Continuous Exposure to Antiretrovirals Imprints a Pro-Inflammatory Senescence State in Alveolar Macrophages

Short title: Antiretrovirals and alveolar macrophage physiology

Vinicius M. Fava,[1,2] Monica Dallmann-Sauer,[1,2,3] Marianna Orlova,[1,2] Wilian Correa-Macedo,[1,2,4] Ron Olivenstein,[5] Cecilia Theresa Costiniuk, [1,6,7] Jean-Pierre Routy,[1,6] Luis B. Barreiro,[8] and Erwin Schurr,[1,2,3]

[1] Program in Infectious Diseases and Immunity in Global Health, The Research Institute of the McGill University Health Centre; Montreal, QC, H4A 3J1, Canada.

[2] McGill International TB Centre, McGill University; Montreal, QC, H4A 3J1, Canada.

[3] Departments of Human Genetics and Medicine, Faculty of Medicine, McGill University, Montreal, QC, H3A 0C7, Canada.

[4] Present address. Canadian Centre for Computational Genomics (C3G) McGill University, Montreal, QC, H3A 0G1, Canada.

[5] Translational Research in Respiratory Diseases Program, The Research Institute of the McGill University Health Centre; Montreal, QC, H4A 3J1, Canada.

[6] Department of Medicine and Chronic Viral Illnesses Service, McGill University Health Centre; Montreal, QC, H4A 3J1, Canada.

[7] Graduate Program in Clinical and Translational Research, McGill University, Montreal, QC, Canada.

[8] Department of Medicine, Genetic Section, University of Chicago, Chicago, Illinois, IL 60637, USA.

Abstract

Advances in antiretroviral therapy (ART) have substantially improved the lives of people with HIV (PWH) and reduced HIV acquisition through pre-exposure prophylaxis (PrEP). However, the long-term effect of ART on the physiological state of cells remains poorly understood. Despite the success of ART in preventing the progression of HIV infection to AIDS, PWH are suffering from a disproportional burden of non-AIDS comorbidities, including lung diseases. Given the central function of alveolar macrophages (AM) in pulmonary immunity, we evaluated the impact of time on ART on AM of PWH and people on PrEP. Employing a retrospective cross-sectional design, we showed that ART imprinted a pro-inflammatory and senescence-like epigenetic and transcriptomic state on AM. This effect of ART was detected irrespective of HIV infection. Increased epigenetic priming and gene expression of cell cycle arrest markers such as *CDKN1A* (p21), senescence associated secretory phenotype (SASP) genes, e.g. *IL6R* and *CXCL8*, and transcription factors subunits of the AP-1 family were a hallmark of ART exposure. The ART-linked epigenetic and transcriptomic changes were strongly dependent on the duration of ART for both the PWH and PrEP groups and consistent with dysregulated AM function. SASP contributes to the pathophysiology of multiple lung diseases such as COPD, pulmonary fibrosis, and asthma. Combined, our data suggested that long-term ART contributes to age-related pulmonary comorbidities in PWH. Given the strong correlation with time on ART and adverse effects in PWH, it is possible that long-term PrEP may result in similar clinical outcomes. While PrEP is critically important for preventing HIV acquisition in the most vulnerable populations, our study advocates for adjuvant therapies and improved drug design to prevent or mitigate potential long-term adverse effects of PrEP.

Introduction

Advances in antiretroviral therapy (ART) have improved the quality of life for people with HIV (PWH). In contrast to the initial regimens, current ART is more effective in the suppression of HIV and has lower toxicity and adverse effects (*1*). In addition to treatment and health care improvements, mitigation of opportunistic infections, and management of HIV comorbidities have greatly increased life expectancy for PWH (*2*). Consequently, HIV infection is nowadays considered a chronic manageable condition. The increased lifespan of PWH coincided with a surge of age-related idiopathic diseases despite well-controlled HIV levels (*3, 4*), and the aging population of PWH now suffers a disproportional burden of cardiovascular, neurological, and pulmonary diseases (*5-7*). The reasons for the increased incidence of non-AIDS diseases in PWH remain poorly understood. HIV persistence, low levels of inflammation and uncharacterized adverse effects of ART are among the potential contributors leading to idiopathic disease in PHW (*8-12*).

With respect to pulmonary involvement, PWH on ART are more susceptible to chronic obstructive pulmonary disease (COPD) (*13*), mortality due to COVID-19 (*14*), tuberculosis (TB) (*15, 16*), pneumonia (*17, 18*), pulmonary fibrosis (*19*), and lung cancer (*20, 21*). Key cells implicated in the pathogenesis of these pulmonary comorbidities are alveolar macrophages (AM) which are long-living and self-renewing sentinel cells of the lung. We previously showed that AM from PWH have an impaired epigenetic and transcriptomic response to *Mycobacterium tuberculosis* (*Mtb*), the causal agent of TB (*16*). Surprisingly, HIV-negative persons taking ART as Pre-exposure Prophylaxis (PrEP) had a similar diminished response to *Mtb*. This finding indicated an unnoticed side effect of ART irrespective of HIV infection. Hence, we investigated the effect of ART on the epigenetic and transcriptomic state of AM from PWH and people on PrEP. We found a highly significant effect of ART on the epigenetic and transcriptomic landscape of AM that was strongly dependent on the duration of ART.

Results

Population samples and study design

Alveolar macrophages isolated from bronchioalveolar lavage (BAL) of 14 HIV-negative healthy controls without ART (HC), 10 HIV-negative persons under PrEP, and 18 PWH on ART were investigated (**Table 1**). Most of the study participants were males and smokers. PWH were on average older than HC and individuals on PrEP (**Table 1**). Participants had uninterrupted use of ART ranging from two months to 5.4 years for PrEP and from seven months to 23.1 years for PWH at the time of BAL cell collection (**Fig. 1A; Table 1**). Most of the PWH had ART introduced in the first two years after HIV diagnosis. However, three participants had ART implemented 10+ years after the initial diagnosis (**Table 1**). To assess the epigenetic and transcriptomic state of AM we used five different assays (**Fig. 1B**). We

evaluated chromatin accessibility, transcription factor enrichment and occupancy via ATACseq, chromatin state via ChIPseq of three histone marks H3K4me3 (active transcription), H3K27ac (active transcription and enhancers), and H3K27me3 (repressed regions), and gene expression via RNAseq

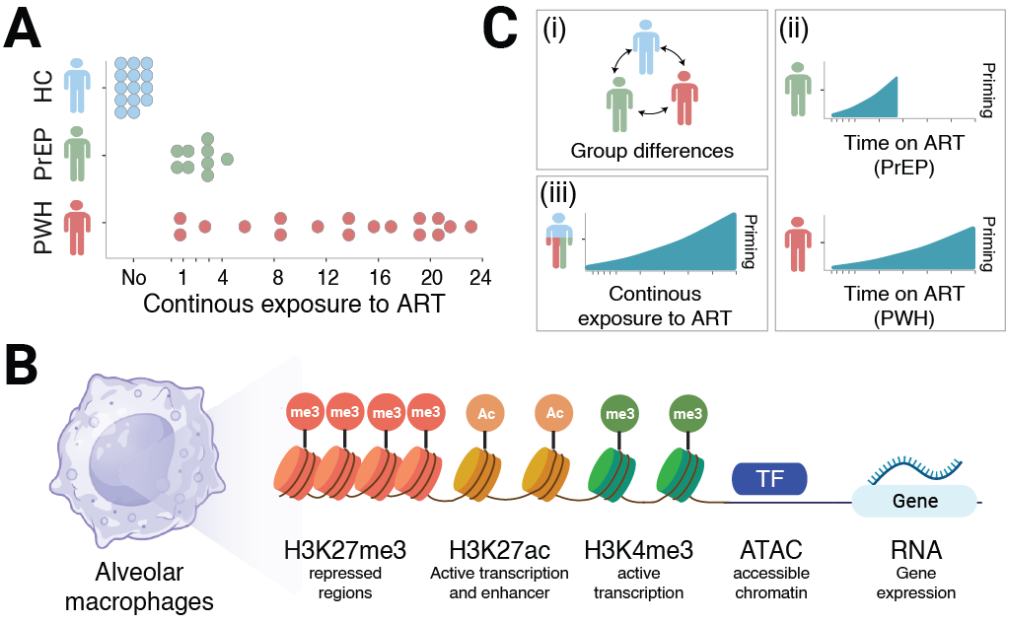

Fig. 1. Study design overview.

(**A**) Continuous exposure to antiretroviral therapy (ART) in years for healthy controls (HC), people on ART as pre-exposure prophylaxis (PrEP), and people with HIV on ART (PWH). Each circle represents a subject enrolled in the study. (**B**) Epigenetic and transcriptomic assays used to evaluate the physiological state of alveolar macrophages. (**C**) Hypotheses tested to assess the biological mechanisms primed by ART in alveolar macrophages: (i) differences between groups independent of duration of HIV infection or time on ART, (ii) Considering time on ART for PrEP and PWH participants, (iii) Exploration of continuous exposure to ART while accounting for group status. (Created with BioRender).

We tested the hypothesis that ART imprinted a physiological state in AM that primed pathways linked to pulmonary diseases. We carried out the testing in three separate stages (**Fig. 1C**). (i) We compared the epigenetic and transcriptomic state of AM between the HC, PrEP and PWH groups independent of time on ART. (ii) We tested the effect of time on ART in PrEP and PWH individuals separately. Accessibility of chromatin regions, histone marks, and gene expression changes with increased time for both groups were denoted as ART effect. (iii) We conditioned the effect of continuous exposure to ART on the PrEP/PWH/HC group status to highlight the main biological mechanisms affected by time on ART.

Differences in AM physiology of PWH, people on PrEP and HC

To assess the correlation of epigenetic and transcriptomic features (i.e., accessible chromatin, H3K27me3, H3K27ac, H3K4me3, and RNA) with study group we performed a principal component analysis (PCA). After accounting for known confounders (smoking, sex, chronological age, and batch) and preserving the phenotype group variable, the top four principal components (PC) captured between 38% to 75% of the remaining variability. Yet, the top components only partially clustered subjects according to their respective groups suggesting strong interindividual variability (**Fig. 2A**). As a result, when performing the differential feature analysis between groups, we observed less than 5% of the tested features passing FDR < 10% on most of the assays (**Fig. S1; Table S1**). Notable exceptions were the number of differentially expressed genes between PrEP and HC AM (**Fig. S1**). Among genes with decreased expression in PrEP, we identified all 13 mitochondrial-encoded genes of the electron transport chain (**Fig. S2**). This indicated a potential mitochondrial dysfunction in AM from PrEP that was also observed to a lesser extent for PWH.

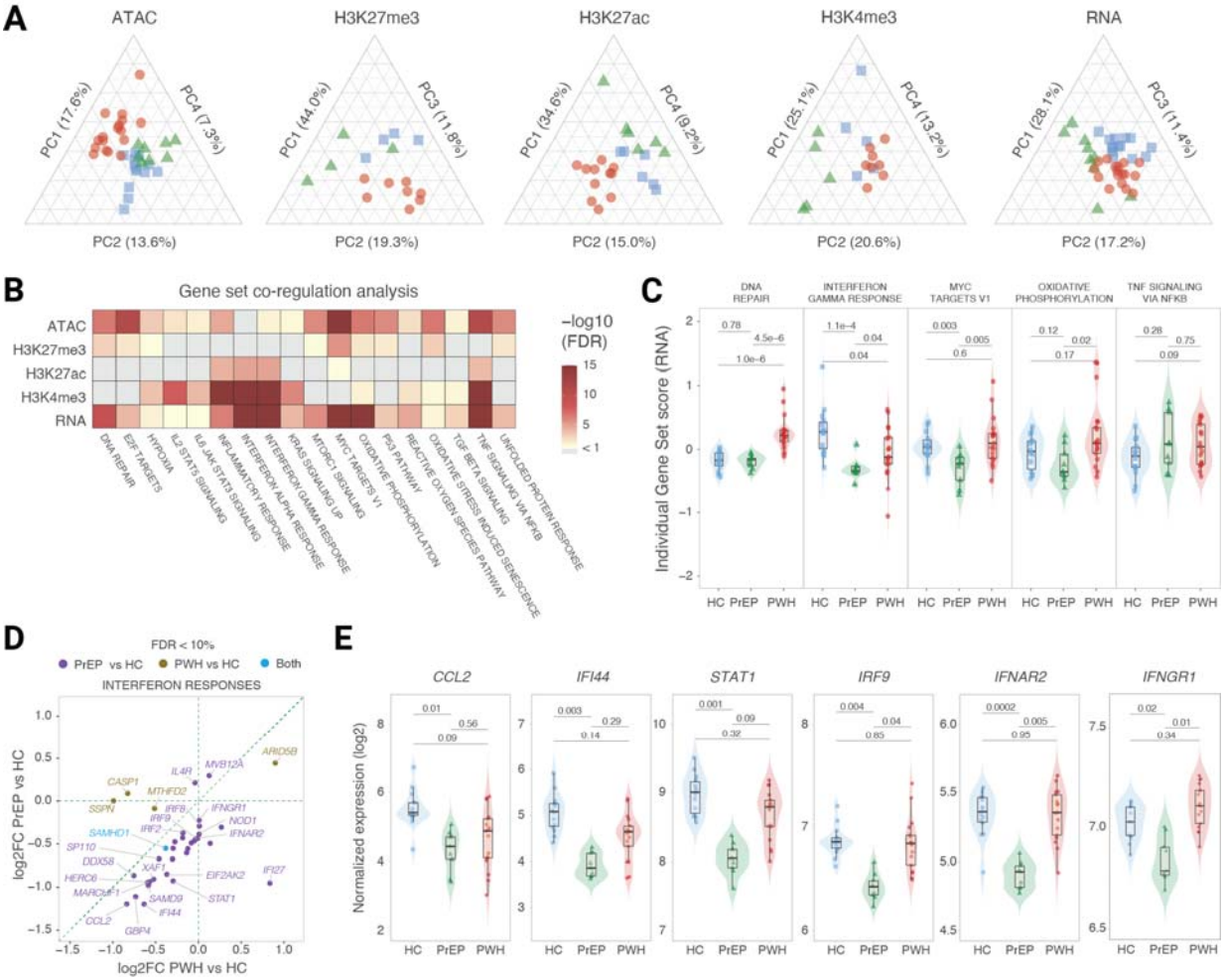

Fig. 2. Epigenetic and transcriptomic differences between AM from people on PrEP, PWH and HC.

