

Global Analysis of Host-Pathogen Interactions that Regulate Early-Stage HIV-1 Replication

Renate König,¹ Yingyao Zhou,^{6,8} Daniel Elleder,^{2,8} Tracy L. Diamond,^{5,8} Ghislain M.C. Bonamy,^{6,7} Jeffrey T. Ireland,⁶ Chih-yuan Chiang,⁶ Buu P. Tu,⁶ Paul D. De Jesus,¹ Caroline E. Lilley,³ Shannon Seidel,² Amanda M. Opaluch,¹ Jeremy S. Caldwell,⁶ Matthew D. Weitzman,³ Kelli L. Kuhen,⁶ Sourav Bandyopadhyay,⁴ Trey Ideker,⁴ Anthony P. Orth,⁶ Loren J. Miraglia,⁶ Frederic D. Bushman,⁵ John A. Young,^{2,*} and Sumit K. Chanda^{1,*}

¹Infectious & Inflammatory Disease Center, Burnham Institute for Medical Research, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA

²Infectious Disease Laboratory

³Laboratory of Genetics

The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA

⁴Department of Bioengineering, University of California, San Diego, La Jolla, CA 92037, USA

⁵Department of Microbiology, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA

⁶The Genomics Institute of the Novartis Research Foundation, 10675 John J. Hopkins Drive, San Diego, CA 92121, USA

⁷Hudson-Alpha Institute for Biotechnology, 127 Holmes Avenue, Huntsville, AL 35801, USA

⁸These authors contributed equally to this work

*Correspondence: jyoung@salk.edu (J.A.Y.), schanda@burnham.org (S.K.C.)

DOI 10.1016/j.cell.2008.07.032

SUMMARY

Human Immunodeficiency Viruses (HIV-1 and HIV-2) rely upon host-encoded proteins to facilitate their replication. Here, we combined genome-wide siRNA analyses with interrogation of human interactome databases to assemble a host-pathogen biochemical network containing 213 confirmed host cellular factors and 11 HIV-1-encoded proteins. Protein complexes that regulate ubiquitin conjugation, proteolysis, DNA-damage response, and RNA splicing were identified as important modulators of early-stage HIV-1 infection. Additionally, over 40 new factors were shown to specifically influence the initiation and/or kinetics of HIV-1 DNA synthesis, including cytoskeletal regulatory proteins, modulators of post-translational modification, and nucleic acid-binding proteins. Finally, 15 proteins with diverse functional roles, including nuclear transport, prostaglandin synthesis, ubiquitination, and transcription, were found to influence nuclear import or viral DNA integration. Taken together, the multiscale approach described here has uncovered multiprotein virus-host interactions that likely act in concert to facilitate the early steps of HIV-1 infection.

INTRODUCTION

Over the course of the last several decades, a number of host cell proteins that influence the early steps of retroviral replication

have been defined (Goff, 2007). However, it is likely that many other cellular factors and processes are exploited by the virus during these stages, which involves uncoating steps that give rise to an active reverse transcription complex (RTC), movement of the viral preintegration complex (PIC) to the cell nucleus, and then integration of the linear viral DNA into a host cell chromosome to generate the provirus (Nisole and Saib, 2004). Recently, a genome-wide siRNA analysis revealed over 250 host cellular factors that influence HIV-1 infection (Brass et al., 2008). Notably, only a small fraction of these factors was proposed to influence the early stages of HIV-1 replication, making it likely that additional cellular factors that regulate these steps remain to be identified. Here, we present a genome-wide analysis of virus-host interactions affecting the early steps of HIV-1 infection.

RESULTS

In this study, we combined several genome-wide analytical methods to characterize the host factors required in the early steps of HIV infection. We performed genome-wide RNAi screens for genes required for infection by utilizing a single-cycle HIV-1 reporter virus engineered to encode luciferase and bearing the Vesicular Stomatitis Virus Glycoprotein (VSV-G) on its surface to facilitate efficient infection (Figure 1A). As controls, parallel screens were performed with other viral vectors encoding luciferase: (1) Moloney murine leukemia virus (MuLV) vector pseudotyped with VSV-G, and (2) an adeno-associated virus (AAV) vector (Figure S1 available online). Prior to infection, human 293T cells were transfected with an arrayed genome-wide siRNA library, which targets ~20,000 human genes. Typically, six unique siRNAs were used to interrogate each gene, with two siRNAs targeting the same gene arrayed in a single well (3 wells/gene)

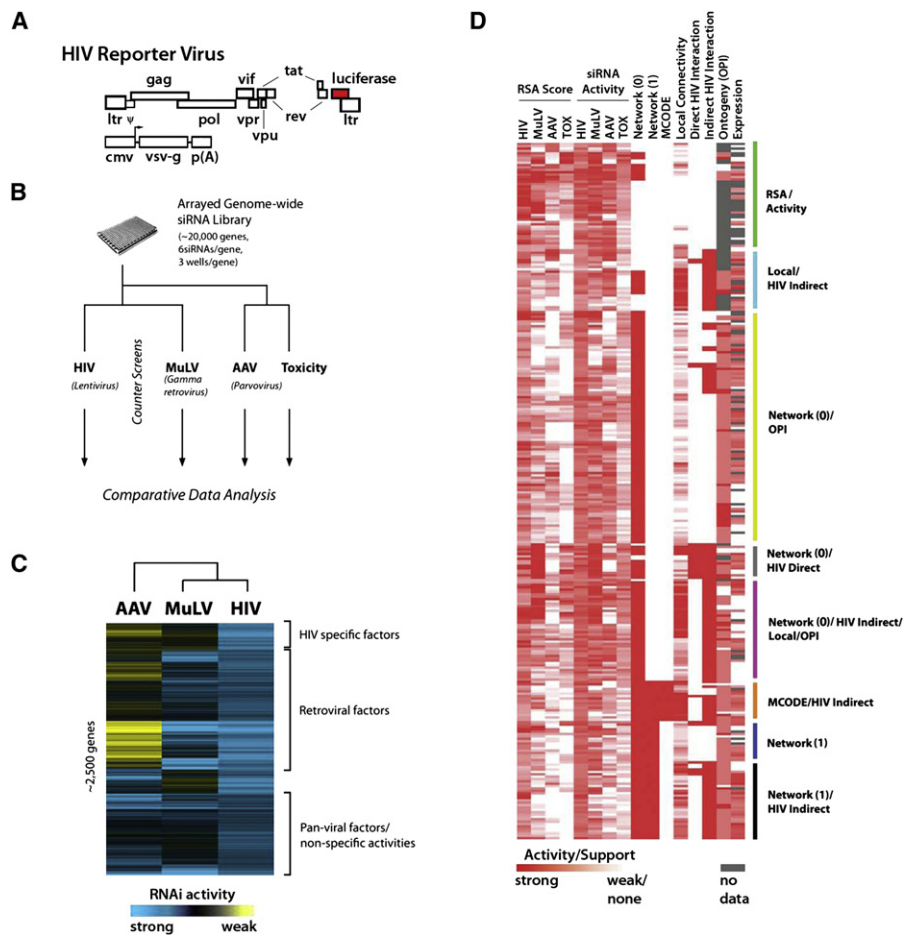


Figure 1. Integrative Analysis of HIV-Host Interactions

(A) To monitor the early stages of HIV replication, we carried out infections by using a single cycle, VSV-G-pseudotyped, HIV-1 vector encoding luciferase (pNL43R⁺E⁻ luc (VSV-G)).

(B) Human 293T cells were transfected with an arrayed genome-wide siRNA library and then challenged with HIV (VSV-G) encoding the luciferase gene. Infection was monitored by measuring luciferase activity. Additionally, counterscreens were performed to identify those host factors affecting gammaretrovirus (MuLV) or parvovirus (AAV) infection, as well as cellular viability (toxicity).

(C) Inhibition data for each gene were normalized, scaled, and enriched for genes that have at least two siRNAs supporting their activity (RSA analysis). Comparative analysis of all viral screens, after filtering for toxicity, was visualized through hierarchical clustering.

(D) To facilitate the validation of biologically relevant functional and biochemical activities, we employed a multiscale approach to select candidate genes for further study. In addition to identifying genes based on either gene (RSA) or single siRNA activity in the genome-wide assay, an “evidence score” was compiled for each gene based upon a variety of other supporting criteria. Specifically, proteins, which were encompassed in statistically significant host-protein interaction networks (Network 0, Network 1, MCODE), or had significant network connectivity with other putative host factors (local connectivity) or HIV-encoded proteins (HIV Direct/Indirect interaction), were given additional consideration. Furthermore, genes identified through the genome-wide assay, which were members of overrepresented functional groups (ontology/OPI), or had coincident expression with CD4/CXCR4 or CD4/CXCR5 were also given additional weight. The plot depicts the activities and support evidence of 295 genes for which activities were subsequently confirmed by two or more siRNAs (Table S3).

