Next-Generation Sequencing Reveals HIV-1-Mediated Suppression of T Cell Activation and RNA Processing and Regulation of Noncoding RNA Expression in a CD4$^+$ T Cell Line

Stewart T. Chang,[a] Pavel Sova,[a] Xinxia Peng,[a] Jeffrey Weiss,[a] G. Lynn Law,[a] Robert E. Palermo,[a] and Michael G. Katze[a,b]

Department of Microbiology, University of Washington, Seattle, Washington, USA[a]; and Washington Regional Primate Research Center, Seattle, Washington, USA[b]

ABSTRACT **Next-generation sequencing (NGS) enables the highly sensitive measurement of whole transcriptomes. We report the first application to our knowledge of this technology to the analysis of RNA from a CD4$^+$ T cell line infected with intact HIV. We sequenced the total mRNA from infected cells and detected differences in the expression of both host and viral mRNA. Viral reads represented a large portion of the total mapped sequencing reads: approximately 20% at 12 h postinfection (hpi) and 40% at 24 hpi. We also detected a small but significant suppression of T cell activation-related genes at 12 hpi. This suppression persisted and expanded by 24 hpi, providing new possible markers of virus-induced T cell cytopathology. By 24 hpi, the expression of over 50% of detectable host loci was also altered, indicating widespread alteration of host processes, including RNA processing, splicing, and transport to an extent not previously reported. In addition, next-generation sequencing provided insights into alternative viral RNA splice events and the expression of noncoding RNAs, including microRNA host genes.**

IMPORTANCE **Recent advances in sequencing technology now allow the measurement of effectively all the RNA in a cell. This approach is especially useful for studying models of virus infection, as it allows the simultaneous measurement of both host and viral RNA. Using next-generation sequencing (NGS), we measured changes in total mRNA from a HIV-infected T cell line. To our knowledge, this is the first application of this technology to the investigation of HIV-host interactions involving intact HIV. We directly measured the amount of viral mRNA in infected cells and detected novel viral RNA splice variants and changes in the host expression of noncoding RNA species. We also detected small changes in T cell activation and other host processes during the early stages of viral replication that increased near the peak of viral replication, providing new candidate biomarkers of T cell death.**

Received 28 June 2011 **Accepted** 19 August 2011 **Published** 20 September 2011

Citation Chang ST, et al. 2011. Next-generation sequencing reveals HIV-1-mediated suppression of T cell activation and RNA processing and regulation of noncoding RNA expression in a CD4$^+$ T cell line. mBio 2(5):e00134-11. doi:10.1128/mBio.00134-11.

Invited Editor Derya Unutmaz, Department of Microbiology and Immunology, Vanderbilt Medical Center **Editor** Terence Dermody, Vanderbilt University Medical Center

Address correspondence to Michael G. Katze, honey@u.washington.edu.

S.T.C. and P.S. contributed equally to this article.

The hallmark of AIDS is the loss of CD4$^+$ T cells, the primary target of human immunodeficiency virus type 1 (HIV-1) infection. Infected CD4$^+$ T cells undergo fundamental changes that eventually result in cell death and the release of new virus particles (reviewed in references 1 and 2). Following the uptake of virus, the viral genome is reverse transcribed and integrated into the host genome. The host machinery is then used to produce viral transcripts that are either spliced into smaller transcripts that serve as the template for viral proteins or left unspliced to be incorporated into new virus particles. Microarray analyses have shown that infected cells respond to these assaults with gene expression changes in a number of pathways, including apoptosis, cell cycle, cholesterol biosynthesis, and inflammation (3–5; reference 1 and references therein).

Many of these cellular responses are also reflected at the level of the host organism. For example, gene expression in the lymph nodes of simian immunodeficiency virus (SIV)-infected Asian pig-tailed macaques (a model of pathogenic HIV infection) reflects strong and sustained type I interferon responses (6). A similar initial interferon response is seen in the natural host, African green monkeys, but it eventually subsides and the infection resolves (6). Despite intense investigation into the molecular events following infection in these and other studies, fundamental gaps in knowledge still remain. For instance, the extent to which host and viral gene expression respond to each other is still poorly understood. New technologies enabling the measurement of both host and viral RNA may help to address such shortcomings.

In this study, we used next-generation sequencing (NGS) to examine changes in the transcriptome of T cells infected with replication-competent HIV. NGS offers a number of benefits for the study of host-pathogen interactions. Increased sensitivity and accuracy over microarrays allow important events such as the initial host response to viral replication to be studied in greater detail (7). Also, host and viral RNA transcripts can be assayed simultaneously, rather than on separate technical platforms, allowing greater reproducibility and increased confidence in the results. Finally, because NGS allows total RNA to be assayed, insights are not limited to annotated transcripts or even protein-coding genes.

Indeed, previous work in our group has shown that many noncoding RNA species are differentially expressed in virus-infected cells (8).

We report here the effects of HIV-1 infection on the expression of polyadenylated RNAs in a $CD4^+$ T cell line. Using NGS, we detected small but significant changes in host gene expression affecting T cell function that coincided with the initiation of viral RNA production at 12 h postinfection (hpi). These changes intensified near peak viral replication at 24 hpi when a multitude of other host processes were disrupted as well. In addition, by using NGS, we observed the dramatic expansion of viral mRNA expression and detected new viral splice events occurring during viral replication and differential expression of noncoding RNA species, including microRNA host genes. These findings provide an unprecedented and comprehensive view into the transcriptome-level changes that occur within T cells infected with replicating HIV.

TABLE 1 Differentially expressed (DE) host genes in HIV-infected SUP-T1 cells[a]

DE gene type and category[b]	No. of DE genes at:	
	12 hpi	24 hpi
Upregulated genes		
$1 < FC \leq 1.5$	37	1,386
$1.5 < FC \leq 2.0$	5	1,040
$FC > 2$	1	246
Total no. of up-regulated genes	43	2,672
Downregulated genes		
$1 < FC \leq 1.5$	36	1,079
$1.5 < FC \leq 2.0$	24	851
$FC > 2$	3	404
Total no. of down-regulated genes	63	2,334

[a] DE genes were identified from NGS data using limma out of 9,992 total gene loci detected in all replicates of at least one biological condition. All genes listed have Benjamini-Hochberg-adjusted *P* values of less than 0.05.
[b] FC, fold change.

RESULTS

Viral mRNA constitutes a large portion of total mRNA of HIV-infected cells. In this study, we investigated changes in host and viral transcription occurring in a T lymphoblast-based model of HIV infection. SUP-T1 cells (5×10^6) were infected in triplicate with HIV-1 strain LAI at a multiplicity of infection (MOI) of 5 which resulted in near-complete infection at 24 hours postinfection (hpi). The phenotype of infected cells (including appearance, viability, and the amount of viral mRNA and protein) was used to guide the selection of samples for NGS. At 12 hpi, viral mRNA could be detected in infected cells at appreciable levels (see Fig. S1 in the supplemental material). Few cells expressed the viral antigen p24 Gag, and visually, the cells appeared to be intact. By 24 hpi, high copy numbers of viral mRNA were present (Fig. S1). Nearly all cells showed the presence of p24 Gag antigen, syncytia were observed, and nonsyncytial cells had increased in volume. By 28 hpi, the effects of viral infection were more pronounced. The amount of viral RNA peaked (Fig. S1), and cell viability decreased. The bulk of cell death occurred between 32 and 48 hpi.

