Network-Based Analysis of OMICs Data to Understand the HIV-Host Interaction

Sergey Ivanov[1,2*], Alexey Lagunin[1,2], Dmitry Filimonov[1], Olga Tarasova[1]

[1]Department of Bioinformatics, Institute of Biomedical Chemistry, Moscow, Russia

[2]Department of Bioinformatics, Pirogov Russian National Research Medical University, Moscow, Russia

* Correspondence:
Dr. Sergey Ivanov
sergey.ivanov@ibmc.msk.ru

Keywords: virus-host interaction, human immunodeficiency virus, protein-protein interactions, OMICs, transcriptomics, network analysis.

Abstract

The interaction of human immunodeficiency virus with human cells is responsible for all stages of the viral life cycle, from the infection of CD4+ cells to reverse transcription, integration, and the assembly of new viral particles. To date, a large amount of OMICs data as well as information from functional genomics screenings regarding the HIV-1-host interaction has been accumulated in the literature and in public databases. We processed databases containing HIV-host interactions and found 2910 HIV-1-human protein-protein interactions, mostly related to viral group M subtype B, 137 interactions between human and HIV-1 coding and non-coding RNAs, essential for viral lifecycle and cell defense mechanisms, 232 transcriptomics, 27 proteomics, and 34 epigenomics HIV-related experiments. Numerous studies regarding network-based analysis of corresponding OMICs data have been published in recent years. We overview various types of molecular networks, which can be created using OMICs data, including HIV-human protein-protein interaction networks, co-expression networks, gene regulatory and signaling networks, and approaches for the analysis of their topology and dynamics. The network-based analysis can be used to determine the critical pathways and key proteins involved in the HIV life cycle, cellular and immune responses to infection, viral escape from host defense mechanisms, and mechanisms mediating different susceptibility of humans to infection. The proteins and pathways identified in these studies may represent a basis for developing new anti-HIV therapeutic strategies such as new small-molecule drugs preventing infection of CD4+ cells and viral replication, effective vaccines, "shock and kill" and "block and lock" approaches to cure latent infection.

1   Introduction

Human immunodeficiency virus (HIV) is one of the most significant pathogens to affect humankind. According to the World Health Organization, approximately 37.9 million people are currently living with HIV, and 770 000 people died from HIV-related

disorders in 2018. Although the existing combination antiretroviral therapy (CART) provides control of the virus and prevents transmission, the HIV infection remains a global health problem (World Health Organization, 2020). Thus, there is an urgent need to understand anti-HIV mechanisms, to develop new strategies for its prophylactics and therapy.

The HIV infection process involves several stages from the binding of virions to receptors on the human CD4+ cell surface to the splicing and export of viral mRNAs from the nucleus to the cytoplasm, and the assembly of virions at the plasma membrane as well as the budding and maturation of the released virions (Kirchhoff, 2013). Like other viruses, HIV cannot complete any aspect of its life cycle without interacting with the host cellular machinery, primarily human proteins and various types of RNA. The human macromolecules, which are required for various stages of HIV life cycle, called host dependency factors (HDFs), whereas macromolecules, which are part of cell defense mechanisms and prevent viral infection, called host restriction factors (HRFs). For instance, the first step of HIV interaction with the CD4+ cell is the binding of the viral envelope GP160 complex, consisting of GP120 and GP41, to CD4 receptor as well as the co-receptors: C-X-C chemokine receptor type 4 (CXCR4), C X C chemokine receptor type 6 (CXCR6) or C-C chemokine receptor type 5 (CCR5). This interaction leads to the activation of various kinase cascades that causes cytoskeleton rearrangement, which favors entry and subsequent nuclear import of the virus, as well as increasing of cell survival (Chen et al., 2018). The HDFs are also involved in the process of HIV-1 integration into host cell genome (Marini et al., 2015). During the HIV attack, several HRFs interfere with viral replication at different steps (Shukla and Chauhan, 2019). For instance, cytidine deaminase, APOBEC3G (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G) induces lethal hypermutations (deamination of C to U) in the HIV genome, which are detrimental to viral replication. Viral proteins nef, vpu and vif decrease expression of HRFs, which increase HIV replication in CD4+ cells. For example, vif binds to and blocks the antiviral activity of APOBEC3 proteins (Yu et al., 2003), vpu downregulates Tim-3, also having antiviral role, from the surface of infected CD4+ T cells (Prévost et al., 2020). The role of tat and rev factors is the positive modulation of the viral expression interacting with host factors as well. For more information on HDFs and HRFs see the reviews (Yamashita and Engelman, 2017; Chen et al., 2018; Engelman and Singh, 2018; Shukla and Chauhan, 2019).

Most of the existing anti-HIV drugs are inhibitors of three HIV enzymes, namely protease, reverse transcriptase and integrase; however, researchers are also studying HDFs as new potential targets (Arhel and Kirchhoff, 2010; Zhan et al., 2016; Zuo et al., 2018; Puhl et al., 2019). The approved drugs maraviroc and ibalizumab have the following mechanism of action: maraviroc is an antagonist of CCR5, and ibalizumab is a monoclonal antibody to the epitope on the CD4 receptor (Kinch and Patridge, 2014; Puhl et al., 2019). Enfuvirtide and albuvirtide block the fusion of viral and cellular membranes by binding to the HIV-1 gp41 subunit of the viral envelope that anchors the gp120 subunit, which normally binds to the CD4 receptor (Fung and Guo, 2004). Some other strategies for development of active molecules targeting HDFs are reported in the literature (Zhang et al., 2015; Tan et al., 2019).

Since HIV is characterized by a high mutation rate (Zanini et al., 2015), it can lead to the development of HIV resistance to some drugs (Weber and Harrison, 2017). To overcome HIV drug resistance, along with the approaches aimed at the HIV drug resistance prediction based on the HIV genotype (Riemenschneider and Heider, 2016; Zazzi et al., 2016; Tarasova and Poroikov, 2018), pharmacogenetics studies of genes metabolizing of ART (Ghodke et al., 2012; McDonagh et al., 2015; Alvarellos et al., 2018) and HIV-host interactions (Huang et al., 2011), some other approaches are currently under development to treat HIV infection, namely the identification of new viral and human targets, the identification of new and safer drug combinations (Tan et al., 2012; Moreno et al., 2019), ART optimization (Castiglione et al., 2007; Tarasova et al., 2020), the creation of anti-HIV vaccines (Hsu and O'Connell, 2017), immunotherapy with broadly neutralizing antibodies (bNAbs) (Lacerda et al., 2013; Sok and Burton, 2018; Haynes et al., 2019) and creation of methods to cure latent HIV infection.

To date neither therapeutic nor preventive anti-HIV vaccine exists (Gray et al., 2016). Out of over 100 vaccine trials, only the RV144 trial has achieved positive yet moderate protection. The difficulty in the creation of anti-HIV vaccine can be explained by the fact that people do not develop natural, protective immunity to HIV infection, whereas almost all successful vaccines were developed for diseases for which natural immunity exists (Brett-Major et al., 2017). The development of effective anti-HIV vaccines gave rise to bioinformatics approaches for computational analysis of viral variants and corresponding host factors (i.e. T-cell epitopes) (Fischer et al., 2007; Korber et al., 2009; Barouch et al., 2013; Hulot et al., 2015; Khan et al., 2017), which can be beneficial for the creation of potential anti-HIV vaccine.

To create effective vaccine, deep understanding of interplay between HIV and human immune system is required. The HIV causes innate immune response and later adaptive T cytotoxic and humoral responses, which cannot completely cure infection due to several reasons (Levy, 2015). First, high mutation rate of virus causes formation of viral variants with antigens' epitopes, which cannot be recognized by HIV-specific antibodies and CD8+ T cells (Gallo, 2015; Brett-Major et al., 2017). Second, HIV has additional mechanisms to escape from immune response. For instance, HIV protein Nef downregulates HLA I by modulating the host membrane trafficking machinery, resulting in the endocytosis and eventual sequestration of MHC-I within the cell (Dirk et al., 2016; van Stigt Thans, 2019). Another HIV-1 protein vpu exerts broad immunosuppressive effects by inhibiting activation of the transcription factor NF-κB (Langer, 2019). Third, 10-25% of HIV-infected individuals exhibit development of bNAbs, which recognize the conserved epitopes of HIV envelope proteins (Wang and Zhang, 2019). These antibodies neutralize a broad range of viral variants and can be used in passive immunization of HIV-infected people; however, they cannot neutralize all viral clones and completely cure the infection (van den Kerkhof et al., 2016). BNAbs emerge only after several years of infection and have high level of somatic hypermutations as a result of multiple rounds of antigen recognition and stimulation. They are also polyreactive that allow promiscuous binding to unrelated self- and foreign antigens. Therefore, the relatively low frequency of HIV bNAbs in natural infection may be due to the clearance and deletion of poly- and auto-reactive B cell clones (Wang and Zhang, 2019). Fourth, in conditions of chronic infection, persistent exposure of T cells to high levels of antigen results in a severe T-cell

dysfunctional state called exhaustion (Fenwick et al., 2019). Immune checkpoint molecules, including PD-1, CTLA-4, TIM-3, CD160, 2B4 and LAG-3, play a critical role in the maintenance of exhaustion and dysfunction. Administration of immune checkpoint inhibitors has therefore attracted considerable interest as a strategy to enhance HIV-specific T cell responses (Seddiki and Lévy, 2018; Mylvaganam et al., 2019). Fifth, since viral cDNA is integrated into the human genome, latent HIV reservoirs such as CD4+ T memory cells are present (Gallo, 2015; Pitman et al., 2018; Sadowski and Hashemi, 2019). These reservoirs do not produce viral particles, but they can give rise to infectious virions following activation by various stimuli, leading to viral rebound when ART is interrupted (Brett-Major et al., 2017). Sixth, mucous membranes are bottleneck for multiple viral variants, because only one (the so-called transmitted/founder variant) or a few HIV variants are able to overcome this barrier. These variants have unique properties allowing escaping from immune response in mucosa and transporting to lymph nodes, e.g. they are relatively more resistant to interferons and bind dendritic cells more efficiently (Iyer et al., 2017; Rios, 2018). Therefore the difficulty in developing effective preventive anti-HIV vaccine is to detect the viral variant, which is able to penetrate the mucosal barrier.

The development of therapeutic strategies to cure latent HIV infection is an important research direction along with the development of vaccines. Three main strategies are currently developed, namely gene therapy and genome editing of CCR5 or integrated HIV genome itself (Wang and Cannon, 2016; Peterson and Kiem, 2018), "shock and kill" and "block and lock" strategies (Gallo, 2016; Darcis et al., 2017; Pitman et al., 2018; Sadowski and Hashemi, 2019). "Shock and kill" approach is based on the application of latency reversing agents, which cause reactivation of viral gene transcription and generation of virions. Latency reversing agents include cytokines agonists, epigenetic modifiers, intracellular signaling modulators, transcriptional elongation regulators and some others. Their application causes elimination of infected cells through immune response and virus-induced apoptosis; however these processes are incapable of eliminating all infected cells, thus, additional methods are used together with the application of latency reversing agents (Pitman et al., 2018; Mylvaganam et al., 2019; Sadowski and Hashemi, 2019). "Block and lock" is an opposite strategy, which is based on prevention of viral RNA expression even in the case of T cell reactivation. It can be achieved by application of HIV Tat inhibitors, small interfering RNA (siRNA) or short hairpin RNA (shRNA) that can target and destroy the viral RNAs, and others (Darcis et al., 2017; Pitman et al., 2018; Sadowski and Hashemi, 2019). The more aggressive strategy is the application of gene therapy and genome editing. Several gene editing approaches including those based on Zinc Finger Nucleases (ZFN), transcription activator-like effector nucleases (TALEN) and Clustered Regularly Interspaced Short Palindromic Repeats/Cas-9 (CRISPR/Cas-9) have been applied to cause CCR5 disruption or remove/modify integrated HIV genome. These approaches demonstrated promising results for *in vitro* and animal experiments as well as in clinical trials (Wang and Cannon, 2016; Peterson and Kiem, 2018; Herrera-Carrillo et al., 2019) but they are still under investigation now (Vansant et al., 2020).

The studies of HIV-host interaction is important because it helps to understand the factors influencing the speed of disease progression and pathogenesis features at individual

4

patients. It is known that some people, called long-term non-progressors, have the ability to suppress viremia to undetectable levels, while maintaining elevated CD4 cell counts in the absence of ART (Poropatich and Sullivan, 2011; Gonzalo-Gil et al., 2017). They are subdivided into several groups. For instance, elite controllers are HIV infected individuals, who suppress viremia to less than 50 copies/ml, while maintaining CD4 cell counts from 200 to 1000/ml. Viremic controllers achieve a lesser degree of virologic control (viral load between 200 to 2000 copies/ml), while usually maintaining CD4 cell counts less than 500/ml, in the absence of ART. The phenomenon of long-term non-progressors can be mainly explained by enhanced cellular immune response and decreased susceptibility of CD4+ T cells to HIV infection. For instance, it is currently known that some genetic alleles of the human leukocyte antigen (HLA) can be responsible for HIV control (Gonzalo-Gil et al., 2017). Investigation of mechanisms of HIV control in these groups of patients is very important, because mimicking similar responses in chronically infected individuals, e.g. by therapeutic vaccine, will lead to functional remission of HIV infection (Seddiki and Lévy, 2018).

