Published in final edited form as:

Future Microbiol. 2010 February ; 5(2): 303–311. doi:10.2217/fmb.09.121.

The use of RNAi-based screens to identify host proteins involved in viral replication

Alec J Hirsch

Vaccine & Gene Therapy Institute, Oregon, Health & Science University, 505 NW 185th Avenue, Beaverton, OR 97006, USA. Tel.: +1 503 418 2784, Fax: +1 503 418 2701, hirschal@ohsu.edu

Abstract

The recent development of RNAi-based techniques for protein knockdown in mammalian cells has allowed for unprecedented flexibility in the study of protein function. Currently, large siRNA libraries are available that allow the knockdown of all proteins known to be encoded by the human genome. These libraries have been used to identify the host proteins required for the replication of several clinically important viruses, including HIV, flaviviruses and influenza. This review summarizes the methods used in RNAi-based screening for host factors involved in virus replication, and discusses published examples of such screens.

Keywords

flavivirus; HIV; influenza; screen; siRNA; viral host factor; virus

Viruses, by definition, are obligate parasites. They require host cell proteins and pathways to carry out many of the phases of their lifecycles. Until recently, elucidation of the contribution of individual mammalian host cell proteins to viral replication faced certain limitations. Direct interactions between viral and host cell proteins could be determined by genetic or biochemical techniques, such as the yeast two-hybrid screen or co-immunoprecipitation. Further analysis of these interactions usually required construction of mutant viruses that lack this interaction. Alternatively, studies made use of small-molecule inhibitors of proteins of interest or knockout mouse lines in which the host gene of interest has been deleted. In 2001, RNAi-mediated knockdown of individual mRNAs in mammalian cells was first described, allowing the transient depletion of individual proteins from target cells [1]. Commercially available siRNA libraries targeting all known and predicted mRNA transcripts in humans, and other organisms, have recently made genome-wide screens practicable. Screening for proteins important in viral replication has been among the first uses of these libraries (also reviewed in [2,3]). Of particular interest is the fact that several screens have been performed in cells infected with the same, or related viruses, allowing analysis of the reproducibility and robustness of this technique.

© 2010 Future Medicine Ltd

For reprint orders, please contact: reprints@futuremedicine.com

Financial & competing interests disclosure

This work was supported by federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, as part of the Pacific Northwest Regional Center of Excellence for Biodefense and Emerging Infectious Disease Research U54 Award AI081680. The author has no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Use of siRNA in protein knockdown & screening

siRNAs are 21–23 nucleotide (nt), dsRNAs containing 2-nt 3' overhangs. One strand of the duplex is complementary to the target mRNA. This targeting (or guide) siRNA strand is incorporated into a multiprotein complex, termed the RNA-induced silencing complex (RISC). The RISC then uses the guide strand as a sequence-specific probe for targeting mRNAs for endonucleolytic cleavage by the RISC component protein argonaute 2. The depletion of a specific mRNA leads to a temporary reduction (knockdown) of the levels of the mRNA-encoded protein in the cell. A complicating technical issue in the use of siRNAs in large screens is the fact that different siRNA sequences, designed against the same mRNA, will result in variable degrees of target degradation. A rigorous analysis of multiple siRNAs designed against each of two target mRNAs (with each siRNA offset by 2 nt from the previous siRNA) demonstrated that siRNA silencing efficiency varied greatly between siRNA sequences, including adjacent siRNAs that varied by only 2 nt [4]. These data suggest that factors intrinsic to the siRNA sequence have a significant impact on the efficacy of a given siRNA. This study further correlated higher degrees of silencing with several features of an siRNA:

- Low G/C content (32–50%)
- Low internal stability at the 3' end of the sense strand (defined as at least 3 A/U bp at positions 15–19)
- Lack of internal repeats
- Specific preferences for certain bases at positions 3, 10 and 19

In addition, regions of secondary structure within the mRNA target also influence the ability of a given siRNA to effect target cleavage [5]. Despite the correlates of siRNA functionality described in this paper and elsewhere [6,7], the extent of protein reduction by an individual siRNA requires empirical determination. Furthermore, off-target effects, in which an siRNA may affect the level of unintentionally targeted mRNAs, often through partial sequence complementarity, are well documented [8–10]. Depletion of non-targeted mRNA may result in misinterpretation of the phenotype produced by an individual siRNA. Therefore, large siRNA libraries are typically designed to contain component siRNAs with minimal sequence complementarity to other mRNAs. These libraries are usually composed of multiple siRNAs targeting each specific mRNA, in order to increase the chance that at least one siRNA will produce adequate target degradation. In some cases, the siRNAs are pooled, requiring analysis of each individual from the pool in a secondary screen. Other factors may also affect the ability of a particular siRNA to substantially reduce target protein expression. Proteins with a long half-life may be less susceptible to knockdown, owing to the propensity of the targeted protein to remain in the cell at close to normal levels, even if the encoding mRNA has been significantly diminished. In order to address this problem, siRNA transfected cells are generally assayed 48–72 h post-transfection. Alternatively, multiple siRNA transfections over several days may be used in order to suppress the level of the mRNA for a longer period of time. Since high-throughput screens containing large numbers of siRNAs prevent optimization of each particular siRNA, certain genes may be inefficiently targeted, potentially resulting in low protein knockdown and false negatives.

