



Published in final edited form as:

*Curr HIV/AIDS Rep.* 2012 March ; 9(1): 34–43. doi:10.1007/s11904-011-0100-x.

## Gene Array Studies in HIV-1 Infection

**Rajeev Mehla and Velpandi Ayyavoo**

Department of Infectious Diseases and Microbiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA

### Abstract

HIV-1–infected individuals exhibit remarkable variation in the onset of disease. Virus replication and disease progression depend on host cellular transcription and gene regulation in virus-specific target cells. Both viral and host factors are implicated in this differential regulation. Gene arrays and transcriptome analyses might shed light on why some infected individuals remain asymptomatic while others progress rapidly to AIDS. Here we review developments in HIV research using gene array technologies and the unifying concepts that have emerged from these studies. Gene set enrichment analysis has revealed gene signatures linked to disease progression involving pathways related to metabolism, apoptosis, cell-cycle dysregulation, and T-cell signaling. Macrophages contain anti-apoptotic signatures. Also, HIV-1 regulates previously under-emphasized cholesterol biosynthesis and energy production pathways. Notably, cellular pathways linked to a subset of HIV-infected individuals known as non-progressors contribute to survival and anti-viral responses.

### Keywords

HIV disease management; Hyperlipidemia; Drug toxicity; Disease progression; Biomarkers; HAART; Monocytes; Macrophages; Gene expression; Gene expression profiling; HIV protease inhibitors; ART; Transcriptome analysis; Gene arrays; AIDS; Lymphocytes; Microarrays

### Introduction

Antiretroviral therapy (ART) has lessened but failed to eradicate the incidence of AIDS and resulted in a continuing search for new drug targets. The goal is to elucidate virus–host interactions, identify genes involved in HIV resistance, and restore functionally active lymphocytes in order to minimize pill burden and facilitate remission. Such a strategy requires a deeper understanding of how HIV infection dysregulates metabolic pathways. HIV has a complex life cycle during which it engages multiple host cellular components, including undermining immune functions by targeting immune cells for virus replication and utilizing host transcription factors and enzymes for virus production and subsequent infection. HIV dysregulates host genes resulting in aberrant immune response, disease progression, and opportunistic infections. Recent developments in gene array technology and high-throughput screening have furthered our understanding of virus–host interaction and genome-wide dysregulation during HIV infection (Fig. 1). Given that single proteins do not work in isolation, gene arrays have revolutionized the way we assess host cellular

© Springer Science+Business Media, LLC 2011

velpandi@pitt.edu.  
ram163@pitt.edu

**Disclosure** No potential conflicts of interest relevant to this article were reported.

pathways in the context of HIV and other diseases. This technology holds the potential to decipher the role of host genes during HIV infection.

## Gene Array Technology in a Nutshell

Gene expression arrays are designed to measure the expression levels of large numbers of genes simultaneously. The array chips hold small DNA or oligonucleotide fragments as probes that will be hybridized to the complementary sequences present in the sample of interest. The development of lithographic techniques for imprinting thousands of oligonucleotide signature sequences for different genes, combined with hybridization principles, resulted in miniaturized blotting surfaces known as biochips, genechips, or DNA chips. They are primarily glass or nylon membrane platforms that can support stable imprinting with oligonucleotides representing a signature sequence from different genes. A single assay can rapidly identify thousands of genes in a sample and can differentiate between expression profiles of two or more sets of samples (such as infected and uninfected). Many variations have been developed—oligonucleotide arrays (Affimetrix chip), can-based microarrays (2-color biotin labeled spotted on glass, 33P labeled nylon filters), amplified RNA arrays, and PCR-based arrays (gene arrays/miRNA arrays). Detection methods have become more sensitive and capable of identifying minor changes in gene expression.

Analytical softwares are available to evaluate the voluminous data and develop significant conclusions (Table 1). Although current approaches for data analysis vary considerably, most use a three-tiered approach. First, differentially expressed genes are identified through Student's t-test and ANOVA or permutation-based significance criteria used in Significance Analysis of Microarrays. Second, false discovery rates are estimated using well-established statistical tests such as Bonferroni, Newman, Kuels, Tukey, Benjamini, and Hochberg. Additional statistical validations are performed using sophisticated multivariate statistics and machine-learning techniques such as Support Vector Machines and Penalized Discriminant Analysis. Third, genes are functionally annotated using public databases such as Gene Ontology, DAVID/Expression Analysis Systematic Explorer, Ingenuity Pathways Analysis, GenMAPP, STRING, Cancer Genome Anatomy Project, and Biocarta. For details of analytical approaches, please refer to other publications [1•, 2, 3]. To validate genes, more sensitive real-time PCR-based assays are being developed into high-throughput PCR array platforms.

## Gene Array Studies of HIV-Specific Target Cells: In Vivo Analysis

HIV-1 mainly targets immune system cells expressing CD4 surface receptors. These include monocytes, macrophages, lymphocytes, and dendritic cells. The virus also directly or indirectly affects uninfected bystander cells such as CD8+ T cells, NK cells, and neurons [4–7]. T-cell defects are thought to be the cause of AIDS pathogenesis [8]. Betts and colleagues [9] have shown that the quality of CD8+ T cells (based on five markers) in long-term non-progressors (LTNPs) is highly active compared with progressors, suggesting that these cells protect the host from developing AIDS. Recent studies have focused on differentially expressed genes in macrophages and CD4+ and CD8+ T cells from HIV-infected individuals. Results showed increased involvement of genes regulating complement activation, actin filaments, proteasome, and proton-transporting ATPase complex. Enriched pathways showed mitochondrial signatures of disease progression and pathways linked to metabolism, energy production, apoptosis, and cell-cycle dysregulation [10•]. Compared to lymphocytes, infected monocytes and macrophages are relatively resistant to apoptotic cell death showing anti-apoptotic gene signatures and serve as a viral reservoir [11]. An adipocytokine NAMPT/visfatin has been identified as a potential contributor to monocyte dysfunction with HIV infection [12•].

## Gene Regulation in HIV-1–Infected Cells: In Vitro Analysis

In vitro gene array studies on cells infected with HIV-1 have identified viral and cellular factors modulating gene expression and highlighted changes in different cell functions. Most report gene modulation associated with immune dysregulation, virus replication, and persistence following in vitro or in vivo infection/exposure to HIV-1 virus or viral proteins. Individual viral proteins such as Env, Vpr, Nef, and Tat induce death in different cell types (reviewed in [13–16]). Temporal analysis of HIV infection in CEM-GFP CD4+ T-cell line showed subjugation of host transcription machinery by viral mRNAs (up to 30% increase) coupled with upregulation of apoptotic genes during later stages of infection, suggesting that overburden of intracellular HIV proteins initiates apoptosis [17]. This study also found increased expression of proapoptotic genes such as p53-induced Bax and activation of caspase-2, -3, and -9 [17].

