Tales from an academic RNAi screening facility; FAQs

Ming Jiang, Rachael Instrell, Becky Saunders, Haakon Berven and Michael Howell

Abstract

RNAi technology is now a well-established and widely employed research technique that has been adopted by many researchers for use in large-scale screening campaigns. Here, we offer our experience of genome-wide siRNA screening from the perspective of a facility providing screening as a service to a wide range of researchers with diverse interests and approaches. We have experienced the emotional rollercoaster of screening from the exuberant early promise of a screen, the messy reality of the data through to the recognition of screen data as a potential information goldmine. Here, we use some of the questions we most frequently encounter to highlight the initial concerns of many researchers embarking on a siRNA screen and conclude that an informed view of what can be reasonably expected from a screen is essential to the most effective implementation of the technology. Along the way, we suggest that for this area of research at least, either centralization of the resources or close and open collaboration between interested parties offers distinct advantages.

Keywords: *RNAi; siRNA; high-throughput screening facility*

INTRODUCTION

The High-Throughput screening facility at the London Research Institute was set up in 2007 in order to make genome-wide RNAi screening readily accessible to the 45 basic research groups in the Institute each with their own diverse research interests. We specialize in genome-wide siRNA screens primarily in human and *Drosophila* cultured cells. This article is not intended to be a comprehensive, 'how to screen' manual since many excellent reviews concerning such matters have already been published [1, 2, 3–5]. Instead, it uses our experiences of establishing and running a central screening facility as a framework for considering some aspects of using genome-wide siRNA screening as a research tool in an academic environment. It is inspired by our frequent interactions with researchers wishing to screen and uses their typical queries as a guide to what anyone contemplating such a screen might want to know and what, if any, misconceptions about what a screen can deliver may still lurk in the broader community. It will not consider shRNA screens although some of the points made here may well be relevant to shRNA as well.

Do you need to have a dedicated facility to screen?

In outline, the process of screening a siRNA library is very simple; all that is required is to introduce RNAi reagent into cells uniformly across a number of multiwell plates and record the outcome. The problems start with the scale and it is this that can put the

Corresponding author. Michael Howell, High-Throughput Screening facility, Cancer Research UK, London Research Institute, 44 Lincoln's Inn Fields, London WC2A 3LY, UK. Tel: +44-20-7269-3170; Fax: +44-20-7269-3581;
E-mail: michael.howell@cancer.org.uk

Ming Jiang obtained her PhD at the Department of Genetics, University of Leeds before joining the HTS facility in 2007.
Rachael Instrell obtained her MSc at the University of Nottingham before working in the High-Throughput Biology division at GlaxoSmithKline. She joined the HTS facility in 2007.
Becky Saunders obtained her Bioinformatics PhD at University College London before joining the HTS facility in 2007.
Haakon Bervan obtained his MSc in Bioinformatics from the University of Newcastle Upon Tyne before joining the HTS facility in 2010.
Michael Howell obtained his PhD at the Biochemistry Department, University of Cambridge before starting the HTS facility in 2007.

technology beyond the reach of any one single researcher. In our case, a genome-wide collection of pooled siRNA oligonucleotides consists of 267, 96-well library plates and so a screen in triplicate consists of 801, 96-well plates. A triplicate screen is equivalent to a stack of 96-well plates 14 m high, consists of almost 77 000 wells (ie 77 000 transfections or immunostainings, etc.) and any manipulation that takes 1 min/plate will equate to 13 h needed to process the entire screen. Despite these numbers (not large by Pharma screening standards but usually disconcerting enough for academic researchers), it is still entirely possible to transfect 800 plates in a working day with only modest automation and a minimum of variability. Adopting smaller formats (384-well plates or spotted arrays) can increase the assay throughput but usually require more specialized infrastructure and may impose some limitations on the assay design [6, 7].

Given the costs of siRNA collections and the infrastructure required to conduct large-scale assays (let alone the time investment needed to establish just one screen) it makes economic sense to provide screening as some form of core or shared resource for a number of interested parties. Such a core enables the assembly of the biological, programming and bioinformatics expertise required in one location, ensures a degree of continuity across many screens and affords the opportunity to create a central repository of data for the wider community. There are many models as to how to organize a central resource, the most common models are a service, as a facility or a collaborative facility hybrid. A service might be characterized as simply operating machinery without necessary knowledge of the sample provenance or what the intention of the experiment is and while such a model is clearly fine for some research activities it is probably less well-suited to RNAi screening. A facility might provide advice and specific machine training but not necessarily provide assistance with the screen itself and would usually require evidence of the reliability and robustness of the proposed assay prior to allocation of time and resources. Many fine facilities operate something along these lines. We operate a more time-consuming version of a facility where collaboration is paramount and interactions with the researcher start at the very inception of the project and we play an active role at all subsequent stages. Why we favour this 'involved' approach will become apparent but in summary, establishing an assay and

performing a screen based on it, requires intimate knowledge of the assay system. For some researchers interested in screening only a restricted set of genes (e.g. kinases or similar themed collections) extensive automation might be helpful but not necessary. However, even under these circumstances there is still much to recommend centralization of the screening knowledge and expertise and data archiving.

Why do researchers screen and what are their expectations?

The reasons why researchers want to perform an siRNA screen are perhaps obvious; a screen can provide fresh insight or a new perspective on a process or area of biology or identify more precisely a general activity that is believed to exist from other work (all of which can be achieved to some extent). However, the hope that a single screen will yield a definitive catalogue of all activities required for a phenotype or identify genes that act universally in all cellular settings is unlikely to be immediately satisfied (see below). Why don't people screen? Even with easy access to screening resources there are a number of reasons why people may not want to screen (aside from the obvious reason that it is an inappropriate tool for their needs). Fears about the possible variability and non-reproducibility of screens as well as considerations of the return for the time/cost investment have an element of truth about them that we shall explore below. Occasionally, there are more generalized misgivings that since screening involves robots, the scientific endeavour is reduced to a soulless mechanical activity requiring little or no thought.