(Figure 1B). The toxicity associated with each pair of siRNAs was also measured by assaying viable cell numbers (Figure 1B).

Since the primary screen was executed in an arrayed format, we were able to employ Redundant siRNA Analysis (RSA) to identify genes that were significantly inhibited by at least two independent siRNAs, many of which possessed reproducible, but moderate, activities (45% or greater reduction of HIV infectivity) (Konig et al., 2007). This analysis revealed that a significant fraction of host proteins required by HIV-1 is also required by MuLV (80%) (Figure 1C). These proteins included a number of

factors that are already known to affect retroviral replication (Brass et al., 2008; Goff, 2007) (Table S1).

To characterize the interactions between HIV-1 and these host cell proteins, we developed an integrative approach that considered additional lines of functional, biochemical, or transcriptional data (collectively called “evidence score”) for candidate gene verification. This approach is based on the assumption that a cellular factor is more likely to be a proximal regulator of viral replication if its activity is supported by multiple independent lines of evidence.

Using a yeast two-hybrid (Y2H) human protein-protein interaction database (Hynet) (Mukherji et al., 2006), we identified an extended interactome network, which contained 2458 putative host cellular factors affecting HIV infection that form 120,211 direct or indirect interactions (Network 0; data not shown). This network map was further refined by removing protein clusters, which influenced infection by the AAV vector or were associated with cellular toxicity (Network 1; $p < 0.001$) (Figure S2A). Additionally, we applied a graph theoretic clustering algorithm (MCODE) to Network 1 to identify regions of locally enhanced connectivity (Figures S2B–S2G) (Bader and Hogue, 2003). Finally, since the Hynet Y2H data set employed in these analyses did not contain virus-host protein-protein interactions, we integrated protein interaction relationships contained in the NIAID HIV-1, Human Protein Interaction Database (<http://www.ncbi.nlm.nih.gov/RefSeq/HIVInteractions/>), as well as newly generated HIV/host Y2H interaction data in which each of the HIV-encoded proteins was tested individually against a library of human proteins (see Experimental Procedures). The resulting host-pathogen interaction map between Network 0 proteins and HIV-encoded proteins was determined to be of high statistical significance (direct interactions, $p = 0.005$; indirect interactions, $p = 10^{-103}$, see Supplemental Data).

In an effort to identify those genes most relevant to HIV-1 infection, the mRNA expression profiles of the identified host genes, extracted from the GNF Expression database (<http://symatlas.gnf.org>), were also correlated to those of the viral receptor CD4 and coreceptors CXCR4 and CCR5. The majority of host proteins in Network 0 displayed strong expression patterns in tissues of lymphoid and neuronal origin (Figure S3). In a subsequent analysis for each candidate host factor, we calculated correlation coefficients with CD4/CXCR4 and CD4/CCR5 coexpression and incorporated this value into the evidence score (see Experimental Procedures). Finally, to correlate gene functions, processes, components, or domains with gene activities across the four assays previously described (siRNA screens against the three viral vectors, and cytotoxicity), we employed an ontology-based pattern identification algorithm (OPI) and identified over 100 enriched activity signatures across the four genome-wide screens (Figure S4; Table S2). Since OPI activity profiles are supported by multiple functionally related genes and are statistically reliable ($p < 0.05$), they were also considered as favorable evidence for the elucidation of physiologically relevant host cellular factors.

We combined these lines of evidence into a decision matrix to prioritize those factors that are likely regulators of HIV-1 replication. Criteria included gene-based (RSA score) and siRNA-based activity, gene expression signatures, gene ontology data (OPI), cellular protein-protein interaction (using Network 0, Network 1, MCODE, and local connectivity), and virus-host interaction data (HIV Direct/Indirect) (Figures 1D and S2; Table S2; data not shown). This enabled us to identify ~800 strongly supported genes based on evidence scores; 295 of these genes have been confirmed to inhibit HIV (VSV-G) infection by more than 45% with at least 2 siRNAs while not significantly affecting cellular viability (Table S3). Most confirmed factors showed broad expression in cell types relevant to HIV pathogenesis (Figure S3), and, in addition, 75 of the confirmed factors share highly coinci-

dent expression patterns with CD4/CXCR4 or CD4/CCR5 (Figures 2A and 3).

Next, we reanalyzed these confirmed factors to identify over-represented biological annotations or protein families based on gene ontology or Interpro (IPR) databases, respectively ($p < 0.01$; Figure S5; Table S4). This analysis elucidated 176 statistically enriched biological terms, including over 25 functional processes or protein domains that influence HIV (VSV-G) infection (Figure 2B). Since functional overrepresentation was one of the criteria used to select genes for further study, many of these classifications reflected functional categories identified in the OPI analysis conducted across the four genome-wide screens (Figure S4; Table S2). Among others, a collection of genes involved in DNA repair and nucleoside diphosphate (NUDIX) hydrolase activity as well as members of the Tripartite Motif (TRIM) family of proteins were implicated as factors likely to be important for early HIV (VSV-G) infection (Figure 2C).

To visualize the global topography of HIV-host interactions, we reconstructed a network based upon the 295 HIV host cellular factors confirmed in secondary assays. Using Hynet, as well as additional protein-protein interaction databases (see Experimental Procedures), we assessed the network connectivity of these confirmed factors and additional host and HIV-encoded proteins (Figure 3A, inset; Figure S6; Table S5). This host-pathogen interaction network contained 213 functionally validated and 169 predicted HIV host cellular nodes, which were internally connected via 2291 binary protein interactions, and 318 interactions to HIV-encoded proteins ($p < 3 \times 10^{-6}$). To identify densely connected local network neighborhoods, we executed a Molecular Complex Detection (MCODE) analysis, which revealed several subnetworks with high local network connectivity (Figure 3A). Strikingly, the functional activities of these protein clusters reflect classifications identified by using the GO and OPI analysis, including the ubiquitin-proteasome pathway, DNA transcription and repair, and nucleic acid binding, further underscoring their importance in the direct regulation of early-stage HIV replication (Figures 3B–3I).

Quantitative PCR analysis of viral DNA products was performed to determine the effects of siRNAs on specific steps of early HIV-1 replication (Figure 4A, left panel). To analyze the accumulation of reverse transcription products, cells transfected with siRNAs targeting confirmed host factors were infected with the VSV-G-pseudotyped HIV-1 vector, and total DNA was extracted for analysis at either 12 hr or 24 hr postinfection. A reduction in viral DNA indicates that the targeted genes were acting either to promote synthesis or inhibit degradation. Control tests performed with inactivated virus or with an inhibitor of reverse transcription (AZT) showed strong inhibition at the expected steps of virus replication (Figure 4B, top panel).

Analysis of the siRNA activities indicates that factors affecting reverse transcription can be segregated into at least two distinct functional classes, those that are absolutely required for viral DNA synthesis (AKAP13 to SPEN) and those that appear to affect the kinetics of viral DNA synthesis (MED7 to SUMO2) (Figure 4A, right panel; Table S6). A third set of proteins was found to affect HIV-1 uncoating or reverse transcription (AP1G2 to TRIAD3), but, based upon statistical criteria, could not be placed into either functional class (Figure 4A, right panel). We hypothesize that factors that fall within the first group influence a step that is