Microarray data have also yielded novel information about potential mechanisms of HIV-mediated pathogenesis, including modulation of cholesterol biosynthetic genes in CD4 T cells (relevant to virus replication and infectivity) and modulation of proteasomes and histone deacetylases in chronically infected cell lines (relevant to virus latency) [1•]. HIV-induced deregulation in host gene responses mimics exposure of those cells to heat shock, interferons, or influenza A virus [18]. One in vitro gene expression profile of HIV-resistant human T-cell clones showed 29 differentially expressed genes when compared to susceptible clones. These include different cell surface adhesion glycoproteins receptors (eg, LFA-1, CD3ε), nuclear receptor (Nup214), and transcription factors (eg, STAT, IRF-2, ErgB) that are important in different stages of viral life cycle [19].

Studies of HIV-infected macrophages show upregulated genes related to inflammation and immune response, transcription factors, and cell cycle. Increased inflammatory genes include β2M, CCL2, CCL8, PKR, OAS, MX1, CD16, MCP-1, and CXCL10; transcription factors such as c-MYC, STAT-1, p38, MAPK, ERK, STAT5A, and IFIT-1 are altered [1•, 20, 21]. Binding of the gp120 envelope has been shown to make gene expression profiles conducive for viral replication. Cell cycle-related genes upregulated in HIV-infected monocytes and MDMs include p21, RBBP, MCC, and YWHAE; down-regulated genes include UbcH6, UbcH7, Ndr, PP2Aalpha, and BM28 [22]. HIV-1 infection produced similar dysregulation of the majority of genes in a line of monocytes (U937) and of lymphocytes (Hut-78); differential regulation was found in genes such as c-myc, CD71, CD69, and β-chemokines [22]. Further, dysregulated genes in infected U937 cells are involved in divergent functions such as apoptosis (FAS, Fas ligand, PIN, HSP90β, bcl-2, bcl-x), cell-signal transduction (Ras, RGS1, IRF-1, STAT3), receptor-mediated signaling transduction (CD71, CD69, CD38), cell cycle and growth (c-myc, cytokines, kinase), transcriptional regulation (EWS, CREB-2), and chemotaxis (β-chemokines, RANTES) [22]. A study of HIV-infected monocyte-derived dendritic cells showed enhanced expression of 20% signal transduction, 14% transcription, 7% cell proliferation and cell cycle, and 7% immune response-related genes. Interferon-stimulated genes (ISGs) including STAT1, MAPK1/ERK2 kinase, and chemokine CXCL3 and SHC1 were differentially upregulated when infected by HIV subtypes C and A/E compared to subtype B [23].

## In Vivo Gene Profiling: HIV-Infected Subjects with Distinct Disease Progression or Phenotypes

HIV-1–infected individuals show a remarkable variation in virus replication and disease progression. Subsets of HIV-1 or AIDS phenotypic groups include rapid progressors (RPs), chronic progressors (CPs), viremic non-progressors (VNPs), LTNPs, and elite controllers (ECs) (previously elite suppressors), based on the decreasing occurrence of phenotypic markers (severity, CD4 counts, and viral load) of disease progression. The latter groups

effectively suppress HIV infection and maintain normal CD4 counts, and are currently the object of research interest. Gene array studies of ECs—who maintain undetectable viral load without using ART—are aimed at identifying genes for immune-mediated control of HIV or vaccine development [24–27]. Both host and viral factors are implicated in HIV-1 disease resistance [28]. It was once thought that the viruses infecting ECs are defective [24]. However, the whole-genome sequencing of six isolates showed that they were functionally intact [24], suggesting that host cellular factors are involved. In contrast to progressors, who have mitochondrial pathways related to apoptosis, nonprogressors have MAPK, WNT, and AKT pathways that contribute to cell survival and anti-viral responses [10•]. The majority of *in vivo* gene array data comes from longitudinal or cross-sectional cohort studies; most are performed on PBMCs partially because not enough patient samples are available. The use of PBMCs may limit interpretations because of dilution and may not reflect the gene profile from individual subpopulations [29]. It has been suggested that researchers limit studies using PBMCs in favor of those within subpopulations [30].

Gene expression profiling using PBMCs isolated from HIV drug-naïve mothers showed a broad spectrum of innate immune response gene-sets, including toll-like receptor, ISGs, and anti-viral RNA response pathways. HIV-specific host genetic profiling is believed to be a useful tool in preventing HIV infection and transmission [31]. Another gene array study on PBMCs from HIV-infected patients found a reduction in IL-7R $\alpha$  and increased expression of perforin in antigen-experienced mature CD8+ T cells [32]. Comparison of CD3+ T cells from LTNPs and matched disease progressors showed distinct profile differences. The LTNPs expressed genes involved in cytokine-cytokine receptor expression, negative regulation of apoptosis, and regulation of actin cytoskeleton at higher levels, whereas progressors showed an increased expression of viral genes interacting with host cellular partners [33].

In contrast to T cells, circulating monocytes from HIV-infected subjects compared to those from healthy individuals show specific anti-apoptotic gene signatures [11, 12•]. These signatures contain enhanced expression of TNF, CD40/CD40L, ERK/MAPKinase, G-protein signaling-related genes, PPAR, and p53 transcription factor. These genes modulate major monocyte functions including inflammatory response, lipid metabolism, and survival, indicating that HIV-1-resistant transcriptome is present in monocytes (Table 2) [11]. Differential regulation of transcriptome in these cell types can provide insight into the HIV-resistant genes. Transcriptome analyses of ECs recently identified subgroups of ECs—one resembling HIV-infected ART-treated aviremic individuals (EC-ART) and another showing a profile similar to that of HIV-negative controls (EC-NC) [34•]. Microarray study of such subgroups is expected to identify genes that confer HIV resistance. It has yet to be determined if these two subgroups experience differences in HIV-related complications later in treatment. That knowledge may provide specific gene data related to an effective immune reconstitution.