Human CD4+ T cells are the most widely recognized and best-described cell type, which can be infected by HIV; however, the virus can also replicate in cells of other types including monocytes and macrophages, and various kinds of dendritic and epithelial cells (Kandathil et al., 2016). In particular, HIV can infect cells of the central nervous system, which leads to neuroAIDS in about two-thirds of patients, and characterized by a decline in brain function and movement skills (Elbirt et al., 2015; Clifford, 2017; Kumar et al., 2018; Olivier et al., 2018; Sillman et al., 2018). The corresponding condition is called HIV-associated Neurocognitive Disorder (HAND), which can be further categorized from asymptomatic HAND to HIV-associated dementia linked with cognitive impairment, motor dysfunction, speech problems, and behavioral changes. The entry of HIV into the brain is mediated by infected monocytes and macrophages from the peripheral circulation. These cells are characterized by an enhanced ability to cross the blood-brain barrier (BBB) and go into the brain. Early HIV infection of brain cells triggers the release of BBB-compromising factors, which induces the influx rate of leukocytes into the brain. Higher accumulation of infected leukocytes increases the progression of HIV infection in different brain cells, such as microglia and astrocytes. The infection leads to astrocytes and microglia activation, which causes severe neuroinflammation. The released proinflammatory factors, including nitric oxide, TNF-α, IL-1β, quinolinic acid, β-chemokines, arachidonic acid, etc., as well as HIV proteins, e.g., gp120, tat, vpr, and nef, increase permeability of BBB, cause functional changes in and death of various brain cells, such as microvascular endothelial cells, microglia, neurons, and astrocytes, as well as dendritic and synaptic damage through multiple mechanisms. The corresponding mechanisms include mitochondrial dysfunction, oxidative stress, ion channel modulation, and excitotoxicity induced by a decrease in glutamate uptake by astrocytes (Sillman et al., 2018). Several human genes related to neurotransmission, the integrity of mitochondrial and nuclear DNA, cytokines and associated receptors of the immune system, matrix metalloproteinases, etc. are associated with the severity of the disorder (Olivier et al., 2018). To date, the administration of ART allows significantly decrease the severity of HAND in HIV-infected subjects; however, low permeability of antiviral drugs through BBB and presence of latent HIV infection considerably reduce the efficacy of therapy. To increase the concentration of anti-HIV compounds in the brain, various drug delivery

approaches and prodrugs are developed (Kumar et al., 2018). Since HIV acquires latency in various brain cells, e.g., perivascular macrophages, microglial cells, and astrocytes, the brain tissues represent an essential reservoir for HIV. Despite ART administration, it can lead to chronic pathological implications because minimum viral genome transcription can continuously produce little virus, and viremia can be rebound upon latency reactivation (Van Lint et al., 2013); therefore, HAND treatment also requires the elimination of latent infection by various approaches described above (Marban et al., 2016; Proust et al., 2017).

Development of all above mentioned approaches require the understanding of the mechanisms of HIV-host interactions. To date, a large amount of these OMICs data about the HIV-host interaction has been recorded in the literature and in public databases. Network-based analysis allows for integrating OMICs data of various types, and it can be used to determine critical pathways and key proteins involved in the HIV life cycle, cellular and immune responses to infection, and different speed of disease progression. The identified proteins and pathways may represent targets for the development of new anti-HIV strategies as well as the optimization of the existing therapy.

This review consists of two parts. The first part contains a description of various types of OMICs data, such as HIV-human protein-protein interactions, RNA-RNA interactions, genomics, transcriptomics, proteomics, and epigenomics information as well as data from functional genomics screenings. We review the content and quantity of the corresponding HIV-related data from public resources, databases and literature. The second part contains a description of the network-based analysis of HIV-human interactions, including the types of networks, the methods of integrating them with the OMICs data, and methods of network topology analysis, which can be used to discover new HDFs and key cellular pathways that form the basis for developing new anti-HIV therapeutic and prophylactic strategies.

2   Types of OMICs data describing the HIV-host interaction

2.1   Interactomics data

2.1.1 Protein-protein interactions

The interaction between HIV and human proteins is the first event that causes subsequent changes in the cellular pathways and processes required for the viral life cycle. Thus, information on protein-protein interactions (PPIs) is the most common type of data, and it is used for the creation of network-based models of HIV infection (see below). The information on HIV-human PPIs can be extracted from at least six public databases that are focused on pathogen-host interactions (Table 1). The information in the specialized databases, in turn, was obtained from publications and more general PPI resources, e.g., BioGRID[1], IntAct[2], MINT[3], DIP[4], and STRING[5]. The proteins in these databases, except

---

[1] https://thebiogrid.org

[2] https://www.ebi.ac.uk/intact

[3] https://mint.bio.uniroma2.it

for the NCBI HIV-1 human interaction database[6], are presented as SwissProt[7] accession numbers, which have two primary features. First, they are gene-centric, so each identifier corresponds to a single gene. However, some HIV genes, such as gag, pol, and env, encode two or more proteins. Second, each protein has several identifiers that correspond to different HIV groups, subtypes and even isolates, e.g., tat protein from HIV-1 group M subtype B refers to 20 SwissProt accession numbers corresponding to 20 different isolates. For illustrative purposes, we calculated the number of pairs listed as "HIV gene symbol – SwissProt identifier of human protein" for six databases, whereas the differences in PPIs between the HIV groups, subgroups, and isolates were ignored (Table 1).

To demonstrate the intersections between the database contents, we created an Upset plot (Lex et al., 2014) (Figure 1). Figure 1 shows that the NCBI HIV-1 human interaction database and the PHISTO database[8] have the highest numbers of unique HIV-1-human PPIs, with 540 and 475 respectively, whereas only 86 PPIs were present in all six databases. Among all the databases, the NCBI HIV-1 human interaction database was created using the manual curation of data taken from the literature. It contains a description of the functional significance of interactions, e.g., "phosphorylation," "acetylation," "activation," and "inhibition." In addition to direct interactions, the NCBI HIV-1 human interaction database also contains indirect ones, which primarily reflect an influence on the gene expression or protein activity (the corresponding indirect interactions are not shown in Table 1 and Figure 1).

By merging the data from six databases, we obtained 2910 unique HIV-1-human PPIs corresponding to 2051 individual human proteins (Table 1). The large number of human proteins interacting with HIV-1 can be explained by disadvantages in the experimental methods used to measure the PPIs (Goodacre et al., 2018). According to the largest database, PHISTO, the most common high-throughput approaches to identifying HIV-human PPIs are affinity-purification coupled to mass spectrometry, pull-down, and co-immunoprecipitation. These methods may address physical interactions in which two proteins are in the same complex but do not interact directly. According to some estimates, 50 to 70 percent of all PPIs may be physical but not direct (Goodacre et al., 2018). Significant numbers of PPIs obtained by high-throughput methods may be false positives (Goodacre et al., 2018). The VirHostNet [9], VirusMentha [10], and Viruses.STRING[11] databases provide confidence scores that reflect the number of studies

---

[4] https://dip.doe-mbi.ucla.edu/dip/Main.cgi

[5] https://string-db.org

[6] https://www.ncbi.nlm.nih.gov/genome/viruses/retroviruses/hiv-1/interactions

[7] https://www.uniprot.org

[8] http://www.phisto.org

[9] http://virhostnet.prabi.fr

[10] https://virusmentha.uniroma2.it/about.php

[11] http://viruses.string-db.org

in which PPI was detected, and a number and quality of experimental techniques were used. The scores vary from 0 to 1 and allow for the selection of the most likely PPIs; however, the medians of the scores (0.21 for VirusMentha, 0.33 for VirHostNet, and 0.44 for Viruses.STRING) indicate that most of the interactions are low-confidence. Thus, the real number of direct interactions is potentially much lower than those presented in Table 1.

Since the PPI profiles for other viruses were shown to be different for different viral variants (Neveu et al., 2012; Goodacre et al., 2018), we calculated the numbers of interactions for particular groups and subtypes of HIV-1 by accounting for the viral isolates (Table 2). Almost all the data belong to HIV-1 group M, whereas the largest number of HIV 1-human PPIs and isolates is associated with subtype B, followed by subtypes A and D. The information about the groups and subtypes for the significant number of interactions (995 PPIs) is not presented in the databases. The merging of data on different groups and subtypes for the network analysis is not strictly correct, because the corresponding interaction patterns may be different (Neveu et al., 2012; Goodacre et al., 2018). Nevertheless, this finding is usually ignored, because the separate analysis of HIV-human PPI data is not possible for most of the groups and subtypes, except for group M subtype B, since they are associated with very few or no interactions (Table 2).

The available number of HIV-2-human interactions is lower than for HIV-1 (Table 1). One may suggest that this is observed because HIV-1 is studied more thoroughly than HIV-2 (Campbell-Yesufu and Gandhi, 2011).

The list of 2910 HIV-1-human PPIs is presented in Table S1.

2.1.2 Interactions between matrix and non-coding RNAs

There is growing evidence regarding the role of non-coding RNAs, such as micro RNA (miRNA) and long non-coding RNA (lncRNA), in HIV-human interactions (Lazar et al., 2016; Fruci et al., 2017; Balasubramaniam et al., 2018). Dozens of human miRNAs are known to bind the matrix RNAs (mRNAs) of HDFs and modulate HIV replication either positively or negatively. Human miRNAs can also provide anti-HIV activity through the direct targeting of the viral genome and viral mRNAs (Balasubramaniam et al., 2018). However, the HIV genome also encodes miRNA targeting human mRNA and miRNA as well as viral mRNAs. These interactions regulate apoptosis and cell survival, the events that give HIV-infected cells a survival advantage, and they have a positive impact on viral replication and provide an escape from the host immune response (Lazar et al., 2016; Fruci et al., 2017). Information about the interactions between HIV-1 and human RNAs can be obtained from three specialized databases, namely ViRBase[12] (Li et al., 2015), VmiReg[13] (Shao et al., 2015), and VIRmiRNA[14] (Qureshi et al., 2014), which integrate

---

[12] http://www.rna-society.org/virbase/index.html

[13] http://210.46.85.180:8080/vmireg

[14] http://crdd.osdd.net/servers/virmirna

the corresponding data from publications. The numbers of various types of interactions in the databases are shown in Table 3.

The ViRBase includes data on a few interactions between human lncRNAs and human or HIV mRNAs. For example, human lncRNAs 7SL and Y3 interact with members of the APOBEC3 family of human proteins, which possess antiviral activity through the deamination of viral RNA (Wang et al., 2007; Zhen et al., 2012). The lncRNA encoded by the NEAT1 human gene is involved in interactions with and posttranscriptional regulation of unspliced HIV-1 transcripts (Zhang et al., 2013).

The list of RNA-RNA interactions from the three databases is presented in Table S2.

2.2 Genomics, transcriptomics, proteomics and epigenomics data

2.2.1 Genomics data

The high mutation rate of the HIV genome is one of the reasons for the formation of multiple HIV variants, which have new antigen properties that allow the virus to escape from an immune response (Rogozin et al., 2005; Cuevas et al., 2015). However, polymorphisms in the human genome can also influence the susceptibility to and severity of HIV infections (Le Clercet al., 2019; Tough et al., 2019). The candidate gene studies revealed a 32-bp deletion in the CCR5 gene (CCR5-Δ32) encoding a co-receptor that is essential for HIV binding to CD4+ cells. The homozygotes of the CCR5-Δ32 mutation are resistant to HIV, whereas heterozygotes demonstrate slower disease progression. Other significant gene loci related to the viral load and HIV progression are located in HLA genes. They encode the major histocompatibility complex, which presents peptides derived from HIV proteins to T-helpers, one of the critical stages of an immune response to infection. The genome-wide association studies (GWASs) allow for the linking of thousands and millions of genome polymorphisms, primarily single nucleotide polymorphisms (SNPs), to various phenotypes, e.g., the viral load, HIV susceptibility or disease progression (Le Clercet al., 2019; Tough et al., 2019). To date, more than 20 GWASs related to HIV were performed; however, they did not allow for the identification of new SNPs, which have strong, statistically significant and reproducible associations with HIV-associated phenotypes, such as viral load and capability to viremic control (Le Clercet al., 2019). Given GWASs confirmed the significance of CCR5 and HLA gene polymorphisms, whereas newly revealed associations were weak and usually non-reproducible. Nevertheless, a few human genes may be considered as potential candidates, e.g., CXCR6 encoding C-X-C chemokine receptor type 6, which is one of the co-receptors for the binding of the HIV-2 and m-tropic HIV-1 strains along with CCR5. A detailed description of HIV-related GWASs explored to date is given in the review (Le Clercet al., 2019). The corresponding associations between human genetic polymorphisms and HIV-related phenotypes can be obtained from the GWAS Catalog[15] or GWAS Central[16] resources.

[15] https://www.ebi.ac.uk/gwas

[16] https://www.gwascentral.org

2.2.2 Transcriptomics data

Information on the expression levels of coding and non-coding RNAs is the most represented OMICs data type in HIV research. The public data from HIV-related transcriptomics experiments can be downloaded from Gene Expression Omnibus[17] (GEO) and ArrayExpress[18]. We searched these databases with the keyword "HIV" and manually inspected the results. We found 232 transcriptomics experiments in which the mRNA, miRNA or lncRNA levels were measured by microarrays or RNA sequencing-based methods under different conditions, in different cell types, *in vivo* or *in vitro* (Table 4 and Table S3). Most of the experiments were focused on measuring mRNA profiles, and, to a lesser degree, on miRNA profiles, whereas only one study was related to the lncRNA profile. Comparisons of RNA transcription profiles between CD4+ cells obtained from HIV-infected or healthy individuals were the most common types of experiment; however, additional conditions were taken into account in most of the experiments. For example, over 30 experiments were related to the comparison of transcriptional profiles between chronic progressors that are in a chronic stage of the infection and develop AIDS without using ART, viremic controllers, and elite controllers (see Introduction). Most of the transcriptome measurements were performed on human cells derived from HIV-infected and healthy donors; however, in some cases, experiments were performed on human cell lines or primary cells infected with HIV *in vitro*. In addition, three experiments were focused on the transcriptomes of human intestinal, lung and oral microbiota from HIV-infected and healthy subjects (see Table 4), because they seem to play an essential role in HIV progression and the development of opportunistic infections (Dang et al., 2012; Vujkovic-Cvijin et al., 2013; Iwai et al., 2014).

In addition to uncovering plenty of experiments in which microarrays and bulk RNA sequencing were used to measure the transcriptome, we found six studies for which single-cell RNA sequencing (Chen et al., 2019; Kulkarni et al., 2019) was applied (Table S4). Single-cell RNA sequencing allows for the identification of clusters of blood cells or cell subtypes with different transcriptional responses to HIV infection. For example, Farhadian S.F. and colleagues performed single-cell RNA sequencing on cerebrospinal fluid and blood from adults with and without HIV. They found a rare (< 5% of cells) subset of myeloid cells that are found only in cerebrospinal fluid and present a gene expression signature that overlaps significantly with neurodegenerative disease-associated microglia. This immune cell subset may perpetuate neuronal injury during HIV infection (Farhadian et al., 2018).