Those involved in RNAi screening have perhaps not helped themselves by inheriting the rather muscular and possibly alien, language of Pharma small molecule screening to describe their activities. For example, a 'hit' is an unnecessarily optimistic name for a statistical outlier and the term 'validation' causes as many problems as it might solve (see below). The question, 'How am I going to be confident that this strange result is in fact real and significant' engenders a different and perhaps more familiar mindset in an academic researcher than does 'How do I validate a hit'. Moreover, the notion that screening is the antithesis of creative scientific thought couldn't be farther from the truth. Like all experiments, screens are, one way or another, based on a central thesis, i.e. that there exists an activity which when

knocked down elicits phenotype X. One can go a stage further and be more explicit in stating the central hypothesis:

> That there exists an activity in these cells (which have been grown in a certain way and whose response characteristics we assume to represent the characteristics of the originating cell), which when reduced below an unknown threshold by a specified RNAi reagent introduced using a specified transfection reagent at a specific concentration can produce phenotype X after a specific period of incubation under specified conditions (assuming that we can measure phenotype X and that it is sufficiently distinct from other phenotypes).

Although this may seem to be pedantry for its own sake, it does focus attention on the assumptions inherent in the assay design, the points where assays might show greatest variation and where answers might not be as universally applicable as first hoped. Like genetic screens in model organisms [8], an siRNA screen faithfully reports the answers to the specific question posed so if the results of the screen do not match expectations, it is always worth reconsidering the assumptions underlying the assay and screen. A central managed resource can provide guidance, helping to define the aims and focussing on the key objective while negotiating the compromises that have to be made to balance the technical feasibility of the undertaking with the return of meaningful biological data. Given that we find ourselves frequently negotiating these issues and discussing screens, what are the types of question we most frequently encounter?

'I have an idea for an assay, when can I screen?'

While transferring an assay from a 'one pipette/coverslip' environment to a 'multi-channel device/multiwell plate' environment might seem to be a simple linear question of re-optimization for a smaller physical chamber, it is often less than straightforward. The many variables that make up an assay are interconnected such that alteration in one can have profound effects on the outcome of the assay (Figure 1). Although, some of the assay steps will be fixed (e.g. only one cell line or cell type or one class of assay readout is possible), the key here is that the development of the assay is an iterative process attempting to maximize the response while
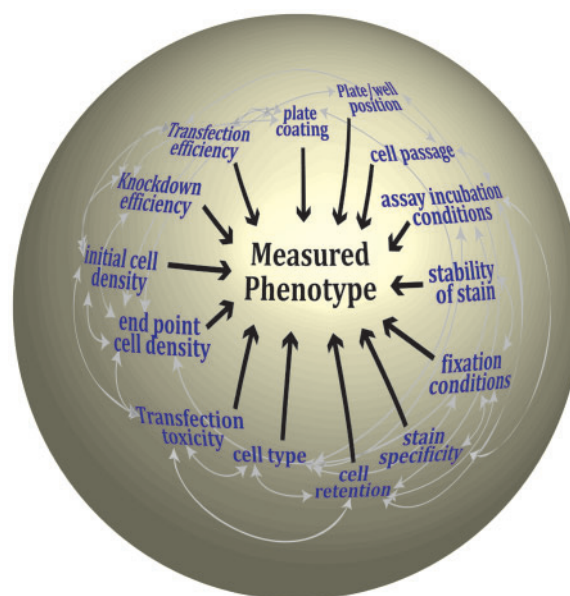

Figure 1: Exploring assay development space. A pictorial representation of the relationships (grey lines) between just some of the variables that impinge on the phenotypic output of an assay (black arrows) and which often have to be balanced against one another when developing an assay for a high-throughput screening campaign. For example, the final cell density in any well is influenced by the initial cell density the transfection efficiency, the transfecton toxicity, the well position, the plate coating, etc.

minimizing the experimental variation and identifying the level of tolerance of the phenotypic output to this variation. This process requires many test plates and pilot (small scale) studies that are best performed using the equipment that will be employed for the primary screen (especially if the assay is in 384-well plates) and moreover, requires constant interaction between data analysts and biologists to maximize the return of data [9]. In our facility, much of our time is spent developing the assay with the researchers in a process that usually takes months rather than weeks but which ultimately equips us with an intimate knowledge of the assay system and the variables influencing the assay outcome.

'What source of siRNA reagents should I use?'

Anyone setting up a facility or venturing into screening will be faced with the decision as to which siRNA collection to use; usually, the choice boils down to a commercial or homemade source. Despite anecdotal reports that siRNAs from supplier

x are 'better' than those from supplier y (declarations often based on quite small sample sizes), we are unaware of any peer-reviewed literature clearly identifying one commercial supplier as preferable over another. Although all suppliers try to improve their design algorithm and bioinformatics filters that inform their siRNA designs [10], most collections will be out of date to some extent as a result of the rapid changes in genome annotation and ever expanding knowledge of small RNA biology [11–14]. However, they are easy to manage and purchase (at a price). An alternative and successfully used approach is to curate your own collection which is very labour-intensive and requires extensive knowledge of the cell line used but has the advantage of increased reagent specificity [15]. The commercial world can be avoided completely by creating your own reagents, e.g. esiRNA collections [16] in a similar manner to *Drosophila* and *Caenorhabditis elegans* collections. However, this approach shifts the emphasis for maintenance and quality control of the collection firmly toward the facility. Another aspect to consider is that as the primary screen is made easier by automation and centralization of expertise, it becomes ever more frustrating to limit the number of candidates that can be taken forward from a primary screen to either a repeat screen or validation step to just the relatively small number that can be afforded by the individual researcher and so a second library of either alternative sequences or individual oligonucleotides is of immense benefit for follow-up studies although 'cherry-picking' from such collections places even greater demands on the facility infrastructure.