Efficiency of transfection is another factor that must be considered when performing or interpreting the results of large-scale siRNA-based screenings. Although transfection of siRNAs into a given cell type is more efficient than transfection of plasmid DNA [11], some cell types remain refractory to siRNA transfection, effectively excluding them from use in such screens. In most of the screens for viral replication factors, discussed in the following sections, the cell types used were derived from easily transfected cell lines, such as HeLa or

HEK293. If hard-to-transfect cell types must be used in a particular assay, introduction of retroviral vectors designed to express shRNA have been used as an alternative [12,13].

Read-out & evaluation of siRNA screening for viral host factors

In the siRNA screens performed to date, expression of a viral or reporter protein is usually used as the read-out to assess viral replication. Visualization of infected cells using antibodies directed against viral proteins, or the use of genetically altered viruses expressing green fluorescent protein if available, allows for quantitation of the percentage of cells infected. Alternatively, viruses engineered to contain enzymatic reporter genes (e.g., luciferase) will exhibit enzyme activity proportional to viral protein production. The chosen read-out, as well as the time postinfection that the assay is performed, can influence which aspects of the viral lifecycle are subject to examination. For example, assaying a viral protein expressed early in infection follows the virus through entry, uncoating or protein expression. (Of course, there may be several additional steps preceding translation – in the case of HIV, reverse transcription and proviral integration.) By contrast, detection of viral proteins whose expression first requires replication of the viral genome (i.e., ‘late’ genes) will add replication to the list of viral lifecycle steps assayed. In contrast, infection of siRNA-transfected cells at low multiplicity and analysis of the spread of infection between cells, or the production of infectious progeny, will allow evaluation of the entirety of the viral replication cycle.

For many cellular proteins, a significant reduction in intracellular levels may result in deleterious effects on the cell itself. Therefore, toxicity associated with protein knockdown must be monitored in the transfected cells. For immunofluorescence-based screens, 4',6-diamidino-2-phenylindole (DAPI) staining of cellular nuclei allows for quantitation of cells in each well, which can be performed in parallel with quantitation of numbers of virus-infected cells. Wells with significantly fewer cells than control wells can then be discarded from further analysis. Alternatively, fluorescent reagents that interact with cellular metabolic pathways, such as alamar blue, can be used to evaluate the overall viability of cells in siRNA-transfected wells.

In the studies described below, several methods were used to classify significant effects on viral replication (i.e., ‘hits’ worthy of further study; for a comprehensive review of these techniques, see [14]). Some of these methods are straight-forward and intuitive, such as the selection of wells displaying a twofold or greater reduction of infected cells compared with controls [15]. By contrast, others require more extensive statistical analysis, such as strictly standardized mean difference [16,17] or analysis of redundant siRNA activity (RSA) [18,19]. The RSA analysis selects targets for which two or more independent siRNAs, targeted to the same mRNA, demonstrate substantial suppression of virus replication. Although this is a robust technique that is less likely than other methods to miss weaker positives, RSA analysis requires that individual siRNA sequences be present in separate wells, and so may not be compatible with libraries in which the siRNAs are arranged in pools [14]. Finally, as Birmingham *et al.* point out, no matter what statistical methodology is applied to the screening results, extensive secondary screening must be employed to validate the hits, and actual ‘hands-on’ bench science, quite possibly without the benefit of multiwell plates and liquid-handling robots, will be needed to figure out which of the proteins identified in high-throughput screens truly contribute significantly to the viral lifecycle. First, secondary screens must confirm that the siRNAs used in the large screen indeed resulted in knockdown of the intended target. Second, experiments designed to elucidate the role of individual host cell proteins will vary depending on the virus and protein of interest, but will likely include identification of the stage of the viral replication cycle that is inhibited by siRNA-mediated knockdown of the protein, examination of proteins interacting with the protein of interest during viral infection, and potential substrates of the protein of interest (viral or cellular), if it possesses enzymatic activity.