(A) Phenotypic group clustering using the top principal components (PCs). Scaled PCs and their corresponding percentage of variance explained are shown on each axis of the ternary plots for the five tested assays. Each shape represents a sample from healthy controls (blue), people on PrEP (green), and PWH (red). **(B)** Gene Set Co-regulation Analysis (GESECA) for Hallmark pathways. Pathways with False Discovery Rate (FDR) < 10% in at least three assays are indicated by a colour scheme. Pathways with FDR > 10% are shown in grey. Results for all tested pathways are given in Table S2. (**C**) GESECA scores for the top five differential pathways. The scores of the transcriptomic assay for each subject are indicated on the y-axis separated by group on the x axis. (**D**) Correlation between differential expression for interferon response genes in PrEP and PWH versus HC. The log2 fold difference (log2FC) between PrEP and HC was plotted against the log2FC for PWH vs HC for the transcriptomic assay. Each circle represents a gene in the Hallmark interferon pathways that had FDR < 10% in PrEP (purple), PWH (yellow) or both (blue) versus HC. **(E)** Normalized gene expression per subject and group for selected genes in HALLMARK interferon pathways. In panels C and E, a Wilcoxon test was used to estimate significant differences in the median between groups. The Benjamini-Hochberg adjusted *p*-values are shown at the top for each group comparison.

Given the pronounced interindividual variability across all studied groups, we performed a Gene Set Co-regulation Analysis (GESECA) to identify pathways associated with the most highly variable features. The GESECA method tests if the highest variable features between samples are enriched in biological pathways and is agnostic to subject grouping. Using this approach, we found that differences between subjects were significantly enriched in inflammatory, cellular homeostasis and stress related pathways both at the transcriptomic and epigenetic levels (**Fig. 2B; Table S2**). Next, we evaluated if the differential pathways between subjects were driven by the phenotypic groups (**Fig. 2C**). We found that genes assigned to the DNA repair pathway had significantly increased expression in AM from PWH compared to PrEP and HC. DNA repair is a known mechanism activated in host response to HIV integration. In contrast, expression of genes assigned to interferon pathways were significantly depleted in AM from PrEP and PWH compared to HC (**Fig. 2C**). Of the 183 genes of hallmark interferon pathways that are expressed in AM, 28 (15.3%) had decreased expression in PrEP compared to HC with most of these genes displaying a trend of decreased expression in PWH (**Fig. 2D**). Among key genes with lower expression in PrEP were chemokines such as *CCL2*, Interferon-Inducible genes (IFIs), Interferon Regulatory Factors (IRFs), and interferon receptors (**Fig. 2E; Table S1**). Interestingly, both type-I (*IFNAR2*) and type-II (*IFNGR1*) interferon receptor genes had lower expression in the PrEP group compared to both the HC and PWH groups (**Fig. 2E; Table S1**). In addition, genes in the MYC targets pathway displayed lower expression in PrEP AM compared to HC and PWH groups (**Fig. 2C**). *MYC*

encodes a pleiotropic transcription factor (TF) that regulates numerous cellular processes, and the effect of MYC responsive genes in PrEP will require further investigation.

Overall, the group comparison indicated an active DNA damage repair (DDR) mechanism in AM from PWH and a dampened expression of interferon-related genes in AM that was most pronounced in PrEP participants. On the other hand, the inter-individual variability associated with oxidative phosphorylation (OXPHOS) and TNF signaling pathways were not captured by group assignment (**Fig. 2C**). The lack of association between gene set scores for these pathways and group labels suggested that unaccounted participant characteristics were likely driving the inter-individual variability. However, when displaying the time on ART for the individuals in Fig. 2C we observed a strong trend for a time dependent effect on TNF signaling for AM from both PrEP and PWH (**Fig. S3**). We therefore hypothesized that time on ART for PrEP and for PWH are important factors to explain differences in AM state between subjects.

The effect of time on ART on the transcriptomic and epigenetic state of AM

While accounting for the same covariates as in the previous analyses, we added a time variable to the model and tested PrEP and PWH groups independently. The time variable captured the fraction of years from the start of the ART regimen to BAL collection (**Fig. 1A; Table 1**). To avoid regressing out the variable of interest, we conducted a PCA on the residuals after removing confounders and preserving the time variable. With this maximized approach we showed that PC1 strongly correlated with time on ART in both PrEP and PWH in all tested assays (**Fig. 3A**). This was reflected in the number of significant differential epigenetic and transcriptomic features at FDR < 10% for time on ART for PrEP and length of time since ART initiation for PWH (**Fig. 3B; Table S3**). Chromatin accessibility and the H3K27ac mark showed the highest numbers of significantly differential features affected by time in both groups. Additionally, PWH presented 5-fold more time-dependent differentially expressed genes compared to the PrEP participants and more active promoters with increased H3K4me3 (**Fig. 3B; Table S3**). This may reflect HIV-specific effects, time-dependent changes not detected in the PrEP group due to the shorter ART regimen, or higher statistical power due to the larger sample size of the PWH group. To compare features affected by time on ART in both groups, we intersected genes and features significantly associated with time in the PrEP and PWH groups (**Fig. 3C**). With exception of the promoter repressor H3K4me3 mark, we identified a significant enrichment of changes occurring in the same direction for time on ART on PrEP and PWH (**Fig. 3C**).

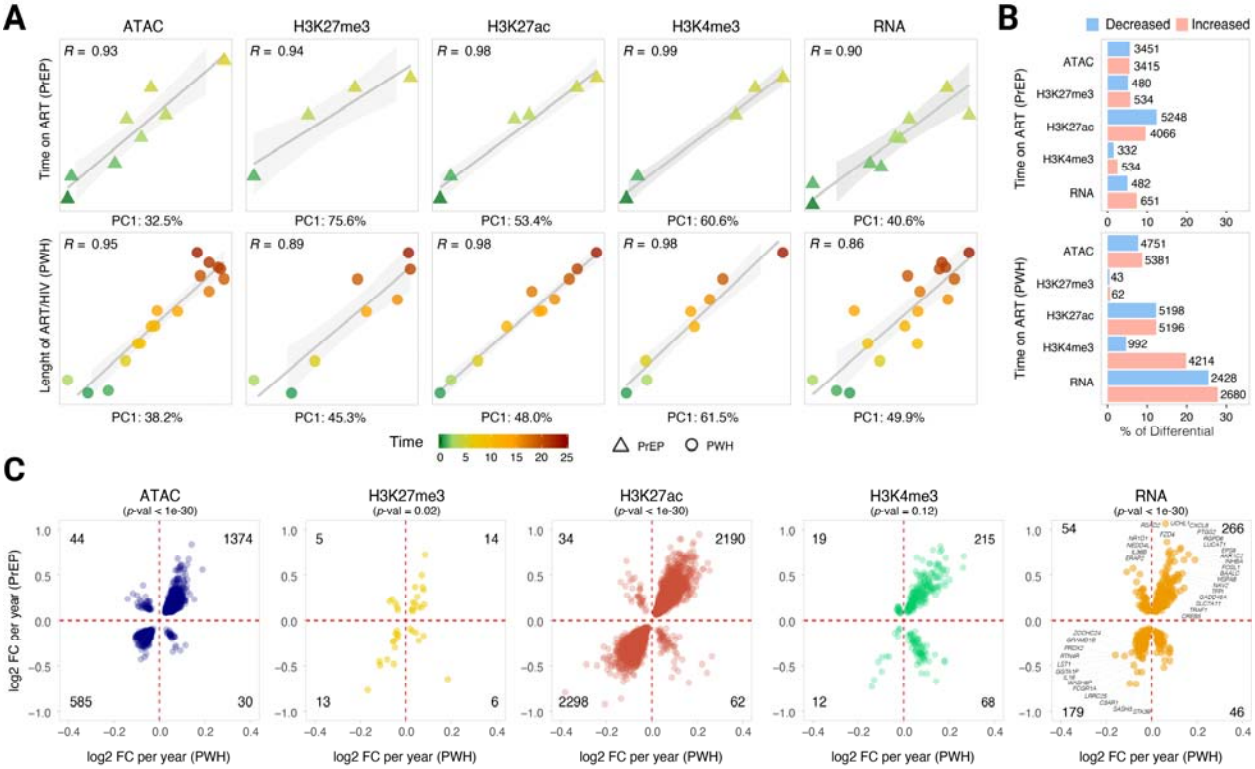

Fig.. 3. The epigenetic and transcriptomic effect of time on ART for people on PrEP and PWH.

(**A**) The proportion of variance captured by the first principal components was plotted against time on ART for people on PrEP (top) and for PWH (bottom). The Pearson correlation R is shown at the left upper corner of each plot. (**B**) Epigenetic and transcriptomic features significantly associated with time on ART in PrEP and PWH participants. The bar plots show the proportion of differential regions/genes at an FDR of 10% relative to the total number of tested regions/genes for time on ART in PrEP (top) and in PWH (bottom). The absolute numbers of differential regions/genes are shown next to the bars. Decreased and increased DNA accessibility, histone marks, and gene expressions are shown in blue and red, respectively. (**C**) Epigenetic and transcriptomic features significantly changed in the same direction with time for PrEP and PWH participants. The log2 FC changes per year for PrEP (y-axis) and PWH groups (x-axis) are plotted against each other for the five tested assays. Only features with FDR < 10% in both groups are plotted. The number of regions and genes are given for each quadrant. *p*-values for the correlation of log2 FC per year direction in PWH and PrEP are given.

Biological mechanisms affected by time on ART in PrEP and PWH

To assess the biological mechanisms affected by time on ART we performed a GESECA analysis separately for the PrEP and PWH groups (**Table S2**). We highlighted 14 biological pathways with the highest variability between samples within the PrEP and PWH groups (**Fig. 4A**). These pathways included OXPHOS, immune mediators, control of cellular stress, cell cycle arrest, and senescence (**Fig.**

4A). Next, we tested if the per individual GESECA score of highlighted pathways correlated with time on ART in PrEP and PWH (**Fig. 4B; Fig. S4**). At the RNA level, 13 out of 14 pathways significantly correlated with time in at least one group ($P \leq 0.01$) (**Fig. 4B; Fig. S4**). Interferon Gamma was the exception where differences between samples did not correlate with time (**Fig. S4**). At the epigenetic level, time on ART correlated ($P \leq 0.01$) with individual GESECA scores in all highlighted pathways for at least one active epigenetic assay (ATAC, H3K27ac, and H3K4me3). Decreased expression of genes assigned to the OXPHOS pathway correlated with time on ART in PWH, but not in the PrEP group (**Fig. 4B**). However, hallmark OXPHOS pathway do not encompass mitochondrial genes of the electron transport chain that had decreased gene expression in PrEP and PWH independent of time on ART (Fig S2). This suggested that time on ART may impact different components of OXPHOS at different levels. This is supported by the decreased expression of autosomal genes encoding members of the electron transport chain complexes I–V that was more pronounced with increased time on ART in PWH (**Fig. 4C**).

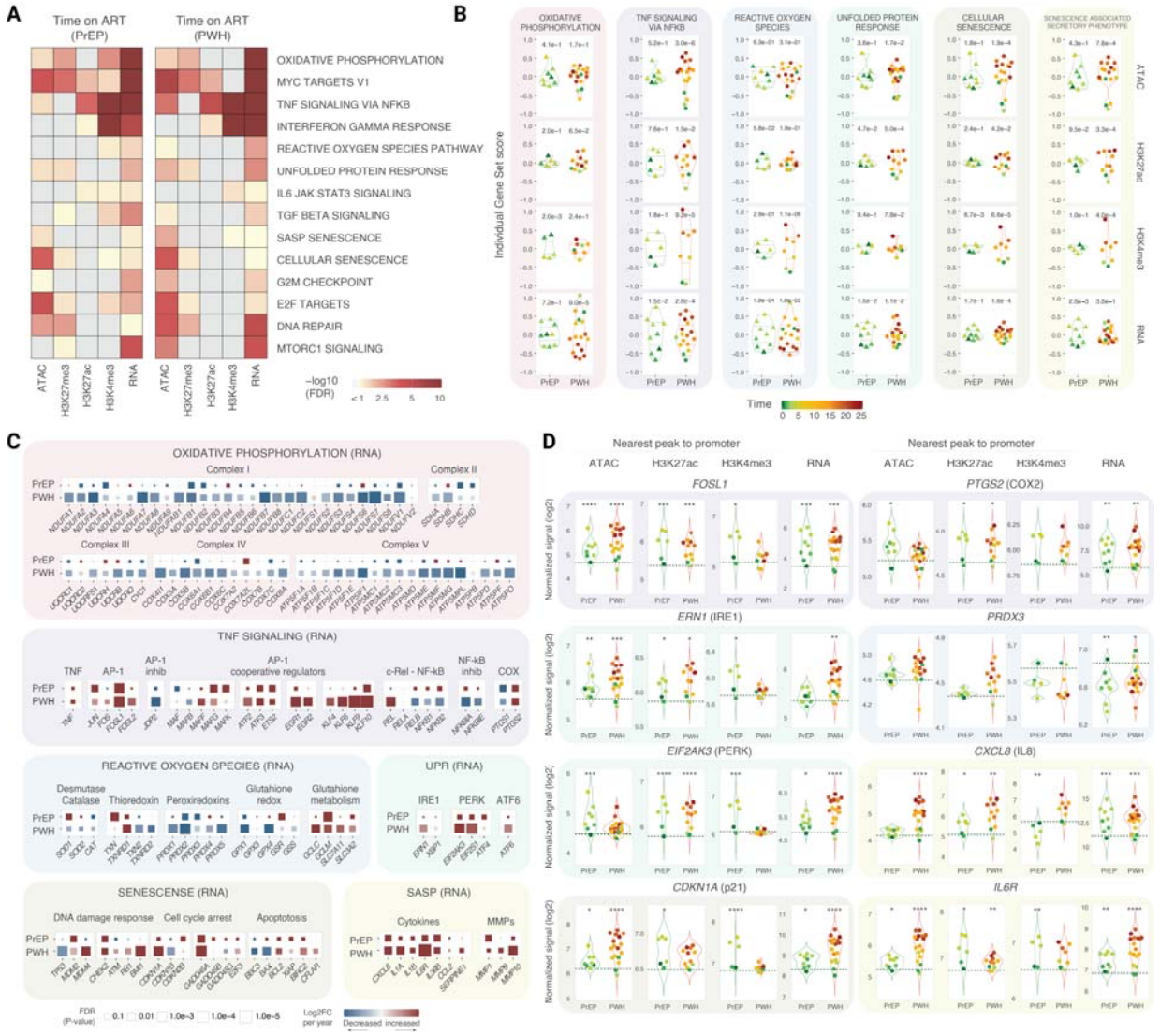

Fig. 4. Biological pathways affected by time on ART in people on PrEP and in PWH.