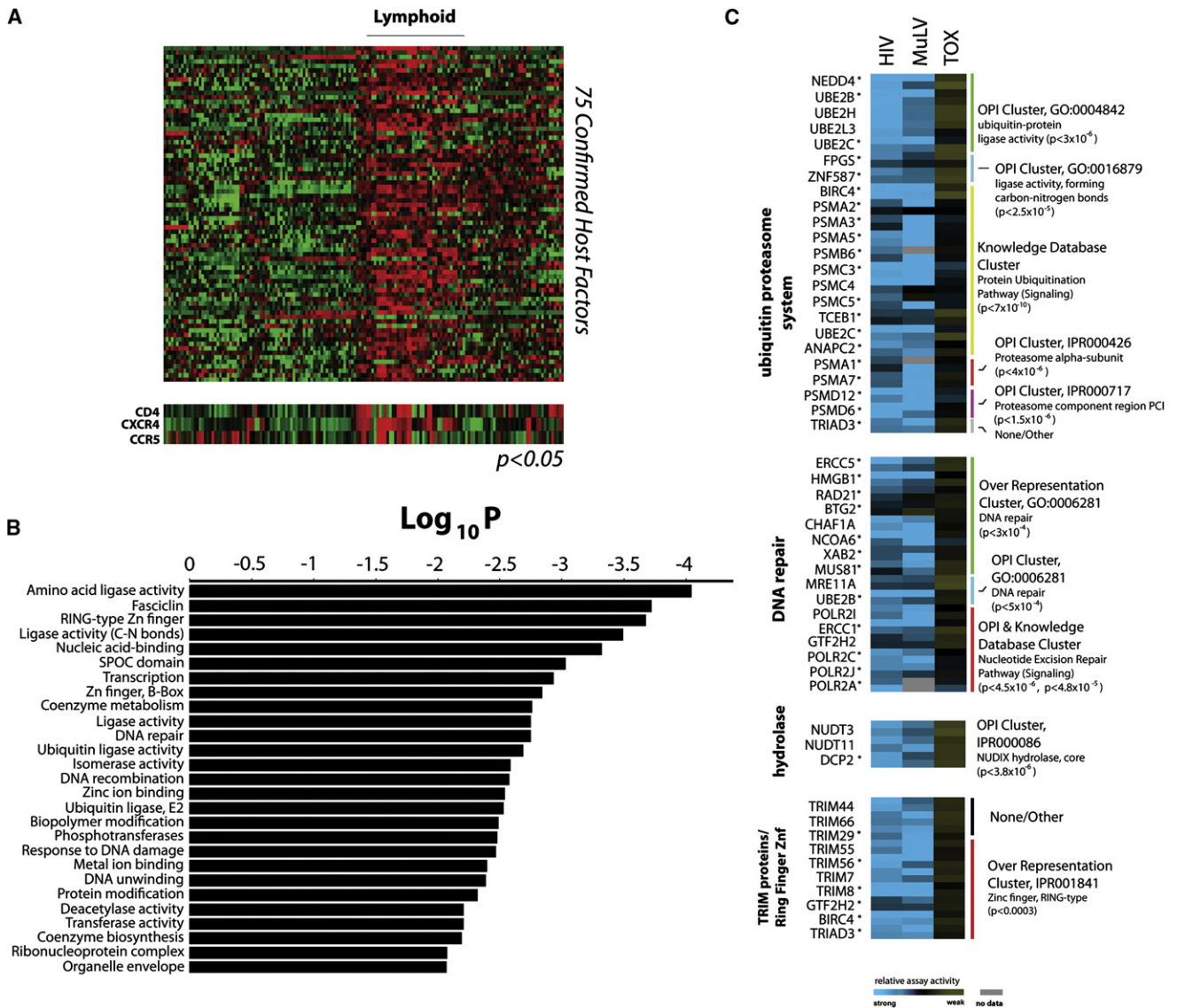


Figure 2. Characterization of Confirmed Factors Required for Infection by the VSV-G-Pseudotyped HIV-1 Vector

(A) Tissue expression of 75 confirmed host cellular factors with statistically coincident expression profiles to HIV receptor/coreceptors CD4/CXCR4 or CD4/CXCR5 ($p < 0.05$, standard chi-square test, $R_{CD4,CXCR4} \geq 0.18$, $R_{CD4,CCR5} \geq 0.13$). Tissues and cell types are shown in the x axis, while genes are displayed on the y axis, using a continuum of green (low expression) to red (high expression). mRNA transcriptional signatures were extracted from the Genomics Institute of the Novartis Research Foundation expression database (<http://symatlas.gnf.org>).

(B) Statistically significant overrepresentation of functional classes and protein families based upon gene ontology (GO) and interpro (IPR) domain mapping of 295 confirmed host factors required for HIV infection (Table S4). P values were calculated by an accumulated hypergeometric distribution function (Zar, 1999).

(C) Relative activities of confirmed genes, each represented by two active siRNAs, across HIV, MuLV, and cytotoxicity assays (TOX) are shown from strong (blue) to weak (yellow) (Figure S4; Table S2). Functional classes are derived from an ontology-based pattern identification algorithm (OPI & Knowledge Database Clusters) or GO/IPR overrepresentation analysis of primary screening data. All measurements represent the mean of at least four assays, and activities not tested or unconfirmed are depicted in gray. To identify those factors most likely to influence the intracellular steps of HIV-1 replication, a subgroup of genes was further assayed and confirmed to inhibit infection by an HIV-1 vector pseudotyped with the 10A1 MuLV envelope protein, a viral envelope protein that directs pH-independent viral entry at the plasma membrane. These genes are indicated with an asterisk (*) (also see Table S3).

important for capsid uncoating or initiation of reverse transcription, or else inhibit degradation of the viral DNA after synthesis. Conversely, we predict that the siRNAs, which alter the kinetics of reverse transcription, target cellular cofactors that facilitate, but are not essential to, the process (Figure 4A). Additional studies will be required to determine if delayed viral DNA synthesis

in the absence of these genes produces PICs that are competent for integration. However, we anticipate that these kinetic factors are likely to play critical roles in maintaining viral fitness in vivo.

Additional PCR-based assays were used to identify those factors that play roles in either nuclear import of the viral PIC or viral DNA integration. The amount of 2-LTR circular forms of viral

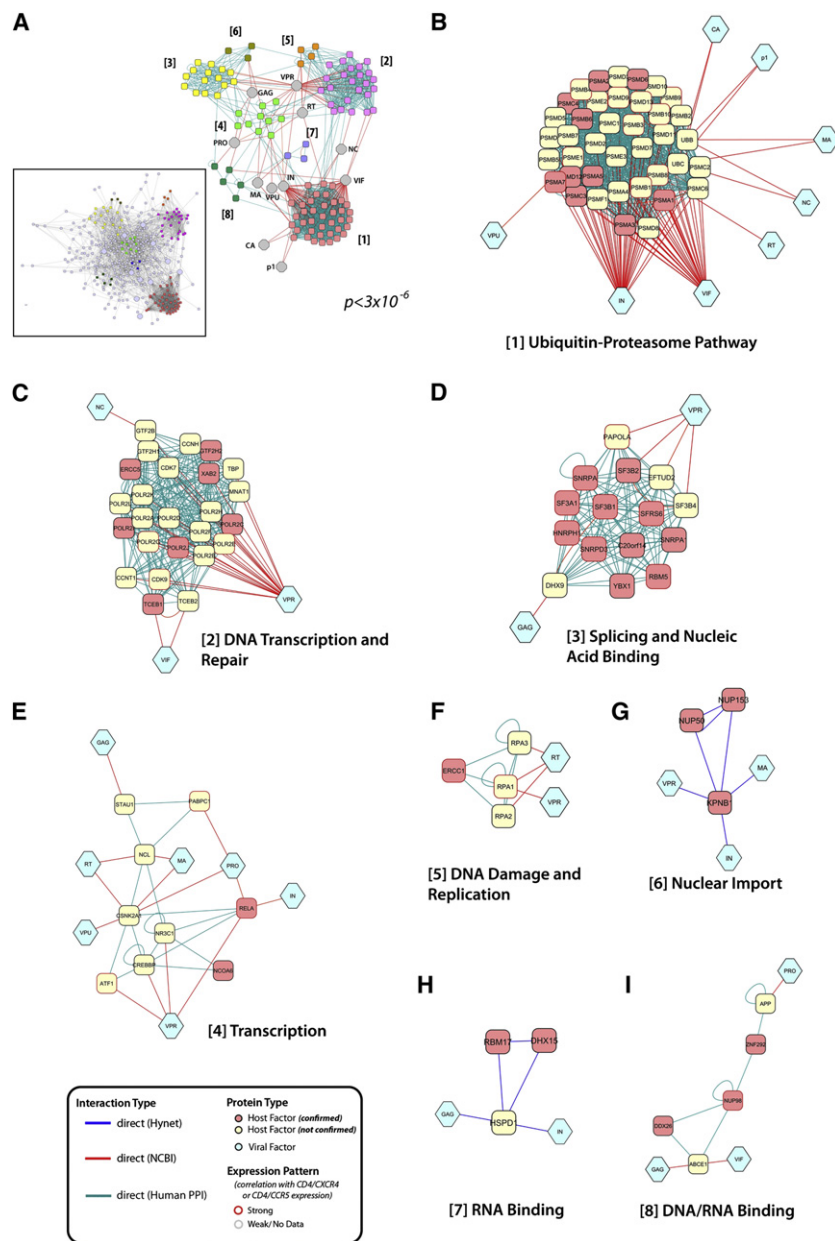


Figure 3. Network Topology of HIV-Host Protein Interactions

(A–I) The interaction network was elucidated based upon protein-protein binding data derived from the Hynet yeast two-hybrid (Y2H) database (blue connectors) and additional various human protein interaction databases (green connectors). Furthermore, connections to HIV-encoded proteins (light blue) were completed by incorporation of data from the HIV-1, Human Protein Interaction Database (NIAID) (red connectors). (A, inset) The resulting network contained 2291 direct protein interactions, connecting 213 confirmed host cellular factors, 11 HIV proteins, and 169 additional human proteins, which interact directly with at least two confirmed HIV host factors and one HIV-encoded protein (Figure S6). Using permutation testing, the density of protein interactions in this network was found to be significantly enriched ($p < 3 \times 10^{-6}$). (A) To identify potential molecular complexes, this network was analyzed for highly connected local network modules (MCODE). (B–I) Importantly, a number of these densely connected areas formed distinct functional subgroups, suggesting that they represent multiprotein complexes, which directly interact with viral factors to facilitate HIV replication. HIV-encoded proteins were abbreviated as follows: CA, Capsid; GAG, gag polyprotein; MA, Matrix; VPR, Vpr; NC, Nucleocapsid; IN, Integrase; VIF, Vif; VPU, Vpu; RT, Reverse Transcriptase; PRO, Protease; P1, p1.