On the basis of these observations, we selected two time points to analyze by NGS: 12 and 24 hpi concurrent with the beginning of viral RNA accumulation and the occurrence of near-peak RNA levels prior to cell death, respectively. We isolated RNA from infected and mock-infected replicate cell samples at these two time points, created cDNA libraries from polyadenylated RNA, and subjected the libraries to analysis by NGS. For each sample, we obtained on average over 21 million 75-nucleotide (nt) reads mapping to either HIV or human genomes. A large number of reads mapped to the viral genome (2.5 to 8 million depending on the time point, see Table S1 in the supplemental material). By 12 hpi, viral reads constituted ~18% of the total mapped reads, and by 24 hpi, this proportion increased to ~38%. Reads mapping to the viral genome indicated the presence of RNA splicing events. In addition to splicing events involving known splice sites, we also observed evidence of five splicing events involving one or more previously unobserved splice sites (see Fig. S2 in the supplemental material). We tested the presence of one of these sites (located 3′ of the annotated splice acceptor site 2) by quantitative PCR (qPCR) and observed an amplicon size consistent with the usage of the site (Fig. S2A to S2C).

HIV-1 infection results in a small set of differentially expressed host genes at 12 hpi. To characterize the host response to HIV infection, we mapped the remaining nonviral reads to RefSeq-annotated human gene loci and quantified them as FPKM (fragments per kilobase of exon per million mapped fragments). We selected a set of 34 host genes whose expression spanned a range of values and measured the expression levels directly by qPCR (see Fig. S3 in the supplemental material). Measurements by NGS showed a high degree of correlation with qPCR results at both time points ($R = 0.97$ and $R = 0.98$ at 12 and 24 hpi, respectively), indicating a close match between the two methods. We then identified specific host genes that were differentially expressed (DE) during infection. Comparing FPKM for each gene in infected cells to time-matched mock-infected cells, we found that a total of 106 genes were DE at 12 hpi (representing all fold changes with Benjamini-Hochberg-adjusted *P* values of less than 0.05; Table 1). By 24 hpi, the number of DE genes had increased to 5,006 representing over 50% of all 9,992 gene loci detected under stringent criteria (Table 1). Closer examination of these values showed that a large proportion of DE genes occurred with small-magnitude changes. In particular, 69% (73 of 106) and 49% (2,465 of 5,006) of the expressed genes exhibited 1- to 1.5-fold changes at 12 and 24 hpi, respectively (Table 1). The observed precision of the sequencing-based measurements (Fig. S3) supported the use of low-fold change thresholds. Other values for the statistical threshold were also used, and for all but the most stringent thresholds (when fewer than 10 genes were observed to be up- or down-regulated at 12 hpi), a similar predominance of small-magnitude changes was observed (see Table S2 in the supplemental material).

Comparing the directionality of change at both time points, we found that nearly all of the DE genes at 12 hpi continued to be regulated in the same manner (i.e., up- or downregulated) at 24 hpi. Specifically, 97 of the 106 DE genes at 12 hpi were also DE at 24 hpi (with Benjamini-Hochberg-adjusted $P < 0.05$), and the majority of these (95 of 97) exhibited the same direction of change relative to mock infection, often with increased degree of change (as indicated by more-intense red or green coloration at 24 hpi than at 12 hpi in a heat map of FPKM fold changes, Fig. 1). This indicated that regulation coincident with the beginning of viral replication at 12 hpi largely persisted and intensified near peak viral replication at 24 hpi, even for initially low fold changes.

HIV-1 infection impacts core T cell functionality by 12 hpi. We determined the possible impact of these DE genes by analyzing

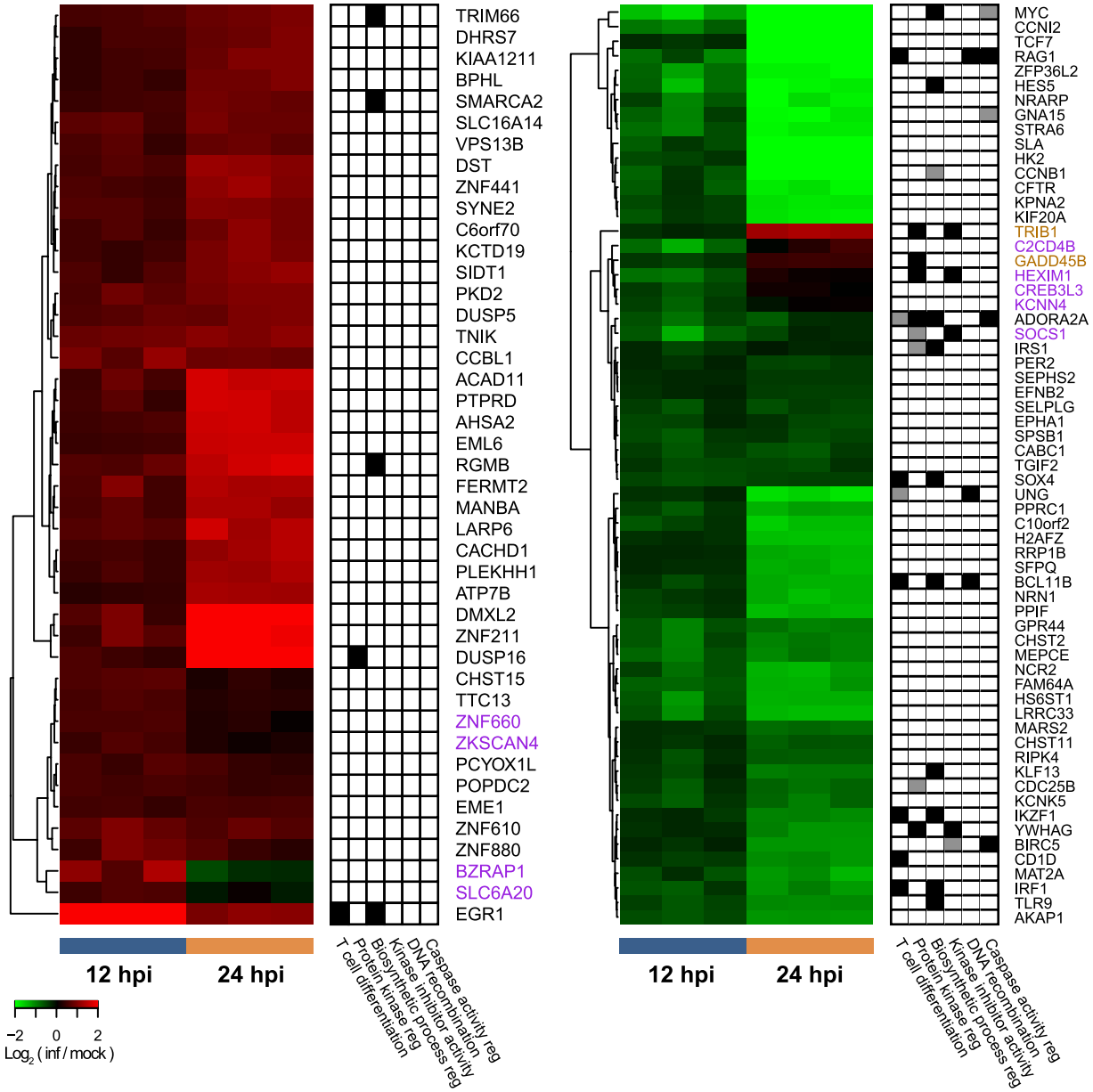

FIG 1 DE genes at 12 hpi and their mRNA expression levels at 12 and 24 hpi. Values shown are $\log_2$ ratios for each individual HIV-infected (inf) replicate cell sample relative to averaged time-matched mock-infected cell samples. Genes were segregated by the direction of change relative to mock infection at 12 hpi, and hierarchical clustering was done within each directional group. Genes that were not also DE at 24 hpi (purple type) and genes that were also DE at 24 hpi but with changed directionality (gold type) are indicated. Annotations indicate overrepresented categories in DAVID (Fig. 2). Matches to top-scoring categories in each DAVID annotation cluster (matching numbers in Fig. 2) (solid black squares) and matches to related categories in the same DAVID cluster as the top-scoring category (solid gray squares) are indicated. reg, regulation.

