High-throughput RNAi screening in cultured cells: a user's guide

Christophe J. Echeverri and Norbert Perrimon‡*

Abstract | RNA interference has re-energized the field of functional genomics by enabling genome-scale loss-of-function screens in cultured cells. Looking back on the lessons that have been learned from the first wave of technology developments and applications in this exciting field, we provide both a user's guide for newcomers to the field and a detailed examination of some more complex issues, particularly concerning optimization and quality control, for more advanced users. From a discussion of cell lines, screening paradigms, reagent types and read-out methodologies, we explore in particular the complexities of designing optimal controls and normalization strategies for these challenging but extremely powerful studies.

Ribozyme
An RNA molecule with catalytic activity.

RNAi
RNA interference refers to the process by which dsRNA molecules silence a target gene through the specific destruction of its mRNA.

dsRNA
Long dsRNAs (usually referring, in this context, to those that are >200 bp in length) that are made from cDNA or genomic DNA templates.

**Cenix BioScience GmbH, Tatzberg 47, Dresden 01307, Germany.*
‡Department of Genetics, Howard Hughes Medical Institute, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA.
e-mails: echeverri@cenix-bioscience.com; perrimon@receptor.med harvard.edu
doi:10.1038/nrg1836
Published online
11 April 2006

Tissue culture cells have provided a powerful system for studying many fundamental problems in signal transduction, cell differentiation and physiology. However, functional studies in cultured cells were hampered in the past by the lack of a powerful method for perturbing gene activities. Several technologies designed to knock down gene function, such as those based on ribozymes and antisense approaches, showed initial promise but ultimately failed to deliver robust protocols.

A turning point came with the discovery of RNAi (REF. 1) and its rapid rise from small-scale experimentation to genome-scale screening in *Caenorhabditis elegans* using dsRNAs[2,3]. Hopes were raised that this method might also be applicable in mammalian cells, providing a direct causal link between gene sequence and functional data in the form of targeted loss-of-function (LOF) phenotypes. The use of long dsRNAs to trigger RNAi was initially hindered in mammals by the fact that these molecules simultaneously activate the interferon response[4]; however, it quickly proved successful in cultured *Drosophila melanogaster* cells[5]. Subsequently, short dsRNAs designed to mimic small interfering RNAs (siRNAs), which were initially identified in plants[6], were shown to elicit a potent and specific RNAi response in cultured human cells, without interferon activation[7]. Several strategies have now been devised to trigger the RNAi pathway, each of which is adapted and optimized for different cell systems. Today, the most commonly used approaches are based on long dsRNA for *D. melanogaster* cells, and either synthetic siRNAs or vector-expressed short hairpin RNAs (shRNAs) for mammalian cells. The fast development of these RNAi tools has been driven by advances in the molecular understanding of the RNAi pathway (BOX 1).

RNAi has accelerated a wide range of small-scale gene characterization studies, but arguably the most important way in which it has transformed biological research is by enabling genome-scale screens in cell culture systems. Driven by genome sequence data, RNAi is now widely used in high-throughput (HT) screens in both basic and applied biology[8]. It is a powerful method for addressing many questions in cell biology, and its amenability for use in modifier screens in addition to direct LOF screening has made it particularly useful for the analysis of signal transduction pathways (BOX 2). RNAi has also become a method of choice for key steps in the development of therapeutic agents, from target discovery and validation to the analysis of the mechanisms of action of small molecules[9]. Although several HT screens have already been carried out in both *D. melanogaster* and mammalian cells[10–32] this is still an area of huge opportunity, especially as new technical advances arise.

Here we provide a guide to carrying out HT RNAi screens in cell systems, focusing on *D. melanogaster* and mammalian cells — the systems in which such HT screens are mainly carried out. Most HT RNAi screens are complex and expensive undertakings, requiring significant automation and computing infrastructures, and a combination of disparate skills, ranging from informatics to cell-culture expertise and HT assay development (BOX 3). In addition to these infrastructure requirements, designing a cell-based HT RNAi screen

Box 1 | **RNAi biology**

The experimental use of RNAi represents the harnessing of endogenous cellular pathways that are present in species ranging from plants to humans. These pathways use two types of small RNA — siRNAs and miRNAs — to direct the sequence-specific downregulation of endogenous or exogenous target genes. In *Drosophila melanogaster* and *Caenorhabditis elegans*, long dsRNAs of a few hundred base pairs are commonly used in RNAi experiments, and silencing is ultimately induced by siRNAs, the key pathway intermediate. In mammalian cells, shorter dsRNAs that closely mimic siRNAs are commonly used to elicit an RNAi response without triggering the interferon pathway, sometimes through a short hairpin (shRNA) construct. As understanding of the miRNA pathway deepens, some efforts have also sought to make further RNAi reagent design improvements, either by directly mimicking miRNA biogenesis, or by learning from their targeting principles.

The siRNA pathway

Long dsRNAs and shRNAs, either ectopically introduced into cells or endogenously generated, are processed by Dicer, a dsRNA-specific RNase III, to form siRNAs. These siRNAs, which are actively maintained in the cytosol by exportin[60], are then loaded into argonaute 2-containing RNA-induced silencing complexes (RISCs). This process imposes a selection, which is based on the relative thermodynamic lability of the two ends of the siRNA, whereby one siRNA strand becomes the 'guide', or targeting co-factor, and the other becomes a temporary 'passenger', which is quickly degraded as a pseudotarget. The guide strand is then used by RISC to direct repeated rounds of target mRNA recognition, cleavage and release, in a powerful processive cycle. A search for clear 'rules' that define target mRNA recognition by the guide strand, which are important for optimizing the specificity of silencing reagents, has proved difficult. Most focus is on the so-called 'seed region' of bases 2–8, which is defined as the primary targeting region for miRNA action and is the region that is least tolerant of mismatches. Nonetheless, siRNA targeting specificity remains incompletely understood.

The miRNA pathway

miRNAs are initially produced as long transcripts (pri-miRNAs) that include hairpin structures and contain one or more miRNAs. Pri-miRNAs are processed in the nucleus by a microprocessor complex that contains the RNase III endonuclease Drosha and an RNA-binding protein Pasha or DGCR8 (DiGeorge syndrome critical region gene 8) (REFS 61–64), which produces 60–70 nt stem-loop intermediates (pre-miRNA). Pre-miRNAs are exported from the nucleus in a process that is dependent on exportin 5 and RAN[65,66], and are processed in the cytoplasm by a complex that contains the enzyme Dicer and RNA-binding protein loquacious or TRBP (TAR RNA binding protein)[67–69], producing an imperfect RNA duplex of the miRNA, the future 'guide strand' and its complement, the so-called 'miRNA*' strand. The miRNA strand is preferentially loaded into the RISC complex, whereas the miRNA* strand is degraded. The miRNA containing RISC complex then associates with target mRNA, leading to cleavage or to translational repression[70,71].

involves many levels of decision-making, including the choice of species and cell line, screening paradigm and format, reagent type and read-out methodology used in phenotypic assays. We discuss each of these considerations, and provide an overview of the necessary controls and optimization procedures for the successful implementation of a cell-based HT RNAi screen.

Choice of cell type

***Drosophila melanogaster* cells.** With a relatively modest but fast-growing list of available cell types, *D. melanogaster* cells are excellent for RNAi screens. They typically grow at or near room temperature under ambient CO_2 levels[33] and several *D. melanogaster* cell lines efficiently take up dsRNA from the medium without the need for transfection reagents[5]. In addition, *D. melanogaster* cells, like mammalian cells, allow high-resolution spatio-temporal observations to be made by microscopy[10].

S2 and Kc cells are the most commonly used lines for *D. melanogaster* RNAi screens (TABLE 1), and both take up dsRNA efficiently by bathing cells in a serum-free medium (for detailed protocols see REFS 5,34). Another popular cell line, clone 8, shows poor uptake with the bathing method, but has been successfully implemented in RNAi screens that use standard lipid-based dsRNA transfection methods[14]. Many other *D. melanogaster* cell lines of various origins[35] can be used for RNAi applications, and are available from the *Drosophila* Genome Resource Center.

RNAi can also be carried out effectively in primary cells that are isolated from *D. melanogaster* embryos. This approach can provide advantages over cell lines, as the differentiation programmes of primary cells follow *in vivo* differentiation patterns more closely. For example, screens for axonal outgrowth and muscle integrity have been completed by simply deriving cells from embryos that express a GFP marker in the cells of interest (K. Sepp, J. Bai and N.P., unpublished observations), and the primary cells tested so far elicit a robust RNAi response after bathing with dsRNAs.