Most of the HIV-related experiments were performed to identify differentially expressed genes (DEGs) between two or more conditions with subsequent functional annotation. The functional annotation of the DEGs is usually performed by pathway enrichment analysis (Khatri et al., 2012; Kuleshov et al., 2016), which allows for the identification of pathways, Gene Ontology[19] biological processes or other functional categories that were

---

[17] https://www.ncbi.nlm.nih.gov/geo

[18] https://www.ebi.ac.uk/arrayexpress

[19] http://geneontology.org

"enriched" by DEGs compared to the background gene set, e.g., all human genes. For example, Devadas K. and colleagues identified the transcriptional changes in the peripheral blood mononuclear cells during HIV-1 and HIV-2 infection (Devadas et al., 2016). HIV-1 caused changes in the expression of 316 genes, whereas HIV-2 changed the expression of only 57 genes. The pathway enrichment analysis allowed for the identification of the pathways and Gene Ontology biological processes associated with the infection. The authors found that the pathways and cellular processes perturbed by HIV-1 and HIV-2 are not the same, e.g., only the HIV-1 virus influenced genes related to the cell cycle and apoptosis. The observed differences explain the different rates of disease progression from HIV-1 and HIV-2 and the observations showed that HIV-2 is generally less pathogenic than HIV-1 (Devadas et al., 2016).

It should be noted that some transcriptomics studies were performed to measure the gene expression changes under the application of anti-HIV vaccine candidates (Zak et al., 2012; Fourati et al., 2019). Currently, there are no effective anti-HIV vaccines despite many attempts to create them; thus, understanding the cellular mechanisms of the immune response at the transcriptome level to more or less effective vaccine candidates, and in individuals with stronger or weaker vaccine effect may help researchers to develop more effective ones (Haynes et al., 2016; Trovato et al., 2018).

Nine transcriptomics studies are related to another important problem: HIV latency. The comparison of transcriptional profiles of latently and productively infected cells, as well as cells treated with different latency reversing agents may allow revealing mechanisms of latency and identifying new more effective solutions for "shock and kill" or "block and lock" approaches. For example, White C.H. and colleagues compared transcription profiles of uninfected and latently infected central memory cells, and identified 826 DEGs, many of which were related to p53 signaling. The authors found that inhibition of the transcriptional activity of p53 during HIV-1 infection reduced the ability of HIV-1 to be reactivated from its latent state. Their observations may help to develop new latency reversing agents (White et al., 2016).

In addition to the differential expression analysis, transcriptomics data may be used for the construction of co-expression and gene regulatory networks (see below).

Beside GEO and ArrayExpress resources, specialized HIVed database[20] also provides access to some HIV-related transcriptomics and proteomics experiments derived from literature. It contains data on gene expression levels during HIV infection and replication as well as at HIV latency (Li et al., 2017).

The list of HIV-related transcriptomics experiments prepared in this study is presented in Table S3.

---

[20] http://hivlatency.erc.monash.edu

2.2.3 Proteomics data

The mass spectrometry-based proteomics approaches allow researchers to measure the absolute or relative amounts of proteins in human cells under various conditions. The public proteomics data describing HIV-human interaction can be obtained from ProteomeXChange[21], which contains information from the PRIDE Archive[22] and some other sources. Several datasets are also published in GEO. We manually inspected the corresponding databases and performed a search in PubMed using the keywords "HIV" and "proteome," and we found 25 published proteomics studies as well as two that have not been linked to publications (see Table S5). Most of these studies are focused on identifying differentially expressed proteins (DEPs) between the same conditions as those in the transcriptomics studies (see above); however, some of the studies have features unique to proteomics experiments. For example, Golumbeanu M. and colleagues measured both the proteomics and transcriptomics responses to HIV-1 infection in SupT1 CD4+ T cells at five time points (Golumbeanu et al., 2019). The simultaneous assessment of changes in the transcriptome and proteome under the same cells and other experimental conditions is extremely important for the building of *in silico* integrative models that provide the most accurate results. Unfortunately, other published studies were only focused on the changes in the proteome. Three studies were related to measurements of the changes in the post-translational modifications of human proteins during HIV infection (Greenwood et al., 2016; Yang et al., 2016; Lapek et al., 2017). Greenwood E.J. and colleagues performed an analysis of more than 6500 HIV and cellular proteins in the CEM-T4 cell line infected by wild type and Vif-deficient viruses. Among others, they measured changes in phosphoproteome and found Vif-dependent hyperphosphorylation in more than 200 cellular proteins, particularly the substrates of the aurora kinases (Greenwood et al., 2016). Since the distinct profile of glycosylated surface proteins can be used for targeting latently infected cells, Yang W. with colleagues measured differences in the glycol-proteome in ACH-2 and A3.01 cell lines, which are models of latently infected and uninfected cells. They identified a change in the levels of 236 glycosite-containing peptides from 172 glycoproteins between two cell lines. These proteins participate in cell adhesion, immune response, glycoprotein metabolism, cell motion, and cell activation (Yang et al., 2016). The study by Zheng D. with colleagues (Zheng et al., 2011) aims at the identification of differences between immune response to different opportunistic infections (Epstein-Barr virus and Kaposi's Sarcoma) based on proteome analysis.

Most proteomics experiments were performed by mass spectrometry-based approaches; however, a few studies used protein arrays, which are usually focused on particular sets of proteins. For example, Yang G. and co-authors aimed at the identification of autoantigens recognized by the bNAbs to 2F5 and 4E10 epitopes of HIV-1 gp41. The corresponding protein array contained more than 9400 recombinant human proteins. As a result, the authors found human kynureninase and splicing factor 3b subunit 3 acting as human self-antigens (Yang et al., 2013).

---

[21] http://www.proteomexchange.org

[22] https://www.ebi.ac.uk/pride/archive

The list of HIV-related proteomic experiments is presented in Table S5.

2.2.4 Epigenomics data

HIV-related epigenomics studies are focused on the changes in DNA methylation profiles (Zhang et al., 2016; Zhang et al., 2018; Liu et al., 2019), post-translational modifications of histones (Marban et al., 2011), genome-wide chromatin accessibility (Johnson et al., 2018), and the identification of DNA binding/occupancy sites for human/HIV proteins (Marban et al., 2011) based on ChIP-chip, ChIP-seq, and ATAC-seq approaches (Table S6). The corresponding data can be obtained through the GEO database. In total, we found 34 studies related to changes in the epigenome under HIV infection. Importantly, some of them have linked transcriptomics data, which was obtained during the same experiments. The list of HIV-related epigenomics experiments is presented in Table S6, whereas some examples are given below.

In a study by Zhang X. and colleagues, the methylation profiles of CpGs were measured in smoking and non-smoking HIV patients. The authors found 137 differentially methylated CpGs between these two groups. The smoking-associated DNA methylation features were used to made prognosis of HIV infection progression to AIDS based on the ensemble-based machine learning approach. The resulting area under the curve (AUC) was 0.73 (Zhang et al., 2018).

Marban C. and colleagues searched the HIV-1 Tat protein binding sites in the human genome using the ChIP-seq approach. In parallel, they performed the corresponding transcriptomic study as well as ChIP-chip experiments to identify histone H3 acetylation sites. All the measurements were taken in Jurkat-Tat and Jurkat cells. The authors found that only ~ 7% of the Tat-bound regions are near transcription start sites at gene promoters, whereas ~ 53% of the Tat target regions are within DNA repeat elements. They also revealed that Tat binding sites are not significantly associated with DEG promoters, whereas changes in histone H3 lysine 9 acetylation are significantly associated (Marban et al., 2011).

Johnson J.S. and colleagues used ATAC-seq technology to identify accessible DNA regions during a time-course on HIV infection in monocyte-derived macrophages. They found that HIV-1 "primes" the chromatin accessibility of innate immune genes, such as IRF family members, NF-B factors, and type I and type III interferons, before and after virus integration (Johnson et al., 2018).

2.2.5 Data on functional genomics screenings

To identify the HDFs, a genome-wide inactivation of gene expression through small interfering RNA (siRNA) and small hairpin RNA (shRNA) can be performed. In three large-scale studies conducted in 2008, 842 human genes were identified as HDFs (Brass et al., 2008; König et al., 2008; Zhou et al., 2008). Brass A.L., König R. and Zhou H. along with their colleagues identified 273, 295 and 230 genes as potential HDFs using corresponding siRNAs transfected into HeLa-derived TZM-bl, HEK-293 and HeLa P4-R5 cell lines; however the percentages of shared genes between the studies were minimal, ranging from 3-6 % (Bushman et al., 2009). A year later, Yeung M.L. with colleagues

performed the corresponding screening on the Jurkat cell line using shRNAs and identified 252 potential HDFs, which also have minimal overlap with genes from three previous studies (Yeung et al., 2009). These results can be explained by differences in the cell lines or type of reporter, experimental noise, and differences between time points and filtering thresholds (Bushman et al., 2009; Tough and McLaren, 2019). Bushman F.D. and co-authors performed a Gene Ontology enrichment analysis for genes from three first screening studies and found that the overlaps between the identified cellular processes are higher than they are at the individual gene level (Bushman et al., 2009). Thus, the revealed potential HDFs may represent at least a basis for further research with more accurate methods and primary human CD4+ cells.

Recently, to identify potential HDFs, Park R.J. and colleagues performed genome-wide knockout screening using CRISPR-Cas9 lentiviral single-guide RNA constructs, which have higher sensitivity and specificity (Park et al., 2017) than screens based on RNA interference. Their research was conducted on GXRCas9 cells and allowed for the identification of 5 HDFs, with HIV co-receptors CD4 and CCR5, which are well-known targets of the anti-HIV drugs maraviroc and ibalizumab (Kinch and Patridge, 2014; Puhl et al., 2019) as well as TPST2, SLC35B2 and ALCAM, which are new potential targets. Genes TPST2 and SLC35B2 encode tyrosylprotein sulfotransferase 2 and solute carrier family 35 member B2, functioning in the same pathway as sulfate CCR5, which facilitates its recognition by the HIV envelope. The ALCAM gene encodes activated leukocyte cell adhesion molecule-mediating cell aggregation, which is required for cell-to-cell HIV transmission. The results were validated in primary human CD4+ T cells through a Cas9-mediated knockout and an antibody blockade. These three human proteins can be used in further research as new potential anti-HIV targets.

3   Network-based integration and analysis of OMICs data describing HIV-host interaction

The general pipeline of network-based analysis of HIV-related OMICs data is shown in the Figure 2.

3.1   Protein-protein interaction network-based analysis

3.1.1 Creation of protein-protein interaction networks

Protein-protein interaction networks (PPI networks) are the most common type of networks used in HIV-related studies. This is due to the high amount and availability of interaction data between human proteins (Csermely et al., 2013; Miryala et al., 2018) as well as human and HIV proteins (see Table 1 and Table S1). The nodes of the PPI network are proteins, whereas the edges represent the direct or physical interactions between them, which are usually obtained by high-throughput *in vitro* experiments (Gillen and Nita-Lazar, 2019). The network may contain only one type of node representing human proteins, or two types of nodes representing both human and HIV proteins. Since PPI networks are derived from *in vitro* experiments, they contain proteins, which are not expressed in CD4+ cells and all the possible interactions that may take place in all the tissues, cell types, and conditions. To create context-specific networks representing molecular interactions under specific HIV-related conditions, the integration

of the PPIs with other types of OMICs data is required (Figure 2). This integration can be performed in at least three ways. First, all the interactions can be measured in specific cells, under particular conditions, e.g., HIV-infected CD4+ T cells (Luo et al., 2016); however, due to their high cost, this type of interactomics experiment is not usually performed. Second, subnetworks containing only human proteins essential for the HIV life cycle can be created. Human proteins encoded by DEGs, immediate DEPs, HDFs derived from functional genomic screens or human proteins that are physically interacting with HIV can be used for this purpose (van Dijk et al., 2010; Xu et al., 2014; Shityakov et al., 2015; Gao et al., 2018; Saha et al., 2018). Third, the context-specific PPI network can be created from a global network by eliminating proteins that are not expressed under particular experimental conditions (Kotlyar et al., 2016; Basha et al., 2017), or by changing the edge weights using transcriptomics, proteomics and epigenomics data (Chen and Li, 2016; Li and Chen, 2018; Csősz et al., 2019). In addition to pure PPI networks, which contain only human/HIV proteins, integrated networks can be created. These networks include various types of nodes representing proteins, mRNAs, non-coding RNAs, or genes, and help in obtaining more accurate results than when only PPIs are studied (Chen and Li, 2016; Li and Chen, 2018). For example, Li C.W. and Chen B.S. created HIV-1-human interspecies protein-protein and miRNA interaction networks for different stages of HIV infection in CD4+ T cells, with reverse transcription, integration/replication, and the late stages of the HIV life cycle. The networks differed by edge weights, which were calculated using the corresponding transcriptomics data and stochastic dynamic modeling (Li and Chen, 2018).

The purposes of analyzing HIV-related PPI networks include the identification of human proteins essential for the HIV life cycle based on the topological characteristics of the networks, the identification of dense communities of proteins in the network, which may perform similar functions in the cell, and predictions of new potential HDFs by gene-phenotype prioritization algorithms.

3.1.2 Identification of essential proteins in a protein-protein interaction network

The degree of protein found in a network is simply the number of direct interactions with other proteins. The degree in real PPI networks has scale-free distribution, so most proteins have a low number of interactions, but small amount of proteins have a high number of interactions. The last type of proteins is called "hubs," and they are critical for cell viability. The disturbance of their function can lead to cell death or carcinogenesis (Csermely et al., 2013; Miryala et al., 2018). In addition to "hubs", the PPI networks contain proteins, called "bottlenecks," which have few interactions but exclusively connect distinct modules and are therefore critical for cell survival. These proteins have high centrality measures (Csermely et al., 2013; Miryala et al., 2018; Li et al., 2019). Two types of centrality are usually used, closeness centrality and betweenness centrality. The closeness centrality is the average length of the shortest paths between the protein and all the other proteins in the network. The betweenness centrality quantifies the number of times a protein acts as a bridge along the shortest path between two other proteins in the network. The degree and centrality measures can be used to identify the essential human proteins influencing HIV-human interactions (Dickerson et al., 2010; van Dijk et al., 2010; Huang et al., 2011; Li et al., 2013; Ma et al., 2013; Bandyopadhyay

et al., 2015; Xie et al., 2015). For example, Huang T. and colleagues compared gene expression profiles from CD4+ T cells between HIV-1-resistant and susceptible subjects using Minimum Redundancy-Maximum Relevance and Incremental Feature Selection algorithms. They identified 185 genes for which the expression levels distinguished between HIV-resistant and susceptible individuals with 85.2 % accuracy. The authors identified 29 proteins from the 185 total based on the calculation of modified betweenness centrality in the HIV-1-human PPI network. This network included both interactions between human proteins as well as interactions between human and HIV-1 proteins. The modified betweenness centrality reflects the number of times a human protein acts as a bridge along the shortest path between two HIV-1 proteins. Twenty-nine identified human proteins may be considered as targets for the disruption of communication between virus-targeted proteins and the prevention of viral infection (Huang et al., 2011). Shityakov S. and colleagues compared transcription profiles in frontal cortex between samples from AIDS patients with and without apparent features of HIV-associated encephalitis and dementia (Shityakov et al., 2015). They identified 1528 DEGs, which are mainly involved in the immune response, regulation of cell proliferation, cellular response to inflammation, signal transduction, and viral replication cycle. The authors created human PPI network containing only proteins encoded by DEGs, and identified hubs, e.g., heat-shock protein alpha, class A member 1, and fibronectin 1, which seem to play essential roles in the pathogenesis of HIV-associated encephalitis (Shityakov et al., 2015).

