Executable models of immune signaling pathways in HIV-associated atherosclerosis

Mukta G. Palshikar[1], Rohith Palli[2], Alicia Tyrell[3], Sanjay Maggirwar[4], Giovanni Schifitto[5, 6], Meera V. Singh[5, 8,] and Juilee Thakar[1, 7, 8, 9**]

**Correspondence: Juilee_Thakar@URMC.Rochester.edu

[1] Biophysics, Structural, and Computational Biology Program, University of Rochester School of Medicine and Dentistry, Rochester, USA. [6] Department of Biostatistics and Computational Biology, University of Rochester School of Medicine and Dentistry, Rochester, USA. [7] Department of Microbiology and Immunology, University of Rochester School of Medicine and Dentistry, Rochester, USA. [8] Department of Biomedical Genetics, University of Rochester School of Medicine and Dentistry, Rochester, USA

Full list of author information is available at the end of the article

Abstract:

Atherosclerosis (AS)-associated cardiovascular disease is an important cause of mortality in an aging population of people living with HIV (PLWH). This elevated risk has been attributed to viral infection, anti-retroviral therapy, chronic inflammation, and lifestyle factors. However, rates at which PLWH develop AS vary even after controlling for length of infection, treatment duration, and for lifestyle factors. To investigate the molecular signaling underlying this variation, we sequenced 9368 peripheral blood mononuclear cells (PBMCs) from eight PLWH, four of whom have atherosclerosis (AS+). Additionally, a publicly available dataset of PBMCs from persons before and after HIV infection was used to investigate the effect of acute HIV infection. To characterize dysregulation of pathways rather than just measuring enrichment, we developed the single-cell Boolean Omics Network Invariant Time Analysis (scBONITA) algorithm.

scBONITA infers executable dynamic pathway models and performs perturbation analysis to identify high impact genes. These dynamic models are used for pathway analysis and to map sequenced cells to characteristic signaling states (attractor analysis). scBONITA revealed that lipid signaling regulates cell migration into the vascular endothelium in AS+ PLWH. Pathways implicated included AGE-RAGE and PI3K-AKT signaling in CD8+ T cells, and glucagon and cAMP signaling pathways in monocytes. Attractor analysis with scBONITA facilitated pathway-based characterization of cellular states in CD8+ T cells and monocytes. In this manner, we identify critical cell-type specific molecular mechanisms underlying HIV-associated atherosclerosis using a novel computational method.

Introduction

Chronic human immunodeficiency virus (HIV) infection increases the risk of atherosclerosis (AS) associated cardiovascular disease (CVD), which is a leading cause of morbidity in persons living with HIV (PLWH) (1-4). PLWH have an increased prevalence of risk factors for AS (5-9), some of which are driven by the off-target effects of antiretroviral drugs (9, 10). However, even after controlling for these risk factors, the risk of CVD remains significantly higher in PLWH (4, 11, 12), suggesting a role of long-term HIV infection.

HIV infection causes metabolic changes leading to a pro-atherogenic inflammatory environment in the vasculature (13-16). HIV infection and long-term antiretroviral therapy mediate signaling dynamics, including inflammasome activation, cell migration and apoptosis, in PBMCs and the vasculature (17, 18). Biomarker studies highlight these atherogenic processes, especially in the context of activated monocyte/macrophages and T cells (14-16). CD8+ T cells contribute to atherogenesis by secretion of cytotoxic granules and formation of the necrotic core of atherosclerotic plaques (19). Monocyte/macrophages migrate into the intima and form apoptotic atherosclerotic plaques (20). CD4+ T cells and B cells have also been previously implicated in

CVD (21-31). The interplay between signaling pathways, immune cell activation and inflammation in HIV infection requires further investigation. Single-cell sequencing (scRNA-seq) allows simultaneous investigation of these perturbations. To investigate the mechanistic link between HIV infection and atherosclerosis, we sequenced PBMCs from PLWH with and without atherosclerosis, matched for length of infection, cART, and other AS risk factors (Supplementary Note 1, Supplementary Figure 1).

Inference of mechanisms from high-dimensional scRNA-seq data is not trivial (32). Typically, scRNA-seq analysis uses clustering to define cell subpopulations, followed by differential expression (DE) and gene set overrepresentation analysis (ORA) to estimate pathway modulation. This approach discounts pathway topology and cannot connect molecular state to cellular state. ORA ignores synergistic interactions among genes by treating genes as independent and equal, and fails to estimate the significance of pathways (33). Our previously published algorithm overcomes above caveats by using discrete-state network modeling to perform pathway analysis using bulk transcriptomic data (34). This algorithm has been rigorously tested and compared to other pathway analysis methods (35). This algorithm employs discrete-state network modeling, which uses Boolean rules to explicitly define signal integration. These rules can be used to simulate dynamic trajectories and to perform *in-silico* perturbation experiments. We have extensively used discrete-state network modeling to investigate virus infections and have experimentally validated the predictions (36, 37). Here, we extensively expand our method and present single-cell Boolean Omics Network Invariant-Time Analysis (scBONITA) to (a) infer Boolean rules describing signal integration for pathway topologies using scRNA-seq data and (b) use these rules to identify dysregulated pathways and to prioritize genes/proteins for further investigation. scBONITA returns precise modes of dysregulation, captured by node-specific scores that quantify the contribution of each node to

the overall dysregulation of a pathway. scBONITA's capability to perform *in silico* simulation and perturbation of molecular pathways shows that it is a powerful hypothesis-generating tool.

Here, we describe the immune cell subpopulations and subpopulation-specific gene expression programs in a novel scRNA-seq dataset obtained from PLWH with and without atherosclerosis. We use scBONITA to identify dysregulated pathways in individual cell subpopulations, focusing on CD8+ T cells and monocytes. scBONITA highlights dysregulated cell migration and lipid metabolism pathways in several subpopulations and influential genes, such as PI3K and PLC, , which have high impacts on signal flow in these networks. Furthermore, we used a publicly available dataset of PBMCs from persons before and after HIV infection (38) to show that cell migration pathways are also dysregulated in the early stages of HIV infection, suggesting modulation of these pathways both by HIV infection and in subsequent AS. We present a novel method ('attractor analysis') that uses the models learned by scBONITA to identify pathway-specific signaling states for CD8+ T cells and monocytes. Additional in silico experiments show that scBONITA can learn a context-specific rule set from the vast possible state space for a Boolean network and that scBONITA's importance score provides novel information about signaling flow. Thus, this work provides insights into the mechanisms of HIV-associated atherosclerosis at the single-cell level by using a novel network modeling algorithm.

Results

PBMC subpopulations in AS+ and AS- PLWH

To investigate dysregulated immune signaling in People Living with HIV (PLWH), we recruited a cohort of eight PLWH, four with atherosclerosis (AS+) and four without atherosclerosis (AS-). Participants were matched for known atherosclerosis risk factors (see Methods, Supplementary Note 1, and Supplementary Figure 1). We transcriptionally profiled ~1200 peripheral blood mononuclear cells (PBMCs) per subject. This data was processed using the Cell Ranger and

Seurat pipelines (39) to identify sixteen subpopulations of immune cells (Methods, Figure 1A) annotated using CIBERSORT (40) (Supplementary Figure 2) and cell-lineage specific markers (Supplementary Table 1). This scRNA-seq dataset is referred to in the text as the HIV/AS dataset.

A population of CD8 T cells/NK resting cells was lower in AS- PLWH and a population of CD14+CD16+ monocytes was higher in AS- PLWH (t-test, $p < 0.05$) (Figure 1B). Cell subpopulation markers were identified using MAST (41) as described in the Methods. The CD14+CD16+ monocytes and CD8+ T cells/NK resting cell markers were enriched for migration-related pathways (Supplementary Table 1). Indeed, these cells are known to migrate into the intima during the formation of atherosclerotic lesions in the vascular wall (42-48).

In addition, we evaluated the proportions of the cell subpopulations identified in scRNA-seq data in a bulk RNA-seq dataset obtained from an independent cohort of matched PLWH with and without AS (49). This dataset was deconvoluted using CIBERSORT (40, 50) to quantify the abundance of the cell subpopulations found in the scRNA-seq dataset (Supplementary Figure 3). We found that most clusters were not significantly different between AS+ and AS- groups, as observed in scRNA-seq data. The subpopulation ‘T cells CD8 NK cells resting’ was more abundant in AS- PLWH (t-test, $p < 0.1$), while there was no significant difference in the abundance of monocytes. As these differences in the abundance of cell subpopulations were not robustly recapitulated in the independent cohort, here we focus on differentially regulated molecular mechanisms between AS+ and AS- PLWH.

Atherosclerosis-associated gene-expression across PBMC subpopulations

Differentially Expressed (DE) genes between AS+ and AS- PLWH (Figure 1C-F and Supplementary Table 2) include upregulation of MHC Class I and Class II genes in multiple cell

subpopulations. ITGB2, which is upregulated in CD8+ T cells from AS+ PLWH, is involved in leukocyte transendothelial migration (Supplementary Table 2). ACTB (β-actin) was upregulated in naïve B cells and CD8+ T cells from AS+ PLWH (Supplementary Table 2). CXCR4 was upregulated in cells from AS+ PLWH in naïve B cells, CD8+ T cells, and a population of resting NK cells. Both CXCR4 and ACTB modulate dynamic actin cytoskeleton remodeling in transendothelial migration. Gene set ORA of DE genes revealed enrichment of cell migration and mobility functions in cells from AS+ PLWH (Supplementary Table 2).

In AS- PLWH, S100A8/S100A9 are upregulated in three cell subpopulations (Figure 1C-F, Supplementary Table 2). S100B is upregulated in CD8+ T cells populations from AS- subjects (Figure 1C, 1F, Supplementary Table 2). Several ribosomal genes were DE in twelve subpopulations; out of which they were only upregulated in naïve B cells from AS- PLWH and naïve T cells from AS+ PLWH (Supplementary Table 2).

While DE and enrichment analysis indicated some mechanisms of HIV-associated atherosclerosis, an integrated model of how these genes jointly regulate signaling cascades did not emerge. Hence, we developed the network-based pathway analysis algorithm scBONITA (Single-Cell Boolean Omics Network Invariant-Time Analysis) to investigate signal integration and flow.

scBONITA learns discrete-state models of signaling pathways

scBONITA is a discrete-state modeling approach to develop executable models of dysregulated immune signaling pathways (35) (Figure 2). The scBONITA algorithm requires two inputs: (a) a scRNA-seq dataset, and (b) a prior knowledge network (PKN) (Figure 2A). In these PKNs, genes and their regulatory interactions are represented by nodes and directed edges respectively. scBONITA leverages the principle that observed states of single cells correspond to states of dynamic biological networks to identify regulatory rules (Figure 2B). A genetic

algorithm is used to identify a minimum-error rule set that is optimized by a node-wise local-search. This returns a set of discrete-state models for pathways as multiple rule sets explain the training data equally well (Equivalent rule set, ERS). A pathway is described as 'optimized' if scBONITA-RD successfully reduced the state space of the possible rules for at least one node in the pathway.

scBONITA models can be simulated to generate dynamic trajectories and *in silico* node perturbations. These models are used to perform node knock-out and knock-in simulations and the difference between network states after these simulations is weighted by the size of the ERS to calculate a node importance score. Thus, node importance score measures the influence of a node over the network and the weight incorporates the uncertainty in rule determination. scBONITA combines these importance scores and comparison-specific fold changes to calculate a pathway modulation score (Methods, Figure 2C).

The simulation trajectories of these discrete-state models fall into steady states known as attractors, which have been hypothesized to correspond to signaling behavior characteristic of specific cell types (51-56). Cells are assigned to the attractor closest to their expression (Figure 2D) to characterize signaling states for a network. In conclusion, scBONITA allows in depth comprehension of signaling pathways by incorporating network topology.

Dysregulated pathways in T cell populations linked to HIV-associated atherosclerosis

The scBONITA pathway analysis algorithm identifies dysregulated pathways in all subpopulations derived from AS+ and AS- PLWH, providing insights into mechanisms of HIV-associated atherosclerosis (Supplementary Table 3). In CD8+ T cells, these pathways included pro-inflammatory, anti-viral, cell migration and apoptosis pathways (Figure 3A, Supplementary Table 3). all of which were downregulated in AS+ PLWH except the Th17 cell differentiation

pathway, which includes genes involved in generic T cell differentiation. All these pathways were identified by scBONITA, but not by enrichr (Supplementary Table 2).

scBONITA identified multiple optimized pathways in CD4+ T cells (subpopulations represented in Figure 1A) as being dysregulated ($p_{adj} < 0.01$) in AS+ PLWH (Figure 3B, Supplementary Table 3). CD4+ T cells may exert either an atherogenic or atheroprotective phenotype, depending on subset and interactions with antigen presenting cells in the adventitia or plaques (57). Our results suggest that CD4+ and CD8+ T cells from AS+ PLWH play a role in cell adhesion, apoptosis and migration processes involved in atherosclerosis, and these atherogenic processes are mediated by the upregulated PI3K-AKT, mTOR and cytoskeletal signaling pathways.

The AGE-RAGE signaling pathway was further investigated in CD8+ T cells because of its role in dysregulated lipid metabolism (19, 23, 58-63) and to demonstrate the additional information obtained from scBONITA in comparison to ORA. This pathway had the highest pathway modulation score (0.8) (Supplementary Table 3). Most genes in this pathway were upregulated in AS+ PLWH (Figure 3C). scBONITA optimized rules for DIAPH1 (uncertainty score = 0.5, see Methods) (Figure 3C), which has a strong influence over the signal flow through the network due to its high centrality. The combination of scBONITA's node importance score and fold change between subject groups was used to identify genes whose activity influences signal flow in AS+ PLWH. These genes include: the class 1 PI3K genes, the P13K regulator PI3KR1, and PLC genes. All of them are highly expressed in AS+ PLWH (Figure 3C). PI3K activates lipid metabolism, macrophage autophagy, phenotypic transition, and the expression of adhesion molecules (reviewed in (64)). In this manner, scBONITA identified dysregulated pathways and genes associated with atherosclerosis-linked migration of T cells derived from PLWH.