Summaries of studies on individual viruses

HIV

With regard to viruses, RNAi-based screens have been used most extensively in the study of HIV. (Table 1 summarizes all of the genome-wide screens discussed in this review). To date, four independent studies have been published identifying host factors important in different phases of the HIV lifecycle [13,17,19,20], as well as a meta-analysis further examining the combined results of three of these studies [21]. Two of these studies performed the screen in HeLa cells expressing the viral co-receptor CD4, although the studies used different siRNA libraries [17,20]. The study by Brass *et al.* examined p24 Gag expression to examine early phases of HIV replication (i.e., viral entry, reverse transcription, integration and steps up to and including protein translation) [20]. Late-acting host factors involved in viral assembly and egress were identified by the transfer of supernatants of infected, siRNA-transfected cells to fresh reporter cells harboring a Tat-responsive β -galactosidase gene, followed by measurement of β -galactosidase activity (thereby measuring infectious progeny). The initial screen included siRNAs covering more than 20,000 genes. Of these, 273 resulted in a significant (more than two standard deviations from the mean) decrease in p24 or β -galactosidase activity, without overt cellular toxicity. In general, identification in the primary screen of multiple genes that are known to interact or participate in the same pathways can be taken as additional validation of the role of these genes in the viral lifecycle (at least, under the cell culture conditions used in the assay). Consistent with this rule of thumb, Brass *et al.* identified several components of the nuclear pore, autophagy, retrograde vesicular transport and the mediator complex (an RNA-polymerase II-associated complex involved in activation of transcription) [20].

A second screen conducted by Zhou *et al.* used a similar strategy to that of Brass *et al.* In this screen, a Tat-responsive β -galactosidase reporter was used following primary infection of siRNA-transfected cells (assaying viral entry through proviral integration and Tat expression), as well as to measure progeny infectious virus by culture of infected cell supernatants with fresh Tat-responsive cells [17]. This screen confirmed 232 genes that exerted an effect on HIV replication. Interestingly, only 15 of these genes were among those identified by Brass *et al.*, although this is a greater degree of overlap than would be expected by chance alone. Seven of the 15 overlapping genes were previously identified as playing a role in HIV replication, including the viral receptors CD4 and CXCR4 (CD4, in fact, was used by as a positive control to optimize the screening conditions). In addition, several components of the mediator complex identified by Brass *et al.*, as well as other components of this complex, were identified in this screen, further confirming the role of the mediator complex in HIV replication.

A third study by König *et al.* narrowed the scope of the screen by focusing on early events of the HIV lifecycle [19]. An HIV construct, deleted of the envelope (Env) sequence and possessing a luciferase reporter gene, was pseudotyped with vesicular stomatitis virus G protein, allowing this virus to complete a single round of replication, without producing infectious progeny. Therefore, this study excludes the steps of Env-dependent entry, but addresses the HIV lifecycle from uncoating to protein synthesis. As in the previous screens, the virus was used to infect 293T cells that had been transfected with an siRNA library. In parallel, Moloney murine leukemia virus (a γ -retrovirus), and adeno-associated virus (a parvovirus), each containing a luciferase reporter, were used to infect similarly transfected cells. Luciferase activity was measured at 24 h postinfection and subjected to an extensive filtering process, including RSA analysis (described previously), removal of siRNAs that showed cellular toxicity, removal of factors shared with adeno-associated virus and delineation of enriched functional networks by comparison of data with human protein-protein interaction databases derived from yeast two-hybrid experiments. Application of these criteria and others resulted in a final pool of 295 genes that act specifically on HIV (and murine leukemia virus,

which showed considerable similarity). Again, little overlap was observed with the previous studies.

In an elegant meta-analysis, Bushman *et al.* addressed this surprising lack of overlap by reexamination of the three siRNA-based screens, together with HIV protein–protein interaction data, previously published interactions of HIV with cellular host proteins, genetic studies of HIV-infected patients, as well as data from siRNA screens with other viruses (West Nile virus [WNV] and influenza, discussed in the following sections) [21]. The authors suggest that variation and lack of overlap may be partially attributed to experimental noise, differences in the timing of sampling and differences in filtering criteria used to select hits. This group had performed one of the previously discussed HIV screens [19], and had data available from a duplicate screen, allowing estimation of experimental variance. Given this variance, it is expected that only 150 of the top 300 hits would be obtained if the screen were repeated using identical experimental conditions. It is then perhaps less surprising that three screens performed under different, albeit similar, conditions produced so little overlap. Owing to the lack of overlap in the siRNA screens, this group looked across the siRNA screens for commonalities in proteins involved in similar cellular processes. Using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) [22] they identified gene ontology groups significantly enriched in each screen. In general, all three screens were highly enriched for the same gene ontology groups, including the nuclear pore, ubiquitin/proteasome, mediator complex, RNA binding and others. Therefore, analysis of enriched networks appears to be more robust than identification of individual proteins involved in viral replication.

