RNA interference in mammals: behind the screen

Eric Campeau and Stéphane Gobeil

Abstract

The discovery of RNA interference (RNAi) and the development of technologies exploiting its biology have enabled scientists to rapidly examine the consequences of depleting a particular gene product in a cell or an animal. The availability of genome-wide RNAi libraries targeting the mouse and human genomes has made it possible to carry out large scale, phenotype-based screens, which have yielded seminal information on diverse cellular processes ranging from virology to cancer biology. Today, several strategies are available to perform RNAi screens, each with their own technical and monetary considerations. Special care and budgeting must be taken into account during the design of these screens in order to obtain reliable results. In this review, we discuss a number of critical aspects to consider when planning an effective RNAi screening strategy, including selecting the right biological system, designing an appropriate selection scheme, optimizing technical aspects of the screen, and validating and verifying the hits. Similar to an artistic production, what happens behind the screen has a direct impact on its success.

Keywords: shRNA; siRNA; esiRNA; RNAi screen; lentivirus; retrovirus

INTRODUCTION

The ability to genetically manipulate model organisms has furnished scientists with an incredible set of tools with which to understand human biology as well as the mechanistic basis of human diseases. From the selective breeding of plants, flies, bacteria and yeast, to genetic complementation using fused cell heterokaryons or cDNA libraries, our grasp of gene regulation and its involvement in development, aging and disease has been increasing exponentially. Recently, the discovery of RNA interference (RNAi), and how it regulates gene expression, has given scientists an additional tool to further our understanding of human biology. Using this instrument, almost any gene product can be selectively depleted and, through the advent of genome-wide RNAi libraries, this process can be accomplished in a high throughput and unbiased manner. Several excellent reviews discussing RNAi screens have been recently published [1–10], some directed towards specific applications, and the reader should refer to them for specific examples and experimental details. Here, we review the various types of RNAi screens, with a focus on the technical considerations and design aspects essential to building a successful loss-of-function screen in mammalian cells. Similar to an artistic production, the success of these screens lies in the technical support and preparation behind the curtain.

Corresponding authors: Eric Campeau, Translational Biology Group, Resverlogix Corp 140, 3553 31st Street NW, Calgary, AB, T2L 2K7, Canada, Tel: +1 403 699 1023, Fax: +1 403 210 1725, E-mail: ecampeau@gmail.com.
Stéphane Gobeil, CHUQ Research Centre, Department of Molecular Medicine, Faculty of Medicine, Laval University, 2705 Laurier boulevard, Québec, G1V 4G2, Canada. Tel: +1 418 654 2296, Fax: +1 418 654 2761, E-mail: stephane.gobeil@crchul.ulaval.ca

Eric Campeau obtained his PhD degree from McGill University (Montréal, Canada). He was a postdoctoral fellow and a scientist at the Lawrence Berkeley National Laboratory (Berkeley, CA) and an instructor in the Program in Gene Function and Expression at the University of Massachusetts Medical School (Worcester, MA). He is now a Research Scientist at Resverlogix Corp. (Calgary, Canada). One of his research interests is to use lentiviral and RNAi strategies to decipher the roles of histone modifications and their regulation in genome stability.
Stéphane Gobeil obtained his PhD from Laval University (Québec, Canada). He was a postdoctoral fellow at the University of Massachusetts Medical School and is now an assistant professor in the department of Molecular Medicine at Laval University (Québec, Canada). One of his research interests focuses on elucidating the molecular basis of cancer metastasis through the use of RNAi screens.

THE ACTORS (TYPES OF RNAi MOLECULES)

siRNA/esiRNAs

Small interfering RNAs (siRNAs) are synthesized duplexes of RNA ranging in size from 19 to 27 nt [11]. Numerous companies offer individual siRNAs as well as libraries either targeting the entire human or mouse genome, or focusing on a subset of molecules, such as kinases or phosphatases. siRNA libraries generally come pre-arrayed in 96- or 384-well format and each gene is usually targeted by several independent siRNAs. Enzymatically generated siRNAs (esiRNAs) are produced from *in vitro* transcribed long dsRNAs (300–500 bp), which are digested to short double-stranded RNAs by the nuclease Dicer or RNase III. This type of resource can be generated by the researcher [12] or acquired commercially. Both siRNAs and esiRNAs are introduced into cells by transient transfection.

shRNAs

Small (or short) hairpin RNAs (shRNAs) are ∼65 nt RNAs that contain complementary sequences at their 3′- and 5′-ends. shRNAs are typically cloned into a viral vector backbone, which, with the assistance of a packaging cell line, can be used to generate viral particles to stably transduce the cell type of choice. Upon transcription, the RNA folds back on itself and forms a short hairpin with a small 4- to 8-nt loop. The shRNA is recognized by the Dicer complex, which cleaves the shRNA to generate siRNAs [13–15]. A variant of shRNAs, the shRNAmir, was introduced by the Hannon and Elledge laboratories [16]. In this variant, the shRNA is embedded in the mir-30 scaffold so it resembles and is processed like an endogenous microRNA. Unlike shRNAs, which can only be transcribed by pol III, the shRNAmir can be transcribed by either RNA pol II or pol III. Similar to siRNAs and esiRNAs, shRNAs can be acquired, either individually or in library format, through several providers. Most companies provide shRNAs as either bacterial clones or viral particles.

miRNAs

MicroRNA (miRNA)-based screens refer to the recently available libraries either overexpressing or depleting all known mouse or human miRNAs. We included these libraries in this review because they rely on the mechanism of RNAi for their action, and therefore similar precautions must be used when designing this type of screen.

Although siRNAs, esiRNAs and shRNAs can all lead to robust depletion of the targeted gene products, there are specific advantages and disadvantages to each one (summarized in Table 1). The major advantages of siRNAs and esiRNAs are that: (i) they can be transfected or electroporated in the target cell of choice with relatively high efficiency [17–19]; (ii) they can be rapidly and, in the case of esiRNAs, economically generated; and (iii) they can be associated with reduced off-target effects (esiRNAs are composed of several different siRNA-like oligonucleotides, thereby diluting off-target effects [20], and siRNAs can be chemically or structurally modified to greatly reduce the likelihood of off-target effects (reviewed in [21]). Their biggest disadvantages are: (i) they produce only transient effects, thereby limiting the screens to phenotypes that can be observed in a short period of time; (ii) some cell lines can be refractory to transfection and electroporation; (iii) siRNA libraries have a finite 'lifespan' and have to be re-ordered/re-synthesized; and (iv) they are not particularly suitable for pooling strategies, and instead require an array format (see below).