Mammalian cells. The vast compendium of publicly available human and rodent cell lines offers a wide range of genotypes, cellular characteristics and tissue derivations, and therefore provides a broad potential for accurately modelling many biological processes. Although RNAi silencing reagents are available for targeting virtually any human, mouse or rat gene, most mammalian cell-based RNAi studies so far have used human cells of various origins. Adherent lines such as HeLa and U2OS offer easy, efficient delivery and fast, robust growth in the well-ordered monolayers that are most desirable for microscopy read-outs. For many of these, the transient transfection of synthetic RNAi-based silencing reagents (for example, siRNAs) has proved highly efficient (>95%) using standard lipid-based transfection reagents, although often not without significant optimization (see the later section on this topic). In such experiments, the doubling time of the cell line directly affects the duration of silencing, which usually does not exceed 5–6 days for most lines[36].

Importantly, using the transfection and culture conditions required for adequate silencing efficiency sometimes comes at the price of increased toxicity or other significant alterations to cell physiology, such that the processes under study might no longer be well represented. The robustness of cell lines varies widely in this respect: certain commonly used lines, such as HeLa, have higher tolerance to conditions that will prove markedly toxic to many others (such as the MCF-7 line) (REF. 37). This highlights the importance of careful optimization of RNAi conditions for each individual cell line, not just for maximal silencing but also to achieve optimal silencing in healthy cells. Beyond these toxicity issues, many cell lines also show genetic instability, which leads to loss of clonality and intra-line heterogeneities in karyotype and physiology. Although too often overlooked, these factors can underlie significant variability in HT RNAi screening results, and might warrant subcloning of the cell line.

Interferon response
A primitive antiviral mechanism that triggers sequence-nonspecific degradation of mRNA and downregulation of cellular protein synthesis.

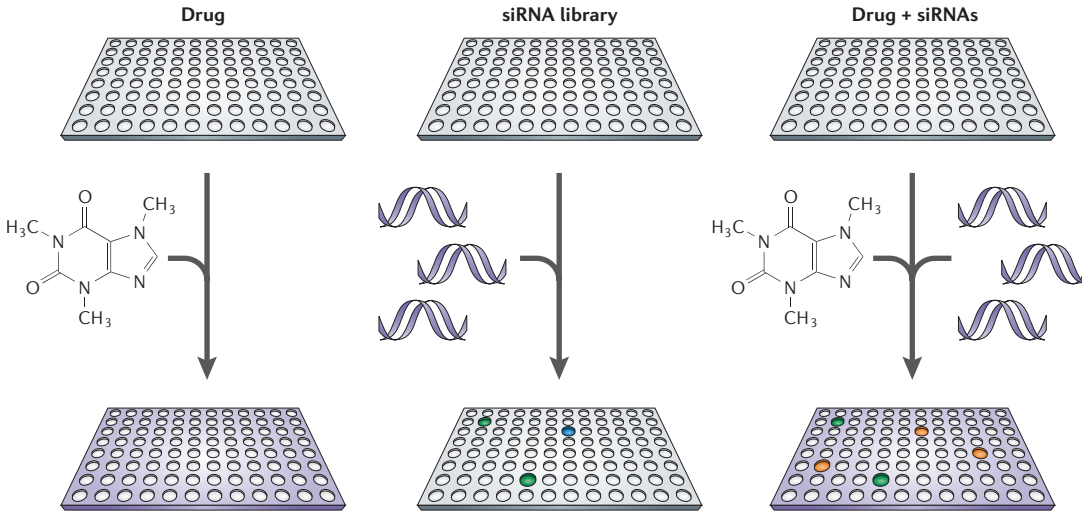
Small interfering RNA
Small RNAs of 21–23 nucleotides in length that engage the complementary mRNA into the RISC complex for degradation.

Short hairpin RNA
Small dsRNA constructs that are usually 22–29 nucleotides long and form a hairpin-like secondary structure.

The accurate modelling of certain biological processes, such as immunological and neurological pathways, remains contentious in transformed cell lines, leading many researchers to preferentially study these in primary cells. With few exceptions (for example, HUVEC cells), primary cells have presented serious obstacles to HT RNAi screening applications, primarily because most are refractory to standard lipid- or peptide-based transfection methods[38]. In some cases, such as B cells and CD34+ haematopoietic progenitors, advanced electroporation-based methods have yielded effective protocols for small-scale work, but these are not yet fully optimized for HT RNAi screening. Most other primary cells have only been accessible to HT RNAi through the use of virally delivered shRNA vectors (see below). This approach has yielded significant successes[24,25,27] despite the generally sub-optimal level of silencing that was observed with early shRNA libraries and the risk that certain viruses might alter key aspects of cellular physiology.

Another important factor to take into account for primary cell screens is the need for a constant supply of biologically homogeneous cells to support a large study

Box 2 | **Direct loss-of-function versus modifier screens**



Loss-of-function screens

The most obvious application of RNAi screening, direct loss-of-function (LOF) screening, involves identifying and functionally characterizing genes of interest on the basis of their LOF phenotypes. Such studies offer the broadest discovery potential, as they simply analyse single-gene LOF phenotypes in otherwise untreated cells. This approach has proved effective for many types of gene, including those that encode structural components, cell-surface receptors, transcription factors and enzymes. It is nonetheless important to remember that RNAi is a method for gene knock down and not knock out. Therefore, the high activity and/or long protein half-life and/or high endogenous expression of some genes might make it difficult to generate detectable LOF phenotypes, especially in the case of certain enzymes, as residual activity might be sufficient to fulfil their cellular roles.

Modifier and synthetic lethal screens

RNAi screens can also be refined through many of the same screening strategies that have been developed and perfected for decades in classic genetic screens. Particularly powerful are modifier screens, whereby RNAi is used to identify genes and pathways that, when silenced, can either enhance or suppress a given phenotype of interest. The phenotype to be modified can be the result of an initial drug treatment (see figure; change in array colour indicates the phenotypic effect of the drug that is to be modified; wells of different colours indicate the effects of siRNAs alone (middle panel) or the combination of drug treatment plus siRNAs (right panel)), in which case the screen will potentially yield insights into both the mechanism of drug action and the drug-targeted molecular pathway(s). The initial phenotype can also be generated by an initial genetic modification, through gene overexpression or even RNAi-mediated pre-silencing. In this case, the screen can potentially shed light on cellular pathways that are relevant to the function of this gene. This principle was recently applied in the context of *in vitro* neoplastic transformation assays to identify novel tumour suppressors[26,27].

In the broader context of developing novel therapeutic agents, these methods are of particular value not only for analysing a compound's mechanism of action and understanding unwanted side-effects, but also for identifying potential gene targets for developing sensitizing agents for existing drugs[72]. By focusing on silencing events that suppress the drug's action, the same approach can also identify and/or validate novel biomarkers to predict non-responsiveness to the compound, an increasingly important tool for optimizing the design of clinical trials. Among the most compelling examples of this approach are synthetic lethal screens, whereby lethal combinations of multiple non-lethal modifications are sought. Here RNAi screening is conducted in cells that are pre-treated to duplicate or mimic naturally occurring genetic lesions that are known to underlie disease states such as cancers. In such studies, the desired RNAi-modified phenotype is cell death, thereby offering a way of specifically killing cancerous cells while preserving healthy ones.

miRNA
Endogenously expressed small dsRNA (21–24 nucleotides), which can either interfere with translation of partially complementary mRNAs (usually through their 3′ end UTRs) or cause small interfering RNA-like degradation of perfectly complementary mRNAs.

Dicer
Refers to members of a highly conserved family of RNase III endonucleases that mediate dsRNA cleavage. This produces the small interfering RNAs or mature miRNAs that direct target silencing in RNAi and miRNA pathways, respectively.

Off-target effect
Any detectible phenotypic change that is triggered by RNAi treatment, other than those that are derived directly or indirectly from silencing the targeted mRNA.

over many weeks. Typical precautions include maintaining a reserve of primary cell lots from pooled donors, and their extensive pre-testing for lot-to-lot variability in both silencing and functional read-out assays before launching a screen.

Reagents for *D. melanogaster* HT RNAi screens

Various libraries, all based on long dsRNAs, are available for RNAi screens in *D. melanogaster* cells (TABLE 2). These have been used in several screens that utilize transcriptional reporter assays or microscopy-based read-outs[10–23]. For *D. melanogaster* cells that do not respond well to the bathing method, standard transfection reagents need to be added[14,34].

Box 3 | **Basic infrastructure needed for high-throughput RNAi screens**

Laboratory automation

Beyond the usual tissue culture facilities, a minimal infrastructure is required for semi-automated cell-based RNAi screening. An arrayer robot (for example, TekBench from TekCel) is required to dispense the reagents to be screened into assay plates (96 or 384 wells), after which a liquid dispenser is needed to add the cells, as well as other reagents such as the culture medium, to the plates (for example, WellMate from Matrix). A plate reader and an inverted fluorescent microscope with automated software are required for data acquisition. Many instruments are available to carry out this last task; in particular, several image-acquisition platforms are available for high-content screens[8]. Finally, a spotter (for example, Genetix) is required for researchers who want to make their own solid-phase optimized transfection RNAi (SPOT-RNAi) arrays.