### Overview of Gene Arrays in HIV Latency

During HAART HIV-1 goes into latency. The viral genome is transcriptionally silenced and the viral load becomes undetectable. HIV-1-infected resting memory T cells are a major latent reservoir for infection and pose a hurdle to eradicating the virus. The use of immune-activating agents such as anti-CD3 antibody or IL-2 to purge the virus from these reservoirs has had limited success and requires identification of other transcription factors as targets to activate cells from latency [35]. A comparative analysis of resting CD4+ T cells from a viremic population and from aviremic and healthy controls ( $n=5$ ) showed a set of 370 differentially upregulated genes upon activation. Mainly related to three functional categories, these genes are essential to sustain virus production. They include genes

involved in early signal transduction molecules, transcription factors known to modulate HIV-1 transcription (YY1, TFCEP2, RUNX1), and heterogeneous nuclear ribonucleoproteins [36]. Other genes upregulated in viremic patients are related to protein/vesicle transport including ER and golgi proteins, vesicle coating associated protein, and ubiquitination-related proteins, suggesting enhanced activity in secretory pathways [36]. Additional gene array studies using ACH-2, a latently infected cell line, indicate that HIV-1 represses genes involved in the glycolytic pathway [37]. This suggests that viremic patients may provide adequate external stimuli, an intracellular environment, and metabolic energy to sustain virus replication. Latency is supported by suppression of various immune-response genes including IL-7, GZM, TLR2, and IFN and transcription factors such as MYB, MYC, and STAT5A [1•]. Genes related to the cell cycle and cytoskeleton were upregulated, while decreased expression of transcription factors and NCoA3, SRC1, p300, RAS, RAF, MAPK, and cytoskeleton-related EIF5 was observed.

### **The Future: The Use of Gene Arrays in Disease Management**

The discovery of biomarkers correlating with the severity of disease will facilitate the diagnosis of AIDS and our understanding of disease development. Although CD4+ T-cell count and viral load remain the gold standard, other immunological correlates including CRP, soluble TLR2, anti-leukocyte antibody,  $\beta$ 2 immunoglobulins, neural markers (sphingomyelin), and HLA provide limited use in predicting disease outcome [38–43]. There is a gradual shift to using molecular instead of immunological biomarkers to determine disease progression. Microarray studies also provide additional targets for antivirals. By identifying novel drug targets and pathways, these studies hold tremendous promise for better health care.

During the initial years of ART, substantial morbidity occurs with advanced disease in the absence of viral-load suppression. HAART patients also develop metabolic abnormalities such as lipodystrophy (40%–70% of patients), insulin resistance, and lactic acidosis associated with NRTI and protease inhibitors [44–46]. In an in vitro model of adipogenesis (3T3-L1 preadipocytes), adipocyte differentiation was inhibited by protease inhibitors (indinavir, saquinavir, and lopinavir) by upregulation of Wnt signaling genes and suppression of genes encoding master adipogenic transcription (C/EBP $\alpha$ , PPAR $\gamma$ ), estrogen receptor  $\beta$ , and adipocyte-specific marker (Adiponectin, leptin, Mrap, cd36, S100A8) [47]. Microarray identified the molecular mechanism behind dyslipidemia caused by protease inhibitors, and elucidated the mechanism of enhanced lipidemia associated with enhanced proteasome system [48]. To identify predictors of drug toxicity and facilitate informed decision making, large-scale gene array projects combined with conventional toxicology (toxicogenomics) have been performed at the InnoMed PredTox Consortium [49–51]. Another study called STALWART revealed that an increase in CD4 count or IL-2 therapy is not sufficient to reduce opportunistic infections or death by AIDS [52], suggesting that additional immunoprotection strategies are required. Raes and coauthors proposed CCL1 and CYP2C19 as potential biomarkers for the development of abacavir hypersensitivity reaction. Other possibilities are cytoplasmic enzymes CA2, transcription factor NFIB, transmembrane receptor NRP2, and an uncharacterized nuclear factor ANP32E [12•]. However, their study lacked statistical power due to its small sample and needs additional validation [12•].

### **The Future: The Use of Gene Arrays to Predict Disease Progression**

It is important to be able to predict how patients will respond to a given drug or what opportunistic infections they may be susceptible to during infection. Studies have been done to identify predictors of disease progression, clinical outcomes, latency, and the emergence of drug-resistant mutants. Existing diagnostic biomarkers offer limited predictive value.



What is needed is the identification of sophisticated molecular multivariate biomarkers, a combination of which would predict disease outcome. A study of 21 HIV patients from Uganda showed that host gene subsets predictive of disease prognosis were mainly related to immune response, T-cell differentiation, apoptosis, and active HIV replication (Table 1), suggesting that active destruction and regeneration of T cells are associated with rapid disease progression [53]. Early detection of these predictors may help physicians monitor overall disease status and drug resistance, change the course of treatment, and avoid worsening of the disease. Soluble CD14 (sCD14), a lipopolysaccharide receptor on monocytes, in plasma is considered predictive of mortality in HIV infection [54, 55]. Although it is a good marker of disease severity, it does not provide information about immune status or disease outcome. Boulware and colleagues [56] reported that increased levels of inflammatory molecules (CRP and IL-6), coagulation (D-dimer), and tissue fibrosis (HA) were directly proportional to disease progression and inversely proportional to plasma sCD14. Patients experiencing immune reconstitution inflammatory syndrome (IRIS) had higher TNF- $\alpha$  and HIV RNA levels, followed by significant increases in CRP, D-dimer, IL-6, IL-8, CXCL10, TNF- $\alpha$ , and IFN- $\gamma$  levels, compared to those who experienced non-IRIS events [56].

Specific gene expression profile has been shown to be an indicator of AIDS progression in HIV-1-infected individuals. Compared to VNPs, RPs upregulate genes encoding apoptosis-related cysteine peptidases (IL-1 $\beta$  convertase), caspase 1, and lymphocyte-activation gene 3 (LAG3) in CD4/CD8+ T cells [57]. Caspase 1 is involved in the proteolytic cleavage of inactive IL-1 that produces inflammation and septic shock. Subsets of regulatory T cells that express LAG3 (CD223) are endowed with potent immunosuppressive activity and release immunosuppressive cytokines such as IL-10 and TGF- $\beta$ 1. Upregulation of these genes could potentially lead to immuno-suppression and disease progression. In contrast, suppression of cytokine signaling 1 (SOCS1) and eukaryotic translation elongation factor 1 delta (EE1FD) gene levels is inversely proportional to disease progression [57]. SOCS1 is involved in negative regulation of JAK-STAT cascade, insulin-receptor signaling pathway, and interferon  $\gamma$ -mediated signaling pathway. The SOCS1 gene encodes STAT-induced STAT inhibitor (SSI), which functions downstream of receptors in a negative feedback loop to attenuate cytokine signaling. HIV is known to interfere with SOCS1 and 3, thus driving immune activation [58–60].