For shRNAs, the advantages include: (i) stable incorporation in the genome of the target cell line, which allows screening for phenotype developing over a long period of time, including in model animals; and (ii) they are amenable to pooling strategies (see below). Their biggest disadvantages are that they require: (i) more experimental setup if viruses need to be generated; (ii) a biosafety level 2 (BSL2) containment for retrovirus screens and a BSL2$^+$ for lentivirus screens.

The miRNA libraries are the latest tools to study the increasingly vast roles played by miRNAs in the cell (reviewed in [22]). The recent addition of these libraries represents the advantage of a relatively untapped resource that could yield significant findings. However, these libraries are most likely incomplete, and potentially more miRNAs will be discovered, as well as other regulatory RNAs such as piwi RNAs (piRNAs) and long intervening non-coding RNAs (lincRNAs) (reviewed in [23, 24]) which are not part of any library yet.

Table 1: Advantages and disadvantages of various types of RNAi screens

Type of RNAi Screen	Advantages	Disadvantages
siRNA	• Can be easily introduced into cells with high efficiency • Easily and rapidly generated. • Chemical modifications possible to reduce off-target effects.	• Non-renewable resource. • Cells might not be transfectable. • Transient effect. • Array format required.
esiRNA	• Can be easily introduced into cells with high efficiency • Can be rapidly produced in the laboratory (<1 day). • Economical. • Reduced off-target effects.	• Cells might not be transfectable • Transient effect. • Array format required.
shRNA (viral)	• High transduction efficiencies. • Transduced cells can be selected through drug resistance or reporter[a]. • Long-term effect. • Can be pooled or arrayed. • Inducible expression[a,b].	• BSL2 facilities required. • More technologically challenging. • Virus-mediated toxic effects.
miRNA (mimic and hairpin)	• Same advantages as viral shRNAs • Uncharted territory for many biological processes.	• Same disadvantages as viral shRNAs. • Possibly incomplete and erroneous libraries.

[a]Not available on all vectors.
[b]Requires compatible cell line.

Table 2: Advantages and disadvantages of various RNAi screen formats

Format of RNAi Screen	Advantages	Disadvantages
Pools	• Low-throughput possible. • More economical.	• Cells need to be isolated or enriched. • Might require deconvolution for weak phenotype.
Arrayed	• Any observable phenotype can be screened.	• High-throughput equipment required. • Thorough statistical analysis required. • More costly.
Positive	• Easier. • More economical.	• Restricted to selectable phenotypes.
Negative	• Can do synthetic lethal screens for drug screenings.	• Requires next generation sequencing or microarray. • Need reference (untreated) control. • More costly.

THE PRODUCERS (RNAi SCREEN FORMAT)
Types of RNAi libraries

RNAi libraries (or sub-libraries) can be arrayed in multiwell plates, each well containing a single shRNA/siRNA species (arrayed format) or for shRNA libraries, in pools containing 5000–10 000 shRNAs (pool format). The arrayed format is the choice for siRNA-based screens, although shRNA/miRNA screens can also be performed in this format [25]. These screens are suitable for any phenotype that can be assayed or visualized. The advantage of the array format is that it does not require the selection or isolation of the desired phenotype, as long as it can be detected or assayed in the well (Table 2). However, the major disadvantage is its requirement for high throughput equipment to transfect/transduce the cells, collect the data

and analyze the results. It also costs significantly more than the pooled format in most cases.

Pools of shRNAs are habitually transduced into a cell line at a high representation (each shRNA is present at approximately 1000 copies within the pool) and a relatively low multiplicity of infection ($MOI = 0.1–1.0$). The high representation ensures that every shRNA present in the library has a chance to exert its role, and the low MOI ensures that the majority of cells receive only one shRNA. Pooled screens can be used to interrogate any phenotype that can be selected for, fractionated, or isolated. Cells harboring the selected phenotype are isolated and the candidate shRNAs identified by microarray or sequencing. The significant advantage of the pooled library approach is certainly its low cost and its feasibility to be used by labs that do not have high throughput screening capabilities. On the other hand, its main disadvantages include the need to isolate or select the desired phenotype from the total population and the difficulty to screen for complex phenotypes. Pool libraries are available from several commercial providers, and recently the DECIPHER open source RNAi screening project offers free of charge access to lentiviral shRNA pools and softwares to academic and non-profit institutions (http://www.decipherproject.net). Arrayed and pooled screens can be subdivided into positive and negative screens.

Positive screens

Positive screens identify siRNAs/shRNAs that confer a cellular phenotype that can be selected for (e.g. increased survival, invasion or migration capacities) or isolated through, for example, cell sorting or a reporter assay [26–31] . A few groups have also used a mouse model for their positive screens, which they term *in vivo* RNAi screen [32–35]. In positive selection screens using pooled libraries, the shRNA of interest can be identified by conventional sequencing of genomic DNA isolated from the positively selected/isolated cell. Alternatively, microarrays (using the barcode present in certain shRNA libraries or the hairpin sequence) and deep sequencing can be used to identify shRNAs that are enriched over the basal population. Different strategies have been proposed to identify the significant hits using the full hairpin, a half hairpin or the barcode present in some libraries [36, 37].

Negative screens

Negative screens rely on the dropout of siRNA/ shRNA sequences from the treated population compared to a control, reference population. For example, the shRNA/siRNA of interest may result in cell death, and would therefore be significantly underrepresented—or absent—relative to the control population [38, 39]. Notably, shRNA pooled negative screens can currently only be accomplished through microarrays or deep sequencing. Synthetic lethal screens, a type of negative screen, originated from classical yeast genetic screens designed to identify two mutations that, in combination (but not alone), result in lethality [40]. Synthetic lethal screens have therapeutic applications in cancer biology, where they may be used to identify 'second hits' that can kill a cell bearing an oncogenic mutation. For example, recent RNAi screens have identified synthetic lethal interactions with the Ras oncogene [41] and the KRas oncogene [42]. A variant approach, referred to as chemical synthetic lethality, screens for genetic factors that modify the effect of a chemical [1]. This approach is extremely powerful to elucidate genetic networks and can have direct clinical applications. For example, synthetic lethal RNAi screens have been used to identify genes that, when knocked down, confer sensitivity to a variety of chemotherapeutic drugs including IKKβ inhibitors [43], olaparib [44], gemcitabine [45], paclitaxel [46] and camptothecin [47]. Synthetic lethal RNAi screens may also be used to identify new pathways involved in angiogenesis, cell growth, migration, or metastasis, which could be harnessed in regenerative medicine or cancer treatments. As pointed out by Mullenders and Bernards [1], an advantage of RNAi over the traditional knockout is that the presence of residual gene activity conferred by RNAi more closely resembles the physiological situation achieved with pharmacological inhibition than the knockout which completely eliminates any gene activity.