Dysregulated pathways in monocytes linked to HIV-associated atherosclerosis

In monocytes, scBONITA identified several dysregulated pathways (Figure 4A, Supplementary Table 3) which are known to be involved in the pro-inflammatory behavior of pro-atherogenic monocytes (43, 65-74). Of these pathways, only cAMP signaling and endocrine resistance are overall upregulated in cells from AS+ PLWH, suggesting that monocytes from AS- are pro-atherogenic.

The leukocyte transendothelial migration pathway was further investigated as monocyte migration outside the vascular compartment plays a crucial role in the inflammatory cascade that leads to an atherosclerotic phenotype (65, 67, 75). The pathway modulation score of 0.45 was the third highest amongst tested pathways (Supplementary Table 3) (Figure 4B). scBONITA optimizes regulatory rules for the influential RHOA gene (uncertainty factor=0.13). High importance scores were assigned to the NCF genes, PLCG genes, MSN, ROCK, CYBA and CYBB. NCF genes are involved in superoxide production and are a positive regulator of P13K signaling (76, 77). The upstream regulator of PLCG1, MSN, is involved in cytoskeletal remodeling during leukocyte migration (78). ROCK2 shows a small change across AS groups, which may be driven by feedback regulation, but has a stronger role in regulating signal flow, as shown by its high importance score. The high importance of the G protein GNAI3, downstream of CXCR4 indicates a role for CXCR4-mediated activation of this pathway. GNAI3 is also present in two other significantly dysregulated networks in monocytes – Cushing Syndrome and cAMP signaling – but does not have a high importance score in those networks, indicating that it is critical in regulating signal coming from CXCR4 only in the leukocyte transendothelial migration network. The downstream effectors of these genes have higher fold changes, possibly due to feedback regulation. These effectors include ACTG1 and EZR, involved in cytoskeletal remodeling (78-80), and ITGA4, ITGB1, and ITGB2, involved in cell adhesion. scBONITA thus

identifies several genes and pathways associated with dysregulated transendothelial migration in the context of HIV-associated atherosclerosis.

Pathways dysregulated by HIV infection and in HIV-associated atherosclerosis

To identify the biological mechanisms modulated by HIV infection that may also contribute to the elevated risk of AS in PLWH, we used a publicly available scRNA-seq dataset of PBMCs from persons before and during acute HIV infection (38). Kazer et al (38) identified gene expression programs in activated T cells, monocytes, and NK cells during HIV infection. We used this dataset with scBONITA to infer Boolean rules for KEGG networks and perform pathway analysis. The results were compared to the pathways dysregulated in HIV-associated atherosclerosis in the above-described HIV/AS dataset.

scBONITA identified 10 optimized pathways dysregulated after 1 year of HIV infection in cytotoxic T cells (38) (Figure 3A-B & 5A, Supplementary Figure 6, Supplementary Tables 5 & 6). Of these pathways, five signaling pathways were dysregulated in AS+ PLWH. Similarly, 5 out of 19 optimized pathways were dysregulated in monocytes upon HIV infection and in AS+ individuals (Figures 4A & 6A, Supplementary Figure 6, Supplementary Table 6). 81 and 41 genes from these overlapping pathways were upregulated both upon HIV infection and in AS+ PLWH in the cytotoxic T cell populations and monocyte population respectively (Supplementary Figure 7, statistical significance was not tested as scBONITA does not depend on DE genes to perform pathway analysis). The genes upregulated after HIV infection and in AS+ PLWH in the CD8+ T cell subpopulation were involved in viral response pathways (Supplementary Figure 7). Similarly, the genes upregulated after HIV infection and in AS+ PLWH in the monocyte subpopulation are involved in cell migration related pathways (Supplementary Figure 7). This

suggests that the modulation of cell migration and inflammation processes upon HIV infection progresses over time, increasing risk of AS in PLWH.

Pathway-specific cellular signaling states associated with atherosclerosis

We used the network models learned by scBONITA to identify attractors for the pathways dysregulated between AS+ and AS- PLWH in CD8+ T cells discussed above (Cluster CD8 T cells -1 in Figure 1A) and evaluated cellular states across subjects and disease groups. The simplest rules, which have the smallest number of “AND” terms, were chosen to simulate the network and identify attractors (Methods). The insulin resistance pathway, which is downstream of the AGE-RAGE and PI3K-AKT signaling pathways, was particularly interesting in CD8+ T cells (Cluster CD8 T cells -1 in Figure 1A). 72 signaling states were identified by network simulation, of which 3 dominant signaling states mapped to 16.5%, 7.5% and 27.4% of cells respectively (Figure 5A-B). There is an association between attractors and subjects (chi-square test, p-value < 0.05) but no association between attractors and atherosclerosis status (chi-square test, p-value > 0.05). Additionally, we also performed attractor analysis for other cell clusters. Among those, interestingly, assigned chemokine signaling attractors of CD4+ T cells were subject-specific (chi-square test, $p < 0.01$). Similarly, we observed subject-specific attractors of the PI3K-AKT signaling pathway for the T cells CD8/CD4/CD4 naive cluster (chi-square test, $p < 0.01$).

The CD8+ T cell states for the insulin resistance pathway were characterized by differences in several key genes (Figure 5C), including PI3K genes and the PI3K regulators that were identified as being influential in the AGE-RAGE signaling pathway (Figure 3B). In addition, the two less abundant attractors differed in the activity of the key TNFR and TNF genes. Hence, these attractors are referred to as the PIK3R+ PI3K+, TNFR+TNF- and TNFR-TNF+ attractors.

The activity of PI3K and AKT genes was higher in the most common signaling state (PI3KR+ PI3K+ attractor). However, the activity of the downstream targets of AKT, such as the CREB genes, NFKB1, FOXO1, were lower in the PI3KR+ PI3K+ attractor. TNF, which is produced at a low level by some subsets of T cells, also mediates a range of pro-inflammatory processes in vascular endothelial cells (81, 82). TNF-TNFR1 signaling mediates an apoptotic process mediated by TRADD and FADD (83). TNFR1 activates PI3K signaling in regulatory T cells (84). Thus, our attractor analysis reveals different T CD8+ cell states. This analysis also indicates differences in TNF production and response in the cells that are assigned to these cell states and suggests the existence of distinct modes of operation across subjects (Figure 5D).

Attractor analysis of the leukocyte transendothelial migration pathway in monocytes revealed 9 attractors that mapped to cells in the dataset. The two dominant signaling modes differed in the activity of the PECAM1 and F11R genes respectively (Figure 6A-C). The attractors are referred to as the PECAM+ and F11R+ attractors and mapped to 51.04% and 30.73 % of cells respectively. F11R is required for platelet adhesion to vascular endothelial cells (85), which occurs prior to infiltration of monocytes into the endothelium and eventual plaque formation (86, 87). PECAM1 has widespread effects on vascular biology and atherosclerosis in particular (88-90). These subject specific attractors (chi-square test, p-value < 0.05, Figure 6B) were not associated with atherosclerosis status. Similarly, the attractor activity of individual genes was not associated with atherosclerosis status (two-sided t-test, p-value > 0.05). Thus, scBONITA identifies two cellular states of monocytes driven by molecular signaling. These states are associated with monocyte migration into the vasculature and reflect inter-subject variability.

Evaluation of scBONITA performance in silico

The BONITA algorithm has already been rigorously validated in our previous study (35). Specifically, comparison with other pathway analysis tools has been performed. Here we

evaluate scRNA-seq specific components of the algorithm. To show that scBONITA rule determination is robust to training set size, we varied the size of the training data provided to scBONITA. The number of cells in the largest cluster of cells (Naïve B cells -1) were varied by random selection from 1% of cells from that cluster to 200% by adding cells from neighboring clusters. The reduced size of the ERS for nodes with in-degree 3 (i.e., the most complex case considered by scBONITA) is a metric for certainty in rule inference by scBONITA (Figure 7A). While there was a significant decline in performance when the data was downsampled to 1% of the original cluster, there was no significant increase in effect once 50% of the cells were used, or when the training dataset was augmented. This indicates that scBONITA is robust to heterogeneity in the training data set.

The node importance scores calculated by scBONITA for KEGG networks trained on the HIV/AS dataset as described above are not correlated to six centrality measures (Figure 7D).We found that the node importance scores between comparable cell subpopulations in the Kazer et al and HIV/AS datasets investigated here were correlated, as shown by a representative comparison between cytotoxic T cells from the Kazer et al dataset and the subpopulations of CD8+ T cells from the HIV/AS dataset (Figure 7C, $0.71 <$ Pearson Correlation Coefficient < 0.91, $p < 0.01$). Similarly, the node importance scores for the populations of monocytes were highly correlated (Pearson correlation coefficient = 0.78, Supplementary Table 4, Supplementary Figure 4). The correlations were lower for other pairs of subpopulations (Supplementary Table 4, Supplementary Figure 4). This indicates that scBONITA learns some characteristic features of network topologies, but node importance scores are still assigned in a context-dependent manner.

Discussion

Among people living with HIV, widespread use of cART has significantly reduced overall mortality. However, the earlier and increased incidence of cardiovascular diseases remains the major cause of mortality in an aging HIV+ population for multiple intersecting reasons (4, 9-12). We and others have attempted to identify the immune signaling mechanisms that lead to this increased incidence of atherosclerosis (49). However, to the best of our knowledge no study has systematically investigated cell-type specific signaling dysregulations at single-cell level. To this end, we sequenced, for the first time, PBMCs from four AS+ and four AS- PLWH. The cohort was closely matched for known AS risk factors. These risk factors will be harder to match exactly in a larger cohort. Although a small sample size in some contexts can limit the power of the study, previous scRNA-seq studies have provided insight into both changes in PBMC composition and signaling with cohort sizes similar to this study in the context of HIV, because of the depth of information on molecular mechanisms that can be obtained from scRNA-seq experiments (38, 91-94). To increase statistical power, we evaluated the proportion of the subpopulations identified from scRNA-seq in a bulk RNA-seq dataset from an independent cohort of PLWH with and without AS (49) (Supplementary Figure 3). Enrollment of more subjects in the future will help further elucidate the signaling dysregulations leading to HIV-associated atherosclerosis. We found that in accordance with previous studies (19, 58-62, 95, 96), a population of CD8+ T cells was increased in PBMCs from AS+ PLWH. Contrary to our expectations (20, 66, 67, 72, 74, 97-100), a population of monocytes was decreased in PBMCs from AS+ PLWH. Here we focus on cell-type specific molecular changes associated with atherosclerosis.

We observed that differentially expressed genes between AS+ and AS- were involved in cell migration; however relevant pathways were not identified using over-representation analysis (ORA). This demonstrated the inefficacy of conventional ORA methods in identifying coherent

molecular processes. For example, CXCR4, a DE gene, leads to the activation of phosphatidylinositol-3-OH kinases (PI3K), which in turn activates the serine-threonine kinase AKT via PIP3 (101). PI3K/AKT signaling leads to processes involved in plaque formation, such as cell migration, intracellular lipid accumulation, and smooth muscle cell proliferation (64). None of these processes were enriched in the DE genes. This may be because of the technical limitations of scRNA-seq leading to nonspecific distortions of expression and identification of fewer differences across conditions (102, 103). Even DE analysis methods that are sensitive to the known characteristic distributions of scRNA-seq data are prone to false discoveries (104). These methods failed to provide insights into how disparate genes involved in different pathways regulate cellular states. To mechanistically characterize signaling dysregulations in HIV-associated atherosclerosis, we developed the scBONITA algorithms for regulatory rule inference, network simulation, pathway analysis, and attractor/steady-state analysis. As a gene is evaluated in the context of its complete signaling pathway, our approach minimizes the impact of the caveats in scRNA-seq technology described in (102, 103).

scBONITA learns condition-specific logic models using scRNA-seq data in conjunction with published prior knowledge networks (PKNs). This study builds on our published BONITA method (35) that inferred logic rules from bulk RNAseq data. scBONITA exploits the bimodal nature of scRNA-seq data (105, 106) and the cell-level resolution of expression to successfully learn regulatory rules and identify attractors for PKNs. These rules can be used to perturb and simulate pathways *in silico*. Unlike other algorithms to infer logic rules and reconstruct gene-regulatory networks on a small subset of genes from scRNA-seq data (107-109), scBONITA does not pre-select genes. In addition, scBONITA is not dependent on time-series data, a significant advantage since time-series data is rarely available in human studies. In lieu of using time-series data, scBONITA hypothesizes that scRNA-seq data represents samples in the state space of a dynamic Boolean network. In addition, the use of known network topologies reduces

the uncertainty of inferred rules. Indeed, we show that scBONITA can successfully restrict the state space of possible rules. scBONITA combines expression information with the importance score to create a unique metric of pathway dysregulation. scBONITA node scores can be directly translated into empirical perturbation studies.

In CD8+ T cells we identified AGE-RAGE signaling, which elicits activation of multiple intracellular signaling pathways such as cell proliferation and apoptosis pathways (Figure 3A) (110-118). The PI3K family of genes, which had the highest importance score and were upregulated in AS+ PLWH, promote intracellular lipid deposition leading to the formation of foam cells and atherosclerotic plaques and can reduce the expression of lipid transporters and reduce the efflux of intracellular cholesterol depending on upstream signals (64). PLC, which also has a high importance score, promotes leukocyte adhesion, VEC apoptosis and plaque development induced by oxidized low-density lipids (78, 119).

In monocytes, dysregulation of lipid-metabolism pathways such as cAMP signaling and leukocyte transendothelial migration suggest the infiltration of monocytes into the intima during the formation of atherosclerotic lesions and progression of atherosclerosis (42-47). scBONITA identified genes critical to the atherosclerotic process in the leukocyte transendothelial migration pathway. ROCK activation by atherogenic stimuli such as oxidized LDLs leads to pathophysiological changes include endothelial dysfunction and vascular remodeling (120, 121). ROCK inhibitors such as statins attenuate atherosclerosis by inhibiting chemotaxis of macrophages and their transformation into foam cells (122). The dysregulation of glucose metabolism pathways in AS+ PLWH can induce expression of adhesion molecules by VECs and increased monocyte transendothelial migration (64, 123). Thus, here we find migratory and lipid-metabolism pathways and key genes driving atherosclerosis in PLWH.

