FRONTIERS IN RNAi

VOLUME 1

Editors:

Ralph A. Tripp

Jon M. Karpilow

RNAi GLOBAL

Bentham e Books

Frontiers in RNAi
(Volume 1)

Edited By

Ralph A. Tripp

Department of Infectious Diseases
University of Georgia
Athens, GA
USA

&

Jon M. Karpilow

Biosciences Division
Thermo Fisher Scientific
Lafayette, CO
USA

CONTENTS

FOREWORD

Great technologies follow similar roads to maturity. A breakthrough discovery leads to rapid early adoption which (more often than not) precedes the identification of limitations. Eventually, solutions are identified and the technology, with all its strengths and weaknesses, is put to the test through wide-spread implementation. This process of maturation takes years and in this regard, RNA Interference (RNAi) is no different than yeast two-hybrid, chip-based gene expression profiling, and monoclonal antibodies.

Once the scientific community overcame the sense of excitement associated with the discovery of non-coding RNA-based posttranscriptional gene regulation, early adopters between 2000-2005 immediately gained insights into the technology's limitations. Not all siRNAs silenced with equal efficiency. Some siRNA designs activated the innate immune response. Other siRNAs exhibited off-target effects and could induce false positive phenotypes. To some, these challenges might have appeared overly daunting, but given the potential of RNAi, researchers in both academic and industrial settings pressed to find solutions. Algorithms were designed to address issues associated with functionality and position specific chemical modifications were adopted to minimize off-target effects and activation of the innate immune system. Self-delivering siRNA were developed to address the need to deliver reagents to cell lines that were refractory to lipid-mediated delivery and viral-based gene silencing was constructed to facilitate experiments in systems that required extended periods of knockdown.

While challenges still persist (*e.g.*, therapeutic delivery), overcoming the first wave of technical hurdles has been paramount to expansion of RNAi into new fields of interest. As evidenced in this ebook, continued mechanistic studies are now combined with a host of new developments where RNAi technology is being merged with miniaturized screening platforms, field-specific database infrastructures, host-pathogen interaction mapping, and new (3D) tissue culture models. At the same time, the technology is reaching into more applied fields. RNAi screening is slowly becoming a module of synthetic biology and cell line engineering, and a key component of therapeutic drug repurposing. This expansion of the technology mirrors the growth and development observed with other platforms (*e.g.*, NG Sequencing, PCR) and is indicative that over the course of little more than a decade, RNAi has grown up!

Devin Leake
Vice President of Research and Development
Gen9 Inc.
Cambridge, MA
USA

PREFACE

RNA interference (RNAi), has developed into an important tool for gene function elucidation by leveraging this endogenous gene-silencing pathway with well-designed triggers for the systematic silencing of particular genes. Incorporation of RNAi into high throughput screening platforms has led to important discoveries in fields such as cancer biology and pathogenesis, successes driven by establishment of best-practices and developments in instrumentation, analysis, and the growing sophistication of cell-based assays. The following chapters represent some of the most important considerations and recent developments in RNAi screening technology. It is certainly our hope that this compendium contains something for everyone, both novice and experienced researchers alike.

The topics in this collection range from recent automation platforms for ultra-high throughput screens (see Chapters 2 and 8) to novel applications of this gene modulation technology to cells grown in 3D culture (see Chapter 9). Chapter 1 provides a review of RNAi as a research tool as well as experimental and bioinformatics approaches to reduce and identify off-target effects in high-throughput screening data. Important considerations regarding deposition of large data sets in public repositories are presented in Chapter 3. Pooled shRNA screening methods are discussed in Chapter 4. Chapters 5 and 6 describe how RNAi technologies are being used to interrogate the host-pathogen interface. RNAi screening in difficult-to-transfect immune cells is addressed in Chapter 7 and applications beyond target discovery are presented in Chapters 10 and 11.

While it is not evident at first glance, a research alliance underlies all the contributions detailed in this book. All authors are affiliated with a lab that participates in the RNAi Global Initiative, a world-wide association of biomedical researchers established to broaden and accelerate the utility and application of RNAi technology. Founded in 2005 with eleven member institutions, the organization focused on developing and disseminating information on the basics of RNAi screening. As the organization grew and technology developed, focus shifted toward more complex, often biological, challenges.

Nine years after its inauguration, the RNAi Global Initiative now includes over 60 academic institutes spread across five continents. Members continue to share their findings in a range of fields (stem cell biology, cancer, and host-pathogen interactions) and as evidenced by this book, collaborate in the development of new bioinfomatic tools, screening methods, and programs in applied biology. The longevity of this alliance is the result of continued contributions on the part of its membership and is a testament to the importance of collaboration in building a cohesive scientific community.

Ralph Tripp
University of Georgia
College of Veterinary Medicine
Department of Infectious Diseases
Athens, GA, USA

&

Jon Karpilow
Thermo Fisher Scientific
Lafayette, CO, USA

List of Contributors

Abhijeet Bakre
University of Georgia, College of Veterinary Medicine, Department of Infectious Diseases, 111 Carlton Street, Athens, GA 30602, USA

Aideen Long
Department of Clinical Medicine, Trinity College Dublin, Dublin, Ireland

Amanda Birmingham
Dharmacon, part of GE Healthcare, 2650 Crescent Dr., Suite, 100, Lafayette, CO 80026, USA

Andreas Kaufmann
LMSC, ETH, Schafmattstrasse 18, 8093 Zurich, Switzerland

Andrew G. Bean
CSIRO Australian Animal Health Laboratory, Geelong 3220, Victoria, Australia

Anja Van Brabant Smith
Dharmacon, part of GE Healthcare, 2650 Crescent Dr., Suite, 100, Lafayette, CO 80026, USA

Annaleen Vermeulen
Dharmacon, part of GE Healthcare, 2650 Crescent Dr., Suite, 100, Lafayette, CO 80026, USA

Arvind Rao
Department of Bioinformatics and Computational Biology, Division of Quantitative Sciences, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77054, USA

Cameron R. Stewart
CSIRO Australian Animal Health Laboratory, Geelong 3220, Victoria, Australia

Caroline E. Shamu
ICCB-Longwood Screening Facility, Harvard Medical School, 250 Longwood Avenue, SGM Building Room 604, Boston, MA 02115-5731, USA

Esther E. Schmidt
German Cancer Research Center, Division Signaling and Functional Genomics and Heidelberg University, Department for Cell and Molecular Biology, Medical Faculty Mannheim, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany

Geoffrey Bartholomeusz
Department of Experimental Therapeutics, Division of Cancer Medicine, The University of Texas M.D. Anderson Cancer Center, Houston, TX. 77054, USA

HiChul Kim
Institute Pasteur Korea, 696 Sampyeong-dong, Bundang-gu, Gyeonggi-do, 463-400, South Korea

Iain D.C. Fraser
Signaling Systems Unit, Laboratory of Systems Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA.

Jennifer A. Smith
ICCB-Longwood Screening Facility, Harvard Medical School, 250 Longwood Avenue, SGM Building Room 604, Boston, MA 02115-5731, USA

JinY. Kim	Institute Pasteur Korea, 696 Sampyeong-dong, Bundang-gu, Gyeonggi-do, 463-400, South Korea
Jinyeoung Heo	Institute Pasteur Korea, 696 Sampyeong-dong, Bundang-gu, Gyeonggi-do, 463-400, South Korea
John Stambas	Deakin University School of Medicine, Waurn Ponds, 3216, Victoria, Australia
John W. Lowenthal	CSIRO Australian Animal Health Laboratory, Geelong 3220, Victoria, Australia
Jon M. Karpilow	Thermo Fisher Scientific, 2650 Crescent Dr., Suite, 202, Lafayette, CO 80026, USA
Karol Kozak	LMSC, ETH, Schafmattstrasse 18, 8093 Zurich, Switzerland
Kaylene J. Simpson	Victorian Centre for Functional Genomics, Peter MacCallum Cancer Centre, East Melbourne, Victoria, 3002, Australia
Kristie A. Jenkins	CSIRO Australian Animal Health Laboratory, Geelong 3220, Victoria, Australia
Leonard H. Izzard	Deakin University School of Medicine, Waurn Ponds, 3216, Victoria, Australia
Mark L. Tizard	CSIRO Australian Animal Health Laboratory, Geelong 3220, Victoria, Australia
Michael A. E. Hansen	Institute Pasteur Korea, 696 Sampyeong-dong, Bundang-gu, Gyeonggi-do, 463-400, South Korea
Michael Boutros	German Cancer Research Center, Division Signaling and Functional Genomics and Heidelberg University, Department for Cell and Molecular Biology, Medical Faculty Mannheim, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany
Michael Freeley	Department of Clinical Medicine, Trinity College Dublin, Dublin, Ireland
Michael S. Banos	Dharmacon, part of GE Healthcare, 2650 Crescent Dr., Suite, 100, Lafayette, CO 80026, USA
Namyoul Kim	Institute Pasteur Korea, 696 Sampyeong-dong, Bundang-gu, Gyeonggi-do, 463-400, South Korea
Olivia Perwitasari	University of Georgia, College of Veterinary Medicine, Department of Infectious Diseases, 111 Carlton Street, Athens, GA 30602, USA
Paula Brooks	University of Georgia, College of Veterinary Medicine, Department of Infectious Diseases, 111 Carlton Street, Athens, GA 30602, USA
Ralph A. Tripp	University of Georgia, College of Veterinary Medicine, Department of Infectious Diseases, 111 Carlton Street, Athens, GA 30602, USA

Roderick L. Beijersbergen   Division of Molecular Carcinogenesis, Cancer Genomics Center and Cancer Systems Biology Center, The Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands

S. Mark Tompkins   Department of Infectious Diseases, University of Georgia, Athens, Georgia, USA

Sabine van der Sanden   Centers for Disease Control and Prevention, 1600 Clifton Road NE, Atlanta, GA, 30333, USA

Samantha Barichievy   Gene Expression and Biophysics Group, Synthetic Biology–Emerging Research Area, Council for Scientific and Industrial Research, Pretoria, South Africa

Sarah B. Anderson   Challenge Technology Inc., 4950 Ward Road, Wheat Ridge, CO 80033, USA

Sean M. Johnston   ICCB-Longwood Screening Facility, Harvard Medical School, 250 Longwood Avenue, SGM Building Room 604, Boston, MA 02115, USA

Sinu P. John   Signaling Systems Unit, Laboratory of Systems Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA.

Steven Oberste   Centers for Disease Control and Prevention, 1600 Clifton Road NE, Atlanta, GA, 30333, USA

TaeKyu Lee   Institute Pasteur Korea, 696 Sampyeong-dong, Bundang-gu, Gyeonggi-do, 463-400, South Korea

Timothy J. Doran   CSIRO Australian Animal Health Laboratory, Geelong 3220, Victoria, Australia

Veronica Soloveva   Institute Pasteur Korea, 696 Sampyeong-dong, Bundang-gu, Gyeonggi-do, 463-400, South Korea

Weilin Wu   University of Georgia, College of Veterinary Medicine, Department of Infectious Diseases, 111 Carlton Street, Athens, GA 30602, USA

Yong-Jun Kwon   Institute Pasteur Korea, 696 Sampyeong-dong, Bundang-gu, Gyeonggi-do, 463-400, South Korea

ACKNOWLEDGEMENTS

CHAPTER 1

RNAi and Off-Target Effects

Amanda Birmingham[1],[†], Andreas Kaufmann[2],[†] and Karol Kozak[2],[3],[†],[*]

[1]*Dharmacon, part of GE Healthcare, 2650 Crescent Dr., suite 100, Lafayette, CO 80026 USA;* [2]*LMSC, ETH, Schafmattstrasse 18, 8093 Zurich, Switzerland and* [3]*Medical Faculty, Technical University Dresden, Photenhauerstr 41, 01307 Dresden, Germany*

Abstract: RNA interference (RNAi) is widely used in high-throughput reverse genetics screens for basic research and drug discovery. However, this technique is known to produce "off-target" effects, or phenotypic results caused by an RNAi reagent's knockdown of unintended genes rather than the gene of interest. Off-target effects are regulated through multiple mechanisms within the cell, and their presence can greatly complicate the interpretation of experimental data. We review the biology of off-targeting and discuss both bioinformatics and experimental approaches to reduce RNAi reagents' off-target effects. Since such techniques cannot completely eliminate off-target effects, we also discuss analysis methods developed to identify off-target effects in RNAi screening data and, in some cases, even leverage them to uncover novel biological functionality.

Keywords: bioinformatics, miRNA, mRNA, off-target effects, off-targets, pools, RNAi, screening, seeds, silencing, siRNA.

INTRODUCTION

RNA interference (RNAi) is a widely used tool for functional genetic studies and genetic screens in a number of organisms. In addition, RNAi-based technologies seem to have significant potential for the treatment of human disease [1, 2]. RNAi targets messenger RNA (mRNA) for degradation or translational attenuation and is mediated by small interfering RNA (siRNA) or microRNA (miRNA) bound to the RNA-induced silencing complex (RISC). Initially, it was believed that this gene silencing is highly specific due to required near-perfect complementarity between the siRNA and its target transcript [3, 4]. However, siRNA-mediated gene silencing can lead to widespread, unintended effects that severely complicate the interpretation of experimental results or disease treatment [5-7]. Non-target-specific side-effects caused by the introduction of siRNAs, such as broad changes

*****Corresponding author Karol Kozak:** Medical Faculty, Technical University Dresden, Photenhauerstr 41, 01307 Dresden, Germany; Tel: 0351 8823 248; E-mail: karkoz1@gmx.de
[†]These authors contributed equally to this work.

in gene expression profiles caused by cationic lipids present in transfection reagents [8] and a non-specific interferon response [9, 10], can be successfully addressed by optimizing delivery methods and siRNA design algorithms. However, RISC-dependent dis-regulation of unintended transcripts containing full or partial sequence complementarity [5-7, 11-14] has proven more challenging to control. This dysregulation leads to sequence-dependent false positive results, or so-called off-target effects. The use of bioinformatics tools alone has been unsuccessful in preventing off-target effects, but a combination of bioinformatics and chemical modifications of the siRNA has been found to reduce unintended targeting, and bioinformatics analysis is increasingly able to identify off-target effects in experimental results [15-21].

Mechanisms of RNAi

RNAi leads either to cleavage of the target mRNA or to its translational repression and/or degradation.

The Cleavage-Based Pathway

Cleavage of the target mRNA is facilitated by small double-stranded RNA molecules having complete or near-complete complementarity to their target. Such molecules can be produced when exogenous (500-1000 nucleotide long) double-stranded RNA (dsRNA) is processed in the cytoplasm by Dicer, an RNAse III enzyme, into siRNAs-approximately 21-nucleotide-long double-stranded RNA fragments with a two-nucleotide overhang at both 3′ ends [22, 23]. Alternatively, synthetic siRNAs that mimic Dicer cleavage products can be directly introduced into cells [3], thus bypassing target-nonspecific inhibitory mechanisms, such as the interferon response, elicited by longer dsRNA in mammalian cells [24]. The siRNA duplexes consist of a guide or antisense strand, which is of complementary sequence to the target mRNA, and a passenger or sense strand. Dicer facilitates RISC formation and entry of the siRNA into RISC [25]. Argonaute 2 (Ago2), a functional subunit of RISC, is required to cleave the passenger strand and unwind the guide strand from the siRNA duplex. The cleaved passenger strand is then released and the guide strand is loaded into the now-activated RISC [26-29]. It is not fully understood how the activated RISC locates the target mRNA within the cell, but binding of the target mRNA occurs by complementary base pairing and Dicer cleaves the mRNA at the point opposite bases 10 and 11 of the guide [30]. The guide strand is not affected; thus, RISC can cleave multiple copies of the target mRNA, leading to effective gene silencing.

Although siRNAs and shRNAs (short hairpin RNAs) produce a similar functional outcome, shRNAs differ significantly in their structure and processing before entering the cleavage-based pathway of RNAi. Experimentally, shRNAs can be produced from plasmids or viral vectors [31] and have a hairpin structure [32]. In the nucleus, shRNAs are expressed as pri-shRNAs (primary shRNAs) and converted into pre-shRNAs (precursor shRNAs) by the RNAase III nuclease Drosha [33-35]. Exportin-5 (Exp5) then binds to and transfers the pre-shRNA to the cytoplasm [36], where Dicer processes it to a functional siRNA [37].

The Non-Cleavage-Based Pathway

The non-cleavage-based RNAi pathway that leads to translational repression and mRNA degradation is mainly facilitated by miRNAs, which comprise a large family of endogenously encoded, non-protein-coding RNAs approximately 21 nucleotides long. In many organisms, miRNAs are vital for post-transcriptional modulation of gene expression. In this non-cleavage RNAi pathway, pri-miRNA molecules of several kilobases in length are transcribed in the nucleus by RNA polymerase II [38]. Drosha, together with DGCR8, which is essential for the recognition of pri-miRNA, processes the pri-miRNA transcripts into approximately 70 nucleotide-long hairpin pre-miRNAs [33, 38] that are actively exported from the nucleus by exportin-5 [36, 39, 40]. In the cytoplasm, the pre-miRNAs are again processed by Dicer into mature miRNA duplexes [22, 41, 42]. Mature miRNAs function together with argonaute, as catalytic component, and GW182, as scaffold protein, as ribonuleoprotein complexes, miRISCs (miRNA-induced silencing complexes), that regulate protein synthesis by binding to target mRNAs.

In animals, miRNAs generally bind to the 3′-untranslated region (3′UTR) of target mRNAs and attenuate protein synthesis by either inhibiting translation or by inducing mRNA degradation through deadenylation, although it has also been observed that some miRNAs target 5′UTR and coding regions of mRNAs and/or may even enhance translation [43, 44]. Binding of a miRNA to its target mRNA does not require a fully complementary sequence. Instead, association is usually mediated by a six-to-eight base 5′-seed region that provides most of the base-pairing specificity (reviewed in [45, 46]). This matching may in some cases accommodate mismatches and bulged nucleotides [47]. Other partial complementarity such as that conferred by "centered seeds" [48] and "3′-compensatory seeds" [45] has been shown to direct targeting in some situations. Because these targeting mechanisms require such limited pairing, a single miRNA

may regulate multiple mRNAs, and conversely a single mRNA may be regulated by multiple miRNAs [49, 50].

Mechanisms of RISC-Dependent Off-Targeting

Gene silencing by RNAi was initially believed to be highly sequence-specific, only tolerating a few mismatches between the guide strand and the target mRNA [4, 51-53]. However, subsequent genome-wide analysis of siRNA efficacy and specificity revealed that siRNAs containing as few as eleven contiguous nucleotides of identity are in some cases sufficient to silence non-targeted genes [5]. Furthermore, since RISC-dependent off-targeting is driven by the strand of the siRNA duplex that has been loaded into RISC, off-target effects can be caused by erroneous loading of either the sense or the antisense strand [5, 14].

The primary cause of off-target effects appears to be siRNAs functioning in translational repression by a mechanism that is similar to that of endogenous miRNAs [11]. When siRNAs silence off-targeted genes through the non-cleavage-based pathway, full complementarity of the siRNA to the off-targeted mRNA is not required. A single complementary binding site for the hexamer or heptamer seed region (positions 2-7 or 2-8) of the antisense strand of the siRNA within the 3′-untranslated region (UTR) of the mRNA can be sufficient for repression [7, 12, 14]. Silencing by this mechanism generally leads to modest changes in expression, with a magnitude of two-fold or less [13].

Although the majority of observed off-target effects appear to be due to siRNAs acting through a miRNA-like mechanism, off-targeting through the cleavage based-pathway can occur when the RISC-loaded strand has perfect or near-perfect complementarity to an unintended mRNA. Since cleavage-based silencing is generally stronger than non-cleavage-based, the magnitude of such off-target effects can rival that of the intended effect, knocking down expression of the unintended target almost completely. Off-target effects of certain siRNAs can also be caused by complementarity with an unintended target that is complete save for G:U wobbles in position 9 and 10 of the anti-sense strand; while wobbles at these key positions likely prevent cleavage of the unintended target, protein levels for that target have been shown to nonetheless suffer considerable reduction [54, 55].

Effects of off-Targeting

When present but unaccounted-for in experimental data, off-targets lead to misinterpretation of biological results; in high-throughput screens, off-targets are particularly problematic because they inflate the false-positive rate of the screen,

thus increasing cost and time spent in validation. While side-effects due to lipid usage and interferon activation can be adequately controlled in almost all cases, some screens have been found to be particularly vulnerable to RISC-dependent off-targets: for example, Lin *et al*. [7] reported a dual luciferase reporter screen designed to identify novel effectors of the HIF-1 pathway in which all three top hits were found to be caused by off-target effects, and Sudbery *et al*. [56] described an alamarBlue screen for sensitivity to TRAIL-induced apoptosis in which off-targets outnumbered true hits by two to one in their top 20 siRNAs. Schultz *et al*. [57] hypothesize that off-target effects may overwhelm on-target ones in assays that measure particularly dose-sensitive pathways such as TGF-β. However, it is not known how many or which assays fall into this category. Even in generally less-affected screens such as those based on dsRNAs, off-targets can considerably inflate the number of false positives [58]. For these reasons, considerable effort has been devoted to the minimization of off-target effects by optimizing reagent design, adoption of assays that are less sensitive to such effects, and incorporation of post-experimental data analysis.

PREVENTING OFF-TARGETS

Reagent Design

Because of the considerable costs and problems posed by off-target effects, reagent designers employ a wide array of strategies to reduce reagent-induced off-targeting. A sampling of the most successful approaches to date is highlighted below.

Avoiding Toxicity and Interferon Effects

While less problematic than RISC-dependent off-target effects, toxic effects caused by an siRNA's activation of the innate immune system must still be addressed during reagent design. Exogenous double-stranded RNA is prone to recognition by Toll-like receptors (TLRs). TLR7 and TLR8, in particular, recognize synthetic siRNAs in a sequence-dependent manner [59, 60] and seem to prefer GU-rich subset of sequences, so siRNA target sites lacking GU-rich regions are preferable. Avoidance of motifs such as 5′-GTCCTTCAA-3′, 5′-TGTGT-3′, and tetrad-forming poly-(G) stretches also helps to overcome Toll-like receptor recognition [59, 61]. Furthermore, sequence motifs such as UGGC can lead to toxicity through RISC-dependent mechanisms [62], and excluding such sequence-specific triggers to cellular stressors where possible can reduce widespread off-target dysregulation.

When designing shRNAs, additional consideration must be given to the choice of scaffold and promoter. microRNA-derived shRNA scaffolds appears to reduce cellular toxicity compared to simple hairpins, presumably because of more efficient processing by the Microprocessor and Dicer [63-66]. Additionally, extremely strong promoters such as the human U6 small nuclear promoter can create toxicity, perhaps due to overexpression of the shRNA that overwhelms the cell's RNAi machinery [67]. They should therefore be avoided in favor of more moderate promoters.

Designing for Efficient RISC Entry of the Antisense Strand

In the early steps of RISC formation, both strands of the siRNA duplex are initially incorporated into the complex. The siRNA strand with the less thermodynamically stable 5′-end is preferentially utilized as the antisense strand [68, 69] and the remaining (sense) strand is subsequently cleaved and released [29, 70, 71]. Because of this, siRNA guide strands with relatively low 5′-end free energy (such as those having at least three (A/U)s in the final five nucleotides at the 3′-end of the sense strand, are preferable in order to prevent off-targeting driven by unintended RISC-loading of the sense strand [72]. A different approach to prevent passenger strand entry uses an antisense strand that is complemented with two shorter, 9 - 13 nucleotide-long sense strands. Because the RISC is unable to incorporate the bipartite sense strands, such a three-stranded construct completely eliminates unintended mRNA targeting by the sense strand [19]. In addition to sequence- and structure-based design parameters, synthetic siRNAs (but not shRNAs expressed from plasmids or viral vectors) can be chemically modified to prevent loading of the sense strand into RISC (see below).

Avoiding Near-Perfect Complementarity with Unintended Targets

As described above, mRNAs other than intended targets that exhibit perfect or near-perfect sequence complementarity with the siRNA can suffer unintended cleavage. Considerable success in predicting such cleavage-based off-targets is possible, since this is largely analogous to predicting on-target cleavage-based functionality of an RNAi reagent. Generally, such off-targets are predicted by checking for perfect or near-perfect sequence matches against the targeted transcriptome; they are then avoided by choosing only reagents that do not have 17 or more bases of sequence complementarity to any unintended mRNA target [73]. However, this approach can sometimes lead to false positive off-target predictions if the anticipated off-targeted transcript is not expressed in the cell line or tissue to be assayed.

Considering Reagent Seed Sequence

Prediction of a reagent's seed-based off-targets has proven much more challenging than prediction of its cleavage-based off-targets. While it is the case that most off-targeted genes have 3' UTR seed complementarity to the off-targeting reagent, it is *not* the case that most genes with such seed complementarity are in fact off-targeted [12]. Since seed-based off-targets are understood to be produced by an miRNA-like mechanism, miRNA target prediction tools (such as TargetScan, miRanda, and DIANA microT, among others) are frequently employed to predict siRNA-mediated off-targets [74-76]. Such tools are usually based on hexamer, heptamer, or octamer seeds, and are usually limited to identifying reverse complements in the 3′ UTR of the transcriptome. Sometimes they also include criteria of limited relevance to off-target prediction applications, like conservation filters. While these approaches perform better than simple seed-match counting, they generally lead to massive over-prediction of targets [77]. Nonetheless, since the precise mechanism of seed-based targeting is an area of intense scientific interest, much research is underway that will lead to improvement in these techniques [78-84].

Until these techniques are perfected, reagent designers will continue to rely on heuristic approaches to reducing an siRNA's potential for miRNA-like off-targeting. Such methods include avoiding siRNAs whose antisense strands contain seeds already believed to cause miRNA-like targeting in the system of interest (such as those shared by known miRNAs in that system) [73] or preferring siRNAs with low-stability seeds, since these have been shown to have reduced off-target activity [85]. Other approaches examine the potential space of seed-based off-targets for a given antisense seed by evaluating the frequency of complements to it in the 3' UTRs of the transcriptome of interest. It has been shown that siRNAs with low seed complement frequencies (SCFs)-and thus a low number of potential seed-matched transcripts-generally off-target a smaller number of unintended transcripts [86]. Conversely, it has also been suggested that siRNAs whose seeds have a particularly high abundance of potential target sites in the transcriptome may not function efficiently as miRNAs due to dilution of their ability to repress any individual unintended transcript [74]. Thus, designers may choose to favor siRNAs containing low and/or high (but not medium) frequency seeds in order to lower their off-targeting potential.

Chemical Modifications

The most promising strategies so far to reduce off-target effects rely on chemical modification of the siRNA, although the utility of such modifications is limited to exogenous siRNAs and not those produced through shRNA processing. Several attempts have been made to chemically modify the guide strand of siRNAs to enhance on-target specificity and decrease off-target recognition. In addition, certain modifications of the passenger strand prevent it from incorporation into RISC.

Chemical Modifications of the Guide Strand

The seed region of the guide strand initiates recognition of the target mRNA when presented to RISC [87], and the thermodynamic stability of this interaction correlates positively with off-targeting results [11]. Thus, destabilizing seed-target interactions is thought to improve on-target specificity and minimize off-target effects by requiring base-pairing of additional residues outside the seed region for the siRNA to bind to a target mRNA. Consistent with this idea, the incorporation of an unlocked nucleic acid (UNA) monomer, a strongly destabilizing modification, at position 7 of the guide strand has been shown to reduce off-targets with a minimal decrease in siRNA efficiency [17]. A similar destabilizing effect has been achieved by substituting residues 1 - 8 of the guide strand with DNA [16]. 2'-O-methyl modifications of position 2 of the guide strand also reduce off-targeting, presumably due to size constraints on Ago2's accommodation of such a modification [15], and a similar effect has been found when introducing an additional nucleotide at position 2 of the antisense strand that forms a bulge when binding the 19 nucleotide target sequence [18].

Chemical Modifications of the Passenger Strand

In addition to the guide strand, the passenger strand can also contribute significantly to off-target effects [5, 14]. Thus, blocking sense strand incorporation into RISC should significantly decrease off-target silencing. As noted above, RISC preferentially loads the siRNA strand that has the less thermodynamically stable 5′ end. Accordingly, selective thermodynamic stabilization of the passenger strand 5′ end by incorporation of locked nucleic acid (LNA) disfavors passenger strand incorporation into RISC [20]. Similarly, blockage of the passenger strand 5′ phosphate by 5′-O-methyl modification [21], 5-nitroindole modification at position 15 of the passenger strand [88], or chemically modified 3′ overhangs of either strand [89] also reduces sense strand selection and thus off-target effects.

Pooling of siRNAs

As both siRNA and miRNA effects are concentration-dependent [5, 53], off-target effects can be reduced by pooling several different synthetic siRNAs targeting the same gene. The strategy behind this approach is to minimize off-target effects from each individual siRNA while maintaining an effective concentration of the total siRNAs for the intended target: the total concentration of siRNA used is divided up amongst multiple individuals, each of which presumably have distinct off-target signatures, and therefore each individual is less capable of causing off-targets while still contributing to effective on-target knockdown. A similar effect is achieved when using esiRNAs, which are mixtures of siRNA oligos generated by cleaving long double-stranded RNA with an endoribonuclease such as *Escherichia coli* RNase III or Dicer [90]. In this case, a heterogeneous mixture of many dozens of different siRNAs contributes to the efficient on-target knockdown. As was the case with pools of synthetic siRNAs, off-target effect caused by a single siRNA in the esiRNA mixture is thought to be "diluted" out to a degree below detection level. Similar to shRNAs, esiRNAs cannot be chemically modified, so chemical modifications that would increase efficiency, specificity, or stability of the molecules are not usable with this technique. Overall, while these strategies may be successful in reducing off-target effects for individual target transcripts, they can also complicate large-scale screening; because strong off-target effects by one siRNA may mask the phenotype of the other siRNAs in the pool, all gene targets identified through pooled screens must be validated *via* techniques such as deconvolution of the individual reagents or testing of pools of reagents with fully independent sequences (and thus separate off-target effects) [91].

IDENTIFYING OFF-TARGETS

While reagent sequence selection and modification choices can considerably increase siRNA specificity, it is not yet possible to eliminate off-target effects entirely for all reagents. For this reason, researchers employ a number of techniques to identify potential off-target effects in their experimental data.

Reannotating RNAi Libraries

For large-scale genetic screens, several companies offer sets of RNAi reagents targeting thousands of genes. Such RNAi libraries target the whole genome or a family of genes (*e.g.* kinome) of a given organism. There is usually a delay of months to years between the design of a library and the analysis of screening

results produced with it. During this delay, the annotation of transcript sequences or genome assemblies may change, providing additional information about the potential targets of siRNAs in the library. This information may reveal a number of problematic situations: 1) A given oligonucleotide's target gene is now believed to be a pseudogene; 2) The oligonucleotide targets a different gene than originally predicted; 3) The oligonucleotide targets multiple genes; 4) The oligonucleotide targets no gene at all; and/or 5) The oligonucleotide has been shown to be likely to function in an miRNA-like fashion through its seed region. To detect such situations and exclude problematic RNAi reagents from the analysis of screening results, there is a need to reanalyze RNAi library sequences from time to time. This information is provided to the siRNA community by resources such as the RNAiAtlas [92] and GenomeRNAi [93] databases; see the accompanying chapter on RNAi Databases for further details.

Detecting Evidence of off-Targets

Bioinformatics techniques designed to detect evidence of off-targets generally have the goal of narrowing the list of intended targets to follow up on from the set of those tested in a broader experiment (although, in some cases, detection of off-target effects may also suggest additional biology of interest and thus become a hypothesis-generation tool in its own right). While it is entirely possible for phenotypic results from a given reagent to be due to a combination of both its on- and off-target effects, several methods are in use to identify those whose results are likely to be *primarily* due to off-target effects. If multiple reagents per target were tested separately during the experiment, then much can be learned by combining their information, using methods ranging from simple counting approaches to more complex weighted rankings. For example, a basic but common heuristic for validating screening hits is to count the number of independent reagents for a target that give the phenotype of interest; if only a minority of them produce the desired phenotype, then it is more likely to be due to off-targeting than to knockdown of the intended target. This approach has been formalized as the "H score" [94]. More nuanced aggregation of multiple reagent results can be achieved with techniques such as RSA (Redundant siRNA Activity) [95] or cSSMD (Collective Strictly Standardized Mean Difference) [96]. These approaches examine the degree of phenotype produced by all reagents for a given target and rank targets by how consistent the phenotype produced across those reagents is; while high-ranking targets are likely real effects (and are usually the desired output of these methods) low-ranking targets can be interpreted as probable victims of off-target effects. A considerable advantage of techniques based on multiple reagents per target is that they do not require any assumptions

about the mechanism of off-targeting. Thus, they can detect evidence of both cleavage-based and seed-based off-target effects, as well as clues to any other potential mode of off-targeting.

If multiple separate reagents per target were not tested during the experiment, then some mechanistic assumptions must be made in order to find signs of off-target effects in the data. Cleavage-based off-targeting is usually not widespread enough to be detectable by general analysis of a screening data set, but seed-based off-targeting can be investigated by checking for recurring association of a seed with a given phenotype. In the first approach, the tested reagents that have been chosen as screening hits are examined. If the reagent sequences are enriched for any particular seeds, this indicates that this seed may be causing an off-target effect leading to the tested phenotype [56]. When using this method, it is important to compare the enrichment against a relevant set of tested but not positive reagents, since enrichment of certain seeds is likely in the overall tested set due to design algorithm bias; for example, the Reynolds *et al*. siRNA design algorithm [72] favors (A/U)-rich seed regions, so seeds with high (A/U) content will likely be enriched in all tested siRNAs designed with this algorithm, regardless of whether or not they are hits. The second approach, Common Seed Analysis (CSA) [97], examines the overall performance of all reagents containing a particular seed that were assayed in the experiment. The expectation is that if a seed has no off-target effect, those reagents containing it will give phenotypes randomly distributed around zero. Those seeds whose reagents show a concerted shift away from zero in either direction are considered likely off-target drivers. Unlike the first method, this approach does not require identifying hits, although it would not be effective on data from a screen in which most reagents were expected to have a real biological effect (such as that of a targeted library).

Identifying off-Targeted Genes

Considerable efforts have been made to develop bioinformatics methods that identify particular gene(s) off-targeted by a given RNAi reagent or set of reagents. As discussed earlier, *de novo* prediction of genes that will be off-targeted by a reagent, based only on its sequence, is relatively successful for cleavage-based off-targets but often disappointing for seed-based ones. However, outcomes can be improved when experiments have already been performed and phenotypic data can be incorporated into the analysis. In such situations, identification of genes affected by cleavage-based off-targeting is largely ad hoc; a reagent that has been identified by one of the above-discussed methods as likely to produce off-target effects is checked for perfect or near-perfect sequence matches against the

assayed transcriptome using string searching, BLAST [98], or other similar means. Identifying a concise set of genes putatively affected by seed-based off-targeting in experimental results is usually more labor-intensive, and is frequently frustrated by the extremely combinatorial nature of seed-based regulation effects [50]. While it is theoretically possible to use microRNA target prediction algorithms to identify potential seed-based off-targets and then narrow the extensive list using one's phenotypic data, this is in general feasible only if the tested reagent shares a seed with a known miRNA because many of these algorithms accept only miRNAs (not arbitrary antisense sequences) as input. However, other approaches have been developed. For example, GESS (Genome-wide Enrichment for Seed Sequences) [99] identifies transcripts in the tested transcriptome that are enriched for complementarity to 7mer seeds of active siRNAs (relative to those of inactive siRNAs). The developers of CSA have also recently proposed an additional method called Haystack [100], which assesses the correlation for each transcript in the tested transcriptome between each tested reagent's phenotype and the off-target effect predicted for that reagent's seed using a linear model. For each significantly correlated transcript, the program estimates how much of the overall screening results is explainable by off-target effects on that transcript and calculates a p-value for it. As the Haystack developers note, such approaches to finding putative off-target genes have the potential to uncover novel biology. Off-targets represent real biological effects on the pathway of interest, and thus may draw attention to genes that are involved with the phenotype of interest but were not tested in the initial screen.

CONCLUSION AND FUTURE DIRECTIONS

Unintended regulation of gene expression generated by off-target effects represents a major limitation of RNAi-based technologies in research, therapeutics, and diagnostics. Strategies such as chemical modifications, reagent pooling, and novel rational design filters can significantly reduce off-target effects without reducing on-target gene knockdown. However, knockdown of unintended genes by off-targeting remains common in RNAi experimental results.

A key goal of the RNAi field is thus to be able to design RNAi reagents that cause no off-target effects. However, since every possible hexamer seed sequence is represented in the transcriptome, some degree of seed-based off-targeting may be unavoidable. Even avoidance of cleavage-based off-targets may not always be possible, especially in systems such as shRNAs that cannot be chemically modified. Therefore, a fallback aim is to reliably predict, without experimental work, which off-targets a given reagent will have. Efforts to date demonstrate that

this is still very difficult, because seed-based off-target effects are demonstrably different for different experimental systems, cell lines, and assays [101-103]. Much more data will be needed to enable more successful predictive systems.

Until such improved systems are developed, results of any computational off-target identification or prediction technique must be treated with some skepticism. It is critical to recall that although these methods identify potential off-targets, they do not prove off-target activity, which can only be accomplished with further experimental work. Techniques like testing alternate reagents against the intended target and/or the putative off-target of a reagent, or evaluating control reagents such as seed chimeras [56, 104] or C911 constructs [105], can confirm or refute computational predictions and expand our understanding of off-target biology.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors confirm that this chapter contents have no conflicts of interest.

REFERENCES

[1]   Pecot CV, Calin GA, Coleman RL, Lopez-Berestein G, Sood AK. RNA interference in the clinic: challenges and future directions. Nat Rev Cancer. 2011 Jan;11(1):59-67.
[2]   Davidson BL, McCray PB Jr. Current prospects for RNA interference-based therapies. Nat Rev Genet. 2011 May;12(5):329-40.
[3]   Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature. 2001 May 24;411(6836):494-8.
[4]   Elbashir SM, Martinez J, Patkaniowska A, Lendeckel W, Tuschl T. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. The EMBO journal. 2001 Dec 3;20(23):6877-88.
[5]   Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, *et al*. Expression profiling reveals off-target gene regulation by RNAi. Nature biotechnology. 2003 Jun;21(6):635-7.
[6]   Scacheri PC, Rozenblatt-Rosen O, Caplen NJ, Wolfsberg TG, Umayam L, Lee JC, *et al*. Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America. 2004 Mar 17;101(7):1892-7.
[7]   Lin X, Ruan X, Anderson MG, McDowell JA, Kroeger PE, Fesik SW, *et al*. siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. Nucleic acids research. 2005 Jan 9;33(14):4527-35.
[8]   Fedorov Y, King A, Anderson E, Karpilow J, Ilsley D, Marshall W, *et al*. Different delivery methods-different expression profiles. Nature methods. 2005 Apr;2(4):241.
[9]   Sledz CA, Holko M, de Veer MJ, Silverman RH, Williams BRG. Activation of the interferon system by short-interfering RNAs. Nature cell biology. 2003 Sep;5(9):834-9.

[10] Bridge AJ, Pebernard S, Ducraux A, Nicoulaz A-L, Iggo R. Induction of an interferon response by RNAi vectors in mammalian cells. Nature genetics. 2003 Jul;34(3):263-4.

[11] Doench JG, Petersen CP, Sharp PA. siRNAs can function as miRNAs. Genes & development. 2003 Feb 15;17(4):438-42.

[12] Birmingham A, Anderson EM, Reynolds A, Ilsley-Tyree D, Leake D, Fedorov Y, *et al*. 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. Nature methods. 2006 Mar;3(3):199-204.

[13] Jackson AL, Burchard J, Schelter J, Chau BN, Cleary M, Lim L, *et al*. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. RNA (New York, N.Y.). 2006 Jul 1;12(7):1179-87.

[14] Clark PR, Pober JS, Kluger MS. Knockdown of TNFR1 by the sense strand of an ICAM-1 siRNA: dissection of an off-target effect. Nucleic acids research. 2008 Mar;36(4):1081-97.

[15] Jackson AL, Burchard J, Leake D, Reynolds A, Schelter J, Guo J, *et al*. Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. RNA (New York, N.Y.). 2006 Jul;12(7):1197-205.

[16] Ui-Tei K, Naito Y, Zenno S, Nishi K, Yamato K, Takahashi F, *et al*. Functional dissection of siRNA sequence by systematic DNA substitution: modified siRNA with a DNA seed arm is a powerful tool for mammalian gene silencing with significantly reduced off-target effect. Nucleic acids research. 2008 Apr;36(7):2136-51.