3.1.3 Identification of clusters (modules) and biclusters in a protein-protein interaction network

At the higher level of network organization, modular structures can be found. The network module (cluster) is a dense community of nodes, which are highly interconnected with one another but weakly connected to other nodes in the network (Csermely et al., 2013; Miryala et al., 2018; Wu et al., 2019). The proteins from the module usually perform the same biological functions, e.g., a module may represent the dense part of the signaling pathway or complex cellular machinery such as one related to the DNA polymerase protein complex. The identification of modules in human PPI networks integrated with HIV-related OMICs data with subsequent functional annotation using pathway enrichment analysis allows researchers to reveal the particular mechanisms of HIV-host interactions (Xu et al., 2014; Amberkar and Kaderali, 2015; Yang et al., 2019). For example, Amberkar S.S. and Kaderali L. identified modules in the human PPI network using the ClusterONE algorithm and found that they are enriched by potential HDFs from functional genomic screens (see above) (Amberkar and Kaderali, 2015). A comparison of various topological characteristics between the enriched and non-enriched modules allows for the identification of two clusters with unique features. Gene Ontology and pathway enrichment analysis revealed that proteins from the first module were involved in gene transcription, whereas proteins from the second module participated in mRNA processing and splicing. Interestingly, the transcriptional regulation was not revealed in enrichment analyses of the individual screens, which may indicate the importance of module identification before enrichment because the inclusion of protein neighborhoods from corresponding clusters in the analysis may increase its power and sensitivity (Amberkar and Kaderali, 2015).

Information on the interactions between human and HIV proteins can be used to identify biclusters (Figure 2), which contain human proteins that share a common set of HIV proteins (MacPherson et al., 2010; Maulik et al., 2011; Maulik et al., 2013). These topological structures are called "biclusters" because they contain two types of nodes, human and HIV proteins. The human proteins from biclusters represent the gateway proteins that are most affected by the HIV infection. Maulik U. and colleagues identified 14 overlapping biclusters using the original algorithm and bipartite network with two types of nodes corresponding to human and HIV-1 proteins. These biclusters formed a strongly connected subnetwork containing 7 HIV-1 and 19 human proteins. The list of corresponding human proteins is enriched with kinases and other proteins from signaling pathways (Maulik et al., 2011).

3.1.4 Gene-phenotype prioritization algorithms for identifying potential HDFs

The proteins from the dense clusters in the PPI network usually perform the same biological functions. This finding underlies many gene-phenotype prioritization algorithms, which are used to predict gene participation in cellular processes as well as to predict new associations between genes and human diseases (Lan et al., 2015; Fiscon et al., 2018). The corresponding methods can be divided into the following two groups: (1) local methods, which are based on the search for direct interactions between candidate genes and genes with a known phenotype; and (2) global methods, which model how the information flow in the PPI network is used to assess the proximity and connectivity between genes with the established phenotype and candidate genes. The general principle of algorithms is "the closer candidate genes to known phenotype genes in the network, e.g., they are in the same module, the higher score of the algorithm will be obtained." Gene-phenotype prioritization algorithms and PPI networks were used to predict new HDFs essential for HIV-host interaction (Murali et al., 2011; Emig-Agius et al., 2014). For example, Murali T.M. used the original SinkSource algorithm and human PPI network to predict new potential HDFs that influenced HIV-host interactions (Murali et al., 2011). They used 545 genes from the three earliest genomic functional screens (Brass et al., 2008; König et al., 2008; Zhou et al., 2008) (see above) that were present in the PPI network as known HDFs. The Gene Ontology enrichment analysis, which was performed for the top 1000 predicted proteins, revealed that the obtained cellular processes are HIV-related, e.g., in RNA splicing, translation initiation, oxidative phosphorylation, and others. The authors also found that a significant number of the obtained potential HDFs interacted with HIV proteins. The predicted HDFs, along with those derived from genomic screens, may represent potential pharmacological targets for treating HIV infections.

3.2   Co-expression network-based analysis

Co-expression is the simultaneous expression of two or more genes so that the transcription of two genes changed similarly under different conditions. The co-expression of two genes may reflect the transcription regulation by the same transcription factors. To determine co-expression, various measures can be employed, e.g., Pearson or Spearman correlation coefficients, mutual information, or Euclidean distance (van Dam et al., 2018). The unweighted co-expression network can be constructed by selecting a threshold on a co-expression measure so that the edge between the two genes exists when

17

the corresponding value is higher than the threshold. The weighted co-expression networks contain all the possible edges between all the genes for which the weights are calculated using several functions from co-expression measures. A co-expression network can be created for particular cell types and conditions using microarray or RNA sequencing-based data (van Dam et al., 2018). The most common type of co-expression analysis is related to the identification of network modules. The co-expression module contains genes that could be regulated by the same transcription factors and have similar biological functions as in modules from the PPI networks. The most popular method for the creation of co-expression networks and the analysis of modules is the weighted gene correlation network analysis (WGCNA) (Zhang and Horvath, 2005; Langfelder et al., 2008; van Dam et al., 2018). The WGCNA can be used to create co-expression networks, identify modules, estimate module preservation between two networks created for different conditions, reveal modules associated with a clinical trait of interest and find intermodular "hubs," which could be the essential genes regulating the expression of the other genes in the module. To date, several studies have been focused on the creation and analysis of co-expression networks for HIV-related conditions (Ma et al., 2011; Levine, Horvath et al., 2013; Levine, Miller et al., 2013; Xu et al., 2013; Ray and Bandyopadhyay, 2016; Mosaddek Hossain et al., 2017; Ray and Maulik, 2017; Quach et al., 2018; Ding et al., 2019; Nguyen et al., 2019). Most of them were based on transcriptomics data from CD4+ and CD8+ cells, which reflect different stages of HIV infection or progression types, e.g., chronic progressors, viremic controllers, elite controllers or subjects who were utterly resistant to the virus (Ma et al., 2011; Xu et al., 2013; Ray and Bandyopadhyay, 2016; Mosaddek Hossain et al., 2017; Ray and Maulik, 2017; Ding et al., 2019). Some other studies were focused on conditions related to HAND (Levine, Horvath et al., 2013; Levine, Miller et al., 2013; Quach et al., 2018). Almost all of the studies applied WGCNA to identify co-expression modules, which are preserved between different conditions, or modules related to a particular state, with subsequent Gene Ontology and pathway enrichment analysis of the modular genes.

Ray S. and colleagues created a co-expression network for the acute phase of HIV infection with the subsequent identification of modules using WGCNA algorithm and transcription data from CD4+ and CD8+ T cells (Ray and Bandyopadhyay, 2016). The analysis on the preservation of modules across chronic and non-progressor stages was performed using the original rank aggregation algorithm, which is based on the comparison of module ranks and determined from multiple module characteristics, between HIV stages. The authors found 30 modules in the network for the acute phase of HIV infection, but not all of them were preserved in other stages. The genes from modules, according to the pathway enrichment analysis, participate in processes related to the immune system, the regulation of transcription, RNA processing, splicing and translation, cell cycle and apoptosis, and energy derivation as well as cytoskeleton regulation. The authors also performed a transcription factor enrichment analysis and found some novel factors such as "FOXO1","GATA3", "GFI1","IRF1", "IRF7", "MAX", "STAT1", "STAT3", "XBP1" and "YY1", which emerged from the modules that showed significant changes in expression patterns over the HIV progression stages (Ray and Bandyopadhyay, 2016).

Ding J. and colleagues created several co-expression networks based on transcription data from the CD4+ and CD8+ T cells as well as whole blood obtained from subjects with different viremic control statuses, e.g., viremic controllers and elite controllers. They found significant positive associations between several modules and clinical parameters describing HIV progression and the loss of viremic control. The genes from these modules participate in immune-related pathways and cellular processes such as type I interferon signaling pathway, complement activation, the positive regulation of B cell activation, cellular responses to stress, leukocyte migration, and responses to cytokines (Ding et al., 2019).

Zak D.E. and colleagues used the time-serial measurement of gene expression profiles in peripheral blood mononuclear cells, which were derived from subjects who were provided with MRKAd5/HIV vaccination (Zak et al., 2012). These researchers used modular analysis framework to deconvolute complex transcriptional profiles into functionally interpretable patterns. They performed pathway enrichment analysis of genes from revealed modules and found increased expression of genes associated with inflammation, interferon response, and myeloid cell trafficking, and decreases in lymphocyte-specific transcripts, which leads to the hypothesis that the vaccine was stimulating an influx of myeloid cells and an efflux of lymphoid cells from the circulation. The authors also compared the gene expression profiles between subjects who have different magnitudes of HIV-specific CD8+ T-cell responses to the vaccine, and they identified 209 DEGs that were associated with cytotoxic responses, including inhibitory killer cell Ig-like receptor KIR2DL1, the NK-cell activating receptor CLEC2D, and the NK-cell signaling adaptor EWS-FLI1-activated transcript 2 (EAT-2). This finding is important because the adenoviral expression of EAT-2 enhanced vaccine-induced T-cell responses as part of a vaccine strategy (Zak et al., 2012).

3.3   Signed network-based analysis

The signed network is a type of molecular network in which (1) all edges have a direction representing a signal flow from the source to the target node; and (2) all edges are either positive (standing for activation) or negative (representing inhibition). Signaling and gene regulatory networks are the most commonly used networks of this type (Csermely et al., 2013). The signaling network consists of signed direct interactions between proteins, RNAs, and secondary messengers, and it represents the signal flow from receptors to transcriptional factors or other effector molecules. The gene regulatory network consists of signed, mostly indirect interactions between genes, where edges represent how one gene can change the transcription of another gene, either up- or down-regulation (Csermely et al., 2013). The signaling pathways are usually manually created by experts based on a great deal of information regarding the PPIs, post-translational modifications, siRNA-based genetic knockdowns, and data types. Information about signaling pathways can be obtained from various databases such as KEGG[23], Reactome[24], NetPath[25],

---

[23] https://www.genome.jp/kegg

[24] https://reactome.org

[25] http://www.netpath.org

SPIKE[26], and Signor 2.0[27]. A gene regulatory network can be created using reverse engineering methods that are applied to gene transcription data perturbed multiple times, e.g., by siRNA to different genes or small molecule inhibitors (Csermely et al., 2013).

Signed networks can be used in HIV-related research for (1) identifying the motifs of directed interactions between HIV and human proteins (van Dijk et al., 2010; Biswas et al., 2019); and (2) creating dynamic models of HIV interaction with human cells (Oyeyemi et al., 2015; Bensussen et al., 2018).

Motifs are chains or contours of 3-6 vertices in a directed network that are much more common than they are in a random network. These building blocks have been used to study the structure and dynamic behavior of networks. Van Dijk D. with colleagues identified several motifs consisting of human and HIV proteins, e.g., positive feedback, positive and negative co-regulation, co-activation motifs, activation, and inhibition cliques. These motifs may have an essential role in HIV-human interactions. For example, the three-node feedback loop motif, which was identified as indirect self-regulation, is a pattern in which an HIV protein regulates or signals a human protein that regulates/signals another HIV protein in turn (van Dijk et al., 2010).

The dynamic of HIV-host interactions can be simulated by creating discrete and continuous models. Oyeyemi O.J. and colleagues applied Boolean discrete modeling to the T-cell activation signaling pathway containing both HIV and human proteins. They found that the model reproduced the expected patterns of T-cell activation, co-stimulation, and co-inhibition. Through *in silico* knockouts, the model identified an additional nine HDFs, including members of the PI3K signaling pathway that are essential for viral replication. The revealed potential HDFs were retrospectively confirmed by comparison with the results of three functional genomic screens (Oyeyemi et al., 2015). Bensussen A. and co-authors created a gene regulatory network of latent proviruses in resting CD4+ T cells containing human and HIV proteins as well as HIV non-coding RNAs. They applied both Boolean and continuous mathematical models, which were based on ordinary differential equations, to simulate the latency reversion. The authors found that viral non-coding RNAs can counteract the activity of latency reversing agents, which may explain the failure of these compounds to reactivate the latent reservoirs of HIV. They also found that inhibitors of histone methyltransferases, together with releasers of the positive transcription elongation factor (P-TEFb), may increase proviral reactivation despite the self-repressive effects of viral non-coding RNAs (Bensussen et al., 2018).