the corresponding annotations (biological processes, molecular functions, and pathways) using the NIH DAVID resource. We found that T cell activation and differentiation were the most overrepresented annotations among DE host genes at 12 hpi overall (Fig. 2A). Other annotations associated with DE genes included protein kinase regulation, DNA recombination, and caspase activity regulation (Fig. 2A). In addition, several of the DE genes at 12 hpi encoded transcriptional regulators; these genes include *EGR1*, *KLF13*, and *MYC*. Interestingly, four of the nine genes DE at 12 hpi but not at 24 hpi also encoded transcriptional regulators (*CREB3L3*, *HEXIM1*, *ZNF660*, and *ZKSCAN4*), indicating regulation specific to early viral replication.

Seven genes contributed to the overrepresentation of T cell activation among DE genes at 12 hpi (*BCL11B*, *CD1D*, *EGR1*, *IKZF1*, *IRF1*, *RAG1*, and *SOX4*), six of which were downregulated, indicating a net suppression of T cell activation (as T cell differentiation in Fig. 2). Five T cell activation-related DE genes (*BCL11B*, *EGR1*, *IRF1*, *IKZF1*, and *SOX4*) encoded transcription factors, some of which had known relevance to HIV infection. For example, the *BCL11B* gene product binds directly to the HIV-1

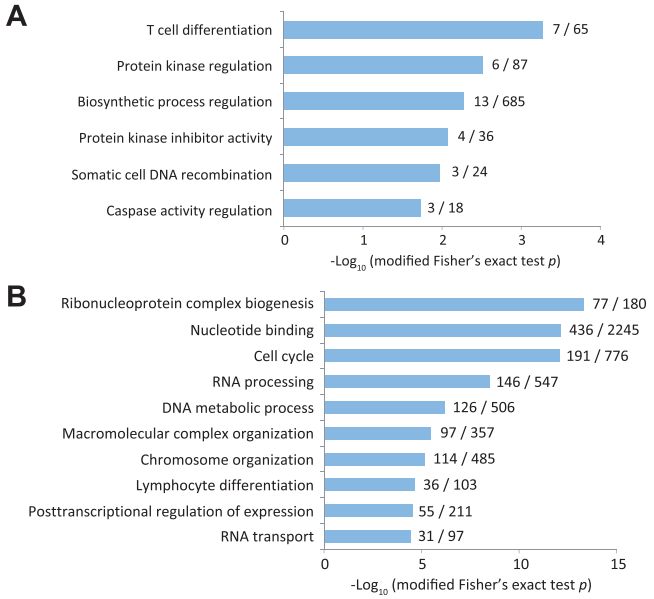

FIG 2 Top annotations associated with DE genes at 12 and 24 hpi as identified by DAVID. Annotations (for biological process, molecular function, or pathway) were identified by Fisher's exact test and then clustered by overlapping gene hits. The negative $\log_{10}$-transformed modified Fisher's exact test *P* values for each cluster are shown with the number of intersecting DE genes and total number of genes in the top-scoring gene set within each cluster indicated to the right of each bar. (A) Annotations at 12 hpi. DE genes were identified by Benjamini-Hochberg-adjusted *P* values of <0.05 with no fold change cutoff. Clustering was performed with high stringency, and all clusters with $-\log_{10} P$ value greater than 1.3 are shown. (B) Annotations at 24 hpi. DE genes were identified as defined above for panel A with average fold change of >1.5 to allow submaximal number of genes for DAVID analysis. Grouping was performed with medium stringency criteria, and the top 10 clusters are shown.

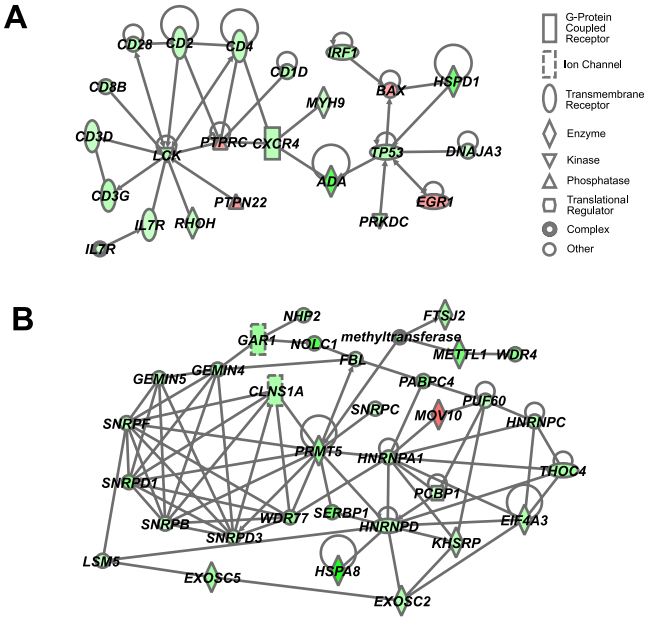

FIG 3 Two networks affected by DE genes at 24 hpi. (A) T cell activation-related DE genes at 24 hpi as connected in a *de novo* network based on known interactions in Ingenuity Pathways Analysis (IPA). A total of 23 genes DE at 24 hpi are shown, selected from the union of T cell activation- and T cell differentiation-related genes as identified in DAVID. Downregulated genes (green) and upregulated genes (red) are indicated. (B) RNA processing-related ribonucleoprotein synthesis DE genes at 24 hpi as shown in a *de novo* network in IPA.

long terminal repeat (LTR) and inhibits LTR expression (9). *BCL11B* downregulation may therefore allow more efficient HIV-1 replication. Other DE genes related to T cell activation not encoding transcription factors also had known relevance to HIV infection. For example, the *CD1D* product presents lipid antigens on the surface and is directed by HIV-1 Nef for internalization and degradation (10). Downregulation of *CD1D* at the mRNA level would further reduce surface expression and facilitate immune evasion.

We complemented functional analysis in DAVID by gene set enrichment analysis (GSEA). This method does not rely on the prior determination of DE genes (e.g., by applying a *P* value threshold) but instead uses a ranked gene list as input and is therefore well suited for identifying annotations among genes with small-magnitude changes in expression (11). GSEA identified additional genes that were downregulated and contributed to the suppression of T cell activation; these genes include the *CD2*, *CD4*, *CD7*, *CD28*, and *SIT1* genes encoding T cell-specific surface molecules (see Fig. S4 in the supplemental material). Many of the genes associated with T cell activation in DAVID and GSEA also had roles in other pathways, indicating possible pleiotropic effects (e.g., *ADORA2A* and *RAG1* which also contribute to caspase activity [Fig. 1]).