'What transfection reagent should I use?'

One of the most frequently encountered and least discussed technical aspects is the toxicity of reagents used to transfer siRNA into cells. We have observed significant batch-to-batch variation among some reagents [6] (and even 'within-batch' variation), reagent and cell type interactions [17], reagent-induced cell-specific morphology effects and reagent-induced assay readout interference. In our view, most reagents induce some form of stress in cells and how this stress manifests itself depends on the cell type and the assay conditions [18–20]. Although alternative technologies such as electroporation exist, they are seldom affordable on the scale of a genome screen (and can themselves suffer from similar toxicity issues) [21, 22]. We deal with this reagent problem empirically by iterative testing of reagent types and doses ideally in an assay mimicking the intended readout as far as possible and constantly monitor reagent performance while developing and perfecting the screen (Figure 2).

'What cells should I use?'

Cells are the central component in any assay and the component about which most assumptions are made with respect to their genotype, response and ability to serve as a model for a cell type *in vivo*, etc. Cell lines differ in their inherent ability to be transfected [22]. We find that how cells are manipulated is crucial to the success of any assay and therefore try to enforce strictly standardized cell culture (including regular growth rate and contamination monitoring). Additionally, the age or passage of the culture *per se* influences the knockdown response and phenotypic
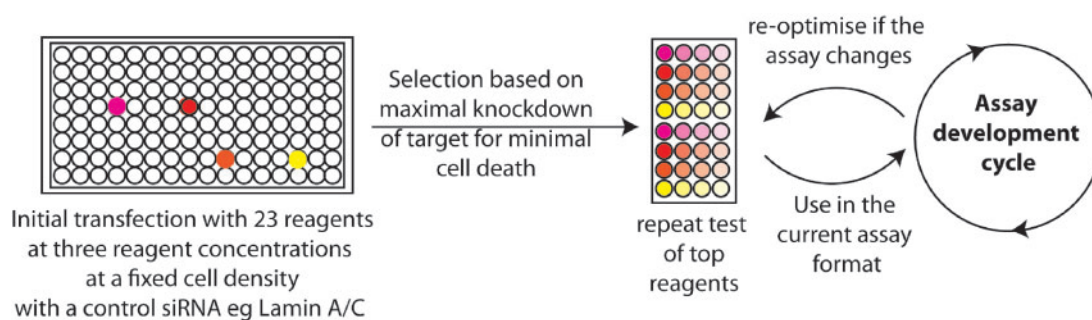

Figure 2: Our approach to optimizing the transfection protocol for a cell line. We initially tested 23 commercial reagents at three concentrations in a reverse transfection format using an siRNA against Lamin A/C and classify reagents by their knockdown efficiency and impact on cell viability. In the second round, the top four or five reagents (with maximal knockdown efficiency and minimal toxicity) are tested over a narrower range of reagent doses and cell densities in either an identical assay or one that is close to the final screen assay if a positive control siRNA exists. The optimal conditions from this test are then used throughout the assay development process but are revisited as the assay develops and changes are made.

outcome [23, 24]. Although not a universal practice, we try to complete genome-wide screens with one batch of cells rather than split into several batches over many weeks, as this can be equivalent to performing several similar but subtly distinct screens. Why might the maintenance and culture history of cells matter so much? It is clear that cells in culture are not uniform and the individual cells in any population display natural variation in cellular states and phenotypic response [25–29]. It is probable that this variation contributes to some of the observed variation in response to siRNA challenge (see below) [30]. Thus anything that might favour one cell state rather than another (like allowing cells to overgrow) can influence the measured phenotype. An alternative approach gaining favour in some circles is to use cells frozen down in assay-ready plates [31]. Cellular responses are also subject to plate positional, edge effects, particularly in 384-well and greater well formats that may not always reveal themselves as changes in viability [32].

A common assay format is to compare the cellular responses to different incubation conditions (e.g. comparing two screens on the same cell line with and without drug) or to compare two related cell lines. For cell line comparisons, assumptions are made about the similarity or differences between them (either with respect to genotype or phenotypic response) and these assumptions may benefit from careful consideration. For example, recent technical advances are revealing extraordinary inter- and intra-tumour genotype and gene expression heterogeneity [33–35] and presumably the same applies to the cell lines derived from them [36]. So far, our experience suggests that screens aimed at such comparisons would benefit from more than one pair-wise comparison.

'What readout should I use?'

Ideally the readout of the assay should be limited only by the imagination of the researcher and what is measured should be as accurate and complete a representation of the desired phenotype as possible. It is in this area that a facility needs maximum technological flexibility. If the cost of the proposed screen readout is likely to be prohibitive or the assay protocol too technically challenging for the degree of automation available, most researchers are willing to consider alternative, substitute, measures which will be good enough (which requires that the system is sufficiently well characterized to recognize

that a readout is good enough). Broadly speaking, we favour fixed-endpoint assays where the readout is captured on a cytometer or microscope rather than homogenous assays, as they afford the greatest opportunity for multiplexing and alternative data gathering [37]. Indeed image analysis and collection of multiple parameters for each cell is an increasingly popular readout for many assays and may provide a robust approach to manage the biological variation seen in screens [38–41]. Although well-averaged measurements can obscure subtle variations in phenotypic response in sub-populations of cells [42] it is often possible to identify simpler population classifications that are sufficiently informative to serve as a screen-readout [43]. Image based readouts also provide a facility with a wealth of bystander data (data not originally planned to be used by the researcher but which can be accumulated and amalgamated across screens to create an useful database, e.g. cell viability or cell shape). A core facility is in the best position to maximize data extraction from different screens and collate them centrally.