THE SCRIPT (THE EXPERIMENTAL DESIGN)

As with any production, the storyline is critical for its success. Careful design and planning of the screen cannot be stressed enough. There are several key questions to consider. What is the biological question? Are there any indications that the cell line or mouse strain to be used in the screen can acquire the
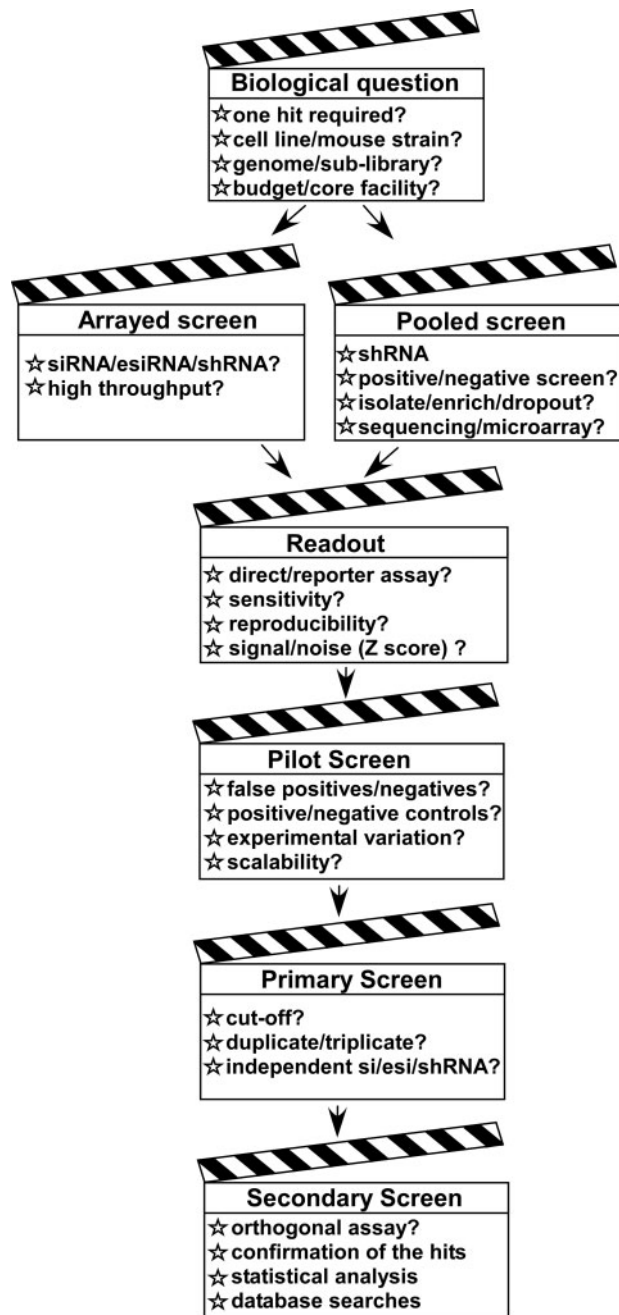
Figure 1: Overview of the planning and execution of a successful RNAi screen. From the biological question to answer, the selection of the RNAi library and screen format, to the execution of the screen and analysis of its results, several technical and practical aspects need to be considered.

desired phenotype? If so, can the phenotype be selected, isolated or observed/assayed in a high throughput fashion? Can the predicted phenotype be efficiently and reproducibly measured significantly above the experimental background? Depending on

the script, certain actors and producers might be better qualified (Figure 1). For example, if the cells that display the desired phenotype can be easily selected or isolated, a pooled library might be preferred. On the other hand, if the desired phenotype can only be observed under certain conditions or needs to be quantified, an arrayed format would be a better choice. If the cells are hard to transfect, a viral-based library might be preferable, and if the cells are slow or non-dividing, lentiviruses are well known for their ability to transduce arrested cells [48, 49]. Another variable is the budget for the production. Large laboratories or core facilities are likely to have the experience, equipment, and financial support necessary to carry out large-scale genomic arrayed screens, whereas smaller laboratories might prefer pooled viral shRNA screens or smaller sub-libraries made from in-house generated esiRNAs. Several screens have been done successfully in smaller settings using either approach [26, 28, 50–53].

An important reminder concerning RNAi is that the depletion of a target protein is never complete, and in many cases, a significant amount of biological activity remains. Therefore, some important players might not be identified if the phenotype to be measured requires a complete loss-of-function. On the other hand, partial depletions are useful to interrogate genes essential for cell viability. Additional criteria that also need to be taken into account in the experimental design are discussed below.

The one hit wonder

Irrespective of the format used, most genome-wide RNAi screens can only assess the effect of single experimental knockdowns on the phenotype of interest; for siRNA screens, one gene is depleted per well, whereas in viral-based screens, cells are infected at a low MOI that averages one shRNA per cell (as discussed above). Consequently, experimental designs that require the knockdown of multiple genes in order to obtain the phenotype of interest are not easily amenable to large genetic screens with today's tools. However, RNAi screens designed around phenotypes controlled by genes involved in redundant pathways can potentially be performed if one of the pathways is chemically or genetically inactivated. Similarly, a cell line could be engineered to be as close as possible to the desired phenotype, for example, by inactivating a tumor suppressor gene in a cancer initiation screen, or pre-selecting a cell line

with an intermediate drug sensitivity for a drug resistance or small molecule screen. However, the rate at which the cell line spontaneously acquires the phenotype of interest should be very low to avoid an unmanageable rate of false positives (see below).