HIV-1 induces large-scale changes in transcription by 24 hpi. By 24 hpi, large-scale changes in the transcriptomes of HIV-infected cells were evident (Table 1; see the heatmap in Fig. S5A in the supplemental material). Consistent with the large number of DE genes at 24 hpi, we found that a wide range of biological processes was affected as determined by further analysis using DAVID (Fig. 2B; complete listing in Table S3 in the supplemental material). T cell activation and differentiation continued to be overrepresented and were associated with an increased number of DE genes (as lymphocyte differentiation [Fig. 2B]). Other fundamental cellular functions associated with DE genes at 24 hpi included DNA metabolism, transcription, mRNA processing and splicing, translation, protein degradation, and cell cycle control (Fig. 2B). This breadth of functional regulation suggests that HIV-1 infection resulted in the effective transcriptional reprogramming of SUP-T1 cells by 24 hpi.

Suppression of T cell functionality persists and expands at 24 hpi. As was the case at 12 hpi, genes related to T cell activation and differentiation were predominantly downregulated at 24 hpi. All of the T cell activation-related DE genes at 12 hpi continued to exhibit the same direction of change at 24 hpi (Fig. 1). Other T cell activation-related genes DE at 24 hpi included *CD3D*, *CD3Q*, and *RHOH*, all of which encode membrane proteins crucial for T cell receptor function. A network depicting known interactions among proteins encoded by T cell activation-related DE genes at 24 hpi showed two highly connected nodes: *LCK* (lymphocyte-specific protein kinase) and *TP53* (p53), both of which were strongly downregulated at 24 hpi (Fig. 3A). Our observed downregulation of *TP53* differed from the upregulation observed in previous microarray studies and may indicate a cell type-specific effect (1). Genes involved in other core T cell functions, such as proliferation, survival, and antigen presentation, were also downregulated at 24 hpi. These genes included *CD1D*, *CD28*, *CXCR4*,

TNFRSF4, and *TREML2*. *TOB1* (transducer of ERBB2), a negative regulator of T cell activation, proliferation, and interleukin 2 (IL-2) production, showed increased expression and may contribute to the downregulation observed in other genes (data not shown). A connection between *TOB1* expression and HIV-1 infection has not previously been mentioned. Overall, the increased number of DE genes related to T cell activation indicated increased suppression of this pathway by 24 hpi.

Ribonucleoprotein complex biogenesis and RNA processing. Gene sets related to ribonucleoprotein complex biogenesis were the most overrepresented annotations among DE genes at 24 hpi (Fig. 2B). Ribonucleoproteins contribute to diverse functions within the cell, including microRNA synthesis, ribosomal assembly, and translation (12). This overlap was reflected at the gene level as well. In particular, ribonucleoprotein complex biogenesis shared many DE genes at 24 hpi with another top annotation cluster, RNA processing. This set of genes encoding the ribonucleoprotein component of the RNA processing machinery was almost entirely downregulated and included many members of the heterogeneous nuclear ribonucleoprotein (hnRNP) and small nuclear ribonucleoprotein (snRNP) families. In a network identifying interactions among the proteins encoded by these genes, a number were found to be highly connected (Fig. 3B). One of these was *HNRNPA1* whose gene product regulates the splicing of eukaryotic and viral mRNA (13). HNRNPA1 and other hnRNP gene products are known to affect the localization of HIV-1 Gag/Pol mRNA, and their overexpression has previously been shown to reduce HIV-1 replication (Fig. 3B) (14). Other DE genes at 24 hpi related to ribonucleoprotein biogenesis and RNA processing are also known to affect HIV-1 replication directly; these genes include *TARDBP* (TAR DNA-binding protein 43) whose protein binds the transactivation response (TAR) element of integrated HIV-1 DNA and represses HIV-1 transcription (15–17). Downregulation of *HNRNPA1*, *TARDBP*, and other related genes may therefore allow more efficient HIV-1 replication in SUP-T1 cells.

RNA transport. The importance of host regulation of RNA fate was underscored by another cluster of gene sets related to RNA transport (Fig. 2B). Host factors for RNA transport are known to be coopted by HIV-1 to allow the export of unspliced, partially spliced, or fully spliced viral mRNAs out of the nucleus (18). In particular, the host factors Crm1 (exportin 1) and Ran GTPase are used for the export of unspliced viral mRNA, while the host factor NXF1 (nuclear RNA export factor 1) facilitates the export of fully spliced viral mRNA in a Ran-independent manner (19). In our data set, we observed no change in the expression of the Crm1-encoding gene *XPO1*, but several members of the Ran signaling pathway were downregulated, suggesting that the overall export of unspliced viral RNA was suppressed (see Fig. S5B in the supplemental material; data not shown for *XPO1*). However, in contrast to other RNA transport-related genes, *NXF1* and *NXF3* were expressed more highly in infected cells, suggesting that HIV infection may have selectively enhanced the export of fully spliced viral RNA (Fig. S5B).

Limited upregulation of host processes at 24 hpi. Despite the presence of an almost equal number of up- and downregulated genes at 24 hpi, relatively few gene sets (functions, processes, or pathways) were associated with upregulated genes. When up- and downregulated DE genes at 24 hpi were analyzed separately, more annotation clusters were observed among downregulated genes (70 versus 13 among upregulated genes with significance defined by modified Fisher's exact test as $P < 0.05$). A similar result was obtained by GSEA (see Fig. S6A and S6B in the supplemental material). These results suggested that upregulated genes were distributed across many gene sets with few occurring in particular functions or pathways. Among the few upregulated pathways were stress-activated/Jnk cascade signaling and ion transport (Fig. S6C). Consistent with these observations, HIV-1 Nef has been shown to activate Jnk signaling, ultimately activating the caspase cascade and triggering cell death (20). The triggering of apoptosis at 24 hpi is consistent with our observation that infected cells began to die following the 24-hpi time point.

HIV-1 infection does not trigger innate sensing at 12 hpi. Notably, innate immunity was absent from the processes associated with DE genes at 12 hpi, suggesting that HIV-1 infection impaired T cell activation while evading virus-sensing mechanisms (Fig. 2A). With the exception of *IRF1*, which was downregulated at 12 hpi (Fig. 1), interferon response factors (IRFs) (*IRF2*, *IRF3*, *IRF7*, and *IRF9*) showed no significant change in expression at 12 hpi (data not shown). Furthermore, specific IRF targets, including IRF3 target genes *IFI35*, *IFI44*, *ISG15*, and *ISG20*, were also not differentially expressed at this time point (21, 22). IRF3 targets were of particular interest, as intact cytoplasmic HIV-1 DNA has been shown to activate IRF3 in $CD4^+$ T cells (23). Our observed lack of IRF3 target gene expression is consistent with previous observations that replicating HIV-1 suppressed IRF3 activity (24). We also examined the levels of inflammation-related genes previously observed to be differentially regulated in HIV infection studies (1): *ANXA1*, *B2M*, and *CD69* (generally upregulated) and *CD53*, *CD71*, *IL-13*, and *IL-16* (generally downregulated). These genes were also either not expressed at detectable levels or unchanged in expression at 12 hpi (data not shown).