(**A**) GESECA identified Hallmark pathways with the largest variability between samples within each group. Pathways with FDR < 10% in at least two assays in each group are shown as a heatmap with the level of significance indicated by a colour scheme. Pathways with FDR > 10% are shown in grey. Data for all tested pathways is given in Table S2. (**B**) Per subject GESECA scores (y-axis), separated by group (x-axis) for six pathways affected by the time variable. Remaining pathways are shown in Fig S4. Colour scheme indicates length of ART exposure. The quantiles are shown as violins for reference of the distribution. A linear model was used to test the correlation of time on ART in PrEP and PWH with the individual GESECA scores for all tested assays. *p*-values for the correlations are shown at the top of each comparison. The repressor mark H3K27me3 did not show a significant correlation with time in any group and was omitted. (**C**) Changes in gene expression for selected genes belonging to pathways shown in panel B. The log2 FC per year and FDR were plotted for time on ART in PrEP (top) and in PWH (bottom). The size of the squares represents the magnitude of the FDR. Scaled colours indicate a decrease (blue) or an increase (red) in gene expression. Genes were grouped by functional contribution to each pathway. As the time on ART effect was more pronounced in AM from PrEP the magnitude of log2FC per year was capped to 0.1 to facilitate visual comparison between PrEP and PWH groups. The absolute values for log2 FC and FDR are shown in Table S3. (**D**) Normalized signal for epigenetic features at the promoter and expression of selected genes affected by time on ART in both groups. Normalized log2 signals are shown on the y-axes. On the x-axes for each assay, subjects are separated by group. The time on ART is indicated by colour shades and circles represent individual participants. A dotted horizontal green line indicates the average signal in healthy controls for reference. *p*-values are given at the top of the group for each comparison with the absolute values and FDR shown in Table S4. Significance per group *, *p*-value < 0.1, ** *p*-value < 0.01, *** *p*-value < 0.001 and, **** *p*-value < 0.0001.

Transcription of genes assigned to the TNF signaling pathway correlated with time on ART in both PrEP and PWH groups (**Fig. 4B**). The increase in the transcription-based GESECA score per year for the TNF pathway was higher in PrEP compared to PWH (**Fig. 4B**). Among the genes assigned to TNF signaling that displayed increased expression with time on ART were transcription factors that are key mediators of immune responses and senescence (**Fig. 4C**). For example, subunits of the AP-1 complex (e.g. *JUN* and *FOSL1*), and AP-1 cooperative regulators (e.g. *EGR1*, *ATF3*, *ETS2*, and *KLF10*) had increased expression with time on ART in AM from subjects on PrEP and PWH (**Fig. 4C**). Moreover, key pro-inflammatory mediators such as *PTGS2* (i.e. COX-2) and *TNF* showed increased expression with time on ART in both groups, albeit *TNF per se* did not surpass FDR correction (**Fig. 4C**). For these key genes we also identified a time-related effect at the epigenetic level. For example, the transcription

starting site (TSS) of *FOSL1* and *PTGS2* had increased chromatin accessibility, as well as H3K27ac and H3K4me3 marks with increased time on ART (**Fig. 4D**).

Time on ART also impacted the AM cellular response to stress via inactivation of reactive oxygen species (ROS) and protein turnover via the unfolded protein response (UPR) (**Fig. 4B**). For the ROS pathway we noted an opposite effect of time for AM from the PrEP and PWH groups at the RNA level (Fig 4B). However, this was due to groups capturing different aspects of the ROS pathway. For example, redoxins (e.g. PRDXs and GPXs) had more decreased expression with time on ART in PWH with a less pronounced effect in PrEP (**Fig. 4C**). Conversely, glutathione metabolism substrates used by GPXs to neutralize free radicals, and glutathione reductase (*GSR*) had increased expression with time on ART in PrEP (**Fig. 4C**). Overall, AM presented an unbalanced expression between effector enzymes and substrate metabolism when coping with cellular stress. The UPR pathway plays a crucial role in maintaining cellular homeostasis and encompasses three main branches via IRE1, PERK and ATF6. Notably, the *ERN1* (IRE1), *EIF2AK3* (PERK) and *ATF6* genes had increased expression with time on ART in PWH (**Fig. 4C**). The PERK branch genes also had increased expression with time on ART in PrEP (**Fig. 4C**). Interestingly, at the epigenetic level, the promoters of *ERN1* and *EIF2AK3* were primed with increased time on ART in both PrEP and PWH groups (**Fig. 4D**). Taken together, our data suggested a dysregulated response to stress in AM with increased time on ART.

Accelerated cellular aging has been suggested as a candidate mechanism contributing to the increased incidence of comorbidities in PWH. We found that the Cellular Senescence and Senescence Associated Secretory Phenotype (SASP) pathways correlated with time on ART (**Fig. 4B**). Senescence is a complex mechanism that involves multiple cellular processes including DDR, cell cycle arrest, cellular stress and immune mediating cytokines. Increased AP-1 activity, and levels of p21 (encoded by *CDKN1A*), as well as Growth Arrest and DNA Damage-inducible genes (i.e. *GADD45A*) are considered a hallmark of cellular senescence via cell cycle arrest (**Fig. 4C**). The *CDKN1A* promoter was primed with increased chromatin accessibility, H3K27ac and H3K4me3 marks, and *CDKN1A* had increased expression with time on ART for both PrEP and PWH groups (**Fig. 4C and D**). Intriguingly, expression of *TP53* (p53) was decreased while *MDM2* (p53 inhibitor) expression was increased with time on ART in PWH suggesting an impairment of DDR in AM from PWH (**Fig. 4C**). Moreover, key inducer of apoptosis such as *BBC3* (i.e., PUMA) and *BAX* had decreased expression with time on ART while apoptosis inhibitors via caspase neutralization (e.g., *BCL2*, *XIAP*, *BIRC2* and *CFLAR*) had increased expression. In SASP high levels of IL-1-family members, IL-8, IL-6, PAI-1 (*SERPINE1*), and MMPs are a sign of accelerated cellular aging. *IL1A*, *IL1B*, *CXCL8* (IL-8), *IL6R* (IL-6 receptor) and MMPs had increased expression with time on ART in AM from both the PrEP and PWH group (**Fig. 4C**). The effect of time on ART on *CXCL8*, *IL6R* and *IL36B* (IL-1 family member) were among the most significant

changes at the transcriptomic and epigenetic levels (**Fig. 4D; Table S4**). Taken together, our analysis of the effect of time on ART found that AM of both groups displayed the characteristic of cells undergoing cell cycle arrest senescence with dysregulated DDR.

When studying time on therapy as a variable there is potential for collinearity with chronological age which may not be fully accounted for in statistical modeling. To address this, we evaluated the correlation between time on ART and the chronological age of participants. No significant correlation was detected between the participants' chronological age and time on ART in PrEP ($R = 0.001$, $p = 1$) with a trend detected for PWH ($R = 0.40$, $p = 0.11$) (**Fig. S5A**). Next, we assessed if the chronological age was a confounder in the association of senescence and TNF signaling pathways with time on ART in PrEP and in PWH We compared individual GESECA scores of these two pathways with chronological age in AM from both groups and failed to detect any significant correlation (**Fig. S5B**). Another possible confounder for the inflammatory and cell cycle arrest senescence-like status of AM is smoking. When comparing individual GESECA scores for senescence and TNF signaling pathways we found no association with smoking status (**Fig. S5C**). These observations reinforced the accuracy of our modeling in accounting for confounders related to pathways associated with ART. Yet, this does not exclude the independent contribution of ageing and smoking on the physiological state of AM, which was not evaluated in the context of this study.

Increased chromatin accessibility in AM with continuous exposure to ART is enriched for motifs and footprints of pro-inflammatory and senescence associated TFs

Given that the effect of time on ART was detected in both the PrEP and PWH groups for multiple pathways, we tested the overall impact of continuous exposure to ART. We treated the HC group as a time zero and assessed the effect of time on ART on AM physiology by combining all samples while adjusting on the HC, PrEP and PWH status. We detected significant changes at FDR < 10% with time on ART in 26.7% of the regions with accessible chromatin, 22.4% of H3K27ac, 21.8% of H3K4me3, and 52.7% of expressed genes (**Fig. 5A; Table S3**). We then repeated the GESECA analysis for the timepoints irrespective of the groups to recapitulate the correlation of biological pathways with time on ART (**Fig. S6**). For example, the TNF signaling pathway correlated with the continuous exposure to ART, recapitulating the per group time effect (**Fig. 5B**). To showcase the effect of time on ART at the epigenetic level, we derived the traces representing promoter accessibility and active histone marks for two AP-1 genes: *FOSL1*, *JUN* (**Fig. 5C**, left), and two AP-1 cooperators: *ATF3*, *ETS2* (**Fig. 5C**, right). These genes showed significantly increased expression with time on ART in AM from both PrEP and PWH (**Fig. 4C)** and consequently in the combined analysis (**Fig. 5C; Table S3**).

The focus of the combined analysis was to evaluate the overall epigenetic regulation of ART on the AM physiological state. Based on our finding that gene expression of TFs, such as members of the AP-1 family, were associated with time on ART in AM from both PrEP and PWH, we tested if TF motifs and footprints were enriched in chromatin regions with differential accessibility with continuous exposure to ART. First, we evaluated if chromatin accessibility changes with time on ART were enriched for motifs of TFs that supported a poised pro-inflammatory and senescent state of AM. We estimated the frequency of TFs motifs cataloged by HOMER in two control groups: (i) 500,000 random genomic regions (Random), and (ii) 20,060 ATAC regions with FDR > 50% (FDR50) for the continuous exposure to ART. Next, we compared motif frequencies in the control groups with the 8177 increased and 8274 decreased differential chromatin sites with continuous exposure to ART (ATAC at **Fig. 5A and Table S4**). We found 283 and 15 motifs significantly enriched (FDR < 1%) in regions with increased and decreased accessibility with time on ART compared to the two control groups, respectively (**Table S4**). Next, we performed TF binding clustering by motif similarities and identified 8 groups encompassing the top 50 motifs with higher frequency in regions with increased accessibility with continuous exposure to ART (**Fig. 5D**). Group 2 included motifs for the Early growth response (EGR) family; Group 3 motifs for members of the KLF family; Group 5 motifs for MYC and MAFK; Group 6 motifs for AP-1 subunits such as Fra1 (i.e *FOLS1*), FOSL2 and JUN, and AP-1 cooperative TFs such as ATF3 and BACH; and Group 8 included members of the CRE and ATF family which crosstalk with multiple members of the AP-1 complex (**Fig. 5E**). The majority of TFs included in the five largest groups displayed coordinated effects where TFs with motifs enriched in chromatin regions with increased accessibility also showed differential gene expression with time on ART (**Table S3; Fig. S7**). The enrichment of motifs in regions with decreasing accessibility with time on ART was low. We found that 13 out of the 15 motifs with higher frequency in regions with decreased chromatin accessibility versus FDR50 were part of a single group including IRF8 and ELFs genes (**Fig. 5E,** right).

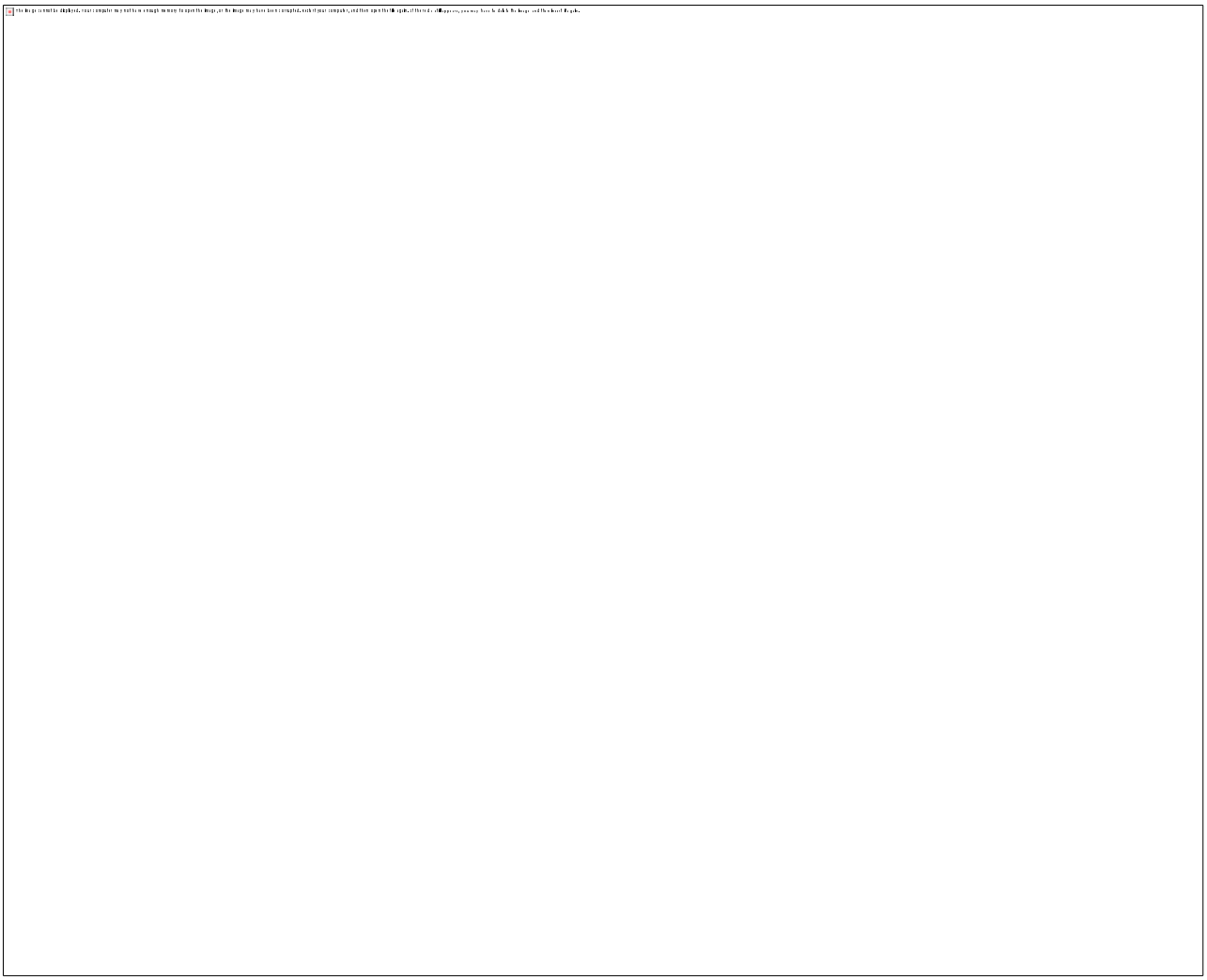

Fig. 5. The effect of continuous exposure to ART on the epigenetic regulation of AM physiology.