The scene for the show (the experimental system)

Careful consideration should be taken when choosing an experimental system for the RNAi screen. The hits obtained from a screen might be more clinically relevant under certain conditions or by using a specific cell line or model system. A critical aspect to consider when choosing a cell line is whether it can be efficiently transfected/transduced to ensure that the cells take up most si/sh/miRNAs. Inadequate transfection/transduction will result in an incomplete representation of the library and increase the likelihood that candidates might be missed. Preliminary experiments using fluorescently labeled siRNA or lenti/retroviral particles encoding a fluorescent protein can quickly indicate the potential of a candidate cell line to be transfected or transduced. Under certain circumstances, the choice of the cell line is mandated by the biological question to be answered. In this case, a thorough optimization of the transfection/transduction protocol is imperative. Another consideration to take in account when choosing the experimental system is that the results can drastically differ depending on the assay conditions, even in the same cell line. For example, a comparison was done between three similar RNAi screens carried out by different groups to identify factors important for HIV replication [54]. All three groups used an siRNA format; two groups used a β-galactosidase reporter assay in HeLa cells [55, 56] whereas the third group used 293T cells [57]. Surprisingly, the overlap between the identified candidates was only 5–7%. It remains to be determined if these differences resulted from different experimental setups or from the statistical analysis of the candidates.

The broadcasting (the readout)

The readout is another crucial aspect of the screen: a broad readout might identify different genes involved in various processes. For example, Neumann *et al.* [58] used time-lapse microscopy of a fluorescently labeled histone coupled to machine learning to screen for genes necessary for mitosis, migration and survival. Any phenotype, measured directly or indirectly through a reporter, could be used as readout. Importantly, the readout should be assessed at a time when the phenotype is maximally expressed. For siRNA-based screens, most of the readouts will occur between 48 and 72 h post-transfection, although survival-based assays could be read at later time points. The timing of the readout is more flexible for shRNA-based screens: proteins that are rapidly depleted, and which therefore confer an immediate effect, can be assayed early in the screen, whereas proteins that have a longer half-life can be assayed at later time points. The format of the screen will influence the readout and its associated costs. For pool-based screens, simpler readouts like cell survival are often used, which help keep their costs to a minimum. For arrayed screens, more complex readouts can be accommodated with correspondingly higher costs. Free publicly available cell imaging software such as the CellProfiler suite, [59–61], CellCognition [62], Advanced Cell Classifier [63], μmanager [64], micropilot [65] and EBImage [66] can be installed on a variety of microscopes to help lower some of the costs.

The false prophets (the false positives/negatives)

Once the readout has been chosen, its level of sensitivity should be experimentally measured and determined. If not sensitive enough, a high frequency of false negatives will be attained, and only a few (or no) candidates will be isolated. Conversely, if the readout is too sensitive, a significant number of false positives will be identified. In the former case, important candidate genes might be missed, and an incomplete or biased picture might emerge, whereas in the latter case the real candidates might be swamped in a sea of irrelevant hits and no significance would be detected. Furthermore, considerable amounts of time and reagents would subsequently be spent to confirm irrelevant candidates. The best way to ascertain the rates of false negatives and positives is to perform a pilot screen.

THE REHEARSAL (THE PILOT SCREEN)

Before spending valuable resources and efforts on a screen, the readout must be fine-tuned to generate reproducible results under normal experimental variations. Several factors influence the levels of false positives and negatives generated by the screen, including the number of cells that will be used, the

concentration of drugs used to select for the anticipated phenotype and/or the stringency of the readout. The use of a negative controls, such as a scrambled, non-silencing or luciferase siRNA/shRNA is imperative, as it will give a direct indication of the false positive rate. Repeating the pilot screen a few (two to three) times will also give an indication of the experimental variation inherent to the assay and allow fine-tuning of the experimental procedure. For example, an arrayed RNAi screen measuring levels of phosphorylated histone H2AX as a marker of DNA damage included several positive and negative control replicates on each plate to measure the daily experimental variability of the assay [67]. A 'dry run' experiment on a small scale and a 'reduced run' experiment (i.e. using a fraction of the library) at the actual scale should give a good indication of the feasibility of the screen and the likelihood of getting good results. This step should also be used to calibrate the assay and standardize experimental variations (for example, by using the same lot of antibody if using an antibody-based method for detection, or the same lot of serum to grow cells). A pilot screen will also help in planning the logistics, such as the time required to run the full screen as well as the equipment and supplies needed. For pooled screens, it is advisable (if possible) to include a positive control diluted at the level of the other shRNAs to verify that it can be recovered from the screen. This test could give an idea of the false negative rate intrinsic to the screen. During the pilot screens, an initial quality assessment of the screen can be made by looking at the candidates identified; some of the candidates should be expected to be recovered, or should fall in similar biological pathways or gene family. Once all the parameters have been optimized, the screen can now proceed.

THE PREMIERE AND THE CRITIQUE (INTERPRETATION AND VALIDATION OF THE SCREEN AND THE CANDIDATES)
The Director's cut
Once the first results are obtained, a number of options are available to identify the valid candidates from the experimental background. In most screens, a range of phenotypes will be obtained due to variations in knockdown efficiency and the inherent phenotypic effect of the gene. There are several excellent reviews on statistical tests and parameters that

can be used to get as much information as possible from the results of a screen [68–73], including free software to analyze high-throughput screen results (CellHTS2 [70]). Although this step might seem trivial at a small scale, it becomes important when more than 100 000 candidates are tested concurrently. It is also suggested to compare the nucleic acid sequences of the different candidates as a recent siRNA screen to identify modulators of TRAIL-induced apoptosis identified a significant number of siRNAs acting through off-target effects due to micro-RNA effects from partial seed complementarity [74]. For siRNA pools (i.e. three to five different siRNAs targeting the same gene in a single well), it is often recommended to deconvolute the pool in order to identify the siRNA(s) responsible for the phenotype.