[17] Bramsen JB, Pakula MM, Hansen TB, Bus C, Langkjær N, Odadzic D, *et al*. A screen of chemical modifications identifies position-specific modification by UNA to most potently reduce siRNA off-target effects. Nucleic acids research. 2010 Sep 1;38(17):5761-73.

[18] Dua P, Yoo JW, Kim S, Lee D. Modified siRNA structure with a single nucleotide bulge overcomes conventional siRNA-mediated off-target silencing. Molecular therapy?: the journal of the American Society of Gene Therapy. 2011 Sep;19(9):1676-87.

[19] Bramsen JB, Laursen MB, Damgaard CK, Lena SW, Babu BR, Wengel J, *et al*. Improved silencing properties using small internally segmented interfering RNAs. Nucleic acids research. 2007 Jan;35(17):5886-97.

[20] Elmén J, Thonberg H, Ljungberg K, Frieden M, Westergaard M, Xu Y, *et al*. Locked nucleic acid (LNA) mediated improvements in siRNA stability and functionality. Nucleic acids research. 2005 Jan;33(1):439-47.

[21] Chen PY, Weinmann L, Gaidatzis D, Pei Y, Zavolan M, Tuschl T, *et al*. Strand-specific 5'-O-methylation of siRNA duplexes controls guide strand selection and targeting specificity. RNA (New York, N.Y.). 2008 Feb;14(2):263-74.

[22] Hutvágner G, McLachlan J, Pasquinelli AE, Bálint E, Tuschl T, Zamore PD. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science (New York, N.Y.). 2001 Aug 3;293(5531):834-8.

[23] Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. 2001 Jan 18;409(6818):363-6.

[24] Williams BR. PKR; a sentinel kinase for cellular stress. Oncogene. 1999 Nov 1;18(45):6112-20.

[25] Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, *et al*. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell. 2004 Apr 2;117(1):69-81.

[26] Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Molecular cell. 2004 Jul 23;15(2):185-97.

[27] Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song J-J, *et al*. Argonaute2 is the catalytic engine of mammalian RNAi. Science (New York, N.Y.). 2004 Sep 3;305(5689):1437-41.

[28] Rand TA, Ginalski K, Grishin NV, Wang X. Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity. Proceedings of the National Academy of Sciences of the United States of America. 2004 Oct 5;101(40):14385-9.

[29] Rand TA, Petersen S, Du F, Wang X. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. Cell. 2005 Nov 18;123(4):621-9.

[30] Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., Tuschl, T. (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell 110:563-574.

[31] Brummelkamp TR, Bernards R, Agami R. A system for stable expression of short interfering RNAs in mammalian cells. Science (New York, N.Y.). 2002 Apr 19;296(5567):550-3.

[32]   Paddison PJ, Caudy AA, Bernstein E, Hannon GJ, Conklin DS. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. Genes & development. 2002 Apr 15;16(8):948-58.

[33]   Denli AM, Tops BBJ, Plasterk RHA, Ketting RF, Hannon GJ. Processing of primary microRNAs by the Microprocessor complex. Nature. 2004 Nov 11;432(7014):231-5.

[34]   Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, *et al*. The nuclear RNase III Drosha initiates microRNA processing. Nature. 2003 Sep 25;425(6956):415-9.

[35]   Gregory RI, Yan K-P, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, *et al*. The Microprocessor complex mediates the genesis of microRNAs. Nature. 2004 Nov 11;432(7014):235-40.

[36]   Lund E, Güttinger S, Calado A, Dahlberg JE, Kutay U. Nuclear export of microRNA precursors. Science (New York, N.Y.). 2004 Jan 2;303(5654):95-8.

[37]   Lee Y, Jeon K, Lee J-T, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. The EMBO journal. 2002 Sep 2;21(17):4663-70.

[38]   Lee Y, Kim M, Han J, Yeom K-H, Lee S, Baek SH, *et al*. MicroRNA genes are transcribed by RNA polymerase II. The EMBO journal. European Molecular Biology Organization; 2004 Oct 13;23(20):4051-60.

[39]   Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes & development. 2003 Dec 15;17(24):3011-6.

[40]   Bohnsack MT, Czaplinski K, Gorlich D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA (New York, N.Y.). 2004 Feb;10(2):185-91.

[41]   Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, *et al*. Genes and Mechanisms Related to RNA Interference Regulate Expression of the Small Temporal RNAs that Control C. elegans Developmental Timing. Cell. 2001 Jul;106(1):23-34.

[42]   Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes & development. 2001 Oct 15;15(20):2654-9.

[43]   Henke JI, Goergen D, Zheng J, Song Y, Schüttler CG, Fehr C, *et al*. microRNA-122 stimulates translation of hepatitis C virus RNA. The EMBO journal. 2008 Dec 17;27(24):3300-10.

[44]   Orom UA, Nielsen FC, Lund AH. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. Molecular cell. 2008 May 23;30(4):460-7

[45]   Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell. 2009 Jan 23;136(2):215-33.

[46]   Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nature reviews. Genetics. 2008 Feb;9(2):102-14.

[47]   Vella MC, Choi E-Y, Lin S-Y, Reinert K, Slack FJ. The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. Genes & development. 2004 Jan 15;18(2):132-7.

[48]   Shin C, Nam JW, Farh KK, Chiang HR, Shkumatava A, Bartel DP.: Expanding the microRNA targeting code: functional sites with centered pairing. Mol Cell. 2010, Jun 25;38(6):789-802.

[49]   Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005 Jan 14;120(1):15-20.

[50]   Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N. Combinatorial microRNA target predictions. Nat Genet. 2005 May;37(5):495-500.

[51]   Tuschl T, Zamore PD, Lehmann R, Bartel DP, Sharp PA. Targeted mRNA degradation by double-stranded RNA *in vitro*. Genes & development. 1999 Dec 15;13(24):3191-7.

[52]   Du Q, Thonberg H, Wang J, Wahlestedt C, Liang Z. A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites. Nucleic acids research. 2005 Jan;33(5):1671-7.

[53]   Semizarov D, Frost L, Sarthy A, Kroeger P, Halbert DN, Fesik SW. Specificity of short interfering RNA determined through gene expression signatures. Proceedings of the National Academy of Sciences of the United States of America. 2003 May 27;100(11):6347-52.

[54]   Holen T, Moe SE, Sørbø JG, Meza TJ, Ottersen OP, Klungland A. Tolerated wobble mutations in siRNAs decrease specificity, but can enhance activity *in vivo*. Nucleic acids research. 2005 Jan;33(15):4704-10.

[55]   Alemán LM, Doench J, Sharp PA. Comparison of siRNA-induced off-target RNA and protein effects. RNA (New York, N.Y.). 2007 Mar;13(3):385-95.

[56]   Sudbery I., Enright AJ, Fraser AG and Dunham I: Systematic analysis of off-target effects in an RNAi screen reveals microRNAs affecting sensitivity to TRAIL-induced apoptosis. BMC Genomics 2010, 11:175.

[57]   Schultz N, Marenstein DR, De Angelis DA, Wang WQ, Nelander S, Jacobsen A, Marks DS, Massagué J, Sander C. Off-target effects dominate a large-scale RNAi screen for modulators of the TGF-ß pathway and reveal microRNA regulation of TGFBR2. Silence. 2011 Mar 14;2:3.

[58]   Kulkarni MM, Booker M, Silver SJ, Friedman A, Hong P, Perrimon N, Mathey-Prévôt B. Evidence of off-target effects associated with long dsRNAs in Drosophila melanogaster cell-based assays. Nat Methods. 2006 Oct;3(10):833-8.

[59]   Judge AD, Sood V, Shaw JR, Fang D, McClintock K, MacLachlan I. Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. Nat Biotechnol. 2005 Apr;23(4):457-62.

[60]   Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. Nature. 2001 Oct 18;413(6857):732-8

[61]   Hornung V, Guenthner-Biller M, Bourquin C, Ablasser A, Schlee M, Uematsu S, Noronha A, Manoharan M, Akira S, de Fougerolles A, Endres S, Hartmann G. Sequence-specific potent induction of IFN-alpha by short interfering RNA in plasmacytoid dendritic cells through TLR7. Nat Med. 2005 Mar;11(3):263-70.

[62]   Fedorov, Y, Anderson, EM, Birmingham A, Reynolds A, Karpilow J, Robinson K, Leake D, Marshall WS and Khvorova A. Off-target effects by sirna can induce toxic phenotype. RNA, 2006, 12(7):1188-96.

[63]   Silva JM, Li MZ, Chang K, Ge W, Golding MC, Rickles RJ, Siolas D, Hu G, Paddison PJ, Schlabach MR, Sheth N, Bradshaw J, Burchard J, Kulkarni A, Cavet G, Sachidanandam R, McCombie WR, Cleary MA, Elledge SJ, Hannon GJ. Second-generation shRNA libraries covering the mouse and human genomes. Nat Genet. 2005 Nov;37(11):1281-8.

[64]   McBride JL, Boudreau RL, Harper SQ, Staber PD, Monteys AM, Martins I, Gilmore BL, Burstein H, Peluso RW, Polisky B, Carter BJ, Davidson BL. Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi. Proc Natl Acad Sci U S A. 2008 Apr 15;105(15):5868-73.

[65]   Siolas D, Lerner C, Burchard J, Ge W, Linsley PS, Paddison PJ, Hannon GJ, Cleary MA. Synthetic shRNAs as potent RNAi triggers. Nat Biotechnol. 2005 Feb;23(2):227-31.

[66]   Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. Cell, 2005, 123 (4): 631-40.

[67]   Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, Marion P, Salazar F, Kay MA. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. Nature. 2006 May 25;441(7092):537-41.

[68]   Schwarz DS, Hutvágner G, Du T, Xu Z, Aronin N, Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. Cell. 2003 Oct 17;115(2):199-208.

[69]   Khvorova A, Reynolds A, Jayasena SD. Functional siRNAs and miRNAs exhibit strand bias. Cell. 2003 Oct 17;115(2):209-16.

[70]   Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell. 2005 Nov 18;123(4):607-20.

[71]   Leuschner PJF, Ameres SL, Kueng S, Martinez J. Cleavage of the siRNA passenger strand during RISC assembly in human cells. EMBO reports. 2006 Mar;7(3):314-20.

[72]   Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorova A. Rational siRNA design for RNA interference. Nat Biotechnol. 2004 Mar;22(3):326-30.

[73]   Birmingham A, Anderson E, Sullivan K, Reynolds A, Boese Q, Leake D, Karpilow J, Khvorova A. A protocol for designing siRNAs with high functionality and specificity. Nat Protoc. 2007;2(9):2068-78.

[74]   Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP. Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. Nat Struct Mol Biol., 2011, 18 (10): 1139-1146.

[75]   Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. Genome Biol. 2003;5(1):R1.

[76]   Reczko M, Maragkakis M, Alexiou P, Papadopoulos GL, Hatzigeorgiou AG. Accurate microRNA Target Prediction Using Detailed Binding Site Accessibility and Machine Learning on Proteomics Data. Front Genet. 2011;2:103.

[77]   Alexiou P, Maragkakis M, Papadopoulos GL, Reczko M, Hatzigeorgiou AG. Lost in translation: an assessment and perspective for computational microRNA target identification. Bioinformatics. 2009 Dec 1;25(23):3049-55.

[78]   Jalvy-Delvaille S, Maurel M, Majo V, Pierre N, Chabas S, Combe C, Rosenbaum J, Sagliocco F, Grosset CF. Molecular basis of differential target regulation by miR-96 and miR-182: the Glypican-3 as a model. Nucleic Acids Res. 2012 Feb;40(3):1356-65.

[79]   Helwak A, Kudla G, Dudnakova T, Tollervey D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell. 2013 Apr 25;153(3):654-65.

[80]   Gan HH, Gunsalus KC. Tertiary structure-based analysis of microRNA-target interactions. RNA. 2013 Apr;19(4):539-51.

[81]   Janas MM, Wang B, Harris AS, Aguiar M, Shaffer JM, Subrahmanyam YV, Behlke MA, Wucherpfennig KW, Gygi SP, Gagnon E, Novina CD. Alternative RISC assembly: binding and repression of microRNA-mRNA duplexes by human Ago proteins. RNA. 2012 Nov;18(11):2041-55.

[82]   Xia Z, Huynh T, Ren P, Zhou R. Large domain motions in Ago protein controlled by the guide DNA-strand seed region determine the Ago-DNA-mRNA complex recognition process. PLoS One. 2013;8(1):e54620.

[83]   Rasmussen SH, Jacobsen A, Krogh A. cWords - systematic microRNA regulatory motif discovery from mRNA expression data. Silence. 2013 May 20;4(1):2.

[84]   Betancur JG, Yoda M, Tomari Y. miRNA-like duplexes as RNAi triggers with improved specificity. Front Genet. 2012;3:127.

[85]   Ui-Tei K, Naito Y, Nishi K, Juni A, Saigo K. Thermodynamic stability and Watson-Crick base pairing in the seed duplex are major determinants of the efficiency of the siRNA-based off-target effect. Nucleic Acids Res. 2008 Dec;36(22):7100-9.

[86]   Anderson EM, Birmingham A, Baskerville S, Reynolds A, Maksimova E, Leake D, Fedorov Y, Karpilow J and Khvorova A. Experimental validation of the importance of seed complement frequency to sirna specificity. RNA, 2008, 14(5):853-61.

[87]   Wang Y, Sheng G, Juranek S, Tuschl T, Patel DJ. Structure of the guide-strand-containing argonaute silencing complex. Nature. 2008 Nov 13;456(7219):209-13.

[88]   Zhang J, Zheng J, Lu C, Du Q, Liang Z, Xi Z. Modification of the siRNA Passenger Strand by 5-Nitroindole Dramatically Reduces its Off-Target Effects. Chembiochem?: a European journal of chemical biology. 2012 Sep 3;13(13):1940-5.

[89]   Bramsen JB, Laursen MB, Nielsen AF, Hansen TB, Bus C, Langkjaer N, *et al*. A large-scale chemical modification screen identifies design rules to generate siRNAs with high activity, high stability and low toxicity. Nucleic acids research. 2009 May;37(9):2867-81.

[90]   Yang D, Buchholz F, Huang Z, Goga A, Chen CY, Brodsky FM, Bishop JM. Short RNA duplexes produced by hydrolysis with Escherichia coli RNase III mediate effective RNA interference in mammalian cells. Proc Natl Acad Sci U S A. 2002 Jul 23;99(15):9942-7.

[91]   Sigoillot FD, King RW. Vigilance and validation: Keys to success in RNAi screening. ACS Chem Biol. 2011 Jan 21;6(1):47-60.

[92]   Mazur S, Csucs G, Kozak K. RNAiAtlas: a database for RNAi (siRNA) libraries and their specificity. Database?: the journal of biological databases and curation. 2012 Jan 10;2012(0):bas027.

[93]   Horn T, Arziman Z, Berger J, Boutros M.: GenomeRNAi: a database for cell-based RNAi phenotypes. Nucleic Acids Res. 2007 Jan;35(Database issue):D492-7.

[94]   Bhinder B, Djaballah H. A simple method for analyzing actives in random RNAi screens: introducing the "H Score" for hit nomination & gene prioritization. Comb Chem High Throughput Screen. 2012 Nov;15(9):686-704.

[95]   König R, Stertz S, Zhou Y, Inoue A, Hoffmann HH, Bhattacharyya S, Alamares JG, Tscherne DM, Ortigoza MB, Liang Y, Gao Q, Andrews SE, Bandyopadhyay S, De Jesus P, Tu BP, Pache L, Shih C, Orth A, Bonamy G, Miraglia L, Ideker T, García-Sastre A, Young JA, Palese P, Shaw ML, Chanda SK: Human host factors required for influenza virus replication. Nature 463:7282 2010 Feb 11 pg 813-7.

[96]   Zhang XD, Santini F, Lacson R, Marine SD, Wu Q, Benetti L, Yang R, McCampbell A, Berger JP, Toolan DM, Stec EM, Holder DJ, Soper KA, Heyse JF, Ferrer M. cSSMD: assessing collective activity for addressing off-target effects in genome-scale RNA interference screens. Bioinformatics. 2011 Oct 15;27(20):2775-81.

[97]   Marine S, Bahl A, Ferrer M, Buehler E. Common seed analysis to identify off-target effects in siRNA screens. J Biomol Screen. 2012 Mar;17(3):370-8.

[98]   Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ. Basic local alignment search tool. J Mol Biol, 1990, 215(3):403-10.

[99]   Sigoillot FD, Lyman S, Huckins JF, Adamson B, Chung E, Quattrochi B, King RW. A bioinformatics method identifies prominent off-targeted transcripts in RNAi screens. Nat Methods. 2012 Feb 19;9(4):363-6.

[100]   Buehler E, Khan AA, Marine S, Rajaram M, Bahl A, Burchard J, Ferrer M. siRNA off-target effects in genome-wide screens identify signaling pathway members. Sci Rep. 2012;2:428

[101]   Burchard J, Jackson AL, Malkov V, Needham RH, Tan Y, Bartz SR, Dai H, Sachs AB, Linsley PS. MicroRNA-like off-target transcript regulation by siRNAs is species specific. RNA. 2009 Feb;15(2):308-15.

[102]   Baum P, Fundel-Clemens K, Kreuz S, Kontermann RE, Weith A, Mennerich D, Rippmann JF. Off-target analysis of control siRNA molecules reveals important differences in the cytokine profile and inflammation response of human fibroblasts. Oligonucleotides, 2010, 20(1):17-26.

[103]   Stone DJ, Marine S, Majercak J, Ray WJ, Espeseth A, Simon A, Ferrer M.: High-throughput screening by RNA interference: control of two distinct types of variance. Cell Cycle. 2007 Apr 15;6(8):898-901.

[104]   Echeverri CJ, Beachy PA, Baum B, Boutros M, Buchholz F, Chanda SK, Downward J, Ellenberg J, Fraser AG, Hacohen N, Hahn WC, Jackson AL, Kiger A, Linsley PS, Lum L, Ma Y, Mathey-Prévôt B, Root DE, Sabatini DM, Taipale J, Perrimon N, Bernards R. Minimizing the risk of reporting false positives in large-scale RNAi screens. Nat Methods. 2006 Oct;3(10):777-9.

[105]   Buehler E, Chen YC, Martin S. C911: A Bench-Level Control for Sequence Specific siRNA Off-Target Effects. PLoS One. 2012;7(12):e51942.

Automation Considerations for RNAi Library Formatting and High Throughput Transfection

Sean M. Johnston, Caroline E. Shamu and Jennifer A. Smith[*]

ICCB-Longwood Screening Facility, Harvard Medical School, 250 Longwood Avenue, SGM Building Room 604, Boston, MA 02115 USA

Abstract: Laboratory automation impacts nearly all aspects of high throughput RNAi screens. It is particularly relevant when considering library format and storage, and also when planning high throughput transfection protocols. In situations where libraries are stored as screening copies that are used for just-in-time dispensing, automation can be utilized as a tool to accommodate diverse assay and cell types and can enable forward or reverse transfections into a variety of different microplates. Automation has the ability to increase the feasibility and decrease consumable costs for assays that require a large number of replicates. It is also an important tool when considering more complex assays, such as those that utilize non-standard plate types or electroporation, as automation increases reliability and can improve assay performance. This chapter highlights important considerations for library formatting and ways in which laboratory automation can be implemented to facilitate RNAi high throughput screening.

Keywords: HTS, laboratory automation, library format, RNAi, robotics, siRNA transfection.

INTRODUCTION

There are many aspects to consider when planning an RNAi-based high throughput screen (HTS). Among the more important is access to laboratory automation. This must be taken into account when planning library formats and high throughput transfection logistics. This chapter will discuss common methods for RNAi library formatting and storage, and illustrate how automation can facilitate diverse assay protocols.

This chapter presents the perspective of one academic screening center, the ICCB-Longwood Screening Facility (ICCB-L) at Harvard Medical School, which has

*****Corresponding author Jennifer A. Smith:** ICCB-Longwood Screening Facility, Harvard Medical School, 250 Longwood Avenue, SGM Building Room 604, Boston, MA 02115 USA; Tel: 617-432-5735; Fax: 617-432-6424; E-mail: jennifer_smith@hms.harvard.edu

supported genome-scale siRNA screening since 2006. ICCB-L is organized as a modular screening facility, with the instrumentation set up in a workstation format rather than integrated into fully automated systems. This enables multiple screening projects to be performed simultaneously. ICCB-L operates under an investigator-initiated, staff-assisted screening model. Investigators provide the scientific rationale and overall assay design for their projects and carry out the bulk of the work (*e.g.*, cell culture, plating, and assay readout) for their own screens. ICCB-L personnel maintain the HTS infrastructure and perform all complex automation tasks, including those that involve the libraries. Critical to each screen's success, ICCB-L personnel provide advice and assistance during all aspects of the screen. This includes training on instrument use, and advising on assay development and optimization as well as data analysis.

The goal of this chapter is to provide examples from ICCB-L demonstrating how laboratory automation can be used for formatting siRNA libraries and in implementing a variety of transfection strategies for high throughput screens.

LIBRARY FORMAT

There are a number of considerations when deciding how to array an siRNA library. One of the first choices to be made is the initial state of the reagents. Depending on the vendor, a new library may arrive lyophilized or already in suspension. Lyophilized libraries are more easily transported and stored because they are stable at room temperature for some length of time [1]. Receiving and storing libraries lyophilized allows the screening facility to control resuspension volume and concentration, but quality control testing may still be necessary to ensure that the appropriate quantity of each library reagent has been received and completely re-suspended. Lyophilized siRNA libraries are received at ICCB-L, which are then resuspended at 10 or 20 μM to generate stock copies. Subsequently, stocks are diluted to make "screening" copies of each library at 1 μM for use in transfections. See the Library Storage section below for a more detailed description of ICCB-L siRNA library stock and screening plates.

High throughput RNAi screening is generally carried out in 96- or 384-well microplate format. The layout of the library in microplates must be carefully planned to ensure it is suitable for the assays that will be conducted in the screening facility. Re-arraying an siRNA library should be avoided whenever possible. Moving sets of wells introduces the potential for costly errors: physical mistakes (transferring reagents to the wrong well or inadvertently combining wells), or misannotation in tracking library contents. In addition, every time a

library is transferred, a small amount of library material is lost to the pipet tips and the source labware. These pitfalls of re-arraying may be avoided by coordinating the desired plate layouts with the vendor.

When possible, it is advisable to avoid plating siRNA reagents in edge wells (*e.g.*, rows A and P and columns 1 and 24 in a 384-well plate), as edge effects can have a major impact on assay robustness and data quality [2]. This edge effect variation can be caused by uneven evaporation and temperature fluctuations, both of which are major concerns for RNAi assays, which usually have incubation times of 72 to 96 hours. Leaving the outer two rows and columns of a library plate free of siRNAs would likely significantly decrease assay edge effects, but this would increase the total number of plates required to hold the library by approximately 25%. Because there are other strategies that can be used to mitigate edge effects (*e.g.*, breathable seals, MicroClime lids, active humidity controlled incubators, *etc.*), a good practice is to leave at least one column and one row empty at each outer edge. Fig. **1** illustrates the current siRNA library plating strategy at ICCB-L.

There are two classes of controls that must be taken into account when planning the library format, "library" controls and "assay-specific" positive and negative controls. The library controls are arrayed onto the library plates and are used to monitor successful transfection and library integrity. Assay-specific controls are added to the assay plates directly at the time of transfection and thus have corresponding empty wells in the appropriate position on the library plate. The library controls should consist of broadly applicable transfection controls, such as death-inducing siRNAs (*e.g.*, PLK1 or KIF11) whose effects can be monitored in many different assay types, non-targeting siRNA negative controls, and even RISC-free fluorescent controls that are useful when quantitating siRNA uptake efficiency. In contrast, assay-specific controls should be selected by individual investigators for each screen. The positive control is an siRNA that produces the desired phenotype sought in the screen. When possible, it is good practice to include multiple different siRNA positive controls, *e.g.*, two to three with strong or medium effects in the assay, so that the dynamic range of the assay might be monitored. Additional negative controls should be included on assay plates if those included as "library" controls are not sufficient. The negative controls should be non-targeting siRNAs that have been demonstrated to have no significant phenotype in the screening assay; this must be tested explicitly as no "universal" siRNA negative control exists that has no phenotype in all cell-based assays. If the library is a focused set of siRNAs (*e.g.*, small library targeting kinases) where many wells might be expected to display some phenotype, it is

crucial that enough wells of the library are left empty for a robust set of negative controls, as the plate median or average will not be a useful baseline for assay quantitation [3].
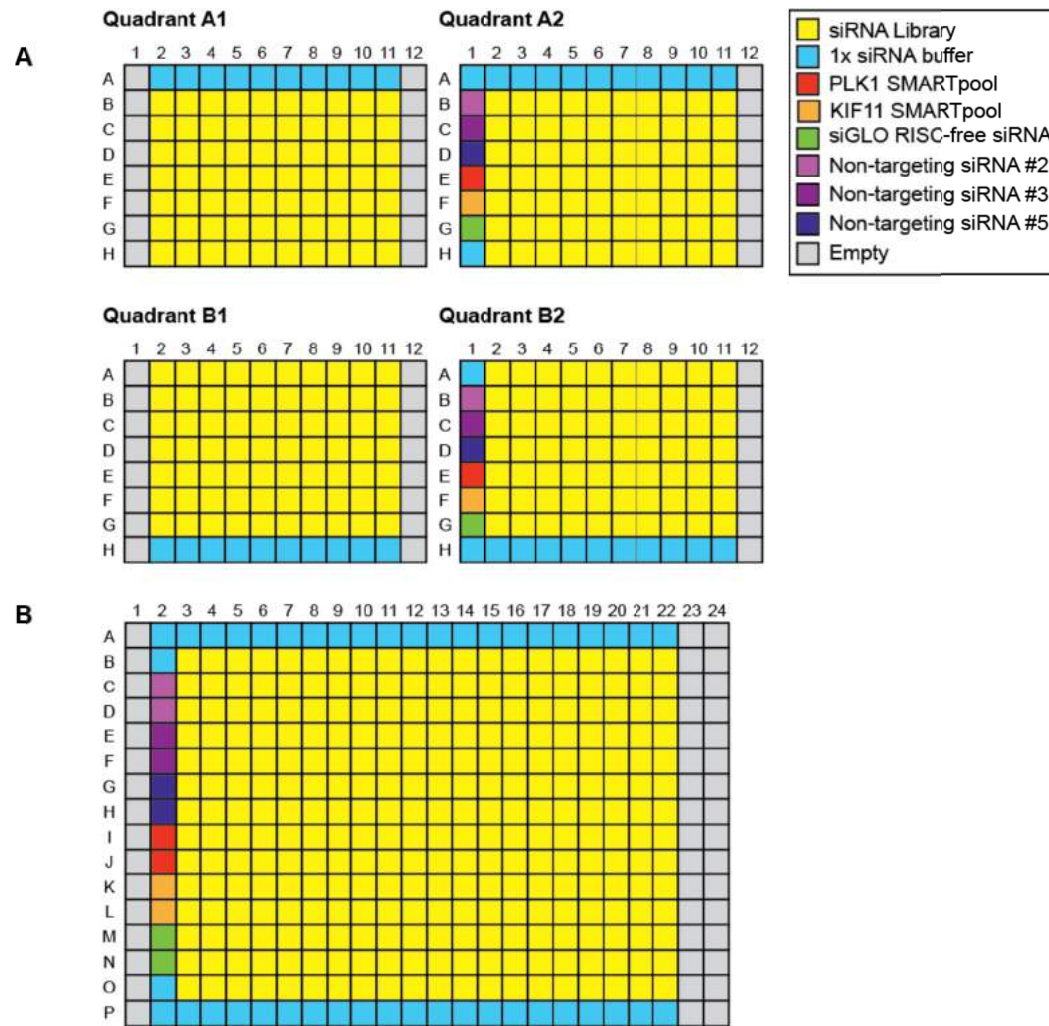
Figure 1: **Current siRNA library plate format at ICCB-Longwood**. **A**) 96-well siRNA library layout. Location of experimental and empty wells was determined by ICCB-L and formatted by the vendor. Library-specific controls were added to original 96-well source plates (Quadrant A2 and Quadrant B2) by ICCB-L staff prior to formatting into 384-well copies. **B**) In the corresponding 384-well siRNA library layout, quadrant A1 maps to wells A1, A3, A5, C1, C3, C5, *etc.* in the 384-well plate layout. Quadrant A2 maps to wells A2, A4, A6, C2, C4, C6, *etc.* in the 384-well format. Quadrant B1 maps to wells B1, B3, B5, D1, D3, D5, *etc.* in the 384-well format. Quadrant B2 maps to B2, B4, B6, D2, D4, D6, *etc.* in the 384-well format. The legend indicates siRNA reagent type in each well.
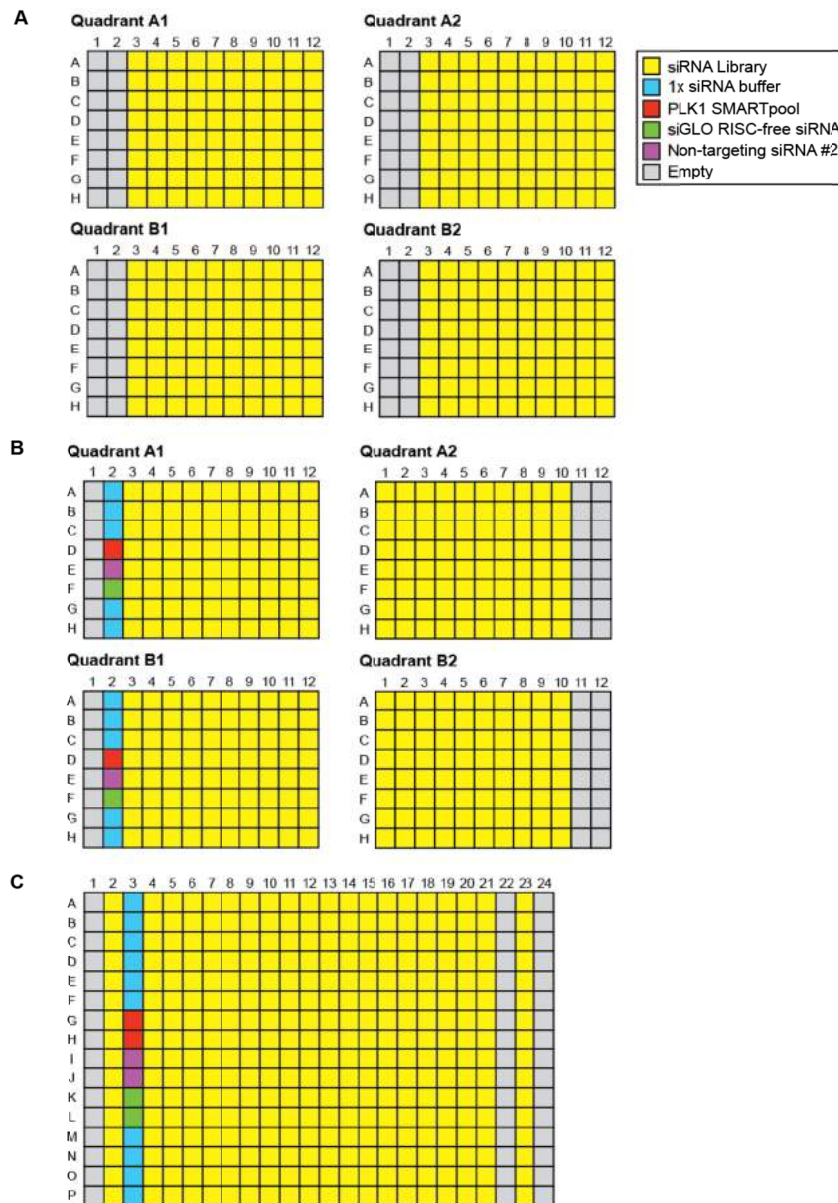
Figure 2: **Formatting 384-well library plate from identical 96-well quadrants**. **A**) 96-well siRNA layout received from vendor, with the location of experimental and empty wells pre-determined by the vendor. **B**) Library-specific controls were added to original 96-well source plates (Quadrant A1 and Quadrant B1) and two of the 96-well quadrants (A2 and B2) were rotated 180° by ICCB-L staff to better distribute empty wells on each side of the resultant 384-well plate prior to 384-well formatting. **C**) In the resulting 384-well format, columns 1, 22 and 24 are empty and can be populated by assay-specific controls. Column 3 contains the library controls. The legend indicates siRNA reagent type in each well.

The standard siRNA library layout at ICCB-L is a 384-well format with the outermost row (A and P) and column (1 and 24) empty, as well as one column (2) designated for a variety of library controls and an additional column (23) empty for assay-specific controls (Fig. **1B**). While the ICCB-L RNAi libraries are stored and screened in 384-well format, they are usually purchased and received in 96-well plates (Fig. **1A**). The larger volume capacity of a 96-well plate is useful when diluting a lyophilized library to an appropriate stock concentration, and 96-well plates are amenable to both automated and manual pipetting. For this reason, many vendors stock products in 96-well plate format. Other formats and densities may require a surcharge or be unavailable entirely. A 384-well plate can be subdivided into four sets of 96 wells, or quadrants [4]. The quadrants interleave with even spacing, such that the 96-well tips will go into every other well in each direction (described in Fig. **1** legend). The quadrants are then designated A1, A2, B1, or B2, depending on which well of the 384-well plate corresponds to well A1 of a particular 96-well plate (Fig. **1A**). The standard ICCB-L 384-well layout required requesting customized 96-well plates from the vendor, such that the 4 quadrants have non-identical layouts. A library vendor may be willing to provide the reagents in the desired format or the purchaser may re-organize the collection by moving samples with a pipettor, although this may be a daunting task across an entire genome. In instances where a library is shipped in 96-well plates with identical layouts (Fig. **2A**), a reasonable option for 384-well formatting may be to rotate two of the 96-well quadrants when transferring them into 384-well format (Fig. **2B**), thus equally distributing the empty wells to each side of the 384-well plate (Fig. **2C**).

For focused libraries or smaller collections, such as ICCB-L's miRNA mimic and inhibitor libraries of less than 1,000 wells, one may customize a layout with less concern for the total number of plates. As seen in Fig. **3**, the miRNA library has the two outer rows (A, B, O and P) and columns (1, 2, 23 and 24) empty, which create a strong buffer against edge effects. There is also a field of 4 columns in the middle of the plate for controls, insulating them from edge effects and allowing for the highest fidelity control data. In this example, 2 of the 4 columns are utilized for library-specific controls (13 and 14) and the other 2 columns remain available for assay-specific controls (11 and 12).

LIBRARY STORAGE

Several factors need to be considered simultaneously when determining how to store the libraries. There are three basic ways in which RNAi libraries are stored: 1) as pre-dispensed ready-to-go assay plates; 2) as screening copies that are used

for just-in-time dispensing; and 3) in individual reagent tubes for the creation of focused libraries on demand. The latter option utilizes a tube storage system (*e.g.*, TTP Labtech comPOUND or HighRes Biosolutions SampleVault) and enables the formation of a custom library for each experiment rather than being restricted to a pre-formatted library. While use of reagent tubes is ideal in terms of possibilities for customization of plate format and contents, this type of storage system is expensive and may be challenging for typical academic screening centers to implement. The former two storage strategies work well for many applications and will be the focus here.
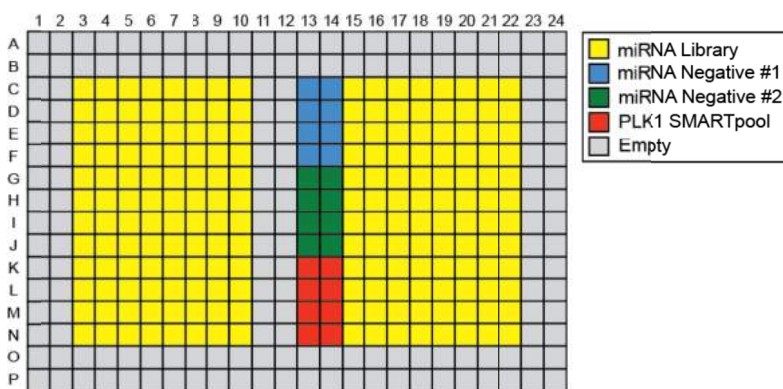
Figure 3: **Small library plate format at ICCB-Longwood**. Location of experimental and empty wells was determined by ICCB-L and formatted by the vendor. Library-specific controls were added to each original source plate by ICCB-L staff prior to formatting into copies. The legend indicates RNAi reagent type in each well.

Several screening centers elect to pre-dispense libraries into ready-to-go assay plates; for all screens, every assay plate holds the same amount of each siRNA reagent. This strategy offers certain advantages, including simplified screener-staff logistics, plate availability whenever a screener is ready for that portion of the library, and the ability to mass produce library aliquots, thus decreasing both consumables cost and staff time. However, ready-to-go assay plate formatting might put limits on assay protocol development. It requires reverse transfection protocols and assay plate types must be determined at the time the library plate is formatted. siRNA amounts would also be fixed at the time of formatting. Thus, while it might be possible to slightly modify the final siRNA concentration in ready-to-go plates by altering assay volumes in some cases, there may be a narrow acceptable volume range that is compatible with the assay protocol, making it impossible to use ready-to-go plates.

Just-in-time dispensing overcomes several of the challenges associated with ready-to-go assay plates. This type of dispensing places few limitations on assay protocols. It allows for a high degree of flexibility in terms of forward or reverse transfection, plate type and well density utilized, final siRNA concentration, and number of replicates included in the screen. It is also easy to automate the addition of assay-specific controls and mixing of siRNA and transfection reagent with tips. Yet there are drawbacks that must be considered, including higher consumables cost, increased staff time for development of custom automation protocols and for siRNA transfection, as well as additional logistics in terms of scheduling (staff, investigator, and equipment).

RNAi libraries at ICCB-L are stored as stock plates and screening copies, with the screening copies utilized for just-in-time dispensing during automated transfections. The majority of the library is stored in high concentration (10 or 20 µM) stock plates. The stocks can be diluted 10- or 20-fold in 1X siRNA buffer to create 1 µM screening copies as necessary. We have found that 1 µM is an ideal library screening copy concentration because 1.5 – 3 µL of 1 µM reagent can be added to a 40 – 50 µL assay well volume in 384-well plates to achieve the recommended siRNA transfection concentration of 25 – 50 µM. In addition, a volume as low as 1 µL can be accurately and precisely dispensed with automated 384-channel pipettors and does not introduce an excessive amount of library buffer salts to the assay.

The plate type used for RNAi library storage at ICCB-L is the Eppendorf twin.tec 384-well PCR plate. These plates are low profile, thus minimizing storage space. They have steep conical wells, promoting high volume recovery, and, unlike most PCR plates, the twin.tec plates have a full rigid polycarbonate skirt that is compatible with a wide array of automation platforms. The plates are available in a variety of colors, allowing color coding of libraries and custom arrays as desired. The maximum comfortable well volume of the twin.tec plates is 30 µL. This allows for volume displacement when pipet tips are inserted and prevents cross contamination when plates are stacked.

At ICCB-L, library plates are heat sealed and stored at -20 °C in plastic containers (*e.g.*, 13" x 7-1/2" x 4-1/4", Container Store catalog #10008759) within non-defrosting standard lab freezers (*e.g.*, Thermo Scientific Revco Ultima II). Heat sealing is an excellent way to preserve the library, as the well chimneys of the storage plate melt slightly and form a bond with the foil seal material. Adhesive seals also work, but they can be more difficult to apply consistently and effectively – they may leave an adhesive residue, thus causing plates to stick

together, or the seals may not bind very well and fall off in storage. However, heat sealing is not without its challenges. After selecting a seal material that is suitable for long-term storage (*e.g.*, aluminum- and plastic-based laminate), one must optimize the sealing time and temperature. These parameters may vary slightly with different types of storage plates and materials, so it is important to thoroughly test the settings.

Whether using heat or adhesive methods, if the seal fails, the sample will evaporate. Differential evaporation is a major concern for RNAi libraries. Over time, seals can become imperfect while in deep -20 °C storage, resulting in partially or completely dried-down samples in some wells across library plates. Partially dried samples have a higher concentration and may yield false toxicity if screened, while fully dried samples will not transfer to assay plates and will potentially cause false negative results. It is important to check libraries for differential evaporation. In instances where this has been observed, we have found that drying the contents of the entire plate (*e.g.*, by placing the plate, unsealed, in a sterile tissue culture hood overnight) and re-suspending the library in nuclease-free water is an effective method for recovery.