DNA was quantified and served as markers of blocks either to nuclear import (reduction in 2-LTR circles) or to integration (accumulation of 2-LTR circles). In addition, viral DNA integration was quantified by taking advantage of the amplification of HIV DNA sequences adjacent to cellular Alu repeats (Butler et al., 2001) (Figure 5A). For control purposes, parallel infections were performed in the presence of an HIV-1 integrase inhibitor (Hazuda et al., 2000).

A total of 15 genes were identified that, when targeted by cognate siRNAs, significantly reduced viral DNA integration ($p < 0.001$) (Figure 5A; Table S7). The activities of all functional siRNAs targeting the same gene were combined to classify these factors into two categories based on the quantitative PCR as-

says. In the first category, RNAi knockdowns resulted in a reduction in integration and an increase in 2-LTR circles, indicating effects on viral DNA integration itself (ANAPC2 to IK) (Figure 5A). In the second category, little or no reduction in 2-LTR circles was seen, indicating a block to nuclear import (TNPO3 to NUP153) (Figure 5A). One factor, CHADL/LOC150356, could initially not be definitely placed, but further analysis suggests that it is likely involved in regulating viral DNA integration (Figure S7).

Several of the identified factors were members of the nucleocytoplasmic transport machinery. Unexpectedly, inhibition of these factors affected HIV-1 infection at two different steps. NUP153 and RANBP2 seemed to be involved in nuclear import of the PIC (Figure 5B, upper panel). In contrast, the nuclear import protein KPNB1 (importin β -1) as well as NUP98 seemed to be required for viral integration. (Figure 5B, middle panel). These results suggest that the karyopherin and the nucleoporin can regulate viral DNA integration at a postnuclear entry step; however, additional studies will be helpful in determining whether they act directly on viral nucleoprotein complexes or regulate a host cellular factor required for integration.

Four out of six of the cellular factors implicated in nuclear import, including NUP358/RANBP2, NUP153, and the importin- β -family protein Transportin-SR2/TNPO3, were selective for HIV (Figure 5A). TNPO3 was identified by Brass et al. (2008) as being

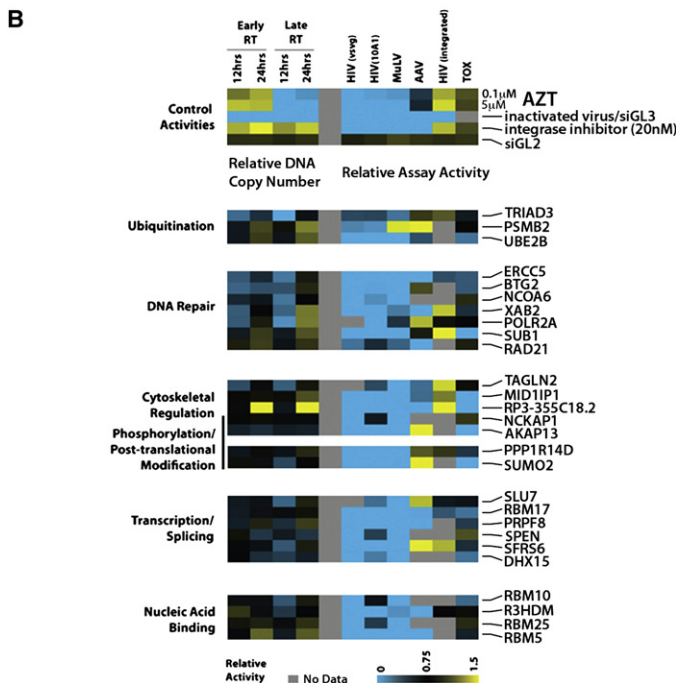
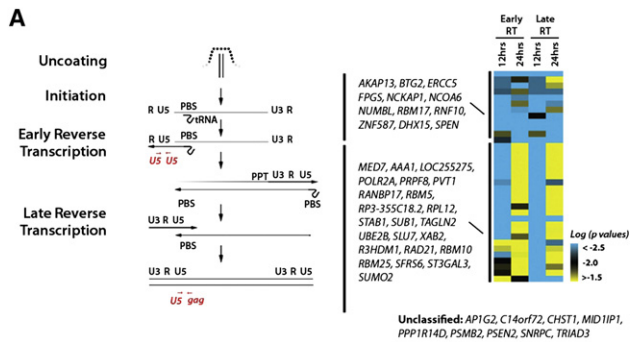


Figure 4. Host Factors Important for HIV-1 Reverse Transcription

(A) 293T cells were transfected with siRNAs targeting indicated genes and, after 48 hr, infected with pNL43R+E-luc(VSV-G). Real-time quantitative PCR amplification analysis on DNA extracts was performed at the indicated time points to assess for the amount of early and late viral RT products. The left panel depicts a schematic of the reverse transcription process and the corresponding primer/probe sets utilized for the quantitative assay (red). Multiple experiments for one of more active siRNAs targeting the same gene were considered to determine statistical significance compared to negative controls (Wilcoxon signed-rank test) (right panel). A pairwise comparison across the assays (data not shown) was also performed, and then the factors were categorized into two functional groups (indicated by the black bars). The black bars in the left panel indicate the potential steps during reverse transcription in which the two separate groups of factors are proposed to act. Additional factors that influenced reverse transcription and could not be statistically segregated into either of these two groups are indicated (unclassified).

(B) As in (A), reverse transcription was monitored with 293T cells that had been transfected with siRNAs targeting indicated host mRNAs (lower panel) or with a negative control siRNA GL2 (upper panel). For control purposes, assays were also performed with the HIV-1 RT inhibitor AZT, an HIV-1 integrase inhibitor, or with heat-inactivated virus. (Top panel) Quantitative PCR values were normalized to an internal control gene, and then to the control siRNA GL2 values. The right section of the heatmap depicts the effects of gene inhibition on various cellular assays, including infection by VSV-G or 10A1-pseudotyped HIV-1 vector, VSV-G-pseudotyped MuLV vector, or by AAV, as indicated, as well as on HIV-1 LTR-mediated transcription (HIV integrated) and on cellular viability (TOX), as described in Figure 1. Luciferase signal values were also normalized to control siRNA GL2. The median relative activity is depicted in a continuum of strong (blue) to weak (yellow). Only the genes that, upon depletion, significantly alter the initiation or rate of reverse transcription are shown ($p < 0.001$, Student's *t* test).

important for HIV-1 infection, and here we demonstrate its involvement in nuclear import of the PIC (Figure 5A). In contrast, all genes that influence integration also affected infection by MuLV (Figure 5A). Thus, the ability of the lentivirus HIV-1 to replicate in nondividing cells (Lewis and Emerman, 1994), as compared with MuLV, is reflected in the differences of host factors needed for nuclear import (De Rijck et al., 2007). Interestingly, the HIV PIC was unable to integrate into the host genome independently of nuclear translocation in cycling cells (postnuclear envelope breakdown), suggesting that HIV nuclear import and integration may be functionally coupled processes (see Discussion).

DISCUSSION

Here, we describe a genome-wide assay with an arrayed siRNA library to identify genes required for early stages of HIV infection. Inhibition of the gene function of putative host factors may induce cellular responses that destabilize the viral cDNA, accel-

erate turnover, or otherwise indirectly interfere with viral replication. To identify those proteins that are more likely to be direct regulators of the viral life cycle, and thus be critical for HIV-1 pathogenesis, we developed further selection criteria, which took into consideration a number of additional, and statistically significant, lines of evidence, including protein interactions, mRNA expression, and gene ontology (Figure 1D; Supplemental Data). Importantly, since this approach does not only rely on identifying host factors based solely on ranking siRNA activities, we were able to more effectively mine the genetic data sets to identify those factors more likely to be immediate regulators of HIV-1 infection, and establish a network of host-pathogen interactions that coordinates HIV-1 infection. Retrospective analysis showed that evidence scores of confirmed genes are, on average, 40% higher than those that did not confirm ($p = 9.7 \times 10^{-15}$), indicating that our approach enriched for biologically relevant activities (Table S3; data not shown).