'What controls should I use?'

Control siRNAs are usually included on every plate in a screen and can be used in different ways at different times in the screening process. We tend to use them as indicators of screen efficiency and for data curation in the primary screen and as tools for normalization in follow-up screens. Negative or null control siRNAs are often designed against non-mammalian proteins or are modified so that they do not enter into the RISC complex [44, 45]. We often find in analysing screen data that such controls cannot be assumed to occupy the central, medial position in the dataset as might be anticipated (since they have no effect) but instead, can often be skewed toward one end of the distribution. This is particularly evident in our hands when looking at viability assays; our interpretation has been that almost all siRNAs in the library can have some effect on viability and perhaps simply the engagement of the RISC complex is sufficient to retard cell growth to some extent in some lines. Often the point of the screen is to identify putative activities so positive control siRNAs (i.e. that can elicit the desired phenotype) may not be readily available. It is possible to develop an assay and prosecute a screening campaign in the absence of a positive control but very much harder to do so. It is also

common practice to use an siRNA which produces a recognizable phenotype as a control to demonstrate that the transfection was effective [1, 4].

'How do I analyse the data?'

Much work has gone into identifying appropriate methods to analyse RNAi data to compensate for assay bias and enrich for true hits by statistical and bioinformatics analysis [46–48]. Since different assays can have distinct response characteristics there is unlikely to be only one best way to analyse the data and it is beyond the available space to discuss all of the possibilities here [49–53]. Certainly we would advise visualizing data graphically as a good way of identifying problems in a dataset [54]. We also find it best to be ruthless and disregard data where there is clearly a problem (e.g. where the control siRNAs in a minority of plates are out of keeping with the rest of the dataset). An integrated facility quickly learns to identify such problems and the information can be fed back to the screeners to modify future screening campaigns if necessary. It is our impression that strong hits or responses usually remain so whatever the type of analysis employed although hit classification for weaker effects can vary enormously depending on the technique employed [55]. From the researchers' point of view, the important question is whether it is possible to employ some measures or analysis of the data that maximizes the return of biologically relevant hits or hits that will in some sense be true. Although there have been a number of additional bioinformatic filters and analyses employed to increase the frequency of selecting genuine biological hits from RNAi data [56, 57], we are of the opinion that ultimately there is probably a limit to what statistics and analysis can provide given the degree of biological variation and pragmatically it might be best to be more generous with thresholds for defining hits and spend more time on further experiments.

'How do I validate my hits?'

The phrase 'hit validation' maybe an example where the choice of language can impact on the perception of screen data and the nature of the subsequent experimentation. Originally from the Pharma chemical screening environment (where it usually referred to the question 'Is the observed effect due to the chemical that I believe to be in that well'), 'validation' has begun to acquire a broader, less precise meaning. It does not imply biological validity (i.e. that the observation is true under all circumstances, or that loss of gene X alone causes the observed phenotype). Moreover, 'validation' isn't a process following a set of immutable rules with a defined end-point; like all research, the observation (the RNAi effect) will be constantly challenged as knowledge evolves and is only as 'valid' as the last experiment. The central question for siRNA screens is 'for a number of candidates from the primary screen, what is the quickest way to become confident that the effects elicited by those reagents are largely or solely the result of the disappearance of the intended protein'. Any approach that increases the confidence would be a 'validation' strategy and perhaps the best such corroborative data would come from techniques other than RNAi, e.g. mouse or zebrafish mutants. Most users agree that one of the best approaches that is amenable to larger numbers of samples is rescue of the phenotype by expression of a siRNA resistant clone [58, 59], although precise modulation of the expression levels of the rescuing clone is essential, but not straightforward, to achieve. Even more straightforward and by far the most popular start point for validation is to show that the effect is produced by more than one oligonucleotide sequence. Such approaches are not without their problems as alternative siRNA reagents with different sequences may be inherently more or less effective than the original siRNA [4, 60] and interpretation of data from these screens can be further complicated by the fact that the relationship between the extent of mRNA decrease, decrease in protein levels and phenotypic change might not be linear for all genes [61]. Validation requires a combination of many different techniques and approaches often in multiwell plate format and it is best if assays designed to classify hits from the primary screen are developed in conjunction with the primary assay. In short, the end of the screen is the start of a very long journey.

'What are the common artefacts?'

The most well-known artefact is that of off-target effects, i.e. unintended silencing of some activity other than the intended target. The mechanisms by which an off-target effect can be produced are many fold and well-reviewed [62–67] and often such effects become apparent only after extensive work. As knowledge about the biology of small and noncoding RNAs increases, it is to be hoped that reagent design can improve further. However, it is possible that no matter what design algorithms are used,

siRNA will always be to some extent 'dirty' reagents, i.e. having effects additional to those for which they were designed. Image-based screens recording multiple parameters per cell might make it easier to identify a true phenotype from apparently similar phenotypes generated by off-target effects [68]. There is an alternative, more positive view of off-target effects: any oligonucleotide eliciting a phenotype solely through an off-target effect is still a reagent that elicits the desired phenotype (and therefore may well be pursuing if the research requires it) but the precise mechanism of its action is unknown, although it probably involves RNA at some point [66].