PI3K-AKT signaling is dysregulated in all cell subpopulations (Figures 4 – 5, Supplementary Table 3). The activation and effector mechanisms of this cascade vary by cell type, as shown by

the different dysregulation of linked signaling pathways in different cell types. For example, in two populations of naïve B cells derived from AS+ PLWH, the apelin signaling pathway was upregulated in one and the adipocytokine signaling pathway was downregulated in the other. The cardioprotective effect of apelin is modulated by (amongst other routes) the PI3K-AKT signaling and MAPK signaling pathways (80-82), which are also dysregulated in these naïve B cells. Apelin is also shown to be upregulated in human atherosclerotic coronary arteries and colocalized with markers for macrophages (124, 125).

We find that HIV infection dysregulates several pathways that are further impacted in PLWH with AS (Figure 3A & 4A, Supplementary Figure 6, and Supplementary Table 6). These pathways are suggestive of changes occurring in cell migration of cytotoxic T cells and monocytes (126-129). The overlap also suggests that cAMP and PI3K-AKT signaling is affected during the course of HIV infection.

To map cells to distinct signaling modes of the pathways described above, we developed scBONITA's attractor analysis functionalities. Attractors are regions in the state space of a dynamic system towards which simulation trajectories are "pulled" and are characteristics of a specific network with a specific set of regulatory rules. These attractors may correspond to observable cell states, or hallmarks of specific phenotypes such as cell type differentiation, disease state, or drug treatment (51-56). These studies show that even simple dynamic models capture rich and nuanced cell behaviors. scRNA-seq allows the study of these dynamic landscapes and their attractors at an unprecedented resolution (56, 130, 131). Attractor analysis with scBONITA allows users to characterize cells based on the dynamic properties of signaling networks, which dictate phenotype. scBONITA identifies these attractors and their master regulators that control the changes between these cell states, providing complex insights into cellular processes.

The importance of cell migration and lipid signaling in the development of HIV-associated atherosclerosis was underscored by attractor analysis in CD8+ T cells and monocytes. In most cases, only one dominant signaling state existed in the cell subpopulation. However, the three dominant signaling modes of the insulin resistance pathway in CD8+ T cells differed in the activity of PI3K and AKT genes and their downstream effectors, such as CREB and FOXO1. This suggests the existence of two distinct modes for this signaling pathway corresponding to a proliferative cell state (activation of PI3K and AKT) and a senescent cell state (transcription of CREB- and FOXO1-controlled genes) (132-134). The insulin signaling pathway exerts immunomodulatory effects on T cells (135). Similarly, we identified two dominant signaling modes (PECAM+ and F11R+) for the leukocyte transendothelial migration pathway in monocytes, suggesting variation in the cell states with respect to this pathway in PLWH. Dysregulation in insulin signaling, which is a risk factor for CVD promotes PECAM1-mediated migration of monocytes into the endothelium (136, 137). PECAM1's loss contributes to atherosclerosis (90). PECAM1+ cells may contribute to the suppression of inflammatory processes driving atherosclerosis. Thus, our analysis connects molecular processes to cellular states, unlike conventional DE and ORA analyses.

Although the BONITA algorithm has been rigorously validated in our prior publication (138) here we wanted to evaluate the scRNA-seq specific parts of the algorithm. We demonstrated that scBONITA can identify characteristic structural properties of networks and use this in conjunction with expression information to identify dysregulated pathways in a specified condition (Figure 7). Moreover, scBONITA is minimally affected by the heterogeneity of training data and can narrow down the vast state space for a Boolean network (Figure 7A). We expect that scBONITA inference will improve when pure cell populations are sequenced. While scBONITA is not strictly dependent on the clustering method used to classify scRNA-seq data into subpopulations, we used pre-classified subpopulations to reduce variability and to improve

the specificity of scBONITA-PA. Additionally, scBONITA-RD requires a longer runtime (<12 hours in our tests) and more powerful computational capabilities than a typical processing and clustering analysis pipeline run on datasets of typical size. These resources are usually available to academic users on computing clusters. In conclusion, we present a novel dynamic network modeling method that yields mechanistic insights into the cellular and immunological processes involved in HIV-associated atherosclerosis.

Materials and Methods

Participant cohort summary, sample collection, and storage

Eight men living with HIV and >= 50 years of age on stable combined antiretroviral therapy (cART) for at least 1 year and with viral load <= 50 copies/mL were recruited. All methods were carried out in accordance with University of Rochester guidelines and regulations, and all experimental and study protocols were approved by the University of Rochester Institutional Review Board (#RSRB00063845). Informed consent was obtained from all subjects. Individuals were classified as having atherosclerosis (AS+) if they had plaques on the carotid arteries on ultrasound imaging. Four of the eight subjects were assigned as AS+ and had plaques in both right and left carotid arteries. AS- subjects were aged between 47 and 57 and AS+ subjects were aged between 51 and 66. AS+ subjects had mean serum cholesterol of 161.5 mg/dl (σ = 40.9) and mean serum high-density lipid HDL of 54.7 mg/dl (σ = 16.3). AS- subjects had mean serum cholesterol of 167.7 mg/dl (σ =57.2) and mean serum high-density lipid HDL of 51 mg/dl (σ = 7.7). AS- subjects and AS+ subjects had a mean CD4+ T cell count of 518.5 cells/µl (σ =347.8 cells/µl) and 838.7 cells/µl (σ = 514.5 cells/µl) respectively. De-identified subject information is available in Supplementary Note 1 and Supplementary Figure 1. Subjects were matched for lipid profiles, hypertension status, smoking status, CD4+ T cell counts, and age. In addition, all subjects were treated with cART for at least one year. 30 mls of blood per study

participant was collected in ACD vacutainers and was processed within 2 - 3 hours of collection. Peripheral Blood Mononuclear Cells (PBMCs) were isolated using Ficoll density gradient centrifugation. 5 million PBMCs were preserved using RNAlater (Thermo Fisher) and were used for scRNA-seq.

Single-cell RNA sequencing and data processing

Frozen vials containing cells in RNAlater were thawed quickly in a 37-degree water bath. Cell suspension was transferred to a 15ml conical tube. 10 ml PBS/2% FBS was slowly added. Samples were centrifuged at 1600rpm for 6 min. Washes were repeated for an additional 2 times for a total of 3 washes. Using the MACS Miltenyi Biotec Dead Cell removal kit (PN130-090-101), dead cells were removed using manufacturer's recommendations. Cells were counted and cellular suspensions were loaded on a Chromium Single-Cell Instrument (10x Genomics, Pleasanton, CA, USA) to generate single-cell Gel Bead-in-Emulsions (GEMs). ScRNA-seq libraries were prepared using Chromium Single-Cell 3' Library & Gel Bead Kit (10x Genomics). The beads were dissolved, and cells were lysed per manufacturer's recommendations. GEM reverse transcription (GEM-RT) was performed to produce a barcoded, full-length cDNA from poly-adenylated mRNA. After incubation, GEMs were broken, and the pooled post-GEM-RT reaction mixtures were recovered, and cDNA was purified with silane magnetic beads (DynaBeads MyOne Silane Beads, PN37002D, ThermoFisher Scientific). The entire purified post GEM-RT product was amplified by PCR. This amplification reaction generated sufficient material to construct a 3' cDNA library. Enzymatic fragmentation and size selection was used to optimize the cDNA amplicon size and indexed sequencing libraries were constructed by End Repair, A-tailing, Adaptor Ligation, and PCR. Final libraries contain the P5 and P7 priming sites used in Illumina bridge amplification. Sequence data was generated using Illumina's NovaSeq 6000. Approximately 2000 cells were sequenced from each subject. Cell Ranger (version 2.1.1; 10x Genomics) was used for demultiplexing and alignment with default parameters. Reads were

aligned to the human reference genome GRCh38 (Ensembl 93). The Seurat R package (version 2.3.4) (39) was used to further process the gene counts obtained from the CellRanger pipeline. Cells that express < 200 genes, > 2500 genes, or > 5% mitochondrial genes were filtered out. Genes expressed in < 3 cells were filtered out. Gene counts were per-cell normalized and $\log_2$-transformed. These preliminary filtering and selection procedures yielded a set of 9368 sequenced cells, approximately equally distributed between subjects (and hence conditions), and 14017 genes. Note that sample collection, processing and sequencing were performed in one batch, leading to extremely high-quality data where no subject specific patterns were observed.

Classification into subpopulations using modularity-optimized Louvain community detection, and cluster labeling

Cells were classified into subpopulations using modularity optimized community detection, implemented in the Seurat R package (39). 664 highly variable genes were used to identify 10 principal components that explained the majority of variance in the data. These principal components were used to cluster the data. Clustering yielded 16 subpopulations. Cluster markers were identified using MAST (105). As suggested in (139), CIBERSORT (40) was used to "deconvolute" the average gene expression of each cluster into the constituent canonical cell types. A reference expression set of 22 immune cell types and 547 genes was used (40). Over-representation analysis was performed using the implementation of the hypergeometric test in the R package clusterprofiler (version 3.12.0) with Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets downloaded from MSigDb (140-142). Gene sets were identified as significantly over-represented if the Bonferroni-adjusted p-value was < 0.05.

scBONITA algorithm for development of discrete-state models of pathways

Network topologies: ScBONITA infers Boolean regulatory rules/ logic gates for directed networks wherein nodes represent genes and edges represent the regulatory relationships between those genes. These networks contain edge annotations denoting activation/inhibition relationships between nodes, which are exploited by scBONITA to restrict the search space for rule inference to sign-compatible canalyzing functions. Such network topologies of biological pathways are commonly obtained from pathway databases such as KEGG and WikiPathways (142-144). ScBONITA offers an interface to KEGG and WikiPathways databases that allows automated download and processing of user-specified networks. Users can also provide custom networks in graphml format.

Boolean rule determination from scRNAseq data: scBONITA assumes that cross-sectional measurements of cells by scRNA-seq data represent steady state of an underlying dynamic biological process. scBONITA's rule determination (scBONITA-RD) algorithm, which has been extended from our previous BONITA algorithm exploits this property to infer Boolean rules for an input biological network, using a combination of a genetic algorithm (GA) and a node-wise local search (138).

The global search using GA infer a single candidate rule set that adequately describes the input data with respect to the network topology with minimum error (35, 145). The function to be minimized is:

$$\sum_{c=1}^{cells} min\,(\sum_{n=1}^{nodes} |E_{c,n} - A_{c,n,a}|\ \forall\, a\ in\ T_c\ )$$

Where *c* from *1* to *cells* iterates over the number of cells in the training dataset, n iterates from 1 to number of *nodes* in the network, $E_{c,n}$ is the binarized expression of node *n* in cell *c*, $A_{c,n,a}$ is the value of node *n* in the attractor *a* reachable from cell *c*, and T_c is the attractor reachable from *c*. Note that T_c may have multiple repeating states in a limit cycle or only one steady state, i.e., it may be a singleton attractor. T_c is obtained after simulating the network with the candidate rule set for 100 time-steps, which causes the simulation to reach an attractor state for all tested networks.

The minimum error rule set identified using the above-described genetic algorithm strategy is further refined by a node-level local search that sequentially optimizes the rule for each node keeping the rules for all other nodes in the network constant. An optimal set of rules for a node *n* is obtained by minimizing the function

$$\sum_{c=1}^{cells} min\ (|E_{c,n}\ -\ A_{c,n,a}|\ \forall\ a\ in\ T_c)$$

where variables and constants are same as described above.

Several rules may satisfy the termination criteria. The local search returns a set of equivalent rules that all satisfactorily explain the observed state in the experimental data. This set of rules is referred to as the equivalent rule set (ERS) in the text.

Pathway analysis (PA) with scBONITA: scBONITA performs pathway analysis in a two-step process. In the first step, importance scores for each node in the biological network under consideration are calculated. In the second step, a pathway modulation metric incorporating both experiment-specific fold changes and the node importance scores calculated in step 1 is calculated.

scBONITA quantifies the influence I_n of a node n over the state of the network by quantifying the overall effect of its perturbation on that network. This is achieved by simulating knock-in and knock-out of that node.

$$I_n = \sum_{c=1}^{cells} |KI_{c,n} - KO_{c,n}| * Uncertainty\ Factor$$

where $KI_{c,n}$ and $KO_{c,n}$ are the discrete expression vectors of network node *n* in the attractors reached after a simulation starting from cell *c* where the node under consideration *n* is knocked in and knocked out respectively. The uncertainty factor is defined as follows:

$$\text{Uncertainty factor} = \frac{|Maximum\ ERS_i| - |Observed\ ERS_i| + 1}{|Maximum\ ERS_i|}$$

Where ERS_i is the ERS for a node *i*, $|Maximum\ ERS_i|$ is the maximum possible size of the ERS for a node *i* and $|Observed\ ERS_i|$ is the size of the ERS for a node *i* upon optimization by scBONITA.

The uncertainty factor weighs I_n relative to the maximum state space for that node, to capture the uncertainty in the rule determination for that node. The importance scores of the nodes in a network are scaled to *[0, 1]* by dividing by the maximum calculated importance score for the network under consideration.

A pathway modulation metric (M_p) is calculated by weighting the node importance score by the difference between the average gene expression in each group (relative abundance, *RA*) and the standard deviation of expression of that gene (*σ*) across cells. A p-value is calculated by bootstrapping, where a contrast-specific distribution of weighted importance scores is generated using randomly resampled *RA* values. Pathways are described in the text as being overall upregulated in a given contrast if the sum of fold changes of all genes in the pathway is positive.

Conversely, pathways are described as being downregulated if the sum of fold changes of all genes in the pathway is negative.

$$M_p = \sum_{n=1}^{nodes} RA_n * \sigma_n * I_n$$

Steady-state analysis with scBONITA: scBONITA assumes that the observed cellular states are defined by states of multiple dynamic cellular processes or signaling pathways. While observed cells are samples along a dynamic trajectory of signaling cascades, analyzing attractors upon randomly sampling the rules from ERS allows us to investigate most common signaling states of a network under consideration. We sample ten network specific from the ERS inferred by scBONITA-RD to identify a set of reachable attractors. This is achieved by simulating the network synchronously as performed in other studies (107, 146-148) starting from an observed state (i.e., a cell expression vector) until a steady state (or an attractor cycle) is reached. By starting simulations from observed expression levels of all cells (i.e., all observed states), we can ensure that these simulations cover a large fraction of available state space for a given network. In this way, all reachable attractor states, corresponding to observable signaling states, can be identified. The similarity between cells and attractors is quantified using the Hamming distance. Cells are assigned to the attractor that is closest to their expression data.