Recent reports suggest that cumulative viral load better correlates with disease severity and should be further validated [61]. PCR-based gene arrays can be used to assess viral loads, genotypes, and variants in patients' samples during HAART. When coupled with biomarkers of disease progression, assessment of a patient's viral load could lead to robust diagnostics and better disease management. Genotyping can be coupled in the same array to find clade-specific disease outcomes and facilitate surveillance. Data integration is the key to producing and disseminating useful information for clinical research.

Led by the nonprofit Critical Path Institute, Predictive Safety Testing Consortium is a successful public-private partnership that focuses on qualifying preclinical safety biomarkers related to drug-induced nephrotoxicity [49, 50]. Similar large-scale collaborations need to be encouraged to find biomarkers for HIV-1 disease progression. Most biomarkers found in such studies require evaluation in larger, long-term trials.

### **Suggestions for Improving Gene Array Analysis**

Many investigators use gene array analyses for biomarker identification, applying widely divergent biological samples, platforms, and analytical approaches to interpret the results. Care must be taken to interpret results in order to reach appropriate conclusions. We offer the following suggestions to improve the quality of microarray analysis.

First, choose an appropriate biological sample. The sample chosen may affect the overall interpretation. For instance, PBMCs show dilutional effect of genes active in monocytes where only highly dysregulated genes in T cells predominate while masking the gene signatures with small changes [29]. PBMCs could be used to identify “highly dysregulated genes,” bypassing unnecessary separation protocols for individual cell types.

Second, take into account that current approaches for imprinting gene chips offer poor reproducibility within different batches and need to be regulated. A particular gene chip may have overrepresented genes for certain pathways that might bias the results and future studies.

Regarding other technical variables, we discourage the use of reagents from different batches and the use of the same samples as those used in gene array for validation in PCR. Use of the same samples will validate the microarray test rather than provide conceptual validation. To avoid false subclasses in analysis, minimize nonspecific biases due to differences in sample processing. For instance, methods of cell isolation such as continuous leukapheresis may induce stress-related genes, so care must be taken in interpreting results [62]. Guidelines for a standard method for imprinting a set of genes need to be implemented to maintain uniformity should gene-array technology become a diagnostic tool.

Our next suggestions have to do with sample size and statistical considerations. It's important to calculate the dataset sample size required in an analysis to classify microarray data. Establish and standardize statistical stringency as part of routine diagnostic use of gene arrays to minimize analytical variables. Use of false discovery rate has become standard in current analysis. Giri and colleagues [1•] cautioned that removal of FDR in an array of 10,000 genes with alpha of 500 means that ~500 differentially regulated genes can skew the interpretations. Pay particular attention to pathway analysis using differentially regulated genes. Consider results with caution, and take care not to overlook underrepresented pathways, which are easy to miss because of over-representation of genes of a highly studied pathway (apoptosis, for instance) in databases. When that happens, not all the pathways resulting from the analysis may be represented.

Be cautious of false interpretations based on genomic data. For instance, CD4+ T cells from HIV-infected patients show high levels of SOCS mRNA despite protein levels remaining low [58]. Another example: modulation of some genes may be at translation levels and not transcriptional level and can be missed in these arrays. In large-scale transcriptome studies, transitory surges of viral loads in elite controllers or nonprogressors may skew the results and interpretations of the dysregulated genes [63]. Therefore, use caution in removing outliers in the estimations, especially for in vivo studies. They might represent a subgroup of a population rather than an experimental anomaly. Whenever possible, do follow-up functional validation studies of selected targets of interest in order to reach valid conclusions in terms of biological relevance.

## Conclusions

Gene array technology in HIV-1 research has opened new avenues and encouraged collaboration across different fields including basic science and clinical, statistical, and bioinformatics approaches. Sophisticated high-throughput platforms have been developed and coupled with various databases to share information efficiently across the globe (Fig. 1). This technology offers tremendous potential for studying host–pathogen interactions and discovering new drug targets in HIV-1 and other diseases. Gene array technology is expected to prevent virus replication, boost anti-viral immune response, and dampen deleterious host cell loss. The literature shows that an HIV/AIDS-specific gene array needs

to be developed to extract meaningful information that can predict disease outcome and help manage treatment. That array would combine HIV viral load, gene expression, host gene markers, AIDS progression markers, and relevant host-with-host geno-type markers.

Current gene profile studies focus on a subset of HIV-infected individuals showing resistance to AIDS progression in order to find the gene signatures that make them resistant. This subset includes LTNPs, ECs, and highly exposed seronegative (HESN) individuals (with no  $\Delta 32$  mutations). Together these studies indicate that the pathways related to HIV resistance appear to be linked to cell survival and anti-inflammation, while disease progression is related to apoptotic, cell-cycle, and metabolic pathways. Understanding the differences in regulation of these genes will further our understanding of the pathogenicity of HIV and enable the discovery of novel therapeutic approaches for AIDS.

Gene modulation studies are a rapidly developing field. The challenge is how to analyze, share, and unify data from a wide array of platforms. With the expanding knowledge of gene profiles and individualized medicine, we must be prepared for future possibilities. In particular, we need an efficient way to share data not only among researchers but also between patients and their health care providers. Towards this end, we support the use of public databases for knowledge discovery and data sharing.

## References

Papers of particular interest, published recently, have been highlighted as:

- Of importance

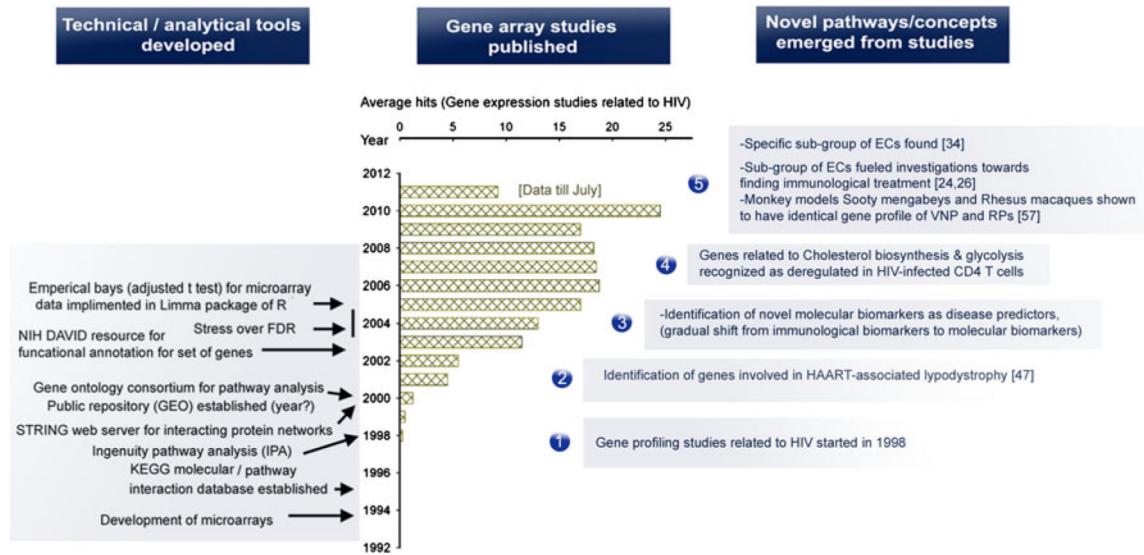
1. Giri MS, Nebozhyn M, Showe L, et al. Microarray data on gene modulation by HIV-1 in immune cells: 2000–2006. *J Leukoc Biol.* 2006; 80:1031–43. [PubMed: 16940334] [The authors have analyzed the gene expression studies and analytical approaches used in the studies prior to 2006.]
2. Doniger SW, Salomonis N, Dahlquist KD, et al. MAPPFinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. *Genome Biol.* 2003; 4:R7. [PubMed: 12540299]
3. t' Wout AB. Gene expression profiling of HIV-1 infection using cDNA microarrays. *Methods Mol Biol.* 2005; 304:455–9. [PubMed: 16061997]
4. Heath SL, Sabbaj S, Bansal A, et al. CD8 T-cell proliferative capacity is compromised in primary HIV-1 infection. *J Acquir Immune Defic Syndr.* 2011; 56:213–21. [PubMed: 21079516]
5. Richard J, Sindhu S, Pham TN, et al. HIV-1 Vpr up-regulates expression of ligands for the activating NKG2D receptor and promotes NK cell-mediated killing. *Blood.* 2010; 115:1354–63. [PubMed: 20008788]
6. Yadav A, Collman RG. CNS inflammation and macrophage/microglial biology associated with HIV-1 infection. *J Neuroimmune Pharmacol.* 2009; 4:430–47. [PubMed: 19768553]
7. Kottlilil S, Shin K, Jackson JO, et al. Innate immune dysfunction in HIV infection: effect of HIV envelope-NK cell interactions. *J Immunol.* 2006; 176:1107–14. [PubMed: 16393999]
8. Sodora DL, Silvestri G. Immune activation and AIDS pathogenesis. *AIDS.* 2008; 22:439–46. [PubMed: 18301056]
9. Betts MR, Nason MC, West SM, et al. HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. *Blood.* 2006; 107:4781–9. [PubMed: 16467198]
10. Wu JQ, Dwyer DE, Dyer WB, et al. Genome-wide analysis of primary CD4+ and CD8+ T cell transcriptomes shows evidence for a network of enriched pathways associated with HIV disease. *Retrovirology.* 2011; 8:18. [PubMed: 21410942] [This study compared the insights from HIV-infected long-term non-progressors.]
11. Giri MS, Nebozhyn M, Raymond A, et al. Circulating monocytes in HIV-1-infected viremic subjects exhibit an antiapoptosis gene signature and virus- and host-mediated apoptosis resistance. *J Immunol.* 2009; 182:4459–70. [PubMed: 19299747]



12. Van den BR, Florence E, Vlieghe E, et al. Transcriptome analysis of monocyte-HIV interactions. *Retrovirology*. 2010; 7:53. [PubMed: 20546557] [This is an important study where authors found gene correlates of HIV disease progression (eg, NAMPT) and therapy-associated complications (eg, CCL1 and CYP2C19).]
13. Majumder B, Venkatachari NJ, Schafer EA, et al. Dendritic cells infected with vpr-positive human immunodeficiency virus type 1 induce CD8+ T-cell apoptosis via upregulation of tumor necrosis factor alpha. *J Virol*. 2007; 81:7388–99. [PubMed: 17475642]
14. Gougeon ML. Apoptosis as an HIV strategy to escape immune attack. *Nat Rev Immunol*. 2003; 3:392–404. [PubMed: 12766761]
15. Cummins NW, Badley AD. Mechanisms of HIV-associated lymphocyte apoptosis: 2010. *Cell Death Dis*. 2010; 1:e99. [PubMed: 21368875]
16. Malim MH, Emerman M. HIV-1 accessory proteins—ensuring viral survival in a hostile environment. *Cell Host Microbe*. 2008; 3:388–98. [PubMed: 18541215]
17. Corbeil J, Sheeter D, Genini D, et al. Temporal gene regulation during HIV-1 infection of human CD4+ T cells. *Genome Res*. 2001; 11:1198–204. [PubMed: 11435401]
18. t' Wout AB, Lehrman GK, Mikheeva SA, et al. Cellular gene expression upon human immunodeficiency virus type 1 infection of CD4+ T-cell lines. *J Virol*. 2003; 77:1392–402. [PubMed: 12502855]
19. Kartvelishvili A, Lesner A, Szponar M, et al. Microarray analysis of differentially expressed genes in cells resistant to HIV-1. *Immunol Lett*. 2004; 93:79–86. [PubMed: 15134903]
20. Woelk CH, Ottonnes F, Plotkin CR, et al. Interferon gene expression following HIV type 1 infection of monocyte-derived macrophages. *AIDS Res Hum Retroviruses*. 2004; 20:1210–22. [PubMed: 15588343]
21. Coberley CR, Kohler JJ, Brown JN, et al. Impact on genetic networks in human macrophages by a CCR5 strain of human immunodeficiency virus type 1. *J Virol*. 2004; 78:11477–86. [PubMed: 15479790]
22. Wen W, Chen S, Cao Y, et al. HIV-1 infection initiates changes in the expression of a wide array of genes in U937 promonocytes and HUT78 T cells. *Virus Res*. 2005; 113:26–35. [PubMed: 15885842]
23. Solis M, Wilkinson P, Romieu R, et al. Gene expression profiling of the host response to HIV-1 B, C, or A/E infection in monocyte-derived dendritic cells. *Virology*. 2006; 352:86–99. [PubMed: 16730773]
24. Blankson JN, Bailey JR, Thayil S, et al. Isolation and characterization of replication-competent human immunodeficiency virus type 1 from a subset of elite suppressors. *J Virol*. 2007; 81:2508–18. [PubMed: 17151109]
25. Blankson JN. Effector mechanisms in HIV-1 infected elite controllers: highly active immune responses? *Antiviral Res*. 2010; 85:295–302. [PubMed: 19733595]
26. Baker BM, Block BL, Rothchild AC, et al. Elite control of HIV infection: implications for vaccine design. *Expert Opin Biol Ther*. 2009; 9:55–69. [PubMed: 19063693]
27. Saez-Cirion A, Pancino G, Sinet M, et al. HIV controllers: how do they tame the virus? *Trends Immunol*. 2007; 28:532–40. [PubMed: 17981085]
28. Hubert JB, Burgard M, Dussaix E, et al. Natural history of serum HIV-1 RNA levels in 330 patients with a known date of infection. The SEROCO Study Group. *AIDS*. 2000; 14:123–31.
29. Bryant PA, Smyth GK, Robins-Browne R, et al. Detection of gene expression in an individual cell type within a cell mixture using microarray analysis. *PLoS One*. 2009; 4:e4427. [PubMed: 19212463]
30. McLaren PJ, Mayne M, Rosser S, et al. Antigen-specific gene expression profiles of peripheral blood mononuclear cells do not reflect those of T-lymphocyte subsets. *Clin Diagn Lab Immunol*. 2004; 11:977–82. [PubMed: 15358662]
31. Montano M, Rarick M, Sebastiani P, et al. Gene-expression profiling of HIV-1 infection and perinatal transmission in Botswana. *Genes Immun*. 2006; 7:298–309. [PubMed: 16691187]
32. Boutboul F, Puthier D, Appay V, et al. Modulation of interleukin-7 receptor expression characterizes differentiation of CD8 T cells specific for HIV, EBV and CMV. *AIDS*. 2005; 19:1981–6. [PubMed: 16260904]