Innate immunity and the inflammatory response continued to be absent from the significantly upregulated gene sets at 24 hpi (by GSEA [see Fig. S6A in the supplemental material]). *IRF1* was more strongly downregulated at 24 hpi than at 12 hpi, but other IRF genes (*IRF2*, *IRF3*, *IRF7*, and *IRF9*) continued to show no significant change in expression (Fig. 1; data not shown). In a separate experiment, we confirmed that *IRF1* was downregulated at 24 hpi by qPCR (by an average of 44%; data not shown). Other interferon (IFN)-inducible genes were differentially expressed at this time point but in different directions, such as *B2M* and *IFI16* (both downregulated) and *IFI30* and *ISG20* (both upregulated). This lack of concerted change may have offset the detection of interferon response gene sets as up- or downregulated at 24 hpi by GSEA. Instead, at 24 hpi, the related gene sets of cellular defense and immune response were found to be primarily downregulated (by GSEA [Fig. S6B]). Several of the downregulated genes associated with these gene sets were also associated with T cell activation; these genes include *CD1D*, *CD2*, *CD28*, *RAG1*, and *TNFRSF4* (Fig. S6D). Additional immune response-specific genes downregulated at 24 hpi included *IL4* and *IL7R*, suggesting that HIV-1 may have impaired the ability of infected cells to signal other cells or respond to extracellular cues.

Regulation of noncoding RNAs by HIV-1 infection. By mapping sequencing reads to RefSeq transcript annotation, we also observed changes in the expression of several non-protein-coding RNA species. Specifically, reads mapping to RefSeq transcripts that began with NR were considered noncoding. As before, expression of these transcripts was compared between infected and mock-infected samples. At 12 hpi, only one annotated noncoding

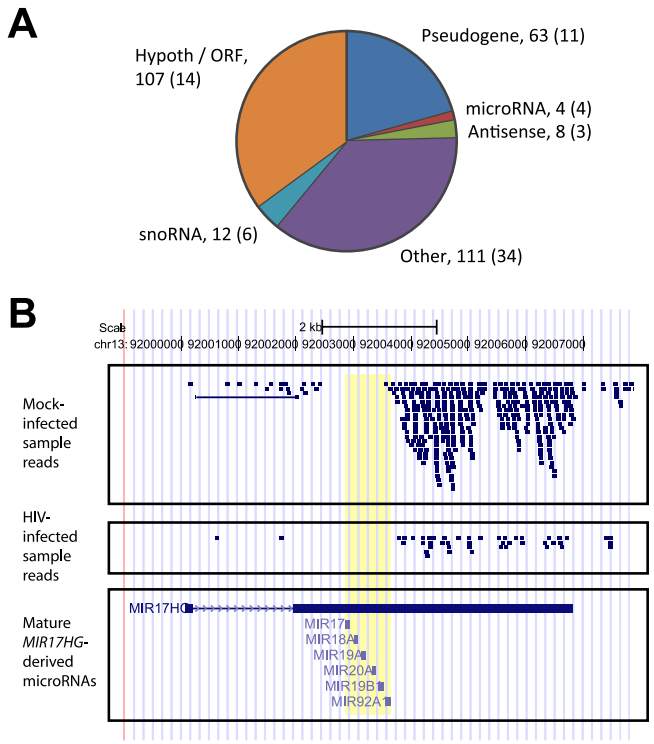

FIG 4 Noncoding RNAs regulated by HIV-1 infection in SUP-T1 cells as detected by NGS and reads mapping to one DE microRNA host gene. (A) Classes of RefSeq-annotated noncoding mRNAs that were differentially expressed at 24 hpi. Reads mapping to RefSeq transcripts beginning with "NR" were tabulated and classified by the indicated terms (Hypoth, hypothetical). Names not including one of the five terms listed were considered "other." The numbers refer to DE RefSeq transcripts (i.e., with Benjamini-Hochberg-adjusted $P < 0.05$), while the numbers in parentheses indicate the subset of DE transcripts with mean fold changes of 2 or greater. (B) Stack diagram showing reads mapping to microRNA host gene *MIR17HG* (NR_027350) as visualized in the UCSC Genome Browser. One representative replicate sample each for the mock- and HIV-infected samples is shown, and each blue square indicates a single mapped read at a given position. The known positions of the mature microRNAs derived from the *MIR17HG* gene is shown below the mapped reads with the yellow background denoting the range of microRNA sequence positions. chr13, chromosome 13.

RNA (NR_003697 encoding C7orf40) was found to be differentially expressed (Benjamini-Hochberg-adjusted $P < 0.05$). However, by 24 hpi, 305 annotated noncoding RNAs were found to be differentially expressed by the same criterion, with 73 changed by an average of 2-fold or greater (Fig. 4A). Several classes of noncoding RNAs were represented, including microRNA host genes, hypothetical genes and open reading frames (ORFs), small nucleolar RNAs (snoRNAs), and pseudogenes (Fig. 4A). DE pseudogenes often matched their protein-coding counterparts in directionality and degree of regulation, which suggests that regulatory regions were maintained and polyadenylated transcripts were produced (data not shown).

We also observed differential expression of four annotated microRNA host genes in infected cells, three of which were downregulated (*MIR17HG*, *MIR142*, and *MIR621*), while the remaining one was upregulated (*MIR518F*), all with 2-fold or greater changes. The downregulation of *MIR17HG* (by 2.25-fold) was of particular interest, as *MIR17HG* encodes a cluster of microRNAs, including miR-17, -18A, -19A, -19B, -20A, and -92 (Fig. 4B); in contrast, the other microRNA host genes encode only single microRNAs. Reads mapped to *MIR17HG* 3′ of the mature microRNA sequences, indicating that the postexcised polyadenylated product may have been detected (Fig. 4B). We also observed a concurrent upregulation of known targets of *MIR17HG*-encoded microRNAs, including *PCAF*, a host factor required for Tat-induced HIV-1 gene expression (25).

DISCUSSION

HIV infection and host noncoding RNA. In this study, we used NGS to measure all of the polyadenylated RNAs in $CD4^+$ T lymphoblasts infected with intact, replication-competent HIV. NGS offers a number of potential benefits to the study of viral infections; among them is the ability to detect non-protein-coding RNAs expressed during infection. In a previous study, we used NGS to detect long noncoding RNAs in the lungs of severe acute respiratory syndrome (SARS) coronavirus (CoV)-infected mice (8). Many of these RNAs were found to be differentially expressed, and unique signatures of infection could be identified (8).