EXAMPLE ACADEMIC RNAi SCREENING FACILITY SET-UP

Screening instruments at ICCB-L are set up in workstation format as modules rather than integrated into a fully automated system (Fig. **4**). There are a number of benefits to modularity in a screening room. The flexibility of having individual workstations allows for the rapid addition of new or special equipment for specific protocols. Equipment that is not needed for a particular assay is available for other users, whereas it may be unavailable in a fully automated system even when it is not being utilized as part of an assay in production. In modular mode, the ability to have all workstations in use simultaneously allows for the interleaving or "dove-tailing" of projects, where at any moment there may be separate screens: 1) preparing the transfection reagent and assay plates; 2) delivering the library and forming the transfection complex; 3) preparing cells and adding them to the transfection complex; 4) incubating; and/or 5) reading out assay endpoints. In modular screening facilities that have instrument redundancy, troubleshooting broken or defective equipment does not need to delay screening efforts as it is frequently possible to quickly transfer screening plates to a functional workstation. The workstation format is not without challenges, though. Transfections and assays can be time sensitive, so it is important that the screener defines the effective temporal window for each step and strives to consistently work within it. This can be made manageable by working in smaller batches and

making advance preparations, such as labeling assay plates, measuring reagents and performing necessary calculations. The logistics are nontrivial in this model, as the screener must reserve all necessary pieces of equipment and coordinate their work at the screening facility with the staff and other screeners.
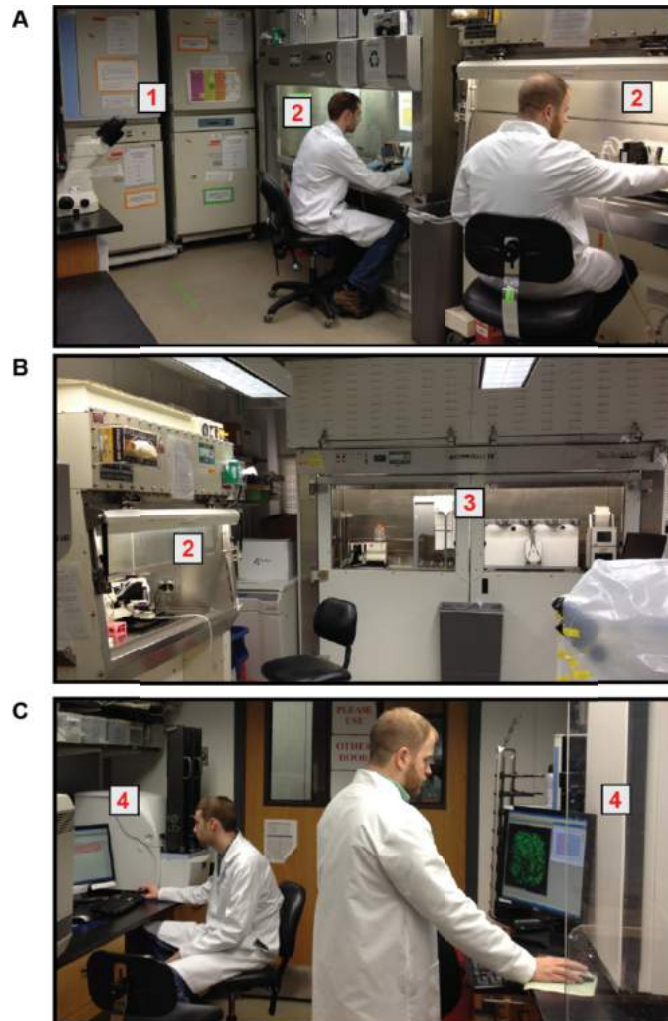
Figure 4: Modular workstation illustration. Several screens have the capacity to operate simultaneously with the use of modular workstations. **A**) Incubators (1) provide space for assays to be incubated. Hoods containing bulk reagent dispensers (2) can be utilized to prepare reagents, fill assay and intermediate plates, prepare cells, plate cells, *etc*. **B**) Workstation (3) dedicated to delivering the library to assay plates and forming the transfection complex. **C**) A variety of instrumentation (4), *e.g.*, plate reader and imager, throughout the lab space is utilized to read out assay end-points.

The ICCB-L RNAi platform employs three types of automated liquid handling during the screening process. Transfection and assay reagent additions are performed using bulk reagent plate fillers, such as the peristaltic pump-based Matrix WellMate (Thermo Scientific), which uses disposable tubing cartridges that can be distributed to users so they have control of their own fluid paths. siRNA library reagents are transferred into assay plates using an automated 384-channel pipettor, the Agilent Bravo. The Bravo is also utilized for all library reformatting. In contrast to the peristaltic pump-based bulk reagent plate fillers, in which wells in rows are filled 8-at-a-time, an automated 384-channel pipettor will transfer fluid to all wells in a microplate simultaneously. A two-channel Tecan EVO75 and an eight-channel Tecan EVO150 pipetting station have been customized to pull screening positives or "cherry picks" from library plates to create custom sets of siRNA reagents for follow-up work. More information about laboratory automation for liquid handling can be found in other overview publications [5].

AUTOMATION OF siRNA TRANSFECTION

The majority of siRNA screens at ICCB-L are performed *via* reverse transfection, where the transfection complex (transfection reagent, Opti-MEM (Life Technologies) and siRNA) is generated within the assay plate and the cell suspension is subsequently added to the complexed siRNA. A typical reverse transfection involves diluting the transfection reagent in Opti-MEM and adding it to assay plates using a WellMate plate filler, then adding the siRNA from library screening plates to assay plates using an automated pipettor, and finally adding the cell suspension with the WellMate. Following 72 to 96 hours of incubation, the assay end-point is reached, samples are processed, and results read out.

The automated pipettor used for library formatting and transfections (*e.g.*, the Agilent Bravo at ICCB-L) should be selected with an emphasis on accuracy at lower volumes and have ample deck space for a variety of applications. It is recommended that disposable tips are utilized with an automated pipettor because they minimize concerns of carry over and library contamination. As disposable automation pipet tips are inherently expensive, it is impractical to change them after siRNA library delivery between assay plate replicates. It is also unreasonable to return to the source plate with potentially dirty pipet tips. Thus serial dispensing is a necessity. Serial dispensing can be defined as aspirating a bulk volume and dispensing a smaller volume to several replicates in series.
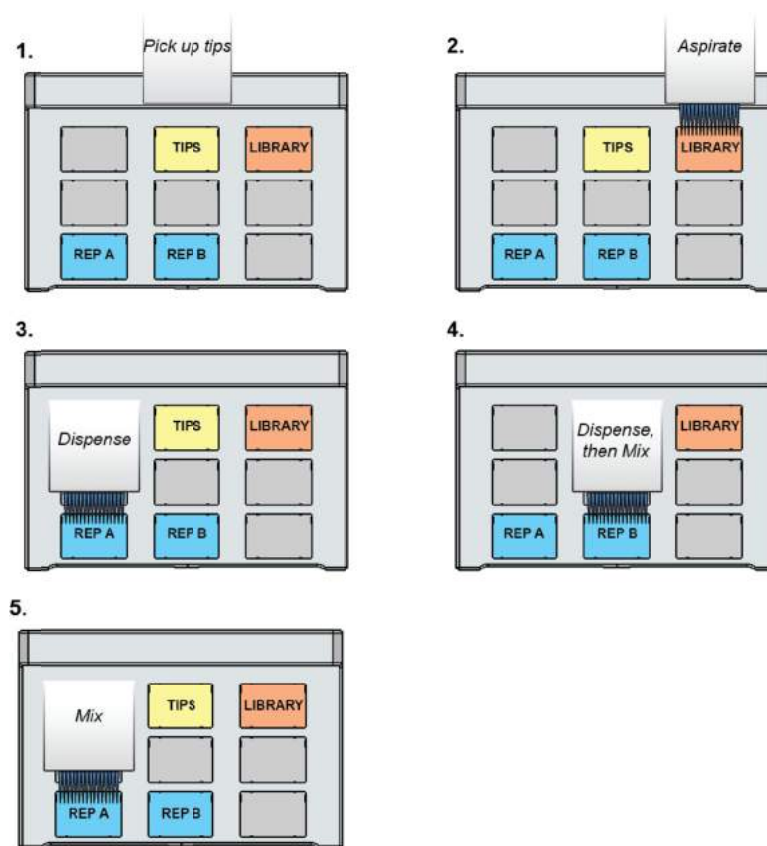
Figure 5: **Schematic of two replicate automated transfection**. **1**) The Bravo (384-channel pipettor) moves to the tip box and attaches tips. **2**) Bravo moves to the Library plate and aspirates 3 µL. **3**) Bravo moves to Replicate (REP) A assay plate (which already contains 8 µL of transfection reagent and Opti-MEM) and dispenses 1.5 µL. **4**) Bravo moves to REP B and dispenses 1.5 µL. It then mixes REP B by aspirating 5 µL (half of the total volume in each well) up and down three times. **5**) Bravo moves back to REP A and mixes by aspirating and dispensing 5 µL three times.

When ICCB-L performed its first automated transfections in 2006, the design closely followed established chemical screening practices. Libraries were screened in duplicate and controls were manually added to assay plates. Most screeners opted for reverse transfection rather than forward transfection because it required fewer steps, was less expensive and easier, and some reported higher transfection efficiency with this method. In one example (Fig. **5**), 3 µL of siRNA was aspirated from the library plate and 1.5 µL was dispensed to each of two assay plates. The transfection complex was then mixed in each assay plate using the automated pipettor. This first generation of screens identified areas for

improvement in the screening strategy. RNAi screens tend to have greater variability compared to chemical screens for a variety of reasons, including differences in cell plating, transfection efficiency, toxicity, and edge effects over longer incubation times [3]. Beginning in 2007, a third replicate was added to all new projects in an effort to improve statistical analyses. The assay-specific control addition step was also automated, thus improving the reproducibility of these controls and enabling screeners to easily increase the number of assay-specific controls per plate. As expected, data quality improved with these changes. In the updated example that illustrates these improvements (Fig. **6**), 4.5 µL of siRNA is aspirated from the library plate and 1.5 µL is dispensed to each of three assay plates. The pipettor then utilizes the same set of tips to aspirate 4.5 µL from a "control" source plate that has a layout complementary to the library and contains only assay-specific siRNA controls (Fig. **6B**). It then dispenses 1.5 µL to each of the three replicates and mixes the transfection complex in each assay plate.

Although most of the screens performed at ICCB-L are run in 384-well format, 96-well screening is also possible. Some assays have technical limitations that necessitate larger well format. For example, assays monitoring rare events require examination of many more cells per test condition or specific plate types may not be available in 384-well format. Many automated liquid handlers offer interchangeable 96- and 384-well pipettor heads, where the 96-well head is able to index to the four quadrants within a 384-well plate, thus it is easy to accommodate a 96-well assay even when library stocks are plated in 384-well format. Screening in 96-well format requires larger volumes of costly reagents and tends to have slower throughput. In addition, library controls may not be evenly distributed amongst all four 96-well plates if siRNAs are aliquoted from 384-well library plates. Thus, this screening strategy is only recommended for assays that have proven to be unsuccessful in 384-well format.

To meet the needs of assay protocols that directly compare two experimental conditions (*e.g.*, effects of siRNA on cancer cells *versus* non-malignant cells), automation protocols have been established for six replicate assays (three replicates per experimental condition). However, automating six assay replicates is not straight forward. The serial dispensing that was used in previously described protocols (Figs. **5** and **6**) proved to be inaccurate across greater numbers of replicates (5+), thus necessitating an alternative approach. Since larger dispense volumes and smaller numbers of replicates tend to be more accurate, an intermediate dilution plate for the siRNA library was created as needed for
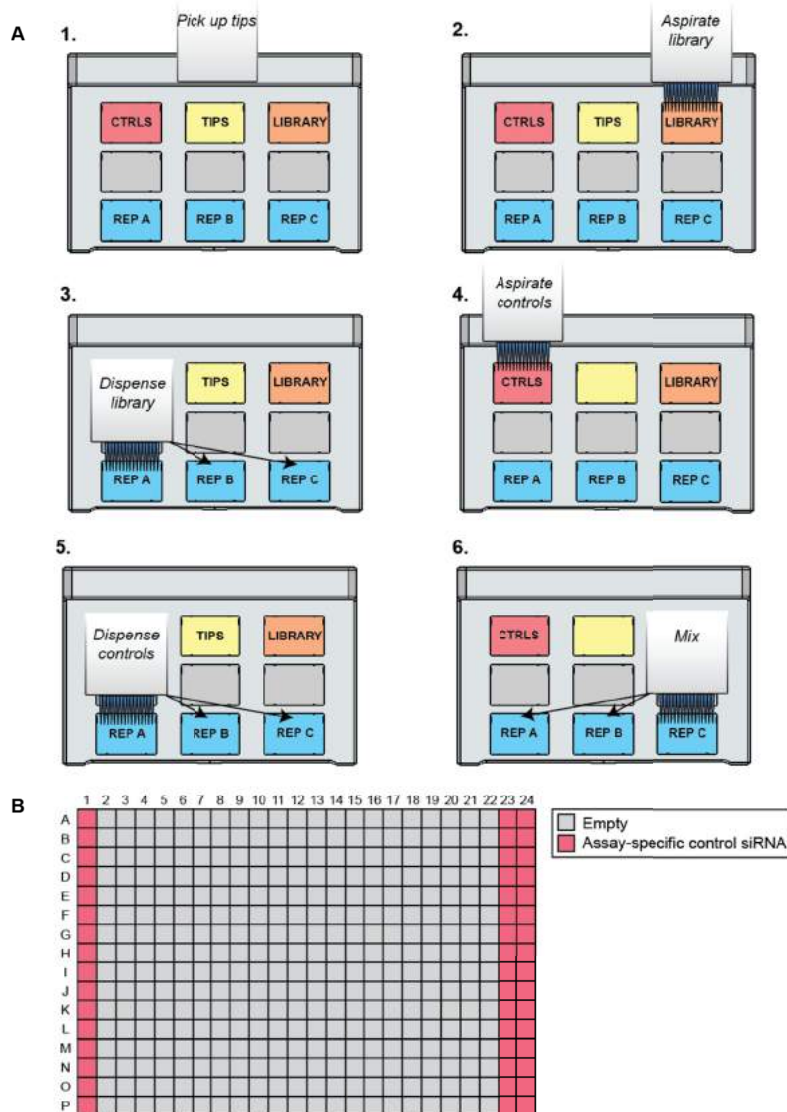
Figure 6: **Schematic of three replicate automated transfection with control plate**. **A) 1**) Bravo moves to the tip box and attaches tips. **2**) Bravo moves to the Library plate and aspirates 4.5 μL. **3**) Bravo moves to REP A, REP B and then REP C assay plates (which already contain 8 μl of transfection reagent and Opti-MEM), dispensing 1.5 μL of the library into each replicate plate. **4**) With the same set of tips, Bravo moves to the Control plate (CTRLS, Fig. **6B**) and aspirates 4.5 μL. **5**) Bravo moves to REP A, REP B and then REP C, dispensing 1.5 μL of the controls into each replicate plate. **6**) Bravo mixes REP C, REP B and then REP A by aspirating and dispensing 5 μL three times. **B**) Control plate, with gray wells indicating empty wells (these correspond to wells that are full in the library plate, as shown in Fig. **1B**) and pink wells representing locations for assay-specific controls.

individual assays. In the two-condition example (Fig. **7**), 10 μL of siRNA is aspirated from the library plate and dispensed into a twin.tec plate containing 10 μL of either 1X siRNA buffer or Opti-MEM, creating a 1:1 dilution of the library.
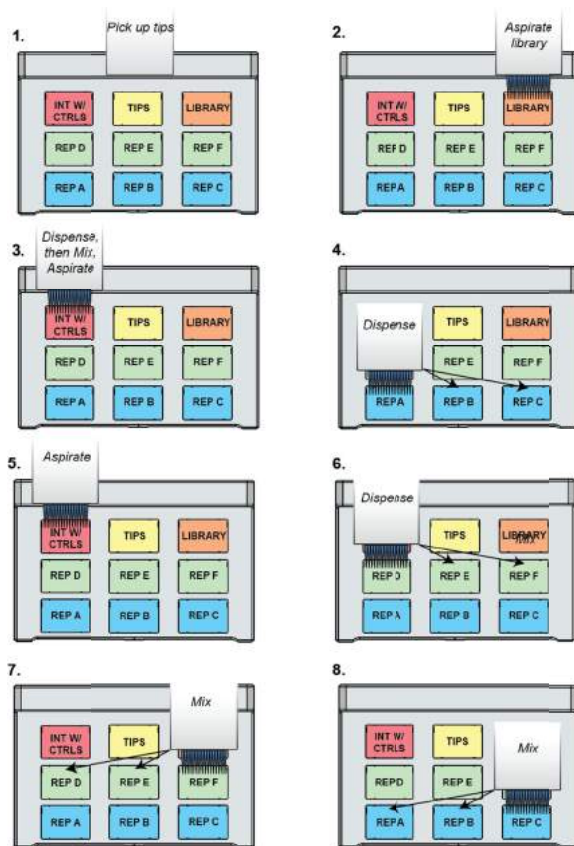
Figure 7: **Three replicate x 2 automated transfection with intermediate plate. 1**) Bravo moves to the tip box and attaches tips. **2**) Bravo moves to the Library plate and aspirates 10 μL. **3**) Bravo moves to the Intermediate plate (INT W/ CTRLS), which contains pre-aliquoted assay-specific controls and 10 μL diluent (Opti-MEM and transfection reagent, Opti-MEM, or 1x siRNA buffer), and dispenses 10 μL of the library. It mixes the Intermediate plate by aspirating and dispensing 10 μL three times and then aspirates 9 μL. **4**) Bravo moves to REP A, REP B and then REP C assay plates (which already contain 7 μl of transfection reagent and Opti-MEM or Opti-MEM), dispensing 3 μL of the diluted library and controls into each plate. **5**) Bravo moves to the Intermediate plate and aspirates another 9 μL. **6**) Bravo moves to REP D, REP E and then REP F assay plates, dispensing 3 μL of the diluted library and controls into each plate. **7**) Bravo mixes REP F, REP E and then REP D by aspirating and dispensing 5 μL three times. **8**) Lastly, Bravo mixes REP C, REP B and then REP A by aspirating and dispensing 5 μL three times. One set of tips is utilized for all of the steps outline above.

This new assay-specific, intermediate copy of the library can be mixed using the automated pipettor and dipped into multiple times with the same set of tips because there is no concern for library contamination. It can also be pre-loaded with assay-specific controls for automated delivery. For transfection, siRNAs from the intermediate copy are dispensed to the six assay plates in two passes, adding twice the volume to each that would have been added from standard library screening plates. The pipettor aspirates 9 µL from the intermediate copy and dispenses 3 µL to each of three replicates, then repeats the process for the remaining three replicates. After all six assay plates have received siRNA, the copies are mixed in reverse order by aspirating and dispensing 50% of the well volume three times.

The initial challenges associated with automating a two-condition assay emphasized the importance of pipettor calibration to validate transfer volumes. Robust calibration can be achieved through a variety of methods, including gravimetric, photometric and fluorometric quantitation. The gravimetric calibration is simple and fast, but it does not provide information about well-to-well variability, and, since the readout is a well average, it is only recommended for larger volumes of 10 µL or more. Fluorometric and photometric calibrations require a series of standards to which each well may be compared and independent values for each well can indicate failing pipettor channels or faulty tips, as well as general accumulation of error across a serial dispense that would not be obvious through gravimetric means. Photometric dyes (*e.g.*, Tartrazine) do not photo-bleach and are more stable than fluorometric markers (*e.g.*, Rhodamine Green, Fluorescein), but they lack the sensitivity that is ideal for quantitating sub-microliter transfers. ICCB-L employs high sensitivity fluorometric calibration for sub-microliter calibrations, a simple photometric method for 1 – 10 µL transfer validations, and a gravimetric approach for any transfer over 10 µL. The gravimetric readout may also be coupled with a photometric reading to identify any outlying wells.

It is possible to perform siRNA transfections on a variety of interfaces beyond traditional tissue culture-treated polystyrene or glass assay plates. Two examples are transwell plates and hydrogel coated plates. A Transwell 96-well plate (Corning) is comprised of two parts: an upper compartment with individual wells and a lower compartment that has either individual wells or a communal trough. These two compartments are separated by a membrane, thus transwell plates can facilitate a wide array of trafficking assays. Hydrogel coated plates [6] can be created with tunable rigidity, more closely simulating the interfaces found in

mammalian tissues compared to glass or polystyrene surfaces. When working with these forms of labware, it is important to carefully consider the pipettor depth and piston speed while aspirating, dispensing and mixing, as close proximity or vigorous liquid handling may disrupt the cell layer or interface.
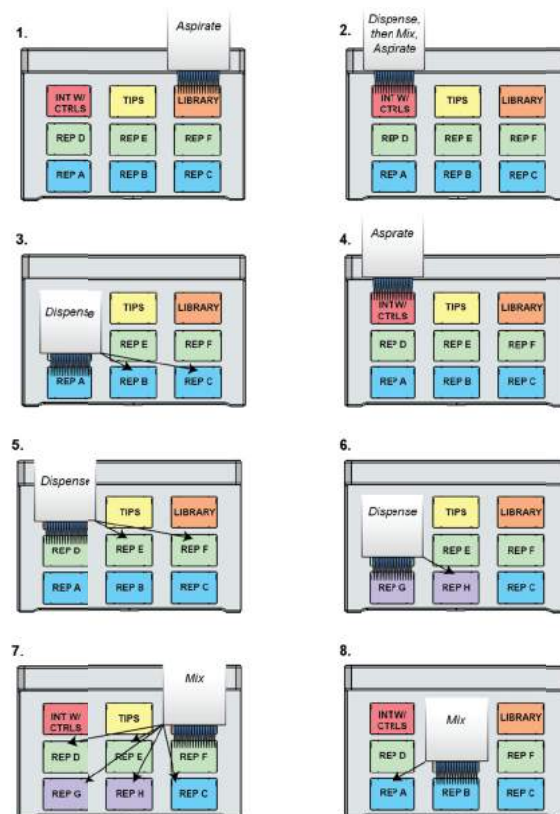
Figure 8: **Three replicate x 2 and two replicate x 1 automated transfection with intermediate plate. 1**) Bravo moves to the tip box, attaches tips (not displayed) and then moves to the Library plate and aspirates 13 μL. **2**) Bravo moves to the Intermediate plate containing assay-specific controls and 13 μL diluent, and dispenses 13 μL of the library. It then mixes the Intermediate plate by aspirating and dispensing 13 μL three times. Bravo aspirates 9 μL. **3**) Bravo moves to REP A, REP B and then REP C, dispensing 3 μL of the diluted library and controls into each plate. **4**) Bravo moves to the Intermediate plate and aspirates another 9 μL. **5**) Bravo moves to REP D, REP E and then REP F, dispensing 3 μL of the diluted library and controls into each plate. **6**) Bravo moves to the Intermediate plate and aspirates another 6 μL (not displayed). REP A and REP B are replaced with REP G and REP H. Bravo moves to REP G and then REP H, dispensing 3 μL of the diluted library and controls into each plate. **7**) Bravo mixes REP F through REP H (F, E, D, C, H and then G) by aspirating and dispensing 5 μL three times. **8**) REP G and REP H are replaced with REP A and REP B. Bravo mixes REP B and then REP A by aspirating and dispensing 5 μL three times. One set of tips is utilized for all of the steps outlines above.

Finally, it should be noted that siRNA screens may also be run without the use of transfection reagents. Electroporation has long been a promising alternative for introducing siRNA into difficult-to-transfect cell lines, but concerns including increased cell death, transfection efficiency and cost of consumables must be addressed. High throughput electroporation platforms have recently become commercially available (*e.g.*, Lonza HT Nucleofector System). Introduction of these platforms into the laboratory automation requires careful logistical planning and testing of each step.

The complexity of the automated transfection can increase as needed. ICCB-L has successfully implemented eight replicate assays, with two experimental conditions in triplicate plus a viability arm in duplicate (Fig. **8**). In addition, a 15 replicate, five cell line forward transfection campaign has been performed. However, these major endeavors are costly, time consuming and require meticulous attention to the timing and logistics of every step.

In conclusion, laboratory automation impacts nearly all aspects of RNAi screening, with some of the greatest influences found in the method of library storage and how siRNA transfections are performed. This important technology can be used to implement simple and complex experimental protocols for a wide variety of RNAi assays.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors confirm that this chapter contents have no conflict of interest.

REFERENCES

[1] Sigma-Aldrich. General siRNA FAQs. How long are siRNAs stable?. Available from: http://www.sigmaaldrich.com/life-science/functional-genomics-and-rnai/sirna/learning-center/mission-sup-reg0/general-sirna-faqs.html#si7.

[2] Lundholt BK, Scudder KM, Pagliaro L. A simple technique for reducing edge effect in cell-based assays. J Biomol Screen. 2003 Oct;8(5): 566-70.

[3] Birmingham A, Selfors LM, Forster T, Wrobel D, Kennedy CJ, Shanks E, *et al*. Statistical methods for analysis of high-throughput RNA interference screens. Nat Methods. 2009 Aug;6(8): 569-75.

[4]   Cheung KH, Hager J, Nelson K, White K, Li Y, Snyder M, *et al*. A dynamic approach to mapping coordinates between microplates and microarrays. J Biomed Inform. 2002 Oct-Dec;35(5-6): 306-12.

[5]   Rudnicki S, Johnston S. Overview of Liquid Handling Instrumentation for High-Throughput Screening Applications. Current Protocols in Chemical Biology: John Wiley & Sons, Inc.; 2009.

[6]   Tse JR, Engler AJ. Preparation of hydrogel substrates with tunable mechanical properties. Curr Protoc Cell Biol. 2010 Jun;Chapter 10: Unit 10 6.

Public Repositories for RNAi Screening Data

Esther E. Schmidt[1,*], Michael S. Banos[2], Jennifer A. Smith[3], Amanda Birmingham[2], Michael Boutros[1] and Caroline E. Shamu[3]

[1]*German Cancer Research Center (DKFZ), Division Signaling and Functional Genomics and Heidelberg University, Department for Cell and Molecular Biology, Medical Faculty Mannheim, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany;* [2]*Dharmacon, part of GE Healthcare, 2650 Crescent Dr., suite 100, Lafayette, CO 80026 USA and* [3]*ICCB-Longwood Screening Facility, Harvard Medical School, 250 Longwood Ave., Seeley G. Mudd Building, Room 604, Boston MA 02115-5731, USA*

Abstract: Public repositories for genomic data, such as sequencing and expression studies, play key roles in the dissemination of large-scale studies. It can be expected that repositories for functional genomic data, such as RNAi screens, will have a similar important role. RNAi data repositories store information about RNAi reagents and results from RNAi screening experiments, and present them in a structured and searchable manner. Implementation and use of robust, public RNAi databases is critical to realizing the potential of RNAi experiments. These databases allow investigators to re-analyze deposited datasets to ask new and different questions, and they are a rich source for functional gene annotation. This chapter describes challenges faced as databases for genome-scale RNAi screening results are developed: the diversity of RNAi assays carried out in multiple cell types and organisms; the variety of identifiers and annotations used to describe RNAi reagents; the lack of an established and accepted ontology to describe RNAi experiments; and the challenge of curating RNAi screening results and collecting complete datasets. Examples of Laboratory Information Management Systems (LIMS) that store RNAi data and of RNAi reagent and result annotation databases are provided.

Keywords: Data annotation, data comparison, data curation, data management, data repository, database, genome-scale RNAi screen, high throughput screening (HTS), RNA interference, RNAi, shRNA, siRNA.

THE POTENTIAL OF RNAi DATA REPOSITORIES

Since the discovery of RNAi in 1998, screening experiments based on this mechanism have moved from experimental applications to highly automated, well-established routine procedures. High-throughput technology enables the

Corresponding author Esther E Schmidt: German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany; Tel: +49 6221 421958; Fax: +49 6221 421959; E-mail: e.schmidt@dkfz.de

efficient production of RNAi knockdown phenotypes on a large scale in a variety of species [1, 2]. The number of large-scale RNAi screening experiments reported in the literature has substantially increased in recent years: a PubMed search for "genome rnai" returns 59 entries for the year 2001 and 552 for the year 2013.

RNAi technology has proven useful in a number of different applications: it has, for example, been used in the elucidation of developmental phenotypes in *Caenorhabditis elegans* [3] and *Drosophila melanogaster* [4], and it serves as a tool for deciphering the structure and regulation of complex signaling pathways [5, 6]. The concept of synthetic lethality-here referring to RNAi double knockdown with the aim to identify such genes that are lethal when abrogated in combination but not individually-is a promising tool in the search for new drug targets in cancer [7, 8].

As more and more RNAi screening experiments are done at genome-wide coverage, they have the potential to provide unbiased answers to research questions and give unexpected new insights into gene functions. For example, Luo *et al.* conducted a genome-wide RNAi screen to identify synthetic lethal interactions with the *KRAS* oncogene (Gene ID: 3845). Among the list of Ras synthetic lethal candidates, they found only a few genes from pathways known to be implicated in Ras-driven oncogenesis. The vast majority of the candidate genes was spread over multiple cellular functions not previously associated with Ras activity, thus opening new avenues in terms of exploiting target genes for therapeutic intervention [9].

Apart from providing answers to a given research question, RNAi screening data may enable other investigators to use an existing dataset to ask new and different questions in addition to what the screen was initially designed for. RNAi data repositories will constitute an important contribution to this end by collecting diverse screening data in one place and presenting them in a structured and searchable manner. Moreover, datasets would be accompanied by well-annotated experimental and data analysis protocols, facilitating reproduction of results as needed.

RNAi screening data are also a rich source for functional gene annotation. The systematic interference with gene expression and the use of a variety of screening assays contributes to the elucidation of gene function across all areas of biology. Collecting this information in a central repository constitutes a valuable task, complementing details currently available in gene catalogues. For example, phenotype data from the GenomeRNAi database [10] have been integrated into

the popular FlyMine data mining tool [11], allowing the user to view and download RNAi phenotypes along with gene information from multiple other data sources.

RNAi technology was first discovered and applied to the model organism *C. elegans* [12], then later to the fruit fly *D. melanogaster* [13]. Such model organisms are far easier to study than mammalian species, especially as they can be subjected to *in vivo* experiments. In addition, their genomes show less redundancy, thus avoiding the difficulty encountered when duplicated genes "make up" for the loss of the gene under examination. Therefore, these model organisms are a popular choice for RNAi experiments and the results are utilized to generate hypotheses for the function of homologous genes in other species. RNAi databases that encompass information on screening data from different species, interlinked *via* gene homology information, represent a very useful resource for cross-species comparison of RNAi results.

Benefits

- Unbiased approach -> unexpected new insights
- Re-analysis of RNAi screening data
- Detailed experimental protocols -> data reproducibility
- Functional gene annotation
- Cross-species comparison
- Meta-analysis across different screens

Challenges

- Variety of biomodels
- Variety of assays
- Complex phenotypes from high-content and *in vivo* experiments
- Data standardization
- Identifier mapping and currency
- Ontology development
- Incomplete data sources
- Manual curation vs. direct data submission

Figure 1: Summary of potential benefits and challenges in providing RNAi screening data repositories.

RNAi screening data collected in data repositories also lend themselves to meta-analyses, spanning multiple screens and/or analysis methods. For example, genes that share a phenotype within a screen could be represented as a network, and overlaid onto networks obtained from other screens, or indeed onto networks based on other genomic data, such as protein-protein interaction or genomic studies. RNAi data could provide new insights into pathway relationships, contributing functional data complementary to other data sources. Another useful application concerns the comparison of RNAi screens. Phenotype results from multiple screens could reveal

central "hub" genes that are involved in many diverse biological mechanisms on the one hand, but comparison across different screens also has the potential to highlight problematic reagents that produce so-called off-target effects, by inactivating partially matching gene transcripts in addition to their intended transcript target. As indicated above, clustering of phenotype data from multiple screens could yield new, possibly unexpected relationships and interactions between genes.

RNAi technology is still comparatively new and its potential is only beginning to be realized. We expect it to continue to grow in significance, making major contributions to the elucidation of gene function and the exploration of interactions across multiple biological pathways. RNAi databases will play a pivotal role in making this data widely available and usable. Fig. **1** summarises the benefits of providing such repositories, as well as challenges faced, as described in the following section.

THE CHALLENGES FACED IN DEVELOPING RNAi DATABASES

The Variety of Assays Makes Data Comparison Difficult

In order for databases to be useful, standardized data representations are needed to allow implementation of efficient search strategies, to enable comparison between different RNAi experiments, and to allow the display of grouped data (*e.g.* all screens performed in a given cell line). The wide variety of assays, which here refers to the experimental strategies to measure phenotypes, being employed in RNAi screening experiments is a major challenge in pursuing data standardization and comparability. RNAi can be performed in whole organisms, *e.g.* in *C. elegans*, *Drosophila* or *Anopheles*, using various reagent delivery methods such as soaking, microinjection, feeding of bacteria or transgenesis [14-16]. Knockdown of gene expression can also be limited to specified tissues by the use of tissue-specific promoters in the vectors utilized for transgenesis [15, 17, 18]. More frequently, RNAi experiments are performed on cultured cells, introducing the reagents directly through the cell culture medium for *Drosophila* cells or *via* siRNA transfection, electroporation or lentiviral infection as required in vertebrate cells [1, 19, 20].

The types of assays used in RNAi experiments vary widely, from single read-out assays (*e.g.* monitoring expression of a luciferase reporter gene) to the observation of multiple complex phenotypic features by high-content microscopy. Microscope images provide a spectrum of measurable features, including cell number, size, shape, arrangement, density, *etc* [21, 22]. Likewise, whole animal experiments also result in diverse sets of data, and can monitor phenotypes such as organism behavior,

organ structure, protein localization in tissues, and many others [23-27]. Even simple read-out results are often combined with other measurements or results from counterscreens carried out in parallel as part of a primary RNAi screen. A typical example would be measurement of inhibition of expression from a specific promoter, combined with the additional condition that the RNAi reagent also meets a certain viability threshold for it to be considered a screening positive. More complex combinations have also been implemented [2, 20]. It should be noted that while most RNAi screens produce quantitative measurements, yielding numerical data, some assay readouts are only textual descriptions of visual observations.

Standardization of Identifiers and Annotations is Needed

RNAi assays cover a range of observations because they constitute individual approaches for any biological question that a researcher might pursue. This results in an abundance of phenotype definitions that is challenging to contain in the sort of structured ontology framework that would be desirable for collection and comparison of RNAi screening data. Moreover, RNAi assay protocols are complex, characterized by a great number of reagents, analysis strategies, and variables that must be well-described in order to enable proper interpretation and comparison of experiments. For example, the Minimum Information About an RNAi Experiment (MIARE) reporting guidelines (http://miare.sourceforge. net/HomePage) include a checklist of more than 50 items required to provide only the minimum information for the description of an RNAi experiment. Another effort to provide a standardized framework for the description of experiments involving cellular assays more generally is Minimum Information About a Cellular Assay (MIACA) (http://miaca.sourceforge.net/). MIACA might also be applied to describing RNAi experiments. However, development of both the MIARE and MIACA guidelines appears to have stalled somewhat, possibly due to the current lack of corresponding data repositories. In the field of microarray expression data, where results are routinely deposited into public databases, the Minimum Information About a Microarray Experiment (MIAME) [28] guidelines represent a very successful example of data standardization. Three major data repositories (ArrayExpress, GEO and CIBEX) accept and distribute MIAME-compliant data, secondary data resources have been developed, and most scientific journals require the deposition of data formatted according to the MIAME standard [29].

To describe the targets of RNAi reagents, a number of commonly accepted identifiers are available, *e.g.* NCBI (National Center for Biotechnology Information) Gene ID [30], Ensembl [31], RefSeq [32], FlyBase [33]. The usage

of multiple identifier types by different authors presents a challenge. When attempting to compare different data sets, it is critical to select "equivalent" identifiers. Ideally, a database uses one type of identifier as its key (typically an internal identifier), which allows the grouping of all relevant data for a particular gene, *e.g.* for display on a single web page. Other identifiers need to be "mapped" to the reference identifier based on known relationships.

Such a mapping procedure is not always straightforward as several issues may complicate matters. These include: a) the other identifier set may not contain an equivalent identifier at all; b) in the case of closely related genes, the other identifier set may contain several possible equivalent identifiers, leaving ambiguity as to which is the correct one; and c) there may be multiple equivalent identifiers due to the type of identifiers used, *e.g.* when mapping a gene identifier to a set of transcript identifiers. Mapping procedures are further complicated by the fact that identifier sets are updated over time, and different identifier resources have different update cycles. Thus, an identifier relationship established at one time point may not be reproducible at a later time point if one or more of the underlying resources have changed. Mapping procedures may have to be repeated regularly in order to avoid too many out-of-date relationships. Similar challenges have been described in the context of array-based expression data, and tools have been developed to address the (re-) mapping of identifiers across different data sets [34-36].

Ontology Development is Essential but Requires a Community Effort

The use of established ontologies and controlled vocabularies to describe experimental reagents and protocols as they are incorporated into databases is essential for data standardization. Ontologies have been actively developed in several areas, each attempting to define terms for the description of data relevant to the respective field. For example, Gene Ontology (GO) [37] is widely used for the annotation of genes, providing term definitions to describe the cellular location, molecular function or biological process of the gene in question. Ontology development relies heavily on community input and consensus in order to assure the quality of the resource and its acceptance by the research community. "The Open Biological and Biomedical Ontologies" (OBO) Foundry has been set up to coordinate such efforts and to provide a framework for efficient integration of the various ontology projects [38]. The RNAi community is in need of a well-structured phenotype ontology in order to facilitate the unambiguous description of phenotypic observations. There is an OBO-candidate ontology, "Mammalian phenotype", under active development [39], which is currently

geared towards the description of mouse knockout phenotypes. The "Phenotypic Quality Ontology" (formerly "Phenotype and Trait Ontology", PATO) [40] provides terms to be used in conjunction with other ontologies to refer to phenotypes, *e.g.* "red", "high temperature", "ectopic". For the annotation of cell-based RNAi phenotypes, however, the challenge remains to develop an ontology, or indeed to define a combination of ontologies, addressing the abundance of phenotypic observations produced by a broad spectrum of biological assays. This task cannot be tackled by an individual research unit alone but would have to be coordinated within a wider representation of the RNAi community.

Comprehensive Data are Critical and Valuable

Comparison of RNAi screening data is impeded by another issue: Many data producers do not make available the complete set of results they have obtained from large-scale screens. Frequently, publications describe large-scale experiments with regard to the technical details, but the actual results reported focus only on "interesting hits". These hits are elaborated on in detailed validation studies and follow-up experiments. Other hits may or may not be mentioned, while "negative" phenotype results are often left out completely. It is not uncommon for a publication on a genome-wide screen, carried out as a high-throughput experiment, to leave out the full list of genes included in the study. As a consequence, valuable data are actually lost. High-content RNAi screens represent a particular challenge due to the difficulty of storing and providing the large amount of raw and processed image files they produce.

Though often not of great interest to the scientist undertaking a high-throughput experiment with a specific biological question in mind, comprehensive results-even negative ones-are valuable assets to the research community as a whole. It makes a considerable difference whether a gene has been abrogated in a study and found to have no effect on a specific biological mechanism, as opposed to a gene not being included in the study at all. Here we perceive a strong need for adequate data repositories to welcome submission of comprehensive data, as such data usually exceed the scope and focus of traditional journals [41].

Comprehensive information about RNAi experiments also includes sequence data on RNAi reagents. These are necessary to assess the specificity of reagents and to annotate RNAi constructs when gene models change; ideally RNAi repositories should require deposition of sequence information for constituent RNAi reagents. A further issue is the appropriate description of controls used in RNAi experiments. Both assay- and library-specific controls are essential for the

interpretation of RNAi screening data. The MIARE and MIACA efforts mentioned above encourage authors to supply sufficient details to ensure proper interpretation and/or re-analysis; this should be supported by RNAi data repositories.

Curation Strategies – How to Populate Data Repositories?

We have stressed the need for comprehensive, well-structured and standardized data repositories for RNAi screening data. A question arises as to which is the best method of populating these with data. Two major approaches can be envisaged, namely manual curation by trained curators or direct submission by data producers.

The first approach relies on dedicated curators who extract information from the literature or review information supplied by data producers or by other means. Curators have been trained in applying annotation guidelines, are frequently involved in developing them in the first place, and are familiar with the complexity of the annotation process. The obvious advantage is a high level of consistency in the data representation, while the disadvantage is the high cost in terms of time and resources, resulting in very slow progress towards a comprehensive data repository. It should be noted that curators can only take into account those details that have actually been published or submitted to them, and there is a risk of mis-interpretation of reported data as the curator has not been involved in the data production at source.