4   Conclusions

HIV/AIDS remains one of the most significant dangers for humankind. Although the existing antiretroviral therapy allows for the control of the virus and prevents transmission, HIV infection remains a global health problem due to the impossibility of eliminating the virus from the human body and problems with creation of anti-HIV

---

[26] https://www.cs.tau.ac.il/~spike

[27] https://signor.uniroma2.it

vaccine. To overcome the limitations of the existing anti-HIV therapy, to improve its efficacy and safety, and to develop new therapeutic approaches, such as therapeutic and preventive vaccines as well as approaches to cure latent infection, a deeper understanding of the mechanisms of HIV-human interaction is required. A network-based analysis of OMICs data may shed light on the corresponding mechanisms and allows for the identification of new points in therapeutic interventions. To date, more than 2900 interactions between human and HIV-1 proteins belonging to different viral groups and subtypes have been added to public databases. Most of these interactions belong to HIV-1 group M subtype B, whereas other groups and subtypes are associated with very few or no interactions, which results in difficulties in the network-based analysis of other HIV variants. We found 232 HIV-related transcriptomics experiments in which the expression profiles of human and HIV coding and non-coding RNAs were measured in various cell types, *in vivo* and *in vitro*, under different conditions. Different individuals may have different levels of susceptibility to HIV, disease progression rates, and different responses to drug and vaccine treatments. The identification of differentially expressed genes between various conditions with the pathway enrichment analysis allows researchers to explain the differences in these conditions. Particularly, the comparison of transcriptional profiles in various immune cells from individuals treated with more or less effective vaccines, with stronger or weaker immune response to particular vaccine allows identifying genes and pathways, which modulation by adjuvants may increase vaccine efficacy. Similarly, comparison of transcriptional profiles between uninfected and latently infected cells allows identifying unknown mechanisms of latency, which may help to develop new more effective latency reversing agents or agents causing transcriptional silencing of integrated HIV genome ("shock and kill", "block and lock" strategies). Comparison of transcriptional profiles between chronically infected individuals, viremic and elite controllers allows identifying mechanisms of decreased susceptibility to HIV infection. Mimicking the corresponding transcriptional profiles of viremic and elite controllers by various chemical and biological agents may increase the efficacy of antiretroviral therapy and vaccines. The creation of co-expression networks with the subsequent identification of dense gene clusters allows for the identification of new HIV-related pathway and cellular processes that cannot be obtained through a simple analysis of differentially expressed genes. In addition to protein-protein interactions and transcriptomics data, a few datasets on the interactions between human and viral RNAs, genomics, proteomics, epigenomics data as well as data from functional genomic screenings are currently available. The integration of human and human-HIV protein-protein interactions with other types of OMICs data allows for the creation of context-specific networks reflecting particular experimental and clinical conditions. Networks reflecting different degree of susceptibility to HIV infection (chronically infected patients, viremic and elite controllers), productive or latent HIV infection can be created. The identification of modules in human context-specific protein-protein interaction networks, as in co-expression network, allows for the identification of more HIV-related pathways and cellular processes than through simple comparison of transcriptomic and proteomic profiles. The analysis of the topology of context-specific human or human-HIV protein-protein interaction networks may assists in identifying the proteins with high degree or centrality, which are essential for HIV-human interaction, and may represent the most

perspective human targets to prevent infection of human CD4+ cells, reactivate or silent latent HIV infection, and model the efficacy of immune response induced by vaccines.

Since there are hundreds of publicly available transcriptomic experiments that were performed under many conditions as well as thousands of known human and HIV-human protein-protein interactions, whereas only a small portion of them have been used in network-based analyses, this gap provides an opportunity to create many novel network-based models and potentially obtain new knowledge on HIV-human interaction mechanisms.

In conclusion, we note that the network-based analysis of OMICs data is a prospective area of research that may help us to understand the mechanisms of HIV-host interactions and improve existing and develop new anti-HIV therapeutic strategies.

5   Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6   Author Contributions

All authors contributed to the conceptualization, writing, reviewing, and editing of this article.

7   Funding

8   References

Ako-Adjei, D., Fu, W., Wallin, C., Katz, K.S., Song, G., Darji, D., et al. (2015). HIV-1, human interaction database: current status and new features. Nucleic Acids Res. 43, D566-570.

Alvarellos, M., Guillemette, C., Altman, R.B., Klein, T.E. (2018). PharmGKB summary: atazanavir pathway, pharmacokinetics/pharmacodynamics. Pharmacogenet. Genomics. 28, 127-137.

Amberkar, S.S., and Kaderali, L. (2015). An integrative approach for a network based meta-analysis of viral RNAi screens. Algorithms Mol. Biol. 10:6. doi: 10.1186/s13015-015-0035-7

Ammari, M.G., Gresham, C.R., McCarthy, F.M., Nanduri, B. (2016). HPIDB 2.0: a curated database for host-pathogen interactions. Database (Oxford). 2016, baw103. doi: 10.1093/database/baw103

Arhel, N., and Kirchhoff, F. (2010). Host proteins involved in HIV infection: new therapeutic targets. Biochim. Biophys. Acta. 1802, 313-321.

Balasubramaniam, M., Pandhare, J., Dash, C. (2018). Are microRNAs Important Players in HIV-1 Infection? An Update. Viruses. 10:E110. doi: 10.3390/v10030110

Bandyopadhyay, S., Ray, S., Mukhopadhyay, A., Maulik, U. (2015). A review of in silico approaches for analysis and prediction of HIV-1-human protein-protein interactions. Brief. Bioinform. 16, 830-851.

Barouch, D.H., Stephenson, K.E., Borducchi, E.N., Smith, K., Stanley, K., McNally, A.G., et al. (2013). Protective efficacy of a global HIV-1 mosaic vaccine against heterologous SHIV challenges in rhesus monkeys. Cell. 155, 531-539.

Basha, O., Barshir, R., Sharon, M., Lerman, E., Kirson, B.F., Hekselman, I., et al. (2017). The TissueNet v.2 database: A quantitative view of protein-protein interactions across human tissues. Nucleic Acids Res. 45, D427-D431.

Bensussen, A., Torres-Sosa, C., Gonzalez, R.A., Díaz, J. (2018). Dynamics of the Gene Regulatory Network of HIV-1 and the Role of Viral Non-coding RNAs on Latency Reversion. Front Physiol. 9:1364. doi: 10.3389/fphys.2018.01364

Biswas, S., Ray, S., Bandyopadhyay, S. (2019). Colored Network Motif Analysis by Dynamic Programming Approach: An Application in Host-Pathogen Interaction Network. IEEE/ACM Trans. Comput. Biol. Bioinform. doi: 10.1109/TCBB.2019.2923173

Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., et al. (2008). Identification of host proteins required for HIV infection through a functional genomic screen. Science. 319, 921-926.

Brett-Major, D.M., Crowell, T.A., Michael, N.L. (2017). Prospecting for an HIV vaccine. Trop. Dis. Travel. Med. Vaccines. 3:6. doi: 10.1186/s40794-017-0050-4

Bushman, F.D., Malani, N., Fernandes, J., D'Orso, I., Cagney, G., Diamond, T.L., et al. (2009). Host cell factors in HIV replication: meta-analysis of genome-wide studies. PLoS Pathog. 5:e1000437. doi: 10.1371/journal.ppat.1000437

Calderone, A., Licata, L., Cesareni, G. (2015). VirusMentha: a new resource for virus-host protein interactions. Nucleic Acids Res. 43, D588-D592.

Campbell-Yesufu, O.T., and Gandhi, R.T. (2011). Update on human immunodeficiency virus (HIV)-2 infection. Clin. Infect. Dis. 52, 780-787.

Castiglione, F., Pappalardo, F., Bernaschi, M., Motta, S. (2007). Optimization of HAART with genetic algorithms and agent-based models of HIV infection. Bioinformatics. 23, 3350-3355.

Chen, B.S., and Li, C.W. (2016). Constructing an integrated genetic and epigenetic cellular network for whole cellular mechanism using high-throughput next-generation sequencing data. BMC Syst. Biol. 10:18. doi: 10.1186/s12918-016-0256-5

Chen, G., Ning, B., Shi, T. (2019). Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. Front. Genet. 10:317. doi: 10.3389/fgene.2019.00317

Chen, L., Keppler, O.T., Schölz, C. (2018). Post-translational Modification-Based Regulation of HIV Replication. Front. Microbiol. 9:2131. doi: 10.3389/fmicb.2018.02131

Clifford, D.B. (2017). HIV-associated neurocognitive disorder. Curr. Opin. Infect. Dis. 30, 117-122.

Cook, H.V., Doncheva, N.T., Szklarczyk, D., von Mering, C., Jensen, L.J. (2018). Viruses.STRING: A Virus-Host Protein-Protein Interaction Database. Viruses. 10:E519. doi: 10.3390/v10100519

Csermely, P., Korcsmáros, T., Kiss, H.J., London, G., Nussinov, R. (2013). Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review. Pharmacol. Ther. 138, 333-408.

Csősz, É., Tóth, F., Mahdi, M., Tsaprailis, G., Emri, M., Tőzsér, J. (2019). Analysis of networks of host proteins in the early time points following HIV transduction. BMC Bioinformatics. 20:398. doi: 10.1186/s12859-019-2990-3

Cuevas, J.M., Geller, R., Garijo, R., López-Aldeguer, J., Sanjuán, R. (2015). Extremely High Mutation Rate of HIV-1 In Vivo. PLoS Biol. 13:e1002251. doi: 10.1371/journal.pbio.1002251

Dang, A.T., Cotton, S., Sankaran-Walters, S., Li, C.S., Lee, C.Y., Dandekar, S., et al. (2012). Evidence of an increased pathogenic footprint in the lingual microbiome of untreated HIV infected patients. BMC Microbiol. 12:153, doi: 10.1186/1471-2180-12-153.

Darcis, G., Van Driessche, B., Van Lint, C. (2017). HIV Latency: Should We Shock or Lock? Trends Immunol. 38, 217-228.

Devadas, K., Biswas, S., Haleyurgirisetty, M., Wood, O., Ragupathy, V., Lee, S., et al. (2016). Analysis of Host Gene Expression Profile in HIV-1 and HIV-2 Infected T-Cells. PLoS One. 11:e0147421. doi: 10.1371/journal.pone.0147421

Dickerson, J.E., Pinney, J.W., Robertson, D.L. (2010). The biological context of HIV-1 host interactions reveals subtle insights into a system hijack. BMC Syst. Biol. 4:80. doi: 10.1186/1752-0509-4-80

Ding, J., Ma, L., Zhao, J., Xie, Y., Zhou, J., Li, X., et al. (2019). An integrative genomic analysis of transcriptional profiles identifies characteristic genes and patterns in HIV-

infected long-term non-progressors and elite controllers. J. Transl. Med. 17:35. doi: 10.1186/s12967-019-1777-7

Dirk, B.S., Pawlak, E.N., Johnson, A.L., Van Nynatten, L.R., Jacob, R.A., Heit, B., et al. (2016). HIV-1 Nef sequesters MHC-I intracellularly by targeting early stages of endocytosis and recycling. Sci. Rep. 6:37021. doi: 10.1038/srep37021

Durmuş Tekir, S., Çakır, T., Ardiç, E., Sayılırbaş, A.S., Konuk, G., Konuk, M., et al. (2013). PHISTO: pathogen-host interaction search tool. Bioinformatics. 29, 1357-1358.

Elbirt, D., Mahlab-Guri, K., Bezalel-Rosenberg, S., Gill, H., Attali, M., Asher, I. (2015). HIV-associated neurocognitive disorders (HAND). Isr. Med. Assoc. J. 17, 54-59.

Emig-Agius, D., Olivieri, K., Pache, L., Shih, H.L., Pustovalova, O., Bessarabova, M., et al. (2014). An integrated map of HIV-human protein complexes that facilitate viral infection. PLoS One. 9:e96687. doi: 10.1371/journal.pone.0096687

Engelman, A.N., and Singh, P.K. (2018). Cellular and molecular mechanisms of HIV-1 integration targeting. Cell Mol. Life Sci. 75, 2491-2507.

Farhadian, S.F., Mehta, S.S., Zografou, C., Robertson, K., Price, R.W., Pappalardo, J., et al. (2018). Single-cell RNA sequencing reveals microglia-like cells in cerebrospinal fluid during virologically suppressed HIV. JCI Insight. 3:121718. doi: 10.1172/jci.insight.121718

Fenwick, C., Joo, V., Jacquier, P., Noto, A., Banga, R., Perreau, M., et al. (2019). T-cell exhaustion in HIV infection. Immunol. Rev. 292, 149-163.

Fischer, W., Perkins, S., Theiler, J., Bhattacharya, T., Yusim, K., Funkhouser, R., et al. (2007). Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. Nat. Med. 13, 100-106.

Fiscon, G., Conte, F., Farina, L., Paci, P. (2018). Network-Based Approaches to Explore Complex Biological Systems towards Network Medicine. Genes (Basel). 9:E437. doi: 10.3390/genes9090437

Fourati, S., Ribeiro, S.P., Blasco Tavares Pereira Lopes, F., Talla, A., Lefebvre, F., Cameron, M., et al. (2019). Integrated systems approach defines the antiviral pathways conferring protection by the RV144 HIV vaccine. Nat. Commun. 10:863. doi: 10.1038/s41467-019-08854-2

Fruci, D., Rota, R., Gallo, A. (2017). The Role of HCMV and HIV-1 MicroRNAs: Processing, and Mechanisms of Action during Viral Infection. Front. Microbiol. 8:689. doi: 10.3389/fmicb.2017.00689

Fung, H.B., and Guo, Y. (2004). Enfuvirtide: a fusion inhibitor for the treatment of HIV infection. Clin. Ther. 26, 352-378.

Gallo, R.C. (2015). Developing a Successful HIV Vaccine. J. Infect. Dis. 212, S40-S41.

Gallo, R.C. (2016). Shock and kill with caution. Science. 354, 177-178.

Gao, L., Wang, Y., Li, Y., Dong, Y., Yang, A., Zhang, J., et al. (2018). Genome-wide expression profiling analysis to identify key genes in the anti-HIV mechanism of CD4+ and CD8+ T cells. J. Med. Virol. 90, 1199-1209.

Ghodke, Y., Anderson, P.L., Sangkuhl, K., Lamba, J., Altman, R.B., Klein, T.E. (2012). PharmGKB summary: zidovudine pathway. Pharmacogenet. Genomics. 22, 891-894.

Gillen, J., and Nita-Lazar, A. (2019). Experimental Analysis of Viral-Host Interactions. Front. Physiol. 10:425. doi: 10.3389/fphys.2019.00425

Golumbeanu, M., Desfarges, S., Hernandez, C., Quadroni, M., Rato, S., Mohammadi, P., et al. (2019). Proteo-Transcriptomic Dynamics of Cellular Response to HIV-1 Infection. Sci. Rep. 9:213. doi: 10.1038/s41598-018-36135-3

Gonzalo-Gil, E., Ikediobi, U., Sutton, R.E. (2017). Mechanisms of Virologic Control and Clinical Characteristics of HIV+ Elite/Viremic Controllers. Yale J. Biol. Med. 90, 245-259.

Goodacre, N., Devkota, P., Bae, E., Wuchty, S., Uetz, P. (2020). Protein-protein interactions of human viruses. Semin. Cell Dev. Biol. 99, 31-39.

Gray, G.E., Laher, F., Lazarus, E., Ensoli, B., Corey, L. (2016). Approaches to preventative and therapeutic HIV vaccines. Curr. Opin. Virol. 17, 104-109.

Greenwood, E.J., Matheson, N.J., Wals, K., van den Boomen, D.J., Antrobus, R., Williamson, J.C., et al. (2016). Temporal proteomic analysis of HIV infection reveals remodelling of the host phosphoproteome by lentiviral Vif variants. Elife. 5:e18296. doi: 10.7554/eLife.18296

Guirimand, T., Delmotte, S., Navratil, V. (2015). VirHostNet 2.0: surfing on the web of virus/host molecular interactions data. Nucleic Acids Res. 43, D583- D587.