Computing infrastructure

It is crucial to establish a solid computing backbone to support genome-scale high-throughput RNAi (HT RNAi) screening as it is not only a matter of reducing repetitive tasks and increasing throughput, but makes the difference between a successful, insightful study and an expensive nightmare. The key elements to focus on are listed.

Relational database with large capacity storage and professional back-up system. These are needed to organize, store and readily retrieve all levels of information that go into and come out of an HT RNAi screen. This includes genome and reagent sequences, tube, plate or slide identification numbers, raw experimental data and processed data. The 'Excel swamp' can be avoided by using, as a bare minimum, a carefully built Filemaker Pro or Access database, both of which can be self-taught. MySQL and Postgres databases provide more solid alternatives, although these require professional programmers.

Laboratory information management system. The risk of data handling errors occurring in such complex studies is difficult to eliminate completely, and even a rare error — either human or robotic — can cause huge losses of data, time and money. A good laboratory information management system (LIMS) is crucial, combined with bar-coded sample labelling, not only to yield an efficient process with minimized error risks, but also to efficiently troubleshoot errors that do occur. As the cheapest solution, a paper-based LIMS, if designed carefully and implemented diligently, can be effective in tracking and managing data flow throughout the most complex screening processes. Off-the-shelf LIMS software solutions typically require significant customization to address HT RNAi workflows, involving nearly as much programmer time as developing solutions from scratch. The construction of a software LIMS is therefore an expensive and complex long-term project that is only justifiable if several HT RNAi screens will definitely be run (for example, in the case of screening facilities).

Data processing, graphing and statistical analyses. Processing, statistically analysing and graphically representing the large and often multi-parametric data sets that are produced from RNAi screens can largely be done within Excel, although larger data sets will stretch the limits of this software. More specialized packages such as Spotfire , offer a broader range of statistical tools, and more powerful, streamlined graphing options. Some image-analysis software for automated microscopy instruments already integrate some of this functionality.

Reagents for mammalian HT RNAi screens

A range of small RNAs have been developed as silencing reagents for use in mammalian HT RNAi screens, each with their own advantages and disadvantages (FIG. 1).

Synthetic siRNA-like molecules. Most mammalian cell-based RNAi studies have used siRNAs that are designed to closely mimic endogenous 21-nt siRNAs with 2-base overhangs at both 3′ ends[7]. Several genome-scale libraries have been built on this template (TABLE 3), incorporating a range of sequence-selection criteria to maximize the probability of potent target mRNA cleavage while minimizing the risk of generating off-target effects (OTEs; see later section)[28,39]. When it has been carried out, experimental validation of these libraries has typically shown approximately 80–90% probabilities of individual siRNAs yielding a >70% reduction in target mRNA levels after 48 hours under standardized conditions in transformed human cells. As discussed earlier, it is important to bear in mind that the silencing threshold needed to generate a detectable LOF phenotype depends both on the gene[40] and on the sensitivity of the phenotypic read-out being used.

Importantly, using multiple siRNAs that target each gene, the combined probabilities of achieving >70% silencing are theoretically increased to ~95% or more. Although the concept of using such a pool of siRNAs is attractive for achieving a higher probability of strong silencing in far fewer experimental samples, it assumes that the silencing performance of the pool is at least as good as the individual siRNAs. In fact, when carefully optimized, such 'low-complexity siRNA pools' (3–6 siRNAs per pool) generally perform better than the worst of their constituent siRNAs, but not quite as well as the best, as poorly performing siRNAs have been found to compete with better ones[41]. Similarly, the specificity profiles of such pools seem to be 'cleaner' (fewer apparent OTEs, as measured by cDNA microarrays) than those of the 'dirtier' siRNAs in the pool, but not as clean as the best ones (A. Khvorova, personal communication). The increased throughput and reduced cost of using such pools (or polyclonal shRNA preparations, see the next section) are therefore likely to come at the cost of higher rates of false negatives compared with using each of the constituent siRNAs individually.

The recent development of endonuclease-prepared siRNAs (esiRNAs)[31,42] takes the pooling concept to a higher level. esiRNAs are produced from 200–500 bp dsRNAs that are transcribed *in vitro* from DNA templates and then digested by either a recombinant Dicer enzyme or bacterial RNase III. The result is a high-complexity 'cocktail' of siRNA-like molecules, all targeting a single gene. So far, just one screen has been carried out using esiRNAs[31], but this suggests that these RNAs can offer silencing efficacies that are comparable to those of siRNAs, with the promise of significantly lower production costs, and perhaps even the hope of cleaner specificity profiles. If large-scale production can be developed to yield ready-to-use esiRNA libraries of reproducibly high quality, the more in-depth characterization of all aspects of their performance (including silencing, specificity, kinetics and toxicity) in a wider range of cells will be crucial to their wider adoption in HT RNAi screening.

Table 1 | **Available cell lines for RNAi in *Drosophila melanogaster***

Cell lines	Description	References
Schneider derivatives (S2, S2*, S2-R+ and DL2)	Embryonic, phagocytic, semi-adherent, round or flat	74,75
Kc	Embryonic, phagocytic, round	76
Clone 8‡	Imaginal discs, epithelial cells	77
Primary cells	Embryonic muscle cells	J. Bai and N.P., unpublished observations
Primary cells	Embryonic neurons	K. Sepp and N.P., unpublished observations

‡Clone 8 cells require transfection for efficient RNAi (REF. 14). For details on transfection reagents see REF. 34.

The use of all synthetic RNAi reagents depends on their efficient delivery into cells. The range of commercially available lipid- and peptide-based transfection reagents offers ample potential for optimizing delivery conditions for most transformed cell lines. Importantly, for large-scale screening applications it is worthwhile not only to test for transfection efficiency and associated cellular toxicity, but also to monitor batch-to-batch variability in the efficiency of the transfection reagents. Beyond lipid- and peptide-based transfection, several devices are now commercially available for electroporation in microwell plate formats (for example, products from Ambion and Cytopulse). Although the overall observed performance has been promising using this technique, including the crucial issue of well-to-well reproducibility, the significantly higher amount of siRNA needed per well (up to tenfold higher) is costly.

When used to deliver siRNAs or esiRNAs, all of the methods described above yield only transient silencing, which typically ceases after 5–7 days in actively dividing cell lines. So, when sustained silencing is desired for more than ~5 days, or where efficient delivery becomes limiting (as is the case with certain terminally differentiated primary cells), the best choice is viral delivery of shRNA constructs.

Vector-based shRNA libraries. The advent of shRNA technology has allowed the development of cheaper, renewable RNAi libraries that can be delivered into almost any transformed or primary cell type, enabling sustained silencing over weeks if necessary. The shRNA approach is most powerful when combined with viral vector-based delivery, which can yield nearly 100% delivery in many cell types. So far, retroviral, adenoviral and lentiviral vectors have been most widely used, with notable successes, and several libraries are now available (TABLE 3).

Several technical hurdles initially slowed down the development of first-generation shRNA technology[24,25]. Some libraries were plagued by the instability of shRNA constructs, which is now addressed by using recombination-deficient host strains, more stringent bacterial growth conditions and inclusion of selectable markers in close proximity to the hairpin construct. Another problem was insufficient expression levels; this contributed to the variable silencing performance that was commonly observed with first-generation constructs, which were designed as hairpin transcripts driven by RNA polymerase III promoters, entering the RNAi pathway as pre-miRNA-like molecules (BOX 1).

Second-generation shRNA libraries now offer significantly better silencing performance than their predecessors[43], benefiting from a combination of multiple design improvements, including many of the same sequence features that are used to optimize siRNA silencing. One new library integrates the use of a 'backbone' sequence that is based on an endogenous miRNA (miR-30). This yields so-called 'shRNAmiRs' that enter the RNAi pathway as pri-miRNAs upstream from simple hairpins, and are thought to undergo more efficient processing by the RNAi machinery[43]. Although the relative contributions of the updated sequence designs and the miR backbone to the improved performance remain individually unclear, the compatibility of these constructs with either RNA polymerase II or III promoters also enables a wider range of choices for controlling expression, including tissue-specific promoters. However, despite these advantages, the inherent cell-to-cell variability of expression that is observed with all shRNA vectors still requires the application of selection and reporter strategies to focus LOF analyses on those cells that express the highest levels of shRNAs.

Finally, it should be noted that the new generation of shRNAs has not yet been fully characterized with respect to specificity or toxicity. All current vector-driven shRNA approaches are inherently limited in the amount of regulation of shRNA-expression levels, making it difficult to control the risk of triggering OTEs. The development of inducible shRNA vectors (for examples see REFS 44,45), promises further refinements in this area.

Screening paradigms and formats

In undertaking a large-scale cell-based RNAi screen, the next question is typically that of the screening paradigm to be applied. This can be a systematic screen, targeting each gene individually, or a selection-based screen, using pooled libraries of shRNAs to target many genes at once. The two approaches have different strengths and limitations, which will determine which approach is used in conjunction with a particular cell type and LOF phenotype (summarized in FIG. 2). Both approaches rely on the exploitation of annotated genome sequences and the accuracy of current gene and transcript predictions.