Further validation studies have identified over 40 host factors that regulate capsid uncoating and reverse transcription steps of early HIV-1 replication (Figure 4). Additionally, we have elucidated 15 cellular factors that facilitate nuclear entry of the HIV-1 PIC and integration of proviral DNA (Figure 5). Importantly, only genes that regulated infection with HIV-1 virus pseudotyped with both VSV-G and 10A1 envelopes were considered for further analysis, thus excluding factors that may regulate

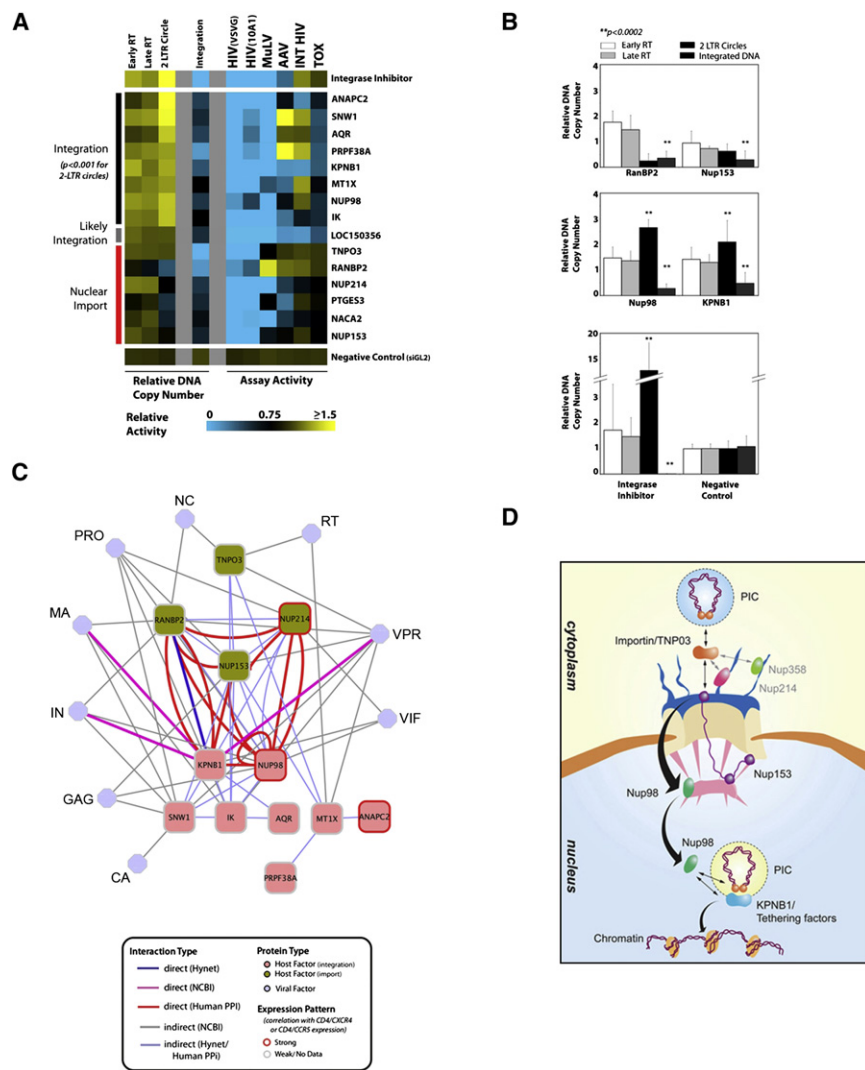


Figure 5. Host Factors Important for Nuclear Import of HIV-1 PICs and Viral DNA Integration

(A) 293T cells were transfected with siRNAs targeting indicated host factors or control siRNA (lower panel) or were treated with an HIV-1 integrase inhibitor and subsequently challenged with pNL43R⁺E⁻luc(VSV-G). Real-time quantitative PCR amplification analysis was then used to measure early and late RT products at 24 hr postinfection, as well as levels of 2-LTR circles and levels of integrated viral DNA at 24 and 48 hr postinfection, respectively. siRNA-transfected 293 T cells were also subjected to the same infection, proviral expression, and toxicity screens as described in Figure 4B (right panel of heatmap). The median of at least three replicate experiments for one or more active siRNAs targeting the same gene is shown in a blue-to-yellow continuum. Genes that, when inhibited, significantly impede viral integration are shown ($p < 0.01$, Student's t test). Segregation of factors into nuclear import, likely integration, and integration classes is primarily based on the level of 2-LTR circles and is also supported by additional statistical significance calculations based upon the collective activity across all PCR assays in comparison to positive (integrase inhibitor) and negative (siGL2) controls (Figure S7). (B) Representative single siRNA activities of nuclear envelope/import-associated genes in quantitative PCR assays reflect the dual categorization of this class of proteins in both nuclear import (upper panel) and integration (middle panel). Relative DNA copy numbers of early RT, late RT, 2-LTR circle formation, and integrated DNA as well as statistical significance for 2-LTR circles and integrated DNA are shown ($p < 0.0001$, Student's t test). Error bars represent standard deviations of the data. (C) Biochemical relationships, based upon the network analysis shown in Figure 3, between proteins involved in integration (red) and nuclear import (green) and direct or indirect interactions among those proteins and with proteins encoded by HIV (blue) are depicted.

(D) A model for the molecular coupling between nuclear import of the viral PIC and viral DNA integration processes. Proteins were organized as predicted from the protein interaction data in (C) and are oriented on the basis of the quantitative PCR data (A and B).

endosomal function associated with VSV-G-mediated entry (see Figure 2C legend). Taken together, these studies indicate that host cellular factors are involved in a variety of different cellular processes that influence HIV-1 reverse transcription, nuclear import, and integration.

Comparison with Reported HIV Host Factors Identified through RNAi-Based Functional Screening

A recent genome-wide RNAi analysis by Brass et al. (2008) has identified ~284 genes as host cellular factors required for HIV replication. Comparison with the 295 confirmed genes presented here reveals a modest, but statistically significant, overlap of 13 genes ($p = 0.00021$). We speculate that this moderate concordance is largely due to differences in the analysis and experimental methodologies used. For example, the criteria Brass et al. (2008) employ to report host cellular factors are genes targeted by one or more siRNAs with activities > 2 stan-

dard deviations from the mean. Since ~154 of the genes in the Brass et al. study were supported by the activity of only a single siRNA, it is likely that a fraction of these reported host factors represent false-positive readouts due to off-target RNAi activity (Echeverri et al., 2006).

If we apply the criteria used by Brass et al. to the data presented here, we can identify 60 genes that the two RNAi studies have in common ($p = 0.024$; Table S8). Further reinforcing the parallels, we find that, based upon protein network analysis, an additional 64 genes reported by Brass et al. directly interact with a confirmed gene in our study ($p = 0.019$; Table S8). In our study, we have also prioritized gene activities not only based upon siRNA activity, but also based on comparative activities in additional screens, HIV-host protein interaction data, as well as gene expression and ontology analysis. We anticipate that this approach enabled us to enrich for the most relevant host cellular factors that promote HIV infection, but this also likely

contributed to the observed differences between the two host factor data sets.

Several experimental differences must also be considered when comparing these studies (also see [Table S8](#)). Two independent RNAi libraries, constructed by using separate design criteria and arrayed in different formats, were used to conduct these screens. Also, we have employed a single-cycle, replication-defective HIV vector pseudotyped with a VSV-G envelope in HEK293 cells and measured viral infection at a 24 hr time point. [Brass et al. \(2008\)](#) monitored the replication of a wild-type X4 strain of HIV in CD4/CXCR4-expressing HeLa cells over the course of 48 hr. Thus, we would not be able to identify factors involved in CD4-mediated viral entry, as well as host molecules that regulate late-stage HIV replication. In contrast, however, we monitored single-cycle virus infectivity at an early (24 hr) time point, which likely enabled us to elucidate a more comprehensive set of host factors specifically involved in the early stages of replication, including uncoating, reverse transcription, and integration. These also encompassed proteins that regulate the kinetics of these processes ([Figure 4A](#)). Thus, whereas false-positive activities are an inherent part of large-scale analyses, it is likely that variations in both experimental and data analysis techniques can largely account for the differences between our results and those reported by [Brass et al. \(2008\)](#).