(**A**) Regions and genes affected by continuous exposure to ART. The bar plots show the proportion of affected regions/genes at an FDR < 10% relative to the total number of tested regions/genes. The absolute number of affected regions/genes are shown next to the bars. Decreased and increased accessibility, histone marks, and gene expression are shown in blue and red, respectively. (**B**) Correlation of individual gene set scores (y-axis) for the TNF signaling pathway with length of ART exposure (x-axis). Pearson correlation R and corresponding *p*-values are shown at the top left of the plots. (**C**) Trace plots for chromatin accessibility, as well as H3K27ac and H3K4me3 marks at the promoter of four transcription factors (TF) of the TNF signaling pathway. Each line on the plot represents the normalized and covariate-corrected coverage for one of the timepoints plotted in B at a 40 base-pair (bp) resolution. (**D**) Hierarchical clustering the top 50 TF motifs found in regions with increased accessibility. (**E**) Motif frequency of the top TF motifs in the Decreased and Increased chromatin accessibility regions with continuous exposure to ART and the two control groups. The first row orders the motif-binding

homology groups from D. The following four rows show the negative log10 FDR for the contrast indicated on the left side with FDR colour scale given on op. The last four rows represent the z-scaled motif frequency for random background including 500 thousand genomic regions with the same length distribution and GC-content as in the tested regions and (ii) nonsignificant regions at FDR > 50% (FDR50) followed by motif frequencies in the regions with decreased and increased chromatin accessibility. Shades of red denote higher frequency and shades of blue denote lower frequency for each motif. (**F**) TF footprint depth in regions with increased accessibility with increased length of time on ART. The volcano plot shows the negative log10 FDR *p*-value on the y-axis and the fold change per year for the footprint score on the x-axis for 919 tested TFs. Each TF is shown as a dot. TF with increased and decreased depth in the footprint are given in pink and green, respectively. (**G**) Examples of aggregated footprint signals for TFs associated with continuous exposure to ART. Each line represents a timepoint of ART exposure as plotted in panel B using the same colour coding. The traces depict the average signal surrounding motifs for four selected TF enriched in the 8177 increased accessibility regions. The signal was centred on the motifs and coverage was estimated within 60bp on each side. The signal was normalized for the minimum signal per sample to start at 0. The number of footprints representing the aggregated signal is shown at the top. (**H**) Trace plots showing footprints of AP-1 subunits and cooperators in the promoter of genes associated with senescence. The bottom trace plots show chromatin accessibility as described in panel C. The top traces represent a zoom of the regions marked in grey at a 1bp resolution to showcase footprints in detail. At the bottom of the 1bp panels, coloured bars show where motifs for TF were located. A dampening in the ATAC trace signal (i.e. valleys) indicates footprints on top of the motifs.

Next, we tested the differential TF occupancy (i.e. footprint) against continuous exposure to ART. Of note, participants with the exact same time on ART were treated as a single timepoint. TF footprints were detected via depletion of ATACseq signal overlapping TF motifs within higher flanking ATACseq signals (i.e. valleys). A footprint score per timepoint was calculated for 919 JASPAR core motifs over the 8177 regions with increased chromatin accessibility with continuous exposure to ART. We then tested if the footprint mean score per TF (which is the average of footprint scores in all motifs for a given TF) correlated with the continuous exposure to ART (**Table S5**). We identified a significant increase in mean footprint score per year for 421 TFs and a significant decrease for 55 TFs (FDR < 1%) (**Fig. 5F**). Among the TF footprints with the most pronounced fold change per year (FC > 0.2) were members of the AP-1 family (**Fig. 5F**). Moreover, members of all motif groups that displayed higher frequency in increased chromatin accessibility with time on ART had significantly stronger footprints with time on ART (**Fig. 5E and F**). Compared to changes in TF footprint, the effect of time on ART was

more apparent for chromatin accessibility (**Fig. 5G**). Yet, most striking was the coordinated combined effect of accessibility and footprints and gene expression. For example, footprints for AP-1 subunits and other TFs associated with time on ART were identified in the core promoter of important genes mediating immune and senescence phenotypes. We detected footprints for the AP-1 JUN/FOS complex and other AP-1 cooperators in the TSS of *CDKN1A* (p21), *IL6R* and *IL36B* that are key mediators of senescence (**Fig. 5H**). These footprints had a higher depth for footprint valleys in longer ART regimens (**Fig. 5H**).

Discussion

The results of our study showed that continuous exposure to ART modulates the homeostatic and cellular state of alveolar macrophages at the epigenetic and transcriptomic levels. AM from PWH and PrEP participants showed an increased pro-inflammatory and senescence-like profile that correlated with time on ART. This constant pro-inflammatory and cellular stress will inevitably trigger epigenetic and metabolic reprogramming (*22, 23*). The epigenetic rewiring is necessary to promote self-tolerance at the expense of a negative influence on immune responses and dysregulated immune checkpoints. We previously showed that AM obtained from PWH and people on PrEP had a diminished transcriptomic and epigenetic response to *Mtb* (*16*). Therefore, it is possible that the altered state of AM with continuous exposure to ART does contribute to the pathogenesis of other pulmonary immune and infectious diseases in the aging PWH population.

A common feature of PWH under ART is increased frequency of chronic airway diseases such as COPD, lung fibrosis and asthma compared to the healthy ART-naive population (*24*). Increased expression of AP-1 subunits in macrophages promoted lung fibrosis and contributed to the pathogenesis of COPD and asthma (*25, 26*). Moreover, PAI-1 (*SERPINE1)* and p21 protein expression that increased with time on ART promoted senescence in fibrotic lung disease (*27*). In addition, COPD and asthma displayed increased levels of IL-6 and IL-8 in serum and BAL fluid (*28-30*). The latter cytokines are a hallmark of the SASP, which is a mechanism with an emerging role in chronic airway diseases (*31-33*). SASP is also characterized by increased oxidative stress by overproduction of ROS by mitochondria or reduced antioxidant capacity (*24, 31*). We detected an impairment of major antioxidant enzymes with longer time on ART. Moreover, effector enzymes such as glutathione peroxidases (GPx) and peroxiredoxins (PRDx) that inactivate H_2O_2 had lower expression with longer time on ART. Oxidative stress-based therapeutics are under investigation as candidates to treat the underlying pathogenic mechanisms of COPD (*34*). For example, GPx mimetics boosted GPx-1 concentrations in animal models of COPD and reduced lung inflammation (*34-36*). Our results reinforced a role for AP-1 and SASP in AM senescence-like state and suggested that ART duration could play a role in chronic airway diseases in

PWH and potentially in PrEP. These mechanisms are good candidates to be target by adjuvant therapies and improved ART multi drug designs.

ART has resulted in a large drop of lung cancers among PWH. However, PWH/ART still suffer a significantly higher burden of lung cancers relative to healthy subjects with the same level of tobacco exposure (*24*). Additionally, lung cancers tend to be diagnosed at a younger age in PWH (*37*). AM have a key role in shaping the tumor microenvironment in non-small cell lung carcinoma (NSCLC), the most common type of lung cancer (*38*). While AM are not a common metastatic cell, we found dysregulated expression of oncogenes with increased time on ART. For example, the tumor suppressor *TP53* was downregulated with increased time on ART in PWH with a corresponding increased expression of its antagonist *MDM2* in PWH participants. *MDM2* overexpression is a common feature in lung cancers and the MDM2-p53 loop contributes to chemotherapeutic resistance in human malignancies (*39, 40*). Another example is *BRD4* which showed increased expression with longer time on ART. *BRD4* is frequently overexpressed in a wide range of human cancers and provides cancer cells with a survival advantage by suppressing apoptosis and correlates with poor prognosis for NSCLC (*41, 42*). More generally, we detected impairment of DDR mechanisms and cell cycle checkpoints in AM from both PWH and PrEP, which suggested that long-term ART may contribute at least in part to increased susceptibility to lung cancer.

Tuberculosis, pneumonia and more recently COVID-19 are major causes of death in PWH under ART (*15, 43*). In infectious diseases, interferons (IFN) are crucial mediators of the host response to both viral and bacterial pathogens (*44, 45*). A decreased expression of key members of the IFN signaling cascade such as the IFN receptors *IFNGR1* and *IFNAR2*, *IFI44*, *STAT1* and *CCL2* was displayed by AM of PrEP subjects independent of time on ART. In PWH, expression of *IFNGR1* and multiple interferon regulatory factors (IRFs) was decreased with time on ART in PWH. A noteworthy example is IRF9, which had been previously identified as a key TF in the AM response to *Mtb* (*16*). In addition, *TYK2* had decreased expression with longer time on ART, which is consistent with an impaired AM response to *Mtb*. *TYK2* is a kinase that mediates downstream IFN signaling. Rare *TYK2* mutations are causal for monogenic susceptibility to TB (*46, 47*). *TYK2* common regulatory variants are also strongly associated with COVID-19 severity (*48*). Overall, a decreased IFN mediated response by AM may lead to an impaired response to intracellular pathogens (*49*). Of particular relevance is the lower IFN gene expression in PrEP, independent of time, and the decreased expression with longer exposure to ART in PWH. This baseline effect combined with the previously reported impairment of IFN transcriptional responses to *Mtb*, suggests a side effect of ART that is relevant for lung infections (*16, 50*). Assuming that an increased and wider scale of PrEP uptake may involve medicating millions of healthy people,

potential long term side effects need to be further addressed, particularly in low HIV endemic settings (*51*).

Our data indicates that time on ART induces AM of both groups to display characteristics of cellular senescence via cell cycle arrest. Genes encoding proteins such as p21, CHEK2, and GADD45A had increased expression with time on ART, while GADD45B expression increased in both PrEP and PWH, independent of ART duration. These proteins are known to arrest cells in the G2 phase, preventing progression to the M phase (*52-54*). Concurrently, key pro-apoptotic inducers such as *BAX* and *BBC3* (PUMA) had decreased expression, while genes that promote cell survival via caspase inhibition such as *BCL2*, *XIAP* and *BIRC2* had increased expression with time on ART. Together, these mechanisms suggest that AM are at least partially trapped in the G2 phase, which combined with the increased cellular and mitochondrial stress and pro-inflammatory SASP profile associated with longer ART duration could be detrimental to the AM homeostatic state. Senescent-like AM endure despite the cytotoxic microenvironment they live in, albeit with an impaired ability to exert their proper role. In this context, senolytics is a new category of compounds shown to mitigate cellular senescence (*55*). Drugs such as dasatinib, quercetin, and fisetin are currently being tested in clinical trials for pulmonary diseases linked to senescence, including COPD, pulmonary fibrosis, and COVID-19 (*55*). Given the targeted action of senolytics in senescent cells, these compounds could be considered candidates for improved antiretroviral therapies or as adjuvants to lessen the ART-induced senescence-like state of AM in PWH and PrEP.

Our study has some limitations. Of the participant in the PrEP and PWH groups only one was female. Therefore, with our dataset it is not possible to test a sex bias of the ART effect. Sex is associated with different levels of drug metabolism, which might impact the ART effect in females. Participants of the PrEP group were more homogeneous using the same ART regimen while in the PWH group, participants had more complex therapies. PWH were exposed to multiple and different types of ART drugs throughout their life, which may impact the ART fold change per year observed for the PWH group. Exposure to older and more toxic ART might have contributed to the DNA damage observed in AM from PWH. Our study did not allow to address the mechanisms and the ART drug concentration that underlie the *in vivo* cellular senescence-like and poised inflammatory state of AM. However, our study has several unique strengths. First, it is the only study to date that evaluated the long-term epigenetic effect of ART in the HIV^+ and HIV^- context. Second, studies of PrEP side effects in healthy subjects have relative short follow-up periods ranging from a few months to about a year. In contrast, our retrospective cross-sectional approach evaluated a much longer effect of ART which would be challenging to implement in a longitudinal design. Additionally, participants of the PrEP and PWH groups did not include drug abusers, and both groups have similar numbers of cannabis users, which minimizes potential confounding effects of drug addiction. Another aspect is that studies on ART focused on PWH, making it

difficult to separate the HIV-effect from the ART-effect. Our study addressed this gap by studying ART effects in both HIV^+ and HIV^- persons allowing us to identify specific ART-effects. Albeit few, our study included PHW with <10 years gap between HIV diagnosis an ART initiation. These participants showed an AM profile that is in line with the time on ART and not with the time from HIV diagnosis, which reinforced the role of ART in the AM physiological state.

Pre-exposure prophylaxis is a highly effective strategy to prevent HIV acquisition (*56*). Our study does not advocate against PrEP rollout in any way as HIV infection dramatically outweigh potential ART side effects. PrEP rollout is offered to patients at increased risk of contracting HIV or residing in high exposure settings, which contributes to the decline of HIV diagnosis at the population level (*57-60*). Discontinuation, intermittent use or low PrEP compliance are known to reduce prevention effectiveness (*61*). Guidelines vary by country, but the most common recommended PrEP regimen is composed of daily oral intake of the nucleoside reverse transcriptase inhibitors (NRTIs) tenofovir disoproxil fumarate (TDF) and emtricitabine (FTC). In our study, all PrEP participants used daily TDF/FTC while most of the PWH participants used TDF/FTC (or other NRTIs) in more complex drug regimen. The mechanisms by which long-term use of NRTIs impacts cell physiology independent of HIV are poorly understood. However, *in vitro* studies have shown that leukocytes exposed to TDF/FTC display increased ROS production and lipid uptake and decreased mitochondrial mass and efferocytosis (*62, 63*). An independent study showed that exposure to Atripla (containing TDF/FTC) and Trimumeq (Abacavir and Lamivudine as NRTIs) caused monocyte-derived macrophages to display mitochondrial dysfunction, higher levels of ROS, IL-6, IL-1β and TNF (*64*). These results are consistent with epigenetic and transcriptomic alterations in AM from PrEP and PWH with increased time on ART observed in our study.