Systematic screening. Systematic screening offers the possibility of working through any selection of genes, from a selected subset to the entire genome. Each gene is silenced individually and an appropriate read-out methodology is applied to characterize and measure the resulting LOF phenotype. This is the most direct approach to RNAi screening, and the most broadly applicable in terms of the range of phenotypes that can be assayed. However, the optimization that is required to make assays both sensitive and robust enough to yield reproducible results throughout the genome-scale screen represents a significant challenge that is often underestimated. In addition, the cost associated with systematic genome-scale screens can be considerable, as large amounts of screening

Table 2 | **Available *Drosophila melanogaster* dsRNA collections**

Coverage	Description of reagents	Availability	Comments	References
Entire *D. melanogaster* genome	21,396 dsRNAs, with an average length of 400 bp	PCR products for dsRNA synthesis are available from Eurogentec	Amplification was carried out using gene-specific primers designed to combine genome annotations that are available from the original BDGP/Celera data (13,672 genes) and the Sanger Center data (20,622 genes)	11
The best annotated *D. melanogaster* genes	13,071 dsRNAs of 300–800 bp	dsRNAs are available from Ambion	Design based on Flybase v3	
Most *D. melanogaster* genes that are phylogenetically conserved with mammalian genes	7,216 dsRNAs of 300–600 bp	dsRNAs are available from Open Biosystems	For each gene, the exonic sequence was amplified using gene-specific primers	18
Genes that are represented in the cDNA set 1 collection from the BDGP	4,923 dsRNAs of variable size	dsRNAs are available from the authors of REFS 14,73		14,73

BDGP, Berkeley *Drosophila* Genome Project.

reagents are required, as well as expensive instrumentation for automation, and extensive computing infrastructure, which is crucial for the management and analysis of the large and complex data sets.

Systematic screens can be carried out using either synthetic siRNA or dsRNA libraries, or vector-based shRNA libraries. These are typically carried out in arrayed formats using microwell plates with 96 or 384 wells (which yield sufficient numbers of cells to achieve statistically relevant data sets), and use either microscopy or plate readers for read-out. Although some cells and assays are more difficult to adapt to the 384-well than to the 96-well format, the 384-well format offers faster throughputs, shorter timelines and lower costs.

Another emerging format used for systematic RNAi screening is based on the 'reverse-transfection' method, which can be applied to cell arrays[46], and is also known as solid-phase optimized transfection RNAi (SPOT-RNAi). Cells are cultured on a glass slide that is printed in known locations with discrete spots of nanolitre volumes of dsRNAs or siRNAs[47–50]. Only those cells that land on the printed spots take up the RNA, forming clusters of 80–200 cells that silence the targeted genes. At only 150–250 μm in diameter, 5,000–10,000 such spots can be printed on a standard microscope slide, significantly reducing the volume requirements for most screening reagents. Although restricted to the use of microscopy read-outs, this screening platform offers the potential of vastly increasing throughputs and significantly reducing costs, while still allowing a wide breadth of multiplexing experiments.

SPOT-RNAi is potentially ideal for modifier screens (BOX 2) to be carried out in cells that have been pre-sensitized through the addition of a single dsRNA, siRNA or cDNA. This can be used to identify shared components or parallel pathway components, synthetic lethal effects and mechanisms of suppressing the over-activation of cellular pathways that results from gain-of-function (GOF) mutations. Microarrays printed with highly selected dsRNAs that target a process of interest could also be combined with small-molecule drug-discovery screens in an effort to speed up target identification.

There are however some important limitations to SPOT-RNAi. First, it works best with cell types that show little or no motility. Second, the spotted transfection mixture must be carefully optimized for each cell type, restricting the use of pre-spotted plates to compatible cell types, although some suppliers are getting around this by spotting only the nucleic acids and letting users add the transfection reagents. Third, spot sizes must also be optimized carefully to allow statistically sufficient numbers of cells to be counted on each spot. Finally, the technology will be difficult to adapt to screens in which the read-out involves secreted factors, or for cells that grow in suspension.

Selection-based screening. Selection-based screening allows an entire library of silencing molecules to be delivered as one pool to a single, large population of cells, without the need for arrayed formats. This relies on the ability to sort cells on the basis of their LOF phenotypes, and then identify the responsible silencing molecules. Although this clearly offers the potential for faster, simpler and less expensive screening, these advantages come at the price of a narrower range of applicability, as the RNAi-induced phenotypes of interest must confer a selectable property to individual cells to allow their sorting. One way of achieving this is by basing the read-out assay on the expression of a fluorescent reporter gene. An even more directly selectable phenotype is cell growth modulation, preferably with RNAi causing a growth advantage[24,27,43,49].

The identification of the responsible silencing molecules has been most elegantly achieved through the use of 'bar-coded' shRNA vector libraries. Here each construct, which expresses a single shRNA molecule, also contains a unique bar-code sequence, which is optimized for hybridization-based detection in a microarray format[24,43,49]. The application of molecular bar-code technology maximizes the sensitivity of these libraries by allowing the identification and efficient 'deconvolution' of even extremely rare silencing events that generate the desired LOF phenotypes[24,43,49].

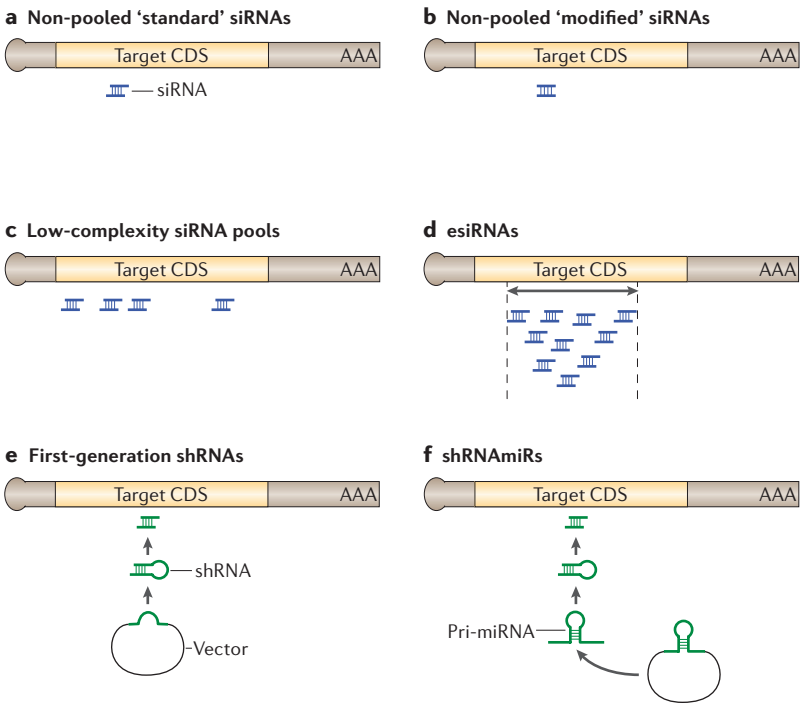


Figure 1 | **Silencing reagents for RNAi screens in mammalian cells.** Publicly available libraries of silencing reagents enable genome-scale RNAi screens in mammalian cells using the following types of molecules. **a** | The most widely used small interfering RNAs (siRNAs) are synthetic molecules with a canonical structure that consists of a 19-bp duplex with 2-base overhangs at the 3′ ends and an unmodified RNA backbone (these are supplied by companies that include Ambion, Dharmacon and Qiagen). **b** | Alternatively, synthetic siRNAs with non-canonical siRNA structures (an siRNA with no overhangs is shown) and/or a modified RNA backbone are also available (for example, the Stealth siRNAs from Invitrogen). **c** | siRNAs can also be used as low-complexity pools of <10 molecules that target the same transcript (for example, SmartPools from Dharmacon). **d** | High-complexity pools of siRNA-like molecules (esiRNAs) can be synthesized by *in vitro* digestion of long dsRNAs using bacterial RNase III or Dicer[42]. **e** | As an alternative to these synthetic molecules, vector-based library reagents are also available, all expressing short hairpin RNA (shRNA) constructs, which are usually delivered virally. Most vector-based shRNA libraries carry a single RNase III-driven shRNA insert[24,25]. **f** | A new vector design now offers an RNase II- or RNase III-driven shRNA insert within a backbone from a known miR (REF. 43), producing so-called 'shRNAmiRs' that enter the RNAi pathway as pri-miRNAs, upstream from conventional shRNAs. CDS, coding sequence.