A less well-described assay artefact we have encountered is the effect of cell death on a number of different readouts primarily luciferase assays and fluorescent protein assays. We observe that RNAi reagents that cause significant cell death can also disproportionately increase the fluorescent or luminescent output of those cells such that the top hits contain an over-representation of siRNAs inducing cell death [69].

'Are siRNA screens reliable and reproducible?'

Increasing numbers of publications clearly attest to the fact that RNAi screening works and identifies novel activities that can be shown by methods other than RNAi to be true. However, our numerous 'corridor conversations' about rumoured screen artefacts or the inability of researchers to reproduce a published observation using similar reagents suggest that there is still some unease about how to interpret results from siRNA screens. The similarity between replicates within any one screen is usually high, i.e. screens are self-consistent. But the frequency with which primary hits can be reproduced in secondary screens is variable (a combination of genuine reproduction failure, identification of the effect as entirely an off-target effect and possibly differences in the analysis of data between genome-wide and smaller scale experiments) [70]. What might make an individual RNAi reagent fail to reproduce from one occasion to another or from one researcher to another? For any one cell, the phenotype resulting from a reduction in the level of a target protein probably depends on:

- the rate and extent of the reduction (and whether the phenotype responds linearly to reduction,

requires a threshold before an effect is observed or shows an alternative more complex dose dependency).
- the degree of functional redundancy of the protein and whether the protein is involved in multiple cellular processes.
- any specific phenotypes elicited by the transfection process *per se*.
- any contributing phenotypic effects of off-target actions of the RNAi reagent.

Within a cell population, the many possible distinct cellular states that co-exist could determine the measured penetrance of the phenotype in that population [28]. All of these aspects could contribute to the day-to-day variability of assays and might therefore contribute to the failure of some primary screen 'hits' to reproduce even in the hands of the same researcher. It would also suggest that hits surviving several rounds of screening are likely to be those showing the least complex of dependencies.

When screening data from different laboratories addressing either similar or apparently identical areas of biology are compared, the overlap can be very disappointing, classic examples being those of screens aimed at understanding the biology of HIV [71, 72] or JAK/STAT signalling [73]. If, as suggested above, each screen provides a precise answer to the question posed, it might be that there are sufficient differences in assay design and execution to explain the poor overlap. In that sense all screen data are 'correct' it is just that the questions posed were different in detail. Such an interpretation therefore focuses attention on the nature of the information sought from a screen. Given the possibility for variability in the phenotypes elicited by individual siRNAs it is unlikely that any one screen will give a complete and full inventory of all of the activities involved in a biological process in either one cell line or across all cell lines. Such data is therefore probably best garnered by multiple screens possibly across many lines or with distinct complementary readouts [70, 74].

These cautious sentiments, however, should not detract from the overwhelming evidence that the careful and intelligent application of genome-wide RNAi screening is an increasingly successful research tool. Even the most cursory of surveys of the literature for 2010 reveals at least 25 publications reporting new and interesting results from genome-wide siRNA, esiRNA or RNAi screens (and an even

larger number of publications reporting genome-wide shRNA screens, miRNA screens or *in vivo* screens in *Drosophila* or *C. elegans*). These screens used assays in both mammalian and *Drosophila* cell systems to identify new components regulating signalling pathways [69, 75–80], mitosis [15, 81, 82], regulation of alternative splicing [83, 84], regulation of cell spreading [85], regulation of apoptosis [86], regulation of basal autophagy [87], regulation of protein secretion [88], innate immunity [89] and stem cell identity [90]. They have addressed disease processes from host–pathogen interactions [91] and viral replication [92–94] to tuberculosis [95] and cancer [81, 96, 97]. In almost all cases, the hits from the primary screen were rescreened in at least one further round of screening using either alternative sequences or individual sequences (if a pool was used) and a subset of these high confidence hits were analysed further using techniques that did not rely on RNAi. Each one of these reports represents the start of new and potentially valuable lines of enquiry and although in many cases only a few genes of interest are highlighted in the publication, it is likely that these hard won datasets contain within them yet more gems waiting to be mined. One conclusion would seem to be that many areas of research can potentially benefit from a well-thought out RNAi screen and in our experience, the most common problem faced by most screeners is that they end up with more 'high confidence' candidates than they can work on at any one time and therefore usually face difficult decisions about prioritizing further work.

What's in the future?

In the short term we anticipate:

- improved and more widespread adoption of smaller screening formats (including spotted arrays) with consequent increased throughput [6].
- Increased understanding of RNA biology [13, 98] [99] and the ways in which its disruption can lead to cellular phenotypes will aid the interpretation of existing and future data.
- Ever increasing demand for, and sophistication of, image based screens and the analysis thereof.
- Increasing use of either real time or dynamic screen formats or assays mimicking more closely *in vivo* settings (3D culture, mixed cell assays).
- Improvements in cell culture and increased awareness of the diversity of cell states and sensitivities probably driven by advances in stem cell research

where considerations of the culture conditions are more to the fore.

We would like to see:

- the evolution of a more cohesive academic screening community with open data sharing and image repositories for screening data to enable virtual screening (e.g. images captured for one screen being re-analysed for an alternative phenotype). Although researchers are often uneasy with the idea of sharing screen data before it has been fully mined, that process can be long and tortuous and we suggest that the community as a whole would benefit from earlier data availability and collaboration. Increased adherence to the proposed MIARE [100] and MIACA standards of curation and reporting would be a prerequisite.
- Publishers recognizing that every screen has its limits and not every screen can be the ultimate description of a particular biology.
- Commercial suppliers to seize the opportunity afforded by the increase in academic screening and collaborate or partner such ventures with increased openness with respect to RNAi reagent design principles, open source or easily accessible software and interfaces, etc.