Most importantly, our study highlights the constant need for improved drug designs and multidrug therapies to mitigate side effects. Treatment and prevention of HIV infection have come a long way since the introduction of the first generation of highly toxic ART (e.g., zidovudine, stavudine, didanosine). However, our study indicates the gaps that still need to be addressed to continue improving treatment and ensuring best quality of life for ART users. If the discovered effect is exclusive to the NRTI class, drugs with distinct mechanisms of action could potentially circumvent the side effects of long-term ART. Long-lasting injectables, such as the HIV integrase inhibitor Cabotegravir and the capsid inhibitor lenacapavir, have been investigated or approved as potential PrEP regimens (*65-68*). Albeit for most people at risk of HIV acquisition the efficacy of PrEP outweighs the risks of potential side effects; our results suggest that lengthy and continuous exposure to ART requires a highly judicious use, particularly for diseases where AM play a critical role.

Materials and Methods

Subjects

Subjects included in this study were enrolled as part of our study on AM responses to *Mtb*. Hence, details of quality control regarding BAL practice, AM processing and preparation were as described previously (*16*). To study the impact of ART on the AM physiological state, we excluded subjects with previous pulmonary infections, no abnormalities in baseline lung functions, pulmonary conditions, chronic cardiovascular or metabolic diseases and illegal substances abuse. In addition, we applied the following technical exclusion criteria: samples with high noise-to-signal ratio (e.g. fraction of reads in peaks) and libraries with less than 20 million reads. Of the 50 subjects where data was available, 42 passed quality control. Of these, 14 were HIV-negative participants without ART (HC); 10 were HIV-negative subjects at a high exposure setting for HIV under prophylactic therapy (PrEP); and 18 were people with HIV under uninterrupted use of ART (PWH) (**Fig. 1A; Table 1**). Due to the limited yield of AM from BALs, some assays could not be tested in all subjects (**Table 1**). All ART regimens included at least one NRTI. For the PrEP group and for 13 out of 18 PWH, the ART regimen included TDF/FTC (**Table 1**). Of the remaining 5 PWH participants, three used Abacavir, one Etravirine, and one Tenofovir plus Rilpivirine as the NRTI on their regimens (**Table 1**). The average time from HIV diagnosis to ART varied between subjects (**Table 1**). None of the participants had developed resistance to ART. The study was approved by the Research Ethics Board of the McGill University Health Center (MUHC) (MP-CUSM-15-406). All participants signed a written informed consent and were recruited at the McGill University Health Centre (MUHC, Montreal, Canada).

Library preparation and final dataset

Data for the RNAseq, ATACseq, and H3K27ac was generated previously and is available at the Gene Expression Omnibus repository (GSE165709). The methodology used to prepare libraries and to process the data for RNA-, ATAC- and H3K27ac was described previously (*16*). For this study, we added two histone marks: H3K4me3 as a mark for active/poised transcription and H3K27me3 as a repressor mark. We also expanded the dataset for the H3K27ac mark. Libraries for the three histone marks were produced using ChIPmentation kits for histones from Diagenode (Cat # C010110009). Briefly, aliquots of crosslinked AM containing 10^5 cells stored at -80°C were thawed and resuspended on 50 µL of ice-cold Sarkosyl buffer. Next, samples were sonicated on Bioruptor Pico to obtain fragments between 200 – 500 bp. ChIPmentation was carried out using the automated IP-Star Compact with Diagenode recommended protocol and the following antibodies for H3K4me3 (Cat # C15410003, Lot # A8034D), H3K27ac (Cat # C15410196, Lot # A1723-00410), and H3K27me3 (Cat # C15410195, Lot # A0824D). After overnight incubation, DNA fragments captured by the immunoprecipitation were reverse crosslinked and unique

barcodes, Diagenode (Cat #C01011034, #C01011036 and, #C01011037), were added to the sequences of each subject and histone mark. Finally, libraries were pooled and sequenced on Illumina NovaSeq 6000 S4 PE100 aiming for an average of 50 million reads per library. The raw sequences were then processed using the same protocol previously reported (*16*). After data processing and quality control, we included 61,541 genomic regions detected with ATACseq, 21,261 (H3K4me3), 42,378 (H3K27ac), and 9,309 (H3K27me3) histone marks, and 9,919 genes in the analysis.

Statistical modeling

Differences in epigenetics and transcriptomics of AM were tested with three models comparing: (i) groups: PrEP, PWH and HC; (ii) time on ART for PrEP and for PWH, and (iii) continuous exposure to ART adjusted by group with the following linear modeling in Limma v3.60.2 (*69*):

$$\text{(i)} \quad Y_{(i,j)} \sim \beta 0_{(i)} + \beta_{Groups} \cdot X_{Groups\ (PrEP,PWH)} + \beta_{Confounders} \cdot X_{(Confounders)} + \beta_{Batches} \cdot X_{(Batches)} + \varepsilon_{(i,j)}$$

$$\text{(ii)} \quad Y_{(i,j)} \sim \beta 0_{(i)} + \beta_{Time \times Groups} \cdot (X_{Groups\ (PrEP,PWH)} \times X_{Time}) + \beta_{Groups} \cdot X_{Groups\ (HC,PrEP,PWH)} + \beta_{Confounders} \cdot X_{(Confounders)} + \beta_{Batches} \cdot X_{(Batches)} + \varepsilon_{(i,j)}$$

$$\text{(iii)} \quad Y_{(i,j)} \sim \beta 0_{(i)} + \beta_{Time} \cdot X_{Time} + \beta_{Group} \cdot X_{Groups\ (HC,PrEP,PWH)} + \beta_{Confounders} \cdot X_{(Confounders)} + \beta_{Batches} \cdot X_{(Batches)} + \varepsilon_{(i,j)}$$

Where $Y_{(i,j)}$ represents the estimate quantification for each feature "i" and sample "j". Briefly, quantification matrices with $Y_{(i,j)}$ were depth-normalized with *calcNormFactors* upperquartile (0.75) for the epigenetic assays and TMM for the transcriptomic data using edgeR v.3.40.2. The mean-variance relationship and log2 count per million was estimated with *voom* v3.60.2 (*70*). In these models, $\beta 0_{(i)}$ represents the intercept for the feature "i". In model (i) the intercept corresponds to the HC group. In models (ii) and (iii) the intercept represents the predicted value of Y when the continuous variables of interest (e.g. Time on ART) are equal to zero. In model (i) β_{Groups} output two β that captures the main effect for PrEP and PWH relative to the intercept (HC), respectively. The contrast between PrEP and PWH was done post-hoc by contrasting their respective β coefficients. In models (ii) and (iii) the β_{Groups} adjusts for the main effect of group membership on the "Time" effect. In model (ii) the $\beta_{Time \times Groups}$ is an interaction term that output two β that captures the effect of time on ART for PrEP and for PWH. In model (iii) the β_{Time} term estimates the effect of continuous exposure to ART adjusting for the main group effect. The coefficients of $\beta_{Confounders}$ included the phenotypic confounders: sex, chronological age, smoking status, *Mtb* infection, and a seasonal effect of BAL collection. The coefficients of $\beta_{Batches}$ included the technical covariates: sequencing batch, Fractions of Reads in Peaks for the epigenetic assays, time of AM incubation, and number of AM seeded per plate. The $\varepsilon_{(i,j)}$ represents the residual error not explained by variables in the model. A Benjamini–Hochberg (BH) procedure was used to estimate FDR for each assay in the three tested models. Genes and peaks presenting FDR < 10% were selected for further follow-up on the pathway and transcription factors analyses.

Peak to gene assignment and Gene-Set Co-Regulation Analysis

Genes were assigned to peaks using *annotatePeak* from ChIPSeeker v.1.40 (*71*). Each peak from the epigenetic assays was assigned to the nearest gene and to any additional flanking gene at < 5 kilobases from a peak. The list of the nearest gene and nearest plus flanking genes per peak is shown in Tables S2 and S4. The gene set co-regulation analysis (GESECA) was performed with *geseca* from *fgsea* v1.30 (*72*). Briefly, GESECA identifies patterns of co-expression, co-accessibility or co-marks, rather than individual genes, enriched within a predefined gene sets (e.g. HALLMARK pathways) across different samples. The advantage of GESECA compared to standard Gene Set Enrichment or Gene Ontology Analysis is that it does not require a prior explicit contrast or ranking via fold changes or statistical significance. Comparisons between groups or time courses can be made later ad hoc. GESECA was developed to compare datasets with a large number of different biological conditions such as timepoints (e.g. time on ART). It uses as input a quantification or a covariate-reduced matrix (gene/peak by samples) to calculate the proportion of variance between samples explained by each gene set. *P*-value significance for each gene set (pathway) is then calculated by permutation (n = 10,000) of genes between lists. For the GESECA we used covariate-regressed matrices obtained via *removeBatchEffect* from *Limma* for transcriptome and epigenetic (nearest plus flanking genes) for all three comparisons (between groups, time effect per group, and time effect with group as a covariate). By design, GESECA is "contrast" independent, which means it does not directly compare phenotypic groups but rather variability between samples (i.e., individuals GESECA scores). To assess if the variability between samples was driven by group or a time effect, we contrasted the GESECA score per subject with the phenotype of interest in each model. For group comparisons, we used a Wilcoxon test to estimate significant differences between the median GESECA score per group. To test the correlation of time and GESECA scores we used linear regression both for the per-group time comparisons (PrEP and PWH) and the overall effect of continuous exposure to ART. An FDR correction for these comparisons was calculated using BH. We ran GESECA using the HALLMARK pathway gene-set collection from MSigDB (*73*). Additionally, as senescence is a hypothesized mechanism contributing to comorbidities in the aging PWH we included the REACTOME "Senescence" branch in the HALLMARK gene-set list (*74*).

Motif enrichment analysis

To test for enrichment of transcription factor binding motifs in differential regions associated with the continuous exposure to ART, we used *findMotifsGenome* from HOMER v4.11 (*75*). Differential accessible chromatin regions (ATAC) were compared to (i) 500,000 random regions matched by size and GC content and (ii) versus nonsignificant 20,060 regions at FDR > 50% in the continuous exposure to

ART analysis. We considered a motif to be significantly enriched if it presented FDR < 1%. Pairwise correlation of TF motif binding by similarities was performed with STAMP v.1.3 default parameters (*76*).

Footprint analysis

Transcription factor Occupancy also known as Footprint analysis was performed with TOBIAS v0.14.0 using covariate-regressed ATAC traces (*77*). Briefly, we adapted the concept of regressing out known confounders (e.g. covariables and batch effects) prior to chromatin modeling as proposed in (*78, 79*). First, signal intensity for regions with increased chromatin accessibility (ATAC) with continuous exposure to ART was calculated from QCed BAM files of each subject using *bamCoverage* from deepTools v.3.5.1 with --binSize 1 and --normalizeUsing CPM (*80*). Next, ATAC signals per sample were summarized using *multiBigwigSummary* (*80*). The output of this procedure was a quantification matrix with ATAC signal intensity at 1bp resolution per subject. This matrix was then exported into R v4.4.0. We used *removeBatchEffect* from *Limma* to calculate the residual signal intensity after regressing the coefficients $\beta_{\text{Group (HC,PrEP,PWH)}}$, $\beta_{\text{Confounders}}$ and β_{Batch} while preserving β_0 and $\beta_{Time.}$ from the model (iii). This residual signal intensity matrix represents ATAC traces mimicking the linear model (iii) tested in the differential analysis. The residual traces were exported from R and converted back to individual BigWigs using *bedGraphToBigWig* from UCSC. Next, we used TOBIAS *ScoreBigwig* to identify regions of protein binding across differential regions with continuous exposure to ART. To predict which TF bind to the footprints detected, we combined footprint scores per subject with information on TF binding motifs using TOBIAS *BINDetect*. We used as reference 919 motifs from the JASPAR CORE database for vertebrates (*81*). *BINDetect* calculates a mean footprint score for each of the 919 TF binding sites split per subject. This score can be compared between samples to access differential changes as TF having high-scoring footprints denotes a clearer and more defined bound. To test if the TF footprint intensity was associated with time, we fit a linear model comparing the footprint mean intensity per TF with the continuous exposure to ART. Fold change per year was calculated via the exponential of the model slopes and FDR *p*-values were adjusted with BH. TOBIAS *PlotAggregate* was used to summarize traces for selected examples of TF (**Fig. 5G**). The residual signal intensities for selected examples of footprints were plotted per timepoint and according to their chromosomal position for each assay using *ggplot2* v.3.4.4 with *geom_spline* from *ggformula* v.0.10.4 (*82, 83*) (**Fig. 5C and H**).

Table 1. Characteristics of enrolled subjects.