In this study, we also detected several noncoding RNA species in HIV-infected cells that were differentially expressed (DE). In most cases, the functions of these noncoding RNAs and their relevance to infection remains unknown. One noncoding RNA of interest, the microRNA host gene *MIR17HG*, encoded a cluster of microRNAs and was strongly downregulated during infection. While our method of RNA library preparation precluded direct detection of mature microRNAs, the NGS data set allowed the expression of both regulators and targets of microRNAs to also be observed. For example, we observed a concurrent upregulation of the known target host gene *PCAF*, a cellular factor required for HIV replication (25). In contrast to Triboulet et al. (25) who observed that *PCAF* was upregulated at the protein level but not the mRNA levels in infected peripheral blood mononuclear cells (PBMCs), we found that *PCAF* was upregulated at the mRNA level in infected SUP-T1 cells. This may indicate alternative modes of *PCAF* regulation in different cell types. In addition, *MIR17HG*-encoded microRNAs have hundreds of other candidate targets in the TargetScan database (26). Several of these candidates were found to be coexpressed with *PCAF* in our data set, including *KLF3* (unpublished data). *KLF3* encodes a zinc finger transcription factor whose precise function is unknown but whose sequence is located proximally to a single-nucleotide polymorphism strongly associated with HIV-1 plasma levels (27). We also observed differential expression of factors that may have mediated *MIR17HG* downregulation. For example, O'Donnell et al. (28) found that c-Myc regulates expression of *MIR17HG*. Consistent with that study, we observed that *MYC* and *MIR17HG* were concordantly expressed, as *MYC* was downregulated at 12 and 24 hpi (Fig. 1). *MYC* suppression by HIV-1 may therefore underlie a number of subsequent of transcriptional events.

Early regulation of host transcription factors. We also observed discordance between the large amount of viral RNA present (~20% of total mappable RNA) and the limited amount of altered host gene expression at 12 hpi (~1% of detectable gene loci). That is, despite the use of host machinery and the presence of foreign mRNA, host transcription remained relatively unchanged. This result suggests that at 12 hpi, viral transcription occurred on top of largely undisturbed transcription of host genes. The host genes that we detected as differentially expressed at 12 hpi showed that T cell activation and differentiation were suppressed in in-

fected cells, likely via the downregulation of T cell activation-specific transcription factors (*BCL11B*, *IRF1*, *IKZF1*, and *SOX4*). The regulation of these and transcription factors specific for other functions (e.g., *EGR1*, *KLF13*, and *MYC*) may explain how HIV initiated replication with minimal disruption to host gene expression at 12 hpi but elicited larger-scale changes later in infection.

The suppression of T cell activation that we observed was consistent with previous studies on the response of T cells to infection by chemokine (C-X-C motif) receptor 4 (CXCR4)-tropic versus chemokine (C-C motif) receptor 5 (CCR5)-tropic HIV-1 strains (29–31). For example, Sirois et al. (31) found that key T cell activation-related genes, including *LCK*, were downregulated at 24 hpi by a CXCR4-tropic strain, whereas these same genes were upregulated by a CCR5-tropic strain. Our study utilized the CXCR4-tropic strain LAI, and it would be interesting to generate NGS data for a CCR5-tropic virus for comparison. Interestingly, our observations of small-magnitude differences in expression were consistent with those of Sirois et al. (31), who found the expression of T cell activation-related genes to be changed by 0.5-fold or less using reverse transcription-PCR (RT-PCR).

Large-scale disruptions to host transcription near peak viral replication. By 24 hpi, extensive reprogramming of the host transcriptome affecting a multitude of pathways had occurred. Remarkably, these large-scale changes occurred without concerted upregulation of innate immunity at the gene set level, suggesting that HIV-1 evaded viral sensing mechanisms. By this time point, the most affected pathways related to RNA fate determination, including RNA processing. While HIV-related RNA processing has been studied intensively, the contribution of most host factors remains incompletely defined (32). Modulation of this pathway by the virus may allow the proportion of unspliced, partially spliced, and fully spliced HIV RNA to be adjusted depending on the stage of the HIV life cycle. Our finding that over 150 genes related to RNA processing were differentially expressed suggests that HIV-1 infection results in a more complete modulation of RNA processing than previously identified (5, 33). However, our observed downregulation of RNA processing was consistent with results from an earlier study using NGS to investigate changes in $CD4^+$ T lymphoblasts infected with an HIV-based, nonreplication-competent vector at 24 hpi (34). Altered regulation of this and other pathways has been observed using microarrays as well, although in general, we observed changes in greater numbers of genes affecting these pathways using NGS (3, 5, 33, 35, 36).

Finally, we compared our results from NGS to other types of data available regarding HIV-infected cells, including protein-protein interaction (PPI) data and small interfering RNA (siRNA)-based screens. In an analysis of HIV-related PPI data, van Dijk et al. (37) identified sets of human proteins highly connected to host and viral proteins (37). Several proteins related to T cell activation were identified as highly connected; these proteins included CD4, CXCR4, and LCK (see Table S4 in the supplemental material). Downregulation of these members, as we observed, would therefore potentially affect the functions of many other proteins as well. Together, these data emphasize the high degree to which HIV-1 suppressed this pathway. In addition, many of the pathways associated with DE genes were identified as essential for HIV-1 replication in a recent meta-analysis of siRNA-based screen data (38). These pathways included DNA binding, RNA splicing, RNA export, nuclear transport, and protein complex assembly (Table S3). While the siRNA data suggest that HIV-1 initially requires these pathways to be active to establish an infection, our data suggest that HIV-1 also later suppresses and inactivates these pathways during active replication.

In conclusion, using NGS, we have reported a number of unique findings in HIV-infected cells: the direct measurement of viral and host RNA, the detection of small changes in the abundance of mRNA transcripts, the identification of novel viral RNA splice events, and the assay of multiple forms of noncoding RNAs, including insights into the regulation of microRNAs. These observations give additional insights into the panoply of changes that occur in host cells infected with HIV and provide the groundwork for using new sequencing technologies in future studies investigating the host response to viral infection.

MATERIALS AND METHODS

Cell lines and virus infection. SUP-T1 cells were obtained from American Type Culture Collection (CRL-1942) and propagated in RPMI 1640 medium (Gibco) supplemented with 10% fetal bovine serum (HyClone), penicillin (100 U/ml), streptomycin (100 μg/ml), and GlutaMAX-I. HIV-1 LAI strain (catalog no. 2522) was obtained from the NIH AIDS Research and Reference Reagent Program (Germantown, MD) and propagated in SUP-T1 cells. U373-MAGI-$CXCR4_{CEM}$ cells were obtained from M. Emerman through the AIDS Research and Reference Reagent Program, and the virus titer in these cells was measured by the protocol of Deminie and Emerman (39). Typical titers reached 10^7 infectious units per ml. Infections were carried out at a multiplicity of infection (MOI) of 5 and performed in triplicate. Mock-infected samples received SUP-T1 cell conditioned medium and were also performed in triplicate. The infectious dose was optimized to achieve 100% infected cells at 24 hpi with ~50% cell viability as measured by trypan blue exclusion assay. Infected cells were visualized by immunofluorescence assay with rabbit HIV-1_{SF2} p24 antiserum kindly provided by BioMolecular Technologies through the AIDS Research and Reference Reagent Program.

RNA preparation and next-generation sequencing. Total RNA was extracted from 5×10^6 cells per sample using a mirVana kit (Applied Biosystems/Ambion, Austin, TX), and the quality and concentration of the RNA were determined by an Agilent 2100 Bioanalyzer. Samples were submitted for sequencing to IlluminaFastTrack Sequencing Services (Hayward, CA). cDNA libraries were generated using Illumina mRNA-SEQ kit. The quality and concentration of these libraries were determined by an Agilent 2100 Bioanalyzer. The libraries were clonally amplified on a cluster generation station using Illumina version 4 cluster generation reagents to achieve a target density of approximately 300,000 (300K)/mm^2 in a single channel of a flow cell. The resulting libraries were then sequenced on a Genome Analyzer IIx using Illumina version 4.0 sequencing reagents which generated single reads of 75 nucleotides (nt). Image analysis, base calling, and error estimation were performed using Illumina Analysis Pipeline (version 2.6).