In theory, the application of pooled shRNA libraries for selection-based screens is expected to be most powerful under conditions that favour the expression of a single construct per cell (that is, low multiplicity of infection), as the simultaneous targeting of multiple genes in the same cell significantly weakens the silencing of each gene[3]. Therefore, although multi-copy delivery conditions (that is, high multiplicity of infection) can be used to maximize the expression of a particular shRNA, their application in the context of a pooled library inevitably favours the targeting of multiple genes per cell, which probably results in individual shRNAs being used significantly below their optimal silencing potential. The resulting dilemma as to which is the bigger risk to the screen's overall sensitivity — inter-gene 'competition' from multi-copy delivery or insufficient shRNA expression from single-copy delivery — can only be resolved through experimental testing on a case-by-case basis. Although recent improvements in shRNA silencing efficacy are beginning to reduce the need for multi-copy conditions, these variables should be thoroughly investigated when optimizing screening conditions, especially for those studies in which the expectations include a low rate of false negatives.

Read-out options

Until recently, assays that are based on the use of plate readers, such as those that use luminescent reporters, were the most favoured read-outs for HT cell-based screening owing to the simplicity of workflow, their strong robustness and generally high reproducibility. In the case of selection-based RNAi screens using pooled libraries, fluorescent reporter-based assays can enable fluorescence-activated cell sorting (FACS) of treated cells, offering powerful read-outs. However, the inherently narrow insight that these methods offer into cellular physiology has driven the fast development in recent years of high-content screens that provide multi-parametric read-outs — that is, they measure multiple phenotypic features simultaneously, usually by microscopy. Several automated microscopy platforms are now available that offer the rapid, robotic acquisition of bright field and/or multi-channel fluorescence microscopy images from both standard slides and microwell plates. A detailed comparison of the many features offered by these systems and some key suppliers of well-tested systems has been published recently[8]. Although auto-focusing and optical resolution are still undergoing important refinements, it is the automated processing and analysis of the enormous wealth of resulting image data that now present the greatest challenges.

Available image-analysis tools are now undergoing rapid and much-needed improvements to upgrade their applicability for cell-based high-content screens. Most image-processing packages currently perform best with intensity-based read-outs (for example, counting all cells above or below a certain intensity threshold) and simple morphometric read-outs from gross changes in sub-cellular localization patterns (for example, cytoplasmic to nuclear translocations, or nuclear morphology). However, more complex shape and structure-based read-outs remain problematic, including accurate segmentation of cells that grow very densely or overlap one another. Therefore, the development of better algorithms that are compatible with high-content screens is eagerly awaited, and perhaps already heralded at least in part by a new wave of object-orientated image-analysis packages that are now emerging (for example, Cellenger and CellProfiler), which offer improved cell segmentation and a powerful set of measurement tools. However, these new tools often require significantly more computing power to achieve adequate processing throughputs, representing a notable entry barrier for 'casual' screeners.

Controls, optimization and quality control

Controlling RNAi specificity. siRNAs, dsRNAs, esiRNAs and shRNAs all achieve their effects through a complex array of molecular interactions, including but not limited to those that confer the desired nucleotide

High-content screens
Screens that apply multi-parametric read-outs, that is, that measure multiple phenotypic features simultaneously, usually by microscopy.

Table 3 | **siRNA libraries for use in mammalian cells**

Company	Species	Coverage	Reagent description
*Synthetic siRNA libraries**			
Ambion	Human, mouse and rat	Genome-wide	21 nt with 3′ overhangs; unmodified RNA
Dharmacon	Human, mouse and rat	Genome-wide	21 nt with 3′ overhangs; unmodified RNA
Qiagen	Human	Genome-wide	21 nt with 3′ overhangs; unmodified RNA
Invitrogen	Human	Kinase genes	25 nt with blunt ends; modified backbone
Vector-based shRNA libraries			
Open Biosystems	Human, mouse	Genome-wide	shRNAs with miR backbone in retroviral vectors, sold as bacterial stocks; second-generation design by the Hannon and Elledge laboratories
Open Biosystems, Sigma-Aldrich	Human, mouse	Genome-wide	shRNAs in lentiviral vectors; second-generation design by the RNAi Consortium

*Note that these and various other suppliers offer pre-designed small interfering RNAs (siRNAs) that target individual genes or small collections. shRNAs, short hairpin RNAs.

sequence-based specificity — that is, they cause OTEs. Although both types of OTE (sequence-dependent and independent) have been observed in many screens, they are manageable through the careful design and correct implementation of appropriate controls.

Sequence-dependent OTEs are those elicited by specific nucleotide sequences within the silencing reagent. Partial homology to the so-called 'seed region' of siRNAs (that is, positions 2–8 of the antisense strand) across sequences as short as 8 contiguous nucleotides can yield detectable cross-silencing through mRNA degradation[51,52]. This might partly account for the surprising complexity of the siRNA reagent-specific signatures that have been reported in microarray experiments[51,52]. siRNAs also have the potential to function as miRNAs[53,54], mediating translational inhibition of unintentionally targeted transcripts through short stretches of partial complementarity. This represents a particularly difficult source of OTEs to detect, as the sequence-matching requirements are degenerate and the effects might not always be measurable at the mRNA level (although see discussion in REF. 52). Finally, sequence-dependent OTEs also include the triggering of the interferon response[55]. The high rates of apparent false positives that are often picked up in the first stage of large-scale screens, and the failure of repeated efforts to confirm some single siRNA results with multiple other siRNAs, strongly supports the idea that such effects are real and relatively common.

Efforts to predict sequence-dependent OTEs using advanced sequence-homology analyses have consistently failed, although with some recent progress coming from an increased focus on the seed region[52]. Although improved design algorithms and modified backbone chemistries are now being explored to further minimize these risks, none of these precautions fully eliminates them. The most effective and straightforward way to address sequence-dependent OTEs — ensuring that observed phenotypes are indeed target gene-specific, and not reagent-specific — is to demonstrate that these phenotypes can be generated by multiple siRNAs or shRNAs with completely different sequences, which all target the same gene. Alternatively, a phenotypic rescue experiment, whereby the LOF phenotype can be eliminated under silencing conditions through expression of a version of the target gene that cannot be silenced, offers the ultimate proof of specificity. This can be achieved using an orthologue from a closely related species that has enough sequence degeneracy over the targeted region. An alternative is to use siRNAs that target the 3′ UTR in combination with an expression construct that lacks that UTR region only. Although such experiments can be technically challenging, the tools to carry them out are improving, including methods for generating BAC-based constructs, which yield expression levels at, or near, endogenous levels[78].

Long dsRNAs are an interesting case, as little is known about their OTE risk profiles. Theoretically, if these molecules are thought of as being composed of several shorter siRNAs, their specificity might be higher when used at an equal total concentration to a single siRNA. This is by virtue of 'diluting out' the OTEs of each constituent siRNA, while maintaining the silencing effect on the common target of all of these. However, one must assume that the OTE risks remain significant and can only be ruled out by the same reagent redundancy strategy described above — that is, obtaining the same phenotype from multiple dsRNAs with completely distinct sequences, or by phenotypic rescue experiments. These observations probably apply equally to dsRNAs that are used in *D. melanogaster* cells and high-complexity esiRNA cocktails. Experimental data to test these hypotheses are eagerly awaited.

Sequence-independent OTEs encompass various ill-defined, unspecific effects that are triggered irrespective of the reagent's nucleotide composition. These include cellular events that are triggered by chemical or structural features of the silencing and/or delivery reagents, few of which are well understood. Among these, the interferon-response pathway, which, as noted above, is triggered by certain sequence motifs, can also be activated in a sequence-independent manner[55]. Many delivery reagents cause significant cellular toxicity, which might differ markedly when they are used in combination with silencing reagents. Finally, there is the risk, especially when using high concentrations of silencing reagents, that these might compete with endogenous miRNAs for the RNAi machinery of the cell[60]. Certain

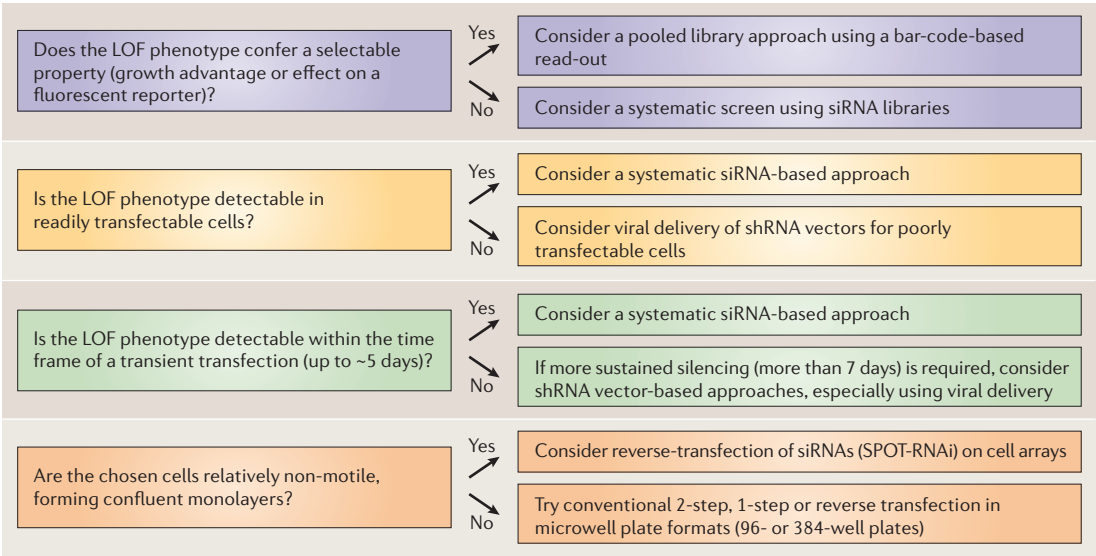


Figure 2 | **Choosing the screening paradigm, experimental format and reagents in RNAi screens.** Technical feasibility is always the first consideration in making the choice of screening paradigm, experimental format and reagents. Some key questions to consider when making these choices are shown. LOF, loss of function; shRNA, short hairpin RNA; siRNA, small interfering RNA; SPOT-RNAi, solid-phase optimized transfection RNAi.

tissues and related cell lines might be more susceptible to this problem if they underexpress key rate-limiting components of the RNAi pathway. Considering the growing evidence that implicates miRNA function in fundamental aspects of maintaining cellular differentiation states and overall physiology, this could account for many of the more complex, pleiotropic, sequence-independent OTEs that are observed in many RNAi experiments.