Sample	Group	Sex	Age range	Smoker [a]	ART regiment	Molecules	Duration of HIV	Time on ART [b]	RNA	ATAC	H3K27ac	H3K27me3	H3K4me3
AMC2	HC	M	56 – 60	YES	.	.	.	0	X	X	.	.	.
AMC7	HC	M	46 – 50	NO	.	.	.	0	X	X	.	.	.
AMC8	HC	M	21 – 25	NO	.	.	.	0	X	.	.	.	.
AMC9	HC	M	56 – 60	YES	.	.	.	0	X	X	X	X	X
AMC10	HC	M	26 – 30	NO	.	.	.	0	X	X	X	.	X
AMC11	HC	M	26 – 30	NO	.	.	.	0	X	X	X	.	.
AMC12	HC	M	36 – 40	NO	.	.	.	0	X	X	X	.	X
AMC13	HC	F	31 – 35	NO	.	.	.	0	X	X	X	X	X
AMC14	HC	F	36 – 40	YES	.	.	.	0	X	X	.	X	.
AMC15	HC	F	41 – 45	NO	.	.	.	0	X	X	.	.	.
AMC16	HC	F	36 – 40	NO	.	.	.	0	X	X	.	.	.
AMC18	HC	M	26 – 30	NO	.	.	.	0	X	X	.	.	.
AMC19	HC	M	51 – 55	NO	.	.	.	0	X	X	.	.	.
AMC22	HC	F	41 – 45	YES	.	.	.	0	X	X	X	X	X
AMP33	PrEP	M	26 – 30	NO	Truvada	Emtricitabine, Tenofovir	.	0.1	X	X	X	.	X
AMP35	PrEP	M	21 – 25	YES	Truvada	Emtricitabine, Tenofovir	.	1.0	X	X	X	X	X
AMP23	PrEP	M	61 – 65	YES	Truvada	Emtricitabine, Tenofovir	.	1.3	X	.	.	.	.
AMP32	PrEP	M	36 – 40	NO	Truvada	Emtricitabine, Tenofovir	.	1.5	X	X	.	.	.
AMP30	PrEP	M	31 – 35	NO	Truvada	Emtricitabine, Tenofovir	.	2.4	X	.	.	.	.
AMP31	PrEP	M	36 – 40	NO	Truvada	Emtricitabine, Tenofovir	.	2.5	X	X	.	.	.
AMP34	PrEP	M	41 – 45	NO	Truvada	Emtricitabine, Tenofovir	.	3.2	X	X	X	.	.
AMP36	PrEP	M	26 – 30	YES	Truvada	Emtricitabine, Tenofovir	.	3.4	X	X	X	X	X
AMP27	PrEP	M	36 – 40	NO	Truvada	Emtricitabine, Tenofovir	.	4.4	X	X	X	X	X
AMP32.2[c]	PrEP	M	41 – 45	NO	Truvada	Emtricitabine, Tenofovir	.	4.8	.	.	X	X	X
AMP29	PrEP	M	36 – 40	YES	Truvada	Emtricitabine, Tenofovir	.	5.4	.	X	.	.	.
AMV23	PWH	M	31 – 35	YES	Genvoya	Emtricitabine, Tenofovir, Cobicistat, Elvitegravir	0.7	0.6	X	X	X	X	X
AMV27	PWH	M	56 – 60	NO	Descovy, Isentress	Emtricitabine, Tenofovir, Raltegravir	1.1	1.0	X	X	.	.	.
AMV17	PWH	M	41 – 45	YES	Viread, Tivicay, Edurant	Tenofovir, Dolutegravir, Rilpivirine	3.7	2.7	X	X	X	X	X
AMV21	PWH	M	61 – 65	YES	Truvada, Tivicay	Emtricitabine, Tenofovir, Dolutegravir	6.2	5.7	X	X	X	X	X
AMV5	PWH	M	56 – 60	NO	Truvada	Emtricitabine, Tenofovir	21.5	8.5	X	X	.	.	.
AMV26	PWH	M	46 – 50	NO	Atripla	Emtricitabine, Tenofovir, Efavirenz	18.4	8.5	X	X	.	.	.
AMV25	PWH	M	56 – 60	NO	Atripla	Emtricitabine, Tenofovir, Efavirenz	11.6	11.3	.	X	X	.	X
AMV19	PWH	M	56 – 60	YES	Triumeq	Abacavir, Dolutegravir, Lamivudine	12.7	11.4	X	X	.	.	.
AMV15	PWH	M	51 – 55	NO	Triumeq	Abacavir, Dolutegravir, Lamivudine	14.1	13.7	X	X	X	X	X
AMV9	PWH	M	56 – 60	YES	Truvada, Isentress	Emtricitabine, Tenofovir, Raltegravir	14.8	13.8	X	X	X	.	.
AMV16	PWH	M	56 – 60	YES	Genvoya	Emtricitabine, Tenofovir, Cobicistat, Elvitegravir	30.1	15.7	X	.	X	X	X
AMV4	PWH	M	56 – 60	YES	Biktarvy	Emtricitabine, Tenofovir, Bictegravir	17.0	17.0	X	X	X	.	.
AMV11	PWH	M	46 – 50	YES	Complera	Emtricitabine, Tenofovir, Rilpivirine	19.9	18.9	X	X	X	X	X
AMV3	PWH	M	51 – 55	NO	Truvada, Isentress	Emtricitabine, Tenofovir, Raltegravir	21.4	19.4	X	X	.	.	.
AMV14	PWH	M	56 – 60	YES	Kivexa, Reyataz, Norvir	Abacavir, Lamivudine, Atazanavir, Ritonavir	21.2	20.5	X	X	X	X	.
AMV20	PWH	M	56 – 60	YES	Truvada, Tybost, Prezista	Emtricitabine, Tenofovir, Cobiscistat, Darunavir	21.3	20.7	X	X	.	.	.
AMV10	PWH	M	56 – 60	NO	Intelence, Prezista, Tivicay, Norvir	Etravirine, Darunavir, Dolutegravir, Ritonavir	25.4	21.5	X	X	.	.	.
AMV13	PWH	F	51 – 55	YES	Complera, Reyataz, Isentress	Emtricitabine, Tenofovir, Rilpivirine, Atazanavir, Raltegravir	23.1	23.1	X	X	X	X	X

[a] Cigarettes and/or cannabis. [b] Fraction of years from the ART start to BAL collection. [c] BALs of the same subject collected at different dates. ART, antiretroviral therapy; BAL, bronchoalveolar lavage, HC, healthy controls; PrEP healthy subject on pre-exposure prophylaxis; PWH, people with HIV under ART; an X indicate samples employed in each of the five tested assays.

Supplementary Materials

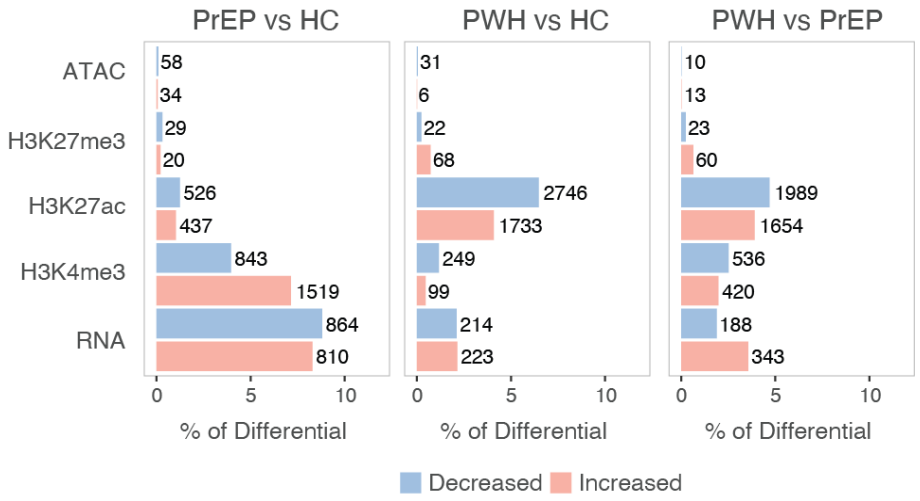

Fig. S1. Differential epigenetic and transcriptomic features between groups.

Differential epigenetic and transcriptomic features between groups are shown for the three baseline comparisons. The bar plots show the proportion of differential regions/genes at an FDR of 10% and absolute log2FC > 0.2 relative to the total number of tested regions/genes for time on ART in PrEP vs HC (left), PWH vs HC (center) and PWH vs PrEP (right). The absolute number of differential regions/genes are shown next to the bars. Decreased and increased DNA accessibility, histone marks, and gene expression are shown in blue and red, respectively.

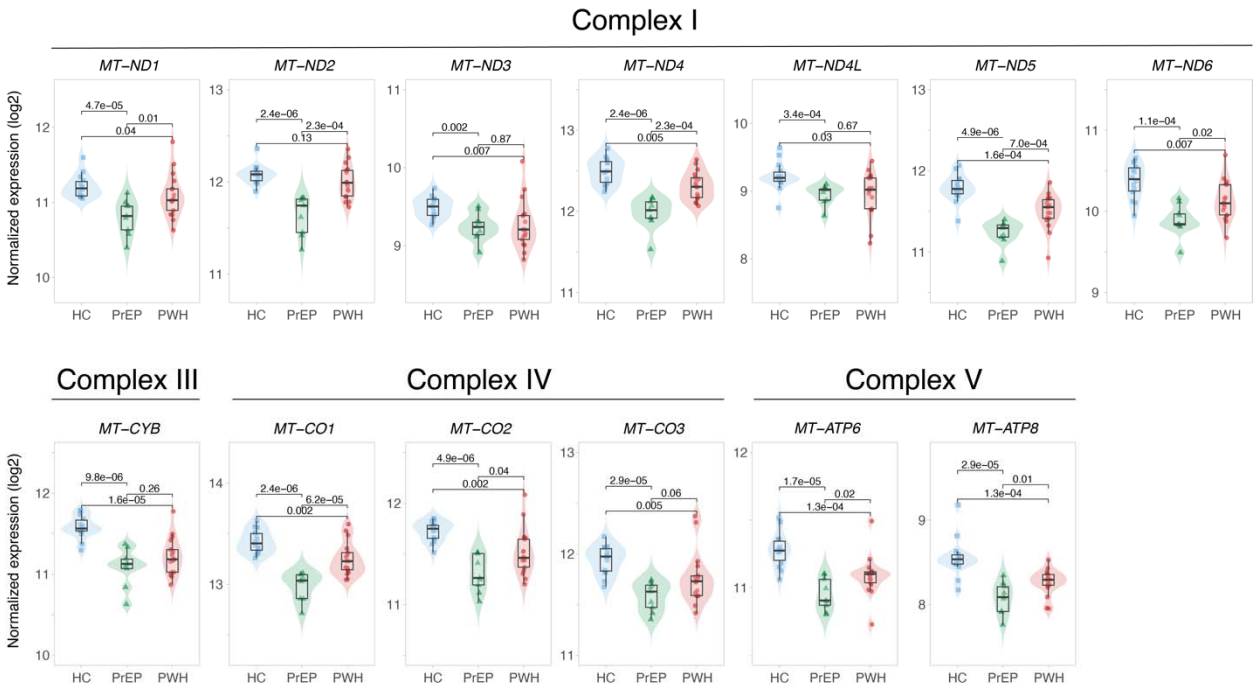

Fig. S2. Reduced expression of mitochondrial genes encoding enzymes of the electron transport chain.

Normalized gene expression per subject and group for mitochondrial genes encoding proteins of the human electron transport chain. *p*-values for the Wilcoxon test comparing differences in median expressions between groups are shown at the top.

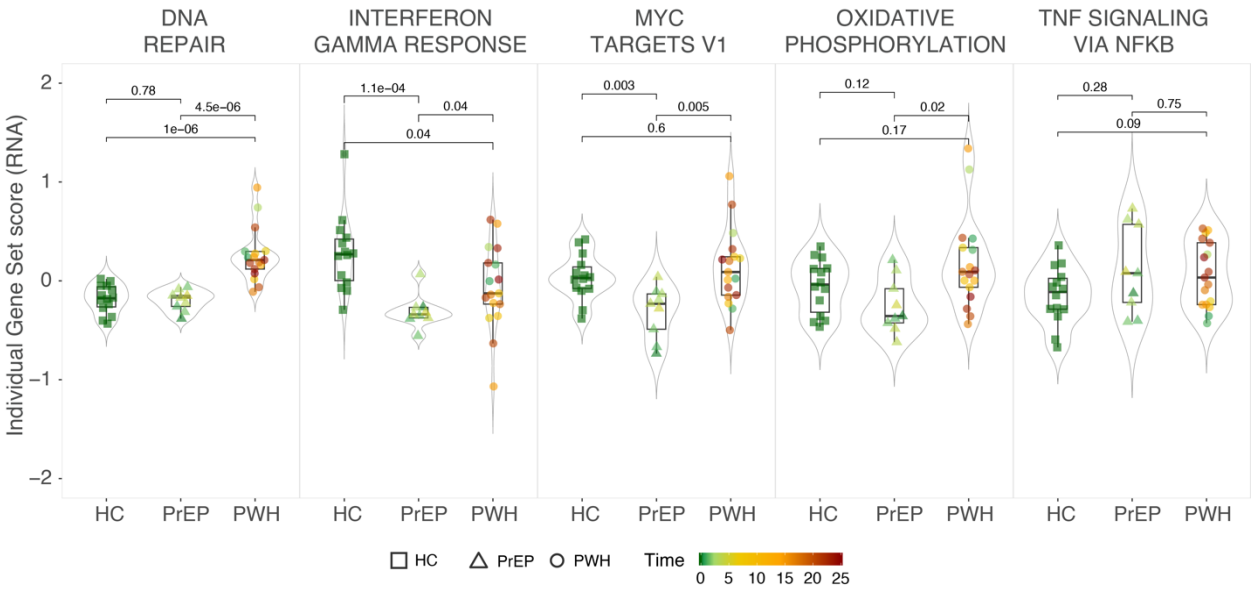

Fig. S3. Comparison between GESECA scores and phenotypic groups as in Fig. 2C indicating time on ART.

GESECA scores for the transcriptomic assay for each subject are plotted on the y-axis separated by phenotypic group on the x-axis. The *p*-values shown at the top are for the group compassions shown in Fig 2C. The color shades indicate the length of time on ART for PrEP and for PWH.