The second approach shifts the task of populating the data repository to the data producers themselves. This requires a suitable infrastructure for uploading the data, well-documented annotation guidelines, and, not least, a strategy to motivate data producers to make the effort to submit their data in the first place. Benefits of this approach are the lower cost involved, and the fact that the data producers know their data best and have the entire data set and all background information at their disposal. Given the variety of experimental setups and the resulting complexity of annotation guidelines, expected drawbacks are a lower level of consistency in the data representation, incompleteness of data submissions, and the relatively steep learning curve for individual investigators to accomplish a limited number of submissions.

Neither of these approaches alone will satisfy the requirements of robust, useful databases, so the likely mode of action will have to be a compromise between the desire for strict data standardization and for simple direct submissions. The

proposed best-case might be a model combining both aspects, namely a staff-assisted data submission procedure. In this model, data producers submit their data themselves, helped along by adequate tools and guidelines, and accompanied by assistance and final review by expert curators. This model has been implemented fairly successfully by the U.S. National Center for Biotechnology Information (NCBI) for their PubChem BioAssay database (http://www.ncbi.nlm.nih.gov/pcassay). GenomeRNAi [10] has also begun to take this approach.

TYPES OF RNAi DATABASES

Databases for storing high throughput screening data can broadly be divided into two types: Laboratory Information Management Systems (LIMS) and result- or reagent-focused annotation databases.

Laboratory Information Management Systems (LIMS)

LIMS store information relevant to laboratory workflow, *e.g.* user and administrative information, in addition to screening results. This type of database is essential for screening facility operations and is usually implemented so that most of the data stored are not publicly accessible, although at least some LIMS systems have been set up so that reagent and screening data stored in them can be made public *via* the LIMS framework when appropriate. Examples of LIMS systems developed in academe that support RNAi screening data management include Screensaver [42], MScreen [43], and FlyRNAi.org [44]. Screensaver and MScreen are open-source projects-software code is freely available for new users to customize-and software for FlyRNAi is available from the developers on request. Several commercially available LIMS packages that support RNAi screening are also available, including ActivityBase (by IDBS) and BioRails (by Accelrys).

Because LIMS systems integrate administrative and experimental workflows within laboratories, they are usually created to be dynamic and customizable. Many facilities that use LIMS employ at least one staff software developer to carry out ongoing improvements and to adapt the LIMS as lab policies and workflows change. On the administrative side, LIMS store and report on information about screens relevant to operation of the screening facility. This can include information about investigators that perform screens (*e.g.*, contact information), critical correspondence pertaining to investigator work in the screening facility, and, if relevant, billing and accounting information. LIMS also store important screen-specific information, such as details about reagents used in

screens (*e.g.*, sequence and catalogue information), well-described assay and data analysis protocols, information about RNAi library copies used for specific screening experiments, dates on which screening activities were performed, and perhaps files that record primary screen hits chosen for follow-up work.

A key function of any LIMS is to match screening results with the contents of library wells that were screened. Thus, a LIMS stores information about RNAi library contents and plate formats in addition to screen data. Many LIMS link out to external, public databases, such as GenomeRNAi or the NCBI Probe database (see below), for additional information describing RNAi reagents. Most LIMS also facilitate library management within the facility, allowing tracking of library plate copies and storing information about plate locations, well volumes, accumulated freeze/thaw cycles, *etc*.

For screening results, LIMS typically store both raw and processed data, recording plate and well layout of assay results for plate-based screens, as well as the identities of and data from experimental controls-including their positions on assay plates. This information enables anyone accessing the screen data to re-analyze the data as necessary-for example, to perform per plate normalization, Z' factor calculations, *etc*-and also to compare results directly with those from other screens carried out in the facility. In public databases where the primary goal is to provide annotation for reagents and screening data (see the section below), raw screen data is usually not available, nor is layout information for library and assay wells or detailed information about controls. Thus, re-analysis of experimental data from these databases is somewhat limited and original analyses cannot be perfectly reproduced from the information provided.

Result- or Reagent-Focused Annotation Databases

While LIMS databases are typically internally facing, a considerable number of public databases aggregating various data about RNAi reagents and their experimental results have also been developed (for a partial listing, see Table **1**). Several of particular interest are highlighted below.

Probe - http://www.ncbi.nlm.nih.gov/probe

The Probe database [45] from NCBI is a registry of nucleic acid screening reagents. Information currently available consists of individual and/or pooled siRNAs and shRNAs offered by providers such as The RNAi Consortium, Dharmacon (part of GE Healthcare), Life Technologies, and Sigma-Aldrich, although reagent listings may be incomplete. For each reagent record, the

expected target location in the current version of the genome is shown and a link to the respective supplier's website is given. Sequence information is also available for a subset of reagents.

PubChem - http://pubchem.ncbi.nlm.nih.gov/

NCBI's PubChem [45] is an aggregation of the databases Substance, BioAssay, and Compound (which is largely irrelevant for RNAi screening). Information on screening samples, including user-submitted RNAi reagents such as siRNAs and shRNAs, is contained in Substance, with each substance record including links to an appropriate Probe record where relevant. Dharmacon (part of GE Healthcare) has begun submitting its RNAi reagents to Substance, making them available for experimenters to use in depositing experimental results generated with these reagents to BioAssay.

BioAssay is "a public repository for biological activity data of small molecules and RNAi reagents" [46] that focuses on information regarding whether or not a tested reagent was determined to be "active" by the testing researcher's criteria of interest. Although predominantly used for small molecule data to date, BioAssay's capacities have been extended to enable it to support submission of information on RNAi screens, such as those done with siRNAs, shRNAs, and dsRNAs against a variety of organisms. BioAssay also now supports deposition of MIARE-compliant reagent and assay protocol information. These data from BioAssay are available to download either on a per assay or a per reagent-of-interest basis, although it lacks full information about plate layouts, tested controls, and raw screen data that would enable further secondary analysis.

GenomeRNAi - http://www.genomernai.org

GenomeRNAi [10] provides information for human and *Drosophila* RNAi libraries from reagent providers. When available either as a result of literature curation or experimenter-submitted data, associated phenotypic information for a gene/reagent is provided, as well as reagent sequences where approved by the provider. Direct submission of experimental data is encouraged, and guidelines for proper annotation of data using a controlled vocabulary are provided-in effect standardizing the way data can later be consumed. A further step towards data standardization is taken by regularly mapping author-provided gene identifiers to the corresponding NCBI Gene ID. Additionally, reagent specificity and efficiency information are updated regularly using NEXT-RNAi [47].

As well as listing RNAi reagents, the GenomeRNAi website provides the ability to browse, search, and download data; one may also view and download frequent hitters (lists of genes frequently showing a phenotype) as well as links to other resources such as FlyBase, FlyMine, UniProt, and GeneCards. Furthermore, GenomeRNAi data can be accessed *via* a DAS server [48], and the website includes a DAS-based genome browser.

RNAiAtlas - http://rnaiatlas.ethz.ch/

RNAiAtlas [49] acts as an annotation resource for commercially available siRNAs. It includes both annotations provided by the manufacturer and those generated by an independent annotation pipeline based on publicly available transcriptomic data. Reagent annotations include target site details for predicted intended target and off-target transcripts, providing a granular view of which genes and individual variants may be affected in an experiment. Versioned copies of older annotations can also be viewed, allowing users to see how expected targets have changed as transcript information has evolved. Reagent sequences are available to registered users who have a product license from the reagent supplier.

RNAi Codex - http://cancan.cshl.edu/cgi-bin/Codex/Codex.cgi

The RNAi Codex database [50] is composed of sets of public shRNA designs, providing full genome coverage for human, mouse, rat, and *Drosophila*, as well as partial coverage of the *Arabidopsis* genome. All designs are based on the miR-30 backbone, and include the full hairpin and a barcode sequence. Registration provides access to additional designs.

WormBase - http://www.wormbase.org/

WormBase [51] is a repository for a large collection of RNAi reagents for worms, predominately for *C. elegans*. For each reagent listed, users have the ability to contribute experimental data pertaining to affected phenotypes. Bi-monthly data releases allow researchers to consume up-to-date repository information offline in addition to exploring the contents through an online search interface. In addition to reagent information, WormBase maintains an active user base, offering general resources for the research community, including a discussion board, meeting announcements, and links to other relevant research resources.

FLIGHT, ParameciumDB, HIVsirDB, and VIRsiRNAdb

Several additional databases have emerged catering to the research communities for specific target organisms. These include, but are not limited to, FLIGHT

(http://flight.icr.ac.uk/) [52], ParameciumDB (http://paramecium.cgm.cnrs-gif.fr/) [53], HIVsirDB (http://crdd.osdd.net/raghava/hivsir/) [54], and VIRsiRNAdb (http://crdd.osdd.net/servers/virsirnadb/) [55] for *Drosophila*, *Paramecium tetraurelia*, HIV, and human viruses, respectively. All provide varying levels of detail (from reagent type and intended target to experimental efficacy) for screening reagents targeted to their respective organism and allow users to contribute their own experimental data, promoting a community-driven approach to building RNAi knowledgebases. In addition to providing data for siRNA intended targets, HIVsirDB and VIRsiRNAdb also offer experimental data for efficacy of reagents against viral escape sequences when available.

Table 1: Summary of selected public RNAi databases

Database	URL	Organism (s)	Reagent(s)	Downloads
FLIGHT	http://flight.icr.ac.uk/	*Drosophila*, human	dsRNA	Yes
FlyRNAi.org	http://www.flyrnai.org/index.html	*Drosophila*	dsRNA	Yes
GenomeRNAi	http://genomernai.org/	Human, *Drosophila*	dsRNA, siRNA, shRNA, UAS-IR construct	Yes
HIVsirDB	http://crdd.osdd.net/raghava/hivsir/	HIV	sRNA	Yes
NIH siRNA database	http://www.biomedcentral.com/content/supplementary/1471-2105-7-65-s4.csv	Human, mouse, *Drosophila*, hamster	siRNA	Yes
NKI shRNA library	http://screeninc.nki.nl/downloads/Supplemental_data_NCB.xls	Human	shRNA	Yes
Novartis siRNA database	http://www.nature.com/nbt/journal/v23/n8/extref/nbt1118-S4.pdf	Human, rat	siRNA	Yes
ParameciumDB	http://paramecium.cgm.cnrs-gif.fr/page/index	*P. tetraurelia*	dsRNA	Yes
PhenomicDB	http://www.phenomicdb.de/	*C. elegans, D. discoideum, Drosophila*, human, mouse, yeast, zebrafish	sRNA, dsRNA	Yes, upon request
Probe	http://www.ncbi.nlm.nih.gov/probe	Human, mouse, *Drosophila*, rat, *C. elegans C. intestinalis*, zebrafish, HIV, hepatitis B, dog, *Gallus gallus*, rice grassy stunt virus,	sRNA, morpholino	Yes

Table 1: contd….

		SARS coronavirus, *Oryza sativa* endornavirus, human rhinovirus, firefly, bull		
PubChem	http://pubchem.ncbi.nlm.nih.gov/	Human, rat, *Drosophila*	sRNA	Yes
RNAi Atlas	http://rnaiatlas.ethz.ch/	Human	siRNA	No
RNAi Codex	http://codex.cshl.org/cgi-bin/Codex/Codex.cgi	Human, mouse, rat, *Drosophila*, *Arabidopsis*	shRNA	Yes
RNAiDB	http://www.rnai.org/	*C. elegans*	dsRNA	No
RNAimmuno	http://rnaimmuno.ibch.poznan.pl	Human	multiple types	No
siRNAdb	http://sirna.sbc.su.se/sirnadb_050915.txt	Human	siRNA	Yes
University of Tokyo siRNA database	http://nar.oxfordjournals.org/content/suppl/2007/03/28/gkl1120.DC1/nar-02060-met-g-2006-File012.xls	Human	siRNA	Yes
VIRsiRNAdb	http://crdd.osdd.net/servers/virsirnadb/	Human virus	sRNA	No
WormBase	http://www.wormbase.org	*C. brenneri, C. briggsae, C. elegans, C. remanei, P. pacificus*	dsRNA	Yes

CONCLUSION

The growth of new high-throughput technologies such as whole-genome RNAi screening (as well as microarrays and next-generation sequencing) over the last decade has exponentially increased the volume of biological data being generated. However, because data-management strategies have not matured as quickly as data-generation technologies, the full potential of this tidal wave of biological information has not yet been realized. Aggregation of these data into robust databases is critical but challenging due to the wide range of assays and the variety of annotations that hamper a standardized description of RNAi data.

LIMS for tracking reagent inventory, raw data, and progress through the experimental process are indispensible to labs and/or screening facilities actually performing high-throughput RNAi experiments, as evidenced by the independent development of several such systems. Likewise, public databases of reagents and published phenotypes from the full range of RNAi screens are crucial to the larger

community, since combination of these data offer unprecedented opportunities for mitigating limitations of the RNAi technique (such as off-target effects) and for generating novel biological insights. Recognizing this need, a considerable number of RNAi-related databases have been developed to serve various constituencies.

The field as a whole is only beginning to appreciate the benefits of submitting full screening datasets-including sequences of RNAi reagents, and both positive and negative results-to public data repositories. Obstacles to establishing widely accepted and utilized repositories for RNAi data (such as the need for more curation efforts, for appropriate ontologies, and for techniques to manage large image collections) are considerable but not insurmountable. We must vigorously pursue the goal of comprehensive data repositories if we hope to ride the wave of RNAi high-throughput data, rather than be left floundering in the surf.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors confirm that this chapter contents have no conflicts of interest.

REFERENCES

[1]	Boutros M, Ahringer J. The art and design of genetic screens: RNA interference. Nat Rev Genet. 2008;9(7):554-66. Epub 2008/06/04.
[2]	Demir K, Boutros M. Cell perturbation screens for target identification by RNAi. Methods Mol Biol. 2012;910:1-13. Epub 2012/07/24.
[3]	Sonnichsen B, Koski LB, Walsh A, Marschall P, Neumann B, Brehm M, *et al*. Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. Nature. 2005;434(7032):462-9. Epub 2005/03/26.
[4]	Aikin R, Cervantes A, D'Angelo G, Ruel L, Lacas-Gervais S, Schaub S, *et al*. A genome-wide RNAi screen identifies regulators of cholesterol-modified hedgehog secretion in Drosophila. PLoS One. 2012;7(3):e33665. Epub 2012/03/21.
[5]	Fisher KH, Brown S, Zeidler MP. Designing RNAi screens to identify JAK/STAT pathway components. Methods Mol Biol. 2013;967:81-97. Epub 2013/01/09.
[6]	Chia J, Goh G, Racine V, Ng S, Kumar P, Bard F. RNAi screening reveals a large signaling network controlling the Golgi apparatus in human cells. Mol Syst Biol. 2012;8:629. Epub 2012/12/06.
[7]	Michiue H, Eguchi A, Scadeng M, Dowdy SF. Induction of *in vivo* synthetic lethal RNAi responses to treat glioblastoma. Cancer Biol Ther. 2009;8(23):2306-13. Epub 2009/11/11.
[8]	Scholl C, Frohling S, Dunn IF, Schinzel AC, Barbie DA, Kim SY, *et al*. Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells. Cell. 2009;137(5):821-34. Epub 2009/06/06.

[9]   Luo J, Emanuele MJ, Li D, Creighton CJ, Schlabach MR, Westbrook TF, *et al*. A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. Cell. 2009;137(5):835-48. Epub 2009/06/06.

[10]   Schmidt EE, Pelz O, Buhlmann S, Kerr G, Horn T, Boutros M. GenomeRNAi: a database for cell-based and *in vivo* RNAi phenotypes, 2013 update. Nucleic Acids Res. 2013;41(D1):D1021-6. Epub 2012/11/30.

[11]   Lyne R, Smith R, Rutherford K, Wakeling M, Varley A, Guillier F, *et al*. FlyMine: an integrated database for Drosophila and Anopheles genomics. Genome Biol. 2007;8(7):R129. Epub 2007/07/07.

[12]   Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998;391(6669):806-11. Epub 1998/03/05.

[13]   Kuttenkeuler D, Boutros M. Genome-wide RNAi as a route to gene function in Drosophila. Brief Funct Genomic Proteomic. 2004;3(2):168-76. Epub 2004/09/10.

[14]   Conte D, Jr., Mello CC. RNA interference in Caenorhabditis elegans. Curr Protoc Mol Biol. 2003;Chapter 26:Unit 26 3. Epub 2008/02/12.

[15]   Perrimon N, Ni JQ, Perkins L. *In vivo* RNAi: today and tomorrow. Cold Spring Harb Perspect Biol. 2010;2(8):a003640. Epub 2010/06/11.

[16]   Blandin S, Shiao SH, Moita LF, Janse CJ, Waters AP, Kafatos FC, *et al*. Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector Anopheles gambiae. Cell. 2004;116(5):661-70. Epub 2004/03/10.

[17]   Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, Fellner M, *et al*. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. Nature. 2007;448(7150):151-6. Epub 2007/07/13.

[18]   Kleinhammer A, Wurst W, Kuhn R. Constitutive and conditional RNAi transgenesis in mice. Methods. 2011;53(4):430-6. Epub 2010/12/28.

[19]   Mohr SE, Perrimon N. RNAi screening: new approaches, understandings, and organisms. Wiley Interdiscip Rev RNA. 2012;3(2):145-58. Epub 2011/09/29.

[20]   Bakal C. Drosophila RNAi screening in a postgenomic world. Brief Funct Genomics. 2011;10(4):197-205. Epub 2011/07/15.

[21]   Fuchs F, Pau G, Kranz D, Sklyar O, Budjan C, Steinbrink S, *et al*. Clustering phenotype populations by genome-wide RNAi and multiparametric imaging. Mol Syst Biol. 2010;6:370. Epub 2010/06/10.

[22]   Rohn JL, Sims D, Liu T, Fedorova M, Schock F, Dopie J, *et al*. Comparative RNAi screening identifies a conserved core metazoan actinome by phenotype. J Cell Biol. 2011;194(5):789-805. Epub 2011/09/07.

[23]   Saj A, Arziman Z, Stempfle D, van Belle W, Sauder U, Horn T, *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(5):862-76. Epub 2010/05/25.

[24]   Mummery-Widmer JL, Yamazaki M, Stoeger T, Novatchkova M, Bhalerao S, Chen D, *et al*. Genome-wide analysis of Notch signalling in Drosophila by transgenic RNAi. Nature. 2009;458(7241):987-92. Epub 2009/04/14.

[25]   Fu C, Whitfield CW. Genes associated with honey bee behavioral maturation affect clock-dependent and -independent aspects of daily rhythmic activity in fruit flies. PLoS One. 2012;7(5):e29157. Epub 2012/05/19.

[26]   Kalis AK, Kroetz MB, Larson KM, Zarkower D. Functional genomic identification of genes required for male gonadal differentiation in Caenorhabditis elegans. Genetics. 2010;185(2):523-35. Epub 2010/03/24.

[27]   Pospisilik JA, Schramek D, Schnidar H, Cronin SJ, Nehme NT, Zhang X, *et al*. Drosophila genome-wide obesity screen reveals hedgehog as a determinant of brown *versus* white adipose cell fate. Cell. 2010;140(1):148-60. Epub 2010/01/16.

[28]   Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, *et al*. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet. 2001;29(4):365-71. Epub 2001/12/01.

[29]   Brazma A. Minimum Information About a Microarray Experiment (MIAME)--successes, failures, challenges. ScientificWorldJournal. 2009;9:420-3. Epub 2009/06/02.

[30]   Maglott D, Ostell J, Pruitt KD, Tatusova T. Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res. 2011;39(Database issue):D52-7. Epub 2010/12/01.

[31]   Flicek P, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, *et al*. Ensembl 2012. Nucleic Acids Res. 2012;40(Database issue):D84-90. Epub 2011/11/17.

[32]   Pruitt KD, Tatusova T, Brown GR, Maglott DR. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. Nucleic Acids Res. 2012;40(Database issue):D130-5. Epub 2011/11/29.

[33]   McQuilton P, St Pierre SE, Thurmond J. FlyBase 101--the basics of navigating FlyBase. Nucleic Acids Res. 2012;40(Database issue):D706-14. Epub 2011/12/01.

[34]   Ballester B, Johnson N, Proctor G, Flicek P. Consistent annotation of gene expression arrays. BMC Genomics. 2010;11:294. Epub 2010/05/13.

[35]   Day RS, McDade KK, Chandran UR, Lisovich A, Conrads TP, Hood BL, *et al*. Identifier mapping performance for integrating transcriptomics and proteomics experimental results. BMC Bioinformatics. 2011;12:213. Epub 2011/05/31.

[36]   Chavan SS, Shaughnessy JD, Jr., Edmondson RD. Overview of biological database mapping services for interoperation between different 'omics' datasets. Hum Genomics. 2011;5(6):703-8. Epub 2011/12/14.

[37]   Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, *et al*. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25-9. Epub 2000/05/10.

[38]   Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, *et al*. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol. 2007;25(11):1251-5. Epub 2007/11/09.

[39]   Smith CL, Eppig JT. The Mammalian Phenotype Ontology as a unifying standard for experimental and high-throughput phenotyping data. Mamm Genome. 2012;23(9-10):653-68.

[40]   Gkoutos GV, Green EC, Mallon AM, Hancock JM, Davidson D. Using ontologies to describe mouse phenotypes. Genome Biol. 2005;6(1):R8. Epub 2005/01/12.

[41]   Shamu CE, Wiemann S, Boutros M. On target: a public repository for large-scale RNAi experiments. Nat Cell Biol. 2012;14(2):115. Epub 2012/02/03.

[42]   Tolopko AN, Sullivan JP, Erickson SD, Wrobel D, Chiang SL, Rudnicki K, *et al*. Screensaver: an open source lab information management system (LIMS) for high throughput screening facilities. BMC Bioinformatics. 2010;11:260. Epub 2010/05/21.

[43]   Jacob RT, Larsen MJ, Larsen SD, Kirchhoff PD, Sherman DH, Neubig RR. MScreen: an integrated compound management and high-throughput screening data storage and analysis system. J Biomol Screen. 2012;17(8):1080-7. Epub 2012/06/19.

[44]   Flockhart IT, Booker M, Hu Y, McElvany B, Gilly Q, Mathey-Prevot B, *et al*. FlyRNAi.org--the database of the Drosophila RNAi screening center: 2012 update. Nucleic Acids Res. 2012;40(Database issue):D715-9. Epub 2011/11/10.

[45]   Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, *et al*. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2006;34(Database issue):D173-80. Epub 2005/12/31.

[46]   Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Zhou Z, *et al*. PubChem's BioAssay Database. Nucleic Acids Res. 2012;40(Database issue):D400-12. Epub 2011/12/06.

[47]   Horn T, Sandmann T, Boutros M. Design and evaluation of genome-wide libraries for RNA interference screens. Genome Biol. 2010;11(6):R61.

[48]   Jenkinson AM, Albrecht M, Birney E, Blankenburg H, Down T, Finn RD, *et al*. Integrating biological data--the Distributed Annotation System. BMC Bioinformatics. 2008;9 Suppl 8:S3. Epub 2008/08/05.

[49]   Mazur S, Csucs G, Kozak K. RNAiAtlas: a database for RNAi (siRNA) libraries and their specificity. Database (Oxford). 2012;2012:bas027. Epub 2012/06/16.

[50]   Olson A, Sheth N, Lee JS, Hannon G, Sachidanandam R. RNAi Codex: a portal/database for short-hairpin RNA (shRNA) gene-silencing constructs. Nucleic Acids Res. 2006;34(Database issue):D153-7. Epub 2005/12/31.

[51]   Yook K, Harris TW, Bieri T, Cabunoc A, Chan J, Chen WJ, *et al*. WormBase 2012: more genomes, more data, new website. Nucleic Acids Res. 2012;40(Database issue):D735-41. Epub 2011/11/10.

[52]   Sims D, Bursteinas B, Jain E, Gao Q, Baum B, Zvelebil M. The FLIGHT Drosophila RNAi database: 2010 update. Fly (Austin). 2010;4(4):344-8. Epub 2010/09/22.

[53]   Arnaiz O, Sperling L. ParameciumDB in 2011: new tools and new data for functional and comparative genomics of the model ciliate Paramecium tetraurelia. Nucleic Acids Res. 2011;39(Database issue):D632-6. Epub 2010/10/19.

[54]   Tyagi A, Ahmed F, Thakur N, Sharma A, Raghava GP, Kumar M. HIVsirDB: a database of HIV inhibiting siRNAs. PLoS One. 2011;6(10):e25917. Epub 2011/10/25.

[55]   Thakur N, Qureshi A, Kumar M. VIRsiRNAdb: a curated database of experimentally validated viral siRNA/shRNA. Nucleic Acids Res. 2012;40(Database issue):D230-6. Epub 2011/12/06.

Pooled shRNA Screening

Annaleen Vermeulen[2], Anja van Brabant Smith[2], Sarah B. Anderson[3], Roderick L. Beijersbergen[1,*] and Kaylene J. Simpson[4,5]

[1]*The Netherlands Cancer Institute, Plesmanlaan 121, 1066CX, Amsterdam, The Netherlands;* [2]*Dharmacon, part of GE Healthcare, 2650 Crescent Dr., Suite 100, Lafayette, CO 80026, United States of America;* [3]*Challenge Technology, Inc., 4950 Ward Road, Wheat Ridge, CO 80033, USA;* [4]*Victorian Centre for Functional Genomics, Peter MacCallum Cancer Centre, East Melbourne, Victoria, 3002, Australia and* [5]*Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, Victoria 3052, Australia*

Abstract: Short hairpin RNA interference (shRNA) screens have earned their place in the technical repertoire of high throughput screening approaches by virtue of their broad applicability to targeting regular and primary cell types and the capacity to perform both positive and negative selection screens both *in vitro* and *in vivo*. This chapter focuses primarily on pooled shRNA screens, outlining the breadth of resources available, important library features and methods to establish effective transduction. We discuss assay development and optimization, followed by strategies for hit identification, principally using Next Generation Sequencing (NGS) approaches. Validation of any screen is essential and our collective experience guides the reader to consider a range of approaches towards confirming targets identified in the screen subsequently recapitulate the biological premise of the screen. We conclude with a thought provoking discussion on the future of shRNA screens, the challenges and the scope we can look forward to.

Keywords: Next Generation Sequencing, pooled screens, RNA interference, shRNA, validation.

INTRODUCTION

Large-scale screens using RNA interference (RNAi) have demonstrated utility in identifying gene targets that play a role in specific biological pathways or disease progression. RNAi screens have been performed using synthetic small interfering RNAs (siRNAs), synthetic microRNA mimics or inhibitors, or expressed shRNAs to trigger the gene silencing event. While RNAi screens using synthetic siRNAs have been widely used [1-12], the applications for RNAi screens using synthetic siRNA reagents are limited by the requirement that cells used in the screen need

*Corresponding author Roderick L. Beijersbergen: The Netherlands Cancer Institute, Plesmanlaan 121, 1066CX, Amsterdam, The Netherlands; Tel: +31 20 512 1960; E-mail: r.beijersbergen@nki.nl

to be relatively easy to transfect and the phenotype of interest should be apparent before the effects of gene silencing diminish. Screens using expressed shRNAs circumvent many of these limitations. Virally-expressed shRNAs can be transduced into most cell types and stable expression of the shRNAs allows analysis of phenotypes that require longer times to develop.

shRNA screens can be performed in either an arrayed or pooled format. In arrayed screening, individual shRNAs are distributed across individual wells of multi-well culture plates and phenotypes are screened on a well-by-well basis. Arrayed shRNA screens have been used to study a variety of processes, including circadian rhythm and mitotic progression [13, 14]. In pooled shRNA screens, hundreds or thousands of different shRNAs are introduced into a population of cells that are then selected for the phenotype of interest; shRNAs that are either enriched or depleted in the selected population are identified by PCR amplification of the shRNA from genomic DNA (gDNA); changes in relative abundance of the individual shRNAs between control and experimental cell populations are evaluated using microarray or NGS technologies (Fig. **1**). Pooled shRNA screens have been used to identify genetic components of various cellular processes that have been similarly interrogated by siRNA screens including cell proliferation [15], tumorigenicity [16, 17], adhesion [18], and migration [19] in cell culture systems. However, there are applications and biological questions that can currently only be addressed using pooled shRNA screening approaches; these include analysis of large gene sets across hundreds of cell lines and *in vivo* screens that will be discussed in the future directions section.

For the purposes of this review, we will focus on pooled shRNA screens and the important factors that should be considered in order to ensure that meaningful results are obtained from these screens. A successful pooled shRNA screen relies on the quality of the library, careful experimental design, hit identification and hit validation. Finally, we will discuss how pooled shRNA screening has been applied in novel areas and where pooled shRNA screening will take us in the future.

AVAILABLE shRNA LIBRARIES

The earliest collections of shRNA vectors were generated by academic efforts. Initially, these targeted a small number of genes with three to five individual shRNAs for each gene [20]. Subsequently, larger collections spanning thousands of genes were generated by the Netherlands Cancer Institute [20, 21] and
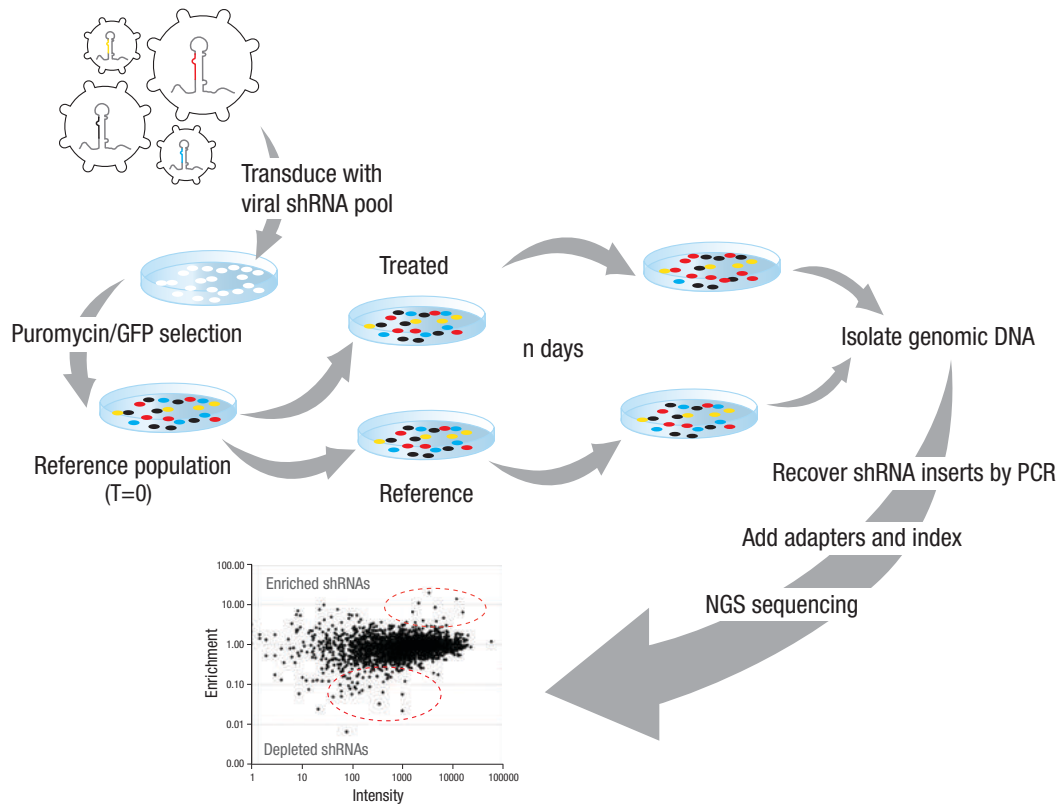
Figure 1: Pooled shRNA screening workflow *in vitro*. A population of cells are transduced with an shRNA library such that on average each cell contains a single shRNA with a recommended 500-1000 copies of each shRNA in the cell population. Depending on the shRNA vector used, cells can be subjected to antibiotic selection or sorted based on fluorescence to generate a population of cells expressing only hairpins. These cells are split into a reference sample that remains untreated (T_0) and a sample that is subjected to a selective pressure (T_1), ranging simply from growth for multiple passages to drug dosage to identify resistant or sensitive shRNA targets. Following a certain time period, the genomic DNA of T_0 and T_1 samples are isolated, shRNA-associated sequences are amplified and NGS is applied to count the number of shRNA sequences in each sample to determine the enriched and depleted shRNA in the experiment, resulting in a statistically ranked hit list.

Hannon/Elledge groups [22, 23] and these were made available to the scientific community. The first generation shRNA vectors were based on the expression of a small hairpin RNA under the control of H1 or U6 pol III promoters. These shRNAs consisted of a gene specific sequence between 19 and 27 bases in length, a loop sequence of four to 12 bases (sometimes including a restriction site),

followed by the reverse complement of the gene specific sequence terminating in a stretch of five T's. Further insight into endogenously expressed siRNAs led to the generation of microRNA-embedded shRNAs. In such systems, the shRNA is placed into the scaffold of an endogenous microRNA, *e.g.* miR-30. The microRNA-based hairpins resemble native microRNA structures and are processed by the endogenous microRNA pathway, potentially leading to more efficient knockdown of gene expression. These microRNAs can be expressed using RNA polymerase II promoters, allowing for promoters with cell- or tissue type specific expression. In addition, inserts can be generated in which the microRNA is fused to a reporter gene (*e.g.* GFP) to monitor expression. For both systems, inducible vectors have been generated to allow for sequential activation and inactivation of gene expression in stable cell lines. These promoter and shRNA structures are usually incorporated into viral vector systems, including retroviral, adeno-associated viral, adenoviral and lentiviral platforms.

Currently, there are several large collections of shRNA vectors available (reviewed in [24]), including the Netherlands Cancer Institute (NKI) libraries [20, 21], Hannon-Elledge libraries [22, 23] (available from Dharmacon, part of GE Healthcare), The RNAi Consortium (TRC) collection [13, 25] (available from Dharmacon, part of GE Healthcare, Sigma-Aldrich) and TransOMIC collections. Large libraries that are already in the pooled shRNA format are available from several vendors including GeneNet libraries from System Biosciences, Decipher libraries from Cellecta [26] and Dharmacon Decode Pooled Libraries. From these large collections, smaller libraries have also been generated to target specific gene families or other genes of interest.

The characteristics of a pooled shRNA library directly impacts the quality of data obtained for a pooled screen. Important characteristics include: (1) the robustness of knockdown, (2) composition of the library including gene coverage and the number of shRNAs per gene, (3) uniformity of the pooled library.

Robust Knockdown

The efficiency of individual shRNAs to suppress gene expression is central to a robust pooled shRNA screen. Over time various algorithms have been developed to enable the selection of highly functional targeting sequences. Although these algorithms are quite successful for siRNA design, achieving on average more than 80% knockdown, this is in general not the case for shRNA vectors. Also, additional criteria have to be implemented for microRNA embedded shRNAs, ensuring proper processing by the endogenous microRNA machinery, including

processing by Drosha and Dicer [22]. Due to the inability to select highly active shRNAs, in general libraries contain multiple, unique shRNAs targeting the same transcript. The number of shRNAs per gene can range from three to 25. The presence of multiple independent shRNAs per gene allows for the calculation of a combined score for each gene that can determine a value for hit selection. The presence of multiple shRNAs per gene also facilitates the discrimination between on- and off-target hits in large scale screens (see Hit Validation).

Library Composition

Depending on the scope of the screen, libraries can be designed to target all genes in the organism of interest (whole genome) or all the genes in a specific biological area of interest (*e.g.* protein kinases, DNA damage pathway or apoptosis collections). Because it is essential to have multiple shRNAs per gene, the size of a collection can become quite significant. If one would include five shRNAs per gene in a human whole genome collection, this would effectively mean a collection of more than 100,000 shRNAs. The use of such large collections in pooled screens is a logistical challenge requiring large numbers of cells and amounts of assay reagents (discussed in Assay Development and Screen Optimization). As such, screens are frequently designed around a limited number of genes and collections between 1,000 and 5,000 shRNAs.

Pool Uniformity

Critical to the success of a pooled shRNA experiment is the requirement that each shRNA is represented at approximately equal levels in the pool. This ensures that each individual shRNA is interrogated in the screen and reaches a sufficient threshold to allow identification of targets, especially in the case of depletion compared to the control population. Several strategies to minimize differences in abundance have been used to create libraries, including growth of individual bacterial cultures or the plating of individual colonies on agar plates. Ideally, one could isolate DNA from individual cultures and add equal amounts to the plasmid pools for the libraries; however, this requires a sophisticated infrastructure, especially when creating large collections. Generally, pooling of individual cultures or scraping and mixing of bacterial colonies is used for plasmid DNA isolation. The relative representation of each individual shRNA can be determined by deep sequencing of the DNA pool used to make the virus [27]. This provides the researcher with a solid starting point to determine subsequent enrichment or depletion under assay conditions.

ASSAY DEVELOPMENT AND SCREEN OPTIMIZATION

Prior to running a screen, it is crucial to optimize all assay and screening parameters. These optimization steps increase reproducibility between samples and hit calling capabilities. Experimental variability can arise in practically all stages of transduction and screening, including packaging of viral particles; transduction medium, additives and duration; cell density at transduction; functional viral titer in the cell line of interest; selection of transduced cells; multiplicity of infection (MOI); average shRNA fold representation during transduction, cell passaging and PCR; number of biological replicates; and application of selective pressure and phenotypic selection. These variables must be considered and addressed in order for a pooled shRNA screen to be run robustly.

Lenti- or retroviral particles are the most commonly used transduction vehicles for pooled shRNA screening because of their ability to transduce cell types that are refractory to transient transfection or other RNAi delivery methods, and their prolonged expression following genomic integration. In addition, viral vectors allow for the introduction of a single shRNA expressing vector per cell, which can be important for phenotypic selection in a screen. Packaging of virus from plasmid pools should be performed using a large number of cells to ensure adequate representation of shRNAs in the pool and to maintain consistent viral stocks. Once the virus is packaged, ideal transduction conditions and efficiency will vary for every cell line and must be determined empirically.

Well characterized positive and negative control shRNAs should be used to fully determine transduction parameters prior to beginning a pooled shRNA screen. Positive control shRNAs with validated knockdown efficiency against a biologically relevant target, can be used to assess the level of knockdown in a given set of conditions and the strength of the selective phenotype under screening conditions. Negative control shRNAs, both technical (*e.g.* a gene target that induces cell death) and biological (on target knockdown but no phenotype in the screen) can be used to assess the effects of transduction conditions on cell health and viability and the variation or noise under phenotypic selection. In addition to these shRNA controls, cell passage number and reagent lot numbers should be closely monitored to ensure consistency between assay optimization experiments and the pooled shRNA screen itself.

Preliminary transduction optimization steps include establishment of standard media additives such as hexadimethrine bromide (commonly known as

Polybrene) and serum. Polybrene is a cationic polymer that is thought to aggregate viral particles and neutralize cell surface charge, facilitating viral access to cell surface receptors [28]. Polybrene is generally used at $1 - 10$ μg/mL and should be tested in every cell line of interest to assess affects on transduction efficiency and viability. Ideally, transductions should be performed in serum-free medium; however, this is not possible in all cell lines, in which case the minimal allowable serum concentration should be assessed to maximize transduction efficiency.

Once medium conditions have been established, cell density and duration of transduction can be determined. Optimal cell density for transduction can be assessed by plating cells at a range of densities and transducing at a single MOI. Since cell number is a critical parameter in MOI calculations, it is important to know the doubling time of the cell line of interest, and therefore the number of cells at the time of transduction. Duration of viral transduction can also vary dramatically between cells lines. Times ranging from 4 to 24 hours are appropriate for viral transduction and should be determined for every cell line.

Functional titer, or infectious titer, is a measure of the transducing units per milliliter of virus in a cell line of interest. This metric must be determined empirically for every cell line and culture condition. Functional titer cannot be directly converted from titers that have been determined by other methods, such as p24 titering assays or colony forming unit assays in a different cell line. To determine the functional titer for a specific cell line, the viral pool should be titered in the cell line of interest using a colony formation assay or a FACS assay. Alternatively, a relative functional titer can be determined using a negative control shRNA virus that has been titered in parallel with the viral pool. This negative control is titered using the exact conditions determined in the optimization experiments, resulting in a conversion factor that can be calculated from the original titer of the negative control and the titer under screening conditions. This conversion factor, or relative functional titer, can then be used to determine the functional titer of the viral pool. For researchers who make their own virus, the former method is recommended, while the later technique is useful for researchers who purchase pre-packaged virus or who plan to use the virus under several different circumstances.

Most viral backbones contain a mammalian antibiotic selection cassette and a fluorescent reporter such as GFP or RFP. Any of these markers can be used to select a population of transduced cells. If using an antibiotic selection cassette, an antibiotic kill curve should be established during assay optimization to determine

the concentration and duration of antibiotic selection. In general, the antibiotic treatment is established as the lowest concentration that kills 100% of untransduced cells in three to six days. If using a fluorescent reporter to analyze transduced cells by FACS, fluorescent gates and time after transduction need to be optimized using cells transduced with control reporter constructs transduced at the same MOI established for the screen.

