado-García Y, Taboada B, Morán P, Rivera-Gutiérrez X, Serrano-Vázquez A, Iša P, et al. Tobamoviruses can be frequently present in the oropharynx and gut of infants during their first year of life. Sci Rep. 2020;10(1):13595.

62. Holtz LR, Cao S, Zhao G, Bauer IK, Denno DM, Klein EJ, et al. Geographic variation in the eukaryotic virome of human diarrhea. Virology. 2014;468–470:556–64.
63. Nishijima S, Nagata N, Kiguchi Y, Kojima Y, Miyoshi-Akiyama T, Kimura M, et al. Extensive gut virome variation and its associations with host and environmental factors in a population-level cohort. Nat Commun. 2022;13(1):5252.
64. Zuo T, Sun Y, Wan Y, Yeoh YK, Zhang F, Cheung CP, et al. Human-gut-DNA virome variations across geography, ethnicity, and urbanization. Cell Host Microbe. 2020;28(5):741-751.e4.
65. Di Bonito P, Della Libera S, Petricca S, Iaconelli M, Sanguinetti M, Graffeo R, et al. A large spectrum of alpha and beta papillomaviruses are detected in human stool samples. J Gen Virol. 2015;96(Pt 3):607–13.
66. Wei F, Gaisa MM, D'Souza G, Xia N, Giuliano AR, Hawes SE, et al. Epidemiology of anal human papillomavirus infection and high-grade squamous intraepithelial lesions in 29 900 men according to HIV status, sexuality, and age: a collaborative pooled analysis of 64 studies. Lancet HIV. 2021;8(9):e531–43.
67. Donà MG, Gheit T, Latini A, Benevolo M, Torres M, Smelov V, et al. Alpha, beta and gamma human papillomaviruses in the anal canal of HIV-infected and uninfected men who have sex with men. J Infect. 2015;71(1):74–84.
68. Mlakar B, Kocjan BJ, Hošnjak L, Fujs Komloš K, Milošević M, Poljak M. Beta-papillomaviruses in the anal canal of HIV positive and HIV negative men who have sex with men. J Clin Virol. 2014;61(2):237–41.
69. Beachler DC, Weber KM, Margolick JB, Strickler HD, Cranston RD, Burk RD, et al. Risk factors for oral HPV infection among a high prevalence population of HIV-positive and at-risk HIV-negative adults. Cancer Epidemiol Biomarkers Prev. 2012;21(1):122–33.
70. Muller K, Kazimiroff J, Fatahzadeh M, Smith RV, Wiltz M, Polanco J, et al. Oral human papillomavirus infection and oral lesions in HIV-positive and HIV-negative dental patients. J Infect Dis. 2015;212(5):760–8.
71. Palefsky J. Human papillomavirus-related disease in people with HIV. Curr Opin HIV AIDS. 2009;4(1):52–6.
72. Yunihastuti E, Teeratakulpisarn N, Jeo WS, Nilasari H, Rachmadi L, Somia IKA, et al. Incidence, clearance, persistence and factors related with high-risk anal HPV persistence in South-East Asian MSM and transgender women. AIDS. 2020;34(13):1933–41.
73. Moustafa A, Xie C, Kirkness E, Biggs W, Wong E, Turpaz Y, et al. The blood DNA virome in 8,000 humans. PLoS Pathog. 2017;13(3):e1006292.
74. Dang X, Hanson BA, Orban ZS, Jimenez M, Suchy S, Koralnik IJ. Characterization of the brain virome in human immunodeficiency virus infection and substance use disorder. PLoS One. 2024;19(4):e0299891.
75. Young JC, Chehoud C, Bittinger K, Bailey A, Diamond JM, Cantu E, et al. Viral metagenomics reveal blooms of anelloviruses in the respiratory tract of lung transplant recipients. Am J Transplant. 2015;15(1):200–9.
76. Sabbaghian M, Gheitasi H, Shekarchi AA, Tavakoli A, Poortahmasebi V. The mysterious anelloviruses: investigating its role in human diseases. BMC Microbiol. 2024;24(1):40.
77. De Vlaminck I, Khush KK, Strehl C, Kohli B, Luikart H, Neff NF, et al. Temporal response of the human virome to immunosuppression and antiviral therapy. Cell. 2013;155(5):1178–87.
78. Madsen CD, Eugen-Olsen J, Kirk O, Parner J, Kaae Christensen J, Brasholt MS, et al. TTV viral load as a marker for immune reconstitution after initiation of HAART in HIV-infected patients. HIV Clin Trials. 2002;3(4):287–95.