Read mapping and transcript quantification. We mapped the 75-nt reads to human ribosomal sequences using the short-read aligner software Bowtie to remove potential rRNA sequences (40). We then mapped the remaining unmapped reads to the HIV genome (GenBank accession no. K02013) using the gapped aligner software TopHat, which predicts HIV splicing junctions and maps intron-spanning reads to known splicing junctions (41). To quantify transcript expression, we mapped all reads that remained unmapped to the human reference genome (hg19, build GRCh37, downloaded from UCSC genome browser, http://genome.ucsc.edu) using the gapped aligner software TopHat. RefSeq transcript annotations were supplied to facilitate the mapping of reads spanning known splicing junctions. On the basis of these human genome mapping results, we then estimated the levels of expression at both the transcript and locus levels for RefSeq-annotated genes using the transcript assembly software Cufflinks (42). Read sequences that mapped to more than one genomic location were excluded from expression quantification. For visualization,

BAM files were generated using TopHat and SAMtools (43) and displayed using the UCSC Genome Browser.

Splice site variant detection and testing. The positions of known viral splice sites were found in the literature (44), and the corresponding sequences were identified in strain pNL4-3 (GenBank accession no. AF003887). Based on sequence identity to strain LAI, a list of known LAI-specific splice sites was generated and compared to the NGS TopHat output. Splice patterns involving splice sites SD 1 and SA 2/2* were tested by quantitative PCR (qPCR) using primers PSK027F (5′ CAGGGACTTG AAAGCGAAAG 3′; locations 193 to 212 in HIVBRUCG accession no. K02013) and PSK027R (5′ TGGGGCTTGTTCCATCTATC 3′; locations 5136 to 5155).

Quantitative reverse transcription (RT-PCR). RNA was reverse transcribed using the QuantiTect reverse transcription kit (Qiagen, Valencia, CA). The resulting cDNA samples were diluted 50×. ABI TaqMan assays were run for each sample in triplicate (see Fig. S3C in the supplemental material). Relative expression was calculated using the $\Delta\Delta C_T$ method with averaged ΔC_T values (where C_T stands for threshold cycle) for the *OAZ1* gene as a calibrator, as the expression of *OAZ1* did not significantly change between 12 and 24 hpi in the NGS data. Intracellular viral RNA load was quantified as previously described (45). Relative change of *IRF1* was determined by qPCR using primers PSK1 (forward primer [5′ TCTG GCTTTTTCCTCTGAGC 3′] and reverse primer [5′ ATGCTTTTCTGG GGTCACTG 3′]).

Differential expression analysis. To compare transcript expression across different conditions, we first normalized transcript abundances by the following methodology. Transcript abundance was quantified as FPKM (fragments per kilobase of exon per million mapped fragments) values estimated by Cufflinks. We chose one sample arbitrarily as a reference. Distributions of $\log_2$-transformed FPKM values between the reference and remaining samples were compared by quantile-quantile plots. We determined the scaling factor for each sample as the median difference of the corresponding quantile values of the sample and reference. Only genes/transcripts with raw FPKM values of ≥1 in all samples were considered in the estimation of scaling factors. We retained those genes/transcripts with nonzero FPKM values in 100% of the samples of at least one biological condition (our detection criterion). An offset of 1 was added to all normalized values to facilitate the comparisons involving one or more FPKM values of zero and to reduce the variability of the log ratios for low expression values. Transcripts were mapped to RefSeq gene loci, resulting in 9,992 loci with detectable reads. The data are available at http://www.viromics.washington.edu or upon request. The normalized expression data were analyzed for differential expression by using linear model methods as implemented in the R package limma (46). *P* values were derived from linear model-based *t* tests between infected and time-matched mock-infected samples. Unless otherwise noted, we defined differential expression by Benjamini-Hochberg-adjusted *P* values of less than 0.05 based on the assumption that a false discovery rate of 5% provided an acceptable balance of false-positive control and statistical power. Fold changes (FC) were derived from comparing the means of these groups, and multiple groupings of fold changes were used (1.0 to 1.5 FC, 1.5 to 2.0 FC, and 2.0+ FC) based on previous observed fold change ranges observed in high-throughput and in particular NGS data in virus-infected systems (8).

DAVID and GSEA analyses. To identify annotations among DE genes, we used the NIH DAVID resource (47, 48). Default settings were used in DAVID with GO_BP_FAT, GO_CC_FAT, GO_MP_FAT, BIOCARTA, and KEGG_PATHWAY gene set annotations. Complementary analysis was performed using gene set enrichment analysis (GSEA) (11). Default settings were used in GSEA with Gene Ontology (GO) categories (c5.all.v2.5 gene sets) and Biocarta and KEGG pathways (c2.all.v2.5 gene sets) and 5,000 gene set-based permutations. Leading-edge analysis was performed within GSEA to derive genes for hierarchically clustering up- and downregulated gene sets.

Network analysis. Interactions between DE genes were identified using Ingenuity Pathways Analysis (IPA) software (Ingenuity Systems, Redwood City, CA). Networks were generated within IPA based on direct, literature-curated interactions. Subsets of genes were selected for input into IPA based on DE gene overlaps with DAVID annotation clusters. Heat maps were generated using the R package gplots. Comparisons to published networks of protein-protein interaction (PPI) data were made using Cytoscape (49).

ACKNOWLEDGMENTS

This work was funded by Public Health Service grants P30DA015625 and P51RR000166 from the National Institutes of Health.

We thank Matthew Thomas and Richard Green for technical assistance and Marcus Korth for editorial assistance. We also thank Peter Sloot and David van Dijk for generously sharing expertise related to protein-protein interaction networks.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00134-11/-/DCSupplemental.

FIGURE S1, PDF file, 0.1 MB.
FIGURE S2, PDF file, 0.1 MB.
FIGURE S3, PDF file, 0.1 MB.
FIGURE S4, PDF file, 0.1 MB.
FIGURE S5, PDF file, 0.1 MB.
FIGURE S6, PDF file, 0.3 MB.
TABLE S1, PDF file, 0.1 MB.
TABLE S2, PDF file, 0.1 MB.
TABLE S3, PDF file, 0.1 MB.
TABLE S4, PDF file, 0.1 MB.