Key Points

- Like all screens, siRNA screens are based on a central question or assertion. Understanding precisely what the question is, what its limitations are and how the real experimental setup reflects the question is of vital importance if the screen data is to be meaningful.
- siRNAs are imperfect reagents. Fortunately, we have a good deal of knowledge about the ways in which they are imperfect. However, assays based on siRNAs will always suffer from some degree of variance that the researcher has to cope with.
- Be kind to your cells and do not assume that they are a homogenous population.
- Deciding which hits have a high confidence value requires as much or more effort than the primary screen that identified them and should utilize a large range of complementary techniques
- A central or shared core facility is an extremely efficient way to organize screening and enables the creation of repositories of knowledge and data that are of value to the wider community.

References

1. 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.

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

3. Echeverri CJ, Perrimon N. High-throughput RNAi screening in cultured cells: a user's guide. *Nat Rev Genet* 2006;**7**:373–84.

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

5. Sachse C, Krausz E, Kronke A, *et al*. High-throughput RNA interference strategies for target discovery and validation by using synthetic short interfering RNAs: functional genomics investigations of biological pathways. *Methods Enzymol* 2005;**392**:242–77.

6. Erfle H, Neumann B, Liebel U, *et al*. Reverse transfection on cell arrays for high content screening microscopy. *Nat Protoc* 2007;**2**:392–9.

7. Erfle H, Neumann B, Rogers P, *et al*. Work flow for multiplexing siRNA assays by solid–phase reverse transfection in multiwell plates. *J Biomol Screen* 2008;**13**:575–80.

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

9. Carralot JP, Kim TK, Lenseigne B, *et al*. Automated high-throughput siRNA transfection in raw 264.7 macrophages: a case study for optimization procedure. *J Biomol Screen* 2009;**14**:151–60.

10. Boese Q, Leake D, Reynolds A, *et al*. Mechanistic insights aid computational short interfering RNA design. *Methods Enzymol* 2005;**392**:73–96.

11. Abdelrahim M, Safe S, Baker C, *et al*. RNAi and cancer: Implications and applications. *J RNAi Gene Silencing* 2006;**2**: 136–45.

12. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. *Nat Rev Genet* 2009;**10**:94–108.

13. Sharp PA. The centrality of RNA. *Cell* 2009;**136**:577–80.

14. van den Berg BH, McCarthy FM, Lamont SJ, *et al*. Re-annotation is an essential step in systems biology modeling of functional genomics data. *PLoS One* 2010;**5**: e10642.

15. 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.

16. Kittler R, Surendranath V, Heninger A-K, *et al*. Genome-wide resources of endoribonuclease-prepared short interfering RNAs for specific loss-of-function studies. *Nat Meth* 2007;**4**:337–44.

17. Fedorov Y, King A, Anderson E, *et al*. Different delivery methods-different expression profiles. *Nat Methods* 2005;**2**: 241.

18. Portis AM, Carballo G, Baker GL, *et al*. Confocal microscopy for the analysis of siRNA delivery by polymeric nanoparticles. *Micros Res Techniq* 2010;**73**:878–85.

19. Bauer M, Kristensen BW, Meyer M, *et al*. Toxic effects of lipidmediated gene transfer in ventral mesencephalic explant cultures. *Basic Clin Pharmacol Toxicol* 2006;**98**:395–400.

20. Spagnou S, Miller AD, Keller M. Lipidic carriers of siRNA: differences in the formulation, cellular uptake, and delivery with plasmid DNA. *Biochemistry* 2004;**43**:13348–56.

21. Zumbansen M, Altrogge LM, Spottke NU, *et al*. First siRNA library screening in hard-to-transfect HUVEC cells. *J RNAi Gene Silencing* 2010;**6**:354–60.

22. Maurisse R, De Semir D, Emamekhoo H, *et al*. Comparative transfection of DNA into primary and transformed mammalian cells from different lineages. *BMC Biotechnol* 2010;**10**:9.

23. Brazas RM, Hagstrom JE. Delivery of small interfering RNA to mammalian cells in culture by using cationic lipid/polymer-based transfection reagents. *Methods Enzymol* 2005;**392**:112–24.

24. Mocellin S, Provenzano M. RNA interference: learning gene knock-down from cell physiology. *J Transl Med* 2004; **2**:39.

25. Altschuler SJ, Wu LF. Cellular heterogeneity: do differences make a difference? *Cell* 2010;**141**:559–63.

26. Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. *Nature* 2010;**467**:167–73.

27. Levsky JM, Singer RH. Gene expression and the myth of the average cell. *Trends Cell Biol* 2003;**13**:4–6.

28. Sacher R, Stergiou L, Pelkmans L. Lessons from genetics: interpreting complex phenotypes in RNAi screens. *Curr Opin Cell Biol* 2008;**20**:483–9.

29. Yuan TL, Wulf G, Burga L, Cantley LC. Cell-to-cell variability in PI3K protein level regulates PI3K-AKT pathway activity in cell populations. *Curr Biol* 2011;**21**:173–83.

30. Snijder B, Pelkmans L. Origins of regulated cell-to-cell variability. *Nat Rev Mol Cell Biol* 2011;**12**:119–25.

31. Swearingen EA, Fajardo F, Wang X, *et al*. Use of cryopreserved cell aliquots in the high-throughput screening of small interfering RNA libraries. *J Biomol Screen* 2010;**15**: 469–77.

32. Lundholt BK, Scudder KM, Pagliaro L. A Simple technique for reducing edge effect in cell-based assays. *J Biomol Screen* 2003;**8**:566–70.