These concerns can be addressed by the inclusion of so-called negative 'unspecific' or 'scrambled' siRNA or shRNA controls. These are usually more useful than 'mock transfection' negative controls in which the silencing reagent is excluded, as these might have effects that are completely unrelated and irrelevant to the experimental samples. However, identifying appropriate negative-control siRNAs is not a trivial matter. Despite being designed to avoid targeting any known transcripts in the target genome, it cannot be ruled out that these might trigger their own sequence-dependent OTEs. It is therefore worthwhile to screen through multiple siRNA candidates before selecting one that accurately represents the baseline for each chosen cell line and assay. Alternatively, the mean read-out values from several unspecific siRNA controls can be used to yield a more reliable baseline. In screens using pooled siRNAs, it might be more appropriate to also use pools of negative-control siRNAs; although in this case, as with individual 'negative' siRNAs, multiple candidate control pools should be first tested to ascertain how faithfully they represent each assay's baseline.

Optimizing silencing and specificity using multi-pass screens. To achieve an optimal balance between comprehensive coverage, minimization of false negatives and elimination of false positives, all at acceptable costs and within reasonable timelines, systematic RNAi screens typically comprise multiple rounds or passes. The first pass usually focuses on maximizing detection sensitivity. Conditions are chosen to favour maximal silencing, including the use of multiple siRNAs (or other silencing reagents), usually at high concentration, to target each gene. If affordable, each silencing reagent is also used individually. Depending on the type of assay and the threshold used for hit selection (usually 2–3 standard deviations from the baseline), up to 10% of genes will show at least one positive hit — that is, the RNAi treatment causes a detectable phenotype — thereby warranting follow-up.

The initial hits will probably include significant numbers of false positives, which are addressed during the second screening pass that establishes the gene specificity of the observed phenotypes. All the candidate genes that were hit during the first stage are re-tested to eliminate false positives that result from reagent-specific OTEs, that is, those genes for which a positive LOF phenotype cannot be confirmed with more than one distinct siRNA or shRNA in the second pass. Because fewer genes are analysed at this stage, secondary assays are often also carried out to further refine the relevance of selected hits with respect to the biological process of interest. Finally, those genes that are confirmed as 'true positives' in the second pass can be subjected to a final pass, wherein the functional phenotypic read-out is repeated in parallel with an assay to measure target gene silencing (usually a quantitative reverse transcriptase PCR (qRT-PCR) or branched DNA assay), therefore directly confirming the link between the two.

Sensitivity, robustness and quality control. During the optimization phase of any HT RNAi screen, experimental conditions are refined such that they offer the best possible screening window[56], reflecting good sensitivity, strong signal-to-noise ratio and low variability (high robustness and reproducibility). The dynamic range of the

Quantitative reverse transcriptase PCR
This reaction is a sensitive method that is used to detect mRNAs.

Branched DNA assay
A signal-amplification technique that detects the presence of specific nucleic acids by measuring the signal that is generated by many branched, labelled DNA probes.

chosen read-out assay therefore represents the difference between 'baseline values' that are obtained from negative-control genes and representative 'positive hit values' that are obtained from positive-control genes. This is equally important in systematic and selection-based screens; in fact, positive controls have a central role in optimization of the latter, for determining whether sufficient, selectable numbers of 'positive' cells can be generated with the pooled format. These controls not only represent important optimization tools, but also should be included in each screening plate, dish or slide to monitor data quality. Negative-control samples primarily allow the evaluation of sequence-independent OTEs and the normalization of all data subsets from different plates, dishes and slides into a single coherent data set (although in some cases positive controls can also help the latter role as well). Positive-control samples primarily offer a measure of quality control to ensure that all screened genes are subjected to the same or at least a similar stringency of testing conditions. It is therefore crucial to select these control genes carefully.

Any selection of siRNA or shRNA sequences to serve as 'unspecific' or 'scrambled' negative controls must be validated carefully during assay optimization. During the primary screen, if the assumption can be made that the hit rate will be similarly low for all plates, dishes or slides, the mean read-out values from the overall population of experimental samples can be used to set the baseline used for data normalization. Alternatively, the read-out values from negative-control-treated samples can be used. During secondary and tertiary screening passes, because the above 'low hit rate' assumption probably no longer holds true, the quality of negative-control samples becomes important, although at this point the results from the first pass should offer numerous candidates for use as negative-control genes for the chosen screen.

During assay optimization, positive-control genes are typically used for two primary purposes: optimizing target silencing and optimizing the signal-to-noise ratio of the functional read-out assay. First, to optimize silencing, many users focus on transfection efficiency, which, in the case of siRNA-based screening, can be misleading if it is only assessed with the use of fluorescently labelled siRNAs. The appearance of an intracellular signal does not guarantee that these siRNAs are actually functional within the cells, as in many cases significant amounts of lipid-transfected siRNAs can become 'trapped' in endocytic compartments without being available to the cytosolic RNAi machinery. The ultimate test is therefore to measure silencing itself in terms of target mRNA levels (using qRT-PCR or branched DNA). In some cases, the goal of optimizing both silencing and the performance of the functional assay is best met by using different positive-control genes. When the desired LOF phenotype includes cell death or a reduction in cell proliferation, silencing measurements often reflect misleadingly high target-mRNA levels from surviving cells that were probably not adequately transfected and outgrew the well-transfected ones. Therefore, the best choices of positive-control genes for optimizing silencing conditions are those for which losses of function are known not to affect cell proliferation or viability.

A further challenge in optimization is choosing positive controls that are most representative of the desired target gene population. As these will form the key quality-control parameter to determine the level of silencing on each screening plate, dish or slide, a gene that is particularly easy to perturb (that is, one with a low LOF threshold) provides an inclusive criterion, whereas a gene with a high LOF threshold might be too restrictive, perhaps causing unacceptable expense owing to the higher numbers of rounds of screening that are required to satisfy this stringency.

The overall quality of the final data set will depend heavily on the screen's robustness: the degree to which all sources of variability affect experimental reproducibility. These factors include experimental design, technical implementation, data processing and, perhaps most fundamentally, inherent biological variability. Beyond issues that are related to the heterogeneity of the cells, screening read-outs that monitor the end-result of long and complex pathways typically represent indirect measures of a silencing event, especially if that event targets the early steps of an enzymatic pathway. In these cases there is inherently more variability than when using read-outs that measure more direct consequences of the biochemical activity of a silenced target. Such variability is often most apparent when read-outs are taken at a single time-point after silencing, as initial differences in the kinetics of target silencing can be greatly magnified through the 'domino effect' of the individual kinetics of downstream steps. This can be countered by either reading the assay at multiple time-points, or, in some cases, using a single, late time-point. It is also worth noting that the uniqueness of the phenotype is an important factor in increasing confidence, as the likelihood that, by random chance, the off-target hits from two completely distinct siRNAs would both yield similar phenotypes is low. This of course depends on the complexity of the phenotypic features being scored.

Finally, it should be noted that the accurate assessment of baseline levels and screening variability to allow an optimal, statistically correct definition of hit-selection thresholds is a crucial but complex and often under-estimated challenge in HT RNAi screens. Data from control genes and 'simple' dogmatic practices such as the 'three-standard-deviations-from-the-mean' rule for hit selection should be viewed as guidelines, rather than strict rules, to help build a strategically optimal and scientifically sound study.

The future of cell-based RNAi screening

We are clearly only at the beginning of exploiting the fruits of HT RNAi screens, and in the next few years current technologies will be improved and new ones developed. Here we describe a few exciting advances that have already started to take place.

The dominant trend in the past few years has been to develop more specific and potent silencing reagents. Although the set of tools that are currently available is impressive, better and cheaper reagents are likely to be developed. To this end, new designs of siRNA-like molecules continue to emerge, exploring different lengths[57],

structures and backbone modifications. As these undergo field-testing by users, their applicability for HT RNAi screening, rather than for small-scale uses, will be defined not only by their experimental performance but also by the quality and cost of large-scale library production.