79. Liu K, Li Y, Xu R, Zhang Y, Zheng C, Wan Z, et al. HIV-1 infection alters the viral composition of plasma in men who have sex with men. mSphere. 2021;6(3):e00081-21.
80. Esser PL, Quintanares GHR, Langhans B, Heger E, Böhm M, Jensen BOLE, Esser S, Lübke N, Fätkenheuer G, Lengauer T, Klein F, Oette M, Rockstroh JK, Boesecke C, Di Cristanziano V, Kaiser R, Pirkl M. Torque Teno Virus Load Is Associated With Centers for Disease Control and Prevention Stage and CD4+ Cell Count in People Living With Human Immunodeficiency Virus but Seems Unrelated to AIDS-Defining Events and Human Pegivirus Load. J Infect Dis. 2024;230(2):e437–e446. https://doi.org/10.1093/infdis/jiae014. Erratum in: J Infect Dis. 2024 Aug 27:jiae420. https://doi.org/10.1093/infdis/jiae420.
81. Gore EJ, Gard L, Niesters HGM, Van Leer Buter CC. Understanding torquetenovirus (TTV) as an immune marker. Front Med (Lausanne). 2023;10:1168400.
82. Jaksch P, Kundi M, Görzer I, Muraközy G, Lambers C, Benazzo A, et al. Torque Teno virus as a novel biomarker targeting the efficacy of immunosuppression after lung transplantation. J Infect Dis. 2018;218(12):1922–8.
83. Gottlieb J, Reuss A, Mayer K, Weide K, Schade-Brittinger C, Hoyer S, et al. Viral load-guided immunosuppression after lung transplantation (VIGILung)-study protocol for a randomized controlled trial. Trials. 2021;22(1):48.
84. Haupenthal F, Rahn J, Maggi F, Gelas F, Bourgeois P, Hugo C, et al. A multicentre, patient- and assessor-blinded, non-inferiority, randomised and controlled phase II trial to compare standard and torque teno virus-guided immunosuppression in kidney transplant recipients in the first year after transplantation: TTVguideIT. Trials. 2023;24(1):213.
85. Moore RD, Keruly JC. CD4+ cell count 6 years after commencement of highly active antiretroviral therapy in persons with sustained virologic suppression. Clin Infect Dis. 2007;44(3):441–6.
86. Zhang W, Yan J, Luo H, Wang X, Ruan L. Incomplete immune reconstitution and its predictors in people living with HIV in Wuhan, China. BMC Public Health. 2023;23(1):1808.
87. Kufa T, Shubber Z, MacLeod W, Takuva S, Carmona S, Bor J, et al. CD4 count recovery and associated factors among individuals enrolled in the South African antiretroviral therapy programme: an analysis of national laboratory based data. PLoS One. 2019;14(5):e0217742.
88. He L, Pan X, Dou Z, Huang P, Zhou X, Peng Z, et al. The factors related to CD4+ T-cell recovery and viral suppression in patients who have low CD4+ T cell counts at the initiation of HAART: a retrospective study of the national HIV treatment sub-database of Zhejiang Province, China, 2014. PLoS One. 2016;11(2):e0148915.
89. Abbate I, Rozera G, Cimini E, Carletti F, Tartaglia E, Rubino M, et al. Kinetics of TTV loads in peripheral blood mononuclear cells of early treated acute HIV infections. Viruses. 2023;15(9):1931.
90. Schmidt L, Jensen BEO, Walker A, Keitel-Anselmino V, di Cristanziano V, Böhm M, et al. Torque Teno Virus plasma level as novel biomarker of retained immunocompetence in HIV-infected patients. Infection. 2021;49(3):501–9.
91. Dandekar S. Pathogenesis of HIV in the gastrointestinal tract. Curr HIV/AIDS Rep. 2007;4(1):10–5.
92. Guadalupe M, Reay E, Sankaran S, Prindiville T, Flamm J, McNeil A, et al. Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy. J Virol. 2003;77(21):11708–17.
93. Guadalupe M, Sankaran S, George MD, Reay E, Verhoeven D, Shacklett BL, et al. Viral suppression and immune restoration in the gastrointestinal mucosa of human immunodeficiency virus type 1-infected patients initiating therapy during primary or chronic infection. J Virol. 2006;80(16):8236–47.
94. Mehandru S, Poles MA, Tenner-Racz K, Jean-Pierre P, Manuelli V, Lopez P, et al. Lack of mucosal immune reconstitution during prolonged treatment of acute and early HIV-1 infection. PLoS Med. 2006;3(12): e484.

Publisher's Note