33. Cooke SL, Temple J, MacArthur S, *et al*. Intra-tumour genetic heterogeneity and poor chemoradiotherapy response in cervical cancer. *Br J Cancer* 2011;**104**:361–8.

34. Gerlinger M, Swanton C. How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine. *Br J Cancer* 2010;**103**:1139–43.

35. Navin N, Krasnitz A, Rodgers L, *et al*. Inferring tumor progression from genomic heterogeneity. *Genome Res* 2010;**20**:68–80.

36. Manning BD. Challenges and opportunities in defining the essential cancer kinome. *Sci Signal* 2009;**2**:pe15.

37. Conrad C, Gerlich DW. Automated microscopy for high-content RNAi screening. *J Cell Biol* 2010;**188**: 453–61.

38. Bjorklund M, Taipale M, Varjosalo M, *et al*. Identification of pathways regulating cell size and cell-cycle progression by RNAi. *Nature* 2006;**439**:1009–13.

39. Carpenter AE, Sabatini DM. Systematic genome-wide screens of gene function. *Nat Rev Genet* 2004;**5**:11–22.

40. Perlman ZE, Slack MD, Feng Y, *et al*. Multidimensional drug profiling by automated microscopy. *Science* 2004;**306**: 1194–8.

41. Tsui M, Xie T, Orth JD, *et al*. An intermittent live cell imaging screen for siRNA enhancers and suppressors of a Kinesin-5 inhibitor. *PLoS One* 2009;**4**:e7339.

42. Gascoigne KE, Taylor SS. Cancer cells display profound intra- and interline variation following prolonged exposure to antimitotic drugs. *Cancer Cell* 2008;**14**:111–22.

43. Slack MD, Martinez ED, Wu LF, *et al*. Characterizing heterogeneous cellular responses to perturbations. *Proc Natl Acad Sci USA* 2008;**105**:19306–11.

44. Bramsen JB, Pakula MM, Hansen TB, *et al*. A screen of chemical modifications identifies position-specific modification by UNA to most potently reduce siRNA off-target effects. *Nucleic Acids Res* 2010;**38**:5761–73.

45. Iorns E, Lord CJ, Turner N, *et al*. Utilizing RNA interference to enhance cancer drug discovery. *Nat Rev Drug Discov* 2007;**6**:556–68.

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

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

48. Malo N, Hanley JA, Cerquozzi S, *et al*. Statistical practice in high-throughput screening data analysis. *Nat Biotechnol* 2006;**24**:167–75.

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

50. Chung N, Zhang XD, Kreamer A, *et al*. Median absolute deviation to improve hit selection for genome-scale RNAi screens. *J Biomol Screen* 2008;**13**:149–58.

51. Zhang XD. An effective method for controlling false discovery and false nondiscovery rates in genome-scale RNAi screens. *J Biomol Screen* 2010;**15**:1116–22.

52. Zhang XD, Heyse JF. Determination of sample size in genome-scale RNAi screens. *Bioinformatics* 2009;**25**:841–4.

53. Zhang XD, Lacson R, Yang R, *et al*. The use of SSMD-based false discovery and false nondiscovery rates in genome-scale RNAi screens. *J Biomol Screen* 2010;**15**:1123–31.

54. Chatterjee S, Firat A. Generating data with identical statistics but dissimilar graphics. *Am Stat* 2007;**61**:248–254.

55. Barrows NJ, Sommer CL, Garcia-Blanco MA, *et al*. Factors affecting reproducibility between genome-scale siRNA-based screens. *J Biomol Screen* 2010;**15**:735–47.

56. Bankhead A, 3rd, Sach I, Ni C, *et al*. Knowledge based identification of essential signaling from genome-scale siRNA experiments. *BMC Syst Biol* 2009;**3**:80.

57. Wiles AM, Ravi D, Bhavani S, *et al*. An analysis of normalization methods for Drosophila RNAi genomic screens and development of a robust validation scheme. *J Biomol Screen* 2008;**13**:777–84.

58. Kittler R, Pelletier L, Ma C, *et al*. RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells. *Proc Natl Acad Sci USA* 2005;**102**:2396–401.

59. Kittler R, Pelletier L, Heninger A-K, *et al*. Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nat Cell Biol* 2007;**9**:1401–12.

60. Parsons BD, Schindler A, Evans DH, *et al*. A direct phenotypic comparison of siRNA pools and multiple individual duplexes in a functional assay. *PLoS One* 2009;**4**:e8471.

61. Martin SE, Jones TL, Thomas CL, *et al*. Multiplexing siRNAs to compress RNAi-based screen size in human cells. *Nucleic Acids Res* 2007;**35**:e57.

62. Fedorov Y, Anderson EM, Birmingham A, *et al*. Off-target effects by siRNA can induce toxic phenotype. *RNA* 2006;**12**:1188–96.

63. Jackson AL, Burchard J, Schelter J, *et al*. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA* 2006;**12**:1179–87.

64. Klinghoffer RA, Magnus J, Schelter J, *et al*. Reduced seed region-based off-target activity with lentivirus-mediated RNAi. *RNA* 2010;**16**:879–84.

65. Kulkarni MM, Booker M, Silver SJ, *et al*. Evidence of off-target effects associated with long dsRNAs in Drosophila melanogaster cell-based assays. *Nat Meth* 2006;**3**:833–8.

66. Sudbery I, Enright A, Fraser A, *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.

67. Aleman LM, Doench J, Sharp PA. Comparison of siRNA-induced off-target RNA and protein effects. *RNA* 2007;**13**:385–95.

68. MacDonald ML, Lamerdin J, Owens S, *et al*. Identifying off-target effects and hidden phenotypes of drugs in human cells. *Nat Chem Biol* 2006;**2**:329–37.