In contrast to these three reports, a fourth, recently published, RNAi-based screen of HIV host factors contained several important differences in methodology. First, this study used Jurkat cells, a T-cell-derived line more closely resembling the natural targets of HIV infection. Second, instead of transfected siRNA pools, shRNA were introduced into target cells by a lentiviral vector. An RNA polymerase III-dependent promoter in the vector drives expression of the shRNA, which must be further processed by the cellular enzyme dicer before incorporation into the RISC. Furthermore, unlike the previously discussed siRNA transfection-based studies, in which siRNAs (or siRNA pools) targeting individual genes are segregated in 384-well plates, cells in this study were transduced by the lentiviral vectors and subjected to antibiotic selection to develop a mixed population of cell clones, each expressing a shRNA. As might be expected, many shRNAs were not tolerated over the course of prolonged selection, or resulted in enough of a growth disadvantage that these shRNAs were eventually eliminated from the population. Therefore, approximately 9300 clones, or 18% of the original library, were subsequently examined for an effect on HIV infection. Because Jurkat cells die 3–5 days postinfection with HIV, this screen used cell survival as a read-out of inhibition of virus replication. Following infection, RNA was isolated from surviving clones and shRNAs in the clones were identified using a microarray approach. Despite the fact that fewer cellular targets were examined than in the previously discussed siRNA-based screens (owing to the removal of many shRNAs because of toxicity or growth disadvantage), a comparable number of hits (252) were obtained in this screen. Not surprisingly, given the lack of overlap between the first three screens, few genes identified by the shRNA library had been found in any of the siRNA screens – none with Brass *et al.* [20], and three each with König *et al.* [19] and Zhou *et al.* [17]. Nevertheless, significant overlap in involvement in cellular signaling pathways and protein complexes was found between proteins identified in this study and those of the other three, including NF- κ B, AKT and PPAR- γ signaling, and the nuclear pore complex.

Flaviviruses

Several recent siRNA-based screens have focused on members of the virus family, *Flaviviridae*. Flaviviruses are single-stranded, positive-sense RNA viruses. Their genome

encodes a single polypeptide that is cleaved into functional subunits by host and virus proteases. These subunits include the structural components of the virus, as well as machinery required for replication of the viral RNA [23]. Using the same collection of pooled siRNAs as Brass *et al.* [20], Krishnan *et al.* examined the effect of host protein knockdown on a single cycle of WNV replication [15]. In this assay, HeLa cells, transfected with the gene-specific siRNA pools, were infected with WNV at low multiplicity (0.3 plaque-forming units/cell). A total of 24 h postinfection, cells were fixed and the viral envelope protein (Env) was detected by indirect immunofluorescent staining. The authors suggest that this assay is focused on early stages of infection – viral entry and RNA translation – while excluding replication, assembly and egress. However, translation of viral proteins occurs at much higher levels following replication and amplification of the genomic RNA [24], so it is therefore likely that RNA replication is also one of the facets of the viral lifecycle assayed by this screen. Setting a threshold of a twofold change compared with controls for selection of hits, the authors identified 305 genes that affect WNV. Knockdown of 283 of these proteins reduced viral gene expression (indicating that the targeted proteins are required for, or enhance, viral replication). These include some previously identified factors, such as components of the vacuolar ATPase, which helps produce the low pH environment of the endosome needed for viral fusion and entry. By contrast, 22 siRNAs enhanced viral gene expression, suggesting that the targeted proteins act to inhibit replication. These include IRF3, a major regulatory component of the cellular response to RNA viruses, including interferon production. As with the HIV screens, this screen identified several proteins that are present in the same macromolecular complexes, or have a previously documented interaction with other identified host factors. These include proteins involved in intracellular trafficking, endoplasmic reticulum-associated degradation and ubiquitination/proteasomal degradation pathways. The authors investigated the roles of some of the identified proteins and found that CBLL1, a ubiquitin ligase involved in endocytosis of cell surface proteins, is required for entry of WNV into the host cell. They further determined that several components of the proteasome function at a later stage in the virus lifecycle. In agreement with these findings, Gilfoy and Mason (who also contributed to the Krishnan *et al.* study) identified proteasome subunits in a smaller scale siRNA-based study (encompassing ~5500 genes) using WNV replicons containing a luciferase reporter [25]. These investigators determined that proteasome activity is required for amplification of the viral genome, although the exact mechanism of action still remains to be elucidated.