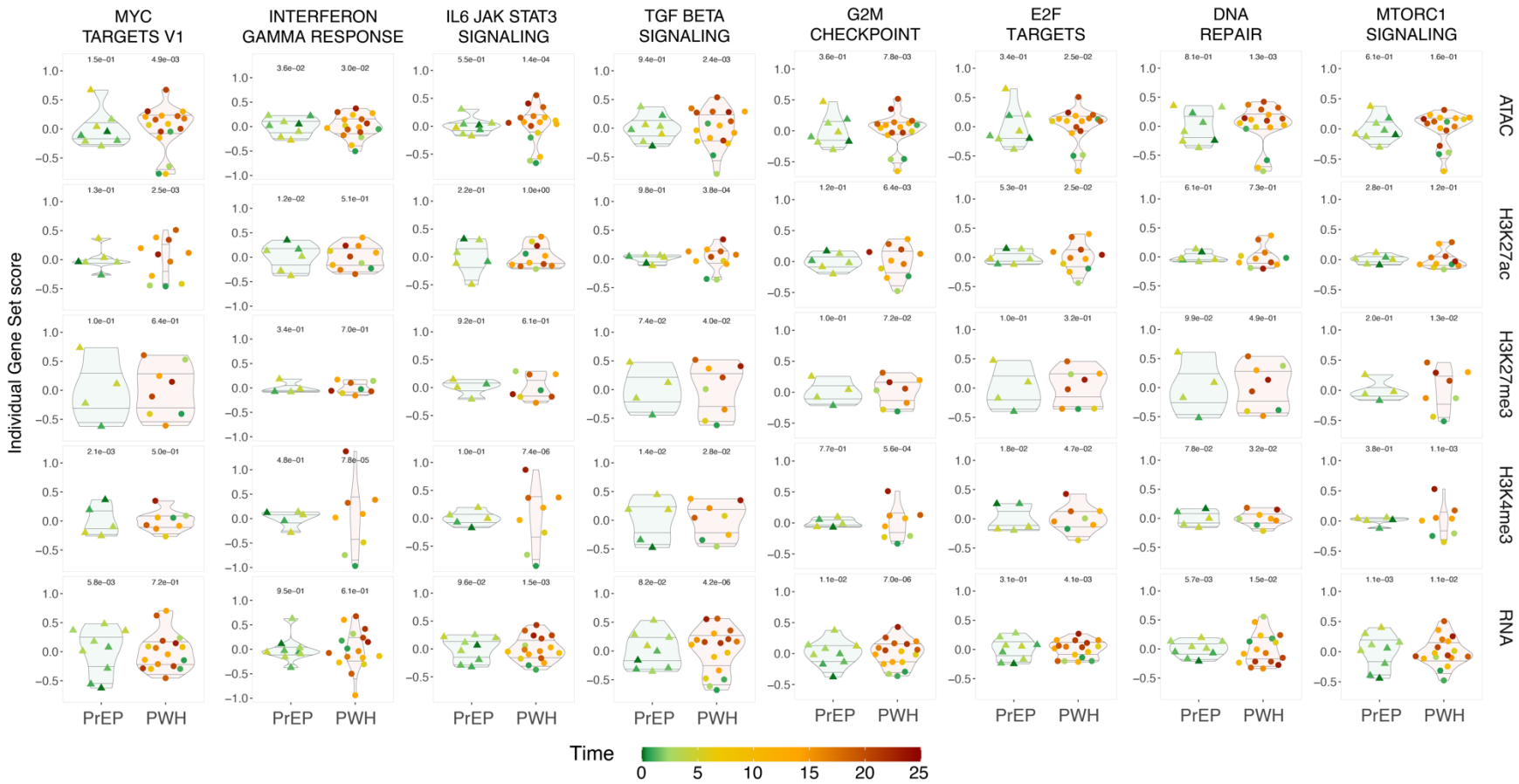

Fig. S4. **Per subject GESECA score for eight pathways affected by time as a complement to Fig. 4B.**

Each dot represents an individual with colour shading indicating time on ART for PrEP or for PWH. GESECA scores per individual were plotted on the y-axis and subjects were separated by groups on the x-axis. The quantiles are shown as violins for reference of the distribution. A linear model was used to test the correlation of time on ART on PrEP or on PWH with the individual GESECA scores for all tested assays. *P*-values for the correlations are shown at the top of each comparison.

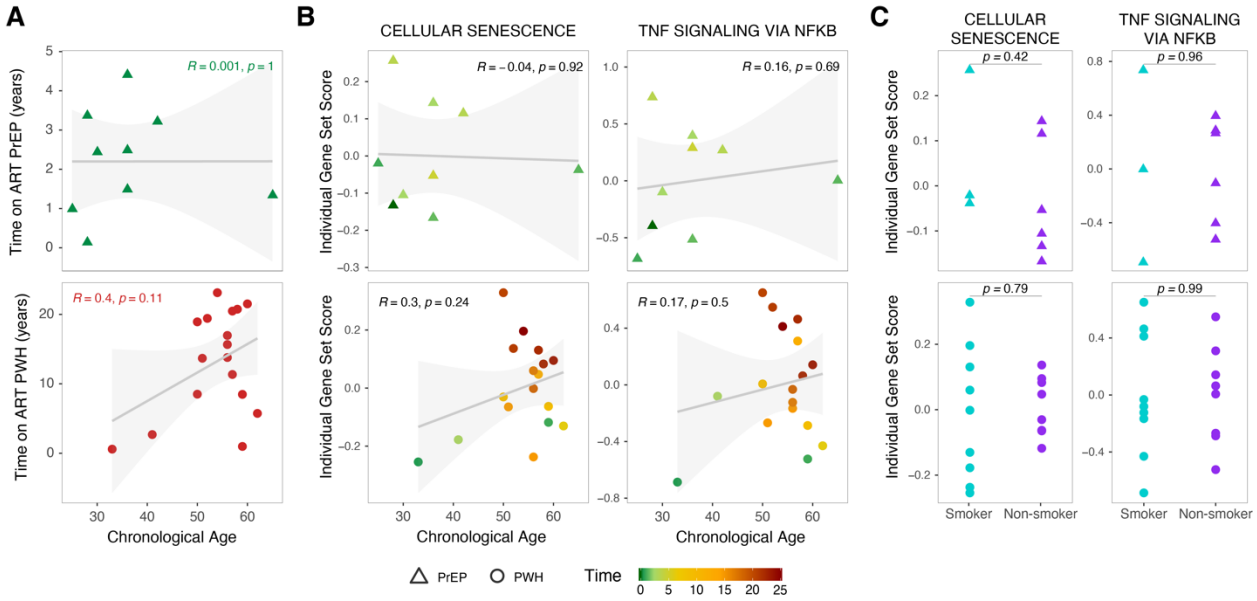

Fig. S5. Effect of chronological age and smoking status on ART associated pathways.

(**A**) Correlation plot between chronological age and time on ART for people on PrEP (top) or for PWH (bottom). Each shape and color represent a subject in each group. (**B**) Correlation plot between individual gene set scores of the RNA assay for two main pathways associated with time and chronological age. Each shape represents a sample, and colour shades indicate the time on ART for PrEP (top) or for PWH (bottom). (**C**) Comparison of individual gene set scores of the RNA assay and smoking status for two main pathways associated with time on ART. PrEP subjects are shown on top and PWH at the bottom graphs. The *p*-values indicate differences in the mean GESECA scores between groups.

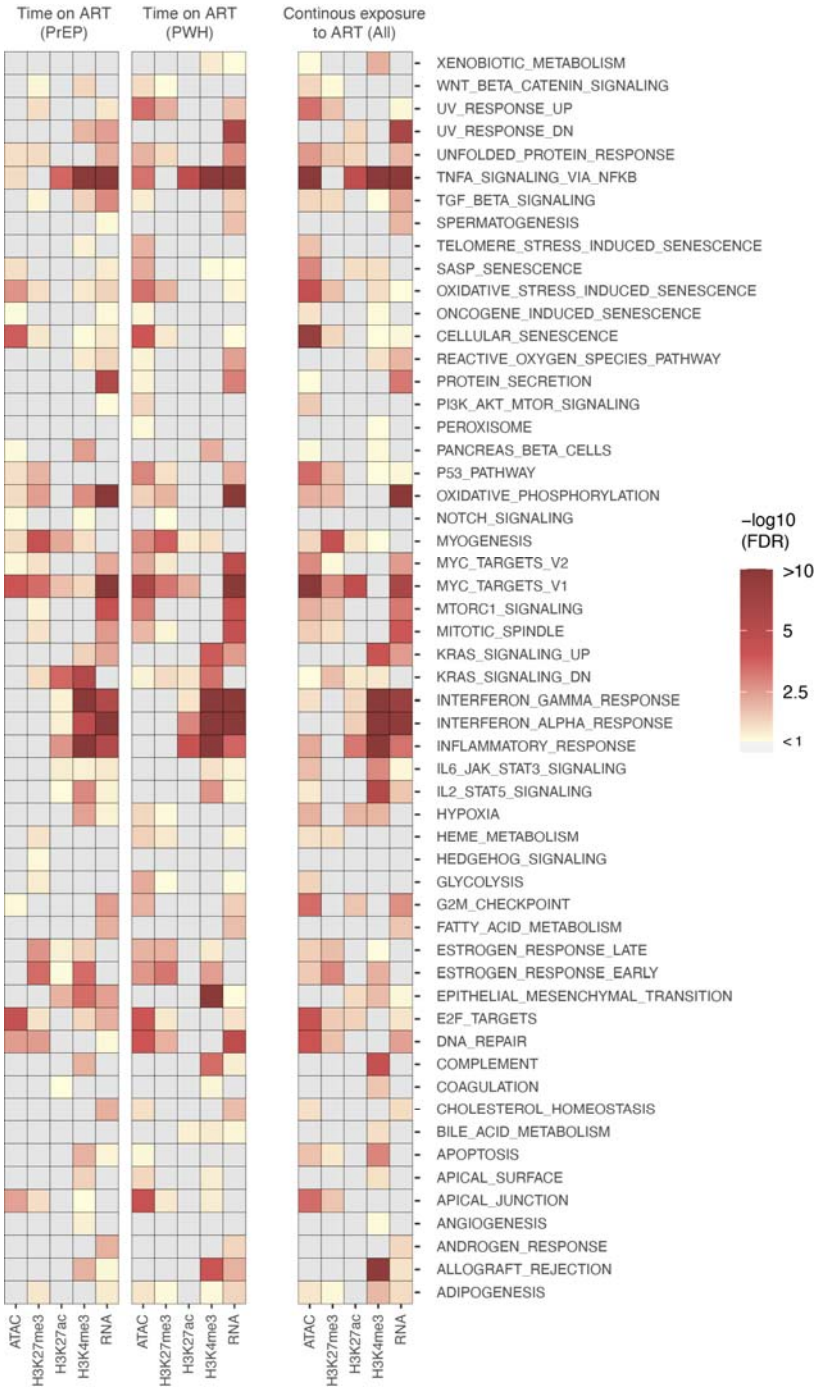

Fig. S6. Comparison of GESECA association with continuous exposure to ART and the effect of time per group.

GESECA for Hallmark pathways for time on ART in PrEP (left), in PWH (center), and continuous exposure to ART irrespective of the group (right). FDR for all tested pathways is shown as a heatmap with increased significance from yellow to red. Pathways with FDR > 10% are shown in grey.

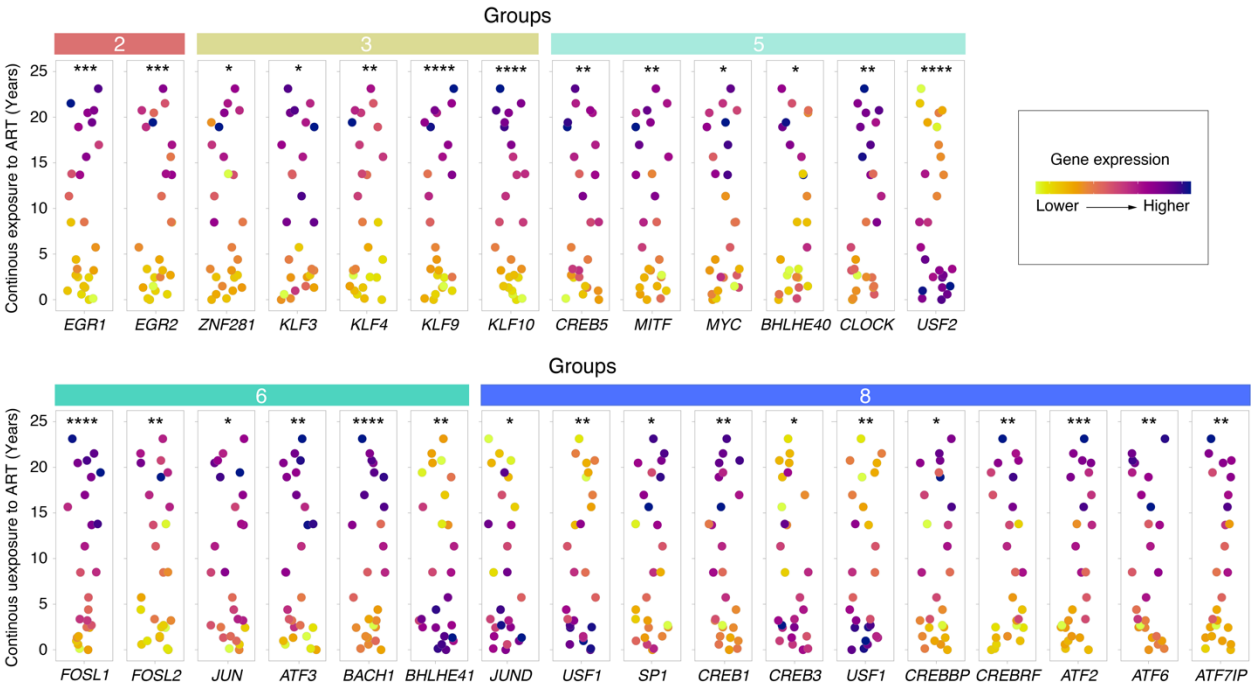

Fig. S7. Coordinated increased gene expression in AM of transcription factors with motifs associated with continuous exposure to ART.

Boxplots showing the time of continuous exposure to ART on the y-axes for genes encoding transcription factors (TF) shown in Fig. 5E. Each dot indicates the continuous exposure to ART for all subjects with the shades indicating the direction of expression changes from lower (yellow) to higher (purple). The boxes at the top of the plots show the corresponding groups for the TF gene. Asterisks indicate the magnitude of FDR significance for the continuous exposure to ART comparison. *, FDR < 10%, ** FDR < 1%, *** FDR < 0.1% and, **** FDR < 0.01%.