33. Salgado M, Lopez-Romero P, Callejas S, et al. Characterization of host genetic expression patterns in HIV-infected individuals with divergent disease progression. *Virology*. 2011; 411:103–12. [PubMed: 21239032]
34. Vigneault F, Woods M, Buzon MJ, et al. Transcriptional profiling of CD4 T cells identifies distinct subgroups of HIV-1 elite controllers. *J Virol*. 2011; 85:3015–19. [PubMed: 21177805]  
[This is an important study where the authors for the first time report two subgroups of elite controllers based on their gene profile. One of the ECs subgroup is closely related to uninfected controls, in contrast to the other subgroup which is more related to HIV-infected treated patients. This has implications in the development of effective therapy against HIV.]
35. Williams SA, Greene WC. Regulation of HIV-1 latency by T-cell activation. *Cytokine*. 2007; 39:63–74. [PubMed: 17643313]
36. Chun TW, Justement JS, Lempicki RA, et al. Gene expression and viral production in latently infected, resting CD4+ T cells in viremic versus aviremic HIV-infected individuals. *Proc Natl Acad Sci*. 2003; 100:1908–13. [PubMed: 12552096]
37. Munier S, Delcroix-Genete D, Carthagena L, et al. Characterization of two candidate genes, NCoA3 and IRF8, potentially involved in the control of HIV-1 latency. *Retrovirology*. 2005; 2:73. [PubMed: 16305739]
38. Lau B, Sharrett AR, Kingsley LA, et al. C-reactive protein is a marker for human immunodeficiency virus disease progression. *Arch Intern Med*. 2006; 166:64–70. [PubMed: 16401812]
39. Heggelund L, Flo T, Berg K, et al. Soluble toll-like receptor 2 in HIV infection: association with disease progression. *AIDS*. 2004; 18:2437–9. [PubMed: 15622324]
40. Fernandes AP, Goncalves MA, Zavanella RB, et al. HLA markers associated with progression to AIDS are also associated with susceptibility to cytomegalovirus retinitis. *AIDS*. 2003; 17:2133–6. [PubMed: 14502022]
41. Carbone J, Sarmiento E, Rodriguez-Molina JJ, et al. Immunoglobulin levels and prediction of progression to AIDS in HIV-infected injection drug users. *AIDS Patient Care STDS*. 2004; 18:685–6. [PubMed: 15659878]
42. Kiprov DD, Anderson RE, Morand PR, et al. Antilymphocyte antibodies and seropositivity for retroviruses in groups at high risk for AIDS. *N Engl J Med*. 1985; 312:1517. [PubMed: 2985997]
43. Lambin P, Lefrere JJ, Doinel C, et al. Neopterin and beta 2-microglobulin in serum of HIV-seropositive subjects during a two-year follow-up. *Clin Chem*. 1988; 34:1367–8. [PubMed: 3288377]
44. Mallewa JE, Wilkins E, Vilar J, et al. HIV-associated lipodystrophy: a review of underlying mechanisms and therapeutic options. *J Antimicrob Chemother*. 2008; 62:648–60. [PubMed: 18565973]
45. Nagy GS, Tsiodras S, Martin LD, et al. Human immunodeficiency virus type 1-related lipodystrophy and lipohypertrophy are associated with serum concentrations of leptin. *Clin Infect Dis*. 2003; 36:795–802. [PubMed: 12627366]
46. Fiorenza CG, Chou SH, Mantzoros CS. Lipodystrophy: pathophysiology and advances in treatment. *Nat Rev Endocrinol*. 2011; 7:137–50. [PubMed: 21079616]
47. Pacenti M, Barzon L, Favaretto F, et al. Microarray analysis during adipogenesis identifies new genes altered by antiretroviral drugs. *AIDS*. 2006; 20:1691–705. [PubMed: 16931933]
48. Waring JF, Ciurlionis R, Marsh K, et al. Identification of proteasome gene regulation in a rat model for HIV protease inhibitor-induced hyperlipidemia. *Arch Toxicol*. 2010; 84:263–70. [PubMed: 20213465]
49. Ellinger-Ziegelbauer H, Adler M, Amberg A, et al. The enhanced value of combining conventional and “omics” analyses in early assessment of drug-induced hepatobiliary injury. *Toxicol Appl Pharmacol*. 2011; 252:97–111. [PubMed: 20888850]
50. Suter L, Schroeder S, Meyer K, et al. EU Framework 6 Project: Predictive Toxicology PredTox-overview and outcome. *Toxicol Appl Pharmacol*. 2011; 252:73–84. [PubMed: 20955723]
51. Bolt HM, Marchan R, Hengstler JG. Gene array screening for identification of drugs with low levels of adverse side effects. *Arch Toxicol*. 2010; 84:253–4. [PubMed: 20229237]