Krishnan *et al.* also asked if proteins that were identified as affecting WNV replication are also involved in the replication of a related flavivirus, Dengue virus (DENV). Using the same assay used for WNV, they determined that approximately one third of the factors also appeared to play a role in DENV infection. Interestingly, all the proteins whose silencing increased WNV infection had a similar effect on DENV, suggesting that these proteins are involved in restriction of virus (or, at least, positive-sense RNA virus) growth. This is certainly expected for host proteins for which an antiviral role has already been clearly defined, such as IRF3.

Sessions *et al.*, also conducted a screen for host factors associated with DENV [26]. Given that DENV replicates in a mosquito host, this group took advantage of the availability of siRNA libraries targeting the genome of *Drosophila melanogaster*. This library had previously been used to identify host factors for *Drosophila C virus* (DCV) [27,28]. In invertebrates, RNAi does not require 22 bp siRNAs, but can be achieved with long (~700 bp) dsRNAs, derived from sequences with the target mRNA. The dsRNAs are easily introduced into *Drosophila* S2 cells, requiring only co-incubation with the cells [29]. Another advantage of RNAi screening in *Drosophila* is the relative absence of redundant genes within the *Drosophila* genome, making it less likely that knockdown of a single host gene will be compensated for by a related protein [30]. Following a duplicate screen and subsequent rescreen of hits, 116 genes were identified as *Drosophila* host factors for DENV. In total, 82 of the 116 host factors had identifiable human homologs, allowing a further screen to determine if these genes are required for DENV growth

in human cells. Following transfection of Huh-7 cells with siRNAs targeting the human homologs, 42 scored as positive for knockdown of DENV growth.

HCV is also a member of the family *Flaviviridae* and, therefore, uses a similar replication strategy to WNV and DENV. Approximately 170 million people are infected with HCV worldwide, making this a virus of considerable importance with regards to public health [31]. Therefore, it is not surprising that many groups have used siRNA-based approaches for identification of HCV-associated host factors, with an eye to the development of novel therapeutics [32–37]. To date, however, only one has used a library covering the entire genome [38]. This screen was carried out in Huh-7 cells harboring an HCV luciferase-expressing replicon. The replicon contains neomycin phospho-transferase and luciferase genes at the 5' end of the genome, followed by the HCV non-structural (NS) proteins NS3, NS4 and NS5. The HCV NS proteins form an active replication complex that maintains the RNA genome in the host cells. Luciferase activity is directly proportional to the copy number of replicon RNA in the cell. Since the replicon is present within the cell at the beginning of the assay, and is unable to produce infectious progeny, this assay examines neither entry nor egress, but focuses on viral protein expression and RNA replication. The initial results yielded 236 hits. This was further pared down to 96 following secondary screening, in which a threshold was set requiring at least two individual siRNAs from the pool to effect reduction of luciferase expression. Little overlap was observed in hits between this screen and other siRNA screens performed earlier [33–35]. Tai *et al.* observed that several previously identified putative HCV host factors may have been identified in their screen if less stringent selection criteria were used [38]. This may also reflect the possibility that certain weaker targets may be more readily identified in smaller screens. Nevertheless, both Tai *et al.* [38] and Berger *et al.* [32], identified a requirement for phosphatidylinositol 4-kinase in HCV replication. Both groups defined a role for phosphatidylinositol 4-kinase in the formation of the membranous web, an intracellular structure present in HCV-infected cells that is thought to be a necessary structure for the formation of HCV replication complexes [39].

Influenza

Influenza viruses are negative stranded, multisegmented viruses of the family *Orthomyxoviridae*. Influenza A represents a seasonal public health threat and periodically has the potential to become a global pandemic, arguably representing the greatest viral public health threat to mankind [40]. Hao *et al.* used a *Drosophila* dsRNA library to identify influenza host factors [41]. Since influenza does not normally infect insect cells, a genetically modified virus in which the vesicular stomatitis virus G protein and a luciferase reporter replace the hemagglutinin and neuraminidase coding regions was constructed for this assay. Furthermore, this virus does not release infectious progeny into the culture supernatant of infected *Drosophila* cells. Therefore, similar to the replicon-based screen, this screen is focused on the middle stages of the virus lifecycle, including the release and nuclear import of the viral ribonucleoprotein complexes, mRNA synthesis and translation. Screening of the dsRNA library identified 110 genes that reduced luciferase expression from this virus, several of which had clear human homologs. These include genes expected to play a role in influenza replication, including a subunit of the v-ATPase, a component of mitochondrial electron transport and an mRNA nuclear export factor. These human homologs were analyzed in human HEK293 cells by knockdown with specific siRNAs. These experiments verified not only reduction of luciferase from the reporter virus in human cells, but also reduction of intact influenza virus.