Haynes, B.F., Burton, D.R., Mascola, J.R. (2019). Multiple roles for HIV broadly neutralizing antibodies. Sci. Transl. Med. 11:eaaz2686. doi: 10.1126/scitranslmed.aaz2686

Haynes, B.F., Shaw, G.M., Korber, B., Kelsoe, G., Sodroski, J., Hahn, B.H., et al. (2016). HIV-Host Interactions: Implications for Vaccine Design. Cell Host Microbe. 19, 292-303.

Herrera-Carrillo, E., Gao, Z., Berkhout, B. (2019). CRISPR therapy towards an HIV cure. Brief. Funct. Genomics. elz021. doi: 10.1093/bfgp/elz021

Hsu, D.C., and O'Connell, R.J. (2017). Progress in HIV vaccine development. Hum. Vaccin. Immunother. 13, 1018-1030.

Huang, T., Xu, Z., Chen, L., Cai, Y.D., Kong, X. (2011). Computational analysis of HIV-1 resistance based on gene expression profiles and the virus-host interaction network. PLoS One. 6:e17291. doi: 10.1371/journal.pone.0017291

Hulot, S.L., Korber, B., Giorgi, E.E., Vandergrift, N., Saunders, K.O., Balachandran, H., et al. (2015). Comparison of Immunogenicity in Rhesus Macaques of Transmitted-Founder, HIV-1 Group M Consensus, and Trivalent Mosaic Envelope Vaccines Formulated as a DNA Prime, NYVAC, and Envelope Protein Boost. J. Virol. 89, 6462-6480.

Iwai, S., Huang, D., Fong, S., Jarlsberg, L.G., Worodria, W., Yoo, S., et al. (2014). The lung microbiome of Ugandan HIV-infected pneumonia patients is compositionally and functionally distinct from that of San Franciscan patients. PLoS One. 9:e95726. doi: 10.1371/journal.pone.0095726

Iyer, S.S., Bibollet-Ruche, F., Sherrill-Mix, S., Learn, G.H., Plenderleith, L., Smith, A.G., et al. (2017). Resistance to type 1 interferons is a major determinant of HIV-1 transmission fitness. Proc. Natl. Acad. Sci. U S A. 114, E590-E599.

Johnson, J.S., Lucas, S.Y., Amon, L.M., Skelton, S., Nazitto, R., Carbonetti, S., et al. (2018). Reshaping of the Dendritic Cell Chromatin Landscape and Interferon Pathways during HIV Infection. Cell Host Microbe. 23, 366-381.

Kandathil, A.J., Sugawara, S., Balagopal, A. (2016). Are T cells the only HIV-1 reservoir? Retrovirology. 13:86. doi: 10.1186/s12977-016-0323-4

Khan, A.M., Hu, Y., Miotto, O., Thevasagayam, N.M., Sukumaran, R., Abd Raman, H.S., et al. (2017). Analysis of viral diversity for vaccine target discovery. BMC Med. Genomics. 10:78. doi: 10.1186/s12920-017-0301-2

Khatri, P., Sirota, M., Butte, A.J. (2012). Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput. Biol. 8:e1002375. doi: 10.1371/journal.pcbi.1002375

Kinch, M.S., and Patridge, E. (2014). An analysis of FDA-approved drugs for infectious disease: HIV/AIDS drugs. Drug Discov. Today. 19, 1510-1513.

Kirchhoff, F. (2013). "HIV Life Cycle: Overview," in the Encyclopedia of AIDS, ed. TJ. Hope, M. Stevenson, D. Richman (New York: Springer New York), 1–9.

König, R., Zhou, Y., Elleder, D., Diamond, T.L., Bonamy, G.M., Irelan, J.T., et al. (2008). Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Cell. 135, 49-60.

Korber, B.T., Letvin, N.L., Haynes, B.F. (2009). T-cell vaccine strategies for human immunodeficiency virus, the virus with a thousand faces. J. Virol. 83, 8300-8314.

Kotlyar, M., Pastrello, C., Sheahan, N., Jurisica, I. (2016). Integrated interactions database: tissue-specific view of the human and model organism interactomes. Nucleic Acids Res. 44, D536-D541.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 44, W90-W97.

Kulkarni, A., Anderson, A.G., Merullo, D.P., Konopka, G. (2019). Beyond bulk: a review of single cell transcriptomics methodologies and applications. Curr. Opin. Biotechnol. 58, 129-136.

Kumar, S., Maurya, V.K., Dandu, H.R., Bhatt, M.L., Saxena, S.K. (2018). Global Perspective of Novel Therapeutic Strategies for the Management of NeuroAIDS. Biomol. Concepts. 9, 33-42.

Lacerda, M., Moore, P.L., Ngandu, N.K., Seaman, M., Gray, E.S., Murrell, B., et al. (2013). Identification of broadly neutralizing antibody epitopes in the HIV-1 envelope glycoprotein using evolutionary models. Virol. J. 10:347. doi: 10.1186/1743-422X-10-347

Lan, W., Wang, J., Li, M., Peng, W., Wu, F. (2015). Computational approaches for prioritizing candidate disease genes based on PPI networks. Tsinghua Science and Technology. 20, 500–512.

Langer, S., Hammer, C., Hopfensperger, K., Klein, L., Hotter, D., De Jesus, P.D., et al. (2019). HIV-1 Vpu is a potent transcriptional suppressor of NF-κB-elicited antiviral immune responses. Elife. 8:e41930. doi: 10.7554/eLife.41930

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 9:559. doi: 10.1186/1471-2105-9-559

Lapek, J.D. Jr., Lewinski, M.K., Wozniak, J.M., Guatelli, J., Gonzalez, D.J. (2017). Quantitative Temporal Viromics of an Inducible HIV-1 Model Yields Insight to Global Host Targets and Phospho-Dynamics Associated with Protein Vpr. Mol. Cell. Proteomics. 16, 1447-1461.

Lazar, D.C., Morris, K.V., Saayman, S.M. (2016). The emerging role of long non-coding RNAs in HIV infection. Virus Res. 212, 114-126.

Le Clerc, S., Limou, S., Zagury, J.F. (2019). Large-Scale "OMICS" Studies to Explore the Physiopatholgy of HIV-1 Infection. Front. Genet. 10:799. doi: 10.3389/fgene.2019.00799

Levine, A.J., Horvath, S., Miller, E.N., Singer, E.J., Shapshak, P., Baldwin, G.C., et al. (2013). Transcriptome analysis of HIV-infected peripheral blood monocytes: gene transcripts and networks associated with neurocognitive functioning. J. Neuroimmunol. 265, 96-105.

Levine, A.J., Miller, J.A., Shapshak, P., Gelman, B., Singer, E.J., Hinkin, C.H., et al. (2013). Systems analysis of human brain gene expression: mechanisms for HIV-associated neurocognitive impairment and common pathways with Alzheimer's disease. BMC Med. Genomics. 6:4. doi: 10.1186/1755-8794-6-4

Levy, J.A. (2015). Dispelling myths and focusing on notable concepts in HIV pathogenesis. Trends Mol. Med. 21, 341-353.

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., Pfister, H. (2014). UpSet: Visualization of Intersecting Sets. IEEE Trans. Vis. Comput. Graph. 20, 1983-1992.

Li, B.Q., Niu, B., Chen, L., Wei, Z.J., Huang, T., Jiang, M., et al. (2013). Identifying chemicals with potential therapy of HIV based on protein-protein and protein-chemical interaction network. PLoS One. 8:e65207. doi: 10.1371/journal.pone.0065207

Li, C., Ramarathinam, S.H., Revote, J., Khoury, G., Song, J., Purcell, A.W. (2017). HIVed, a knowledgebase for differentially expressed human genes and proteins during HIV infection, replication and latency. Sci. Rep. 7:45509. doi: 10.1038/srep45509

Li, C.W., and Chen, B.S. (2018). Investigating HIV-Human Interaction Networks to Unravel Pathogenic Mechanism for Drug Discovery: A Systems Biology Approach. Curr. HIV Res. 16, 77-95.

Li, X., Li, W., Zeng, M., Zheng, R., Li, M. (2019). Network-based methods for predicting essential genes or proteins: a survey. Brief. Bioinform. bbz017. doi: 10.1093/bib/bbz017

Li, Y., Wang, C., Miao, Z., Bi, X., Wu, D., Jin, N., et al. (2015). ViRBase: a resource for virus-host ncRNA-associated interactions. Nucleic Acids Res. 43, D578- D582.

Liu, D., Zhao, L., Wang, Z., Zhou, X., Fan, X., Li, Y., et al. (2019). EWASdb: epigenome-wide association study database. Nucleic Acids Res. 47, D989-D993.

Luo, Y., Jacobs, E.Y., Greco, T.M., Mohammed, K.D., Tong, T., Keegan, S., et al. (2016). HIV-host interactome revealed directly from infected cells. Nat. Microbiol. 1:16068. doi: 10.1038/nmicrobiol.2016.68

Ma, C., Dong, X., Li, R., Liu, L. (2013). A computational study identifies HIV progression-related genes using mRMR and shortest path tracing. PLoS One. 8:e78057. doi: 10.1371/journal.pone.0078057

Ma, C., Zhou, Y., Huang, S.H. (2011). Inequalities and duality in gene coexpression networks of HIV-1 infection revealed by the combination of the double-connectivity approach and the Gini's method. J. Biomed. Biotechnol. 2011:926407. doi: 10.1155/2011/926407

MacPherson, J.I., Dickerson, J.E., Pinney, J.W., Robertson, D.L. (2010). Patterns of HIV-1 protein interaction identify perturbed host-cellular subsystems. PLoS Comput. Biol. 6:e1000863. doi: 10.1371/journal.pcbi.1000863

Marban, C., Forouzanfar, F., Ait-Ammar, A., Fahmi, F., El Mekdad, H., Daouad, F., et al. (2016). Targeting the Brain Reservoirs: Toward an HIV Cure. Front. Immunol. 7:397. doi: 10.3389/fimmu.2016.00397

Marban, C., Su, T., Ferrari, R., Li, B., Vatakis, D., Pellegrini, M., et al. (2011). Genome-wide binding map of the HIV-1 Tat protein to the human genome. PLoS One. 6:e26894. doi: 10.1371/journal.pone.0026894

Marini, B., Kertesz-Farkas, A., Ali, H., Lucic, B., Lisek, K., Manganaro, L., et al. (2015). Nuclear architecture dictates HIV-1 integration site selection. Nature. 521, 227-231.

Maulik, U., Bhattacharyya, M., Mukhopadhyay, A., Bandyopadhyay, S., et al. (2011). Identifying the immunodeficiency gateway proteins in humans and their involvement in microRNA regulation. Mol. Biosyst. 7, 1842-1851.

Maulik, U., Mukhopadhyay, A., Bhattacharyya, M., Kaderali, L., Brors, B., Bandyopadhyay, S., et al. (2013). Mining quasi-bicliques from HIV-1-human protein interaction network: a multiobjective biclustering approach. IEEE/ACM Trans. Comput. Biol. Bioinform. 10, 423-435.

McDonagh, E.M., Lau, J.L., Alvarellos, M.L., Altman, R.B., Klein, T.E. (2015). PharmGKB summary: Efavirenz pathway, pharmacokinetics. Pharmacogenet. Genomics. 25, 363-376.

Miryala, S.K., Anbarasu, A., Ramaiah, S. (2018). Discerning molecular interactions: A comprehensive review on biomolecular interaction databases and network analysis tools. Gene. 642, 84-94.

Moreno, S., Perno, C.F., Mallon, P.W., Behrens, G., Corbeau, P., Routy, J.P., et al. (2019). Two-drug vs. three-drug combinations for HIV-1: Do we have enough data to make the switch? HIV Med. 20, 2-12.

Mosaddek Hossain, S.M., Ray, S., Mukhopadhyay, A. (2017). Preservation affinity in consensus modules among stages of HIV-1 progression. BMC Bioinformatics. 18:181. doi: 10.1186/s12859-017-1590-3

Murali, T.M., Dyer, M.D., Badger, D., Tyler, B.M., Katze, M.G. (2011). Network-based prediction and analysis of HIV dependency factors. PLoS Comput. Biol. 7:e1002164. doi: 10.1371/journal.pcbi.1002164

Mylvaganam, G., Yanez, A.G., Maus, M., Walker, B.D. (2019). Toward T Cell-Mediated Control or Elimination of HIV Reservoirs: Lessons From Cancer Immunology. Front. Immunol. 10:2109. doi: 10.3389/fimmu.2019.02109

Neveu, G., Cassonnet, P., Vidalain, P.O., Rolloy, C., Mendoza, J., Jones, L., et al. (2012). Comparative analysis of virus-host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase. Methods. 58, 349-359.

Nguyen, S., Deleage, C., Darko, S., Ransier, A., Truong, D.P., Agarwal, D., et al. (2019). Elite control of HIV is associated with distinct functional and transcriptional signatures in lymphoid tissue CD8+ T cells. Sci. Transl. Med. 11:eaax4077. doi: 10.1126/scitranslmed.aax4077

Olivier, I.S., Cacabelos, R., Naidoo, V. (2018). Risk Factors and Pathogenesis of HIV-Associated Neurocognitive Disorder: The Role of Host Genetics. Int. J. Mol. Sci. 19:E3594. doi: 10.3390/ijms19113594

Oyeyemi, O.J., Davies, O., Robertson, D.L., Schwartz, J.M. (2015). A logical model of HIV-1 interactions with the T-cell activation signalling pathway. Bioinformatics. 31, 1075-1083.

Park, R.J., Wang, T., Koundakjian, D., Hultquist, J.F., Lamothe-Molina, P., Monel, B., et al. (2017). A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors. Nat. Genet. 49, 193-203.

Peterson, C.W., and Kiem, H. (2018). Cell and Gene Therapy for HIV Cure. Curr. Top. Microbiol. Immunol. 417, 211-248.

Pitman, M.C., Lau, J.S.Y., McMahon, J.H., Lewin, S.R. (2018). Barriers and strategies to achieve a cure for HIV. Lancet HIV. 5, e317-e328.

Poropatich, K., and Sullivan, D.J. Jr. (2011). Human immunodeficiency virus type 1 long-term non-progressors: the viral, genetic and immunological basis for disease non-progression. J. Gen. Virol. 92, 247-268.