Selecting the appropriate MOI for a screen is a crucial step in the optimization process. To identify shRNAs that are depleted from the population, it is important to have no more than one shRNA integration per cell. However, for screens based on the positive selection this is less prohibitive. The proportion of cells that are infected at any given MOI follows a Poisson Distribution where low MOIs decrease the likelihood that any given cell is infected with more than one shRNA virus. The highest recommend MOI is 0.3, where more than 70% of infected cells are likely to have only one shRNA integrant and less than 10% of the infected cells are likely to have more than one shRNA integrant. MOIs greater than 0.3 increase the likelihood that any cell will have more than one shRNA integrant. In practice, MOIs of 0.3 to 5.0 are commonly used and the choice of MOI depends on the specific assay being used, toxicity of shRNA hairpins and available cell number.

A critical and necessary consideration of pooled lentiviral shRNA screening is the extent to which any given shRNA construct in a pooled library will be represented in the screen (*i.e.*, the number of cells that contain an independent genomic integration of any given shRNA or the number of biological replicates of each shRNA integration event). High shRNA representation results in high reproducibility between biological replicates and ensures that there is a sufficient window for detection of changes in shRNA representation after phenotypic selection [27]. shRNA fold representation between 500 and 1,000 is recommended and increasing this number can further enhance the reproducibility of a screen [29]. These recommendations are particularly important if you are interested in observing shRNA depletion hits or more subtle hits with less fold-change between reference and experimental samples.

Once a shRNA fold representation has been established, it should be maintained throughout all steps of the pooled shRNA screen, including cell passages, PCR and NGS. For example, if the shRNA pool contains 1,000 constructs and the shRNA fold representation is 1,000, then $1x10^6$ transduced cells should be obtained, this means the transduction of at least $3x10^6$ cells at an MOI of 0.3. The number of $1x10^6$ cells should be maintained for each biological replicate, it should

be maintained at each passage and 1×10^6 genomes worth of genomic DNA ($\sim$ 6 μg gDNA) should be used to recover the shRNA inserts by PCR. If shRNA representation is not maintained through all steps of the screen, biological reproducibility decreases and hits with smaller changes between reference and experimental samples are lost.

Each pooled shRNA screen requires independent biological replicates to ensure detection of biological variation as result of treatment or selection. Multiple biological replicates are required to perform rigorous statistical analysis of the results and hit selection. In general three biological replicates are used but if resources are limited, higher shRNA fold representation is more economical than running additional biological replicates; however, at least two biological replicates should be performed in all screens. In addition to replicates, multiple samples obtained at different time points can be used to establish a temporal pattern in the changes of relative abundance of individual shRNAs.

Pooled shRNA screens depend on selective enrichment or depletion of cells based on a phenotype induced after gene knockdown. The types of selective pressure and phenotypic selection are screen-specific and techniques vary widely. A classic approach involves selection of cells based on their rate of proliferation or cell survival, assayed in the absence of treatment (straight lethality screens) [30-32] or under selective pressure of a specific drug (resistance or enhancer screens) [33-38] However, cell selection can also report on other phenotypes such as expression of cell surface proteins [39-41], intracellular fluorescent reporters [42, 43], migration or adhesion [44]. If a screening approach allows for generation of a reference and selected population, one can identify shRNAs that are specifically enriched or depleted in selected samples compared to a reference sample.

To determine the relative abundance of each individual shRNA in cell populations obtained under assay conditions (including reference and experimental), genomic DNA (gDNA) is extracted and hairpins are recovered by PCR amplification. PCR primers, PCR conditions, amount of gDNA and the number of PCR cycles should be optimized to ensure that reactions remain in the linear phase of log amplification. Through the addition of adaptor sequences and index tags to permit multiplexing per lane, the samples can be analyzed by NGS. As mentioned above (assay optimization), it is important to maintain shRNA fold representation during PCR amplification. In general this means that several independent PCR reactions must be performed for each sample and then combined for sequencing. The number of PCR reactions can be reduced by enrichment of shRNA inserts by

restriction digests and fragment isolation of gDNA or by the use of DNA capture technologies.

NGS AND HIT IDENTIFICATION

Measuring the relative abundance of each shRNA within different cell populations is fundamental to the process of pooled screening. For every population, the abundance of each shRNA should be determined in a quantitative manner. Originally, DNA microarrays containing complementary oligonucleotides to shRNA sequences or associated barcodes were used for abundance detection. Although initially these approaches were sufficient for detection, increasing shRNA library size and more advanced design rules for effective shRNAs (which reduced sequence differences between unique shRNAs) resulted in reduction in specificity and high background hybridization. This higher background hybridization made it challenging to identify a decrease in abundance. Recently, NGS techniques were used to detect abundance of each individual shRNA in each sample. NGS allows for detection of each shRNA in complex libraries even when shRNA sequences are similar. Direct counting of sequence reads has the advantage of dramatically increasing the dynamic range thereby widening the screening window. In addition, multiple screening samples can be combined in one sequencing lane, eliminating artifacts due to different microarray hybridization experiments.

The analysis of individual shRNA abundance is based on sampling from the entire population with the goal of estimating the real frequency of each shRNA in that population. This sampling results in errors, known as Poisson sampling errors and caused by insufficient representation of each shRNA in the different populations. To reduce sampling error effects, each shRNA should be represented multiple times, usually at least 1,000 times. However, as pooled shRNA libraries are not normalized with respect to the abundance of each shRNA, a 1,000 fold representation does not avoid sampling error on low frequency shRNAs in the population. The screening system can also strongly influence noise and the correlation between biological replicates. For example, a low frequency of background colonies present in a screen designed to identify genes that upon knockdown cause resistance against drug-induced apoptosis, will produce a small number of colonies with different shRNAs. However, these shRNAs will not be shared among replicates because the number of these colonies is too small to represent sufficient sampling from the entire library. To be able to resolve real hits from background, it is essential to compare biological replicates. The different

types of sampling errors must be taken into account for the analysis of pooled shRNA screening results.

Data generated by NGS of pooled shRNA libraries resembles count data, similar to the output from RNA sequencing experiments. Laboratories may set up their own NGS analysis pipelines, and several methods have been developed including DESeq [45] and EdgeR [46] to generate hit lists. A data normalization step is recommended to correct for differences in read numbers between the different samples. Following this step, every construct in each population can be compared as a relative ratio.

As a general concept, these methods estimate variation based on large numbers of different shRNAs with similar count frequencies, and result in a list of shRNAs that are either enriched or depleted. Cut-offs can then be based on significance and false discovery rates in combination with a fold change threshold. However, it is difficult, due to the nature of shRNA reagents, to translate this directly into a gene hit list without applying additional criteria for selection. All RNAi-based reagents have two characteristics that complicate the performance and analysis of large scale screens; (1) the varying degree of knockdown and (2) the existence of off-target effects [47]. At this moment both the efficiency of knockdown and the degree of off-target effects for individual RNAi reagents are largely unpredictable. It has been suggested that off-target effects due to seed complementarity are less abundant, though not eliminated, with shRNA reagents compared to siRNA. To address the challenges of limited efficiency and off-target effects, most shRNA libraries contain multiple, different shRNAs targeting the same transcript. Hits can be prioritized based on the number of shRNAs per target scoring as a hit in the assay. Both concerns of silencing efficiency and off-target effects are diminished for hits that confirm with multiple, independent reagents. If sufficiently large numbers of different shRNAs per gene are included in the library, a score or metric for each shRNA can be calculated and all shRNAs targeting the same gene can subsequently be tested for a significant change in distribution compared to non-targeting controls [29]. Alternatively, a criterion of the presence of two individual shRNAs targeting the same gene among the hit list is also used, especially for libraries in which five or more shRNAs are present per gene. Another approach based on the same concept is a selection based on the second best hairpin. This method is based on the assumption that when at least two shRNAs for one gene are significantly enriched, the score for both shRNAs should also be significant. If only one shRNA scores statistically as a hit, the gene may still warrant further investigation if there is a strong biological rationale. It is

clear that a pooled shRNA screen will yield a list of potential shRNAs and their corresponding genes. Because of the nature of the RNAi reagents, including variable knockdown efficiency and off-target effects, validation of these hits is still required and strategies are outlined below.

HIT VALIDATION

All high throughput screening approaches require a validation step, whereby the statistically defined primary screen target list is reduced to a more manageable high confidence gene list. Validation strategies can take multiple paths and depend on: 1) the biological question being asked, *e.g.* is there only one assay for your biological question or are there additional assays that can reliably report on your phenotype; 2) the resources available, *e.g.* can you screen in additional cell lines, or can you afford orthogonal assays; and 3) the RNAi reagents available, *e.g.* do you have access to additional shRNA or siRNA sequences. The goal is to perform experiments that progressively increase your confidence that the target and the phenotype are related.

Validation of individual shRNAs can be quite laborious due to the requirement to prepare DNA for each construct, make and titer virus, perform appropriate assays and evaluate knockdown. Therefore, validation strategies can be highly influenced by the magnitude of the hit list. Small numbers of hits (less than 20 targets) from a primary shRNA screen are not usually observed (unless using an especially stringent assay), although it is possible that several targets will exhibit a higher magnitude of enrichment or depletion after the primary screen. The screener is often faced with a fairly large list that can be triaged on the number of shRNA sequences scoring for the same target gene as described above, and also on bioinformatic pathway analysis implicating target genes involved in the same signaling pathways. As mentioned, gene targets where multiple shRNA constructs have high scores can be ranked as high priority hits, however where only a single shRNA construct is significantly represented in the experiment, additional constructs must be experimentally tested in order to validate the target as a true hit.

For small numbers of primary hits, individual shRNA constructs can easily be obtained and processed in low throughput to verify phenotype and knockdown. In addition, new shRNA constructs can be designed and tested to confirm a gene target. For a larger number of hits, where the screener has no immediate sense of which targets to pursue, one option is to create a sub-pool library of the shRNAs targeting the hit genes and additional constructs, if required, with the aim of

identifying three hairpins with the same phenotype. In this case, the sub-pool library must be created with a similar number of non-enriched targets. This approach is relatively high throughput and affords the ability to review a reasonable number of high ranking targets. The sub-pool library is created and assayed as per the primary screen to identify the relative representation of the hairpins (still using NGS) within the context of a smaller pool. Those hairpins that might have influenced the phenotype, but may not have scored significantly due to their effect being statistically diluted by the vast numbers of other constructs, may be more identifiable in a smaller pool. For a depletion screen, a smaller pool translates to the possibility of working with a higher representation at the outset of the screen, requiring less cells (relative to that required when working with a large pool) and the opportunity for more closely timed assay points to more definitively identify targets regulating death. Again, the endpoint of this step is to refine the target list down to a workable number of genes.

Validation at the level of knockdown and confirmation that gene knockdown correlates with phenotype is critical. Once a gene target has been verified by the extensive phenotypic methods identified above, the level of gene knockdown must be confirmed. Reagent suppliers guarantee that target gene knockdown should be greater than 75%, confirmed at the mRNA level after short term knockdown or at the protein level after longer knockdown. A non-silencing negative control shRNA construct can be used as a reference for comparing the extent of knockdown. While knockdown alone cannot completely verify that the intended target gene is causing the phenotypic outcome, it remains a very important step in ensuring that there is on-target activity.

The gold standard for confirming that a particular siRNA or shRNA sequence is responsible for a phenotype is to develop a phenotype rescue assay [48]. Usually, a silent mutation is introduced into the target or a target ortholog is used that cannot be targeted by the RNAi reagent, but preserves the phenotype. The constructs are transfected into cells and assayed under screening conditions. Complete reversal of the phenotypic effect should be observed by specific gene knockdown. While phenotypic rescue is perhaps the ultimate means of verification, it is very low throughput and can usually only be performed for a few targets. The success of this approach however, can also be limited depending on the extent of the effect of over-expression of the target.

A common alternative validation strategy is the use of chemical compounds that have target activity against your gene of interest. Standard cell culture and assay conditions can be utilized in this approach, although drug sensitivity will need to

be calculated over a dose curve. Not all drugs are completely specific for one protein, but this offers a very attractive means of verifying the functional effect of the gene of interest, particularly when the project has a translational component. Confirmation of targeting by western analysis is required to verify the phenotypic outcome. This is also an attractive strategy in the scenario where only a single shRNA has been identified as a hit.

As with other experimental systems, verification that the biological effect can be repeated in additional cell lines and with additional experimental approaches is critical to confirming that the biological effect you observed during the screen is robust and reproducible. When using different cell lines, it is always prudent to confirm gene knockdown

Following the validation process, the screener will have arrived at a relatively small number of targets with which to continue assessing in a biological context. This will include standard approaches such as alternative assays, complementing the shRNA strategies with shorter term siRNA approaches, identification of synthetic compounds that can phenocopy the knockdown result and *in vivo* strategies. The latter can include using inducible shRNA constructs (particularly for genes causing cell death) in orthotopic mouse models, or *C.elegans* and *Drosophila* knockout models.

FUTURE DIRECTIONS

Over the past ten years pooled shRNA screens have become recognized as a powerful method for functional genomic screening in mammalian cells. During the same period, screening models have evolved from relatively straightforward phenotypes such as proliferation or survival to more complex phenotypes including adhesion, migration and even gene expression. In addition, pooled shRNA screens are no longer restricted to *in vitro* models, but are also used for *in vivo* screening (described below). Although the number of novel targets for therapy identified by large scale shRNA screening is limited, it has proven powerful in the identification of biomarkers or potential combination therapies.

Although genome wide screens were appealing at first, it has become clear that such approaches require a tremendous effort in validation and follow-up of the primary screen, as discussed above. Also, the interpretation of results from primary shRNA screens is not straightforward due to the inherent nature of RNAi technology, including off-targets effects and low knockdown efficiency. As an alternative, screeners are moving toward smaller, pre-selected gene sets based on gene-families

(protein kinases, phosphatases, metabolic enzymes, chromatin modifiers, *etc.*) or other types of information such as expression, mutation or pathways analysis. Another strategy is to use larger numbers of different shRNAs per gene to allow for the identification of multiple independent shRNAs per gene as hits. This type of approach can yield higher confidence in the genes selected as hits, potentially allowing integrated analysis of these genes with other data types such as gene set enrichment, protein-protein interaction or pathway analysis. Indeed, technical improvements in oligonucleotide synthesis on glass slides has greatly facilitated the custom generation of libraries with up to 50 individual shRNAs per gene [43]. Although this increase in number of shRNAs per gene has advantages in hit selection, a potential complicating factor is the representation of all individual shRNAs in a screen. This can become a major hurdle when more complex phenotypes are used as a selection method in a screen. A possible solution to this problem is the use of shRNA collections displaying more efficient or validated knockdown for the genes targeted and thereby reducing the necessity for large numbers of shRNAs in a library. One strategy to generate validated collections is the use of a target site reporter linked to a fluorescent marker present in the same construct driving the expression of the target specific shRNA [49]. While the short target site cannot fully predict shRNA activity on native mRNA, selecting those cells that have reduced expression of the reporter can enrich for active shRNAs in the population. Subsequently these cell collections can be used for screening or the generation of improved algorithms for better prediction of active shRNAs. Together these different technologies will undoubtedly result in better technology platforms for pooled shRNA screening and thereby increase the possible screening models to be used. In particular, *in vivo* screening will become more feasible.

Large scale RNAi screening has moved screens in mammalian cells towards the characteristics of genetic screens in model organisms. However, major differences still exist. In general, RNAi screens score phenotypes based on partial depletion of protein expression. Although at first this can complicate the interpretation of the effects of knockdown it also creates the opportunity to observe phenotypes based on partial or incomplete knockdown. In the case of lethal genes, this can be sufficient to rescue the lethal phenotype and at the same time cause a biological phenotype (*e.g.* resistance to drug treatment). In addition, one has to deal with potential off- target effects obscuring a biological phenotype. In model organisms gene deletions can be introduced, *e.g.* the yeast deletion collection [50] generation of null-alleles for individual genes without affecting other transcripts. As mentioned above, this cannot be used for essential genes and either hypo-morph or temperature-sensitive alleles are needed to study these essential genes.

Recently, new technologies have been developed enabling gene-editing in mammalian cells, which could potentially be used for large scale screening. The Brummelkamp group has developed a screening system based on haploid cells [51]. These cells carry only one copy of each chromosome and can be used in combination with gene trap technologies to randomly inactivate genes. Analogous to pooled shRNA screens, millions of independent integrations can be generated, selected for a biological phenotype of interest and recovered by NGS [51]. A current limitation for haploid screens in mammalian cells is the limited availability of haploid cell line types and the inability to pre-select cell lines with characteristics associated with the biological question, *e.g.* sensitive and resistant cell line pairs. A potential solution is the use of novel gene-editing technologies such as Zinc-finger-nucleases (ZFNs) [52], transcription activator-like effector nucleases (TALENs) [53] and RNA-guided endonucleases (RGENs) (reviewed in [54]). These technologies use either proteins or protein-RNA complexes to sequence specifically introduce double-stranded DNA breaks that have been harnessed to engineer mammalian genomes. Although at this moment, the application of these technologies in large scale screening is still challenging, recently there has been considerable progress with the use of the CRISPR (clustered regularly interspaced short palindromic repeats)-associated nuclease CAS9 to modify specific genomic loci on a genome wide scale. The specificity of the CAS9 nuclease is determined by short guide RNA sequences. Consequently, large scale libraries can be constructed using array-based oligonucleotide synthesis followed by cloning into lentiviral vectors which is suitable for pooled screening. Indeed, this approach has been applied successfully in positive "resistance" [55, 56] and negative "lethality" screens in mammalian cells, including human embryonic stem cells [57]. A significant difference between CRISPR/CAS9 based screens *versus* shRNA screens is the complete loss *versus* partial knock-down of the expression of the targeted genes. Although this would certainly aid in the identification of straight lethal genes [58], it could also hinder the identification of genes that only display a phenotype at reduced levels rather than complete absence. Finally, the problem of off-target effects is not eliminated with these new gene-editing technologies [59]. Partial similarity of genomic sequences to the guide RNA can result in mutations or deletions in other genes than the intended target. As consequence, also for CRISPR-CAS9 screening systems multiple independent sequences should be used to confirm the on-target phenotype.

Despite these recent advances in genome editing technologies, the application of large scale pooled shRNA screening is still expanding. There is a clear trend of

moving from screens based on differential viability to more advanced screening models coupled with sophisticated read-outs. The latter include flow cytometry-based high content screening technologies where cell surface or intracellular marker expression can be used to select cells or even selection based on quantitative analysis of translocation of cytoplasm to the nucleus. In addition, screening models are moving from 2D culture systems to 3D culture systems and *in vivo* screening. One of the first examples of successful large scale *in vivo* shRNA screens were based on models of leukemia [60]. In these models the grafting efficiency is sufficiently high to screen large collections of shRNA vectors. It is possible to re-generate full organ systems, *e.g.* the liver, under conditions compatible with shRNA screening. Upon ablation of most of the liver with a drug, it can be reconstituted with a drug resistant population of progenitor cells carrying large collections of different shRNA vectors [61]. Another approach is the use of orthotopic transplantation in which an shRNA library containing cell population is transplanted into a specific organ. An elegant example of such approach is a screen for key regulators of neural and malignant glioma stem cells [62].

Besides the development of more advanced screening models, the relative ease at which large scale pooled shRNA screens can be performed allows for the generation of large numbers of screens in many different (tumor) cell lines under different conditions. This allows for the generation of a compendium of genes required for the survival of tumor cells with specific genetic alterations, *e.g.* RAS mutations, PTEN loss or receptor over-expression. The results of such screening efforts can be integrated with genome scale analyses for genomic alterations, gene expression and protein expression and modification thus generating a platform to discover specific dependencies and novel targets for treatment. An interesting strategy is the generation of interaction maps using shRNA vectors that contain 2 different shRNAs, each targeting a specific gene. These can also be applied in a pooled format and the combined effect of knock down of both genes can be addressed [43] thereby identifying genetic dependencies, analogous to the yeast genetic interaction maps [50, 63]. Finally, the combination of pooled shRNA screening with many different compounds can provide data allowing for the clustering of functional classes of compounds and a better understanding of their mechanism of action and potential effect modifiers.

It is without doubt that pooled shRNA screening will find its further application in many areas of research. The development of more advanced screening models, better quality libraries combined with cumulative experience will enable new discoveries that will find their way into the clinic.

COMPETING INTERESTS STATEMENT

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors confirm that this chapter contents have no conflicts of interest.

REFERENCES

[1]   T. Zhou, J. Aumais, X. Liu, L. Yu-Lee, and R. Erikson, A role for Plk1 phosphorylation of NudC in cytokinesis. Dev Cell 5 (2003) 127-38.
[2]   L. Zender, S. Hutker, C. Liedtke, H.L. Tillmann, S. Zender, B. Mundt, M. Waltemathe, T. Gosling, P. Flemming, N.P. Malek, C. Trautwein, M.P. Manns, F. Kuhnel, and S. Kubicka, Caspase 8 small interfering RNA prevents acute liver failure in mice. PNAS 100 (2003) 7797-7802.
[3]   A.C. Hsieh, R. Bo, J. Manola, F. Vazquez, O. Bare, A. Khvorova, S. Scaringe, and W.R. Sellers, A library of siRNA duplexes targeting the phosphoinositide 3-kinase pathway: determinants of gene silencing for use in cell-based screens. Nucleic Acids Res 32 (2004) 893-901.
[4]   Z.Y. Jiang, Q.L. Zhou, K.A. Coleman, M. Chouinard, Q. Boese, and M.P. Czech, Insulin signaling through Akt/protein kinase B analyzed by small interfering RNA-mediated gene silencing. PNAS 100 (2003) 7569-7574.
[5]   K. Kurisaki, A. Kurisaki, U. Valcourt, A.A. Terentiev, K. Pardali, P. ten Dijke, C.-H. Heldin, J. Ericsson, and A. Moustakas, Nuclear Factor YY1 Inhibits Transforming Growth Factor {beta}- and Bone Morphogenetic Protein-Induced Cell Differentiation. Mol. Cell. Biol. 23 (2003) 4494-4510.
[6]   W.B. Struwe, and C.E. Warren, High-throughput RNAi screening for N-glycosylation dependent loci in Caenorhabditis elegans. Methods Enzymol 480 (2010) 477-93.
[7]   A.W. Whitehurst, B.O. Bodemann, J. Cardenas, D. Ferguson, L. Girard, M. Peyton, J.D. Minna, C. Michnoff, W. Hao, M.G. Roth, X.J. Xie, and M.A. White, Synthetic lethal screen identification of chemosensitizer loci in cancer cells. Nature 446 (2007) 815-9.
[8]   K.H. Kok, T. Lei, and D.Y. Jin, siRNA and shRNA screens advance key understanding of host factors required for HIV-1 replication. Retrovirology 6 (2009) 78.
[9]   Q. Li, A.L. Brass, A. Ng, Z. Hu, R.J. Xavier, T.J. Liang, and S.J. Elledge, A genome-wide genetic screen for host factors required for hepatitis C virus propagation. PNAS 106 (2009) 16410-16415.
[10]   B.E. Gewurz, F. Towfic, J.C. Mar, N.P. Shinners, K. Takasaki, B. Zhao, E.D. Cahir-McFarland, J. Quackenbush, R.J. Xavier, and E. Kieff, Genome-wide siRNA screen for mediators of NF-{kappa}B activation. PNAS 109 (2012) 2467-2472.
[11]   A. Lipovsky, A. Popa, G. Pimienta, M. Wyler, A. Bhan, L. Kuruvilla, M.-A. Guie, A.C. Poffenberger, C.D.S. Nelson, W.J. Atwood, and D. DiMaio, Genome-wide siRNA screen identifies the retromer as a cellular entry factor for human papillomavirus. PNAS 110 (2013) 7452-7457.

[12]   D. Nickles, C. Falschlehner, M. Metzig, and M. Boutros, A Genome-Wide RNA Interference Screen Identifies Caspase 4 as a Factor Required for Tumor Necrosis Factor Alpha Signaling. Mol. Cell. Biol. 32 (2012) 3372-3381.

[13]   J. Moffat, D.A. Grueneberg, X. Yang, S.Y. Kim, A.M. Kloepfer, G. Hinkle, B. Piqani, T.M. Eisenhaure, B. Luo, J.K. Grenier, A.E. Carpenter, S.Y. Foo, S.A. Stewart, B.R. Stockwell, N. Hacohen, W.C. Hahn, E.S. Lander, D.M. Sabatini, and D.E. Root, A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell 124 (2006) 1283-98.

[14]   B. Maier, S. Wendt, J.T. Vanselow, T. Wallach, S. Reischl, S. Oehmke, A. Schlosser, and A. Kramer, A large-scale functional RNAi screen reveals a role for CK2 in the mammalian circadian clock. Genes & Dev. 23 (2009) 708-718.

[15]   K. Berns, E.M. Hijmans, J. Mullenders, T.R. Brummelkamp, A. Velds, M. Heimerikx, R.M. Kerkhoven, M. Madiredjo, W. Nijkamp, B. Weigelt, R. Agami, W. Ge, G. Cavet, P.S. Linsley, R.L. Beijersbergen, and R. Bernards, A large-scale RNAi screen in human cells identifies new components of the p53 pathway. Nature 428 (2004) 431-7.

[16]   I.G. Kolfschoten, B. van Leeuwen, K. Berns, J. Mullenders, R.L. Beijersbergen, R. Bernards, P.M. Voorhoeve, and R. Agami, A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. Cell 121 (2005) 849-58.

[17]   T.F. Westbrook, E.S. Martin, M.R. Schlabach, Y. Leng, A.C. Liang, B. Feng, J.J. Zhao, T.M. Roberts, G. Mandel, G.J. Hannon, R.A. Depinho, L. Chin, and S.J. Elledge, A genetic screen for candidate tumor suppressors identifies REST. Cell 121 (2005) 837-48.

[18]   X. Huang, J.Y. Wang, and X. Lu, Systems analysis of quantitative shRNA-library screens identifies regulators of cell adhesion. BMC Syst Biol 2 (2008) 49.

[19]   G.A. Smolen, J. Zhang, M.J. Zubrowski, E.J. Edelman, B. Luo, M. Yu, L.W. Ng, C.M. Scherber, B.J. Schott, S. Ramaswamy, D. Irimia, D.E. Root, and D.A. Haber, A genome-wide RNAi screen identifies multiple RSK-dependent regulators of cell migration. Genes Dev 24 (2010) 2654-65.

[20]   K. Berns, E. Hijmans, J. Mullenders, T. Brummelkamp, A. Velds, M. Heimerikx, R. Kerkhoven, M. Madiredjo, W. Nijkamp, B. Weigelt, R. Agami, W. Ge, G. Cavet, P. Linsley, R. Beijersbergen, and R. Bernards, A large-scale RNAi screen in human cells identifies new components of the p53 pathway. Nature 428 (2004) 431-7.

[21]   R. Bernards, T.R. Brummelkamp, and R.L. Beijersbergen, shRNA libraries and their use in cancer genetics. Nat Methods 3 (2006) 701-6.

[22]   K. Chang, S.J. Elledge, and G.J. Hannon, Lessons from Nature: microRNA-based shRNA libraries. Nat Methods 3 (2006) 707-14.

[23]   J.M. Silva, M.Z. Li, K. Chang, W. Ge, M.C. Golding, R.J. Rickles, D. Siolas, G. Hu, P.J. Paddison, M.R. Schlabach, N. Sheth, J. Bradshaw, J. Burchard, A. Kulkarni, G. Cavet, R. Sachidanandam, W.R. McCombie, M.A. Cleary, S.J. Elledge, and G.J. Hannon, Second-generation shRNA libraries covering the mouse and human genomes. Nat Genet 37 (2005) 1281-8.

[24]   G. Hu, and J. Luo, A primer on using pooled shRNA libraries for functional genomic screens. Acta Biochim Biophys Sin 44 (2012) 103-112.

[25]   D.E. Root, N. Hacohen, W.C. Hahn, E.S. Lander, and D.M. Sabatini, Genome-scale loss-of-function screening with a lentiviral RNAi library. Nat Methods 3 (2006) 715-9.

[26]   Z. Zhu, and D. Huangfu, Human pluripotent stem cells: an emerging model in developmental biology. Development 140 (2013) 705-717.

[27]   Z. Strezoska, A. Licon, J. Haimes, K. Spayd, K. Patel, K. Sullivan, K. Jastrzebski, K. Simpson, D. Leake, A. van Brabant Smith, and A. Vermeulen, Optimized PCR conditions and increased shRNA fold representation improve reproducibility of pooled shRNA screens. PLoS One 7 (2012) e42341.

[28]   B.M. Davis, L. Humeau, V. Slepushkin, G. Binder, L. Korshalla, Y. Ni, E.O. Ogunjimi, L.-F. Chang, X. Lu, and B. Dropulic, ABC transporter inhibitors that are substrates enhance lentiviral vector transduction into primitive hematopoietic progenitor cells. Blood 104 (2004) 364-373.

[29]   M.C. Bassik, M. Kampmann, R.J. Lebbink, S. Wang, M.Y. Hein, I. Poser, J. Weibezahn, M.A. Horlbeck, S. Chen, M. Mann, A.A. Hyman, E.M. Leproust, M.T. McManus, and J.S. Weissman, A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. Cell 152 (2013) 909-22.

[30]   B. Luo, H.W. Cheung, A. Subramanian, T. Sharifnia, M. Okamoto, X. Yang, G. Hinkle, J.S. Boehm, R. Beroukhim, B.A. Weir, C. Mermel, D.A. Barbie, T. Awad, X. Zhou, T. Nguyen, B. Piqani, C. Li,

T.R. Golub, M. Meyerson, N. Hacohen, W.C. Hahn, E.S. Lander, D.M. Sabatini, and D.E. Root, Highly parallel identification of essential genes in cancer cells. Proc Natl Acad Sci U S A 105 (2008) 20380-5.

[31]   J.M. Silva, K. Marran, J.S. Parker, J. Silva, M. Golding, M.R. Schlabach, S.J. Elledge, G.J. Hannon, and K. Chang, Profiling essential genes in human mammary cells by multiplex RNAi screening. Science 319 (2008) 617-20.

[32]   M.R. Schlabach, J. Luo, N.L. Solimini, G. Hu, Q. Xu, M.Z. Li, Z. Zhao, A. Smogorzewska, M.E. Sowa, X.L. Ang, T.F. Westbrook, A.C. Liang, K. Chang, J.A. Hackett, J.W. Harper, G.J. Hannon, and S.J. Elledge, Cancer Proliferation Gene Discovery Through Functional Genomics. Science 319 (2008) 620-624.

[33]   R. Marcotte, K.R. Brown, F. Suarez, A. Sayad, K. Karamboulas, P.M. Krzyzanowski, F. Sircoulomb, M. Medrano, Y. Fedyshyn, J.L.Y. Koh, D. van Dyk, B. Fedyshyn, M. Luhova, G.C. Brito, F.J. Vizeacoumar, F.S. Vizeacoumar, A. Datti, D. Kasimer, A. Buzina, P. Mero, C. Misquitta, J. Normand, M. Haider, T. Ketela, J.L. Wrana, R. Rottapel, B.G. Neel, and J. Moffat, Essential Gene Profiles in Breast, Pancreatic, and Ovarian Cancer Cells. Cancer Discovery 2 (2012) 172-189.

[34]   J. Mullenders, W. von der Saal, M.M.W. van Dongen, U. Reiff, R. van Willigen, R.L. Beijersbergen, G. Tiefenthaler, C. Klein, and R. Bernards, Candidate Biomarkers of Response to an Experimental Cancer Drug Identified through a Large-scale RNA Interference Genetic Screen. Clin. Cancer Res. 15 (2009) 5811-5819.

[35]   K. Berns, H.M. Horlings, B.T. Hennessy, M. Madiredjo, E.M. Hijmans, K. Beelen, S.C. Linn, A.M. Gonzalez-Angulo, K. Stemke-Hale, M. Hauptmann, R.L. Beijersbergen, G.B. Mills, M.J. van de Vijver, and R. Bernards, A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. Cancer Cell 12 (2007) 395-402.

[36]   A. Prahallad, C. Sun, S. Huang, F. Di Nicolantonio, R. Salazar, D. Zecchin, R. Beijersbergen, A. Bardelli, and R. Bernards, Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. Nature 483 (2012) 100-3.

[37]   B. O'Connell, B. Adamson, J. Lydeard, M. Sowa, A. Ciccia, A. Bredemeyer, M. Schlabach, S. Gygi, S. Elledge, and J. Harper, A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. Mol Cell 40 (2010) 645-57.

[38]   T.R. Brummelkamp, A.W. Fabius, J. Mullenders, M. Madiredjo, A. Velds, R.M. Kerkhoven, R. Bernards, and R.L. Beijersbergen, An shRNA barcode screen provides insight into cancer cell vulnerability to MDM2 inhibitors. Nat Chem Biol 2 (2006) 202-6.

[39]   H. Tsujii, Y. Eguchi, A. Chenchik, T. Mizutani, K. Yamada, and Y. Tsujimoto, Screening of cell death genes with a mammalian genome-wide RNAi library. J. Biochem. 148 (2010) 157-170.

[40]   A.B. Mak, K.M. Blakely, R.A. Williams, P.-A. Penttila, A.I. Shukalyuk, K.T. Osman, D. Kasimer, T. Ketela, and J. Moffat, CD133 Protein N-Glycosylation Processing Contributes to Cell Surface Recognition of the Primitive Cell Marker AC133 Epitope. J. Biol. Chem. 286 (2011) 41046-41056.

[41]   C. Gazin, N. Wajapeyee, S. Gobeil, C.M. Virbasius, and M.R. Green, An elaborate pathway required for Ras-mediated epigenetic silencing. Nature 449 (2007) 1073-7.

[42]   M. Coussens, P. Davy, L. Brown, C. Foster, W.H. Andrews, M. Nagata, and R. Allsopp, RNAi screen for telomerase reverse transcriptase transcriptional regulators identifies HIF1{alpha} as critical for telomerase function in murine embryonic stem cells. PNAS 107 (2010) 13842-13847.

[43]   M. Bassik, R. Lebbink, L. Churchman, N. Ingolia, W. Patena, E. LeProust, M. Schuldiner, J. Weissman, and M. McManus, Rapid creation and quantitative monitoring of high coverage shRNA libraries. Nat Methods 6 (2009) 443-5.

[44]   G.A. Smolen, J. Zhang, M.J. Zubrowski, E.J. Edelman, B. Luo, M. Yu, L.W. Ng, C.M. Scherber, B.J. Schott, S. Ramaswamy, D. Irimia, D.E. Root, and D.A. Haber, A genome-wide RNAi screen identifies multiple RSK-dependent regulators of cell migration. Genes & Dev. 24 (2010) 2654-2665.

[45]   R.C. Gentleman, V.J. Carey, D.M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, and J. Gentry, Bioconductor: open software development for computational biology and bioinformatics. Genome biology 5 (2004) R80.

[46]   M.D. Robinson, D.J. McCarthy, and G.K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26 (2010) 139-40.

[47]   A.L. Jackson, J. Burchard, J. Schelter, B.N. Chau, M. Cleary, L. Lim, and P.S. Linsley, Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. RNA 12 (2006) 1179-1187.

[48]   C.J. Echeverri, P.A. Beachy, B. Baum, M. Boutros, F. Buchholz, S.K. Chanda, J. Downward, J. Ellenberg, A.G. Fraser, N. Hacohen, W.C. Hahn, A.L. Jackson, A. Kiger, P.S. Linsley, L. Lum, Y. Ma, B. Mathey-Prevot, D.E. Root, D.M. Sabatini, J. Taipale, N. Perrimon, and R. Bernards, Minimizing the risk of reporting false positives in large-scale RNAi screens. Nat Methods 3 (2006) 777-9.

[49]   C. Fellmann, J. Zuber, K. McJunkin, K. Chang, C. Malone, R. Dickins, Q. Xu, M. Hengartner, S. Elledge, G. Hannon, and S. Lowe, Functional identification of optimized RNAi triggers using a massively parallel sensor assay. Mol Cell 41 (2011) 733-46.

[50]   A.H.Y. Tong, M. Evangelista, A.B. Parsons, H. Xu, G.D. Bader, N. Page, M. Robinson, S. Raghibizadeh, C.W.V. Hogue, H. Bussey, B. Andrews, M. Tyers, and C. Boone, Systematic Genetic Analysis with Ordered Arrays of Yeast Deletion Mutants. Science 294 (2001) 2364-2368.

[51]   J.E. Carette, C.P. Guimaraes, M. Varadarajan, A.S. Park, I. Wuethrich, A. Godarova, M. Kotecki, B.H. Cochran, E. Spooner, H.L. Ploegh, and T.R. Brummelkamp, Haploid Genetic Screens in Human Cells Identify Host Factors Used by Pathogens. Science 326 (2009) 1231-1235.

[52]   D. Carroll, Genome Engineering With Zinc-Finger Nucleases. Genetics 188 (2011) 773-782.

[53]   A.J. Bogdanove, and D.F. Voytas, TAL Effectors: Customizable Proteins for DNA Targeting. Science 333 (2011) 1843-1846.

[54]   J.v.d. Oost, New Tool for Genome Surgery. Science 339 (2013) 768-770.

[55]   T. Wang, J.J. Wei, D.M. Sabatini, and E.S. Lander, Genetic screens in human cells using the CRISPR-Cas9 system. Science 343 (2014) 80-4.

[56]   H. Koike-Yusa, Y. Li, E.P. Tan, C. Velasco-Herrera Mdel, and K. Yusa, Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat Biotechnol 32 (2014) 267-73.

[57]   O. Shalem, N.E. Sanjana, E. Hartenian, X. Shi, D.A. Scott, T.S. Mikkelsen, D. Heckl, B.L. Ebert, D.E. Root, J.G. Doench, and F. Zhang, Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343 (2014) 84-7.

[58]   T. Hart, K.R. Brown, F. Sircoulomb, R. Rottapel, and J. Moffat, Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. Mol Syst Biol 10 (2014) 733.

[59]   Y. Fu, J.A. Foden, C. Khayter, M.L. Maeder, D. Reyon, J.K. Joung, and J.D. Sander, High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31 (2013) 822-6.

[60]   A. Bric, C. Miething, C.U. Bialucha, C. Scuoppo, L. Zender, A. Krasnitz, Z. Xuan, J. Zuber, M. Wigler, J. Hicks, R.W. McCombie, M.T. Hemann, G.J. Hannon, S. Powers, and S.W. Lowe, Functional identification of tumor-suppressor genes through an *in vivo* RNA interference screen in a mouse lymphoma model. Cancer Cell 16 (2009) 324-35.

[61]   T. Wuestefeld, M. Pesic, R. Rudalska, D. Dauch, T. Longerich, T.W. Kang, T. Yevsa, F. Heinzmann, L. Hoenicke, A. Hohmeyer, A. Potapova, I. Rittelmeier, M. Jarek, R. Geffers, M. Scharfe, F. Klawonn, P. Schirmacher, N.P. Malek, M. Ott, A. Nordheim, A. Vogel, M.P. Manns, and L. Zender, A Direct *in vivo* RNAi screen identifies MKK4 as a key regulator of liver regeneration. Cell 153 (2013) 389-401.

[62]   G. Gargiulo, M. Cesaroni, M. Serresi, N. de Vries, D. Hulsman, S.W. Bruggeman, C. Lancini, and M. van Lohuizen, *In Vivo* RNAi Screen for BMI1 Targets Identifies TGF-beta/BMP-ER Stress Pathways as Key Regulators of Neural- and Malignant Glioma-Stem Cell Homeostasis. Cancer Cell 23 (2013) 660-76.

[63]   H. Yu, P. Braun, M.A. Yildirim, I. Lemmens, K. Venkatesan, J. Sahalie, T. Hirozane-Kishikawa, F. Gebreab, N. Li, N. Simonis, T. Hao, J.-F. Rual, A. Dricot, A. Vazquez, R.R. Murray, C. Simon, L. Tardivo, S. Tam, N. Svrzikapa, C. Fan, A.-S. de Smet, A. Motyl, M.E. Hudson, J. Park, X. Xin, M.E. Cusick, T. Moore, C. Boone, M. Snyder, F.P. Roth, A.-L. Barabasi, J. Tavernier, D.E. Hill, and M. Vidal, High-Quality Binary Protein Interaction Map of the Yeast Interactome Network. Science 322 (2008) 104-110.