The Role of Cytoskeletal Proteins in Early HIV Replication

The actin cytoskeleton was previously implicated in regulating the initiation of HIV-1 reverse transcription ([Bukrinskaya et al., 1998](#)) as well as in the movement of intracellular viral nucleoprotein complexes (NPCs) ([Arhel et al., 2006](#)). Consistently, the present study has revealed important roles in the earliest steps of HIV-1 infection for AKAP13, a RhoA-specific guanine nucleotide exchange factor (GEF) that regulates actin stress fiber formation; for NCKAP1, which associates with WAVE proteins that regulate actin nucleation/organization; and for TAGLN-2, a putative actin crosslinking/gelling protein ([Table 1](#)). Microtubules, previously shown to be involved in the intracellular movement of HIV-1 RTCs ([Arhel et al., 2006](#); [Naghavi et al., 2007](#)), may also play a regulatory role in reverse transcription, since viral DNA levels were influenced by RP3-355C18.2, a predicted tubulin tyrosine-ligase, and MID1IP1, involved in bundling and stabilizing microtubules ([Table 1](#)). MID1IP1 was also identified recently by Elledge and colleagues ([Brass et al., 2008](#)), but its role in HIV-1 replication was not defined.

Involvement of the DNA-Damage Repair Pathway in HIV Uncoating/Reverse Transcription

Cellular DNA repair machinery has been implicated in playing roles in viral DNA integration and in the completion of viral DNA synthesis after integration ([Goff, 2007](#)). An unanticipated finding here was that proteins involved in DNA-damage response and repair also influence the initiation of reverse transcription and the accumulation of HIV-1 DNA products prior to integration ([Table 1](#)). Moreover, we have found two locally dense networks of proteins containing host factors that participate in DNA repair ([Figure 3C](#), DNA transcription/repair; [Figure 3F](#), DNA damage/replication). Both clusters contain multiple confirmed factors, in-

cluding POLR2A, XAB2, and ERCC5, which have been mapped to early steps in the viral life cycle ([Table 1](#)). The viral interface for this host-pathogen interaction is mediated by Vpr, a component of the RTC/PIC that was recently being linked to the DNA-damage response pathway ([Schrofelbauer et al., 2007](#)). Several other DNA repair proteins, including MUS81, ERCC1, and MRE11, involved in nucleotide excision repair were also implicated in the early events of HIV-1 infection ([Figure 2C](#)).

Nucleic Acid-Binding Proteins Participate in the Early Stages of HIV Replication

Retroviral reverse transcription presumably involves the unwinding of RNA-RNA, RNA-DNA, and DNA-DNA strands, suggesting, that one or more cellular helicase may participate in viral DNA synthesis. Our results indicate that DHX15, a putative ATP-dependent RNA helicase that plays a role in pre-mRNA splicing, is important for early viral DNA synthesis in target cells ([Table 1](#)). Other nucleic acid-binding proteins, including RBM17, were also implicated in regulating HIV-1 DNA synthesis ([Table 1](#)). DHX15 and RBM17 are contained in a densely connected network ([Figure 3H](#)), supporting their status as proximal regulators of HIV-1 replication. In addition, several factors involved in transcription or mRNA splicing that regulate viral uncoating or reverse transcription were also identified ([Table 1](#)). Correspondingly, two MCODE protein complexes, which regulate transcription, splicing, and nucleic acid binding, were identified through our network analysis ([Figures 3D and 3E](#)). The latter cluster consists of 11 confirmed factors, of which 2 were required for reverse transcription, suggesting that these proteins collectively coordinate either the uncoating or DNA synthesis steps of HIV replication. It remains to be determined whether these host proteins act by interacting directly with the viral nucleic acids, or whether they regulate other host factors required for HIV infection.

The Ubiquitin-Proteasome Pathway Is Required for Capsid Uncoating and Viral DNA Synthesis

The ubiquitin-proteasome pathway has previously been associated with the early steps of HIV replication, where it acts negatively to destroy incoming viral replication complexes ([Butler et al., 2002](#); [Schwartz et al., 1998](#)). Our studies have also revealed a positive role for this pathway. The ubiquitin ligases UBE2B (RAD6) and TRIAD3, as well as the proteasome component PSMB2, were each important for HIV-1 reverse transcription ([Table 1](#)). Network analysis also revealed that the viral Integrase and Vif proteins have multiple interactions with a cluster of proteins that function in the ubiquitin-proteasome pathway ([Figure 3B](#)), indicating that these viral factors may play a structural role in the HIV RTC to recruit the proteasomal machinery and facilitate uncoating or reverse transcription.

Posttranslational Modifications in Early HIV Replication

Previously, the cAMP-dependent protein kinase (PKA) was implicated in regulating early steps of HIV-1 replication ([Cartier et al., 2003](#)). AKAP13 is a PKA scaffold protein ([Table 1](#)) and as such may be involved in mediating PKA-dependent regulation of HIV-1 reverse transcription. Protein dephosphorylation events may also act to regulate HIV-1 reverse transcription,

Table 1. Selected siRNA Phenotypes in Early Steps of HIV-1 Replication

Gene Names	Functional Category ^a	Proposed Role in HIV Replication	Replication Block
ERCC5, BTG2, NCOA6	DNA-damage repair	DNA repair during RT	RT initiation
DHX15	Nucleic acid binding	Putative RNA helicase	RT initiation
RBM17	Nucleic acid binding (RNA splicing)	RNA binding	RT initiation
AKAP13	Regulation of cytoskeleton	PKA scaffold/actin organization	RT initiation
NCKAP1	Regulation of cytoskeleton	Actin organization	RT initiation
SPEN	Signal transduction	Wnt-signaling regulation	RT initiation
NUMBL	Signal transduction	Notch-signaling regulation	RT initiation
POLR2A, SUB1, XAB2, RAD21	DNA-damage repair	DNA repair during RT	RT kinetics or viral DNA stability
R3HDM1	Nucleic acid binding	Single-stranded RNA/DNA binding	RT kinetics or viral DNA stability
RBM5, RBM10, RBM25	Nucleic acid binding (RNA splicing)	RNA binding	RT kinetics or viral DNA stability
SFRS6, PRPF8, SLU7	Nucleic acid binding (RNA splicing)	RNA binding	RT kinetics or viral DNA stability
RANBP17	Nuclear Import	RNA binding	RT kinetics or viral DNA stability
SUMO2	Posttranslational modification	SUMOylation	RT kinetics or viral DNA stability
RP3-355C18.2	Regulation of cytoskeleton	Putative tubulin tyrosine ligase	RT kinetics or viral DNA stability
TAGLN2	Regulation of cytoskeleton	Actin organization	RT kinetics or viral DNA stability
UBE2B	Ubiquitin/proteasome	Ubiquitin ligase	RT kinetics or viral DNA stability
SNRPC	Nucleic acid binding (RNA splicing)	RNA binding	RT (unclassified)
PPP1R14D	Posttranslational modification	Phosphatase activity	RT (unclassified)
MID1IP1	Regulation of cytoskeleton	Microtubule organization	RT (unclassified)
PSEN2	Signal transduction	Notch signaling	RT (unclassified)
TRIAD3	Ubiquitin/proteasome	Ubiquitin ligase	RT (unclassified)
PSMB2	Ubiquitin/proteasome	Proteasome component	RT (unclassified)
TNPO3	Nuclear import	Translocation of PIC	Nuclear import
RANBP2	Nuclear import/posttranslation modification	Translocation of PIC	Nuclear import
NUP153	Nuclear pore component	Translocation of PIC	Nuclear import
NUP214	Nuclear pore component	Translocation of PIC	Nuclear import
PTGES3	Prostaglandin E synthesis	Translocation of PIC	Nuclear import
KPNB1	Nuclear import	Integration coupling factors	Integration
NUP98	Nuclear pore component	Integration coupling factors	Integration
ANAPC2	Anaphase promotion/ubiquitin ligase activity	Putative tethering factors	Integration
PRPF38A	Nucleic acid binding	Putative tethering factors	Integration
SNW1	Transcriptional regulation	Putative tethering factors	Integration
AQR	DNA repair	Putative integration cofactor	Integration

^aReferences are contained in Table S9.

as indicated by the importance of protein phosphatase 1 regulatory subunit 14D (PPP1R14D), a negative regulator of the catalytic subunit of the serine/threonine protein phosphatase 1 (Table 1).

SUMOylation events have been proposed to be important in the early steps of MuLV infection (Yueh et al., 2006). We have found that SUMO-2, one of three small, ubiquitin-related modifier proteins, is important during the late stage of HIV-1 (and MuLV) reverse transcription (Table 1). In addition, RANBP2, a SUMO-1 E3 ligase that is a component of the cytoplasmic filaments of the nuclear pore complex, was required for nuclear import of the HIV-1 DNA (Table 1), perhaps through the sumoylation of viral proteins in the PIC or host factors required

in the PIC. Since RANBP2 influenced HIV, but not MuLV, infection (Figure 4B), different SUMO conjugating systems may be important for each of these two viruses.