Conclusion & future perspective

Although individual host proteins that play highly significant roles in the viral lifecycle will certainly be identified in RNAi-based screens (e.g., CD4 in assays of HIV infection), the fact that several of the screens of the same or similar viruses identified components of the same

complexes (e.g., the mediator complex and the nuclear pore) or pathways (e.g., proteasomal degradation) suggests that the strength of the RNAi-based screens lies in this technique's ability to identify protein networks and macromolecular complexes involved in viral replication. Nevertheless, this assertion needs to be verified by careful analysis of the networks and complexes in question in smaller screens, under conditions in which the efficacy of each siRNA can be verified. In addition, as demonstrated in several of the papers discussed [19,21], comparison of the siRNA data sets to data sets obtained by other experimental protocols (e.g., 'interactomes' derived from two-hybrid or proteomic analysis that define the interactions of host cell and viral proteins) may provide valuable clues as to how individual host proteins exert their effect on viral replication.

As the popularity of this technique increases and more groups add to the 'public data set' (especially for viruses that are of significant public health interest), a clearer picture of which cellular processes contribute to the viral lifecycle will emerge. As these leads are further investigated, the mechanistic details of how these proteins function to affect viral replication will be elucidated. This, in turn, may lead to new therapies for the treatment of viral infection.

Executive summary

Use of siRNA in protein knockdown & screening

- siRNAs are 21–23 bp dsRNAs that induce cleavage of target mRNAs in a sequence-specific manner.
- Individual siRNAs vary in their ability to cause protein knockdown or induce off-target effects.
- Large screens usually include multiple siRNAs for each target.

Read-out and evaluation of siRNA screening for viral host factors

- Evaluation of viral replication is usually performed by immunofluorescent staining or with an enzymatic reporter gene, such as luciferase.
- Cellular toxicity of individual siRNAs must also be evaluated.
- Statistical methods and filters currently in use for evaluation data vary widely in terms of complexity and stringency.

Summaries of studies on individual viruses

- To date, genome-wide RNAi-based screens for host factors have been carried out for HIV (four studies), West Nile virus, Dengue virus, hepatitis C virus and influenza.
- Overlap between screens of individual proteins identified as host factors for individual viruses tends to be surprisingly low, but many of the same host cell pathways or macromolecular complexes are represented across all screens.

Acknowledgments

The author thanks Jessica Smith and David Stein for helpful discussions and critical reading of this manuscript.

Bibliography

Papers of special note have been highlighted as:

- of interest

•• of considerable interest

1. Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 2001;411(6836):494–498. [PubMed: 11373684]
2. Cherry S. What have RNAi screens taught us about viral–host interactions? *Curr. Opin. Microbiol* 2009;12(4):446–452. [PubMed: 19576842]
3. Goff SP. Knockdown screens to knockout HIV-1. *Cell* 2008;135(3):417–420. [PubMed: 18984154]
4. Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorova A. Rational siRNA design for RNA interference. *Nat. Biotechnol* 2004;22(3):326–330. [PubMed: 14758366]
5. Westerhout EM, Berkhout B. A systematic analysis of the effect of target RNA structure on RNA interference. *Nucleic Acids Res* 2007;35(13):4322–4330. [PubMed: 17576691]
6. Elbashir SM, Harborth J, Weber K, Tuschl T. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* 2002;26(2):199–213. [PubMed: 12054897]
7. Holen T, Amarzguoui M, Wiiger MT, Babaie E, Prydz H. Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. *Nucleic Acids Res* 2002;30(8):1757–1766. [PubMed: 11937629]
8. Aleman LM, Doench J, Sharp PA. Comparison of siRNA-induced off-target RNA and protein effects. *RNA* 2007;13(3):385–395. [PubMed: 17237357]
9. Birmingham A, Anderson EM, Reynolds A, et al. 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat. Methods* 2006;3(3):199–204. [PubMed: 16489337]
10. Jackson AL, Burchard J, Schelter J, et al. Widespread siRNA 'off-target' transcript silencing mediated by seed region sequence complementarity. *RNA* 2006;12(7):1179–1187. [PubMed: 16682560]
11. McManus MT, Sharp PA. Gene silencing in mammals by small interfering RNAs. *Nat. Rev. Genet* 2002;3(10):737–747. [PubMed: 12360232]
12. Paddison PJ, Silva JM, Conklin DS, et al. A resource for large-scale RNA-interference-based screens in mammals. *Nature* 2004;428(6981):427–431. [PubMed: 15042091]
13. Yeung ML, Houzet L, Yedavalli VS, Jeang KT. A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication. *J. Biol. Chem* 2009;284(29):19463–19473. [PubMed: 19460752] • Screen for host factors influencing HIV replication using a T-cell line stably transduced with a shRNA library.
14. Birmingham A, Selfors LM, Forster T, et al. Statistical methods for analysis of high-throughput RNA interference screens. *Nat. Methods* 2009;6(8):569–575. [PubMed: 19644458] •• Comprehensive review of statistical methods for analysis of RNAi-based screens.
15. Krishnan MN, Ng A, Sukumaran B, et al. RNA interference screen for human genes associated with West Nile virus infection. *Nature* 2008;455(7210):242–245. [PubMed: 18690214] • Genome-wide siRNA screen for factors associated with the flaviviruses West Nile virus and Dengue virus.
16. Zhang XD, Yang XC, Chung N, et al. Robust statistical methods for hit selection in RNA interference high-throughput screening experiments. *Pharmacogenomics* 2006;7(3):299–309. [PubMed: 16610941]
17. Zhou H, Xu M, Huang Q, et al. Genome-scale RNAi screen for host factors required for HIV replication. *Cell Host Microbe* 2008;4(5):495–504. [PubMed: 18976975] • Genome-wide siRNA screen for factors associated with early and late stages of HIV replication.
18. Konig R, Chiang CY, Tu BP, et al. A probability-based approach for the analysis of large-scale RNAi screens. *Nat. Methods* 2007;4(10):847–849. [PubMed: 17828270]
19. Konig R, Zhou Y, Elleder D, et al. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* 2008;135(1):49–60. [PubMed: 18854154] •• Genome-wide siRNA screen for factors associated with early and late stages of HIV replication. Probably the most rigorous statistical analysis of screens discussed herein.
20. Brass AL, Dykxhoorn DM, Benita Y, et al. Identification of host proteins required for HIV infection through a functional genomic screen. *Science* 2008;319(5865):921–926. [PubMed: 18187620] • Genome-wide siRNA screen for factors associated with early and late stages of HIV replication
21. Bushman FD, Malani N, Fernandes J, et al. Host cell factors in HIV replication: meta-analysis of genome-wide studies. *PLoS Pathog* 2009;5(5):e1000437. [PubMed: 19478882] •• Analysis of