The re-run (duplicate/triplicate screen)
It is becoming standard procedure to run screens in duplicate or triplicate [47, 75–81]—an aspect that should be taken into account when designing a screen as it can significantly affect the price tag. To obtain reliable hits, it might be better to perform a screen in triplicate using a sublibrary than carry out a full genome screen with no replicate screen. Performing a screen multiple times has several advantages. For example, it can help lower the false-positive rate by identifying those candidates isolated in several screens. It also gives an indication of the 'depth' of the screen, i.e. how likely the major pathways have been found through the screen and that minor pathways have not been missed. In some cases, repeating the screen in an alternative cell line or mouse strain might help to narrow down the number of candidates by focusing on only those candidates that are conserved between the two cell lines/strains. However, it is important to keep in mind that a candidate's inability to reproduce the phenotype in the alternate cell line could be due to trivial factors such as sequence or splice variants of the target mRNA or differences in the abundance of the mRNA, which could affect the efficiency of the depletion. Alternatively, it could suggest that the phenotype under study can be regulated through different genetic networks depending on the cell lines or mouse strains investigated. This explanation was proposed for the little overlap obtained amongst similar screens that were done to identify factors involved in either hepatitis C [82–84] or HIV

[55–57] replication, even though the same RNAi library was used in some cases.

The remake (alternate screen as secondary screen)

Some candidates identified through an RNAi screen might reproducibly generate the phenotype of interest, however this might be due to the type of assay used for the readout or, as discussed below, to an off-target effect. For example, a screen relying on the activation of a reporter gene for the readout might pick up valid candidates but also general transcriptional repressors, which might not be specific to the process under study. To overcome such problems, several groups are now using an alternative, or orthogonal, screening procedure. For example, an RNAi screen aimed at identifying the factors that govern the identity of human embryonic stem cells (hESCs) used the H1 cell line stably expressing a reporter construct composed of the *GFP* gene driven by the *POU5F1* promoter as a primary readout. The candidate siRNAs were retested in a secondary screen measuring the expression levels of the stemness markers *OCT4* and *NANOG* in various hESC cell lines [85]. The secondary screen does not need to be as high throughput as the first round, as the number of candidates should have been reduced dramatically. The initial script should be designed with the secondary screen already in mind. In some cases, the RNAi library is already formatted towards a predicted secondary screen. For example, Bauer *et al.* [86], who performed a screen using a siRNA library against the 'druggable genome' (i.e. a library targeting gene products that can be inhibited with a drug) to identify agents that enhance paclitaxel activity in breast cancer cells, used pharmacological agents to confirm their top scoring hits/genes. Candidates arising from such approach represent interesting pharmaceutical targets for clinical applications. Although some potentially good candidates might be weeded out at this step, the primary screen can later be revisited with a different type of secondary screen to potentially pick up candidates left behind by the first secondary screen. Some researchers will elect to carry out their duplicate/triplicate screen at this step instead of at the primary screen step, depending on the number of candidates obtained in the first screen as well as the costs associated with a re-run of the primary versus secondary screen.

Remix of the greatest hits

Once the list of candidates has been narrowed down, each candidate needs to be validated by redoing the assay with the same si/sh/miRNA identified in the screen to make sure it recapitulates the result. A second (and sometimes third) si/sh/miRNA targeting the same gene is also required to rule out off-target effects (see below). The second RNAi reagent could be from a different source; for example, an esiRNA could be used to confirm a hit that was obtained using a shRNA, as long as the sequence used to generate the esiRNAs does not overlap with that of the shRNA. However, it is important to keep in mind that different si/shRNAs targeting the same gene can give rise to different phenotypes; for example, they might either knockdown the target gene with different efficiencies thus leading to a different cellular response, they might target different splice isoforms of the mRNA, or the corresponding mRNA target sequence might harbor a polymorphism that affects its recognition by the RISC complex. Phenotypic differences can also be induced by an off-target effect associated with one of the reagents tested. Finally, another approach to validate candidate hits is to rescue the phenotype by expressing an RNAi-resistant version of the cDNA in the presence of the candidate si/shRNA. Zhou *et al.* [56] opted for a variant of this option in their screen for factors required for HIV replication by transfecting siRNAs targeting the 3′-UTR of their candidate gene while overexpressing the corresponding cDNA that lacked the 3′-UTR. A similar approach was also taken by Tai *et al.* [87] to confirm the requirement of the phosphatidylinositol 4-kinase PI4KA in the replication of the hepatitis virus. However, complications from detrimental effects associated with overexpression of the cDNA could arise in some cases. Neumann *et al.* [58] circumvented that potential pitfall by complementing their human cell line with a BAC expressing the mouse ortholog of their candidate gene under the control of its endogenous promoter. Due to the natural sequence divergence between human and mouse, most of the siRNAs targeting human genes should not target the orthologous mouse gene.

The off-screen (off-target and saturation) effects

Two specific problems associated with RNAi screens can be encountered. The first is when the si/shRNA recognizes another unintended mRNA target and

impedes its translation, either through cleavage of the mRNA or translation inhibition, the so-called off-target effect. The second problem arises when the RNAi machinery is oversaturated with high levels of exogenous si/sh/miRNA which compete with endogenous sh/miRNAs for the available nuclear export complex exportin-5 or RISC complex and prevents them from exerting their normal function. This problem has been reported both *in vitro* [88, 89] and *in vivo* [90] and can result in upregulation of genes normally regulated by endogenous miRNAs, causing toxic effects and even lethality in mice [90]. In both cases, the results obtained from the screen would identify genes that have no relevance to the original biological question. Off-target effects can often be ruled-out by targeting the same gene with a second independent shRNA/siRNA, as described above. An alternate si/shRNA might also solve the saturation problem in some cases, if they happen not to overload the machinery, but in most cases, a complementary secondary screen or complementation with an RNAi resistant cDNA as described above would be the best procedure. A recent example to circumvent the off-target effect was developed by Collinet and colleagues [91] which used three different genome-wide libraries (two siRNA and one esiRNA), at an average of seven to eight siRNAs/esiRNAs per gene, for a quantitative multiparametric image analysis (QMPIA) screen to identify factors involved in endocytosis. By monitoring 46 biological parameters of endocytosis for each esi/siRNA, they were able to obtain an average gene profile, thereby greatly reducing off-target effects of single esi/siRNAs.

The cast away candidates

For some RNAi libraries, not all the candidates obtained during a screen will map to a known gene. For example, shRNA libraries from Open Biosystems were built against all the transcribed sequences, regardless if they were annotated to a validated gene. With the recent discovery of several classes of non-coding RNAs such as piRNAs, lincRNAs and possibly other types yet to be uncovered, it is quite possible that the cast away candidates might truly be involved in the phenotype studied in the screen. It could be interesting to retest and confirm the role of these candidates in the phenotype studied, which could bring insight onto new gene regulation pathways.