52. Tavel JA, Babiker A, Fox L, et al. Effects of intermittent IL-2 alone or with peri-cycle antiretroviral therapy in early HIV infection: the STALWART study. *PLoS One*. 2010; 5:e9334. [PubMed: 20186278]
53. Motomura K, Toyoda N, Oishi K, et al. Identification of a host gene subset related to disease prognosis of HIV-1 infected individuals. *Int Immunopharmacol*. 2004; 4:1829–36. [PubMed: 15531298]
54. Brenchley JM, Price DA, Schacker TW, et al. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat Med*. 2006; 12:1365–71. [PubMed: 17115046]
55. Sandler NG, Wand H, Roque A, et al. Plasma levels of soluble CD14 independently predict mortality in HIV infection. *J Infect Dis*. 2011; 203:780–90. [PubMed: 21252259]
56. Boulware DR, Hullsiek KH, Puroton CE, et al. Higher Levels of CRP, D-dimer, IL-6, and Hyaluronic Acid Before Initiation of Antiretroviral Therapy ART Are Associated With Increased Risk of AIDS or Death. *J Infect Dis*. 2011; 203:1637–46. [PubMed: 21592994]
57. Rotger M, Dalmau J, Rauch A, et al. Comparative transcriptomics of extreme phenotypes of human HIV-1 infection and SIV infection in sooty mangabey and rhesus macaque. *J Clin Invest*. 2011; 121:2391–400. [PubMed: 21555857] [This is the first study that emphasized the rarely reported in literature HIV-1 phenotypes and highlighted the importance of using simian models of HIV infection.]
58. Miller RC, Schlaepfer E, Baenziger S, et al. HIV interferes with SOCS-1 and -3 expression levels driving immune activation. *Eur J Immunol*. 2011; 41:1058–69. [PubMed: 21337543]
59. Whyte CS, Bishop ET, Ruckerl D, et al. Suppressor of cytokine signaling SOCS1 is a key determinant of differential macrophage activation and function. *J Leukoc Biol*. 2011 doi:10.1189/jlb.1110644.
60. Nakagawa R, Naka T, Tsutsui H, et al. SOCS-1 participates in negative regulation of LPS responses. *Immunity*. 2002; 17:677–87. [PubMed: 12433373]
61. Marconi VC, Grandits G, Okulicz JF, et al. Cumulative Viral Load and Virologic Decay Patterns after Antiretroviral Therapy in HIV-Infected Subjects Influence CD4 Recovery and AIDS. *PLoS One*. 2011; 6:e17956. [PubMed: 21625477]
62. Moir S, Donoghue ET, Pickeral OK, et al. Continuous flow leukapheresis induces expression of stress genes in lymphocytes: impact on microarray analyses. *Blood*. 2003; 102:3852–3. [PubMed: 14597611]
63. Wu JQ, Wang B, Saksena NK. Transitory viremic surges in a human immunodeficiency virus-positive elite controller can shift the cellular transcriptome profile: a word of caution for microarray studies. *J Virol*. 2008; 82:10326–7. [PubMed: 18701597]
64. Smyth, GK. Informatics and computational biology solutions using R and Bioconductor. Springer; New York: 2011. *Limma: linear models for microarray data.*; p. 397-420.
65. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B*. 1995; 57:289.
66. Li L. Dimension reduction for high-dimensional data. *Methods Mol Biol*. 2010; 620:417–34. [PubMed: 20652514]
67. Ferre AL, Hunt PW, Critchfield JW, et al. Mucosal immune responses to HIV-1 in elite controllers: a potential correlate of immune control. *Blood*. 2009; 113:3978–89. [PubMed: 19109229]
68. Pereyra F, Addo MM, Kaufmann DE, et al. Genetic and immunologic heterogeneity among persons who control HIV infection in the absence of therapy. *J Infect Dis*. 2008; 197:563–71. [PubMed: 18275276]
69. Wu JQ, Dwyer DE, Dyer WB, et al. Transcriptional profiles in CD8+ T cells from HIV+ progressors on HAART are characterized by coordinated up-regulation of oxidative phosphorylation enzymes and interferon responses. *Virology*. 2008; 380:124–35. [PubMed: 18692859]
70. Geiss GK, Bumgarner RE, An MC, et al. Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. *Virology*. 2000; 266:8–16. [PubMed: 10612655]



**Fig. 1.** Chronological analysis of developments in gene array studies related to HIV and the outcome/novel concepts that emerged from these studies. Graph represents average year-wise published HIV-related gene array studies, retrieved from database searches, manual literature search, and cross-referencing