RNAi for Viral Disease Control

Cameron R. Stewart[1],*, S. Mark Tompkins[2], Kristie A. Jenkins[1], Leonard H. Izzard[3], John Stambas[3], Andrew G. Bean[1], Mark L. Tizard[1], Timothy J. Doran[1] and John W. Lowenthal[1]

[1]*CSIRO Australian Animal Health Laboratory, Geelong 3220, Victoria, Australia;* [2]*Department of Infectious Diseases, University of Georgia, Athens, Georgia, USA and* [3]*Deakin University School of Medicine, Waurn Ponds, 3216, Victoria, Australia*

Abstract: Zoonotic viruses emerging from wildlife and domesticated animals pose a serious threat to human and animal health and are recognised as the most likely source of the next pandemic. Containment of emerging infectious disease (EID) outbreaks is often difficult due to their unpredictability and the absence of effective control measures, such as vaccines, therapies and diagnostics. RNA interference (RNAi) provides a novel and effective therapeutic strategy to combat infectious diseases through modulation of pathogen and/or host gene expression. In this chapter we discuss the applications of RNAi to combat EIDs. We discuss how RNAi has furthered understanding of virus lifecycles by making possible genome-wide functional genomics studies to discover host functions that are essential for virus replication, and in the process, identify new targets for antiviral therapies. We also discuss the advantages and hurdles associated with the use of RNAi as antiviral therapeutics, in addition to the engineering of disease-resistant livestock using RNAi to protect both humans and animals from EIDs.

Keywords: Functional genomics, host-pathogen interactions, RNAi, RNAi delivery.

THE IMPACT OF RNAi DISCOVERY ON HOST-VIRUS STUDIES

RNA Interference

Fire and Mello received the Nobel Prize in 2006 for their discovery of RNA interference (RNAi) in the worm, *C. elegans* [1]. They found that injection of double-stranded RNA into worms was much more effective in interfering with gene expression than injection of single-stranded antisense RNA. With the discovery that the RNAi process is triggered by dsRNA, the intricacies of the

---

*Corresponding author Cameron R. Stewart: CSIRO Australian Animal Health Laboratory, Geelong 3220, Victoria, Australia; E-mail: Cameron.Stewart@csiro.au

RNAi pathway were quickly unravelled. Work from numerous groups subsequently confirmed that RNAi is a natural, sequence-specific post-transcriptional gene silencing pathway, and is an ancient phenomenon that is shared across kingdoms from fungi, to plants, insects and animals [2-4].

At the time of the original discovery, it is unlikely that the researchers would have appreciated the scale and breadth that the application of RNAi technology would make in little over a decade. The recent applications of RNAi in plants and animals have revolutionized our understanding of gene regulation and have proven an invaluable research tool to understand the function of specific genes in numerous organisms. Furthermore, it provides a novel and effective therapeutic strategy to combat infectious diseases through modulation of pathogen and/or host gene expression.

One Health Approach to Fighting Zoonotic Viruses

Over 70% of emerging infectious diseases (EIDs) in humans are zoonotic, that is they originate from animals. Despite scientific advancements over the last three decades, EIDs continue to inflict substantial social and economic costs. Furthermore, they will continue to pose serious threats into the future due to the effects of climate change and increased volumes of trade and human travel. The term One-Health refers to the combined use of human and animal health disciplines, and it is widely agreed that this offers us the best chance to reduce the global impact of EIDs on people, animals and the environment. Perhaps the most impactful examples include H5N1 (avian) and H1N1 (swine) influenza. Since 2003, persistent outbreaks of highly pathogenic H5N1 avian influenza viruses have decimated poultry production in some regions of the world and show 60% mortality rates when transmitted from chickens to humans [5]. In 2009, the H1N1 pandemic virus spread from pigs to humans in over 200 countries and quickly became the dominant circulating strain in the human population [6]. These re-occurring outbreaks highlight the persistent and devastating nature of influenza infections and increase the risk of new pandemics. The prospect of highly pathogenic strains of influenza virus becoming transmissible between humans has recently raised concerns [7, 8]. In addition to avian influenza virus, we have recently witnessed the emergence of pathogenic zoonotic viruses from previously unappreciated reservoir hosts. Bats harbour a large range of viruses, a subset of which are very highly pathogenic in humans; including rabies virus [9], Hendra virus [10], Nipah virus [11], Ebola virus [12] and SARS-like coronaviruses [13].

Novel Strategies for Anti-Viral Drug Discovery

Current control strategies for influenza are largely ineffective. Protection *via* vaccination is complicated by the ability of the virus to rapidly mutate and re-assort and is largely non-existent or ineffective against the highly-pathogenic forms. Furthermore, viral resistance to current antiviral therapeutics is increasing (reviewed in [14], clearly signaling that new strategies to defeat EIDs such as influenza are urgently required. An important first step will be to increase our understanding of host-virus interactions at a cellular and molecular level. RNAi technology provides an opportunity to greatly expand our knowledge in this area.

RNAi SCREENS OF VIRUS-HOST INTERACTIONS

Introduction

Viruses are among the most important causal agents of animal and human disease. For example, influenza virus infection is currently the principal source of combined morbidity and mortality in the world [15] - each year affecting up to 15% of the human population, causing acute illness in millions of people, and resulting in almost 500,000 deaths globally [16].

The rapid evolution of many viruses, coupled with the consistent emergence of new ones, highlight a demand for increased understanding of virus biology and the need for new strategies for antiviral therapies. All viruses lack the full complement of proteins required for the production of infectious virus – for instance the influenza A genome consists of only 8 segments encoding 10-12 proteins. The limited genomes of viruses, particularly RNA viruses, result in elements of host cells being 'high-jacked' and used to facilitate the viral life cycle. Some of the more complex viruses also express proteins that help to evade the host's anti-viral processes. Understanding the host contribution to viral replication and immune evasion is essential for discovering new therapeutic strategies.

The determination of the human genome sequence, coupled with tremendous gains in understanding of RNAi design, has made possible genome-wide functional genomic screens. This technology is an unbiased means of discovery and has led to the identification of previously unappreciated or unknown cellular pathways involved in health and disease. In the context of virus life cycle, we can now identify hundreds of host genes required for all stages of the virus replication cycle – from cellular virus entry and genome replication to assembly/budding. Data can be considered in accompaniment to analogous technologies such as 2-hybrid screens, transcriptomics and compound library screens – collectively

providing a comprehensive view of the host-virus interactome. In addition to greatly enhancing our understanding of host-virus interactions, RNAi screens identify candidates for new antiviral therapeutics due to the druggable nature of many virus-essential host genes [17].

In this section we will briefly discuss genome-wide RNAi screens of host-virus interactions, and profile a select few host genes whose role in host virus interactions has been made possible by this technology.

Virus-Host Screens

The use of RNAi screens to identify host factors associated with influenza virus replication has been reviewed recently by our group and by others [18, 19]. We will therefore only cover this topic briefly in this chapter. In total, five RNAi screens have been performed to detect host contributions to influenza virus replication. Some screens focused on early replication events only by employing viruses incapable of replication [20, 21] while others included both early and late events [22-24]. Together, the five screens generated a list of potentially relevant genes that represented approximately 2% of the screened genes. Meta-analysis of candidate genes across the 5 screens demonstrates a poor degree of candidate gene overlap, however, this feature might be expected given the different methods, reagents, cell types (immortalized human lung epithelial cells used in [17, 25], primary human brochial epithelial cells used in [24], Drosophila D-Mel2 cells used in [20], and a human osteosarcoma **cell** line used in [22]) and viruses employed (A/WSN (H1N1) used in [17, 25], A/PR8 (H1N1) used in [22, 24] and a VSV-influenza pseudovirus constructed to facilitate influenza replication in Drosphilia cells [20]) among the 5 studies. For example, one screen was conducted using reagents in Drosophila, 2 studies focused on early virus replication events, and 3 on the entire replication cycle. It is important to note that the 5 influenza screens did identify common pathways relevant to virus infection – specifically, receptor tyrosine kinases (RTK), protein kinase C (PKC), phosphatidylinositol 3-kinase (PI3-K), Raf/MEK/ERK signaling NF-κB signaling.

There have been a variety of other partial and full-genome screens against other virus infections, including human immunodeficiency virus (HIV) [26-28], human cytomegalovirus (HCV) [29-33], Dengue Virus [34, 35], West Nile virus [36], and more recently, vesicular stomatitis virus [37]. Each of these screens has used different approaches, and like influenza, even where the same virus was used for the screen, there has been limited overlap of genes identified as important for virus replication [38, 39]. However, despite the lack of overlap of specific genes,

common pathways have been identified through meta-analysis [39]. Specific host genes that influence virus replication have also been validated in some of these virus screens. For example, a screen of kinases involved in HCV replication identified carboxylterminal Src kinase (Csk) as important; when the Csk was silenced by RNAi, HCV replication was reduced HCV subgenomic replicon system. This was confirmed using a small molecule inhibitor of Csk, JK239 [30].

Host Factors Impacting Virus Life Cycles Identified by RNAi Screens

Interferon-Inducible Trans-Membrane Proteins

One of the most exciting aspects of RNAi screens is the characterization of host proteins with previously unappreciated roles in virus replication. One example of this is the interferon-inducible trans-membrane proteins, IFITMs, which as the name suggests, were first identified due to their up-regulation in response to interferon treatment in human neuroblastoma cells [40]. The IFITM protein family has over 30 member proteins. A genome-wide RNAi screen of influenza A H1N1 virus identified IFITM3 as a gene inhibiting influenza replication, which was subsequently validated in a secondary screen using deconvoluted siRNAs. IFITM1, 2 and 3 were shown to disrupt the early stages of influenza replication, while IFITM3 also inhibited replication of flaviviruses such as West Nile virus and dengue virus. Demonstrating the critical nature of this protein in antiviral immune responses, IFITM3 was required by interferon (IFN)-α and IFN-γ to confer antiviral resistance in U20S and HeLa cells challenged with influenza virus. It can therefore be stated that IFITM3 in particular plays an important role in the IFN response.

Subsequent work demonstrated that IFITM proteins 1, 2 and 3 inhibited the early replication stages of HIV [41], a function that relied on the intracellular domains of the proteins. Studies were subsequently expanded to show that IFITM proteins display antiviral properties against Marburg and Ebola filoviruses and SARS coronavirus [42]. Here it was shown that IFITM proteins differentially restrict early virus replication depending on virus types − for instance, IFITM1 inhibits Marburg and Ebola viruses more so than IFITM3, while IFITM3 was again shown to be more important for blocking influenza virus than IFITM 1 or 2. The exact mechanisms of IFITM antiviral activities are still being elucidated. As they are relevant for a diverse range of viruses, it is currently thought that IFITMs antagonise viruses during the late endocytotic pathway [42]. Collectively these studies demonstrate a role for IFITM proteins in the early antiviral response against a broad range of enveloped viruses.

The Coatamer Proteins

Proteins associated with the COP-I coatamer complex are perhaps the best-known example of a host factor required for the replication of a diverse range of viruses. In spite of the poor overlap of hit candidates between screens of the same virus – COP-I complex members have been identified as required for influenza virus replication, and for other viruses such as HIV, Dengue and West Nile virus. The COP-I complex is comprised of 7 protein subunits, α-, β-, β'-, γ-, δ-, ε- and ζ-COP [43]. COP-I coated vesicles are approximately 75-100 nm, and were first identified localised to the Golgi apparatus [44, 45] and to the endoplasmic reticulum [46]. The precise role of COP-I complexes in the early secretory pathway is not clear. COPI- coated vesicles are involved in the retrograde transport of transmembrane proteins to the endoplasmic reticulum [47, 48], while prospective roles in anterograde transport are less clear. A recent model proposes that COPI-I vesicles move both in retrograde and anterograde directions to mediate anterograde cargo transport [49].

The protein components of these COPI-coated vesicles have been identified in a number of high throuput siRNA screening studies as being an important host factor during virus infection [31, 36, 50]. The α-COP protein in particular is of interest as an important protein involved during virus infection [51]. RNAi-mediated inhibition of COP proteins, including COPA in PC3 prostate cancer cells [52], and COP-B [53], COPZ-1 [52] and COPZ-2 [52] in other cell types,causes a marked dispersion or "collapse" of Golgi structure. Cytotoxicity associated with COP protein depletion is considered a consequence of modulating the autophagy pathway as a result of ER stress caused by Golgi dispersion [54]. Treatment of cells with the fungal metabolite Brefeldin A (BFA) inhibits COP-I vesicle function in an analogous manner to RNAi interference, and also results in Golgi dispersion. This can be attributed to the effect of BFA on ADP ribosylation factor 1 (Arf1), a GTP-binding protein required for COPI complex formation. The activation of Arf1 by a guanine nucleotide exchange factor (GEF) containing a Sec7 domain [55] coaxes the exchange of bound GDP to GTP, leads to the engagement of the subunits (subcomplexes) from the cytosol to Golgi membranes [56]. Subsequently, assembly takes place to unite these subcomplexes into one structure. The effect of BFA on Arf1 potently inhibits anterograde membrane transport, COP proteins binding to membranes [57], and culminates in Golgi apparatus dispersion and redistribution of Golgi proteins to the ER. BFA impacts replication of a range of viruses, including poliovirus [58], vaccinia virus, and HIV-1 [59].

The mechanism by which COPA and COPI modulation impacts virus replication may be multi-pronged. Gross impairment of the trans-Golgi network observed for COP depletion would be reasonably expected to impair maturation of the virus proteins requiring glycosylation, such as the haemagglutanin (HA) glycoprotein, therefore inhibiting assembly of infectious virus particles. Indeed, BFA treatment decreases the expression of HA on the plasma membrane of influenza-infected cells [60]. However, COPI coatomer proteins are also associated with endocytosis [61, 62]. Relatedly, depleting the temperature-sensitive epsilon-COP inhibits endocytosis of Semliki Forest virus and vesicular stomatitis virus (VSV) [63]. A more recent study has demonstrated that COPI inhibition causes a primary inhibition of VSV internalisation, with a secondary impairment of viral RNA replication [64].

Rab6 GTPase

Rab6 is a Golgi-localised GTPase that regulates endosome to-Golgi and Golgi-to-ER retrograde transport. Rab6 was identified in a genome-wide RNAi screen of host factors required for HIV [28]. While the depletion of Rab6 in HeLa cells restricted HIV replication, this inhibitory effect was not observed when HIV was pseudotyped with VSV G envelope protein, which alters the mode of HIV cellular entry from direct cell membrane fusion to virus endocytosis followed by endosome membrane fusion.

Interestingly, while this first report suggested that Rab6 facilitates HIV entry, subsequent work indicated a role for Rab6 in human CMV assembly. The intracellular formation of CMV virions occurs at perinuclear assembly compartments where viral proteins – including the CMV tegument protein pp150 – localise. The localisation of pp150 to assembly compartments requires an interaction with Bicaudal D1 (BicD1), a host effector protein of Rab6. BicD1 co-localizes with Rab6 in the trans-Golgi network [65]. Disruption of Rab6 function in CMV-infected cells interrupts pp150 intracellular movement and virus production without disturbing formation of assembly complexes [66].

Calcium/Calmodulin-Dependent Protein Kinase (CaM Kinase) II Beta

Calcium/calmodulin-dependent protein kinase (CaM kinase) II beta (CAMK2B) is a widely expressed calcium-binding protein that modulates a range of cellular functions, including those tied to actin-mediated cytoskeletal control and CREB-facilitated transcription. A genome-wide RNAi screen identified CAMK2B as a host gene impacting influenza A (WSN/33) replication [17]. Inhibition of CAMK2B by the chemical compound KN-93 inhibited replication of influenza

A/WSN/33 and swine-origin influenza virus, but not VSV, in MDCK cells, hinting that pharmacological modulation of CAMK2B may be a valuable antiviral strategy. How influenza virus replication is facilitated by CAMK2B is not fully understood. Cells with depleted CAMK2B show impaired nuclear translocation of the influenza NP protein, but not defects in cellular virus entry or the formation of viral ribonucleoproteins in the nucleus [17]. The authors suggested that CAMK2B was regulating viral RNA transcription.

CAMK2B has been implicated in the replication of another virus genome – the double stranded DNA virus African swine fever virus (ASFV) [67]. Replication of ASFV involves the restructuring of scaffolding proteins known as type III intermediate filaments, and a major protein component of such structures known as vimentin [68]. CAMK2B is required for the phosphorylation of vimentin which triggers the disassembly and movement of vimentin filaments on microtubules. Both viral DNA replication and vimentin post-translational modifications (*e.g.*, phosphorylation) are prevented by KN-93 in Vero cells [67].

Members of the calcium/calmodulin kinase networks were also over-represented in an RNAi screen of druggable genes regulating the replication of Ebola virus [69]. Treatment of HEK 293 cells with KN-93 inhibited infection efficiency of Ebola virus, Zaire strain, pseudotyped with lentivirus. The authors of this study point out that while CAMK2B inhibitors have not been tested clinically as antivirals, studies that implicate calcium modulation in virus replication identify additional targets for therapeutic development.

Towards Broad-Spectrum Antiviral Strategies

The plethora of data generated from RNAi screens of host-virus interactions not only furthers understanding of host pathways impacting virus replication cycles, it also provides an opportunity to drive ahead translational outcomes such as new antiviral therapeutics, diagnostics and vaccines. One example of this, the enhancement of influenza virus vaccine production in embryonated chicken eggs, resulting from RNAi-mediated impairment of the host antiviral response, has received considerable commercial interest, and will be discussed later in this book.

While strategies to treat emerging infectious diseases benefit from virus-specific genome wide RNAi screens, it also raises the case for compiling existing information from completed RNAi screens for the development of broad-

spectrum antivirals. This is particularly relevant for sudden disease outbreaks, where a lack of time and resources hinders the management of outbreaks during the crucial early stages. For the approximately fifty FDA-approved antiviral drugs, over 80 % of these have viral targets, particularly targeting viral enzymes [70]. While this strategy offers benefits in regards to minimising side effects, it encounters obvious limitations in regards to the application of drugs to other viruses. Few if any of these drugs are broadly active. The alternate strategy – targeting the host – requires a detailed knowledge of virus-essential host factors for a wide range of viruses, and furthermore, knowledge of the consequences of modulating host gene expression.

The concept of meta-analysing screen data from many RNAi screens presents several challenges. Firstly, there is the issue of a poor degree of agreement in candidate genes observed when screens are conducted by different research groups. However, taking influenza virus as a case in point, there are obvious explanations for these perceived anomalies. For example, for the 5 RNAi screens performed on this virus, there were differences in the type of virus used, multiplicity of infection, duration of infection, cell type used, readout method deployed, among others. Due to these technical differences from screen to screen, it has been proposed that entire pathways be the subject of meta-analysis, rather than individual genes. Performing this analysis clearly identifies pathways relevant to influenza infection [71-73]. We have discussed these pathways in a recent review [73].

Another issue for consideration is the safety of modulating host gene expression – however briefly – to combat a viral infection. Firstly, the consequences of modulating any host protein may be deleterious. One example in this context are the COP-I complex proteins, mentioned in the previous section. While COPI-I proteins are required for the replication for most viruses tested by RNAi screening, their integral role in ER to Golgi cargo trafficking, and the effect of COP-I depletion on Golgi apparatus structure (and function), calls to question whether it would be tolerable for the host to impair this pathway, however briefly. In such instances, a relevant discussion point may be the pathogenicity of the viral pathogen. For example, an individual infected with Hendra virus or Ebola virus would likely be more willing to be treated with a therapy that impairs COP-I function – a potentially life-threatening scenario – more than someone infected with a low pathogenic virus such as seasonal influenza virus.

BIOLOGICAL DELIVERY OF RNAi

Antiviral siRNAs Targeting Host and Virus Genes

The objective of deploying RNAi to control viral disease in a human disease scenario inevitably requires the use of synthetically produced siRNA as a form of antiviral drug. Options that employ a transgenic approach for humans will be very restricted in technical applicability and will encounter substantial regulatory hurdles. Since the first recognition of the mechanism of RNAi its application to control viral infection by delivery of siRNA has been recognized and tested. The most obvious targets were those viruses that initially or predominantly infect accessible tissues, principally those of the respiratory tract including respiratory syncitial virus (RSV), parainfluenza virus (PIV) and influenza virus. Early studies showed promise for the activity of siRNA against all of these viruses [74, 75]. For such an accessible location it may be possible to use siRNA as it is without formulating or complexing reagents. However the entry of siRNA molecules into the target cells in which their antiviral action is required is a complex process that is not fully understood. Although siRNAs are relatively small as nucleic acids, only 19-21 base pairs on average, as drug molecules they are very large, being in the range of 15 kDa and highly negatively charged. This means they will not passively transit the lipid bilayer of the cell membrane. If entry is not *via* direct membrane transit then the uptake of siRNA into the cell is expected to be *via* one or more of a series of mechanisms, including receptor mediated endocytosis, non-specific endocytosis, clathrin-mediated endocytosis, caveolae, micro-pinocytosis, macro-pinocytosis or phagocytosis. This leads to siRNA within the cell entrapped in a vesicular body which may be on a pre-programmed pathway through the cell, possibly on a destructive journey to a lysosome. At some point the siRNA has to escape its membrane bound chamber in order to access the cytoplasm, to interact with the RNA Induced Silencing Complex (RISC) to deliver the desired effect of gene specific silencing. Mediating this outcome is the design objective of various biomaterials, from simple preparations of polyethyleneimine (PEI), modifications of chitosan, through to more complex Dynamic Polyconjugates [76] and other biopolymers [77].

There are a number of research reagents available for transfection of siRNA which work to great effect on a number of cell types *in vitro*. Some have been used *in vivo* experimentally but none of these have made the transition to clinical use in humans. Bitko *et al.* (2004) [74] employed TransIT-TKO (Mirus Bio Corp, USA) in a mouse model administered intranasally and showed significant knockdown of both RSV and parainfluenza virus. Influenza A virus replication

has been controlled in mouse models of infection using hydrodynamic (short duration-high pressure intravenous) delivery of siRNAs targeting the NP or PA genes, in combination with intranasal administration of siRNA formulated in oligofectamine (Invitrogen) [78] or by intravenous delivery of PEI complexed with siRNAs targeting the NP gene [75].

Systemic or parenteral delivery may be required for non-respiratory virus infection or when systemic spread has occurred. The tissues may be more difficult to access especially if this involves the brain or other privileged sites beyond the normal vasculature, *e.g.,* retinal blood barrier, testis blood barrier. Probably the most accessible tissue subject to morbid viral disease is the liver. Its role as part of the reticuloendothelial system (RES or mononuclear phagocytic system) scavenging wastes is a double edged sword. The Kuppfer cells of the liver pick up a great deal of material from circulation in their role in the MPS, but it is often the liver hepatocytes that are the target of drug delivery or viral infection. Hepatitis B virus (HBV) is a widespread infection and significant burden in public health and a beckoning target for siRNA therapy. However, the need to deliver intravenously and get the majority of the dose to hepatocytes in the liver remains a significant hurdle [79]. Early applications of lipid-encapsulated, chemically modified siRNA showed promise in terms of controlling HBV in a mouse model of virus replication, but required repeated daily dosing to a achieve between a 1 and 2 log reduction in circulating HBV DNA [80]. Two methods with similar approaches using cationic lipid/cholesterol formulations have had some success in controlling replication of this virus in a mouse model. One utilized co-formulation with apolipoprotein A-1 to target siRNA delivery to hepatocytes and showed reduction in HBV markers for 8 days [81]. Another study used co-formulation with polyethyleneglycol (PEG) as a stealth agent and hydrodynamic i.v. delivery of the siRNA to increase the duration of the anti-HBV effect [82]. More recently this formulation was used to treat hepatitis C virus infection in the mouse, achieving 65-75% inhibition of viral gene expression with a 2mg/kg dose of siRNA, with an improved efficiency of 95% and an increased duration of effect to 6 days by using a 2'-O-methyl modified siRNA [83].

A wide range of delivery approaches have been employed *in vivo* with varying degrees of success against a variety of significant viruses [84]. A number of delivery approaches are discussed by Guzman-Villanueva *et al.* [85]. These include lipoplexes (particularly those containing cationic lipids), PEI (alone or complexed with PEG), dendrimers, polymeric carriers (including Dynamic Polyconjugates), gold nanospheres, direct cholesterol conjugates and stable

nucleic acid lipid particles (SNALPs). siRNA with chemical modifications have been used in combination with many of these systems but they have also been used alone, so-called naked siRNA delivery. These modifications are generally based around the 2' hydroxyl of the ribose sugars in the RNA backbone and include modifications such as 2'O-methyl, 2'-fluoro, 2'-O-methoxyethyl, phosphorothioate, locked nucleic acid, and unlocked nucleic acid. In many mouse models of naked siRNA or formulated siRNA hydrodynamic delivery has been used, which involves rapid (5-7 sec) injection of up to 10% of animal body weight in fluid. This is a method that clearly is not acceptable for human use.

Of the methods for delivery of siRNA that have made it through to clinical assessment it is predominantly SNALPs and naked (modified) siRNA that have been targeted to viral disease [85]. TKM-Ebola by Tekmira intravenous SNALP formulation, ALN-RSV01 by Alnylam intranasal are the key contenders.

The use of siRNA offers great potential as a treatment for viruses that inflict a variety of human afflictions. Our ability to demonstrate the effectiveness of therapeutic siRNA against viral target genes remains relatively trivial *in vitro*, but the translation of this to humans in the clinic currently remains tantalizingly out of reach for most of our targets. A simple, effective formulation of siRNA that will get it to the appropriate target cell with sufficient penetration and duration remains the Holy Grail. As we learn more it is clear that simplicity may be a vain hope. RNAi has the potential to be the one-size-fits-all solution to knocking down these unwanted or exogenous gene expression. But it is most likely that each virus, with its characteristic tissue tropism and mode of spread, will require a tailor-made delivery vehicle to bring the effector to the target. In the long term the current advances in lipid bases and polymeric delivery vehicles point to great hope for a future in which RNAi can make the impact on viral disease that antibiotics made on bacterial disease in the middle of the last century.

The biosecurity needs of a burgeoning world population and the growing threat of EIDs put tremendous pressure on the medical system as it struggles to combat a variety of viral diseases. Capability now exists to rapidly identify and characterize new viral pathogens, including extracting their nucleotide sequence. This provides the Achilles Heel to attack these pathogens through RNAi in a way that can be rapidly adapted to these new viruses. With the devastating mortality and economic consequences of outbreaks in recent decades of H5N1, SARS, Nipah virus, Hendra virus and others, it is clear that emphasis on getting RNAi delivery through to the clinic should be a priority research objective.

Biological Delivery of RNAi

For a siRNA delivery system to be successful, it must have a number of specific properties. To begin with, it must be able to protect the siRNA from degradation. It must also be able to bind to, and enter the target cell, and must be able to efficiently deliver the siRNA to the cell nucleus. For this reason, scientists have utilised virus delivery systems as they have evolved to display all of these properties (4). Current literature has described successful delivery of both miRNAs and siRNAs using a variety of different virus vectors, from lentiviruses and adenoviruses to replication-competent avian with splice acceptor (RCAS) retroviruses. These systems are discussed briefly below.

Lentivirus Systems

Lentiviruses are members of the retrovirus family and contain two copies of a single stranded RNA genome [86]. Members of the retrovirus family are often highly pathogenic and include HIV. As such they are typically engineered to form 'self inactivating' (SIN) vectors. This usually involves removing the U3 region in the LTR and replacing it with a heterologous promoter such as CMV [87]. The key benefits of lentivirus delivery systems are that they are capable of infecting both dividing and non-dividing cells, making them prime candidates for neurological RNAi studies. In addition, they are able integrate into the genome of the host and stably persist and can be modified to broaden their tropism [86].

Limitations associated with lentiviral delivery include multiple insertions into the genome, no control of the location of insertion which may impact gene expression, a restriction on cell lineages and types infected with the virus, and the insertion of viral elements that may be unwanted into a genome.

Adenovirus Systems

Adenoviruses are one of the most popular virus delivery systems, with over 20 % of gene therapy clinical trials utilizing adenovirus transport systems (Wiley Gene Therapy   Database,   2012; http://www.wiley.com/legacy/wileychi/genmed/clinical/).   Adenoviruses   are double-stranded DNA viruses with the ability to package up to 36 kb of foreign DNA [88]. Of note is their ability to express high levels of siRNA and the ease of vector construction. These systems have been used for various RNAi therapies including; cancer targeting [89], neurodegenerative disease treatment (motor neuron disease) [90] and the targeting of other viruses such as HBV [91]. However, as with all delivery systems a number of limitations have been noted.

Studies have shown that in some cases, high levels of immunostimulation have been observed following adenovirus administration. Moreover, relatively short siRNA expression periods have also been problematic [86].

Other Delivery Systems

Replication-Competent Avian with Splice Acceptor

Replication-competent avian with splice acceptor (RCAS) retroviruses are used to permanently integrate siRNAs into the genome of avian models, most typically chickens [92]. However, modified versions have also previously been used in mouse models [93].

Adeno-Associated Viruses

Adeno-associated viruses (AAV) have small genomes (4.8 kb) comprised of ssDNA. The viruses depend upon a second virus (for example adenovirus) for productive infection to occur [94]. The entire genome of AAV contains only two genes that can be replaced allowing up to 5 kb of foreign DNA to be added. Unlike adenoviruses, AAVs are considered to have a low immunostimulatory response and can be modified to change tropism, thereby the cells it targets. However, similar to adenoviruses, AAVs are not capable of integrating with the host genome and therefore are not a viable option for long term gene silencing [87].

Influenza virus and Herpes Simplex Virus

Influenza A viruses have been engineered to express a neuron-specific microRNA (miR-124) *in vitro* and *in vivo* [95] demonstrating that it is possible for a single stranded RNA viruses to express miRNAs. Herpes simplex viruses have also been effectively used as a shRNA delivery system [96]. They have silenced the expression of an amyloid precursor protein gene and have reduced levels of amyloid-β peptide (a cleavage product of amyloid precursor protein) believed to be important in Alzheimer's disease pathogenesis.

Bacterial Delivery Systems

While extensive preclinical and clinical studies have been largely focused on viral delivery of shRNAs, bacteria have also been modified to function as delivery systems. Xiang *et al.* transfected *Escherichia coli* with a plasmid encoding genes to allow it to not only enter target cells in the colon but to also release a shRNA

targeting the CTNNB1 cancer gene [97]. The same gene has also been targeted using Salmonella enterica using shRNA [97].

RNAi and Clinical Trials

The ability of RNAi to regulate endogenous gene expression is currently being harnessed by numerous research groups, in collaboration with commercial partners, for clinical application in the treatment of human disease. Herein we describe a number of completed, ongoing or pending trials and discuss future directions of this technology.

Non-biological delivery systems have been based on nanoparticles, aptamers, cholesterol and stable nucleic acid lipid particles. Details of the individual systems and appropriate references have been discussed in detail in previous sections. Many of these have now been deployed in Phase I, II and in one case, Phase III clinical trials. The pathway to clinical translation has been difficult. A number of significant obstacles have been encountered and must be considered in current and future applications. These include: (i) delivery, (ii) RNAse-mediated degradation and renal clearance; chemical modification of RNAi molecules to ensure bioreactivity, (iii) targeting to specific tissues and (iv), careful analysis of safety profiles to reduce or remove unwanted off target effects. Having said this, the NIH clinical trials database lists 33 current, completed or pending siRNA studies (see http://www.clinicaltrials.gov). These trials cover a number of conditions including macular degeneration, glaucoma, numerous cancers (*e.g.,* myeloid leukaemia, solid tumours, melanoma, liver and pancreatic), kidney transplantation and renal failure (see Table **1** below).

Table 1: National Institutes of Health RNAi Clinical Trials

Clinical Trial/Sponsor	Target	Medical condition	Status
NCT01591356 M.D. Anderson Cancer Center	EphA2	Advanced cancers	Pending Phase I
NCT00716014 Pachyonychia Congenita Project	PC keratin, K6a	Pachyonychia Congenita	Completed Phase I
NCT00363714 Allergan	Vascular endothelial growth factor receptor-1	a) Macular degeneration b) Choroidal Neovascularization	Completed Phase I Phase II
NCT00672542 Duke University	Immunoproteasome beta subunits LMP2, LMP7, and MECL1	Melanoma	Ongoing Phase I

Table 1: contd…

NCT00395057 **Allergan**	**Vascular endothelial growth factor receptor-1**	**a) Macular degeneration** **b) Choroidal Neovascularization**	**Terminated Phase II Phase II**
NCT00257647 Hadassah Medical Organization	Fusion genes	Chronic Myeloid Leukeamia	Not stated
NCT00689065 Calando Pharmaceuticals	M2 subunit of ribonucleotide reductase	Solid tumour cancers	Ongoing Phase I
NCT00938574 Silence Therapeutics AG	Protein kinase N3	Advanced Solid tumours	Completed Phase I
NCT01064505 Quark Pharmaceuticals	Caspase 2	Optic atrophy	Ongoing Phase I
NCT00306904 Opko Health, Inc.	VEGF	Diabetic Macular Edema	Completed Phase II
NCT00927459 Tekmira Pharmaceuticals Corporation	Apolipoprotein B	Hypercholesterolemia	Terminated Phase I
NCT01437007 National Cancer Institute (NCI)	PLK1	Liver cancer	Completed Phase I
NCT00154934 National Taiwan University Hospital	IL-10	Preeclampsia	Not stated
NCT01075360 National Taiwan University Hospital	Aromatic hydrocarbon receptor	Neuroblastoma	Not stated
NCT01676259 Silenseed Ltd	KRAS oncogene	Advanced pancreatic cancer	Pending Phase II
NCT01438281 Sylentis, S.A.	TRPV1 Receptor	Ocular pain	Completed Phase I
NCT01188785 Silenseed Ltd	KRAS oncogene	Adenocarcinoma of the pancreas	Ongoing Phase I
NCT00554359 Quark Pharmaceuticals	p53	Acute renal failure	Completed Phase I
NCT00802347 Quark Pharmaceuticals	p53	Delayed Graft Function in Kidney Transplantation	Ongoing Phase I
NCT01227291		Glaucoma Ocular Hypertension	Completed

Table 1: contd…

Sylentis, S.A.	beta 2 adrenergic receptors		Phase I Phase II
NCT01058798 National Taiwan University Hospital	β1,4-N-acetylgalactosaminyltransferase III	Neuroblastoma	Not stated
NCT00259753 Opko Health, Inc.	VEGF	Macular degeneration	Completed Phase II
NCT01445899 Quark Pharmaceuticals	RTP801	Diabetic macular edema	Ongoing Phase II
NCT00557791 Opko Health, Inc.	VEGF	Macular degeneration	Withdrawn Phase III
NCT00496821 Alnylam Pharmaceuticals	N protein	Respiratory syncytial virus	Completed Phase II
NCT01617967 Alnylam Pharmaceuticals	transthyretin	Amyloidosis	Ongoing Phase II
NCT01437059 Alnylam Pharmaceuticals	PCSK9	Elevated LDL-Cholesterol	Completed Phase I
NCT01559077 Alnylam Pharmaceuticals	transthyretin	Amyloidosis	Completed Phase I
NCT01158079 Alnylam Pharmaceuticals	KSP/VEGF	Solid Tumours	Completed Phase I
NCT00882180 Alnylam Pharmaceuticals	KSP/VEGF	Solid Tumours	Completed Phase I
NCT01148953 Alnylam Pharmaceuticals	transthyretin	Amyloidosis	Completed Phase I
NCT01065935 Alnylam Pharmaceuticals	N protein	Respiratory syncytial virus	Completed Phase IIb
NCT00658086 Alnylam Pharmaceuticals	N protein	Respiratory syncytial virus	Completed Phase II

Clinical Trial Highlights

It is important to note here that regulatory authorities have not as yet recorded serious adverse reactions following RNAi delivery. At the moment, the most advanced therapeutic in the RNAi pipeline is ALN-RSV (Alnylam Pharmaceuticals), which is currently undergoing Phase IIb trials and targets the expression of the RSV N-protein. The target endpoint focused on the occurrence of new or progressive bronchiolitis obliterans syndrome (BOS) at 180 days post

viral infection. BOS is major source of morbidity and mortality in lung transplant patients, with re-transplantation being the only definitive BOS treatment. RSV is responsible for up to 25 % of viral lung infections occurring in BOS transplant patients. Human experimental infection study showed ALN-RSV01 significantly decreased infection rate in healthy volunteers infected with RSV intranasally [98]. The phase II clinical trial fell short in an intent-to-treat (ITTc) program, but showed greater promise in prospectively defined analysis of ITTc recipients [99].

APPLICATION OF RNAi AND TRANSGENIC TECHNOLOGIES TO DEVELOP VIRUS RESISTANT ANIMALS

Introduction

Transgenic (Tg) species will play a vital role in meeting global dietary and caloric needs and transgenic technology has been successfully applied to both agricultural & aquatic animals (see Tables in [100-102]). Proof-of-principle studies have demonstrated the potential application of RNAi in: (1) production of predominantly female progeny (dairy and egg industries), thereby minimizing *e.g.*, castration and elimination of males; (2) Tg swine to improve meat output by minimizing fatality and increasing the rate at which young animals mature; (3) Tg cows whose mammary glands express an anti-bacterial agent against pathogens responsible for mastitis in dairy cattle; (4) Tg swine that synthesize omega-3 fatty acids, and (5) "Enviro" pigs which release reduced amounts of phosphorous into the surroundings.

In aquatic species, similar enhancements have been achieved (reviewed in [103]) that augment: (1) growth; (2) bacterial resistance; (3) nutritional appeal; and (4) temperature tolerance.

This section focuses on how RNAi in combination with transgenic technologies would control of viral diseases and thus facilitate animal production. It was almost a decade ago when Clark and Whitelaw proposed in their review "A future for transgenic livestock" that the advent of the then new method of RNA interference (RNAi) for modifying genomes will underpin a resurgence of research using transgenic livestock [104]. They suggested this may be an important alternative to traditional breeding and could lead to the generation of farm animals that are more resistant to infectious disease (*e.g.*, influenza resistant poultry). With the advent of RNAi technology and the newest generation of gene editing tools, genetically engineered animals become an alternative to vaccines and small molecule drugs

[105] circumventing the costs associated with these controls, which often limits their uptake and effective implementation.

Applications of RNAi for Disease Resistant Animals

There are a number of approaches that take advantage of combined RNAi and transgenic techniques to ultimately develop disease resistant animals. These include (1) targeting host genes for functional analysis of host pathogen interactions; (2) targeting host genes required by the virus for infection; and (3) targeting of viral genes to prevent replication and spread of infection.

The biggest impact of RNAi so far has been in the study of gene function. RNAi is now a common tool for studying biological processes including host pathogen interactions. This work has primarily been conducted *in vitro* using siRNAs (see RNAi screening section in this chapter) or shRNA expression constructs. From these screens key genes of interest are identified, however, the transfer of this technology to laboratory and production animals has been limited due to difficulties in the delivery of RNAi molecules *in vivo* (see delivery section in this chapter) – this can now be circumvented using transgenic technology which allows integration of shRNA/miRNA sequences into the genomes of target animal species. This also allows tissue specific, constitutive or inducible expression of shRNA/miRNAs depending on the host-pathogen model. Whilst the outcome of this application may not directly lead to the development of viral resistant animals, it will lead to advances in the development of new vaccines and antiviral therapeutics. It may also lead to the identification of key regulatory sequences within genomes which can be used in conventional selective breeding programs for viral disease control.

A more direct application of RNAi and transgenics is to target key genes involved in virus interaction with host cells to thereby limit or prevent the attachment of virus to cells, replication of virus once in the cell and subsequent spread of infection from cell to cell. Luo *et al.* recently used shRNAs to target the porcine integrin αv subunit, which is the foot and mouth disease virus (FMDV) receptor, in a porcine cell line [106]. They reported that the inhibitory effect of the receptor knock down on FMDV growth was >3-fold and was accompanied by a > 99% decrease in virus titre when cells were exposed to 10^2 TCID$_{50}$ of FMDV. This study identifies shRNA as plausible reagents to control FMD infection and dissemination in pigs and other susceptible livestock species. For this strategy, receptor gene knock down may be preferable to alternative knock out approaches as complete abrogation of these proteins may be detrimental to the host. Although

this will not completely eliminate the virus, it should reduce virus levels to allow the host immune response to work in synergy with the RNAi activity and successfully fight infection.