The no-show star (the missing known gene)

Why do certain genes known to be involved in the phenotype being investigated end up missing from the list of candidates generated by a screen? None withstanding various biological implications such as redundant pathways, one possible technical reason could be that the design of the screen is not compatible with the detection of this type of candidates. It should also be noted that screens are almost never saturating, which means it is unlikely that all the candidates that *can* be isolated in fact *will* be detected. In addition, using the current RNAi technologies, some genes might not be knocked down with sufficient efficiency to generate a phenotype. Finally, in the case of pooled screens, it is possible that certain shRNAs might be overrepresented or lost from the population, due to variations occurring during the plasmid and/or the virus preparation steps.

The IMDb (Internet 'multi'-database) searches

With the list of validated candidate genes in hand, one can search the various available databases (public and/or private) for an independent confirmation of the validity of the candidates or to elucidate their mechanism(s) of action. In this regard, a vast amount of high-throughput data is now available online, including data from microarrays, proteomic and interactome studies, genetic networks, and small molecule and RNAi screens, as well as through various online computational resources such as gene ontology [92]. For example, the gene *Gas1*, identified through an RNAi screen aimed at identifying metastasis suppressor genes, was found to be downregulated in patients with metastatic cancer [51] using the database Oncomine [93], which is a compendium of cancer transcriptome profiles with analysis engine.

EPILOG

The advent of RNAi technologies granted researchers a powerful tool to study the consequences of depleting a gene in a wide variety of living organisms. The availability of genome-wide RNAi libraries opened the door to captivating functional genomics studies and, as demonstrated by the ever-increasing number of publications, the possibilities of RNAi screens are almost limitless. However, RNAi libraries are not perfect tools, and one major issue associated with RNAi screening is the off-target

effect. Another concern is the fact that not all of the si/sh/miRNA reagents have been validated to generate a significant knockdown, although some companies are in the process of generating validated libraries. It is thus crucial to continue improving the current RNAi libraries and generate new ones showing less off-target effects and giving a better genome-wide coverage. These improvements will undoubtedly increase the quality of the data retrieved from RNAi screens. Furthermore, the increasing availability of reagents from various laboratories and companies will allow RNAi screening to be performed at more affordable prices. As a result, RNAi screenings will soon become one of the standard tools scientists can use to answer biological questions, alongside cDNA library screenings or microarray studies. For the latter, a consortium of scientists has suggested the use of 'Minimal Information About a Microarray Experiment' (MIAME, [94]) standards to allow other scientists to interpret the results of a microarray experiment. Similarly, a 'Minimal Information About an RNAi Experiment' (MIARE) set of reporting guidelines has also been suggested [95], with the aim of helping researchers evaluate the quality of published screens. In order to perform a successful genome-wide or targeted RNAi screen, several considerations presented in this review need to be taken into account, and with the proper choice of actors, producers and a good script, RNAi can reveal its full potential and might be the first chapter of a successful series.

Key Points

- RNAi screens can be accomplished via siRNAs, esiRNAs, shRNAs or miRNAs under an arrayed or pooled format.
- Positive (enrichment) or negative (depletion) RNAi screens can be performed using direct visualization or assays to select for the phenotype of interest.
- The design of the screen has to be thoroughly planned and tested; performing a screen multiple times can help reducing the rate of false positives.
- Careful validation and analysis of the results is imperative to reveal the true candidates.
- Choosing the appropriate RNAi reagent, type of screen and model system is critical for a successful RNAi screen.

References

1. Mullenders J, Bernards R. Loss-of-function genetic screens as a tool to improve the diagnosis and treatment of cancer. *Oncogene* 2009;**28**:4409–20.

2. Hirsch AJ. The use of RNAi-based screens to identify host proteins involved in viral replication. *Future Microbiol* 2010;**5**:303–11.

3. Wolters NM, MacKeigan JP. From sequence to function: using RNAi to elucidate mechanisms of human disease. *Cell Death Differ* 2008;**15**:809–19.

4. Mohr S, Bakal C, Perrimon N. Genomic screening with RNAi: results and challenges. *Annu Rev Biochem* 2010;**79**:37–64.

5. Mattila J, Puig O. Insights to transcriptional networks by using high throughput RNAi strategies. *Brief Funct Genomics* 2010;**9**:43–52.

6. Boutros M, Ahringer J. The art and design of genetic screens: RNA interference. *Nat Rev Genet* 2008;**9**:554–66.

7. Lord CJ, Martin SA, Ashworth A. RNA interference screening demystified. *J Clin Pathol* 2009;**62**:195–200.

8. Zhang XD, Espeseth AS, Johnson EN, *et al.* Integrating experimental and analytic approaches to improve data quality in genome-wide RNAi screens. *J Biomol Screen* 2008;**13**:378–89.

9. Falschlehner C, Steinbrink S, Erdmann G, *et al.* High-throughput RNAi screening to dissect cellular pathways: a how-to guide. *Biotechnol J* 2010;**5**:368–76.

10. Sharma S, Rao A. RNAi screening: tips and techniques. *Nat Immunol* 2009;**10**:799–804.

11. Elbashir SM, Harborth J, Lendeckel W, *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 2001;**411**:494–8.

12. Yang D, Buchholz F, Huang Z, *et al.* Short RNA duplexes produced by hydrolysis with Escherichia coli RNase III mediate effective RNA interference in mammalian cells. *Proc Natl Acad Sci USA* 2002;**99**:9942–7.

13. Brummelkamp TR, Bernards R, Agami R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* 2002;**296**:550–3.

14. McManus MT, Petersen CP, Haines BB, *et al.* Gene silencing using micro-RNA designed hairpins. *RNA* 2002;**8**:842–50.

15. Paddison PJ, Caudy AA, Bernstein E, *et al.* Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 2002;**16**:948–58.

16. Silva JM, Li MZ, Chang K, *et al.* Second-generation shRNA libraries covering the mouse and human genomes. *Nat Genet* 2005;**37**:1281–8.

17. Kittler R, Buchholz F. Functional genomic analysis of cell division by endoribonuclease-prepared siRNAs. *Cell Cycle* 2005;**4**:564–7.

18. Gopalakrishnan B, Wolff J. siRNA and DNA transfer to cultured cells. *Methods Mol Biol* 2009;**480**:31–52.

19. Fabrizio dAdf, Marzia F, Raffaella DM. Oligofection of small interfering RNAs in senescent human fibroblasts. *Nat Protoc (Protoc Exch)* 2006. doi:10.1038/nprot.2006.365.