Implementation and availability: scBONITA is implemented in Python 3 and C. Source code, documentation, and tutorials are available onhttps://github.com/Thakar-Lab/scBONITA.

Application of scBONITA on a publicly available data set

A scRNA-seq dataset obtained from four persons living with HIV (PLWH) before and during infection was selected to demonstrate the utility of the scBONITA pipeline on other datasets and

to compare signaling dysregulations upon atherosclerosis in PLWH to signaling dysregulations upon HIV infection (38). Log2-transformed TPM data and metadata processed and curated by the study authors was collected from the Single-Cell Portal database (https: //singlecell.broadinstitute.org/single_cell/study/SCP256.). The complete scBONITA pipeline was used to compare samples collected before infection to samples collected 1 year after infection. We retained the cluster labels assigned by the authors of the original study. A set of 210 KEGG networks was used with the scBONITA pipeline.

In silico evaluation of scBONITA

To show that scBONITA-RD is robust to training set size, we selected a cluster of B cells from the HIV/AS dataset. This subset of the dataset was manipulated to either downsample or augment the size of the training dataset (number of cells) presented to scBONITA-RD. The training dataset was downsampled to 1% and 50% of the original number of cells for cluster 0 (“B cells naïve – 1”). To augment the dataset and thereby introduce heterogeneity, the dataset was increased to 200% of its original size by adding in cells from a neighboring cluster of B cells. A set of 210 KEGG networks was used to evaluate the sizes of the ERS obtained by scBONITA-RD using these manipulated training datasets. The size of the ERS is used as a proxy for scBONITA’s ability to successfully cut down the state space of the possible rules for each node using cross-sectional scRNA-seq data.

List of abbreviations

AS: atherosclerosis; HIV: human immunodeficiency virus; PLWH: people living with HIV; PBMC: peripheral blood mononuclear cell; scBONITA: single-cell Boolean Omics Network Invariant Time Analysis; single-cell RNA sequencing: scRNA-seq; CVD: cardiovascular disease; ORA: overrepresentation analysis; cART: combined antiretroviral therapy; KEGG: Kyoto Encyclopedia of Genes and Genomes; scBONITA-RD: single-cell Boolean Omics Network Invariant Time

Analysis rule determination; scBONITA-PA: single-cell Boolean Omics Network Invariant Time Analysis pathway analysis; GA: genetic algorithm; LS: local search; RA: relative abundance; ERS: equivalent rule set; PKN: prior knowledge network

Availability of data and materials

The HIV/AS scRNA-seq dataset presented in this manuscript has been deposited in NCBI's Gene Expression Omnibus (149, 150) and is accessible through GEO Series accession number GSE198339 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE198339). Due to the sensitive nature of HIV data, we have not made the raw data public. All results presented in this manuscript may be recapitulated from the raw count data, processed data, and metadata in GEO. We also analyzed a previously published dataset that is freely accessible at https://singlecell.broadinstitute.org/single_cell/study/SCP256. All source code, tutorials, and documentation for the scBONITA Python package are available at https://github.com/Thakar-Lab/scBONITA.

Article and Author Information

Authors' information

Affiliations

[1] Biophysics, Structural, and Computational Biology Program, University of Rochester School of Medicine and Dentistry, Rochester, USA. [2] Medical Scientist Training Program, University of Rochester School of Medicine and Dentistry, Rochester, USA. [3] University of Rochester Clinical & Translational Science Institute. [4]Department of Microbiology, Immunology and Tropical Medicine, George Washington University School of Medicine and Health Sciences, Washington, D.C., USA. [5]Department of Neurology, University of Rochester School of Medicine and

Dentistry, Rochester, USA. [6]Department of Imaging Sciences, University of Rochester School of Medicine and Dentistry, Rochester, USA. [7]Department of Biostatistics and Computational Biology, University of Rochester School of Medicine and Dentistry, Rochester, USA. [8]Department of Microbiology and Immunology, University of Rochester School of Medicine and Dentistry, Rochester, USA. [9]Department of Biomedical Genetics, University of Rochester School of Medicine and Dentistry, Rochester, USA

Authors' contributions

Conceptualization of the study: JT, conceptualization of BONITA: MGP, RP, JT; Data curation: MGP, AT, JT; Formal Analysis: MGP, JT; Funding acquisition: SM, GS, JT; Investigation: MGP, JT; Methodology: MGP, JT; Project administration: AT, JT; Resources: SM, GS, MVS, JT; Software: MGP, JT; Supervision: JT; Validation: MGP, JT; Visualization: MGP, JT; Writing – original draft: MGP, JT; Writing – review & editing: All authors

Competing interests

The authors declare that they have no competing interests.

Funding

The study was supported by U.S. National Institutes of Health. MGP is supported by R01 AI134058. RP is supported by T32 GM07356. JT was supported by UM1 AI069511, P30 AI078498, R01 AI134058 and R21 AI136668. SBM, GS and AT were supported by R01 HL123346, SBM and MVS are supported by R01 HL128155, R01 NS066801. The University of Rochester Center for AIDS Research (UR-CFAR; P30 AI078498) provided support and core facilities.

Acknowledgements

We would like to thank Alan Grossfield, Andrew McDavid, Gourab Ghoshal, and all past and present members of the Thakar Lab for helpful discussions on scBONITA implementation and functionality. The Center for Integrated Research Computing at the University of Rochester provided high-performance computing resources and computing expertise. We are especially grateful to the participants in the HIV/AS study, their families, and the clinical team.

Ethics Declarations

All methods were carried out in accordance with University of Rochester guidelines and regulations, and all experimental and study protocols were approved by the University of Rochester Institutional Review Board (#RSRB00063845). The project does not qualify as human subjects research (45 CFR 46.102) in that the activities do not involve human subjects as defined in the federal regulations because this project utilizes anonymous information.

References

1. McLaughlin MM, Ma Y, Scherzer R, Rahalkar S, Martin JN, Mills C, et al. Association of Viral Persistence and Atherosclerosis in Adults With Treated HIV Infection. JAMA Netw Open. 2020;3(10):e2018099.
2. Shah ASV, Stelzle D, Lee KK, Beck EJ, Alam S, Clifford S, et al. Global Burden of Atherosclerotic Cardiovascular Disease in People Living With HIV: Systematic Review and Meta-Analysis. Circulation. 2018;138(11):1100-12.
3. Althoff KN, McGinnis KA, Wyatt CM, Freiberg MS, Gilbert C, Oursler KK, et al. Comparison of risk and age at diagnosis of myocardial infarction, end-stage renal disease, and non-AIDS-defining cancer in HIV-infected versus uninfected adults. Clin Infect Dis. 2015;60(4):627-38.
4. Drozd DR, Kitahata MM, Althoff KN, Zhang J, Gange SJ, Napravnik S, et al. Increased Risk of Myocardial Infarction in HIV-Infected Individuals in North America Compared With the General Population. J Acquir Immune Defic Syndr. 2017;75(5):568-76.
5. Mdodo R, Frazier EL, Dube SR, Mattson CL, Sutton MY, Brooks JT, et al. Cigarette smoking prevalence among adults with HIV compared with the general adult population in the United States: cross-sectional surveys. Ann Intern Med. 2015;162(5):335-44.
6. Tien PC, Schneider MF, Cox C, Karim R, Cohen M, Sharma A, et al. Association of HIV infection with incident diabetes mellitus: impact of using hemoglobin A1C as a criterion for diabetes. J Acquir Immune Defic Syndr. 2012;61(3):334-40.
7. Schouten J, Wit FW, Stolte IG, Kootstra NA, van der Valk M, Geerlings SE, et al. Cross-sectional comparison of the prevalence of age-associated comorbidities and their risk factors between

HIV-infected and uninfected individuals: the AGEhIV cohort study. Clin Infect Dis. 2014;59(12):1787-97.
8. Cotter AG, Satchell CS, O'Halloran J A, Feeney ER, Sabin CA, Mallon PW. High-density lipoprotein levels and 10-year cardiovascular risk in HIV-1-infected patients. AIDS. 2011;25(6):867-9.
9. Fahme SA, Bloomfield GS, Peck R. Hypertension in HIV-Infected Adults: Novel Pathophysiologic Mechanisms. Hypertension. 2018;72(1):44-55.
10. Hatleberg CI, Ryom L, d'Arminio Monforte A, Fontas E, Reiss P, Kirk O, et al. Association between exposure to antiretroviral drugs and the incidence of hypertension in HIV-positive persons: the Data Collection on Adverse Events of Anti-HIV Drugs (D:A:D) study. HIV Med. 2018;19(9):605-18.
11. Triant VA, Lee H, Hadigan C, Grinspoon SK. Increased acute myocardial infarction rates and cardiovascular risk factors among patients with human immunodeficiency virus disease. J Clin Endocrinol Metab. 2007;92(7):2506-12.
12. Freiberg MS, Chang CC, Kuller LH, Skanderson M, Lowy E, Kraemer KL, et al. HIV infection and the risk of acute myocardial infarction. JAMA Intern Med. 2013;173(8):614-22.
13. Vos AG, Idris NS, Barth RE, Klipstein-Grobusch K, Grobbee DE. Pro-Inflammatory Markers in Relation to Cardiovascular Disease in HIV Infection. A Systematic Review. PLoS One. 2016;11(1):e0147484.
14. Subramanian S, Tawakol A, Burdo TH, Abbara S, Wei J, Vijayakumar J, et al. Arterial inflammation in patients with HIV. JAMA. 2012;308(4):379-86.
15. Burdo TH, Lentz MR, Autissier P, Krishnan A, Halpern E, Letendre S, et al. Soluble CD163 made by monocyte/macrophages is a novel marker of HIV activity in early and chronic infection prior to and after anti-retroviral therapy. The Journal of infectious diseases. 2011;204(1):154-63.
16. Dirajlal-Fargo S, Sattar A, Kulkarni M, Funderburg N, McComsey GA. Soluble TWEAK may predict carotid atherosclerosis in treated HIV infection. HIV Clin Trials. 2017;18(4):156-63.
17. Longenecker CT, Sullivan C, Baker JV. Immune activation and cardiovascular disease in chronic HIV infection. Curr Opin HIV AIDS. 2016;11(2):216-25.
18. Corbeau P, Reynes J. Immune reconstitution under antiretroviral therapy: the new challenge in HIV-1 infection. Blood. 2011;117(21):5582-90.
19. Schafer S, Zernecke A. CD8(+) T Cells in Atherosclerosis. Cells. 2020;10(1):37.
20. Jaworowski A, Hearps AC, Angelovich TA, Hoy JF. How Monocytes Contribute to Increased Risk of Atherosclerosis in Virologically-Suppressed HIV-Positive Individuals Receiving Combination Antiretroviral Therapy. Front Immunol. 2019;10:1378.
21. Saigusa R, Winkels H, Ley K. T cell subsets and functions in atherosclerosis. Nature reviews Cardiology. 2020;17(7):387-401.
22. Tay C, Kanellakis P, Hosseini H, Cao A, Toh BH, Bobik A, et al. B Cell and CD4 T Cell Interactions Promote Development of Atherosclerosis. Front Immunol. 2019;10:3046.
23. Tse K, Tse H, Sidney J, Sette A, Ley K. T cells in atherosclerosis. Int Immunol. 2013;25(11):615-22.
24. Abdolmaleki F, Gheibi Hayat SM, Bianconi V, Johnston TP, Sahebkar A. Atherosclerosis and immunity: A perspective. Trends Cardiovasc Med. 2019;29(6):363-71.
25. Perry HM, Bender TP, McNamara CA. B cell subsets in atherosclerosis. Front Immunol. 2012;3:373.
26. Srikakulapu P, McNamara CA. B cells and atherosclerosis. Am J Physiol Heart Circ Physiol. 2017;312(5):H1060-H7.
27. Tay C, Liu YH, Hosseini H, Kanellakis P, Cao A, Peter K, et al. B-cell-specific depletion of tumour necrosis factor alpha inhibits atherosclerosis development and plaque vulnerability to rupture by reducing cell death and inflammation. Cardiovasc Res. 2016;111(4):385-97.
28. Tay C, Liu YH, Kanellakis P, Kallies A, Li Y, Cao A, et al. Follicular B Cells Promote Atherosclerosis via T Cell-Mediated Differentiation Into Plasma Cells and Secreting Pathogenic Immunoglobulin G. Arterioscler Thromb Vasc Biol. 2018;38(5):e71-e84.