Host Factors Required for Nuclear Import of the Viral Preintegration Complex

The mechanism of HIV nuclear import is controversial, with multiple proteins and nucleic acids proposed to play a role (Suzuki and Craigie, 2007). Our studies, combined with those of Brass et al. (2008), indicate the involvement of NUP153, RANBP2, and TNPO3 as factors involved in HIV-1 PIC import (Table 1). We have also uncovered roles for NUP214; the nascent polypeptide-associated complex alpha subunit 2, NACA2; and

prostaglandin E synthase, PTGES3 (Table 1 and Figure 5). The potential role of prostaglandins in HIV-1 nuclear import is particularly intriguing, because these factors are already known to regulate the import of other types of nuclear cargo (Gomez et al., 2005; Malki et al., 2005), and they may represent a new therapeutic target for HIV-1 infection.

Host Proteins and Viral DNA Integration

Our studies have also revealed several cellular factors important for HIV-1 DNA integration. The first is ANAPC2, a component of the anaphase-promoting complex that promotes the metaphase-anaphase transition. ANAPC2 is a cullin protein with a role in ubiquitin ligase activity (Table 1), suggesting that ubiquitin modification and/or proteolysis of virus PIC components may play a regulatory role during the integration stage of viral DNA replication. Recent reports have suggested that passage through mitosis may promote HIV as well as MLV DNA integration (Mannioui et al., 2004; Roe et al., 1993), providing a possible role for ANAPC2. The second factor, SNW1, is a transcriptional coactivator that associates with a cyclophilin-like protein, peptidyl-prolyl isomerase-like 1 (PPIL1) (Table 1). The structure of PPIL1 resembles that of other members of the cyclophilin family, in particular Cyclophilin A (Xu et al., 2006), suggesting that cyclophilin proteins might be involved in regulating both late (integration) and early (uncoating) (Luban, 2007) replication steps. A third factor, aquarius or AQR, also regulates HIV-1 DNA integration (Figure 5A). This factor associates with XAB2, a protein involved in transcription-coupled repair (Kuraoka et al., 2008), suggesting that this DNA repair pathway might also play a role during viral DNA integration.

A Model for Import-Coupled Integration

Protein networks that are implicated in nuclear import and integration of the HIV PIC are shown in Figure 5C. This network was constructed based upon known or experimentally determined interactions between host factors and HIV-encoded proteins, and proteins are organized and oriented on the basis of the quantitative PCR mapping data (Figure 5A). Based upon these data, we propose a model wherein nuclear import of the PIC and proviral DNA integration are molecularly coupled events mediated by nucleoporins, karyopherin, and putative tethering factors (Figure 5D). Specifically, we hypothesize that nuclear transport of the viral PIC through the nuclear pore complex is mediated by soluble transport receptors, such as TNPO3, and nuclear pore components (Stewart, 2007). Next, the PIC cargo is transferred consecutively to phenylalanine-glycine (FG) repeat domains of variant nucleoporins, for example NUP358/RANBP2. Subsequently, in this model, the cargo is delivered to the FG repeats of NUP214 (anchored to the cytoplasmic ring) and NUP153 (distal ring). The FG repeats of both NUP214 and NUP153 are highly flexible and can either reach down to the nuclear basket (NUP214) or toward the cytoplasmic periphery of the central pore (NUP153) (Paulillo et al., 2005), thereby possibly acting to accelerate PIC translocation. As an alternative, we cannot rule out that RNAi silencing of nuclear import machinery components may be acting indirectly by altering the localization of protein(s) required for HIV nuclear import.

Upon traversing the nuclear pore, we hypothesize that the PIC is then released into the nucleus through interaction with the FG-containing NUP98 located near the nuclear basket, which dynamically associates with and dissociates from the nuclear pore (Griffis et al., 2004). Our data indicate that NUP98 is essential for viral integration, suggesting that NUP98 likely directs the viral PIC from the nuclear pore to the proximity of the chromatin. The intranuclear mobility of NUP98 has previously been linked to active transcription sites, possibly through direct interactions with the transcriptional machinery or with newly produced transcripts and RNP complexes (Griffis et al., 2004). Additionally, the PIC is anchored to chromatin through potential “tethering” factors. Targeted integration into active transcription units is mediated by PSIP1/LEDGF/p75 (Ciuffi and Bushman, 2006) (and references therein). Additional factors may help direct integration to transcriptionally active regions, as suggested by the finding that histone posttranslational modifications associated with expression positively correlate with HIV integration frequency (Wang et al., 2007). We have identified several proteins that are required for viral integration and are linked to transcription, signaling, and splicing (ANAPC2, MT1X, SNW1, IK, PRPF38A, and AQR). We hypothesize that these proteins might act to target the PIC to the chromatin and/or act as enzymatic cofactors for DNA integration. Our data indicate that karyopherin KPNB1 is also required for viral integration, suggesting that it may either cooperate with NUP98 to direct the PIC to transcriptionally active chromatin or, alternatively, could serve as an intermediate to couple the viral PIC to the tethering factors. These data support a model in which components of nuclear import machinery facilitate both HIV nuclear entry and integration, and they suggest that the coupling of these steps is required for the establishment of the HIV provirus.

Taken together, this integrative approach toward genome-wide host-pathogen interaction analysis has revealed cellular factors that coordinately regulate the early steps of HIV viral replication, including those required for reverse transcription and proviral establishment. To our knowledge, altogether new pathways, such as the possible involvement of Notch signaling in reverse transcription and prostaglandins in nuclear import, are suggested for the first time. An understanding of their role in cell types targeted by HIV during infection, such as helper T cells, macrophages, and dendritic and microglial cells, will be critical for characterizing the contribution of these factors toward disease progression. The development of small molecules that modulate the activity of these proteins may provide novel strategies for the treatment of HIV/AIDS, particularly since the inhibition of early stages in the viral life cycle has already proven to be therapeutically effective.

EXPERIMENTAL PROCEDURES

For detailed Experimental Procedures, see the Supplemental Data.

High-Throughput and Reconfirmation siRNA Screens

Genome-wide siRNA libraries targeting 19,628 human genes (2 siRNAs/well, ~3 wells/gene) were screened in duplicate or triplicate runs by using a high-throughput robotic system to identify siRNAs that influence the infection of human 293T cells by either a VSV-G-pseudotyped HIV-1 reporter virus encoding luciferase, by a VSV-G-pseudotyped MLV vector encoding luciferase,

or by an AAV vector encoding luciferase. In addition, the library was counterscreened to identify siRNAs that influence cell viability. siRNAs for reconfirmation were individually rearranged in 384 wells in duplicate and were screened in at least 4 independent experiments and were analyzed as described in the [Supplemental Data](#).

Expression Vectors and Reporter Viruses

For the HIV and MLV screens, pNL43-Luc-E⁻R⁺ (HIV-1 wild-type Δ env, encoding firefly luciferase GL3) vector was used to generate VSV-G-pseudotyped lentiviral supernatant, and the MLV-based retroviral vector pVGP3, a derivative of pBabe, was used to generate a Moloney-based virus. For the AAV screen, GL3 luciferase was introduced into an ITR-flanked AAV expression vector to generate AAV particles encoding luciferase. For secondary confirmation assays, an amphotropic 10A1 envelope-pseudotyped HIV virus was generated.

Secondary Confirmation Analysis

Genes subject to staging of their effects in the HIV life cycle were also screened in parallel for their activities on HIV-1 (VSV-G), HIV-1 (10A1), MLV (VSV-G), and AAV reporter viruses as well on HIV-1 LTR-mediated transcription and cell viability in 96-well-based assays.

Viral DNA Quantitation by Real-Time PCR

HEK293T cells were transfected with siRNAs and were infected with HIV-1 (VSV-G) 2 days later. Total DNA was isolated at 12, 24, and 48 hr postinfection (hpi). For control purposes, a reverse transcription inhibitor or integrase inhibitor was added at the time of infection. DNA was subject to real-time PCR to specifically quantify HIV Early reverse transcription (RT) products, Late RT products, 2-LTR circle forms, and Proviral DNA by using Alu-based nested PCR. To normalize the amount of DNA in each PCR assay, the copy number of the cellular gene porphobilinogen deaminase (PBGD) was quantified. With each experiment, a standard curve was generated.

Yeast Two-Hybrid Screen with HIV Proteins

A yeast two-hybrid (Y2H) screen was performed by using the entire HIV HXB2 proteome as bait and screening against a human leukocyte cDNA library cloned into the GAL4 expression vector.

Data Sets and Network Analysis

Multiple human protein-protein interaction databases, such as Y2H databases and other curated literature-based protein-protein interaction databases (Bind, HPRD, MINT, Reactome), were incorporated in the network analysis. In addition, two human-HIV interaction databases (the NIAID HIV-1, Human Protein Interaction Database and the Y2H in-house database) and the GNF Tissue Atlas Gene Expression database were used in this study. Detailed methodology of the multiscale strategy for hit selection and network analysis can be found in [Supplemental Data](#).

SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures, eight figures, and nineteen tables and are available with this article online at <http://www.cell.com/cgi/content/full/135/1/49/DC1/>.