Prévost, J., Edgar, C.R., Richard, J., Trothen, S.M., Jacob, R.A., Mumby, M.J., et al. (2020). HIV-1 Vpu downregulates Tim-3 from the surface of infected CD4+ T cells. J. Virol. 94:e01999-19. doi: 10.1128/JVI.01999-19

Proust, A., Barat, C., Leboeuf, M., Drouin, J., Tremblay, M.J. (2017). Contrasting effect of the latency-reversing agents bryostatin-1 and JQ1 on astrocyte-mediated neuroinflammation and brain neutrophil invasion. J. Neuroinflammation. 14:242. doi: 10.1186/s12974-017-1019-y

Puhl, A.C., Garzino Demo, A., Makarov, V.A., Ekins, S. (2019). New targets for HIV drug discovery. Drug Discov. Today. 24, 1139-1147.

Quach, A., Horvath, S., Nemanim, N., Vatakis, D., Witt, M.D., Miller, E.N., et al. (2018). No reliable gene expression biomarkers of current or impending neurocognitive impairment in peripheral blood monocytes of persons living with HIV. J. Neurovirol. 24, 350-361.

Qureshi, A., Thakur, N., Monga, I., Thakur, A., Kumar, M. (2014). VIRmiRNA: a comprehensive resource for experimentally validated viral miRNAs and their targets. Database (Oxford). 2014:bau103. doi: 10.1093/database/bau103

Ray, S., and Bandyopadhyay, S. (2016). Discovering Condition Specific Topological Pattern Changes in Coexpression Network: An Application to HIV-1 Progression. IEEE/ACM Trans. Comput. Biol. Bioinform. 13, 1086-1099.

Ray, S., and Maulik, U. (2017). Identifying differentially coexpressed module during HIV disease progression: A multiobjective approach. Sci. Rep. 7:86. doi: 10.1038/s41598-017-00090-2

Riemenschneider, M., and Heider, D. (2016). Current Approaches in Computational Drug Resistance Prediction in HIV. Curr. HIV Res. 14, 307-315.

Rios, A. (2018). Fundamental challenges to the development of a preventive HIV vaccine. Curr. Opin. Virol. 29, 26-32.

Rogozin, I.B., Malyarchuk, B.A., Pavlov, Y.I., Milanesi, L. (2005). From context-dependence of mutations to molecular mechanisms of mutagenesis. Pac. Symp. Biocomput. 10, 409-420.

Sadowski, I., and Hashemi, F.B. (2019). Strategies to eradicate HIV from infected patients: elimination of latent provirus reservoirs. Cell. Mol. Life Sci. 76, 3583-3600.

Saha, J.M., Liu, H., Hu, P.W., Nikolai, B.C., Wu, H., Miao, H., et al. (2018). Proteomic Profiling of a Primary CD4+ T Cell Model of HIV-1 Latency Identifies Proteins Whose Differential Expression Correlates with Reactivation of Latent HIV-1. AIDS Res. Hum. Retroviruses. 34, 103-110.

Seddiki, N., and Lévy, Y. (2018). Therapeutic HIV-1 vaccine: time for immunomodulation and combinatorial strategies. Curr. Opin. HIV AIDS. 13, 119-127.

Shao, T., Zhao, Z., Wu, A., Bai, J., Li, Y., Chen, H., et al. (2015). Functional dissection of virus-human crosstalk mediated by miRNAs based on the VmiReg database. Mol. Biosyst. 11, 1319-1328.

Shityakov, S., Dandekar, T., Förster, C. (2015). Gene expression profiles and protein-protein interaction network analysis in AIDS patients with HIV-associated encephalitis and dementia. HIV AIDS (Auckl). 7, 265-276.

Shukla, E., and Chauhan, R. (2019). Host-HIV-1 Interactome: A Quest for Novel Therapeutic Intervention. Cells. 8:E1155. doi: 10.3390/cells8101155

Sillman, B., Woldstad, C., Mcmillan, J., Gendelman, H.E. (2018). "Neuropathogenesis of human immunodeficiency virus infection," in The Handbook of Clinical Neurology, ed. B.J. Brew (Amsterdam: Elsevier), 21–40.

Sok, D., and Burton, D.R. (2018). Recent progress in broadly neutralizing antibodies to HIV. Nat. Immunol. 19, 1179-1188.

Tan, S., Li, W., Li, Z., Li, Y., Luo, J., Yu, L., et al. (2019). A Novel CXCR4 Targeting Protein SDF-1/54 as an HIV-1 Entry Inhibitor. Viruses. 11:E874. doi: 10.3390/v11090874

Tan, X., Hu, L., Luquette, L.J. 3rd., Gao, G., Liu, Y., Qu, H., et al. (2012). Systematic identification of synergistic drug pairs targeting HIV. Nat. Biotechnol. 30, 1125-1130.

Tarasova, O., Biziukova, N., Kireev, D., Lagunin, A., Ivanov, S., Filimonov, D., et al. (2020). A Computational Approach for the Prediction of Treatment History and the Effectiveness or Failure of Antiretroviral Therapy. Int. J. Mol. Sci. 21:E748. doi: 10.3390/ijms21030748

Tarasova, O., and Poroikov, V. (2018). HIV Resistance Prediction to Reverse Transcriptase Inhibitors: Focus on Open Data. Molecules. 23:E956. doi: 10.3390/molecules23040956

Tough, R.H., and McLaren, P.J. (2019). Interaction of the Host and Viral Genome and Their Influence on HIV Disease. Front. Genet. 9:720. doi: 10.3389/fgene.2018.00720

Trovato, M., D'Apice, L., Prisco, A., De Berardinis, P. (2018). HIV Vaccination: A Roadmap among Advancements and Concerns. Int. J. Mol. Sci. 19:E1241. doi: 10.3390/ijms19041241

van Dam, S., Võsa, U., van der Graaf, A., Franke, L., de Magalhães, J.P. (2018). Gene co-expression analysis for functional classification and gene-disease predictions. Brief. Bioinform. 19, 575-592.

van den Kerkhof, T.L., de Taeye, S.W., Boeser-Nunnink, B.D., Burton, D.R., Kootstra, N.A., Schuitemaker, H., et al. (2016). HIV-1 escapes from N332-directed antibody neutralization in an elite neutralizer by envelope glycoprotein elongation and introduction of unusual disulfide bonds. Retrovirology. 13:48. doi: 10.1186/s12977-016-0279-4

van Dijk, D., Ertaylan, G., Boucher, C.A., Sloot, P.M. (2010). Identifying potential survival strategies of HIV-1 through virus-host protein interaction networks. BMC Syst. Biol. 4:96. doi: 10.1186/1752-0509-4-96

Van Lint, C., Bouchat, S., Marcello, A. (2013). HIV-1 transcription and latency: an update. Retrovirology. 10:67. doi: 10.1186/1742-4690-10-67

van Stigt Thans, T., Akko, J.I., Niehrs, A., Garcia-Beltran, W.F., Richert, L., Stürzel, C.M., et al. (2019). Primary HIV-1 Strains Use Nef To Downmodulate HLA-E Surface Expression. J. Virol. 93:e00719-19. doi: 10.1128/JVI.00719-19

Vansant, G., Bruggemans, A., Janssens, J., Debyser, Z. (2020). Block-And-Lock Strategies to Cure HIV Infection. Viruses. 12:E84. doi: 10.3390/v12010084

Vujkovic-Cvijin, I., Dunham, R.M., Iwai, S., Maher, M.C., Albright, R.G., Broadhurst, M.J., et al. (2013). Dysbiosis of the gut microbiota is associated with HIV disease progression and tryptophan catabolism. Sci. Transl. Med. 5:193ra91. doi: 10.1126/scitranslmed.3006438

Wang, C.X., and Cannon, P.M. (2016). The clinical applications of genome editing in HIV. Blood. 127, 2546-2552.

Wang, Q., and Zhang, L. (2020). Broadly neutralizing antibodies and vaccine design against HIV-1 infection. Front. Med. 14, 30-42.

Wang, T., Tian, C., Zhang, W., Luo, K., Sarkis, P.T., Yu, L., et al. (2007). 7SL RNA mediates virion packaging of the antiviral cytidine deaminase APOBEC3G. J. Virol. 81, 13112-13124.

Weber, I.T., and Harrison, R.W. (2017). Decoding HIV resistance: from genotype to therapy. Future Med. Chem. 9, 1529-1538.

White, C.H., Moesker, B., Beliakova-Bethell, N., Martins, L.J., Spina, C.A., Margolis, D.M., et al. (2016). Transcriptomic Analysis Implicates the p53 Signaling Pathway in the Establishment of HIV-1 Latency in Central Memory CD4 T Cells in an In Vitro Model. PLoS Pathog. 12:e1006026. doi: 10.1371/journal.ppat.1006026

World Health Organization. (2020). HIV/AIDS. https://www.who.int/news-room/fact-sheets/detail/hiv-aids [Accessed March 20, 2020].

Wu, Z., Liao, Q., Liu, B. (2019). A comprehensive review and evaluation of computational methods for identifying protein complexes from protein-protein interaction networks. Brief. Bioinform. bbz085. doi: 10.1093/bib/bbz085

Xie, D., Han, L., Luo, Y., Liu, Y., He, S., Bai, H., et al. (2015). Exploring the associations of host genes for viral infection revealed by genome-wide RNAi and virus-host protein interactions. Mol. Biosyst. 11, 2511-2519.

Xu, C., Ye, B., Han, Z., Huang, M., Zhu, Y. (2014). Comparison of transcriptional profiles between CD4+ and CD8+ T cells in HIV type 1-infected patients. AIDS Res. Hum. Retroviruses. 30, 134-141.

Xu, W.W., Han, M.J., Chen, D., Chen, L., Guo, Y., Willden, A., et al. (2013). Genome-wide search for the genes accountable for the induced resistance to HIV-1 infection in activated CD4+ T cells: apparent transcriptional signatures, co-expression networks and possible cellular processes. BMC Med. Genomics. 6:15. doi: 10.1186/1755-8794-6-15

Yamashita, M., and Engelman, A.N. (2017). Capsid-Dependent Host Factors in HIV-1 Infection. Trends Microbiol. 25, 741-755.

Yang, G., Holl, T.M., Liu, Y., Li, Y., Lu, X., Nicely, N.I., et al. (2013). Identification of autoantigens recognized by the 2F5 and 4E10 broadly neutralizing HIV-1 antibodies. J. Exp. Med. 210, 241-256.

Yang, S., Fu, C., Lian, X., Dong, X., Zhang, Z. (2019). Understanding Human-Virus Protein-Protein Interactions Using a Human Protein Complex-Based Analysis Framework. mSystems. 4:e00303-18. doi: 10.1128/mSystems.00303-18

Yang, W., Jackson, B., Zhang, H. (2016). Identification of glycoproteins associated with HIV latently infected cells using quantitative glycoproteomics. Proteomics. 16, 1872-1880.

Yeung, M.L., Houzet, L., Yedavalli, V.S., Jeang, K.T. (2009). A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication. J. Biol. Chem. 284, 19463-19473.

Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P., et al. (2003). Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. Science. 302, 1056-1060.

Zak, D.E., Andersen-Nissen, E., Peterson, E.R., Sato, A., Hamilton, M.K., Borgerding, J., et al. (2012). Merck Ad5/HIV induces broad innate immune activation that predicts CD8[+] T-cell responses but is attenuated by preexisting Ad5 immunity. Proc. Natl. Acad. Sci. U S A. 109, E3503- E3512.

Zanini, F., Brodin, J., Thebo, L., Lanz, C., Bratt, G., Albert, J., et al. (2015). Population genomics of intrapatient HIV-1 evolution. Elife. 4:e11282. doi: 10.7554/eLife.11282

Zazzi, M., Cozzi-Lepri, A., Prosperi, M.C. (2016). Computer-Aided Optimization of Combined Anti-Retroviral Therapy for HIV: New Drugs, New Drug Targets and Drug Resistance. Curr. HIV Res. 14, 101-109.

Zhan, P., Pannecouque, C., De Clercq, E., Liu, X. (2016). Anti-HIV Drug Discovery and Development: Current Innovations and Future Trends. J. Med. Chem. 59, 2849-2878.

Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. Stat. Appl. Genet. Mol. Biol. 4:Article17. doi: 10.2202/1544-6115.1128

Zhang, D., Li, W., Jiang, S. (2015). Peptide fusion inhibitors targeting the HIV-1 gp41: a patent review (2009 - 2014). Expert. Opin. Ther. Pat. 25, 159-173.

Zhang, Q., Chen, C.Y., Yedavalli, V.S., Jeang, K.T. (2013). NEAT1 long noncoding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. MBio. 4:e00596-12. doi: 10.1128/mBio.00596-12

Zhang, X., Hu, Y., Aouizerat, B.E., Peng, G., Marconi, V.C., Corley, M.J., et al. (2018). Machine learning selected smoking-associated DNA methylation signatures that predict HIV prognosis and mortality. Clin. Epigenetics. 10:155. doi: 10.1186/s13148-018-0591-z

Zhang, X., Justice, A.C., Hu, Y., Wang, Z., Zhao, H., Wang, G., et al. (2016). Epigenome-wide differential DNA methylation between HIV-infected and uninfected individuals. Epigenetics. 11, 750-760.

Zhen, A., Du, J., Zhou, X., Xiong, Y., Yu, X.F. (2012). Reduced APOBEC3H variant anti-viral activities are associated with altered RNA binding activities. PLoS One. 7:e38771. doi: 10.1371/journal.pone.0038771

Zheng, D., Wan, J., Cho, Y.G., Wang, L., Chiou, C.J., Pai, S., et al. (2011). Comparison of humoral immune responses to Epstein-Barr virus and Kaposi's sarcoma-associated herpesvirus using a viral proteome microarray. J. Infect. Dis. 204, 1683-1691.

Zhou, H., Xu, M., Huang, Q., Gates, A.T., Zhang, X.D., Castle, J.C., et al. (2008). Genome-scale RNAi screen for host factors required for HIV replication. Cell Host Microbe. 4, 495-504.

Zuo, X., Huo, Z., Kang, D., Wu, G., Zhou, Z., Liu, X., et al. (2018). Current insights into anti-HIV drug discovery and development: a review of recent patent literature (2014-2017). Expert Opin. Ther. Pat. 28, 299-316.

Table 1. Public databases containing data on interactions between HIV and human proteins.