29. Tsiantoulas D, Diehl CJ, Witztum JL, Binder CJ. B cells and humoral immunity in atherosclerosis. Circ Res. 2014;114(11):1743-56.
30. Weber C, Zernecke A, Libby P. The multifaceted contributions of leukocyte subsets to atherosclerosis: lessons from mouse models. Nat Rev Immunol. 2008;8(10):802-15.
31. Zernecke A, Winkels H, Cochain C, Williams JW, Wolf D, Soehnlein O, et al. Meta-Analysis of Leukocyte Diversity in Atherosclerotic Mouse Aortas. Circ Res. 2020;127(3):402-26.
32. Ding H, Blair A, Yang Y, Stuart JM. Biological process activity transformation of single cell gene expression for cross-species alignment. Nature Communications. 2019;10(1):4899.
33. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. PLoS computational biology. 2012;8(2):e1002375-e.
34. Palli R, Palshikar MG, Thakar J. Executable pathway analysis using ensemble discrete-state modeling for large-scale data. Plos Computational Biology. 2019;15(9):e1007317.
35. Palli R, Palshikar MG, Thakar J. Executable pathway analysis using ensemble discrete-state modeling for large-scale data. PLoS Comput Biol. 2019;15(9):e1007317.
36. Thakar J, Pilione M, Kirimanjeswara G, Harvill ET, Albert R. Modeling systems-level regulation of host immune responses. PLoS Comput Biol. 2007;3(6):e109.
37. Van Twisk D, Murphy SP, Thakar J. Optimized logic rules reveal interferon-gamma-induced modes regulated by histone deacetylases and protein tyrosine phosphatases. Immunology. 2017;151(1):71-80.
38. Kazer SW, Aicher TP, Muema DM, Carroll SL, Ordovas-Montanes J, Miao VN, et al. Integrated single-cell analysis of multicellular immune dynamics during hyperacute HIV-1 infection. Nat Med. 2020;26(4):511-8.
39. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36(5):411-20.
40. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453-7.
41. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol. 2015;16(1):278.
42. Back M, Yurdagul A, Jr., Tabas I, Oorni K, Kovanen PT. Inflammation and its resolution in atherosclerosis: mediators and therapeutic opportunities. Nature reviews Cardiology. 2019;16(7):389-406.
43. Negreiros-Lima GL, Lima KM, Moreira IZ, Jardim BLO, Vago JP, Galvao I, et al. Cyclic AMP Regulates Key Features of Macrophages via PKA: Recruitment, Reprogramming and Efferocytosis. Cells. 2020;9(1).
44. Wahlang B, McClain C, Barve S, Gobejishvili L. Role of cAMP and phosphodiesterase signaling in liver health and disease. Cell Signal. 2018;49:105-15.
45. Sugimoto MA, Vago JP, Perretti M, Teixeira MM. Mediators of the Resolution of the Inflammatory Response. Trends Immunol. 2019;40(3):212-27.
46. Dalli J, Serhan CN. Pro-Resolving Mediators in Regulating and Conferring Macrophage Function. Front Immunol. 2017;8:1400.
47. Lorenowicz MJ, Fernandez-Borja M, Hordijk PL. cAMP signaling in leukocyte transendothelial migration. Arterioscler Thromb Vasc Biol. 2007;27(5):1014-22.
48. Gewaltig J, Kummer M, Koella C, Cathomas G, Biedermann BC. Requirements for CD8 T-cell migration into the human arterial wall. Hum Pathol. 2008;39(12):1756-62.
49. Cornwell A, Palli R, Singh MV, Benoodt L, Tyrell A, Abe JI, et al. Molecular characterization of atherosclerosis in HIV positive persons. Sci Rep. 2021;11(1):3232.
50. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;37(7):773-82.
51. Huang S. Reprogramming cell fates: reconciling rarity with robustness. BioEssays : news and reviews in molecular, cellular and developmental biology. 2009;31(5):546-60.

52. Huang S, Eichler G, Bar-Yam Y, Ingber DE. Cell fates as high-dimensional attractor states of a complex gene regulatory network. Physical review letters. 2005;94(12):128701.
53. Huang S, Ernberg I, Kauffman S. Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. Semin Cell Dev Biol. 2009;20(7):869-76.
54. Taherian Fard A, Ragan MA. Modeling the Attractor Landscape of Disease Progression: a Network-Based Approach. Front Genet. 2017;8:48.
55. Uthamacumaran A. A review of dynamical systems approaches for the detection of chaotic attractors in cancer networks. Patterns (N Y). 2021;2(4):100226.
56. Trapnell C. Defining cell types and states with single-cell genomics. Genome Res. 2015;25(10):1491-8.
57. Gronberg C, Nilsson J, Wigren M. Recent advances on CD4(+) T cells in atherosclerosis and its implications for therapy. Eur J Pharmacol. 2017;816:58-66.
58. Saigusa R, Winkels H, Ley K. T cell subsets and functions in atherosclerosis. Nature reviews Cardiology. 2020;17(7):387-401.
59. van Duijn J, Kritikou E, Benne N, van der Heijden T, van Puijvelde GH, Kroner MJ, et al. CD8+ T-cells contribute to lesion stabilization in advanced atherosclerosis by limiting macrophage content and CD4+ T-cell responses. Cardiovasc Res. 2019;115(4):729-38.
60. van Duijn J, Kuiper J, Slutter B. The many faces of CD8+ T cells in atherosclerosis. Current opinion in lipidology. 2018;29(5):411-6.
61. Cochain C, Zernecke A. Protective and pathogenic roles of CD8(+) T cells in atherosclerosis. Basic research in cardiology. 2016;111(6):71.
62. Castilho JL, Shepherd BE, Koethe J, Turner M, Bebawy S, Logan J, et al. CD4+/CD8+ ratio, age, and risk of serious noncommunicable diseases in HIV-infected adults on antiretroviral therapy. AIDS. 2016;30(6):899-908.
63. Kyaw T, Winship A, Tay C, Kanellakis P, Hosseini H, Cao A, et al. Cytotoxic and proinflammatory CD8+ T lymphocytes promote development of vulnerable atherosclerotic plaques in apoE-deficient mice. Circulation. 2013;127(9):1028-39.
64. Zhao Y, Qian Y, Sun Z, Shen X, Cai Y, Li L, et al. Role of PI3K in the Progression and Regression of Atherosclerosis. Front Pharmacol. 2021;12:632378.
65. Flynn MC, Pernes G, Lee MKS, Nagareddy PR, Murphy AJ. Monocytes, Macrophages, and Metabolic Disease in Atherosclerosis. Front Pharmacol. 2019;10:666.
66. Gupta RM, Lee-Kim VS, Libby P. The March of Monocytes in Atherosclerosis: One Cell at a Time. Circ Res. 2020;126(10):1324-6.
67. Ley K, Miller YI, Hedrick CC. Monocyte and macrophage dynamics during atherogenesis. Arterioscler Thromb Vasc Biol. 2011;31(7):1506-16.
68. Llodra J, Angeli V, Liu J, Trogan E, Fisher EA, Randolph GJ. Emigration of monocyte-derived cells from atherosclerotic lesions characterizes regressive, but not progressive, plaques. Proc Natl Acad Sci U S A. 2004;101(32):11779-84.
69. Merino A, Buendia P, Martin-Malo A, Aljama P, Ramirez R, Carracedo J. Senescent CD14+CD16+ monocytes exhibit proinflammatory and proatherosclerotic activity. J Immunol. 2011;186(3):1809-15.
70. Ong SM, Hadadi E, Dang TM, Yeap WH, Tan CT, Ng TP, et al. The pro-inflammatory phenotype of the human non-classical monocyte subset is attributed to senescence. Cell Death Dis. 2018;9(3):266.
71. Rafieian-Kopaei M, Setorki M, Doudi M, Baradaran A, Nasri H. Atherosclerosis: process, indicators, risk factors and new hopes. Int J Prev Med. 2014;5(8):927-46.
72. Randolph GJ. The fate of monocytes in atherosclerosis. J Thromb Haemost. 2009;7 Suppl 1(Suppl 1):28-30.
73. Wilson HM. Macrophages heterogeneity in atherosclerosis - implications for therapy. J Cell Mol Med. 2010;14(8):2055-65.

74. Woollard KJ, Geissmann F. Monocytes in atherosclerosis: subsets and functions. Nature reviews Cardiology. 2010;7(2):77-86.
75. Gerhardt T, Ley K. Monocyte trafficking across the vessel wall. Cardiovasc Res. 2015;107(3):321-30.
76. Taylor JP, Tse HM. The role of NADPH oxidases in infectious and inflammatory diseases. Redox Biol. 2021;48:102159.
77. Zhang J, Li H, Wu Q, Chen Y, Deng Y, Yang Z, et al. Tumoral NOX4 recruits M2 tumor-associated macrophages via ROS/PI3K signaling-dependent various cytokine production to promote NSCLC growth. Redox Biol. 2019;22:101116.
78. Hao JJ, Liu Y, Kruhlak M, Debell KE, Rellahan BL, Shaw S. Phospholipase C-mediated hydrolysis of PIP2 releases ERM proteins from lymphocyte membrane. J Cell Biol. 2009;184(3):451-62.
79. Gupta N, Wollscheid B, Watts JD, Scheer B, Aebersold R, DeFranco AL. Quantitative proteomic analysis of B cell lipid rafts reveals that ezrin regulates antigen receptor-mediated lipid raft dynamics. Nat Immunol. 2006;7(6):625-33.
80. Parameswaran N, Matsui K, Gupta N. Conformational switching in ezrin regulates morphological and cytoskeletal changes required for B cell chemotaxis. J Immunol. 2011;186(7):4088-97.
81. Bradley JR. TNF-mediated inflammatory disease. J Pathol. 2008;214(2):149-60.
82. Mehta AK, Gracias DT, Croft M. TNF activity and T cells. Cytokine. 2018;101:14-8.
83. Jones SJ, Ledgerwood EC, Prins JB, Galbraith J, Johnson DR, Pober JS, et al. TNF recruits TRADD to the plasma membrane but not the trans-Golgi network, the principal subcellular location of TNF-R1. J Immunol. 1999;162(2):1042-8.
84. Zhang Q, Cui F, Fang L, Hong J, Zheng B, Zhang JZ. TNF-alpha impairs differentiation and function of TGF-beta-induced Treg cells in autoimmune diseases through Akt and Smad3 signaling pathway. J Mol Cell Biol. 2013;5(2):85-98.
85. Azari BM, Marmur JD, Salifu MO, Ehrlich YH, Kornecki E, Babinska A. Transcription and translation of human F11R gene are required for an initial step of atherogenesis induced by inflammatory cytokines. J Transl Med. 2011;9(1):98.
86. Theilmeier G, Michiels C, Spaepen E, Vreys I, Collen D, Vermylen J, et al. Endothelial von Willebrand factor recruits platelets to atherosclerosis-prone sites in response to hypercholesterolemia. Blood. 2002;99(12):4486-93.
87. Massberg S, Brand K, Gruner S, Page S, Muller E, Muller I, et al. A critical role of platelet adhesion in the initiation of atherosclerotic lesion formation. J Exp Med. 2002;196(7):887-96.
88. Woodfin A, Voisin MB, Nourshargh S. PECAM-1: a multi-functional molecule in inflammation and vascular biology. Arterioscler Thromb Vasc Biol. 2007;27(12):2514-23.
89. Stevens HY, Melchior B, Bell KS, Yun S, Yeh JC, Frangos JA. PECAM-1 is a critical mediator of atherosclerosis. Dis Model Mech. 2008;1(2-3):175-81; discussion 9.
90. Caligiuri G. CD31 as a Therapeutic Target in Atherosclerosis. Circ Res. 2020;126(9):1178-89.
91. Wang X-M, Zhang J-Y, Xing X, Huang H-H, Xia P, Dai X-P, et al. Global transcriptomic characterization of T cells in individuals with chronic HIV-1 infection. Cell Discovery. 2022;8(1):29.
92. Wang S, Zhang Q, Hui H, Agrawal K, Karris MAY, Rana TM. An atlas of immune cell exhaustion in HIV-infected individuals revealed by single-cell transcriptomics. Emerging Microbes & Infections. 2020;9(1):2333-47.
93. Golumbeanu M, Cristinelli S, Rato S, Munoz M, Cavassini M, Beerenwinkel N, et al. Single-Cell RNA-Seq Reveals Transcriptional Heterogeneity in Latent and Reactivated HIV-Infected Cells. Cell Rep. 2018;23(4):942-50.
94. Collora JA, Liu R, Pinto-Santini D, Ravindra N, Ganoza C, Lama JR, et al. Single-cell multiomics reveals persistence of HIV-1 in expanded cytotoxic T cell clones. Immunity. 2022;55(6):1013-31.e7.
95. Reed JC, Preston-Hurlburt P, Philbrick W, Betancur G, Korah M, Lucas C, et al. The receptor for advanced glycation endproducts (RAGE) modulates T cell signaling. PLoS One. 2020;15(9):e0236921.
96. Schäfer S, Zernecke A. CD8(+) T Cells in Atherosclerosis. Cells. 2020;10(1).

97. Flynn MC, Pernes G, Lee MKS, Nagareddy PR, Murphy AJ. Monocytes, Macrophages, and Metabolic Disease in Atherosclerosis. Front Pharmacol. 2019;10(666):666.
98. Cochain C, Zernecke A. Macrophages and immune cells in atherosclerosis: recent advances and novel concepts. Basic research in cardiology. 2015;110(4):34.
99. Shrestha S, Irvin MR, Grunfeld C, Arnett DK. HIV, inflammation, and calcium in atherosclerosis. Arterioscler Thromb Vasc Biol. 2014;34(2):244-50.
100. Crowe SM, Westhorpe CL, Mukhamedova N, Jaworowski A, Sviridov D, Bukrinsky M. The macrophage: the intersection between HIV infection and atherosclerosis. J Leukoc Biol. 2010;87(4):589-98.
101. Bianchi ME, Mezzapelle R. The Chemokine Receptor CXCR4 in Cell Proliferation and Tissue Regeneration. Front Immunol. 2020;11:2109.
102. Kharchenko PV, Silberstein L, Scadden DT. Bayesian approach to single-cell differential expression analysis. Nat Methods. 2014;11(7):740-2.
103. Brorson IS, Eriksson A, Leikfoss IS, Celius EG, Berg-Hansen P, Barcellos LF, et al. No differential gene expression for CD4(+) T cells of MS patients and healthy controls. Mult Scler J Exp Transl Clin. 2019;5(2):2055217319856903.
104. Squair JW, Gautier M, Kathe C, Anderson MA, James ND, Hutson TH, et al. Confronting false discoveries in single-cell differential expression. Nat Commun. 2021;12(1):5692.
105. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol. 2015;16:278.
106. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature. 2013;498(7453):236-40.
107. Schwab JD, Ikonomi N, Werle SD, Weidner FM, Geiger H, Kestler HA. Reconstructing Boolean network ensembles from single-cell data for unraveling dynamics in the aging of human hematopoietic stem cells. Comput Struct Biotechnol J. 2021;19:5321-32.
108. Trinh HC, Kwon YK. A novel constrained genetic algorithm-based Boolean network inference method from steady-state gene expression data. Bioinformatics. 2021;37(Suppl_1):i383-i91.
109. Lim CY, Wang H, Woodhouse S, Piterman N, Wernisch L, Fisher J, et al. BTR: training asynchronous Boolean models using single-cell expression data. BMC Bioinformatics. 2016;17(1):355.
110. Egana-Gorrono L, Lopez-Diez R, Yepuri G, Ramirez LS, Reverdatto S, Gugger PF, et al. Receptor for Advanced Glycation End Products (RAGE) and Mechanisms and Therapeutic Opportunities in Diabetes and Cardiovascular Disease: Insights From Human Subjects and Animal Models. Front Cardiovasc Med. 2020;7:37.
111. Kosmopoulos M, Drekolias D, Zavras PD, Piperi C, Papavassiliou AG. Impact of advanced glycation end products (AGEs) signaling in coronary artery disease. Biochim Biophys Acta Mol Basis Dis. 2019;1865(3):611-9.
112. Narumi K, Miyakawa R, Ueda R, Hashimoto H, Yamamoto Y, Yoshida T, et al. Proinflammatory Proteins S100A8/S100A9 Activate NK Cells via Interaction with RAGE. J Immunol. 2015;194(11):5539-48.
113. Hofmann Bowman MA, Schmidt AM. S100/calgranulins EN-RAGEing the blood vessels: implications for inflammatory responses and atherosclerosis. Am J Cardiovasc Dis. 2011;1(1):92-100.
114. Barlovic DP, Soro-Paavonen A, Jandeleit-Dahm KA. RAGE biology, atherosclerosis and diabetes. Clin Sci (Lond). 2011;121(2):43-55.
115. Yan SF, Ramasamy R, Schmidt AM. The receptor for advanced glycation endproducts (RAGE) and cardiovascular disease. Expert Rev Mol Med. 2009;11:e9.
116. Sparvero LJ, Asafu-Adjei D, Kang R, Tang D, Amin N, Im J, et al. RAGE (Receptor for Advanced Glycation Endproducts), RAGE ligands, and their role in cancer and inflammation. J Transl Med. 2009;7(1):17.