69. Ribeiro PS, JosuÈ F, Wepf A, *et al*. Combined functional genomic and proteomic approaches identify a PP2A complex as a negative regulator of hippo signaling. *Mol Cell* 2010;**39**:521–34.

70. DasGupta R, Nybakken K, Booker M, *et al*. A case study of the reproducibility of transcriptional reporter cell-based RNAi screens in Drosophila. *Genome Biol* 2007;**8**:R203.

71. Cherry S. What have RNAi screens taught us about viral-host interactions? *Curr Opin Microbiol* 2009;**12**:446–52.

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

73. Muller P, Boutros M, Zeidler MP. Identification of JAK/STAT pathway regulators–insights from RNAi screens. *Semin Cell Dev Biol* 2008;**19**:360–9.

74. Liu T, Sims D, Baum B. Parallel RNAi screens across different cell lines identify generic and cell type-specific regulators of actin organization and cell morphology. *Genome Biol* 2009;**10**:R26.

75. Valanne S, Myllymaki H, Kallio J, *et al*. Genome-wide RNA interference in Drosophila cells identifies G protein-coupled receptor kinase 2 as a conserved regulator of NF-kappaB signaling. *J Immunol* 2010;**184**:6188–98.

76. Saj A, Arziman Z, Stempfle D, *et al*. A combined ex vivo and in vivo RNAi screen for notch regulators in Drosophila reveals an extensive notch interaction network. *Dev Cell* 2010;**18**:862–76.

77. Mourikis P, Lake RJ, Firnhaber CB, *et al*. Modifiers of notch transcriptional activity identified by genome-wide RNAi. *BMC Dev Biol* 2010;**10**:107.

78. Kockel L, Kerr KS, Melnick M, *et al*. Dynamic switch of negative feedback regulation in Drosophila Akt-TOR signaling. *PLoS Genet* 2010;**6**:e1000990.

79. Kallio J, Myllymaki H, Gronholm J, *et al*. Eye transformer is a negative regulator of Drosophila JAK/STAT signaling. *FASEB J* 2010;**24**:4467–79.

80. Dekanty A, Romero NM, Bertolin AP, *et al*. Drosophila genome-wide RNAi screen identifies multiple regulators

of HIF-dependent transcription in hypoxia. *PLoS Genet* 2010;**6**:e1000994.

81. Leber B, Maier B, Fuchs F, *et al*. Proteins required for centrosome clustering in cancer cells. *Sci Transl Med* 2010; **2**:33ra38.

82. 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.

83. Moore MJ, Wang Q, Kennedy CJ, *et al*. An alternative splicing network links cell-cycle control to apoptosis. *Cell* 2010;**142**:625–36.

84. Ashton-Beaucage D, Udell CM, Lavoie H, *et al*. The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in Drosophila. *Cell* 2010; **143**:251–62.

85. D'Ambrosio MV, Vale RD. A whole genome RNAi screen of Drosophila S2 cell spreading performed using automated computational image analysis. *J Cell Biol* 2010; **191**:471–8.

86. Wang Q, Silver PA. Genome-wide RNAi screen discovers functional coupling of alternative splicing and cell cycle control to apoptosis regulation. *Cell Cycle* 2010;**9**:4419–21.

87. Lipinski MM, Hoffman G, Ng A, *et al*. A genome-wide siRNA screen reveals multiple mTORC1 independent signaling pathways regulating autophagy under normal nutritional conditions. *Dev Cell* 2010;**18**:1041–52.

88. Wendler F, Gillingham AK, Sinka R, *et al*. A genome-wide RNA interference screen identifies two novel components of the metazoan secretory pathway. *EMBO J* 2010;**29**: 304–14.

89. Kuttenkeuler D, Pelte N, Ragab A, *et al*. A large-scale RNAi screen identifies Deaf1 as a regulator of innate immune responses in Drosophila. *J Innate Immun* 2010;**2**: 181–94.

90. 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.

91. Akimana C, Al-Khodor S, Abu Kwaik Y. Host factors required for modulation of phagosome biogenesis and proliferation of Francisella tularensis within the cytosol. *PLoS One* 2010;**5**:e11025.

92. Moser TS, Sabin LR, Cherry S. RNAi screening for host factors involved in Vaccinia virus infection using Drosophila cells. *J Vis Exp* 2010;**42**:2137.

93. Smith JA, White EA, Sowa ME, *et al*. Genome-wide siRNA screen identifies SMCX, EP400, and Brd4 as E2-dependent regulators of human papillomavirus oncogene expression. *Proc Natl Acad Sci USA* 2010;**107**:3752–7.

94. Karlas A, Machuy N, Shin Y, *et al*. Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* 2010;**463**:818–22.

95. Kumar D, Nath L, Kamal MA, *et al*. Genome-wide analysis of the host intracellular network that regulates survival of Mycobacterium tuberculosis. *Cell* 2010;**140**:731–43.

96. Sudo H, Tsuji AB, Sugyo A, *et al*. Knockdown of COPA, identified by loss-of-function screen, induces apoptosis and suppresses tumor growth in mesothelioma mouse model. *Genomics* 2010;**95**:210–16.

97. Chen S, Blank JL, Peters T, *et al*. Genome-wide siRNA screen for modulators of cell death induced by proteasome inhibitor bortezomib. *Cancer Res* 2010;**70**:4318–26.

98. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev* 2009;**23**:1494–504.

99. Faghihi MA, Kocerha J, Modarresi F, *et al*. RNAi screen indicates widespread biological function for human natural antisense transcripts. *PLoS One* 2010;**5**:e13177.

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