A second important area of continuing development has been the search for novel delivery reagents, methods and instruments that will broaden the range of applicability of HT RNAi to difficult or intractable cell types. Beyond the plethora of new lipofection or peptide-based reagents, most of which only offer incremental advances, one exciting development is the successful combination of cell arrays with second-generation lentiviral shRNA libraries[58]. Although still requiring relatively immotile, 'well-behaved' monolayers of cells (as does its predecessor, SPOT-RNAi), this new approach — which could be termed solid-phase infection RNAi (SPI-RNAi) — now promises broader applicability in primary cells and extremely high throughputs.

Also of interest, especially for researchers studying embryonic stem cell differentiation, are siRNAs that allow delayed, inducible silencing. As potent silencing in differentiated cell types that are poorly transfectable is currently difficult, an alternative approach is to use siRNAs that remain stable and inactive in their easily transfectable precursor cells. These can then be activated once the cells have reached the desired state. Such a reagent has already been developed in the form of light-activatable, caged siRNAs (by Panomics). However, building libraries from these reagents currently seems to be financially unfeasible, and might remain the domain of more focused or smaller-scale applications. An alternative, cheaper approach is emerging in the form of next-generation, inducible shRNA vectors (for example, pPRIME[44]).

The range of applications of HT RNAi screens also relies on the range and quality of the read-out assays. Although not specific to RNAi applications, improvements in read-out methods to diversify the phenotypes that can be scored and allow more multi-parametric phenotyping will be important, allowing more information to be extracted from the same data sets while maintaining reasonable screening costs, throughputs and timelines. Cell-based high-content screens that rely on cellular phenotypes as the read-out are widely used because they produce data sets that are rich in information, particularly when many cellular parameters are scored in a single assay. However, this approach is hampered by various technical issues, including image quality, processing time and reproducibility.

Automated microscopy systems and associated software packages are evolving rapidly, driven by increased competition from the growing number of instruments that are now on the market. Priced for different entry levels, the systems are now offering faster and more reliable image acquisition, better optics (including the confocal imaging long-awaited by some users), and an ever-broadening range of image-analysis solutions. Further improvements in several areas will have particularly strong effects on HT RNAi screening, including more powerful image analysis to allow accurate scoring for a wider range of phenotypes, and better integration of system software with the third-party database, storage and high-performance computing solutions that are, of necessity, widely used. In addition, future advances promise to allow temporal effects of silencing to be monitored in the form of time-lapse read-outs. Initially shown to be feasible over the entire genome in *C. elegans*[59], this approach, which is already being applied to cell division in cultured human cells (see the MitoCheck project homepage), can provide both temporal and spatial information. In this context, SPOT-RNAi and SPI-RNAi have advantages over the multi-well format for image acquisition and auto-analysis algorithms.

There is also interest in moving beyond the monitoring of RNAi-induced phenotypes at the levels of cell morphology and behaviour to examine effects using proteomics approaches. This is currently limited by the availability of antibodies, so a quantitative protein-analysis platform that relies on mass spectrometry would be a major advance. Although such technologies, when applied at their broadest scale (proteome profiling), are far from being applicable to HT, they could be used productively in secondary analyses.

Finally, it will be crucial to carefully define data-exchange standards to ensure that groups generating large-scale RNAi screening data sets use common annotation guidelines for disseminating these data sets online. This could follow the example of the minimum information about a microarray experiment (MIAME) standards that are already in place for microarray experiments. The adoption in the future of a minimal annotation for RNAi experiments (MARIE) would ensure that users fully understand RNAi data sets from disparate groups, biological sources and experimental paradigms, and are able to easily compare these data sets.

1. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
Reports the discovery of RNAi that is mediated by long dsRNA in *C. elegans*.
2. Fraser, A. G. *et al.* Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature* **408**, 325–330 (2000).
3. Gönczy, P. *et al.* Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature* **408**, 331–336 (2000).
4. Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H. & Williams, B. R. Activation of the interferon system by short-interfering RNAs. *Nature Cell Biol.* **5**, 834–839 (2003).
5. Clemens, J. C. *et al.* Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc. Natl Acad. Sci. USA* **97**, 6499–6503 (2000).
The first report that addition of long dsRNAs on cultured cells triggers a potent RNAi effect.
6. Hamilton, A. J. & Baulcombe, D. C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**, 950–952 (1999).
7. Elbashir, S. M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498 (2001).
The discovery that transfection of 21-nucleotide siRNAs cells triggers a potent RNAi effect.
8. Carpenter, A. E. & Sabatini, D. M. Systematic genome-wide screens of gene function. *Nature Rev. Genet.* **5**, 11–22 (2004).
A comprehensive review of HT screens with a particularly useful overview of commercially available image-acquisition platforms.
9. Kramer, R. & Cohen, D. Functional genomics to new drug targets. *Nature Rev. Drug Discov.* **3**, 965–972 (2004).
10. Kiger, A. A. *et al.* A functional genomic analysis of cell morphology using RNA interference. *J. Biol.* **2**, 27 (2003).
11. Boutros, M. *et al.* Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* **303**, 832–835 (2004).