117. Nienhuis HL, Westra J, Smit AJ, Limburg PC, Kallenberg CG, Bijl M. AGE and their receptor RAGE in systemic autoimmune diseases: an inflammation propagating factor contributing to accelerated atherosclerosis. Autoimmunity. 2009;42(4):302-4.
118. Jandeleit-Dahm K, Watson A, Soro-Paavonen A. The AGE/RAGE axis in diabetes-accelerated atherosclerosis. Clin Exp Pharmacol Physiol. 2008;35(3):329-34.
119. Li H, Zhang L, Yin D, Zhang Y, Miao J. Targeting phosphatidylcholine-specific phospholipase C for atherogenesis therapy. Trends Cardiovasc Med. 2010;20(5):172-6.
120. Zhou Q, Liao JK. Rho kinase: an important mediator of atherosclerosis and vascular disease. Curr Pharm Des. 2009;15(27):3108-15.
121. Hartmann S, Ridley AJ, Lutz S. The Function of Rho-Associated Kinases ROCK1 and ROCK2 in the Pathogenesis of Cardiovascular Disease. Front Pharmacol. 2015;6:276.
122. Cai A, Zhou Y, Li L. Rho-GTPase and Atherosclerosis: Pleiotropic Effects of Statins. J Am Heart Assoc. 2015;4(7).
123. Jin SY, Kim EK, Ha JM, Lee DH, Kim JS, Kim IY, et al. Insulin regulates monocyte trans-endothelial migration through surface expression of macrophage-1 antigen. Biochim Biophys Acta. 2014;1842(9):1539-48.
124. Zhou Y, Wang Y, Qiao S, Yin L. Effects of Apelin on Cardiovascular Aging. Front Physiol. 2017;8:1035.
125. Pitkin SL, Maguire JJ, Kuc RE, Davenport AP. Modulation of the apelin/APJ system in heart failure and atherosclerosis in man. Br J Pharmacol. 2010;160(7):1785-95.
126. Ahrens TD, Bang-Christensen SR, Jorgensen AM, Loppke C, Spliid CB, Sand NT, et al. The Role of Proteoglycans in Cancer Metastasis and Circulating Tumor Cell Analysis. Front Cell Dev Biol. 2020;8:749.
127. Edwards IJ. Proteoglycans in prostate cancer. Nat Rev Urol. 2012;9(4):196-206.
128. Aberle H. Axon Guidance and Collective Cell Migration by Substrate-Derived Attractants. Front Mol Neurosci. 2019;12:148.
129. Russell SA, Bashaw GJ. Axon guidance pathways and the control of gene expression. Dev Dyn. 2018;247(4):571-80.
130. Munsky B, Neuert G, van Oudenaarden A. Using gene expression noise to understand gene regulation. Science (New York, NY). 2012;336(6078):183-7.
131. Liu S, Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges. F1000Res. 2016;5:F1000 Faculty Rev-182.
132. Zhao Y, Shao Q, Peng G. Exhaustion and senescence: two crucial dysfunctional states of T cells in the tumor microenvironment. Cell Mol Immunol. 2020;17(1):27-35.
133. Arasanz H, Gato-Canas M, Zuazo M, Ibanez-Vea M, Breckpot K, Kochan G, et al. PD1 signal transduction pathways in T cells. Oncotarget. 2017;8(31):51936-45.
134. Singh MV, Kotla S, Le NT, Ae Ko K, Heo KS, Wang Y, et al. Senescent Phenotype Induced by p90RSK-NRF2 Signaling Sensitizes Monocytes and Macrophages to Oxidative Stress in HIV-Positive Individuals. Circulation. 2019;139(9):1199-216.
135. Tsai S, Clemente-Casares X, Zhou AC, Lei H, Ahn JJ, Chan YT, et al. Insulin Receptor-Mediated Stimulation Boosts T Cell Immunity during Inflammation and Infection. Cell Metab. 2018;28(6):922-34 e4.
136. Rattan V, Shen Y, Sultana C, Kumar D, Kalra VK. Glucose-induced transmigration of monocytes is linked to phosphorylation of PECAM-1 in cultured endothelial cells. Am J Physiol. 1996;271(4 Pt 1):E711-7.
137. Okouchi M, Okayama N, Imai S, Omi H, Shimizu M, Fukutomi T, et al. High insulin enhances neutrophil transendothelial migration through increasing surface expression of platelet endothelial cell adhesion molecule-1 via activation of mitogen activated protein kinase. Diabetologia. 2002;45(10):1449-56.
138. Palli R, Palshikar MG, Thakar J. Executable pathway analysis using ensemble discrete-state modeling for large-scale data. Plos Computational Biology. 2019;15(9).

139. Diaz-Mejia JJ, Meng EC, Pico AR, MacParland SA, Ketela T, Pugh TJ, et al. Evaluation of methods to assign cell type labels to cell clusters from single-cell RNA-sequencing data. F1000Res. 2019;8.
140. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (N Y). 2021;2(3):100141.
141. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417-25.
142. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. Nucleic acids research. 2021;49(D1):D545-D51.
143. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, et al. WikiPathways: connecting communities. Nucleic acids research. 2021;49(D1):D613-D21.
144. Palshikar MG, Hilchey SP, Zand MS, Thakar J. WikiNetworks: translating manually created biological pathways for topological analysis. Bioinformatics. 2021.
145. Fortin FA, De Rainville FM, Gardner MA, Parizeau M, Gagne C. DEAP: Evolutionary Algorithms Made Easy. Journal of Machine Learning Research. 2012;13(1):2171-5.
146. Mendez A, Mendoza L. A Network Model to Describe the Terminal Differentiation of B Cells. PLoS Comput Biol. 2016;12(1):e1004696.
147. Martinez-Mendez D, Villarreal C, Mendoza L, Huerta L. An Integrative Network Modeling Approach to T CD4 Cell Activation. Front Physiol. 2020;11:380.
148. Palma A, Jarrah AS, Tieri P, Cesareni G, Castiglione F. Gene Regulatory Network Modeling of Macrophage Differentiation Corroborates the Continuum Hypothesis of Polarization States. Front Physiol. 2018;9:1659.
149. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic acids research. 2002;30(1):207-10.
150. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic acids research. 2013;41(D1):D991-D5.

Supplementary Material

1. **Supplementary File 1:**
    - **Filename:** supplementary_file_1.pdf
    - **Description:** Contains Supplementary Note 1 and Supplementary Figures 1 – 7 and corresponding captions.
        i. **Supplementary Note 1:** De-identified subject information and description of subject matching in our HIV-associated atherosclerosis study.
        ii. **Supplementary Figure 1:** Description of subjects in HIV-associated atherosclerosis (AS) study. There were no significant differences in age, cholesterol levels, high-density lipids (HDL) levels or CD4+ T cell counts between subjects with and without atherosclerosis (two-sided t-test, $p > 0.05$ for all three comparisons). (A) Distribution of serum lipids (cholesterol and HDL in mg/dl) for subjects with (AS+) and without (AS-) atherosclerosis. (B.) Distribution of CD4+ T cells/µl for subjects with (AS+) and without (AS-) atherosclerosis. (C.) Table showing age, cholesterol, and HDL for each subject.
        iii. **Supplementary Figure 2:** Cluster labels were assigned using CIBERSORT in conjunction with a reference set of 22 immune cell types. The 16 clusters are identified on the x-axis as 0 – 15. Each bar represents the proportions of the immune cell types in each cluster.
        iv. **Supplementary Figure 3:** Deconvolution of a bulk RNA-seq dataset from persons with and without AS to show the proportion of immune cell types. A bulk RNA-seq dataset obtained from matched PLWH with and without AS was deconvoluted using CIBERSORT to quantify the abundance of the cell subpopulations in the scRNA-seq dataset. Subpopulation-level

differences in the percentage of sequenced cells from a subject corresponding to each cell type in Figure 1A (main text) between AS+ and AS- PLWH are identified using a two-sided t-test. The mean of each group is represented by a red asterisk. The population ‘T cells CD8 NK cells resting’ was significantly more abundant in AS- PLWH (t-test, $p < 0.1$).

v. S**upplementary Figure 4:** Correlation between importance scores for networks trained on subpopulations of PBMCs from the Kazer et al dataset and trained on the corresponding subpopulations from the HIV/AS dataset. (A) T cells from the Kazer et al dataset and subpopulations containing CD4+ T cells from the HIV/AS dataset. (B) B cells from the Kazer et al dataset and subpopulations containing B cells from the HIV/AS dataset. (C) Monocytes from the Kazer et al dataset and from the HIV/AS dataset.

vi. **Supplementary Figure 5:** (A) Distribution of ERS sizes for 130 KEGG networks trained on PBMCs from subjects prior to infection and 1 year after HIV infection (Kazer et al). (B) Distribution of ERS sizes for 130 KEGG networks trained on PBMCs from HIV+ subjects with and without atherosclerosis.

vii. **Supplementary Figure 6:** scBONITA infers biologically meaningful dysregulated (Bonferroni-adjusted p-value < 0.1) pathways for subpopulations of PBMCs derived from HIV- subjects and subjects after 1 year of HIV infection (Kazer et al). Panels A – E show the pathways dysregulated in the HIV- vs 1 year post-infection contrast for clusters labeled as (A) 'Plasmablast’, (B) 'Dendritic Cell, (C) 'Natural Killer Cell, (D) 'T Cell’ and (E) ‘B Cell’ subpopulations.

viii. **Supplementary Figure 7.** Comparison of genes upregulated in dysregulated pathways (as identified by scBONITA) in the context of HIV-associated atherosclerosis and HIV infection (A) Number of genes upregulated in monocytes derived from AS+ PLWH and in subjects after HIV infection. (B) KEGG gene sets overrepresented in these genes (enrichr, top 10 gene sets shown) (C) Number of genes upregulated in CD8+ T cells derived from AS+ PLWH and in subjects after HIV infection. (D) KEGG gene sets overrepresented in these genes (enrichr, top 10 gene sets shown).

2. **Supplementary Table 1:**
    - Filename: *supplementary_table_1.txt*
    - **Description of Supplementary Table 1:** Cluster markers for each subpopulation identified in PBMCs derived from AS+ and AS- PLWH are listed in the file "*supplementary_table_1.csv*". Cluster markers were identified using methods implemented in the Seurat R package, as described in the methods.
3. **Supplementary Table 2:**
    - Filename: supplementary_table_2.xlsx
    - Description of supplementary table 2: supplementary_table_2 contains 2 worksheets, "DE genes in AS+ vs AS-" and "enrichr_kegg". "DE genes in AS+ vs AS-" contains a table of genes differentially expressed between cells derived from AS+ and AS- PLWH for each subpopulation in the HIV/AS dataset. The sheet "enrichr_kegg" contains a table of KEGG gene sets enriched (identified using the enrichr R package) in the DE genes from "DE genes in AS+ vs AS-".
4. **Supplementary Table 3:**
    - **Filename:** *supplementary_table_3.txt*

- **Description of Supplementary Table 3:** scBONITA infers biologically meaningful dysregulated pathways for subpopulations of PBMCs derived from AS+ and AS- PLWH in the HIVAS/HIVAS- contrast. The table of dysregulated pathways identified by scBONITA in the HIV+/AS+ - HIV+/AS- contrast in all cell clusters is presented in the CSV file titled "*supplementary_table_3.txt*".

5. **Supplementary Table 4**
    - **Filename:** *supplementary_table_4.txt*
    - **Description:** Pearson correlation coefficients between importance scores for networks trained on subpopulations of PBMCs from the Kazer et al dataset and trained on the corresponding subpopulations from the HIV/AS dataset. All p values are < 0.01.
6. **Supplementary Table 5**
    - **Filename:** *supplementary_table_5.txt*
    - **Description:** scBONITA infers biologically meaningful dysregulated pathways for subpopulations of PBMCs derived from HIV- subjects and subjects after 1 year of HIV infection (Kazer et al). The CSV file *supplementary_table_5.txt* lists the dysregulated pathways and p-values from scBONITA for every subpopulation.
7. **Supplementary Table 6**
    - **Filename:** *supplementary_table_6.txt*
    - **Description:** Comparison of dysregulated pathways, as identified by scBONITA, between subpopulations of PBMCs derived from HIV- subjects and subjects after 1 year of HIV infection (Kazer et al) and subpopulations of PBMCs derived from HIV+ subjects with and without atherosclerosis. The file *supplementary_table_6.txt* lists the dysregulated pathways for subpopulations from the Kazer et al dataset and the corresponding subpopulations from the

HIV/AS dataset, along with the intersections between the pathways dysregulated in the two contrasts. We speculate that intersecting pathways and the pathways dysregulated in the HIV/AS contrast are driven by HIV-associated inflammatory processes. Similarly, pathways that are dysregulated only in the HIV-/HIV+ contrast are assumed to be driven by the immediate antiviral response to HIV infection.