- combined dataset from HIV screens of [17,19,20]. Also compares these screens to data from protein–protein interaction studies.
22. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc* 2009;4(1):44–57. [PubMed: 19131956]
 23. Lindenbach, BD.; Rice, CM. *Flaviviridae: the viruses and their replication*. In: Knipe, DM.; Howley, PM., editors. *Fields Virology*. PA, USA: Lippincott Williams & Wilkins; 2001. p. 991-1041.
 24. Puig-Basagoiti F, Deas TS, Ren P, Tilgner M, Ferguson DM, Shi PY. High-throughput assays using a luciferase-expressing replicon, virus-like particles, and full-length virus for West Nile virus drug discovery. *Antimicrob. Agents Chemother* 2005;49(12):4980–4988. [PubMed: 16304161]
 25. Gilfoy F, Fayzuln R, Mason PW. West Nile virus genome amplification requires the functional activities of the proteasome. *Virology* 2009;385(1):74–84. [PubMed: 19101004]
 26. Sessions OM, Barrows NJ, Souza-Neto JA, et al. Discovery of insect and human dengue virus host factors. *Nature* 2009;458(7241):1047–1050. [PubMed: 19396146] • dsRNA-based screen for genes affecting dengue virus replication in *Drosophila* cells and analysis of human homologs of identified genes.
 27. Boutros M, Kiger AA, Armknecht S, et al. Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* 2004;303(5659):832–835. [PubMed: 14764878]
 28. Cherry S, Doukas T, Armknecht S, et al. Genome-wide RNAi screen reveals a specific sensitivity of IRES-containing RNA viruses to host translation inhibition. *Genes Dev* 2005;19(4):445–452. [PubMed: 15713840]
 29. Clemens JC, Worby CA, Simonson-Leff N, et al. Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc. Natl Acad. Sci. USA* 2000;97(12):6499–6503. [PubMed: 10823906]
 30. Perrimon N, Friedman A, Mathey-Prevot B, Eggert US. Drug-target identification in *Drosophila* cells: combining high-throughput RNAi and small-molecule screens. *Drug Discov. Today* 2007;12(1–2):28–33. [PubMed: 17198970]
 31. Houghton, M. *Hepatitis C Virus*. 3rd Edition. PA, USA: Lippencott-Raven; 1996.
 32. Berger KL, Cooper JD, Heaton NS, et al. Roles for endocytic trafficking and phosphatidylinositol 4-kinase III α in hepatitis C virus replication. *Proc. Natl Acad. Sci. USA* 2009;106(18):7577–7582. [PubMed: 19376974]
 33. Ng TI, Mo H, Pilot-Matias T, et al. Identification of host genes involved in hepatitis C virus replication by small interfering RNA technology. *Hepatology* 2007;45(6):1413–1421. [PubMed: 17518369]
 34. Randall G, Panis M, Cooper JD, et al. Cellular cofactors affecting hepatitis C virus infection and replication. *Proc. Natl Acad. Sci. USA* 2007;104(31):12884–12889. [PubMed: 17616579]
 35. Supekova L, Supek F, Lee J, et al. Identification of human kinases involved in hepatitis C virus replication by small interference RNA library screening. *J. Biol. Chem* 2008;283(1):29–36. [PubMed: 17951261]
 36. Trotard M, Lepere-Douard C, Regeard M, et al. Kinases required in hepatitis C virus entry and replication highlighted by small interference RNA screening. *FASEB J* 2009;23(11):3780–3789. [PubMed: 19608626]
 37. Vaillancourt FH, Pilote L, Cartier M, et al. Identification of a lipid kinase as a host factor involved in hepatitis C virus RNA replication. *Virology* 2009;387(1):5–10. [PubMed: 19304308]
 38. Tai AW, Benita Y, Peng LF, et al. A functional genomic screen identifies cellular cofactors of hepatitis C virus replication. *Cell Host Microbe* 2009;5(3):298–307. [PubMed: 19286138] • Screen for factors affecting HCV replication using a subgenomic HCV replicon.
 39. Gosert R, Egger D, Lohmann V, et al. Identification of the hepatitis C virus RNA replication complex in Huh-7 cells harboring subgenomic replicons. *J. Virol* 2003;77(9):5487–5492. [PubMed: 12692249]
 40. Gatherer D. The 2009 H1N1 influenza outbreak in its historical context. *J. Clin. Virol* 2009;45(3):174–178. [PubMed: 19540156]
 41. Hao L, Sakurai A, Watanabe T, et al. *Drosophila* RNAi screen identifies host genes important for influenza virus replication. *Nature* 2008;454(7206):890–893. [PubMed: 18615016] • dsRNA-based screen for genes affecting influenza replication in *Drosophila* cells.

Table 1

RNAi-based screens for viral host factors.

Virus	siRNA library	Cell type	Read-out	Genes identified (n)	Ref.
HIV (IIIB)	21,121 genes (four siRNAs/ pool)	TZM-bl (HeLa expressing CD4, CCR5, and integrated Tat-dependent β -galactosidase)	IF for p24 expression, 48 h pi B-gal expression in naive cells incubated with culture supernatant	273	[20]
HIV (HXB2)	19,709 genes (three siRNAs/ pool)	HeLa P4/R5 (CD4 expressing, integrated Tat-dependent β -galactosidase)	β -gal expression	232	[17]
HIV (VSV-pseudotyped HIV-1-luciferase vector)	19,628 genes (six siRNA/gene, two siRNA/pool)	293T	Luciferase activity	295	[19]
HIV (NL4-3)	Lentiviral shRNA library targeting 54,509 transcripts (producing 9357 viable clones)	Jurkat	Cell survival	252	[13]
WNV (2471) DENV (NGC)	21,121 genes (four siRNAs/ pool)	HeLa	IF	305 (WNV) 124 shared with DENV	[15]
DENV-2 (adapted to <i>Drosophila</i> cells)	22,632 dsRNAs targeting <i>Drosophila</i> genome	<i>Drosophila</i> S2 cells	IF	116 (confirmed 42 human homologs)	[26]
HCV (genotype 1b Luc replicon)	21,121 genes (four siRNAs/ pool)	Huh7 containing replicon	Luciferase activity	96	[38]
Flu FVG-R (a genetically modified virus derived from influenza A/WSN/33 by replacement of HA and NA with VSV-G and luciferase).	13,071 dsRNAs targeting <i>Drosophila</i> genes	<i>Drosophila</i> DLI	Luciferase activity	110	[41]

DENV: Dengue virus; FVG-R: Flu-VSV-G-Renilla; HA: Hemagglutinin; IF: Immunofluorescence; NA: Neuraminidase; pi: Post-infection; VSV-G: Vesicular stomatitis virus G protein; WNV: West Nile virus.