REFERENCES

1. **Giri MS, Nebozhyn M, Showe L, Montaner LJ.** 2006. Microarray data on gene modulation by HIV-1 in immune cells: 2000–2006. J. Leukoc. Biol. **80**:1031–1043.
2. **Wang WK, Chen MY, Chuang CY, Jeang KT, Huang LM.** 2000. Molecular biology of human immunodeficiency virus type 1. J. Microbiol. Immunol. Infect. **33**:131–140.
3. **Geiss GK, et al.** 2000. Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. Virology **266**:8–16.
4. **Li Y, Chan EY, Katze MG.** 2007. Functional genomics analyses of differential macaque peripheral blood mononuclear cell infections by human immunodeficiency virus-1 and simian immunodeficiency virus. Virology **366**:137–149.
5. **van't Wout AB, et al.** 2003. Cellular gene expression upon human immunodeficiency virus type 1 infection of CD4(+)-T-cell lines. J. Virol. **77**:1392–1402.
6. **Lederer S, et al.** 2009. Transcriptional profiling in pathogenic and nonpathogenic SIV infections reveals significant distinctions in kinetics and tissue compartmentalization. PLoS Pathog. **5**:e1000296.
7. **Bradford JR, et al.** 2010. A comparison of massively parallel nucleotide sequencing with oligonucleotide microarrays for global transcription profiling. BMC Genomics **11**:282.
8. **Peng X, et al.** 2010. Unique signatures of long noncoding RNA expression in response to virus infection and altered innate immune signaling. mBio **1**:e00206-10.
9. **Cismasiu VB, et al.** 2008. BCL11B is a general transcriptional repressor of the HIV-1 long terminal repeat in T lymphocytes through recruitment of the NuRD complex. Virology **380**:173–181.
10. **Chen N, et al.** 2006. HIV-1 down-regulates the expression of CD1d via Nef. Eur. J. Immunol. **36**:278–286.
11. **Subramanian A, et al.** 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. **102**:15545–15550.
12. **Cooper TA, Wan L, Dreyfuss G.** 2009. RNA and disease. Cell **136**:777–793.
13. **Damgaard CK, Tange TO, Kjems J.** 2002. hnRNP A1 controls HIV-1 mRNA splicing through cooperative binding to intron and exon splicing silencers in the context of a conserved secondary structure. RNA **8**:1401–1415.

14. **Jablonski JA, Caputi M.** 2009. Role of cellular RNA processing factors in human immunodeficiency virus type 1 mRNA metabolism, replication, and infectivity. J. Virol. **83**:981–992.
15. **Buratti E, Baralle FE.** 2010. The multiple roles of TDP-43 in pre-mRNA processing and gene expression regulation. RNA Biol. **7**:420–429.
16. **Ou SH, Wu F, Harrich D, García-Martínez LF, Gaynor RB.** 1995. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. J. Virol. **69**:3584–3596.
17. **Sendtner M.** 2011. TDP-43: multiple targets, multiple disease mechanisms? Nat. Neurosci. **14**:403–405.
18. **Swanson CM, Puffer BA, Ahmad KM, Doms RW, Malim MH.** 2004. Retroviral mRNA nuclear export elements regulate protein function and virion assembly. EMBO J. **23**:2632–2640.
19. **Cullen BR.** 2003. Nuclear mRNA export: insights from virology. Trends Biochem. Sci. **28**:419–424.
20. **Lee SB, Park J, Jung JU, Chung J.** 2005. Nef induces apoptosis by activating JNK signaling pathway and inhibits NF-kappaB-dependent immune responses in *Drosophila*. J. Cell Sci. **118**:1851–1859.
21. **Andersen J, VanScoy S, Cheng T-F, Gomez D, Reich NC.** 2008. IRF-3-dependent and augmented target genes during viral infection. Genes Immun. **9**:168–175.
22. **Grandvaux N, et al.** 2002. Transcriptional profiling of interferon regulatory factor 3 target genes: direct involvement in the regulation of interferon-stimulated genes. J. Virol. **76**:5532–5539.
23. **Yan N, Regalado-Magdos AD, Stiggelbout B, Lee-Kirsch MA, Lieberman J.** 2010. The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. Nat. Immunol. **11**:1005–1013.
24. **Doehle BP, Hladik F, McNevin JP, McElrath MJ, Gale M, Jr..** 2009. Human immunodeficiency virus type 1 mediates global disruption of innate antiviral signaling and immune defenses within infected cells. J. Virol. **83**:10395–10405.
25. **Triboulet R, et al.** 2007. Suppression of microRNA-silencing pathway by HIV-1 during virus replication. Science **315**:1579–1582.
26. **Friedman RC, Farh KK-H, Burge CB, Bartel DP.** 2009. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. **19**: 92–105.
27. **Dalmasso C, et al.** 2008. Distinct genetic loci control plasma HIV-RNA and cellular HIV-DNA levels in HIV-1 infection: the ANRS Genome Wide Association 01 study. PLoS One **3**:e3907.
28. **O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT.** 2005. c-Myc-regulated microRNAs modulate E2F1 expression. Nature **435**: 839–843.
29. **Cicala C, et al.** 2006. R5 and X4 HIV envelopes induce distinct gene expression profiles in primary peripheral blood mononuclear cells. Proc. Natl. Acad. Sci. U. S. A. **103**:3746–3751.
30. **Locher CP, et al.** 2005. Differential effects of R5 and X4 human immunodeficiency virus type 1 infection on $CD4^+$ cell proliferation and activation. J. Gen. Virol. **86**:1171–1179.
31. **Sirois M, et al.** 2008. R5 and X4 HIV viruses differentially modulate host gene expression in resting $CD4^+$ T cells. AIDS Res. Hum. Retroviruses **24**:485–493.
32. **Stoltzfus CM.** 2009. Regulation of HIV-1 alternative RNA splicing and its role in virus replication. Adv. Virus Res. **74**:1–40.
33. **Imbeault M, Ouellet M, Tremblay MJ.** 2009. Microarray study reveals that HIV-1 induces rapid type-I interferon-dependent p53 mRNA up-regulation in human primary $CD4^+$ T cells. Retrovirology **6**:5.
34. **Lefebvre G, et al.** 2011. Analysis of HIV-1 expression level and sense of transcription by high-throughput sequencing of the infected cell. J. Virol. **85**:6205–6211.
35. **Corbeil J, et al.** 2001. Temporal gene regulation during HIV-1 infection of human $CD4^+$ T cells. Genome Res. **11**:1198–1204.
36. **Vahey MT, et al.** 2002. Impact of viral infection on the gene expression profiles of proliferating normal human peripheral blood mononuclear cells infected with HIV type 1 RF. AIDS Res. Hum. Retroviruses **18**: 179–192.
37. **van Dijk D, Ertaylan G, Boucher CA, Sloot PM.** 2010. Identifying potential survival strategies of HIV-1 through virus-host protein interaction networks. BMC. Syst. Biol. **4**:96.
38. **Bushman FD, et al.** 2009. Host cell factors in HIV replication: meta-analysis of genome-wide studies. PLoS Pathog. **5**:e1000437.
39. **Deminie CA, Emerman M.** 1995. Quantitation of virus stocks produced from cloned human immunodeficiency virus DNA. Methods Mol. Genet. **7**:195–208.
40. **Langmead B, Trapnell C, Pop M, Salzberg SL.** 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. **10**:R25.
41. **Trapnell C, Pachter L, Salzberg SL.** 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25**:1105–1111.
42. **Trapnell C, et al.** 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. **28**:511–515.
43. **Li H, et al.** 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**:2078–2079.
44. **Purcell DF, Martin MA.** 1993. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. J. Virol. **67**:6365–6378.
45. **Li C-C, Seidel KD, Coombs RW, Frenkel LM.** 2005. Detection and quantification of human immunodeficiency virus type 1 p24 antigen in dried whole blood and plasma on filter paper stored under various conditions. J. Clin. Microbiol. **43**: 3901–3905.
46. **Smyth GK.** 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. **3**: Article3.
47. **Huang DW, Sherman BT, Lempicki RA.** 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. **37**:1–13.
48. **Huang DW, Sherman BT, Lempicki RA.** 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. **4**:44–57.
49. **Shannon P, et al.** 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. **13**: 2498–2504.