Figures

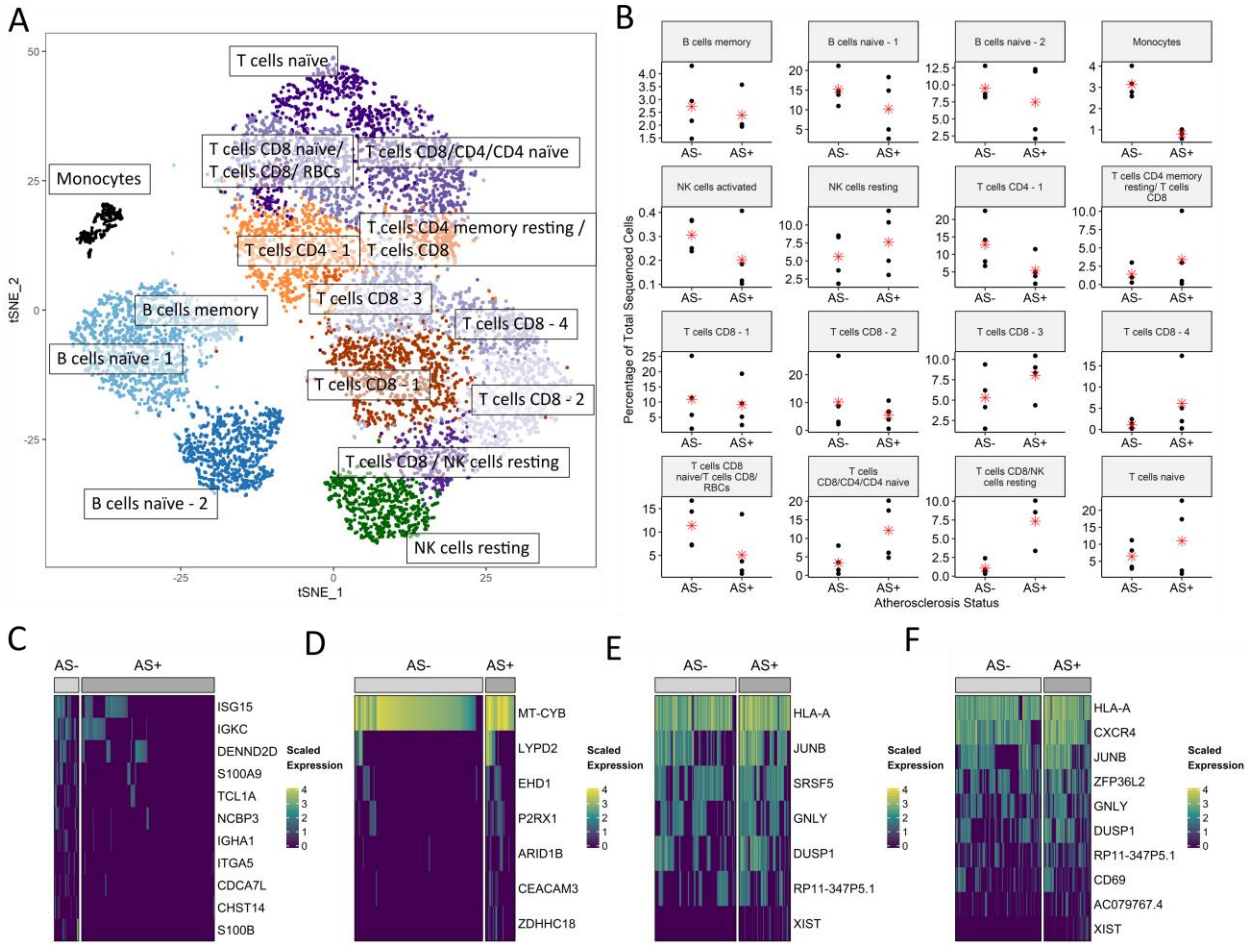

Figure 1: Characterization of PBMC subpopulations in people living with HIV (PLWH) with (AS+) or without atherosclerosis (AS-) (A) t-SNE projection of 16 transcriptionally distinct cell subpopulations, shown in distinct colors. (B) Subpopulation-level differences in the percentage of sequenced cells corresponding to each cell type in panel A from a subject between AS+ and AS- PLWH are identified using a t-test. The mean of each group is represented by a red asterisk. Panels C - F show the expression of genes that are differentially expressed (DE) between cells derived from AS+ and AS- subjects. DE genes were identified using the Wilcoxon test (Bonferroni-adjusted p-value < 0.1, absolute log2 fold change > 0.3.) DE genes in (C) CD8 T cells/NK resting cells, (D) monocytes, (E) naïve B cells referred to as "B cells naïve - 2" in panels A and B, and (F) T cells referred to as "T cells CD8/CD4/CD4 naïve" in panels A and B.

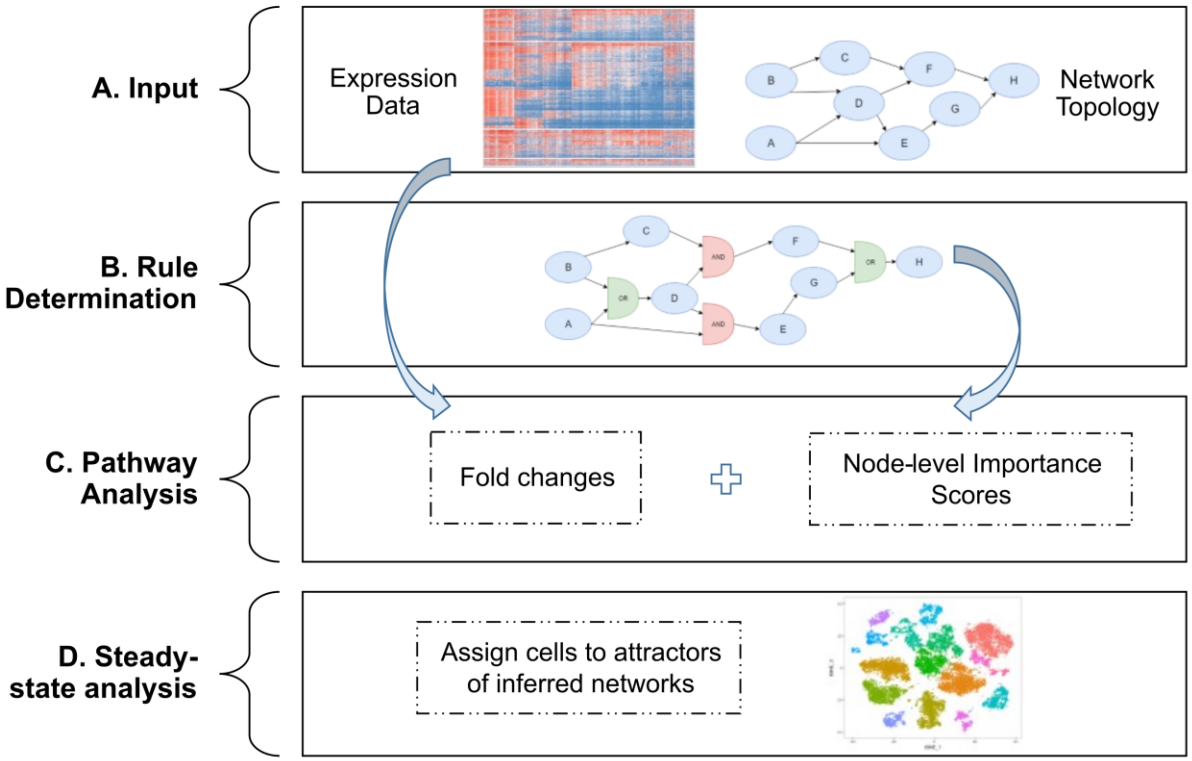

Figure 2: scBONITA pipeline to infer Boolean rules and perform pathway analysis using single cell expression measurements (A) **Input**: a binarized single-cell RNA-seq dataset as a text file, and a prior knowledge network (PKN) describing the activating or inhibitory relationships between genes (B) **Rule determination:** inference of logic rules that describe the regulatory relationships between nodes in the PKN by a global search followed by node-level rule refinement (C) **Pathway analysis:** scBONITA calculates a gene importance score calculated by simulating network perturbations with inferred rules and combines these scores with fold-changes from scRNA-seq to identify differentially regulated pathways in a specified contrast (D) **Steady-state analysis**: scBONITA simulates networks using learned rules to identify steady states which correspond to observed cellular states.

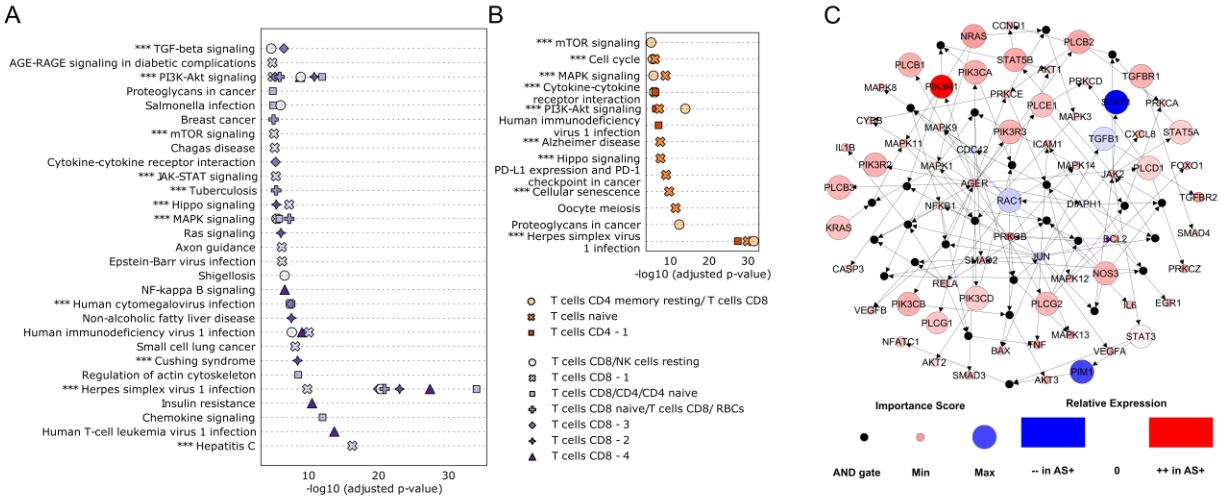

Figure 3: scBONITA identifies dysregulated pathways in T cells derived from AS+ and AS- PLWH. Pathways (y-axis) dysregulated in the AS+ vs AS- contrast in PLWH in clusters of (A) CD8+ T cells and (B) CD4+ T cells and naïve T cells. Clusters are differentiated by point shape, as shown in the legend. Pathways that have Bonferroni-corrected p-value < 0.01 (x-axis) and a reduced ERS (see Methods for details) are shown. Pathways labeled with "***" were also dysregulated between cytotoxic T cells (A) and T cells (B) derived from HIV- subjects and subjects after 1 year of HIV infection (38) (C) Network representation of the AGE-RAGE signaling pathway (Bonferroni-corrected p-value < 0.01) in a cluster of CD8+ T cells referred to as CD8 T cells -1 in Figure 1A. Small black intermediate nodes indicate that the downstream nodes are controlled by an AND function of the upstream nodes. The size of nodes corresponding to genes is proportional to their importance score calculated by scBONITA. Nodes are colored according to the magnitude of their fold change between the HIV+AS+ and HIV+AS- groups. Violet edges indicate inhibition edges and black edges indicate activation edges.

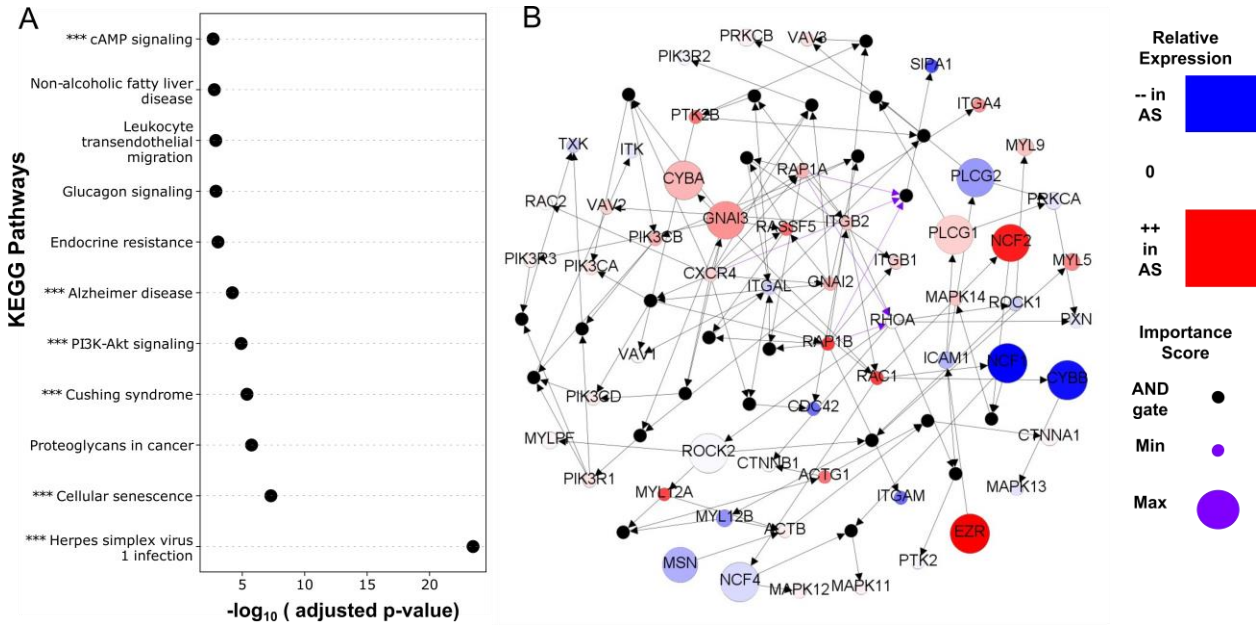

Figure 4: scBONITA identifies dysregulated pathways in monocytes derived from AS+ and AS- PLWH: (A) Pathways (y-axis) dysregulated in the AS+ vs AS- contrast in monocytes derived from PLWH. Only pathways that have Bonferroni-corrected p-value < 0.01 (x-axis) and which have a reduced ERS (see Methods for details) are shown. Pathways labeled with "***" were also significantly dysregulated in monocytes after one year of HIV infection (38) (B) Network representation of the leukocyte transendothelial migration pathway. Small black intermediate nodes indicate that the downstream nodes are controlled by an AND function of the upstream nodes. The size of nodes corresponding to genes is proportional to their importance score as calculated by scBONITA. Nodes are colored according to the magnitude of their fold change between the HIV+AS+ and HIV+AS- groups. Violet edges indicate inhibition edges and black edges indicate activation edges.