20. Theis M, Buchholz F. High-throughput RNAi screening in mammalian cells with esiRNAs. *Methods* 2011;**53**:424–9.

21. Gaglione M, Messere A. Recent progress in chemically modified siRNAs. *Mini Rev Med Chem* 2010;**10**:578–95.

22. Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet* 2010;**11**:597–610.

23. Zhang C. Novel functions for small RNA molecules. *Curr Opin Mol Ther* 2009;**11**:641–51.

24. Taft RJ, Pang KC, Mercer TR, *et al*. Non-coding RNAs: regulators of disease. *J Pathol* 2010;**220**:126–39.

25. Moffat J, Grueneberg DA, Yang X, *et al*. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* 2006;**124**:1283–98.

26. Gazin C, Wajapeyee N, Gobeil S, *et al*. An elaborate pathway required for Ras-mediated epigenetic silencing. *Nature* 2007;**449**:1073–7.

27. Smolen GA, Zhang J, Zubrowski MJ, *et al*. A genome-wide RNAi screen identifies multiple RSK-dependent regulators of cell migration. *Genes Dev* 2010;**24**:2654–65.

28. Sheng Z, Li L, Zhu LJ, *et al*. A genome-wide RNA interference screen reveals an essential CREB3L2-ATF5-MCL1 survival pathway in malignant glioma with therapeutic implications. *Nat Med* 2010;**16**:671–7.

29. Mullenders J, Fabius AW, Madiredjo M, *et al*. A large scale shRNA barcode screen identifies the circadian clock component ARNTL as putative regulator of the p53 tumor suppressor pathway. *PLoS One* 2009;**4**:e4798.

30. Coussens M, Davy P, Brown L, *et al*. RNAi screen for telomerase reverse transcriptase transcriptional regulators identifies HIF1alpha as critical for telomerase function in murine embryonic stem cells. *Proc Natl Acad Sci USA* 2010;**107**:13842–7.

31. Popov N, Wanzel M, Madiredjo M, *et al*. The ubiquitin-specific protease USP28 is required for MYC stability. *Nat Cell Biol* 2007;**9**:765–74.

32. Gumireddy K, Li A, Gimotty PA, *et al*. KLF17 is a negative regulator of epithelial-mesenchymal transition and metastasis in breast cancer. *Nat Cell Biol* 2009;**11**:1297–304.

33. Bric A, Miething C, Bialucha CU, *et al*. Functional identification of tumor-suppressor genes through an in vivo RNA interference screen in a mouse lymphoma model. *Cancer Cell* 2009;**16**:324–35.

34. Meacham CE, Ho EE, Dubrovsky E, *et al*. In vivo RNAi screening identifies regulators of actin dynamics as key determinants of lymphoma progression. *Nat Genet* 2009;**41**:1133–18.

35. Zender L, Xue W, Zuber J, *et al*. An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer. *Cell* 2008;**135**:852–64.

36. Boettcher M, Hoheisel JD. Pooled RNAi Screens - technical and biological aspects. *Curr Genomics* 2010;**11**:162–7.

37. Boettcher M, Fredebohm J, Moghaddas Gholami A, *et al*. Decoding pooled RNAi screens by means of barcode tiling arrays. *BMC Genomics* 2010;**11**:7.

38. Ngo VN, Davis RE, Lamy L, *et al*. A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* 2006;**441**:106–10.

39. Schlabach MR, Luo J, Solimini NL, *et al*. Cancer proliferation gene discovery through functional genomics. *Science* 2008;**319**:620–4.

40. Hartwell LH, Szankasi P, Roberts CJ, *et al*. Integrating genetic approaches into the discovery of anticancer drugs. *Science* 1997;**278**:1064–8.

41. Luo J, Emanuele MJ, Li D, *et al*. A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell* 2009;**137**:835–48.

42. Scholl C, Frohling S, Dunn IF, *et al*. Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells. *Cell* 2009;**137**:821–34.

43. Lam LT, Davis RE, Ngo VN, *et al*. Compensatory IKKalpha activation of classical NF-kappaB signaling during IKKbeta inhibition identified by an RNA interference sensitization screen. *Proc Natl Acad Sci USA* 2008;**105**:20798–803.

44. Wiltshire TD, Lovejoy CA, Wang T, *et al*. Sensitivity to poly(ADP-ribose) polymerase (PARP) inhibition identifies ubiquitin-specific peptidase 11 (USP11) as a regulator of DNA double-strand break repair. *J Biol Chem* 2010;**285**:14565–71.

45. Azorsa DO, Gonzales IM, Basu GD, *et al*. Synthetic lethal RNAi screening identifies sensitizing targets for gemcitabine therapy in pancreatic cancer. *J Transl Med* 2009;**7**:43.

46. Whitehurst AW, Bodemann BO, Cardenas J, *et al*. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature* 2007;**446**:815–9.

47. O'Connell BC, Adamson B, Lydeard JR, *et al*. A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. *Mol Cell* 2010;**40**:645–57.

48. Bukrinsky MI, Haggerty S, Dempsey MP, *et al*. A nuclear localization signal within HIV-1 matrix protein that governs infection of non-dividing cells. *Nature* 1993;**365**:666–9.

49. Lewis P, Hensel M, Emerman M. Human immunodeficiency virus infection of cells arrested in the cell cycle. *EMBO J* 1992;**11**:3053–8.

50. Wajapeyee N, Serra RW, Zhu X, *et al*. Oncogenic BRAF induces senescence and apoptosis through pathways mediated by the secreted protein IGFBP7. *Cell* 2008;**132**:363–74.

51. Gobeil S, Zhu X, Doillon CJ, *et al*. A genome-wide shRNA screen identifies GAS1 as a novel melanoma metastasis suppressor gene. *Genes Dev* 2008;**22**:2932–40.

52. Fazzio TG, Huff JT, Panning B. An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* 2008;**134**:162–74.

53. Palakurthy RK, Wajapeyee N, Santra MK, *et al*. Epigenetic silencing of the RASSF1A tumor suppressor gene through HOXB3-mediated induction of DNMT3B expression. *Mol Cell* 2009;**36**:219–30.