Perhaps the most promising approach is targeting of viral genes within host cells. The basic concept of this technology is that the transgenic animal has an RNAi transgene inserted into the genome leading to expression in every host cell. This transgene expresses shRNA(s) specifically targeting conserved regions within key viral genes. By this approach, essential viral functions, including viral replication, can be eliminated or slowed, thereby disrupting the viral lifecycle. Furthermore, by designing the constructs to target multiple viral genes, this strategy can allow broad protection across a range of serotypes and strains while helping to avert the appearance of resistant viral mutants [105]. There are numerous examples of virus-resistant transgenic animals currently in development including; avian influenza resistant chickens [107], Foot and Mouth Disease resistant ruminants [108], viral hemorrhagic septicaemia resistant zebrafish as a model for salmon infection [109] and the development of transgenic swine capable of inhibiting porcine retrovirus replication [110]. Circuitously, the latter pathogen has been particularly important for humans since the virus can replicate in human cells *in vitro*, and is thus an obstacle to swine xenotransplant advancements.

Design of RNAi Transgenes

There are a number of considerations to take into account when designing RNAi transgenes for this purpose. These include (1) targeting choice; (2) shRNA design (3) promoter choice.

A combinatorial RNAi approach is desirable (Burkout *et al.*) targeting conserved sequence regions within a number of essential viral genes such as polymerases or key structural genes. Sometimes referred to as a "multi-warhead transgene" this also enables targeting of strategic genes at various time points of the virus infection cycle, maximising the opportunity to interfere with and prevent virus spread.

The design of the RNAi hairpin against the identified target is also critical to enable effective processing *via* the RNAi pathway in the cell. There are a number of algorithms available for siRNA design however these do not always translate to effective shRNA sequences. One algorithm to screen siRNAs to see if they may be effective as shRNAs was developed by Taxman *et al.* and is commonly used for this purpose [111]. It is still essential, however, to validate shRNAs against the

target virus in an *in vitro* virus replication assay prior to undertaking the considerable task of generating a transgenic animal. It is also crucial to ensure that the shRNAs do not inhibit host gene expression in an off target manner. Homology between the shRNAs and the host genome should be examined (*e.g.,* BLASTN) and sequences with near perfect identity should be excluded.

Approximately 60 % of studies use the Brummel-Kamp loop [112]; however, a recent development in the design of antiviral shRNAs is mimicking the structure on naturally occurring miRNA sequences, this has been shown by a number of studies to have benefits such as increased processing to the mature form. To date studies which have looked at shRNAs with a miRNA structure have primarily focused on changing the loop to a mir loop sequence *e.g.,* mir30 but not on introducing bulges into the stem which miRNAs also contain [113, 114]. An example of this has been used by Boden *et al.* who enhanced gene silencing efficacy using an anti-HIV shRNA that was designed with a mir30 structure. In general shRNAs and miRNAs act in different ways to knock down gene expression (mRNA degradation *vs* translational repression) and this also must be taken into consideration when designing and developing an RNAi transgene. A study by Lebbink *et al.* [114] comparing miRised shRNA driven by polII promoter were less effective in achieving target knockdown compared to conventional shRNAs expressed from the RNA polymerase III promoter U6.

The third consideration when designing an RNAi transgene is promoter choice. Promoter strength is vital as too much expression can be detrimental or lethal to particular cell types or to embryo development. This is thought to be as a result of saturation of the RNAi pathway meaning key miRNAs that are required for cell differentiation are out competed by the antiviral shRNAs. The limiting step is thought to be transport of expressed hairpins from the nucleus to the cytoplasm *via* Exportin 5 protein [115]. RNA polymerase III (PolIII) promoters are commonly used for shRNA expression. PolIII family members include U6, 7SK and H1 promoters. The H1 promoter is the weakest in strength and is therefore often the promoter of choice for transgenic animal production. Although U6 is commonly used *in vitro*, it is too strong for *in vivo* use, although this can be overcome by site directed mutagenesis of key promoter elements leading to attenuated promoter strength. For particular viral targets it may be preferable to choose tissue specific or inducible promoters such as RNA polymerase II promoters. Species specific promoters are also an important consideration for the

development of optimised RNAi transgenes. This may also be an advantage when considering regulatory approval and consumer acceptance of transgenic animals in the future.

Transgenic Technologies

We have seen many advances in transgenesis technology in animals over recent years, including the use of retroviruses (*e.g.*, lentivirus), transposons (*e.g.*, piggy bac and Tol 2) and non-homologous recombination. These methods all result in the random integration of the transgene into the target genome, often in multiple copies and sometimes into important regulatory sequences within the genome. Therefore there is quite often a large amount of screening work required to identify and characterise the number of insertions and their locations, the level of transgene expression and ultimately the founder animals that will be used to breed transgenic offspring for viral challenge experiments to validate resistance. The recent advances in megenuclease technologies (zincfinger nucleases, TALENs) will allow the targeted integration of RNAi transgenes into selected regions of host genomes (often termed "safe harbours"). We see this as a major advance in the development of transgenic animals with RNAi transgenes conferring resistance to viral pathogens.

CONCLUSION

The ability to produce viral resistant livestock will increase the welfare status of production animals, contribute to increasing the quality and safety of food production particularly in intensively reared animals such as poultry, and most importantly serve to enhance food security worldwide. Perhaps more importantly, developing animals that are resistant to zoonotic viruses with pandemic potential such as H5N1 and H1N1 influenza is a key strategy for reducing the risk of pandemic emergence in humans.

ACKNOWLEDGEMENTS

Declared none.

CONFLICT OF INTEREST

The authors confirm that this chapter content have no conflict of interest.

REFERENCES

[1]   Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998;391(6669):806-11. Epub 1998/03/05.

[2]   Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Curr Biol. 2001;11(13):1017-27. Epub 2001/07/27.

[3]   Catalanotto C, Azzalin G, Macino G, Cogoni C. Gene silencing in worms and fungi. Nature. 2000;404(6775):245. Epub 2000/04/05.

[4]   Fagard M, Boutet S, Morel JB, Bellini C, Vaucheret H. AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. Proc Natl Acad Sci U S A. 2000;97(21):11650-4. Epub 2000/10/04.

[5]   World Health Organization. Seasonal influenza, retrieved 10th March 2010 <http://www.euro.who.int/influenza/20080618_1>. 2010.

[6]   Taubenberger JK, Morens DM. Influenza: the once and future pandemic. Public Health Rep.125 Suppl 3:16-26. Epub 2010/06/24.

[7]   Herfst S, Schrauwen EJ, Linster M, Chutinimitkul S, de Wit E, Munster VJ, *et al*. Airborne transmission of influenza A/H5N1 virus between ferrets. Science.336(6088):1534-41. Epub 2012/06/23.

[8]   Imai M, Watanabe T, Hatta M, Das SC, Ozawa M, Shinya K, *et al*. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. Nature.486(7403):420-8. Epub 2012/06/23.

[9]   Pawan JL. The transmission of paralytic rabies in Trinidad by the vampire bat (Desmodus rotundus murinus Wagner. Caribb Med J. 1959;21:110-36. Epub 1959/01/01.

[10]   Halpin K, Young PL, Field HE, Mackenzie JS. Isolation of Hendra virus from pteropid bats: a natural reservoir of Hendra virus. J Gen Virol. 2000;81(Pt 8):1927-32. Epub 2000/07/19.

[11]   Chua KB, Koh CL, Hooi PS, Wee KF, Khong JH, Chua BH, *et al*. Isolation of Nipah virus from Malaysian Island flying-foxes. Microbes Infect. 2002;4(2):145-51. Epub 2002/03/07.

[12]   Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, Yaba P, *et al*. Fruit bats as reservoirs of Ebola virus. Nature. 2005;438(7068):575-6. Epub 2005/12/02.

[13]   Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, *et al*. Bats are natural reservoirs of SARS-like coronaviruses. Science. 2005;310(5748):676-9. Epub 2005/10/01.

[14]   Ison MG. Antivirals and resistance: influenza virus. Curr Opin Virol.1(6):563-73. Epub 2012/03/24.

[15]   WHO. Overview of pandemic (H1N1) 2009 in the WHO European Region, retrieved 11th March 2010, <http://www.euro.who.int/influenza/AH1N1/20091026_1>. 2009a.

[16]   WHO. Seasonal influenza, retrieved 10th March 2010, <http://www.euro.who.int/influenza/20080618_1>. 2009b.

[17]   Konig R, Stertz S, Zhou Y, Inoue A, Hoffmann HH, Bhattacharyya S, *et al*. Human host factors required for influenza virus replication. Nature. 2010;463(7282):813-7. Epub 2009/12/23.

[18]   Meliopoulos VA, Andersen LE, Birrer KF, Simpson KJ, Lowenthal JW, Bean AG, *et al*. Host gene targets for novel influenza therapies elucidated by high-throughput RNA interference screens. FASEB J.26(4):1372-86. Epub 2012/01/17.

[19]   Stertz S, Shaw ML. Uncovering the global host cell requirements for influenza virus replication *via* RNAi screening. Microbes Infect.13(5):516-25. Epub 2011/02/01.

[20]   Hao L, Sakurai A, Watanabe T, Sorensen E, Nidom CA, Newton MA, *et al*. Drosophila RNAi screen identifies host genes important for influenza virus replication. Nature. 2008;454(7206):890-3. Epub 2008/07/11.

[21]   Konig R, Stertz S, Zhou Y, Inoue A, Hoffmann HH, Bhattacharyya S, *et al*. Human host factors required for influenza virus replication. Nature. 2009;463(7282):813-7. Epub 2009/12/23.

[22]   Brass AL, Huang IC, Benita Y, John SP, Krishnan MN, Feeley EM, *et al*. The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus. Cell. 2009;139(7):1243-54.

[23]   Karlas A, Machuy N, Shin Y, Pleissner KP, Artarini A, Heuer D, *et al*. Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. Nature. 2010. Epub 2010/01/19.

[24]   Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM, Wu L, *et al*. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. Cell. 2009;139(7):1255-67. Epub 2010/01/13.

[25]   Karlas A, Machuy N, Shin Y, Pleissner KP, Artarini A, Heuer D, *et al*. Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. Nature.463(7282):818-22. Epub 2010/01/19.

[26]   Zhou H, Xu M, Huang Q, Gates AT, Zhang XD, Castle JC, *et al*. Genome-scale RNAi screen for host factors required for HIV replication. Cell Host Microbe. 2008;4(5):495-504.

[27]   Konig R, Zhou Y, Elleder D, Diamond TL, Bonamy GM, Irelan JT, *et al*. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Cell. 2008;135(1):49-60. Epub 2008/10/16.

[28]   Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, Xavier RJ, *et al*. Identification of host proteins required for HIV infection through a functional genomic screen. Science. 2008;319(5865):921-6. Epub 2008/01/12.

[29]   Randall G, Panis M, Cooper JD, Tellinghuisen TL, Sukhodolets KE, Pfeffer S, *et al*. Cellular cofactors affecting hepatitis C virus infection and replication. Proceedings of the National Academy of Sciences. 2007;104(31):12884-9.

[30]   Supekova L, Supek F, Lee J, Chen S, Gray N, Pezacki JP, *et al*. Identification of human kinases involved in hepatitis C virus replication by small interference RNA library screening. J Biol Chem. 2008;283(1):29-36.

[31]   Tai AW, Benita Y, Peng LF, Kim SS, Sakamoto N, Xavier RJ, *et al*. A functional genomic screen identifies cellular cofactors of hepatitis C virus replication. Cell Host Microbe. 2009;5(3):298-307.

[32]   Ng TI, Mo H, Pilot-Matias T, He Y, Koev G, Krishnan P, *et al*. Identification of host genes involved in hepatitis C virus replication by small interfering RNA technology. Hepatology. 2007;45(6):1413-21.

[33]   Berger KL, Cooper JD, Heaton NS, Yoon R, Oakland TE, Jordan TX, *et al*. Roles for endocytic trafficking and phosphatidylinositol 4-kinase III alpha in hepatitis C virus replication. Proc Natl Acad Sci U S A. 2009;106(18):7577-82. Epub 2009/04/21.

[34]   Ang F, Wong AP, Ng MM, Chu JJ. Small interference RNA profiling reveals the essential role of human membrane trafficking genes in mediating the infectious entry of dengue virus. Virol J. 2010;7:24.

[35]   Sessions OM, Barrows NJ, Souza-Neto JA, Robinson TJ, Hershey CL, Rodgers MA, *et al*. Discovery of insect and human dengue virus host factors. Nature. 2009;458(7241):1047-50. Epub 2009/04/28.

[36]   Krishnan MN, Ng A, Sukumaran B, Gilfoy FD, Uchil PD, Sultana H, *et al*. RNA interference screen for human genes associated with West Nile virus infection. Nature. 2008;455(7210):242-5. Epub 2008/08/12.

[37]   Panda D, Das A, Dinh PX, Subramaniam S, Nayak D, Barrows NJ, *et al*. RNAi screening reveals requirement for host cell secretory pathway in infection by diverse families of negative-strand RNA viruses. Proceedings of the National Academy of Sciences. 2011;108(47):19036-41.

[38]   Pache L, König R, Chanda SK. Identifying HIV-1 host cell factors by genome-scale RNAi screening. Methods. 2011;53(1):3-12.

[39]   Bushman FD, Malani N, Fernandes J, D'Orso I, Cagney G, Diamond TL, *et al*. Host cell factors in HIV replication: meta-analysis of genome-wide studies. PLoS Pathog. 2009;5(5):e1000437. Epub 2009/05/30.

[40]   Friedman RL, Manly SP, McMahon M, Kerr IM, Stark GR. Transcriptional and posttranscriptional regulation of interferon-induced gene expression in human cells. Cell. 1984;38(3):745-55. Epub 1984/10/01.

[41]   Lu J, Pan Q, Rong L, He W, Liu SL, Liang C. The IFITM proteins inhibit HIV-1 infection. J Virol.85(5):2126-37. Epub 2010/12/24.

[42]   Huang IC, Bailey CC, Weyer JL, Radoshitzky SR, Becker MM, Chiang JJ, *et al*. Distinct patterns of IFITM-mediated restriction of filoviruses, SARS coronavirus, and influenza A virus. PLoS Pathog.7(1):e1001258. Epub 2011/01/22.

[43]   Beck R, Rawet M, Wieland FT, Cassel D. The COPI system: molecular mechanisms and function. FEBS Lett. 2009;583(17):2701-9. Epub 2009/07/28.

[44]   Serafini T, Stenbeck G, Brecht A, Lottspeich F, Orci L, Rothman JE, *et al*. A coat subunit of Golgi-derived non-clathrin-coated vesicles with homology to the clathrin-coated vesicle coat protein beta-adaptin. Nature. 1991;349(6306):215-20. Epub 1991/01/17.

[45]   Duden R, Griffiths G, Frank R, Argos P, Kreis TE. Beta-COP, a 110 kd protein associated with non-clathrin-coated vesicles and the Golgi complex, shows homology to beta-adaptin. Cell. 1991;64(3):649-65. Epub 1991/02/08.

[46]   Orci L, Perrelet A, Ravazzola M, Amherdt M, Rothman JE, Schekman R. Coatomer-rich endoplasmic reticulum. Proc Natl Acad Sci U S A. 1994;91(25):11924-8. Epub 1994/12/06.

[47]   Cosson P, Letourneur F. Coatomer interaction with di-lysine endoplasmic reticulum retention motifs. Science. 1994;263(5153):1629-31. Epub 1994/03/18.

[48]   Letourneur F, Gaynor EC, Hennecke S, Demolliere C, Duden R, Emr SD, *et al*. Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. Cell. 1994;79(7):1199-207. Epub 1994/12/30.

[49]   Orci L, Ravazzola M, Volchuk A, Engel T, Gmachl M, Amherdt M, *et al*. Anterograde flow of cargo across the golgi stack potentially mediated *via* bidirectional "percolating" COPI vesicles. Proc Natl Acad Sci U S A. 2000;97(19):10400-5. Epub 2000/08/30.

[50]   Li Q, Brass AL, Ng A, Hu Z, Xavier RJ, Liang TJ, *et al*. A genome-wide genetic screen for host factors required for hepatitis C virus propagation. Proc Natl Acad Sci U S A. 2009;106(38):16410-5. Epub 2009/09/01.

[51]   Watanabe T, Watanabe S, Kawaoka Y. Cellular networks involved in the influenza virus life cycle. Cell Host Microbe.7(6):427-39. Epub 2010/06/15.

[52]   Shtutman M, Baig M, Levina E, Hurteau G, Lim CU, Broude E, *et al*. Tumor-specific silencing of COPZ2 gene encoding coatomer protein complex subunit zeta 2 renders tumor cells dependent on its paralogous gene COPZ1. Proc Natl Acad Sci U S A.108(30):12449-54. Epub 2011/07/13.

[53]   Styers ML, O'Connor AK, Grabski R, Cormet-Boyaka E, Sztul E. Depletion of beta-COP reveals a role for COP-I in compartmentalization of secretory compartments and in biosynthetic transport of caveolin-1. Am J Physiol Cell Physiol. 2008;294(6):C1485-98. Epub 2008/04/04.

[54]   Platini F, Perez-Tomas R, Ambrosio S, Tessitore L. Understanding autophagy in cell death control. Curr Pharm Des.16(1):101-13. Epub 2010/03/11.

[55]   Chardin P, Paris S, Antonny B, Robineau S, Beraud-Dufour S, Jackson CL, *et al*. A human exchange factor for ARF contains Sec7- and pleckstrin-homology domains. Nature. 1996;384(6608):481-4. Epub 1996/12/05.

[56]   Bethune J, Kol M, Hoffmann J, Reckmann I, Brugger B, Wieland F. Coatomer, the coat protein of COPI transport vesicles, discriminates endoplasmic reticulum residents from p24 proteins. Mol Cell Biol. 2006;26(21):8011-21. Epub 2006/08/31.

[57]   Donaldson JG, Klausner RD. ARF: a key regulatory switch in membrane traffic and organelle structure. Curr Opin Cell Biol. 1994;6(4):527-32. Epub 1994/08/01.

[58]   Maynell LA, Kirkegaard K, Klymkowsky MW. Inhibition of poliovirus RNA synthesis by brefeldin A. J Virol. 1992;66(4):1985-94. Epub 1992/04/01.

[59]   Pal R, Mumbauer S, Hoke GM, Takatsuki A, Sarngadharan MG. Brefeldin A inhibits the processing and secretion of envelope glycoproteins of human immunodeficiency virus type 1. AIDS Res Hum Retroviruses. 1991;7(8):707-12. Epub 1991/08/01.

[60]   Ciampor F, Zavodska E, Cmarko D, Cmarkova J, Vareckova E. Effects of brefeldin A on the expression and transport of influenza A virus haemagglutinin, M1 and M2 proteins within the cell. Acta Virol. 1997;41(2):83-91. Epub 1997/04/01.

[61]   Whitney JA, Gomez M, Sheff D, Kreis TE, Mellman I. Cytoplasmic coat proteins involved in endosome function. Cell. 1995;83(5):703-13. Epub 1995/12/01.

[62]   Aniento F, Gu F, Parton RG, Gruenberg J. An endosomal beta COP is involved in the pH-dependent formation of transport vesicles destined for late endosomes. J Cell Biol. 1996;133(1):29-41. Epub 1996/04/01.

[63]   Daro E, Sheff D, Gomez M, Kreis T, Mellman I. Inhibition of endosome function in CHO cells bearing a temperature-sensitive defect in the coatomer (COPI) component epsilon-COP. J Cell Biol. 1997;139(7):1747-59. Epub 1998/02/07.

[64]   Cureton DK, Burdeinick-Kerr R, Whelan SP. Genetic inactivation of COPI coatomer separately inhibits vesicular stomatitis virus entry and gene expression. J Virol.86(2):655-66. Epub 2011/11/11.

[65]   Fuchs E, Short B, Barr FA. Assay and properties of rab6 interaction with dynein-dynactin complexes. Methods Enzymol. 2005;403:607-18. Epub 2006/02/14.

[66]   Indran SV, Britt WJ. A role for the small GTPase Rab6 in assembly of human cytomegalovirus. J Virol.85(10):5213-9. Epub 2011/03/18.

[67]   Stefanovic S, Windsor M, Nagata KI, Inagaki M, Wileman T. Vimentin rearrangement during African swine fever virus infection involves retrograde transport along microtubules and phosphorylation of vimentin by calcium calmodulin kinase II. J Virol. 2005;79(18):11766-75. Epub 2005/09/06.

[68]   Heath CM, Windsor M, Wileman T. Aggresomes resemble sites specialized for virus assembly. J Cell Biol. 2001;153(3):449-55. Epub 2001/05/02.

[69]   Kolokoltsov AA, Saeed MF, Freiberg AN, Holbrook MR, Davey RA. Identification of novel cellular targets for therapeutic intervention against Ebola virus infection by siRNA screening. Drug Dev Res. 2009;70(4):255-65. Epub 2009/06/01.

[70]   de Chassey B, Meyniel-Schicklin L, Aublin-Gex A, Andre P, Lotteau V. New horizons for antiviral drug discovery from virus-host protein interaction networks. Curr Opin Virol.2(5):606-13. Epub 2012/10/03.

[71]   Watanabe T, Watanabe S, Kawaoka Y. Cellular networks involved in the influenza virus life cycle. Cell Host Microbe. 2010;7(6):427-39. Epub 2010/06/15.

[72]   Shaw ML. The host interactome of influenza virus presents new potential targets for antiviral drugs. Rev Med Virol. 2011;21(6):358-69. Epub 2011/08/09.

[73]   Meliopoulos VA, Andersen LE, Birrer KF, Simpson KJ, Lowenthal JW, Bean AG, *et al*. Host gene targets for novel influenza therapies elucidated by high-throughput RNA interference screens. FASEB J. 2012;26(4):1372-86. Epub 2012/01/17.

[74]   Bitko V, Musiyenko A, Shulyayeva O, Barik S. Inhibition of respiratory viruses by nasally administered siRNA. Nat Med. 2005;11(1):50-5. Epub 2004/12/28.

[75]   Ge Q, Filip L, Bai A, Nguyen T, Eisen HN, Chen J. Inhibition of influenza virus production in virus-infected mice by RNA interference. Proc Natl Acad Sci U S A. 2004;101(23):8676-81. Epub 2004/06/03.

[76]   Rozema DB, Lewis DL, Wakefield DH, Wong SC, Klein JJ, Roesch PL, *et al*. Dynamic PolyConjugates for targeted *in vivo* delivery of siRNA to hepatocytes. Proc Natl Acad Sci U S A. 2007;104(32):12982-7. Epub 2007/07/27.

[77]   Hinton TM, Guerrero-Sanchez C, Graham JE, Le T, Muir BW, Shi S, *et al*. The effect of RAFT-derived cationic block copolymer structure on gene silencing efficiency. Biomaterials.33(30):7631-42. Epub 2012/07/27.

[78]   Tompkins SM, Lo CY, Tumpey TM, Epstein SL. Protection against lethal influenza virus challenge by RNA interference *in vivo*. Proc Natl Acad Sci U S A. 2004;101(23):8682-6. Epub 2004/06/03.

[79]   Ivacik D, Ely A, Arbuthnot P. Countering hepatitis B virus infection using RNAi: how far are we from the clinic? Rev Med Virol.21(6):383-96. Epub 2011/09/14.

[80]   Morrissey DV, Lockridge JA, Shaw L, Blanchard K, Jensen K, Breen W, *et al*. Potent and persistent *in vivo* anti-HBV activity of chemically modified siRNAs. Nat Biotechnol. 2005;23(8):1002-7. Epub 2005/07/26.

[81]   Kim SI, Shin D, Choi TH, Lee JC, Cheon GJ, Kim KY, *et al*. Systemic and specific delivery of small interfering RNAs to the liver mediated by apolipoprotein A-I. Mol Ther. 2007;15(6):1145-52. Epub 2007/04/19.

[82]   Carmona S, Jorgensen MR, Kolli S, Crowther C, Salazar FH, Marion PL, *et al*. Controlling HBV replication *in vivo* by intravenous administration of triggered PEGylated siRNA-nanoparticles. Mol Pharm. 2009;6(3):706-17. Epub 2009/01/23.

[83]   Kim SI, Shin D, Lee H, Ahn BY, Yoon Y, Kim M. Targeted delivery of siRNA against hepatitis C virus by apolipoprotein A-I-bound cationic liposomes. J Hepatol. 2009;50(3):479-88. Epub 2009/01/22.

[84]   Behlke MA. Progress towards *in vivo* use of siRNAs. Mol Ther. 2006;13(4):644-70. Epub 2006/02/17.

[85]   Guzman-Villanueva D, El-Sherbiny IM, Herrera-Ruiz D, Vlassov AV, Smyth HD. Formulation approaches to short interfering RNA and MicroRNA: Challenges and implications. J Pharm Sci.101(11):4046-66. Epub 2012/08/29.

[86]   Grimm D, Kay MA. RNAi and gene therapy: a mutual attraction. Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program. 2007:473-81. Epub 2007/11/21.

[87]   Sliva K, Schnierle BS. Selective gene silencing by viral delivery of short hairpin RNA. Virol J. 2010;7:248. Epub 2010/09/23.

[88]   Mowa MB, Crowther C, Arbuthnot P. Therapeutic potential of adenoviral vectors for delivery of expressed RNAi activators. Expert opinion on drug delivery. 2010;7(12):1373-85. Epub 2010/11/16.

[89]   Shen W, Wang CY, Wang XH, Fu ZX. Oncolytic adenovirus mediated Survivin knockdown by RNA interference suppresses human colorectal carcinoma growth *in vitro* and *in vivo*. J Exp Clin Cancer Res. 2009;28:81. Epub 2009/06/17.

[90]   Towne C, Aebischer P. Lentiviral and adeno-associated vector-based therapy for motor neuron disease through RNAi. Methods Mol Biol. 2009;555:87-108. Epub 2009/06/06.

[91]   Varble A, Chua MA, Perez JT, Manicassamy B, Garcia-Sastre A, tenOever BR. Engineered RNA viral synthesis of microRNAs. Proc Natl Acad Sci U S A. 2010;107(25):11519-24. Epub 2010/06/11.

[92]   Bell EJ, Brickell PM. Replication-competent retroviral vectors for expressing genes in avian cells *in vitro* and *in vivo*. Molecular biotechnology. 1997;7(3):289-98. Epub 1997/06/01.

[93]   Mayr U, von Werder A, Seidler B, Reindl W, Bajbouj M, Schmid RM, *et al*. RCAS-mediated retroviral gene delivery: a versatile tool for the study of gene function in a mouse model of pancreatic cancer. Hum Gene Ther. 2008;19(9):896-906. Epub 2008/09/02.

[94]   Goncalves MA. Adeno-associated virus: from defective virus to effective vector. Virol J. 2005;2:43. Epub 2005/05/10.

[95]   Varble A, Chua MA, Perez JT, Manicassamy B, Garcia-Sastre A, tenOever BR. Engineered RNA viral synthesis of microRNAs. Proc Natl Acad Sci U S A.107(25):11519-24. Epub 2010/06/11.

[96]   Hong CS, Goins WF, Goss JR, Burton EA, Glorioso JC. Herpes simplex virus RNAi and neprilysin gene transfer vectors reduce accumulation of Alzheimer's disease-related amyloid-beta peptide *in vivo*. Gene Ther. 2006;13(14):1068-79. Epub 2006/03/17.

[97]   Xiang S, Fruehauf J, Li CJ. Short hairpin RNA-expressing bacteria elicit RNA interference in mammals. Nat Biotechnol. 2006;24(6):697-702.

[98]   DeVincenzo J, Lambkin-Williams R, Wilkinson T, Cehelsky J, Nochur S, Walsh E, *et al*. A randomized, double-blind, placebo-controlled study of an RNAi-based therapy directed against respiratory syncytial virus. Proc Natl Acad Sci U S A.107(19):8800-5. Epub 2010/04/28.

[99]   Alnylam Announces Top-Line Results of Phase IIb Clinical Trial of ALN-RSV01, an Inhaled RNAi Therapeutic for the Treatment of Respiratory Syncytial Virus (RSV) Infection. http://phxcorporate-irnet/phoenixzhtml?c=148005&p=irol-newsArticle2&ID=1700544&highlight=.

[100]   Fahrenkrug SC, Blake A, Carlson DF, Doran T, Van Eenennaam A, Faber D, *et al*. Precision genetics for complex objectives in animal agriculture. J Anim Sci.88(7):2530-9. Epub 2010/03/17.

[101]   Piedrahita JA, Olby N. Perspectives on transgenic livestock in agriculture and biomedicine: an update. Reprod Fertil Dev.23(1):56-63. Epub 2011/03/04.

[102]   Kues WA, Niemann H. Advances in farm animal transgenesis. Prev Vet Med.102(2):146-56. Epub 2011/05/24.

[103]   Dunham RA. Transgenic fish resistant to infectious diseases, their risk and prevention of escape into the environment and future candidate genes for disease transgene manipulation. Comp Immunol Microbiol Infect Dis. 2009;32(2):139-61. Epub 2008/02/06.

[104]   Clark J, Whitelaw B. A future for transgenic livestock. Nat Rev Genet. 2003;4(10):825-33. Epub 2003/10/04.

[105]   Chen J, Chen SC, Stern P, Scott BB, Lois C. Genetic strategy to prevent influenza virus infections in animals. J Infect Dis. 2008;197 Suppl 1:S25-8. Epub 2008/03/28.

[106]   Luo J, Du J, Gao S, Zhang G, Sun J, Cong G, *et al*. Lentviral-mediated RNAi to inhibit target gene expression of the porcine integrin alphav subunit, the FMDV receptor, and against FMDV infection in PK-15 cells. Virol J.8:428. Epub 2011/09/09.

[107]   O'Neill G. Australia tackles bird flu using RNAi. Nat Biotechnol. 2007;25(6):605-6. Epub 2007/06/09.

[108] Wang H, Wu J, Liu X, He H, Ding F, Yang H, *et al*. Identification of short hairpin RNA targeting foot-and-mouth disease virus with transgenic bovine fetal epithelium cells. PLoS One.7(8):e42356. Epub 2012/08/21.

[109] Clarke BD, Cummins DM, McColl KA, Ward AC, Doran TJ. Characterization of Zebrafish Polymerase III Promoters for the Expression of Short-Hairpin RNA Interference Molecules. Zebrafish. Epub 2012/10/04.

[110] Dieckhoff B, Petersen B, Kues WA, Kurth R, Niemann H, Denner J. Knockdown of porcine endogenous retrovirus (PERV) expression by PERV-specific shRNA in transgenic pigs. Xenotransplantation. 2008;15(1):36-45. Epub 2008/03/13.

[111] Taxman DJ, Livingstone LR, Zhang J, Conti BJ, Iocca HA, Williams KL, *et al*. Criteria for effective design, construction, and gene knockdown by shRNA vectors. BMC Biotechnol. 2006;6:7. Epub 2006/01/26.

[112] McIntyre GJ, Yu YH, Lomas M, Fanning GC. The effects of stem length and core placement on shRNA activity. BMC Mol Biol.12:34. Epub 2011/08/09.

[113] Boden D, Pusch O, Silbermann R, Lee F, Tucker L, Ramratnam B. Enhanced gene silencing of HIV-1 specific siRNA using microRNA designed hairpins. Nucleic Acids Res. 2004;32(3):1154-8. Epub 2004/02/18.

[114] Lebbink RJ, Lowe M, Chan T, Khine H, Wang X, McManus MT. Polymerase II promoter strength determines efficacy of microRNA adapted shRNAs. PLoS One.6(10):e26213. Epub 2011/10/28.

[115] Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, *et al*. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. Nature. 2006;441(7092):537-41. Epub 2006/05/26.

Host-Encoded miRNAs Involved in Host-Pathogen Interactions

Samantha Barichievy[1] and Abhijeet Bakre[2,*]

[1]*Gene Expression and Biophysics Group, Synthetic Biology–Emerging Research Area, Council for Scientific and Industrial Research, Pretoria, South Africa and* [2]*University of Georgia, College of Veterinary Medicine, Department of Infectious Diseases, 111 Carlton Street, Athens, GA 30602, USA*

Abstract: Humans display a remarkably diverse susceptibility to infection, the foundation of which lies in our genetic variation and ability to respond to selective pressures applied by various infectious agents. The evolution of our complex and multi-player immune system underlines the dominance of the human host following a microbial infection. However, given the nature of obligate intracellular pathogens, their complete reliance on host gene expression machinery has led to the evolution of complex interplays between the two, such that pathogens actively and strategically maneuver their way through the host terrain. Our traditional view of this terrain as being comprised of protein-coding genes, translation intermediates (mRNAs) and protein counterparts is far too simplistic, particularly in the context of infection. The discovery of the RNA interference (RNAi) pathway has greatly enhanced our understanding of the host terrain. Small noncoding RNAs (ncRNAs) termed microRNAs (miRNAs) were shown to be key regulators of gene expression that function within the RNAi pathway to post-transcriptionally modulate mRNA stability and subsequent translation [1]. Indeed, it is now understood that miRNAs are able to rapidly, and with exquisite specificity, modulate gene expression in response to numerous environmental cues in a highly coordinated, complex and tissue-specific manner. Given the reliance of intracellular pathogens on host gene expression machinery, the RNAi pathway, and specifically miRNAs, are now understood to lie at the nexus of the host-pathogen interplay. The focus of this chapter will be on the characteristics and roles of these small noncoding RNAs in host-pathogens interactions.

*Corresponding author Abhijeet Bakre: University of Georgia, College of Veterinary Medicine, Department of Infectious Diseases, 111 Carlton Street, Athens, GA 30602, USA; Tel: 706-542-2205; Fax: 706-583-0176; E-mail: bakre@uga.edu

AN INTRODUCTION TO miRNAs

miRNAs in the Noncoding RNA Space

Only 2% of the metazoan genome encodes protein, yet more than 50% is transcribed and we have little knowledge regarding these transcripts that function in the absence of protein production. In fact, stable ncRNA transcripts have been referred to as 'dark matter' within the cellular environment [2]. Despite improvements in the human draft genome sequence, ncRNAs remain difficult to define and thus quantify [3, 4]. However, numerous evolutionary studies have revealed that ncRNAs are estimated to be expressed at 4-fold excess compared to their protein-coding counterparts and are highly conserved across eukaryotic genomes [3]. Currently, ncRNAs have been systematically organized (according to length) into long ncRNAs (lncRNAs; > 200 bp) or small ncRNAs (< 200 bp) and make up greater than twenty-six functional categories that reflect the breadth of ncRNA function and diversity [3, 5]. In the case of humans, lncRNAs such as autonomous and non-autonomous retrotransposons, retrovirus-like elements, and DNA transposon fossils comprise about 48% of the genome [6]. Small ncRNAs including miRNAs, repeat associated small interfering RNA (rasiRNAs) and piwi associated RNAs (piRNAs) constitute a small but significant portion of the genome. While new ncRNAs are rapidly being uncovered, functional data remains sparse particularly at the host-pathogen interface. Analysis of miRNA diversity in the animal kingdom revealed that miRNAs co-evolved with eukaryotic genomes [7, 8] with less than 12 miRNA genes lost during the entire deuterostome evolution, and significant miRNA gene expansion at major evolutionary nodes. Ancestral miRNAs were typically located in inter-genic regions but have since adapted to colonize the introns of eukaryotic genomes (intronic exaptation) presumably to overcome dependence on host transcription machinery and promote generation of novel miRNAs [9]. About 50% of human miRNA genes are located within introns [9], while the remainder are found in either intergenic regions, coding regions or untranslated regions (UTRs). miRNAs may also be produced through the processing of structural RNAs such as tRNAs [10], snoRNAs [11] or transposable element encoded direct and indirect repeats [12].

miRNA Biogenesis

Endogenous miRNAs are usually transcribed from RNA polymerase II promoters [13] as primary miRNA transcripts (pri-miRNAs) several kb in length, containing various stem-loop structures and ssRNA flanking segments. Pri-miRNAs are processed in two compartmentalized steps *via* the actions of distinct protein complexes (Fig. **1** and reference [14]). The initial step involves an RNAse-III

protein family member, Drosha, which cleaves at the base of the stem of the pri-miRNA to release a hairpin structure termed a pre-miRNA [15, 16]. The DiGeorge syndrome critical region gene 8 (DGCR8) protein co-interacts with Drosha and the pri-miRNA to form a 650 kiloDalton (kDA) Microprocessor complex [15, 17]. DGCR8 specifically recognizes the junction between the single stranded (ss) and dsRNA regions of the pri-miRNA stem and directs Drosha-mediated cleavage 11 bp from this site [18]. The resultant pre-miRNA is approximately 70 bp in length and includes a 5' phosphate and two-nucleotide 3' overhang. Interestingly, pri-miRNA processing may occur co-transcriptionally and prior to intronic splicing in the nucleus [19]. Drosha-mediated cleavage of intronic miRNAs does not impair transcript splicing [20] but if the miRNA is located exonically, Drosha processing destabilizes the transcript and affects downstream protein translation [21]. Cellular splicing machinery also plays a role in the biogenesis of a distinct class of miRNAs that bypass Drosha-mediated cleavage [22, 23]. Termed mirtrons, these intronic miRNAs are derived from Pol II-transcribed primary-mirtron precursors that encode canonical splice sites (Fig. **1**). Host splicing machinery recognizes the splice sites and cleaves the transcript generating a hairpin configuration that resembles a Drosha product [24].

Following nuclear cleavage, pre-miRNAs are conveyed to the cytoplasm *via* exportin-5 (Exp5) in a RanGTP-dependent manner [25]. Upon entering the cytoplasm a second RNAse III enzyme, Dicer, cleaves the loop of the pre-miRNA and releases a miRNA duplex. Dicer is comprised of two copies of a conserved RNAse III domain, a dsRBD in the carboxy-terminus, an amino-terminus helicase domain, a domain of unknown function (DUF) and a PAZ (Piwi-Argonaute-Zwille) domain [26]. The gap between the PAZ and RNAse III domains corresponds with the average length (25 bp) of a pre-miRNA stem duplex. Dicer thus functions as a 'molecular ruler' [27] by binding the 3' end of the pre-miRNA duplex in the PAZ domain and cutting at a set distance to generate miRNA products 21 bp in length with characteristic 5' phosphate groups and two-nucleotide 3' overhangs [28, 29]. Dicer interacts with TRBP (TAR RNA binding protein) [30]. PACT (Protein Kinase R activator) [31] and a member of the Argonaute (AGO) family to create the RNA-induced silencing complex (RISC) that is required for RNA interference [32-35]. One strand of the duplex remains bound within RISC as the mature miRNA or 'guide strand' while the remaining passenger strand is degraded [36]. The relative thermodynamic stability of the 5' ends of the miRNA duplex determines which strand is selected as the mature miRNA [37].