ACKNOWLEDGMENTS

Tracy L. Diamond is a Leukemia and Lymphoma Society Fellow (Grant #5217-06). We thank U. O'Doherty and J.J. Yu for technical advice with the Alu-PCR assay and Zhang J. for excellent technical assistance; Walter Tian and Eric Lader (Qiagen GMBGH) for facilitating the acquisition of reagents used for this study; and John Hogenesch for critical feedback on the manuscript. This work was supported by a grant from the United States National Institutes of Health (1 RO1 AI072645-01 to J.A.Y. and RO1 AI052845 to F.D.B.) and the University of Pennsylvania Center for AIDS Research.

Received: May 1, 2008

Revised: June 12, 2008

Accepted: July 18, 2008

Published: October 2, 2008

REFERENCES

- Arhel, N., Genovesio, A., Kim, K.A., Miko, S., Perret, E., Olivo-Marin, J.C., Shorte, S., and Charneau, P. (2006). Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes. *Nat. Methods* 3, 817–824.
- Bader, G.D., and Hogue, C.W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4, 2.
- Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J., and Elledge, S.J. (2008). Identification of host proteins required for HIV infection through a functional genomic screen. *Science* 319, 817–824.
- Bukrinskaya, A., Brichacek, B., Mann, A., and Stevenson, M. (1998). Establishment of a functional human immunodeficiency virus type 1 (HIV-1) reverse transcription complex involves the cytoskeleton. *J. Exp. Med.* 188, 2113–2125.
- Butler, S.L., Hansen, M.S., and Bushman, F.D. (2001). A quantitative assay for HIV DNA integration in vivo. *Nat. Med.* 7, 631–634.
- Butler, S.L., Johnson, E.P., and Bushman, F.D. (2002). Human immunodeficiency virus cDNA metabolism: notable stability of two-long terminal repeat circles. *J. Virol.* 76, 3739–3747.
- Cartier, C., Hemonnot, B., Gay, B., Bardy, M., Sanchiz, C., Devaux, C., and Briant, L. (2003). Active cAMP-dependent protein kinase incorporated within highly purified HIV-1 particles is required for viral infectivity and interacts with viral capsid protein. *J. Biol. Chem.* 278, 35211–35219.
- Ciuffi, A., and Bushman, F.D. (2006). Retroviral DNA integration: HIV and the role of LEDGF/p75. *Trends Genet.* 22, 388–395.
- De Rijck, J., Vandekerckhove, L., Christ, F., and Debysers, Z. (2007). Lentiviral nuclear import: a complex interplay between virus and host. *Bioessays* 29, 441–451.
- Echeverri, C.J., Beachy, P.A., Baum, B., Boutros, M., Buchholz, F., Chanda, S.K., Downward, J., Ellenberg, J., Fraser, A.G., Hacohen, N., et al. (2006). Minimizing the risk of reporting false positives in large-scale RNAi screens. *Nat. Methods* 3, 777–779.
- Goff, S.P. (2007). Host factors exploited by retroviruses. *Nat. Rev. Microbiol.* 5, 253–263.
- Gomez, P.F., Pillinger, M.H., Attur, M., Marjanovic, N., Dave, M., Park, J., Bingham, C.O., 3rd, Al-Mussawir, H., and Abramson, S.B. (2005). Resolution of inflammation: prostaglandin E2 dissociates nuclear trafficking of individual NF- κ B subunits (p65, p50) in stimulated rheumatoid synovial fibroblasts. *J. Immunol.* 175, 6924–6930.
- Griffis, E.R., Craige, B., Dimaano, C., Ullman, K.S., and Powers, M.A. (2004). Distinct functional domains within nucleoporins Nup153 and Nup98 mediate transcription-dependent mobility. *Mol. Biol. Cell* 15, 1991–2002.
- Hazuda, D.J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J.A., Espeseth, A., Gabryelski, L., Schleif, W., Blau, C., et al. (2000). Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. *Science* 287, 646–650.
- Konig, R., Chiang, C.Y., Tu, B.P., Yan, S.F., DeJesus, P.D., Romero, A., Bergauer, T., Orth, A., Krueger, U., Zhou, Y., et al. (2007). A probability-based approach for the analysis of large-scale RNAi screens. *Nat. Methods* 4, 847–849.
- Kuraoka, I., Ito, S., Wada, T., Hayashida, M., Lee, L., Saijo, M., Nakatsu, Y., Matsumoto, M., Matsunaga, T., Handa, H., et al. (2008). Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair. *J. Biol. Chem.* 283, 940–950.
- Lewis, P.F., and Emerman, M. (1994). Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *J. Virol.* 68, 510–516.

- Luban, J. (2007). Cyclophilin A, TRIM5, and resistance to human immunodeficiency virus type 1 infection. *J. Virol.* *81*, 1054–1061.
- Malki, S., Nef, S., Notarnicola, C., Thevenet, L., Gasca, S., Mejean, C., Berta, P., Poulat, F., and Boizet-Bonhoure, B. (2005). Prostaglandin D2 induces nuclear import of the sex-determining factor SOX9 via its cAMP-PKA phosphorylation. *EMBO J.* *24*, 1798–1809.
- Mannioui, A., Schiffer, C., Felix, N., Nelson, E., Brussel, A., Sonigo, P., Gluckman, J.C., and Canque, B. (2004). Cell cycle regulation of human immunodeficiency virus type 1 integration in T cells: antagonistic effects of nuclear envelope breakdown and chromatin condensation. *Virology* *329*, 77–88.
- Mukherji, M., Bell, R., Supekova, L., Wang, Y., Orth, A.P., Batalov, S., Miraglia, L., Huesken, D., Lange, J., Martin, C., et al. (2006). Genome-wide functional analysis of human cell-cycle regulators. *Proc. Natl. Acad. Sci. USA* *103*, 14819–14824.
- Naghavi, M.H., Valente, S., Hatzioannou, T., de Los Santos, K., Wen, Y., Mott, C., Gundersen, G.G., and Goff, S.P. (2007). Moesin regulates stable microtubule formation and limits retroviral infection in cultured cells. *EMBO J.* *26*, 41–52.
- Nisole, S., and Saib, A. (2004). Early steps of retrovirus replicative cycle. *Retrovirology* *1*, 9.
- Paulillo, S.M., Phillips, E.M., Koser, J., Sauder, U., Ullman, K.S., Powers, M.A., and Fahrenkrog, B. (2005). Nucleoporin domain topology is linked to the transport status of the nuclear pore complex. *J. Mol. Biol.* *351*, 784–798.
- Roe, T., Reynolds, T.C., Yu, G., and Brown, P.O. (1993). Integration of murine leukemia virus DNA depends on mitosis. *EMBO J.* *12*, 2099–2108.
- Schrofelbauer, B., Hakata, Y., and Landau, N.R. (2007). HIV-1 Vpr function is mediated by interaction with the damage-specific DNA-binding protein DDB1. *Proc. Natl. Acad. Sci. USA* *104*, 4130–4135.
- Schwartz, O., Marechal, V., Friguet, B., Arenzana-Seisdedos, F., and Heard, J.M. (1998). Antiviral activity of the proteasome on incoming human immunodeficiency virus type 1. *J. Virol.* *72*, 3845–3850.
- Stewart, M. (2007). Molecular mechanism of the nuclear protein import cycle. *Nat. Rev. Mol. Cell Biol.* *8*, 195–208.
- Suzuki, Y., and Craigie, R. (2007). The road to chromatin - nuclear entry of retroviruses. *Nat. Rev. Microbiol.* *5*, 187–196.
- Wang, G.P., Ciuffi, A., Leipzig, J., Berry, C.C., and Bushman, F.D. (2007). HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res.* *17*, 1186–1194.
- Xu, C., Zhang, J., Huang, X., Sun, J., Xu, Y., Tang, Y., Wu, J., Shi, Y., Huang, Q., and Zhang, Q. (2006). Solution structure of human peptidyl prolyl isomerase-like protein 1 and insights into its interaction with SKIP. *J. Biol. Chem.* *281*, 15900–15908.
- Yueh, A., Leung, J., Bhattacharyya, S., Perrone, L.A., de los Santos, K., Pu, S.Y., and Goff, S.P. (2006). Interaction of moloney murine leukemia virus capsid with Ubc9 and PIASy mediates SUMO-1 addition required early in infection. *J. Virol.* *80*, 342–352.
- Zar, J. (1999). *Biostatistical Analysis*, Fourth Edition (Upper Saddle River, NJ: Prentice Hall).

Note Added in Proof

After acceptance of this manuscript, another study has also described a role for TNPO3 in HIV-1 nuclear import (Christ et al., 2008, *Curr. Biol.* *18*, 1192). That study demonstrated a direct interaction between TNPO3 and HIV-1 Integrase.