54. Goff SP. Knockdown screens to knockout HIV-1. *Cell* 2008;**135**:417–20.

55. Brass AL, Dykxhoorn DM, Benita Y, *et al*. Identification of host proteins required for HIV infection through a functional genomic screen. *Science* 2008;**319**:921–6.

56. Zhou H, Xu M, Huang Q, *et al*. Genome-scale RNAi screen for host factors required for HIV replication. *Cell Host Microbe* 2008;**4**:495–504.

57. Konig R, Zhou Y, Elleder D, *et al*. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* 2008;**135**:49–60.

58. Neumann B, Walter T, Heriche JK, *et al*. Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature* 2010;**464**:721–7.

59. Carpenter AE, Jones TR, Lamprecht MR, *et al*. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* 2006;**7**:R100.

60. Lamprecht MR, Sabatini DM, Carpenter AE. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques* 2007;**42**:71–5.

61. Jones TR, Kang IH, Wheeler DB, *et al*. CellProfiler analyst: data exploration and analysis software for complex image-based screens. *BMC Bioinformatics* 2008;**9**:482.

62. Held M, Schmitz MH, Fischer B, *et al*. CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nat Methods* 2010;**7**:747–54.

63. Wild T, Horvath P, Wyler E, *et al*. A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60 S subunit export. *PLoS Biol* 2010;**8**: e1000522.

64. Edelstein A, Amodaj N, Hoover K, *et al*. Computer control of microscopes using microManager. *Curr Protoc Mol Biol* 2010; Chapter 14:Unit14 20.

65. Conrad C, Wunsche A, Tan TH, *et al*. Micropilot: automation of fluorescence microscopy-based imaging for systems biology. *Nat Methods* 2011;**8**:246–9.

66. Pau G, Fuchs F, Sklyar O, *et al*. EBImage–an R package for image processing with applications to cellular phenotypes. *Bioinformatics* 2010;**26**:979–81.

67. Paulsen RD, Soni DV, Wollman R, *et al*. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol Cell* 2009;**35**:228–39.

68. Birmingham A, Selfors LM, Forster T, *et al*. Statistical methods for analysis of high-throughput RNA interference screens. *Nat Methods* 2009;**6**:569–75.

69. Lee S, Howell BJ. High-content screening: emerging hardware and software technologies. *Methods Enzymol* 2006;**414**: 468–83.

70. Pelz O, Gilsdorf M, Boutros M. web cellHTS2: a web-application for the analysis of high-throughput screening data. *BMC Bioinformatics* 2010;**11**:185.

71. Gilbert DF, Meinhof T, Pepperkok R, *et al*. DetecTiff: a novel image analysis routine for high-content screening microscopy. *J Biomol Screen* 2009;**14**:944–55.

72. Kaushal D, Naeve CW. Analyzing and visualizing expression data with Spotfire. *Curr Protoc Bioinformatics* 2004; Chapter 7:Unit 7 9.

73. Boutros M, Bras LP, Huber W. Analysis of cell-based RNAi screens. *Genome Biol* 2006;**7**:R66.

74. Sudbery I, Enright AJ, Fraser AG, *et al*. Systematic analysis of off-target effects in an RNAi screen reveals microRNAs affecting sensitivity to TRAIL-induced apoptosis. *BMC Genomics* 2010;**11**:175.

75. Cole KA, Huggins J, Laquaglia M, *et al*. RNAi screen of the protein kinome identifies checkpoint kinase 1 (CHK1) as a therapeutic target in neuroblastoma. *Proc Natl Acad Sci USA* 2011;**108**:3336–41.

76. Holzel M, Huang S, Koster J, *et al*. NF1 is a tumor suppressor in neuroblastoma that determines retinoic acid response and disease outcome. *Cell* 2010;**142**:218–29.

77. Hurov KE, Cotta-Ramusino C, Elledge SJ. A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability. *Genes Dev* 2010;**24**:1939–50.

78. Schmitz MH, Held M, Janssens V, *et al*. Live-cell imaging RNAi screen identifies PP2A-B55alpha and importin-beta1 as key mitotic exit regulators in human cells. *Nat Cell Biol* 2010;**12**:886–93.

79. Smogorzewska A, Desetty R, Saito TT, *et al*. A genetic screen identifies FAN1, a Fanconi anemia-associated nuclease necessary for DNA interstrand crosslink repair. *Mol Cell* 2010;**39**:36–47.

80. Yoon JC, Ng A, Kim BH, *et al*. Wnt signaling regulates mitochondrial physiology and insulin sensitivity. *Genes Dev* 2010;**24**:1507–18.

81. Zhu YX, Tiedemann R, Shi CX, *et al*. RNAi screen of the druggable genome identifies modulators of proteasome inhibitor sensitivity in myeloma including CDK5. *Blood* 2011; **117**:3847–57.

82. 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**:1413–21.

83. Randall G, Panis M, Cooper JD, *et al*. Cellular cofactors affecting hepatitis C virus infection and replication. *Proc Natl Acad Sci USA* 2007;**104**:12884–9.

84. 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**: 29–36.

85. Chia NY, Chan YS, Feng B, *et al*. A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* 2010;**468**:316–20.

86. Bauer JA, Ye F, Marshall CB, *et al*. RNA interference (RNAi) screening approach identifies agents that enhance paclitaxel activity in breast cancer cells. *Breast Cancer Res* 2010;**12**:R41.

87. 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**:298–307.

88. Castanotto D, Sakurai K, Lingeman R, *et al*. Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. *Nucleic Acids Res* 2007;**35**:5154–64.

89. Khan AA, Betel D, Miller ML, *et al*. Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. *Nat Biotechnol* 2009;**27**:549–55.

90. Grimm D, Streetz KL, Jopling CL, *et al*. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* 2006;**441**:537–41.

91. Collinet C, Stoter M, Bradshaw CR, *et al*. Systems survey of endocytosis by multiparametric image analysis. *Nature* 2010; **464**:243–9.

92. Consortium GO. Creating the gene ontology resource: design and implementation. *Genome Res* 2001;**11**:1425–33.

93. Rhodes DR, Yu J, Shanker K, *et al*. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia* 2004;**6**:1–6.

94. Brazma A, Hingamp P, Quackenbush J, *et al*. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 2001;**29**: 365–71.

95. Haney SA. Increasing the robustness and validity of RNAi screens. *Pharmacogenomics* 2007;**8**:1037–49.