Reference

1. A. Tseng, J. Seet, E. J. Phillips, The evolution of three decades of antiretroviral therapy: challenges, triumphs and the promise of the future. *Br J Clin Pharmacol* **79**, 182-194 (2015).
2. T. A. Premeaux, L. C. Ndhlovu, Decrypting biological hallmarks of aging in people with HIV. *Curr Opin HIV AIDS* **18**, 237-245 (2023).
3. B. H. Balderson *et al.*, Chronic illness burden and quality of life in an aging HIV population. *AIDS Care* **25**, 451-458 (2013).
4. N. G. A. Nanditha *et al.*, Excess burden of age-associated comorbidities among people living with HIV in British Columbia, Canada: a population-based cohort study. *BMJ open* **11**, e041734 (2021).
5. A. Chawla *et al.*, A Review of Long-Term Toxicity of Antiretroviral Treatment Regimens and Implications for an Aging Population. *Infect Dis Ther* **7**, 183-195 (2018).
6. M. E. Fitzpatrick, K. M. Kunisaki, A. Morris, Pulmonary disease in HIV-infected adults in the era of antiretroviral therapy. *AIDS* **32**, 277-292 (2018).
7. K. Crothers *et al.*, HIV infection and risk for incident pulmonary diseases in the combination antiretroviral therapy era. *Am J Respir Crit Care Med* **183**, 388-395 (2011).
8. I. C. Schoepf *et al.*, Epigenetic ageing accelerates before antiretroviral therapy and decelerates after viral suppression in people with HIV in Switzerland: a longitudinal study over 17 years. *Lancet Healthy Longev* **4**, e211-e218 (2023).
9. M. Dube *et al.*, Spontaneous HIV expression during suppressive ART is associated with the magnitude and function of HIV-specific CD4(+) and CD8(+) T cells. *Cell Host Microbe* **31**, 1507-1522 e1505 (2023).
10. S. Fastenackels *et al.*, HIV-mediated immune aging in young adults infected perinatally or during childhood. *AIDS* **33**, 1705-1710 (2019).
11. C. X. Yang *et al.*, Occurrence of Accelerated Epigenetic Aging and Methylation Disruptions in Human Immunodeficiency Virus Infection Before Antiretroviral Therapy. *J Infect Dis* **223**, 1681-1689 (2021).
12. M. J. Corley *et al.*, Abrupt and altered cell-type specific DNA methylation profiles in blood during acute HIV infection persists despite prompt initiation of ART. *PLoS Pathog* **17**, e1009785 (2021).
13. J. J. Bigna, A. M. Kenne, S. L. Asangbeh, A. T. Sibetcheu, Prevalence of chronic obstructive pulmonary disease in the global population with HIV: a systematic review and meta-analysis. *Lancet Glob Health* **6**, e193-e202 (2018).
14. Y. Wang *et al.*, Systematic Review and Meta-Analyses of The Interaction Between HIV Infection And COVID-19: Two Years' Evidence Summary. *Front Immunol* **13**, 864838 (2022).
15. WHO, Global tuberculosis report 2023. (2023).
16. W. Correa-Macedo *et al.*, Alveolar macrophages from persons living with HIV show impaired epigenetic response to Mycobacterium tuberculosis. *J Clin Invest* **131**, (2021).
17. D. R. Feikin, C. Feldman, A. Schuchat, E. N. Janoff, Global strategies to prevent bacterial pneumonia in adults with HIV disease. *Lancet Infect Dis* **4**, 445-455 (2004).
18. J. S. Zifodya, K. Crothers, Treating bacterial pneumonia in people living with HIV. *Expert Rev Respir Med* **13**, 771-786 (2019).

19. T. Maitre *et al.*, Increasing burden of noninfectious lung disease in persons living with HIV: a 7-year study using the French nationwide hospital administrative database. *Eur Respir J* **52**, (2018).
20. C. B. Haas *et al.*, Trends and risk of lung cancer among people living with HIV in the USA: a population-based registry linkage study. *Lancet HIV* **9**, e700-e708 (2022).
21. K. Sigel, A. Makinson, J. Thaler, Lung cancer in persons with HIV. *Curr Opin HIV AIDS* **12**, 31-38 (2017).
22. M. G. Netea *et al.*, Defining trained immunity and its role in health and disease. *Nat Rev Immunol* **20**, 375-388 (2020).
23. M. Divangahi *et al.*, Trained immunity, tolerance, priming and differentiation: distinct immunological processes. *Nat Immunol* **22**, 2-6 (2021).
24. I. Konstantinidis *et al.*, HIV-associated lung disease. *Nat Rev Dis Primers* **9**, 39 (2023).
25. A. C. Ucero *et al.*, Fra-2-expressing macrophages promote lung fibrosis in mice. *J Clin Invest* **129**, 3293-3309 (2019).
26. A. Birnhuber, V. Biasin, D. Schnoegl, L. M. Marsh, G. Kwapiszewska, Transcription factor Fra-2 and its emerging role in matrix deposition, proliferation and inflammation in chronic lung diseases. *Cell Signal* **64**, 109408 (2019).
27. C. Jiang *et al.*, Serpine 1 induces alveolar type II cell senescence through activating p53-p21-Rb pathway in fibrotic lung disease. *Aging Cell* **16**, 1114-1124 (2017).
28. M. Rincon, C. G. Irvin, Role of IL-6 in asthma and other inflammatory pulmonary diseases. *Int J Biol Sci* **8**, 1281-1290 (2012).
29. F. Liu *et al.*, Diagnosis values of IL-6 and IL-8 levels in serum and bronchoalveolar lavage fluid for invasive pulmonary aspergillosis in chronic obstructive pulmonary disease. *J Investig Med* **69**, 1344-1349 (2021).
30. Y. P. Chang *et al.*, Upregulated microRNA-125b-5p in patients with asthma-COPD overlap mediates oxidative stress and late apoptosis via targeting IL6R/TRIAP1 signaling. *Respir Res* **25**, 64 (2024).
31. C. Lopez-Otin, M. A. Blasco, L. Partridge, M. Serrano, G. Kroemer, Hallmarks of aging: An expanding universe. *Cell* **186**, 243-278 (2023).
32. M. Kumar, W. Seeger, R. Voswinckel, Senescence-associated secretory phenotype and its possible role in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* **51**, 323-333 (2014).
33. C. Xie, M. M. Ya Likun, Q. L. Luo, J. C. Dong, Role of cellular senescence in inflammatory lung diseases. *Cytokine & growth factor reviews* **70**, 26-40 (2023).
34. P. J. Barnes, Oxidative stress-based therapeutics in COPD. *Redox Biol* **33**, 101544 (2020).
35. E. Zinellu *et al.*, Glutathione Peroxidase in Stable Chronic Obstructive Pulmonary Disease: A Systematic Review and Meta-analysis. *Antioxidants (Basel)* **10**, (2021).
36. L. C. Oostwoud *et al.*, Apocynin and ebselen reduce influenza A virus-induced lung inflammation in cigarette smoke-exposed mice. *Sci Rep* **6**, 20983 (2016).
37. A. E. Coghill *et al.*, Advanced stage at diagnosis and elevated mortality among US patients with cancer infected with HIV in the National Cancer Data Base. *Cancer* **125**, 2868-2876 (2019).
38. M. Casanova-Acebes *et al.*, Tissue-resident macrophages provide a pro-tumorigenic niche to early NSCLC cells. *Nature* **595**, 578-584 (2021).

39. H. Hou, D. Sun, X. Zhang, The role of MDM2 amplification and overexpression in therapeutic resistance of malignant tumors. *Cancer Cell Int* **19**, 216 (2019).
40. D. Zong *et al.*, BRD4 Levels Determine the Response of Human Lung Cancer Cells to BET Degraders That Potently Induce Apoptosis through Suppression of Mcl-1. *Cancer Res* **80**, 2380-2393 (2020).
41. J. Hu, D. Pan, G. Li, K. Chen, X. Hu, Regulation of programmed cell death by Brd4. *Cell Death Dis* **13**, 1059 (2022).
42. Y. F. Liao *et al.*, High level of BRD4 promotes non-small cell lung cancer progression. *Oncotarget* **7**, 9491-9500 (2016).
43. Y. Alexandrova, C. T. Costiniuk, M. A. Jenabian, Pulmonary Immune Dysregulation and Viral Persistence During HIV Infection. *Front Immunol* **12**, 808722 (2021).
44. J. L. Casanova *et al.*, The ouroboros of autoimmunity. *Nat Immunol* **25**, 743-754 (2024).
45. F. McNab, K. Mayer-Barber, A. Sher, A. Wack, A. O'Garra, Type I interferons in infectious disease. *Nat Rev Immunol* **15**, 87-103 (2015).
46. S. Boisson-Dupuis *et al.*, Tuberculosis and impaired IL-23-dependent IFN-gamma immunity in humans homozygous for a common TYK2 missense variant. *Sci Immunol* **3**, (2018).
47. G. Kerner *et al.*, Homozygosity for TYK2 P1104A underlies tuberculosis in about 1% of patients in a cohort of European ancestry. *Proc Natl Acad Sci U S A* **116**, 10430-10434 (2019).
48. E. Pairo-Castineira *et al.*, GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19. *Nature* **617**, 764-768 (2023).
49. D. I. Kotov *et al.*, Early cellular mechanisms of type I interferon-driven susceptibility to tuberculosis. *Cell* **186**, 5536-5553 e5522 (2023).
50. K. Kgoadi *et al.*, Alveolar macrophages from persons with HIV mount impaired TNF signaling networks to M. tuberculosis infection. *Nature communications* **16**, 2397 (2025).
51. V. Pilkington, A. Hill, S. Hughes, N. Nwokolo, A. Pozniak, How safe is TDF/FTC as PrEP? A systematic review and meta-analysis of the risk of adverse events in 13 randomised trials of PrEP. *J Virus Erad* **4**, 215-224 (2018).
52. F. B. Charrier-Savournin *et al.*, p21-Mediated nuclear retention of cyclin B1-Cdk1 in response to genotoxic stress. *Mol Biol Cell* **15**, 3965-3976 (2004).
53. X. W. Wang *et al.*, GADD45 induction of a G2/M cell cycle checkpoint. *Proc Natl Acad Sci U S A* **96**, 3706-3711 (1999).
54. L. Zannini, D. Delia, G. Buscemi, CHK2 kinase in the DNA damage response and beyond. *J Mol Cell Biol* **6**, 442-457 (2014).
55. S. Chaib, T. Tchkonia, J. L. Kirkland, Cellular senescence and senolytics: the path to the clinic. *Nat Med* **28**, 1556-1568 (2022).
56. M. S. Cohen *et al.*, Antiretroviral Therapy for the Prevention of HIV-1 Transmission. *N Engl J Med* **375**, 830-839 (2016).
57. J. Marrazzo *et al.*, HIV Preexposure Prophylaxis With Emtricitabine and Tenofovir Disoproxil Fumarate Among Cisgender Women. *JAMA* **331**, 930-937 (2024).
58. S. N. Cox *et al.*, Impact of HIV self-testing for oral pre-exposure prophylaxis scale-up on drug resistance and HIV outcomes in western Kenya: a modelling study. *Lancet HIV* **11**, e167-e175 (2024).

59. K. Choopanya *et al.*, Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet* **381**, 2083-2090 (2013).
60. J. M. Baeten *et al.*, Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. *N Engl J Med* **367**, 399-410 (2012).
61. H. Jourdain, S. B. de Gage, D. Desplas, R. Dray-Spira, Real-world effectiveness of pre-exposure prophylaxis in men at high risk of HIV infection in France: a nested case-control study. *Lancet Public Health* **7**, e529-e536 (2022).
62. E. R. Bowman *et al.*, In Vitro Exposure of Leukocytes to HIV Preexposure Prophylaxis Decreases Mitochondrial Function and Alters Gene Expression Profiles. *Antimicrob Agents Chemother* **65**, (2020).
63. J. Cohen, L. D'Agostino, F. Tuzer, C. Torres, HIV antiretroviral therapy drugs induce premature senescence and altered physiology in HUVECs. *Mech Ageing Dev* **175**, 74-82 (2018).
64. J. Wallace *et al.*, Anti-HIV Drugs Cause Mitochondrial Dysfunction in Monocyte-Derived Macrophages. *Antimicrob Agents Chemother* **66**, e0194121 (2022).
65. C. G. Rivera *et al.*, HIV Pre-Exposure Prophylaxis: New and Upcoming Drugs to Address the HIV Epidemic. *Drugs* **83**, 1677-1698 (2023).
66. R. J. Landovitz *et al.*, Cabotegravir for HIV Prevention in Cisgender Men and Transgender Women. *N Engl J Med* **385**, 595-608 (2021).
67. S. Delany-Moretlwe *et al.*, Cabotegravir for the prevention of HIV-1 in women: results from HPTN 084, a phase 3, randomised clinical trial. *Lancet* **399**, 1779-1789 (2022).
68. J. Marrazzo, Lenacapavir for HIV-1 - Potential Promise of a Long-Acting Antiretroviral Drug. *N Engl J Med* **386**, 1848-1849 (2022).
69. M. E. Ritchie *et al.*, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
70. C. W. Law, Y. Chen, W. Shi, G. K. Smyth, voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29 (2014).
71. G. Yu, L. G. Wang, Q. Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383 (2015).
72. A. A. Sergushichev, An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv*, 060012 (2016).
73. A. Liberzon *et al.*, The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425 (2015).
74. M. Milacic *et al.*, The Reactome Pathway Knowledgebase 2024. *Nucleic Acids Res* **52**, D672-D678 (2024).
75. S. Heinz *et al.*, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**, 576-589 (2010).
76. S. Mahony, P. E. Auron, P. V. Benos, DNA familial binding profiles made easy: comparison of various motif alignment and clustering strategies. *PLoS Comput Biol* **3**, e61 (2007).
77. M. Bentsen *et al.*, ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nature communications* **11**, 4267 (2020).
78. J. Zou, J. Ernst, Chromatin state modeling across individuals reveals global patterns of histone modifications. *bioRxiv*, 2022.2008.2002.502571 (2022).

79. F. Grubert *et al.*, Genetic Control of Chromatin States in Humans Involves Local and Distal Chromosomal Interactions. *Cell* **162**, 1051-1065 (2015).
80. F. Ramirez *et al.*, deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165 (2016).
81. I. Rauluseviciute *et al.*, JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **52**, D174-D182 (2024).
82. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. (Springer Publishing Company, Incorporated, 2009), pp. 216.
83. D. Kaplan, R. Pruim, ggformula: Formula Interface to the Grammar of Graphics. *https://github.com/ProjectMOSAIC/ggformula*., (2023).

Acknowledgments

The authors declare that they have no competing interests. We thank all volunteers who participated in this study. We thank all members of the Schurr lab for their help and discussions. This work was supported by grant 487046 from the Canadian Institutes of Health Research (CIHR) to ES. This research was supported through a resource allocation in the Cedar high performance computing cluster by Compute Canada and WestGrid. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The raw sequencing datasets were deposited or appended to GEO with accession number GSE165709.