Table 1

## Methods applicable to gene array data analysis and online resources

Resource/method	Description of usage	Online link/reference
Maths XT software package	Processing of microarray datasets	(Applied Maths, St.-Martens-Latem, Belgium)
CodeLink software package	Processing of CodeLink microarray chip datasets	(Applied biosystems)
Search Tool for Retrieval of Interacting Genes/proteins (STRING)*	To find the neighboring interacting proteins in the pathway	<a href="http://string-db.org/">http://string-db.org/</a>
NIH DAVID*	Functional annotation of enriched set of genes	<a href="http://david.abcc.ncifcrf.gov/">http://david.abcc.ncifcrf.gov/</a>
BioCarta and Kyoto encyclopedia of genes and genomes (KEGG) pathway	Molecular and pathway interaction, predict interaction of differentially expressed genes in gene array data	<a href="http://www.genome.jp/kegg/">http://www.genome.jp/kegg/</a>
MetaCore (from GeneGo)	Integrated software tool for analysis of high throughput datasets in context of pathways, networks, and maps	<a href="http://www.genego.com/metacore.php">http://www.genego.com/metacore.php</a>
ExPlain (BioBase)	Pathway analysis, functional analysis, and promoter analysis (gene transcription regulation and signaling networks)	<a href="http://www.biobase-international.com/">http://www.biobase-international.com/</a>
GO (gene ontology)*	Defines gene products in terms of biological process, cellular components, and cellular functions in a species-independent manner	<a href="http://www.geneontology.org/">http://www.geneontology.org/</a>
Gene Map Annotator and Pathway Profiler (GenMAPP) software package*	Classification and visualization of data (gene maps) into functional groups?	<a href="http://www.genmapp.org/">http://www.genmapp.org/</a>
MAPPFinder*	Tool to create global gene expression profile by integrating annotations of GO using GenMAPP package	<a href="http://www.GenMAPP.org">http://www.GenMAPP.org</a> [2]
Bioconductor*	Tools for microarray analysis, annotation, and statistics -Limma package of R language: Statistical linear model for microarray data (empirical Bayes analysis for differential gene expression) -Benjamini-Hochberg false discovery rate (FDR)	<a href="http://www.bioconductor.org/">http://www.bioconductor.org/</a> [64, 65]
OmicsOffice (Integromics)	Tools for microarray and qPCR data analysis	<a href="http://www.integromics.com/omicsoffice">http://www.integromics.com/omicsoffice</a>
Gene expression omnibus (GEO) database*	Database repository for submission of microarray data and retrieval of old data from previous studies	<a href="http://www.ncbi.nlm.nih.gov/geo/">http://www.ncbi.nlm.nih.gov/geo/</a>
Principal component analysis (PCA)	To more closely compare gene expression signatures for individual samples in a small group; computational dimension reduction method uses orthogonal transformation of correlated variables to generate new set of de-correlated variables (principal components) that can be plotted in a new coordinate system in which axes follow the direction of greatest variance in the dataset	Partek softwares [66]
IPA (from Ingenuity Systems)	miRNA data analysis, integration of miRNA, and mRNA data for biomarker discovery	<a href="http://www.ingenuity.com/index.html">http://www.ingenuity.com/index.html</a>
TargetScan	Prediction of potential miRNA targets	<a href="http://www.targetscan.org/">http://www.targetscan.org/</a>

\* Open resources.



**Table 2**

## Molecular pathways deregulated in in vitro and in vivo studies

	Cells	Pathways involving differentially regulated genes in HIV infection/disease progression	Subgroups	References
<i>In vivo</i>	CD4+	-Interferon-stimulated genes (ISGs), SYNCRIP, EPSTI1, MRPS18B	RP vs EC/ VNP	[57•]
		-Apoptosis, lymphocyte activation, and IL1 activation		
		-Chemokines/cytokines and their receptors	V-ART vs C	[34•]
		-Genes impacting cell cycle		
		-Regulation of immune activation	EC vs C	[34•]
	CD8+	-Components of mitogen-activated protein kinase (MAPK)/ focal adhesion kinase (FAK) pathway, transcripts involved in cell cycle regulation	EC-ART vs EC-NC	[34•]
		-Genes required to produce sustained virus production, assembly, and release	V/AV vs C	[36]
		-Transcriptional modulation and RNA processing		
		-Protein modification and trafficking		
		-Multiple members of proteasome	RP vs EC	[9, 57•, 67, 68]
Monocytes	-Interferon-stimulated immuno-proteasome and ISGs, weak CTL response			
	-Cell cycle and cell division, metabolic genes indicating diminished cell proliferation			
	-Apoptosis/DNA damage and cell cycle	V (therapy naïve) vs C	[12•]	
	-Lipid metabolism			
	-Proteasome activity (cysteine-type endopeptidase activity, proteolysis ubiquitin cycle, ubiquitin-protein ligase activity)			
	-Protein trafficking (protein import into nucleus)			
	-Transcriptional regulation (DNA binding, negative regulation of transcription, transcription regulator activity)			
	-Genes modulated by HIV, persistent (PTGER2, KLF10, FCGR3A)	V (therapy naïve) vs ART-treated	[12•]	
	-Genes modulated by HIV, reversible (CD83, HLA-DRA, BCL6, CDKN1A, MARCKS, STAT1a, STAT1b, NAMPT, PDGFC, CCL2)			
	-Immune activation, anti-apoptotic (cis), pro-apoptotic (trans) and cell cycle	V (therapy naïve) vs infected	[12•]	
PBMCs	-Protein trafficking and transcriptional regulation and metabolic dysregulation			
	-Anti-apoptotic gene signatures	V vs C	[11]	
	-TNF- $\alpha$ signaling and CD40L/CD40 signaling, MAPK signaling, p53 modulation			
	-MAPK signaling pathway (JUND), T-cell trafficking and differentiation (ID1, KLF2)	V vs AV	[32]	
	-Cytokine signaling and interferon signaling (EIF2AK2, IFITM1)			
	-Apoptosis related (PRF1)			
	-Immune response, immature T-cell differentiation, changing immune homeostasis	P (grouped)	[53]	
	-Apoptosis, HIV replication			

	Cells	Pathways involving differentially regulated genes in HIV infection/disease progression	Subgroups	References
<i>Ex vivo</i>	CD3+/CD4 + & CD8+	<ul style="list-style-type: none"> <li>-Cell cycle (specifically to mitosis phase, nucleic acid metabolic processes, and regulation to cell cycle progression), cell apoptosis and cell cycle dysregulation</li> <li>-Endogenous stimulus response (DNA damage stimulus and DNA repair)</li> <li>-Interferon responses as a signature of progression</li> <li>-Metabolism and energy production (TCA, OXOPHOS)</li> </ul>	P vs C/ LTNP	[10,33, 69]
<i>In vitro</i>	T cells	<ul style="list-style-type: none"> <li>-Cell surface adhesion glycoprotein</li> <li>-Interferon signaling, transcription factors</li> <li>-Cell division and transcription family of DEAD box helicases, transcriptional regulation</li> <li>-Enzymes involved in cholesterol biosynthesis, cell activation and transcription factors, proapoptotic genes, expression of p53-induced product Bax, activation of caspase 2, 3, and 9</li> <li>-Genes involved in translation and splicing</li> <li>-T-cell signaling, subcellular trafficking</li> </ul>	Resistant vs Wt	[19]
	Mono/ MDMs	<ul style="list-style-type: none"> <li>-Anti-apoptotic gene signatures, p53 modulation (p21)</li> <li>-TNF-<math>\alpha</math> signaling (CCL2, IER3)</li> <li>-CD40L/CD40 signaling (MTs, CD153),</li> <li>-ERK/MAPK network associated, including genes induced by CCR5, signaling (NRAS, LYN)</li> </ul>	HIV-infected/ gp120 exposed vs C	[17, 18, 70] [11, 20, 21]

AV, aviremic; C, healthy/negative controls; EC, elite controllers; LTNP, long-term non-progressors; P, progressors; RP, rapid progressors; V, viremic; VNP, viremic non-progressors