12. Bettencourt-Dias, M. *et al.* Genome-wide survey of protein kinases required for cell cycle progression. *Nature* **432**, 980–987 (2004).
13. Eggert, U. S. *et al.* Parallel chemical genetic and genome-wide RNAi screens identify cytokinesis inhibitors and targets. *PLoS Biol.* **2**, e379 (2004).
An example of how a genome-wide RNAi screen can allow drug-target identification.
14. Lum, L. *et al.* Identification of Hedgehog pathway components by RNAi in *Drosophila* cultured cells. *Science* **299**, 2039–2045 (2003).
15. Nybakken, K., Vokes, S. A., Lin, T. Y., McMahon, A. P. & Perrimon, N. A genome-wide RNA interference screen in *Drosophila melanogaster* cells for new components of the Hh signaling pathway. *Nature Genet.* **37**, 1323–1332 (2005).
16. Baeg, G. H., Zhou, R. & Perrimon, N. Genome-wide RNAi analysis of JAK/STAT signaling components in *Drosophila*. *Genes Dev.* **19**, 1861–1870 (2005).
17. Muller, P., Kuttenkeuler, D., Gesellchen, V., Zeidler, M. P. & Boutros, M. Identification of JAK/STAT signalling components by genome-wide RNA interference. *Nature* **436**, 871–875 (2005).
18. Foley, E. & O'Farrell, P. H. Functional dissection of an innate immune response by a genome-wide RNAi screen. *PLoS Biol.* **2**, e203 (2004).
19. Gesellchen, V., Kuttenkeuler, D., Steckel, M., Pelte, N. & Boutros, M. An RNA interference screen identifies inhibitor of apoptosis protein 2 as a regulator of innate immune signalling in *Drosophila*. *EMBO Rep.* **6**, 979–984 (2005).
20. DasGupta, R., Kaykas, A., Moon, R. T. & Perrimon, N. Functional genomic analysis of the Wnt-wingless signaling pathway. *Science* **308**, 826–833 (2005).
An example of a genome-wide RNAi screen in *D. melanogaster* cells using a transcriptional reporter assay.
21. Cherry, S. *et al.* Genome-wide RNAi screen reveals a specific sensitivity of IRES-containing RNA viruses to host translation inhibition. *Genes Dev.* **19**, 445–452 (2005).
22. Philips, J. A., Rubin, E. J. & Perrimon, N. *Drosophila* RNAi screen reveals CD36 family member required for mycobacterial infection. *Science* **309**, 1251–1253 (2005).
23. Agaisse, H. *et al.* Genome-wide RNAi screen for host factors required for intracellular bacterial infection. *Science* **309**, 1248–1251 (2005).
24. Berns, K. *et al.* A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**, 431–437 (2004).
25. Paddison, P. J. *et al.* A resource for large-scale RNA-interference-based screens in mammals. *Nature* **428**, 427–431 (2004).
26. Kolfschoten, I. G. *et al.* A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. *Cell* **121**, 849–858 (2005).
27. Westbrook, T. F. *et al.* A genetic screen for candidate tumor suppressors identifies REST. *Cell* **121**, 837–848 (2005).
28. Huesken, D. *et al.* Design of a genome-wide siRNA library using an artificial neural network. *Nature Biotechnol.* **23**, 995–1001 (2005).
29. Rottmann, S., Wang, Y., Nasoff, M., Deveraux, Q. L. & Quon, K. C. A TRAIL receptor-dependent synthetic lethal relationship between MYC activation and GSK3β/FBW7 loss of function. *Proc. Natl Acad. Sci. USA* **102**, 15195–15200 (2005).
30. Pelkmans, L. *et al.* Genome-wide analysis of human kinases in clathrin- and caveolae/raft-mediated endocytosis. *Nature* **436**, 78–86 (2005).
31. Kittler, R. *et al.* An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. *Nature* **432**, 1036–1040 (2004).
32. Nicke, B. *et al.* Involvement of MINK, a Ste20 family kinase, in Ras oncogene-induced growth arrest in human ovarian surface epithelial cells. *Mol. Cell* **20**, 673–685 (2005).
33. Echalier, G. *Drosophila Cell in Culture* (Academic Press, New York, 1997).
34. Armknecht, S. *et al.* High-throughput RNA interference screens in *Drosophila* tissue culture cells. *Methods Enzymol.* **392**, 55–73 (2005).
35. Ui, K. *et al.* Newly established cell lines from *Drosophila* larval CNS express neural specific characteristics. *In Vitro Cell. Dev. Biol. Anim.* **30A**, 209–216 (1994).
36. Song, E. *et al.* Sustained small interfering RNA-mediated human immunodeficiency virus type 1 inhibition in primary macrophages. *J. Virol.* **77**, 7174–7181 (2003).
37. Liu, G. & Chen, X. The ferredoxin reductase gene is regulated by the p53 family and sensitizes cells to oxidative stress-induced apoptosis. *Oncogene* **21**, 7195–7204 (2002).
38. Ovcharenko, D., Jarvis, R., Hunicke-Smith, S., Kelnar, K. & Brown, D. High-throughput RNAi screening *in vitro*: from cell lines to primary cells. *RNA* **11**, 985–993 (2005).
39. Reynolds, A. *et al.* Rational siRNA design for RNA interference. *Nature Biotechnol.* **22**, 326–330 (2004).
40. Huang, F., Khvorova, A., Marshall, W. & Sorkin, A. Analysis of clathrin-mediated endocytosis of epidermal growth factor receptor by RNA interference. *J. Biol. Chem.* **279**, 16657–16661 (2004).
41. McManus, M. T. *et al.* Small interfering RNA-mediated gene silencing in T lymphocytes. *J. Immunol.* **169**, 5754–5760 (2002).
42. Yang, D. *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* **99**, 9942–9947 (2002).
43. Silva, J. M. *et al.* Second-generation shRNA libraries covering the mouse and human genomes. *Nature Genet.* **37**, 1281–1288 (2005).
44. Stegmeier, F., Hu, G., Rickles, R. J., Hannon, G. J. & Elledge, S. J. A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells. *Proc. Natl Acad. Sci. USA* **102**, 13212–13217 (2005).
45. Dickins, R. A. *et al.* Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nature Genet.* **37**, 1289–1295 (2005).
46. Ziauddin, J. & Sabatini, D. M. Microarrays of cells expressing defined cDNAs. *Nature* **411**, 107–110 (2001).
47. Mousses, S. *et al.* RNAi microarray analysis in cultured mammalian cells. *Genome Res.* **13**, 2341–2347 (2003).
48. Erfle, H., Simpson, J. C., Bastiaens, P. I. & Pepperkok, R. siRNA cell arrays for high-content screening microscopy. *Biotechniques* **37**, 454–458, 460, 462 (2004).
49. Silva, J. M., Mizuno, H., Brady, A., Lucito, R. & Hannon, G. J. RNA interference microarrays: high-throughput loss-of-function genetics in mammalian cells. *Proc. Natl Acad. Sci. USA* **101**, 6548–6552 (2004).
50. Wheeler, D. B. *et al.* RNAi living-cell microarrays for loss-of-function screens in *Drosophila melanogaster* cells. *Nature Methods* **1**, 127–132 (2004).
51. Jackson, A. L. *et al.* Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnol.* **21**, 635–637 (2003).
52. Birmingham, A. *et al.* 3′ UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nature Methods* **3**, 199–204 (2006).
53. Doench, J. G., Petersen, C. P. & Sharp, P. A. siRNAs can function as miRNAs. *Genes Dev.* **17**, 438–442 (2003).
54. Scacheri, P. C. *et al.* Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. *Proc. Natl Acad. Sci. USA* **101**, 1892–1897 (2004).
55. Marques, J. T. & Williams, B. R. Activation of the mammalian immune system by siRNAs. *Nature Biotechnol.* **23**, 1399–1405 (2005).
56. Zhang, J. H., Chung, T. D. & Oldenburg, K. R. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen* **4**, 67–73 (1999).
57. Kim, D. H. *et al.* Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nature Biotechnol.* **23**, 222–226 (2005).
58. Bailey, S. N., Ali, S. M., Carpenter, A. E., Higgins, C. O. & Sabatini, D. M. Microarrays of lentiviruses for gene function screens in immortalized and primary cells. *Nature Methods* **3**, 117–122 (2006).
59. Sonnichsen, B. *et al.* Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* **434**, 462–469 (2005).
60. Ohrt, T., Merkle, D., Birkenfeld, K., Echeverri, C & Schwille, P. *In situ* fluorescence analysis demonstrates active siRNA exclusion from the nucleus by exportin 5. *Nucleic Acid Res.* **34**, 1369–1380 (2006).
61. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231–235 (2004).
62. Gregory, R. I. *et al.* The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235–240 (2004).
63. Landthaler, M., Yalcin, A. & Tuschl, T. The human DiGeorge syndrome critical region gene 8 and its *D. melanogaster* homolog are required for miRNA biogenesis. *Curr. Biol.* **14**, 2162–2167 (2004).
64. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).
65. Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear export of microRNA precursors. *Science* **303**, 95–98 (2004).
66. Bohnsack, M. T., Czaplinski, K. & Gorlich, D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **10**, 185–191 (2004).
67. Chendrimada, T. P. *et al.* TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740–744 (2005).
68. Saito, K., Ishizuka, A., Siomi, H. & Siomi, M. C. Processing of pre-microRNAs by the Dicer-1–Loquacious complex in *Drosophila* cells. *PLoS Biol.* **3**, e235 (2005).
69. Forstemann, K. *et al.* Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol.* **3**, e236 (2005).
70. Pillai, R. S. *et al.* Inhibition of translational initiation by Let-7 microRNA in human cells. *Science* **309**, 1573–1576 (2005).
71. Bagga, S. *et al.* Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* **122**, 553–563 (2005).
72. MacKeigan, J. P., Murphy, L. O. & Blenis, J. Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. *Nature Cell Biol.* **7**, 591–600 (2005).
73. Ivanov, A. I. *et al.* Genes required for *Drosophila* nervous system development identified by RNA interference. *Proc. Natl Acad. Sci. USA* **101**, 16216–16221 (2004).
74. Schneider, I. Cell lines derived from late embryonic stages of *Drosophila melanogaster*. *J. Embryol. Exp. Morphol.* **27**, 353–365 (1972).
75. Yanagawa, S., Lee, J. S. & Ishimoto, A. Identification and characterization of a novel line of *Drosophila* Schneider S2 cells that respond to wingless signaling. *J. Biol. Chem.* **273**, 32353–32539 (1998).
76. Echalier, G. & Ohanessian, A. *In vitro* culture of *Drosophila melanogaster* embryonic cells. *In Vitro* **6**, 162–172 (1970).
77. Peel, D. J., Johnson, S. A. & Milner, M. J. The ultrastructure of imaginal disc cells in primary cultures and during cell aggregation in continuous cell lines. *Tissue Cell* **22**, 749–758 (1990).
78. Kittler, R. *et al.* RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells. *Proc. Natl Acad. Sci. USA* **102**, 2396–2401 (2005).

Acknowledgements
We thank B. Mathey-Prevot, A. Friedman, S. Elledge, R. Zhou, K. Echeverri, C. Sachse and M. Mirotsou for helpful discussions and critical reading of the manuscript. N.P. is an investigator of the Howard Hughes Medical Institute, USA.

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION
Ambion: http://www.ambion.com
Berkeley *Drosophila* Genome Project: http://www.bdgp.org
CellProfiler cell image analysis software: http://jura.wi.mit.edu/cellprofiler
Cenix Bioscience: http://www.cenix-bioscience.com
Cytopulse: http://www.cytopulse.com
Definiens Cellenger: http://www.definiens.com/applications/cellenger.php
Dharmacon RNA Technologies: http://www.dharmacon.com
***Drosophila* Genome Resource Center:** http://dgrc.cgb.Indiana.edu
***Drosophila* RNAi Screening Center:** http://flyrnai.org
Genetix: http://www.genetix.com
Invitrogen: http://www.invitrogen.com
Matrix Technologies Corporation: http://www.matrixtechcorp.com
Open Biosystems: http://www.openbiosystems.com
Panomics homepage: http://www.panomics.com
Qiagen homepage: http://www.qiagen.com
Sigma-Aldrich homepage: http://www.sigmaaldrich.com
Spotfire — Interactive Visual Analysis: http://www.spotfire.com
TekCel: http://www.tekcel.com
The Eurogentec web site: http://www.eurogentec.com
The MitoCheck project homepage: http://www.mitocheck.org
The RNAi Consortium: http://www.broad.mit.edu/genome_bio/trc
Access to this links box is available online.