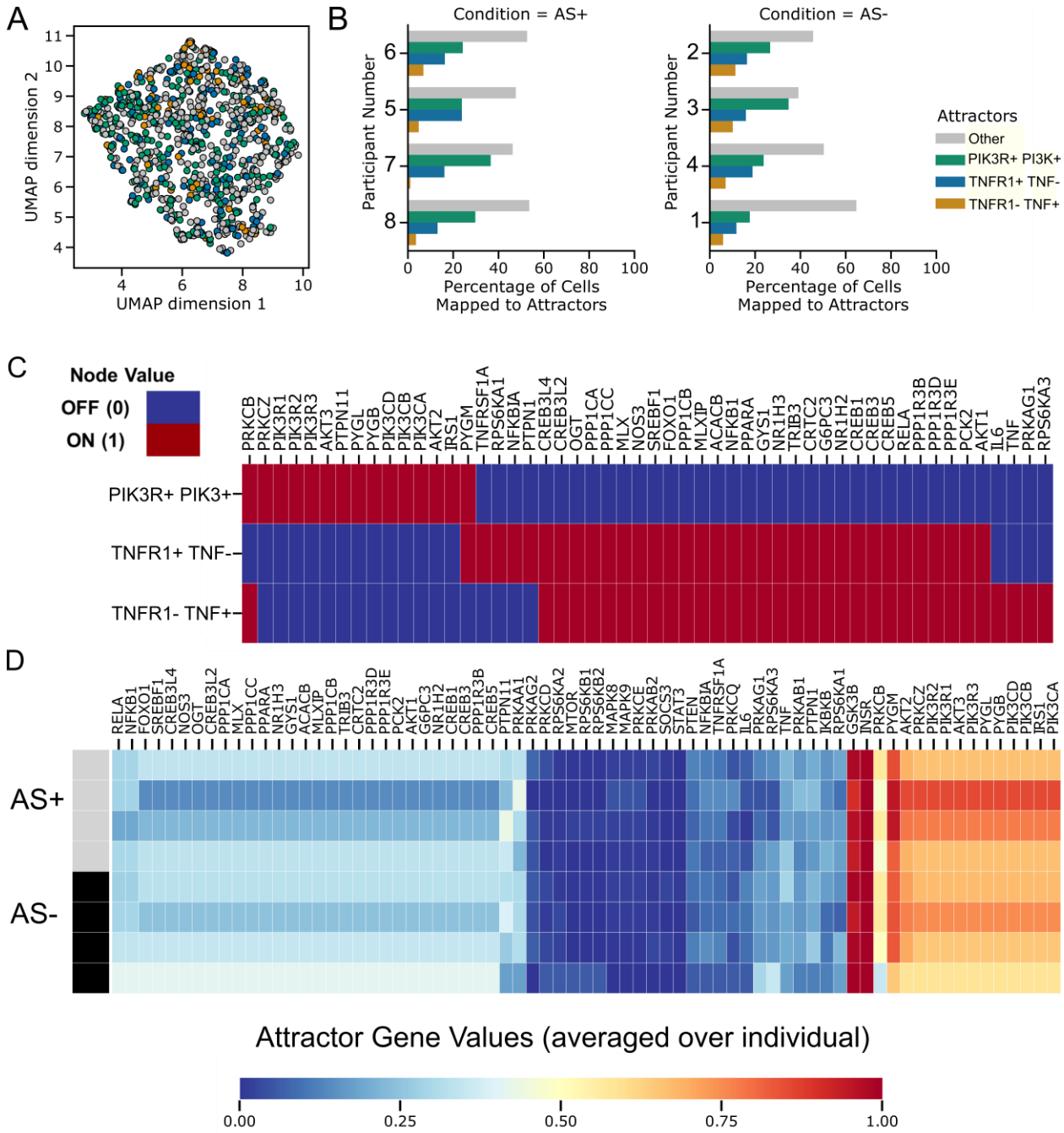

Figure 5: CD8+ T cell states with respect to the insulin resistance pathway identified by attractor analysis with scBONITA. (A) UMAP representation of a cluster of CD8+ T cells (CD8+ T cells – 1 in Figure 1A) colored by the attractor to which they are assigned, based on their similarity. The three dominant states (PI3KR+ PI3K+, TNFR1+TNF- and TNFR1-TNF+

attractors) are represented by green, blue and orange. All other attractors are collectively labeled in grey. (B) Percentages of CD8+ T cells derived from each subject, mapping to the three dominant and all other attractors. (C) Gene activity (ON- red, OFF- light blue) in the three dominant attractors. Only genes that are different between these states are shown. (D) Attractor gene values ranging from 0 (blue) to 1 (red) averaged for each individual subject. The top bar indicates AS+ (grey) and AS- (black) subjects.

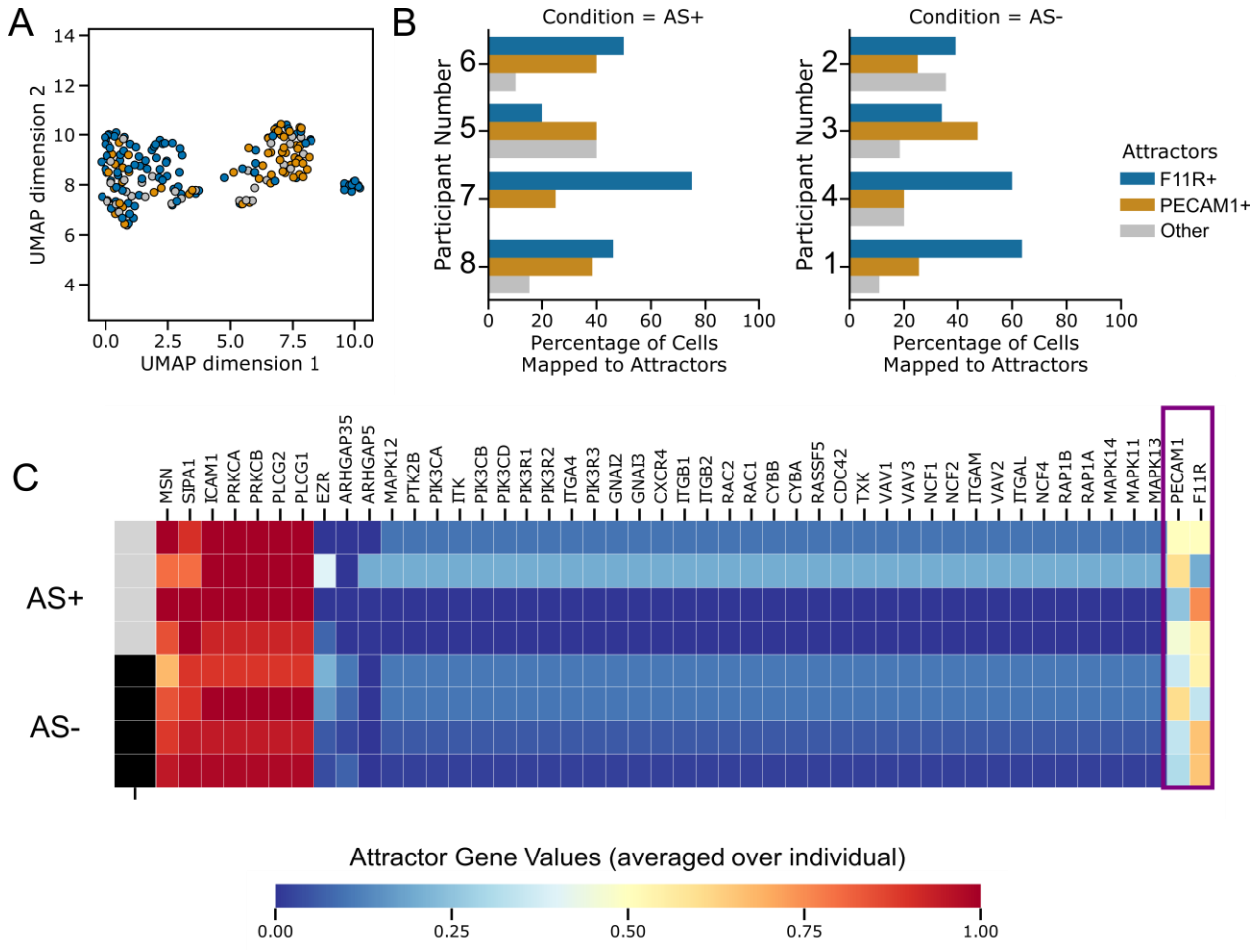

Figure 6: Monocyte states with respect to the leukocyte transendothelial migration pathway identified by attractor analysis with scBONITA. (A) UMAP representation of the cluster of monocytes colored by the attractor to which they are assigned, based on their similarity. The two dominant modes (F11R+ and PECAM+ attractors) are represented by blue and orange. All other attractors are collectively labeled in grey. (B) Percentages of monocytes derived from each subject, mapping to the two dominant attractors and all other attractors for the leukocyte transendothelial migration pathway. (C) Attractor gene values for the leukocyte transendothelial migration pathway trained on monocytes, ranging from 0 (blue) to 1 (red), averaged for each individual subjects. The top bar indicates AS+ (grey) and AS- (black) subjects. The genes that differ between the two dominant attractors F11R+ and PECAM+ are highlighted by a violet box.

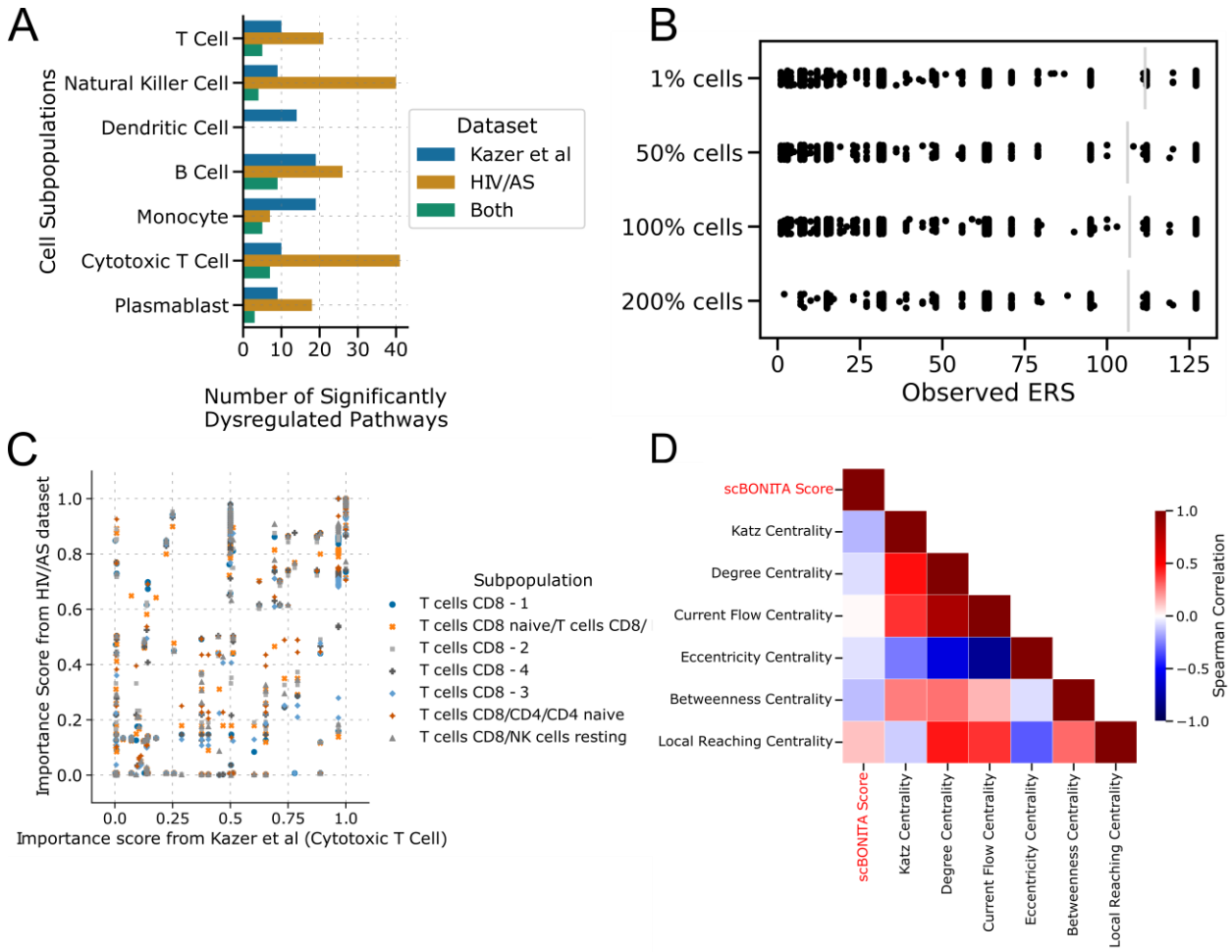

Figure 7: Performance of scBONITA rule determination. (A) The number of pathways identified as significantly dysregulated (Bonferroni-adjusted p value < 0.05) one year upon HIV infection (38), between AS+ and AS- PLWH, and the intersections between these sets. Subpopulations from the two datasets were matched as shown in Supplementary Table 5. (B) Effects of number of cells on the ERS size evaluated by downsampling and augmentation using the largest cluster ("B cells naïve – 1") from the HIV/AS dataset (C) Relation between importance scores in 130 KEGG networks evaluated using CD8+ T cells from two comparisons - AS+ and AS- PLWH and from persons before and one year after HIV infection (38). (D) Spearman correlations (p < 0.01 for all comparisons) between scBONITA's node importance score (labeled as 'scBONITA score') and 6 measures of node centrality (along x and y axis). Correlation coefficients are depicted by colors ranging from blue (-1) to red (+1).