Database	Number of HIV-1-human interactions	Number of human proteins interacting with HIV-1 [1]	Number of HIV-2-human interactions	Number of human proteins interacting with HIV-2 [1]	Reference
NCBI database [2]	1037	842	-	-	(Ako-Adjei et al., 2015)
HPIDB	1668	1390	25	24	(Ammari et al., 2016)
PHISTO	1978	1460	27	26	(Durmuş Tekir et al., 2013)
VirHostNet	1077	985	13	13	(Guirimand et al., 2015)
Viruses.STRING	929	827	3	3	(Cook et al., 2018)
VirusMentha	1206	1052	14	13	(Calderone et al., 2015)
Total [3]	2910	2051	34	30	

[1] The number of unique SwissProt identifiers that corresponds to human proteins interacting with HIV proteins.

[2] The database contains both direct and indirect functional interactions. We considered direct interactions based on the following keywords: "acetylated by", "binds", "cleaved by", "cleaves", "degraded by", "dephosphorylated by", "glycosylated by", "interacts", "is monomethylated", "is polyubiquitinylated by", "isomerized by", "methylated by", "modified by", "myristoylated by", "palmitoylated by", "phosphorylated by", "processed by", "sulfated by", "sumoylated by", "ubiquitinated by".

[3] Total numbers of unique interactions.

Table 2. Distribution of the numbers of PPIs between different HIV-1 groups and subtypes.

	Number of isolates	Range of PPI numbers for different isolates [1]	Number of unique PPIs [2]	Databases [3]
HIV-1 group M subtype A	3	162-350	353	b, d
HIV-1 group M subtype B	27	6-1875	1910	a, b, d, f
HIV-1 group M subtype C	1	5	5	b, d, f
HIV-1 group M subtype D	5	167-347	356	b, d, f
HIV-1 group M subtype F1	1	1	1	b, d
HIV-1 group M subtype G	1	1	1	b, d
HIV-1 group M subtype H	1	1	1	b, d
HIV-1 group M subtype U	1	139	139	d
HIV-1 group N	1	1	1	b, d
HIV-1 (group unknown)	-	995	995	a, b, c, d, e, f

[1] Number of "SwissProt identifier of HIV gene – SwissProt identifier of human protein" pairs.

[2] Number of unique "HIV gene symbol – SwissProt identifier of human protein" pairs. Information about different isolates was merged.

[3] The data originated from the databases: (a) VirHostNet, (b) HPIDB, (c) NCBI database, (d) PHISTO, (e) Viruses.STRING, (f) VirusMentha.

Table 3. Various types of interactions between HIV-1 and human RNAs presented in public databases.

Type of interaction	ViRBase	VmiReg	VIRmiRNA	Total

human miRNA-viral mRNA	43	7	21	49
human miRNA-human mRNA	50	-	3	51
viral miRNA-human mRNA	21	13	2	21
viral miRNA-viral mRNA	5	2	-	5
viral miRNA-host miRNA	2	-	-	2

Table 4. An overview of HIV-related transcriptomic experiments. The numbers of corresponding experiments calculated as numbers of GEO and ArrayExpress identifiers are given in brackets.

Experiment type	Cell types	Conditions
in vivo (140)	***Primary immune cells and tissues:*** peripheral blood mononuclear cells (29), CD4+ T cells (24), whole blood (19), CD8+ T cells (13), monocytes (12), B cells (4), NK cells (3), lymph nodes (2), plasma (2), and others ***Other primary cells and tissues:*** frontal cortex (6), jejunum, colon or rectal cells (4), lymphoma cells (3), adipose cells (2), retinal cells (2), skeletal muscle cells (2), and others	***HIV infection status:*** infected, uninfected (65)
		HIV progression types: resistant to HIV, long-term non-progressors, elite controllers, chronic progressors, and others (36)
		HIV infection stages: acute, chronic (latent), AIDS (4)
		Quantitative characteristics of HIV infection: CD4+ cells count and HIV RNA level (5)
		HIV associated diseases: HIV-associated neurocognitive disorders, HIV-associated tumors, and others (22)
		HIV co-infections: mycobacterium tuberculosis, neisseria gonorrhoeae, pneumococcal meningitis, hepatitis C virus, kaposi's sarcoma-associated herpesvirus, and others (17)
		Small molecule treatment including antiretroviral drugs: treated or not treated patients, before or after therapy, before and after treatment interruption, and other (23)
		Vaccine trials (6)

|		
		Microbiome changes: transcriptional changes in gut, tongue and lung microbiome induced by HIV (3)
		Cell subtypes: comparison of transcription profiles between immune cell subtypes (9)
in vitro (93)	***Primary cells:*** CD4+ T cells (15), monocyte-derived macrophages (15), peripheral blood mononuclear cells (8), monocyte-derived dendritic cells (6), macrophages (2), primary human neurons (2), and others ***Cell lines:*** Jurkat cells (12), SupT1 cells (7), HEK293T cells (4), CEM-SS cells (3), ACH-2 and A3.01 cells (2), SH-SY5Y cells (2), U1 cells (2), U-937 cells (2), WE17/10 cells (2), and others	***HIV infection status:*** infected, uninfected (51)
		HIV infection time-series: different times after infection *in vitro* (22)
		HIV latency: viral latency in various CD4+ cells and reactivation by different agents (9)
		Treatment of cells with small molecule substances including antiretroviral drugs (11)
		Treatment of HIV-infected cells with cytokines (4)
		Infection of cells with various HIV variants and vectors expressing wild type or mutated HIV proteins (30)
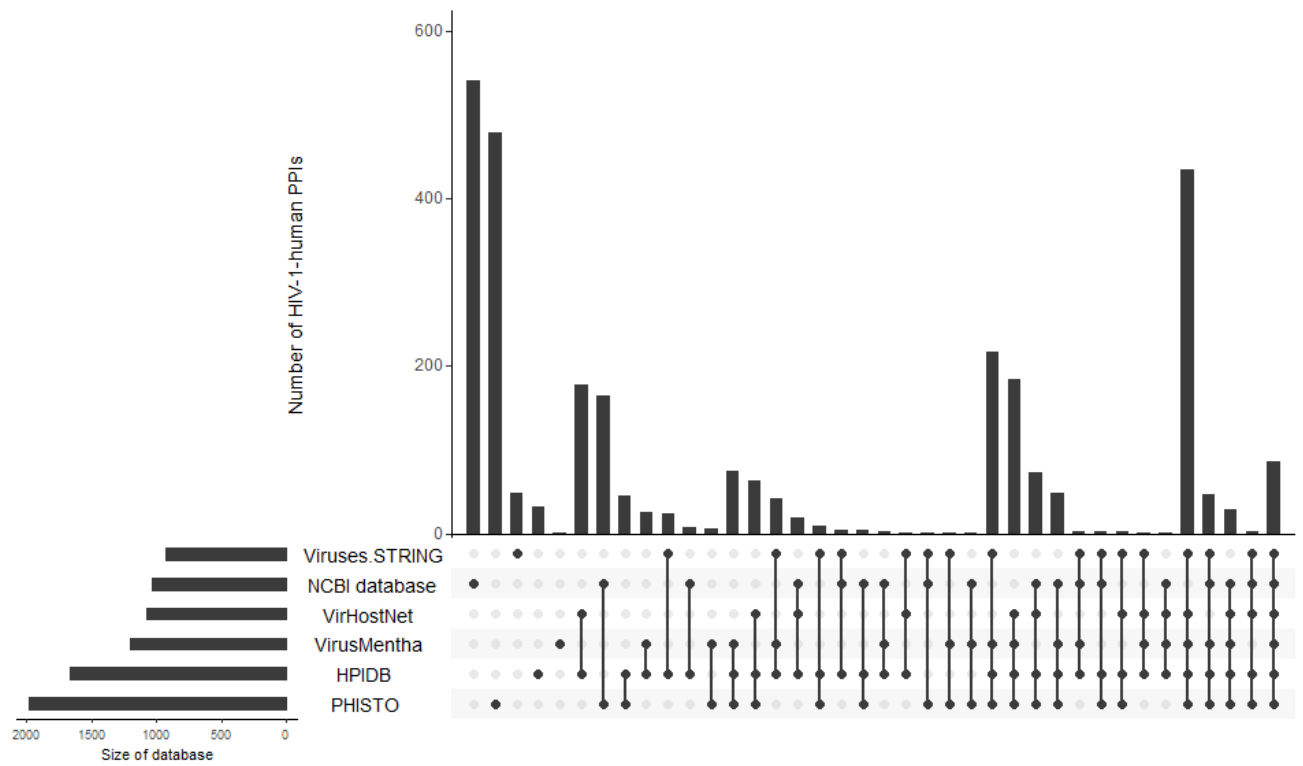
Figure 1. Intersections of HIV-1-human PPIs between six databases. Y-axis represents the numbers of HIV-1-human PPIs, which are either unique for a particular database or

shared by two, three, four, five and six databases. The connections between circles at the bottom part of figure represent intersections of PPIs between databases. The unconnected circles represent PPIs, which are unique for a particular database. The horizontal bars represent the total numbers of HIV-1-human PPIs in each database.
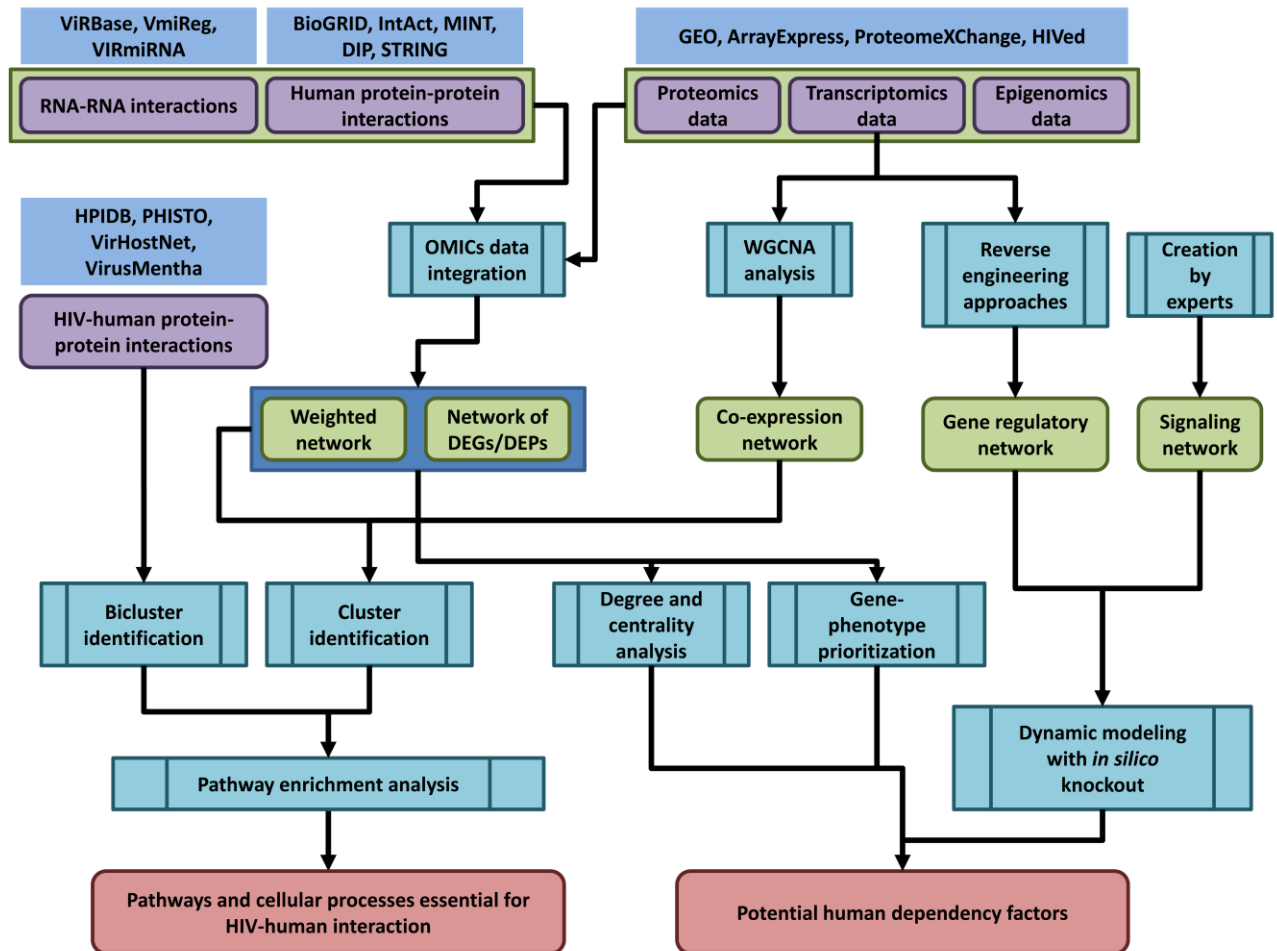
Figure 2. The general pipeline of network-based analysis of HIV-related OMICs data. Public databases provide access to OMICs data on human and HIV-human protein-protein interactions (HPIDB, PHISTO, VirHostNet, VirusMentha, and others, see Table 1, Table S1), interactions between human and viral coding/non-coding RNAs (ViRBase, VmiReg, VIRmiRNA databases, Table 3, Table S2), transcriptomics, proteomics and epigenomics data (GEO, ArrayExpress, ProteomeXChange, HIVed databases, Tables S3-S6) (nodes of purple color in the figure). The HIV-related OMICs data can be used to create context-specific protein-protein interaction networks, co-expression, gene regulatory and signaling networks (nodes of green color in the figure). The context-specific networks can be constructed by weighting protein-protein interactions using transcriptomics, proteomics or epigenomics data, or by taking into account only differentially expressed gene/proteins (DEGs/DEPs). Co-expression networks can be created using transcriptomics data and weighted gene correlation network analysis (WGCNA). Gene regulatory networks can be inferred from transcriptomics data using reverse engineering approaches, whereas signaling networks are usually manually created by experts based on a great deal of information regarding the protein interactions, post-translational modifications, and other data types. The created networks can be used for different types of analysis (nodes of blue color in the figure): (1) identification of dense communities in human protein-protein interaction and co-expression networks (clusters

or modules), or in HIV-human interaction networks (biclusters). The pathway enrichment analysis applied to clusters and biclusters allows identifying pathways and cellular processes, which are essential for HIV-human interaction; (2) degree and centrality analysis, gene phenotype prioritization analysis, as well as dynamic modeling with *in silico* gene knockout allows identifying proteins, which are the most essential for HIV-human interaction (host dependency factors), and can be considered as potential targets for new anti-HIV therapeutic approaches.