Mature miRNAs regulate post-transcriptional protein synthesis by base pairing to cognate mRNAs. Depending on the level of complementarity between the mature miRNA and its mRNA target sequence, the associated protein payload, as well as the specific AGO protein involved in RISC loading, multiple mRNA silencing mechanisms can occur including enhanced exonucleolytic mRNA decay, site-specific endonucleolytic cleavage or translational repression [38, 39]. The initial bases from positions 2 to 7 of the mature miRNA are termed the 'seed' sequence and they provide most of the pairing specificity. In some cases, complete pairing between the seed region and target mRNA is sufficient to mediate AGO2-associated RISC cleavage of the cognate phosphodiester backbone [40, 41]. More typically for mammalian and viral mRNA targets however, cleavage activity of RISC is severely impaired by mismatched pairing in the seed and other regions and translational inhibition occurs [42, 43]. In this pathway, RISC-bound mRNAs associate with cytoplasmic processing bodies (P-bodies) that exclude translational machinery and incorporate proteins required for mRNA remodeling [44]. All four human AGO proteins bind miRNAs, although endonuclease activity is a property of AGO2 alone [40]. Thus RISC complexes containing any of the four AGO proteins can catalyze a translational block but only complexes containing AGO2 can trigger transcript decay in the case of perfect complementarity between a miRNA and its target mRNA [45]. Proteins of the TNRC6A (trinucleotide repeat containing 6A; also known as the GW182 protein family) are also essential for transcript decay activity in the above situation [46]. Intriguingly, since the complementary length of seed sequence required for miRNAs to target cognate mRNAs is short, each miRNA can target and modulate hundreds of transcripts. Indeed, current estimates predict that thousands of human transcripts are regulated by miRNAs [47-49]. Furthermore, a single miRNA can regulate multiple mRNA molecules that can in turn also be acted upon by numerous distinct miRNAs [50, 51]. Importantly, most miRNAs decrease target protein levels by less than 2-fold [52], but this non-linear tuning mechanism can still exert a large physiological effect [53]. Thus, the endogenous miRNA pathway represents a highly efficient system to simultaneously fine-tune the expression of numerous genes as well as modulate specific functional pathways.

miRNAs as Part of the Host Innate Immune Response

Animals have evolved a cellular and organismal response to environmental challenges including infections with intra- and extra-cellular pathogens. The primary response to such a challenge is the innate immune response which is non-specific and functions to identify sites of infection and contain them. The innate
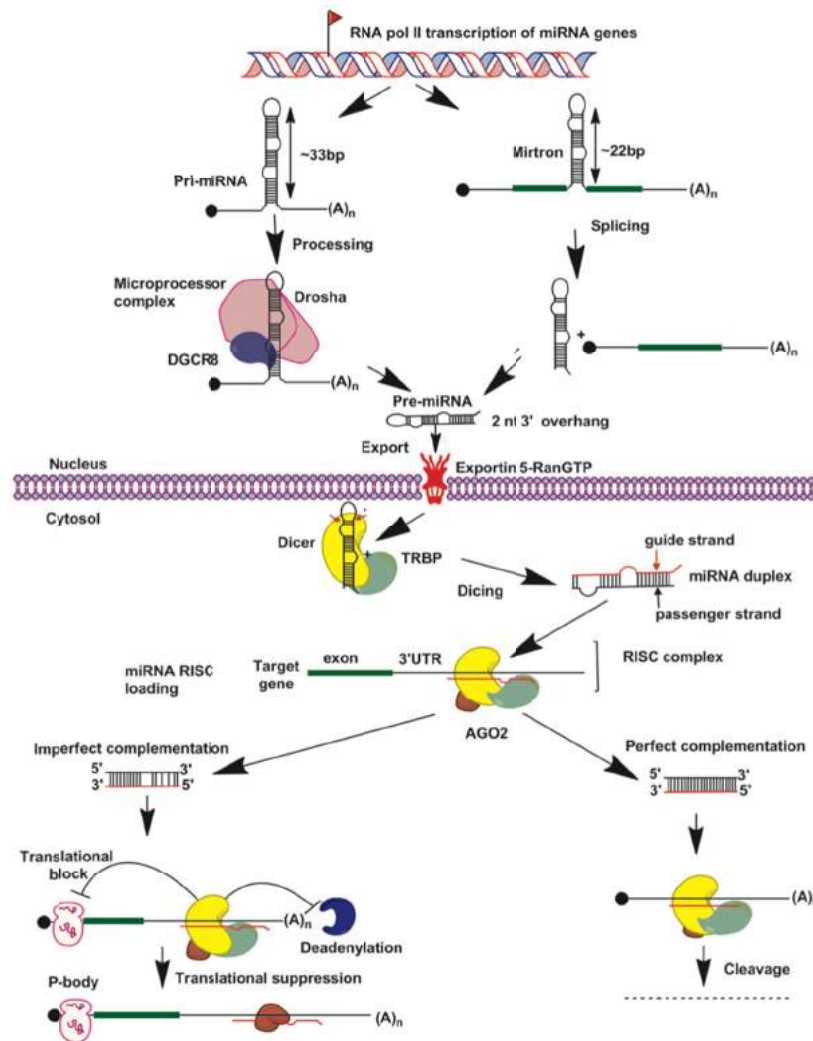
Figure 1: The mammalian RNAi pathway. Endogenous miRNAs are transcribed from RNA Pol II promoters to generate pri-miRNAs with a stem-loop structure approximately 33 bp in length flanked by single stranded regions. The RNAse III protein Drosha complexed to DGCR8 forms a Microprocessor complex that binds the pri-miRNA and cleaves to release a pre-miRNA hairpin structure with characteristic 2 nucleotide 3' overhangs. Pre-miRNAs can also be generated without Drosha from intronic sequences termed mirtrons that contain splice donor and acceptor sites recognized by host splicing machinery. Both forms of pre-miRNA are exported from the nucleus by Exp5, which recognizes and binds the 2 nucleotide 3' overhang. In the cytoplasm, the loop of the pre-miRNA is cleaved by Dicer and its dsRBD partner TRBP to liberate a miRNA duplex with 2 nucleotide 3' overhangs. AGO2 binds the miRNA duplex to form RISC wherein the miRNA duplex is unwound leaving the guide strand (mature mRNA) to pair covalently with its cognate mRNA. If this pairing is incompletely complementary, translational suppression occurs, usually in P-bodies. If the mature miRNA pairs 100 % with the target mRNA, AGO2- mediated cleavage of the mRNA occurs.

immune system primes a pathogen-specific potent adaptive response which functions to 1) eliminate the infection, 2) generate a memory response that can respond to future challenges from the same pathogen, and 3) restore immune homeostasis. Innate immunity depends on the identification of specific pathogen associated molecular patterns (PAMPs) by cellular pattern recognition receptors (PRRs) that lead to induction of specific cellular responses. Toll like receptors (TLRs), Retinoic acid inducible gene (RIG-I/DDX-58) like receptors (RLRs), Nucleotide oligomerization domain (NOD) like receptors (NLRs) and Scavenger receptors form the four main families of PRRs that are located either on the cell surface, internal membranes or secreted into extracellular milieu. TLRs and RLRs form the major PRR class for the recognition of intra-cellular pathogens such as viruses. Recent studies suggest that miRNAs are also integral components of the innate immune response and both the host cell and the pathogen can utilize miRNA-mediated gene regulation toward their own survival and propagation.

Multiple studies that analyzed the expression of the human miRnome in response to viral, bacterial and protozoal infection identified subsets of host miRNAs that are differentially expressed [54-60]. While mechanisms and target gene repertoires regulated by these miRNAs are yet to be fully elucidated, evidence from Drosha and Dicer ablation studies suggest that components throughout the entire miRNA pathway participate in innate and adaptive immune responses, as well as cellular development. For example, Drosha silencing significantly diminishes the capacity of endothelial cells to form tubes and sprout capillaries [61]. Drosha silencing also alters normal development and egg hatching in the nematode *Meloidogyne incognita* [62], inhibits cellular proliferation in *Drosophila melanogaster* [63], and is essential for survival of vascular smooth muscle cells (VSMCs) [64]. Drosha-generated small RNAs (DDRNAs) were also recently shown to be essential for the DNA damage response [65]. Similarly, Dicer knock-out (KO) and knock-down (KD) studies showed that T regulatory (Treg) cell development and function is regulated *via* miRNA-mediated modulation of the crucial transcription factor Foxp3 [66], and Dicer-deficient Treg cells lack suppressor function [67]. Dicer is also important for the development of invariant natural killer T (iNKT) cells [68], and bone marrow derived macrophages (BMDMs) from Dicer KO animals show increased expression of inflammatory cytokines [69]. Furthermore, Dicer-2 (dcr2) regulates the citric acid cycle, mitochondrial oxidative phosphorylation and energy metabolism in fruit flies [70].

Changes in Dicer expression levels have also been linked to altered immune responses to various pathogenic infections. Dicer expression is significantly reduced in clinical cases of Respiratory Syncytial Virus (RSV) [71] and controls Hepatitis C Virus (HCV) RNA accumulation *via* regulating the abundance of miR-122 specifically [72]. Respiratory epithelial cells with reduced Dicer exhibit increased susceptibility to influenza virus infection [73] and protozoan parasite *Leishmania donovani* down-regulates Dicer 1 to regulate miR-122 [74]. *Listeria monocytogenes* and the human vaccine strain used against tuberculosis infection, *Mycobacterium bovis* bacillus Calmette-Guérin (BCG), both decrease miR-29, which controls interferon gamma (IFN- Υ) in natural killer (NK) and T cells [75]. *Salmonella*, an invasive bacterium, down-regulates members of the Let-7 family of miRNAs causing de-repression of target interleukin 10 (IL-10) mRNAs which attenuate pro-inflammatory cytokines [76]. Similarly, bacterial lipopolysaccharides (LPS) stimulate miR-21 expression leading to up-regulation of IL10 mRNAs and subsequent attenuation of inflammatory activation [77]. Finally, mice with reduced Dicer expression in the gut are more susceptible to worm infestations [78] highlighting the central role of Dicer and cellular immunity within intestinal mucosa homeostasis.

Considering that human cells encode >1000 miRNA species, many of which function in innate immunity and apoptosis, it is unsurprising that pathogens (and viruses in particular) have evolved mechanisms to subvert these cellular components [79]. The particular mechanisms by which viruses manipulate the host immune system are as varied as the viruses themselves but if one focuses on viral interactions with cellular microRNA machinery, the options are surprisingly minimized and constrained to a fairly limited number of human viruses. However, the above observations strongly support the role of the miRNAs and proteins involved in their biogenesis in regulating crucial aspects of host responses to environmental stimuli including infection. Furthermore, as miRNAs are expressed in a very tissue-specific manner that can vary depending on the cell cycle stage, the interactions between host miRNAs and pathogens is clearly complex. Understanding which components are involved, how the networks are regulated and what tips the host-pathogen balance in favor of infection or not will shed valuable light on the role of miRNAs in the first-line of defense innate immune response.

DEVELOPMENT OF miRNA-RELATED TECHNOLOGIES

A host organism's ability to build an innate immune response against a pathogen is vital and many cellular mRNAs that control host defenses are regulated by

miRNAs. The promiscuity of miRNAs in regulating their mRNA targets coupled with their importance in post-transcriptional regulation of host gene expression make unraveling the role of miRNAs at the host-pathogen interface extremely challenging. Resolving these interactions requires identification of the specific pathogen-encoded stimuli that induce changes in the host miRnome following infection, assessment of which transcripts are targeted by miRNAs as well as which miRNAs are responsible, quantification of the miRNA-induced changes to the infection transcriptome, analysis of downstream effects on related protein outputs, and validation of each step to ensure a robust understanding of such a complex network of interactions. To achieve these goals with miRNAs, a range of strategies and reagents have been developed over the course of the last decade.

Identification and Detection

The ever-expanding list of miRNAs is comprised of both predicted and validated sequences. As might be expected, the number of predicted sequences is exhausting and primarily generated through the analysis of large quantities of information derived from deep sequencing efforts. Unsurprisingly, only a fraction of the predicted miRNAs has been validated with experimental data [80]. This discrepancy is primarily due to the fact that complex genomes contain large numbers of inverted repeats capable of folding into miR-like hairpin structures [81]. Refinement of the list of predicted miRNA genes has generally involved incorporation of filters that eliminate well-characterized recurring sequences (*e.g.,* long interspersed nuclear elements (LINEs) and alu repeats). Additionally, researchers have incorporated computational techniques that identify defined nucleotide content associated with stem length or Drosha cleavage sites, and evolutionary conservation [82-88]. Approaches have led to the development of a host of public and private bioinformatic tools for computational identification of miRNA sequences including, but not limited to, CID-miRNA (http://mirna.jnu.ac.in/cidmirna/), miRScan (http://genes.mit.edu/mirscan/), and miRFinder (http://www.bioinformatics.org/mirfinder/).

In addition to bioinformatic approaches, there are a number of technologies that have been adopted for the discovery and validation of miRNA sequences. In general, the overall usefulness of each is judged on the basis of specificity, sensitivity, and throughput. Early on, extraction of total RNA followed by Northern Blot Analysis [89] led to quick determination of the presence of a predicted sequence. Yet due to the relatively labour intensive nature of the technique, the need for significant amounts of starting material, the comparatively poor sensitivity, and overall low throughput, alternative technologies and services

have quickly replaced this approach. One such technology that provides a fairly global picture of a cell's miRNA signature is based on microarray gene expression profiling [90]. Gene expression chips containing the entire complement of an organism's known or predicted miRNA content can be probed with labelled sequences generated from as little as 50-100 nanograms of total RNA. Current chip designers (*e.g.* Agilent and Affymetrix) and service providers (*e.g.*, Asuragen and Exiqon) update miRNA content on a regular basis using frequent miRBase releases, thereby offering researchers the ability to rapidly obtain a cell's miRNA profile. Not surprisingly, the greatest limitations associated with array-based profiling relates to issues surrounding specificity and sensitivity. Since arrays are based on hybridization, cross-hybridization of sequences to closely related probes (*e.g.*, miRNA family members) can lead to false positives. Given the additional limitation of arrays to accurately detect i) poorly expressed sequences, or ii) small changes in miRNA expression, the overall effectiveness of the technology may be limited to a subset of a cell's entire miRnome.

The most recent advances in miRNA quantitation utilize qRT-PCR and Next Generation Sequencing (NGS). For qRT-PCR, several companies in the research tool space (including Life Technologies, Qiagen, and Exiqon) have developed dedicated kits to facilitate quantitation of miRNAs. In most cases, workflows combine novel systems for small RNA purification with unique amplification strategies (*e.g.* adaptor ligation) specifically designed for quantitation of this of this small collection of uncommon targets. Assays can be performed on as little as 1 picogram of RNA and in many cases allows for the distinction between closely related family members. Most recently, sequencing has become the gold standard for miRNA quantitation. As with qRT-PCR, the sequencing workflow begins with small RNA purification and adaptor ligation followed by RT and PCR amplification. Subsequent sequencing using any number of available platforms is greatly simplified by the fact that miRNAs are generally shorter than 25 nucleotides in length, thus eliminating the need for sequence alignment. As such, the relative signature or profile of a cell can be generated from the overall abundance of individual sequences within the total population of reads.

Modulation

Modulation of endogenous genes often provides essential clues into an entitie's function or role in cell biology. To that end, both academic and industrial research groups have focused much of their effort on developing collections of miRNA mimics and inhibitors to mediate up- and down-regulation of miRNA concentrations (respectively). To date, miRNA mimics are available in multiple

forms including synthetic and expressed constructs. Synthetic miRNA mimics such as those present in the Dharmacon miRIDIAN collection closely resemble native mature miRNA sequences with the exception that duplexes contain chemical modifications that bias strand entry into RISC. Expressed constructs, such as the shMIMIC collection (Dharmacon, part of GE Healthcare) embed the DNA encoding the desired mature miRNA sequence into a lentiviral backbone. Subsequent packaging of the construct into a viral particle and transduction into target cells leads to integration of the expression construct into the host genome and consistent, long-term expression of the desired miRNA mimic sequence.

Synthetic miRNA inhibitors have also been developed and successfully used. Early inhibitor designs described by Ebert [91] included a collection of molecules comprised of multiple, tandem (antisense) miRNA binding sites. Referred to as "sponges", the authors successfully demonstrated that this collection of competitive inhibitors could efficiently suppress miRNA function in a seed-specific fashion. Vermeulen and colleagues [92] have described additional inhibitor designs. In contrast to the molecular sponges described by Ebert, Vermeulen successfully combined a single miRNA reverse complement sequence with unrelated (flanking) secondary structural elements. Through an unknown mechanism, this combination of traits significantly increased overall inhibitor potency and allowed for multi-miRNA knockdown at sub-nanomolar inhibitor concentrations.

While experiments employing mimic and inhibitor strategies have successfully broadened our understanding of the contribution miRNAs make to cell biology (including host-pathogen interactions) there are caveats to working with these reagents. As described in an article by Robertson [93] inhibitors designed against one member of a miRNA family can silence related family members. This cross-reactivity can significantly confound interpretation of results particularly as researchers move downstream in their attempts to identify relevant miRNA targets. Similar issues of cross-reactivity may complicate the interpretation of results that employ miRNA mimics. As noted by several researchers, the intracellular concentrations of mimics resulting from transfection are predicted to far exceed endogenous concentrations [94]. If non-physiological concentrations of miRNAs result in the silencing of endogenous mRNAs that under normal circumstances would not be targeted, the risk of false positive phenotypes may be heightened. Considering events upstream of actual miRNA targeting, it is conceivable that an over-abundance of a miRNA mimic (particularly those expressed as pri-miRNAs) may saturate one or more steps of the RNAi pathway.

If this sort of over-saturation hinders endogenous miRNA biogenesis, false positive phenotypes may be observed.

Target Identification

As is the case with miRNA gene identification, tools for identifying miRNA targets include both *in silico* and benchtop resources. Computational target prediction began with a large body of work from the Bartel laboratory [95, 96]. Predictions primarily focus on the identification of complementary miRNA seed matches (nucleotides 2-7) in the 3' UTR of target mRNA. By combining seed-related parameters with constraints that demand the presence of position-specific nucleotides [48], evolutionary preservation of target sites, and target mRNA structure [97], researchers have significantly improved the target prediction field. However, while these parameters have been incorporated into multiple search programs including miRBase [98] and miRanda [99], it is acknowledged that the number of mRNA targets identified by these methods exceeds the true target number for any individual miRNA [100]. These shortcomings have led researchers to combine a range of experimental approaches with *in silico* predictions. These include differential gene expression profiling using microarrays, sequencing following treatment of cells with miRNA mimics and/or inhibitors [49, 101], identification of miRNA-mRNA target pairs through immunoprecipitation of RISC [102], mimic target validation using 3' UTR-reporter constructs [103] well as other techniques that focus on correlating miRNA expression with a particular transcript(s).

mIRNAs INVOLVED IN INFECTIOUS DISEASE

While there are a large number of studies that have analyzed miRNA deregulation during infection, studies involving viruses are predominant owing to the intra-cellular nature of the pathogen. For this reason, we limit our summary of miRNA deregulation during infectious diseases to those caused by viral pathogens in order to highlight the unique features of miRNA regulation at the host-virus interface.

Hepatitis C Virus

Hepatitis C virus (HCV) infects about 17 thousand people annually in the US and an estimated 3.9 million people live with HCV infection [104]. Though anti-HCV treatment is effective, lack of HCV diagnosis and effective care prevent effective HCV containment. A significant proportion of HCV-infected people (>85%) are chronically infected of which ~50% develop chronic liver disease (CLD) [105, 106].

HCV is a positive strand RNA virus. The entire genome of HCV is translated into a single viral polyprotein that is subsequently cleaved to generate the core protein, two envelope proteins E1 and E2, ion-channel protein p7, and five non-structural proteins NS1, NS2, NS3, NS4A, NS4B, NS5A, and NS5B.

Infection of Huh-7 cells with three different HCV clones was found to induce the expression of miR-142-3p and downregulate miR-128a and -196a [107]. Expression of these three miRNAs was inversely correlated with the expression profiles of 37 genes deregulated during HCV infection of A549 cells and 4 genes (HNMT, XPO1, PMPCB and HMGB1) were subsequently validated. Similarly, HCV chronically infected peripheral monocytes (cHCV) showed an elevation in miR-155 expression (a positive regulator of multiple cytokines including TNF-α [108]) and downregulation of miR-125b. Naïve monocytes induced miR-155 significantly upon treatment with HCV core or NS3 and NS5 proteins, and this miRNA was found to be elevated in sera taken from cHCV patients [108]. HCV-induced miRNA miR-130a was independently shown to regulate the expression of IFITM1 [109], resulting in reduction of cellular IFITM1 expression. IFITM1 proteins are small transmembrane proteins and overexpression of IFITM1 has an anti-viral effect reducing HCV infectivity [110]. Recently it was shown that IFITM1 induced during HCV infection interacts with HCV co-receptor CD81 and occludin and inhibits viral entry [111].

miR-122 plays a vital function in HCV infection. miR-122 expression is negatively regulated by HCV core protein leading to decreased miR-122 at later stages of infection [112]. miR-122 expression has been shown to increase HCV RNA concentration [113]. It was also shown that miR-122 inhibition abrogated HCV replication but did not alter viral RNA translation or RNA stability. miR-122-mediated upregulation of HCV replication is affected by the abundance of Dicer and TRBP [72] and TRBP was found to be essential for HCV RNA accumulation [72]. Additionally, AGO1 and AGO2 proteins play critical roles in recruitment of miR-122 to the HCV 5' UTR but not the 3'UTR [114], and these proteins delay HCV RNA degradation. Indeed miR-122 prevents HCV RNA degradation by cytosolic 5' exonuclease Xrn1 [115]. Current studies suggest that miR-122-HCV complexes form tertiary RNA structures that recruit other cellular proteins to decrease HCV transcript decay and promote translation [116]. Similar studies with GB virus B (GBV-B), another (+) ssRNA virus which infects chimpanzees, also validated the interaction between HCV 5' UTR and miR-122 and the formation of a tertiary structural feature [117]. Later it was demonstrated that ectopic expression of miR-122 in non-hepatic cells that do not express HCV

receptors not only confers susceptibility to HCV infection but also enables completion of the viral life cycle [118-121].

miR-122-HCV complex interaction utilizes sequences in the 5' non-coding region (NCR) of the HCV genome and spacing individual miR-122 binding sites also determines efficacy of miR-122 mediated HCV RNA upregulation [122]. This interaction is crucial to viral replication since *in vivo* suppression of miR-122 expression in mice, green monkeys [123] and chimpanzees [124] leads to a significant long-term reduction in serum cholesterol and reduced viremia. Indeed, recent human clinical trials with Miravirsen, a locked nucleic acid antisense oligonucleotide against miR-122, showed a long-term reduction of HCV RNA levels in chronically infected HCV patients without emergence of miR-122 viral escape mutants or adverse effects [125]. Recently it was shown that miR-122 also regulates the expression of suppressor of cytokine gene 1 (SOCS1) thus reducing the secretion of anti-viral IFN-β [126]. miRNA-mediated translational blocks occur in cytosolic P-bodies and HCV infection reduces the frequency of cytosolic P-bodies while inducing the formation of stress granules that are sites of active viral RNA synthesis [127].

In addition to miR-122, a number of other miRNAs have been shown to have a role in HCV biology. HCV-induced miR-21 targets two important mediators of the innate immune response, MyD88 and IRAK1, causing suppression of host innate immunity and mediating viral persistence during chronic infection [128]. miR-27a has been shown to regulate the metabolism of lipids in Huh cells infected with HCV by regulating the expression of transcription factor RXRα, and lipid metabolism associated genes FASN, SREBP1, SREBP2, PPARα, PPARγ, Apo1, ApoB100, and ApoE3. miR-27a inhibition increased cellular lipid production and reduced interferon signalling, while miR-27a upregulation increased IFN activity and efficacy of pegylated IFN and Ribavirin therapy against HCV [129]. This suggests that native miR-27a induction plays a role in establishment of a chronic, low viral load state. Finally, the HCV core protein also induced expression of miR-345 which regulates apoptosis by targeting p21 [130].

Human Immunodeficiency Virus

Despite the discovery of the human immunodeficiency virus (HIV-1) nearly three decades ago and the expanding use of potent combinatorial antiretroviral regimens, HIV-1 currently infects 34 million people globally [131]. A deeper understanding of how this pathogen manipulates human cells during infection is paramount to the development of more targeted therapies and an efficacious

vaccine. As an obligate intracellular pathogen, HIV-1 relies on host cellular machinery to complete its life cycle. Integral to this is the modulation of host gene expression to ensure a coordinated regulation of pro- and anti-viral host factors [94, 132, 133]. HIV-1 actively manipulates expression of some host miRNAs during infection although differences in pseudovirus transfection *vs*. molecular infectious clone infections, viral titres, viral subtypes, infection durations and cell types used make comparisons unreliable. As such, robust identification and characterization of host miRNAs that can modulate HIV expression in a natural infection setting, remains to be elucidated. Furthermore, it is unclear whether HIV-1 affects global miRNA expression by altering miRNA biogenesis or processing, or if the virus affects individual miRNAs through changes in transcription or miRNA maturation, or a combination of these. What is clear however is that HIV-1 infection, regardless of the model used, perturbs the cellular miRnome compared to uninfected control cells and thus continued research in this arena remains important.

HIV-1 and the miRNA Biogenesis Pathway

The ability of HIV-1 to affect global miRNA expression would not be uncommon among viruses. Plant and insect viruses in particular have been shown to encode proteins that act as suppressors of RNA silencing (SRS) [134-137], thought to have developed as a counter-strategy to the anti-viral effects of the RNAi pathways in these organisms. The human T-cell lymphotropic virus (HTLV), which encodes a SRS protein, Rex, was shown to interact with Dicer to suppress RNAi-mediated silencing of the virus in mammalian cells [138]. Similarly, the VP35 protein of Ebola was shown to suppress RNAi-mediated silencing [139]. The HIV-1 Tat protein, required for transactivation of viral transcription, may act as a SRS protein but the data related to this are conflicting. Exogenous Tat in HeLa cells [140], and viral-derived Tat in Jurkat and CEM-SS cells [141] is known to work with Dicer and function as a SRS protein. In contrast however, Tat failed to inhibit global RNAi in persistently infected T cells even when stably expressed and assessed at varying time points over the course of infection [142].

During infection, Tat binds to the TAR (*trans*-activation response) element, which is an RNA structure present in the 5' LTR of all HIV-1 transcripts. This interaction recruits cellular factors including TRBP (TAR RNA binding protein) and the P-TEFb complex (CDK9 and cyclinT1) resulting in phosphorylation of stalled RNA polymerase II at the 5' LTR. Importantly, TRBP is a co-factor required for Dicer processing of pre-miRNAs as well as RISC loading (Fig. **1**), and thus titration of this protein by the Tat/TAR interaction was suggested as a

mechanism for global RNAi repression [143]. Supporting data on the sequestration of TRBP by Tat remains to be independently validated, while contrasting data show that global suppression of RNAi does not occur in cells infected with replicating HIV-1 [144]. Another HIV-1 protein, Vpr (viral protein R) was suggested to act as a SRS by binding to Dicer in monocytes concomitant with a global decrease in miRNA production [145]. This study was complicated by the fact that monocytes generally express low levels of most miRNA biogenesis-related proteins, and that Dicer is usually only detectable once the cells differentiate into macrophages. Furthermore, these results remain to be independently validated.

Not only does the validity of HIV-1-encoded SRS proteins continue to be unresolved, the effect of the miRNA pathway on HIV-1 also remains uncertain. In one study using VSVG-pseudotyped pNL4-3Luc reporter vectors in HEK293T cells, HIV-1 infection was enhanced in cells with decreased Drosha and Dicer levels [146]. A second study showed that knockdown or either Drosha or Dicer in peripheral blood mononuclear cells (PBMCs) from HIV-1 infected donors resulted in faster replication kinetics of the virus [147]. In addition, the repression of Dicer and Drosha also negatively affected HIV replication in Jurkat cells as well as latently infected U1 cells [147]. An important caveat from both studies is that Drosha and Dicer knockdown mediated by exogenous siRNAs may severely complicate the interpretation of such experiments. Given the critical role of the RNAi pathway in normal cellular gene expression, it is unsurprising that knockdown of these proteins would affect HIV-1 replication in cells depleted of such factors. The ability to tease out a direct *vs*. off-target effect of depleted Drosha or Dicer on HIV-1 replication is very difficult and until solid, robust observations can be made, the effects of the miRNA pathway on HIV-1 replication remain obscure.

HIV-Mediated Perturbations of Host miRNA Profiles

The dysregulation of host miRNA profiles in response to HIV-1 infection has been examined in various cell lines as well as physiologically relevant primary cells although to date, only one quarter (~ 320) of the known human miRnome covering ~ 1200 miRNAs has been investigated. Overall, changes in the miRNA profiles do not correlate well between different studies but this is not surprising given the cell and tissue-specific expression of miRNAs in combination with varied experimental plans. In 2005, the first of such published studies compared changes in the miRNA profiles between uninfected HeLa cells and those transfected with pNL4-3 pseudovirus using miRNA microarrays that detected 312

miRNAs [148]. The dominant pattern showed HIV-mediated down-regulation of these host miRNAs, but this was unsupported in a later study using the same miRNA microarrays [147]. Eleven miRNAs were up-regulated while expression of a polycistronic miRNA cluster (miR-17/92) was significantly decreased in pNL4-3 infected *vs.* uninfected Jurkat cells [147]. Intriguingly, the miR-17/92 cluster was shown to target PCAF (P300/CBP-associated factor), a histone acetyltransferase that binds Tat to enhance HIV-1 gene expression. Suppression of the miR-17/92 cluster resulted in increased expression of PCAF and enhanced viral replication in PBMCs from HIV-infected donors.

The PBMC miRNA profiles derived from HIV-1 seropositive individuals have been fairly well characterized although the total number of patients sampled was small [149, 150]. In one study, expression levels of 327 miRNAs were assessed by miRNA microarrays in PBMCs isolated from four classes of HIV-1 donors stratified according to CD4+ T cell counts and viral loads [149]. Overall 62 miRNAs were dysregulated in response to HIV-1 infection, 59 of which were down-regulated. Of these, 12 miRNAs were common to all four classes with the remainder separated into class-specific miRNA signatures. These profiles overlapped significantly with those identified in a second study using miRNA microarrays probing for 518 mature miRNAs in HIV-positive donor PBMCs, pNL4-3-infected PBMCs and pNL4-3-infected CEMx174 cells [150]. In total 62 miRNAs were dysregulated in HIV-infected PBMCs, 32 of which were down-regulated, and 33 of which were similarly expressed in CEMx174 cells. Expression of two members of the miR-17/92 cluster, miR-17-5p and miR-20a were suppressed by HIV-1 infection, in line with previous studies [147, 149]. The results of these studies would only benefit from additional data collected across significantly larger populations of HIV-positive donors, encompassing different HIV subtypes, and profiling of the entire known miRnome. However, the current data already suggests that miRNAs may fine-tune HIV's changeover from latency to activation [151], clearing the viral reservoirs in T cells, and reduction in HIV replication [152].

HIV Regulation by Host-Encoded miRNAs

Host miRNAs could potentially regulate HIV-1 replication by targeting cellular factors required by the virus during infection or by directly inhibiting viral transcripts. As described above, suppression of the miR-17/92 cluster caused an increase in PCAF binding to Tat and subsequent enhanced viral replication [147]. Similarly, targeting of cyclinT1 by miR-198 reduced pNL4-3-Luc expression in the promonocytic cell line, Mono Mac 6 (MMC) following stimulation with PMA

[153]. CyclinT1 forms part of the P-TEFb complex required for RNA Pol II-mediated transcription elongation, thus the mechanism of HIV-1 down-regulation seems to be *via* miR-198 inhibition of this cellular factor that is required by the virus. Notably, monocytes express high levels of miR-198 but this is substantially reduced following their differentiation into macrophages, and this oscillating miRNA expression was suggested to be a host-encoded mechanism for restricting HIV-1 infection of monocytes [153].

The Nef sequence of HIV-1 acts as the 3' UTR for the majority of viral transcripts due to its location in the genome. Nef is believed to play a part in pathogenesis [154], and viruses encoding Nef mutations have been linked to slower disease progression [155], suggesting that miRNAs targeted to this region may negatively affect HIV-1 replication. Both miR-29a and miR-29b were shown to target Nef when expressed in a luciferase reporter in HEK293 cells [152]. These two miRNAs were also identified in PBMCs and their over-expression reduced HIV-1 replication in Jurkat and HEK293 cells. In a separate study focused on miRNA profiles in activated *vs*. resting CD4+ T cells, miR-125b, miR-150, miR-223 and miR-382 were highly expressed in the resting cells only and were predicted to target Nef although this was not experimentally validated [151].

Viral replication in H9 T lymphocytes was also decreased in response to miR-29a expression although the mechanism did not involve targeting Nef [146]. Instead, miR-29a enhanced the association of HIV-1 mRNA with RISC proteins and processing body (P-body) complexes. P-bodies include proteins such as Rck/p54 (RNA helicase), GW182 (interacts with Argonaute proteins), LSm-I (RNA binding protein) and XRN1 (5' to 3' exonuclease) that sequester mRNA transcripts as part of a translation control mechanism [156]. Down-regulation of any of these proteins by siRNAs caused a decrease in of pNL4-3-Luc expression in HeLa cells, and knock-down of RCK/p54 released the translation suppression in H9 T lymphocytes [156], as well as virus re-activation in PBMCs from patients on combinatorial antitretrovirals [157].

Taken together, these limited studies provide evidence that some host miRNAs regulate HIV-1 replication, but many outstanding questions remain. Which additional miRNAs regulate HIV-1 expression? What roles do differentially expressed host miRNAs play in HIV-1 pathogenesis? How do P-bodies (and potentially stress granules) assist in miRNA-mediated translational control of HIV-1? Can HIV-1 alter the host RNA content of P-bodies and if so, what are the molecular players involved? These questions, among others will need to be

answered before we have a confident understanding of how host miRNAs guide HIV-1 infection.

HIV-Encoded miRNAs

Plants regularly process virus-encoded dsRNAs into siRNAs that target viral transcripts as part of a highly effective mechanism to control infection [158]. Numerous mammalian viruses have been shown to encode miRNAs including EBV, members of the Herpes family, CMV, MDV and SV40 [79, 159]. Contrasting data exist on whether or not HIV-1 encodes miRNAs. A single publication revealed an HIV-encoded miRNA, termed vsiRNA1, which targeted Env mRNA and reduced subsequent protein levels [141]. One group published data on HIV-encoded miR-H1 that targets cellular AATF thereby initiating apoptosis in human mononuclear cells [160]. A separate group showed miR-N367 to be encoded within *nef* and suggested this miRNA could reduce HIV-1 expression by binding to the negative response element within the U3 region of the 5' LTR promoter [161]. A major limitation of these studies is the lack of independent validation and the observation that none of these sequences were detected in persistently infected T cells [142]. In contrast to this, two independent groups reported two miRNAs encoded within the TAR loop of HIV-1 that aid in chromatin remodeling at the viral LTR, and prevent apoptosis in infected cells by down-regulating cellular ERCC1 and IER3 transcripts [162-165]. Importantly, only TAR-3p was detectable by Northern blot and both miRNAs were 17-18 nt in length which is smaller than the previously established Dicer products (21-22 nt).

With increased access to deep sequencing, additional HIV-encoded small RNAs have been identified. A highly abundant 18 nt RNA was identified in HIV-infected MT4 T cells that is antisense to the viral primer binding site (PBS) and associates with AGO2 [166]. A more in-depth sequencing analysis revealed viral-encoded small RNAs that target both host and viral transcripts in HIV-infected T lymphocytes [167]. Numerous small RNAs corresponding to the HIV genome were detected and due to their positive polarities were defined as viral miRNAs (vmiRNAs). Antagomirs targeted to the vmiRNAs stimulated viral production suggesting a 'classical' anti-viral RNAi effect related to these transcripts. However, a small proportion of viral-encoded RNAs had negative polarities, corresponded to the 3' UTR of the viral genome, and potently suppressed HIV transcription (so-called viral siRNAs or vsiRNAs). Importantly, while the vmiRNAs were transcribed from the integrated HIV promoter, the vsiRNAs were transcribed from cellular promoters positioned downstream of the inserted HIV provirus, thereby providing two separate sets of transcripts that act in concert to

direct host and/or viral gene expression. Overall, while other RNA viruses are known to encode miRNAs that regulate their own gene expression, HIV-1 may not have such control due to its exceptionally high mutation rate. The TAR loop is the most highly conserved region suggesting the miRNAs encoded within this sequence may indeed regulate HIV-1 but additional supportive data is required.

Respiratory Syncytial Virus

Respiratory syncytial virus (RSV) leads to considerable morbidity and severe lower respiratory tract disease in both the young and senior populations. RSV infections account for >14 000 deaths per annum [168] and despite nearly 6 decades of research on RSV, vaccine and therapeutic approaches are limited. RNAi efforts targeting crucial host/viral pathways have been recently shown to limit viral replication [169-171] and siRNAs against the RSV N gene have progressed to Phase III clinical trials [172]. Like many other RNA viruses, RSV modulates the host immune response *via* surface and internal viral proteins. Chief among these are the non-structural proteins (NS1/2), which work together to inhibit activation and nuclear translocation of IFN regulatory factor 3 (IRF3) [173, 174], and play a role in the suppression of cytokine production by proteasome-mediated breakdown of the signal transducer and activator of transcription factor 2 (STAT2) [175, 176]. Together with nucleolin, the F and G proteins participate in attachment and entry [177]. Additionally, RSV G associates with Toll-like receptors (TLR) [178-180], and negatively affects type I IFN [181, 182], and cytokine and chemokine expression [183], in part through the activation of suppressor of cytokine signaling (SOCS) proteins in normal human bronchoepithelial (NHBE) cells and mouse lung epithelial cells [178, 181]. Furthermore, the G protein's CX3C motif imitates fractalkine (CX3CL1), thereby altering immune responses mediated by this chemokine [184].

In vitro and *in vivo* RSV infections stimulate early, middle and late genome-wide transcription profile changes in the host [20, 185], however, these modifications are not precisely mirrored in the host proteome [186]. RSV infection results in G1/S arrest in A549 [187] and HEp-2 cells [188], and a G2/M arrest in primary human bronchial epithelial cells through stimulation of TGF-β1 and diminution of p53 both *in vitro* and *in vivo*. In HEp2 cells, the virus similarly triggers stress granule formation in a protein kinase R-dependent manner, prompting enhanced viral replication in cytosolic viral inclusion bodies [189, 190], which share components with P-bodies in the cytosol [191, 192].

Two studies have recently reported the deregulation of miRNA expression following infection with RSV. Othumpungat *et al* reported the repression of 24 and upregulation of two miRs (miR-886-3p and miR-375) with ~2 fold deregulation observed following HCV infection [193]. Of the miRNAs uncovered, six were expected to influence the neurotrophin nerve growth factor (NGF) gene, which is involved in regulating the NGF-TrKA axis and may allow airway epithelium to tolerate RSV challenge. Transfection with miR-221 mimics reduced NGF transcripts and protein although the miR-221 seed does not have a direct target in the NGF 5'- or 3' UTR and shows insignificant matches within the NGF coding region. miR-221 is a negative regulator of several tumor suppressor genes including PTEN [194], Bim [195], PUMA [196] and transcription factor Foxo3a [197], suggesting that pre-miR-221 treatment enhances apoptosis by reducing the activity of these tumor suppressors (also observed by Othumpangat [193]). We recently reported that when A549 cells are infected with RSV, five miRNAs (let-7f, miR-337, miR-520a, miR-24, miR-26b) are upregulated while two miRs (miR-198 and miR-595) are downregulated Furthermore, we observed that RSV G induced let-7f miRNA in the context of RSV infection [198]. Computational predictions and comparison with existing gene expression data for RSV in A549 cells showed that a substantial number of genes deregulated during RSV infection were predicted to be regulated by let-7f and other RSV deregulated miRNAs. By transfecting miRIDIAN miRNA mimics or inhibitors along with CMV-driven Luciferase-UTR constructs of a subset of let-7f targets, we showed let-7f mediated regulation of many RSV deregulated genes suggesting that let-7f may be serving a pro-viral role. Inhibition of let-7f reduced viral replication while upregulation of let-7f activity increased viral replication relative to control transfections.

Herpesviruses

Herpesviruses are a group of large, enveloped dsDNA viruses comprised of α, β and γ subfamilies based on genomic sequence. Intriguingly, of the >200 viral-encoded miRNAs currently known, the vast majority have been identified in Herpesvirus members (see Table 1 of [199]). Typically dsDNA viruses utilize bi-directional transcription to generate mRNAs and thus viral-targeted miRNAs can be readily encoded within the antisense transcripts. Furthermore, the location of dsDNA viruses within host nuclei provides them access to nuclear RNAi processing machinery allowing them to take advantage of this evolutionarily conserved pathway. This is particularly important for Herpesviruses as they, like HIV, establish long-term latent infections and thus require mechanisms to control the host immune response in an ongoing manner. Given the flexibility in miRNA-

mediated mRNA targeting and the ability of miRNAs to modulate expression of numerous discrete transcripts, miRNAs represent potent tools used by Herpesviruses, among others, to control host gene expression and subsequent immune responses to infection.

α-Herpesviruses

Human herpes simplex virus 1 (HSV-1) and 2 (HSV-2) are both α-herpesviruses that establish latent infections in neuronal cells; causing cold sores and genital herpes respectively. During latent infection, noncoding latency associated transcripts (LATs) predominate, and massively parallel sequencing has uncovered numerous miRNAs within the LAT region that down-regulate viral mRNAs [200-202]. Overall, nine miRNAs are conserved between these two viruses, particularly in the seed sequences [200]. A single miRNA (miR-H1) which is encoded roughly 300 bp upstream of the LAT region in HSV-1 only, exhibits elevated expression during productive infection [94, 200], and a functional analog (miR-H6) that shares seed sequence conservation with miR-H1 has been identified in cells productively infected with HSV-2 [200]. The combination of a single miRNA expressed during active infection and multiple distinct miRNAs expressed during latency suggests that these α-herpesviruses utilize these self-encoded sequences to facilitate and maintain stable latent infections in host cells.

β-Herpesviruses

Human cytomegalovirus (CMV) is a β-herpesvirus that causes apathogenic infection of up to 90% of individuals worldwide but maintains a high morbidity rate in immunocompromised patients and is the primary source (1%) of congenital abnormalities [94]. Typically, CMV establishes a latent infection in haematopoetic progenitor cells but despite the discovery of 12 functional miRNAs scattered throughout the viral genome [203, 204], none seem to be expressed during latent infection. While no animal models exist to investigate the roles of these miRNAs *in vivo*, a recent study conducted in human foreskin fibroblasts revealed that intergenic RNAs located within a 15 kb segment of the genome function as a host-directed microRNA decay element [205]. This decay element directed specific degradation of host miR-17 and miR-20a and resulted in accelerated virus production during lytic infection. Interestingly, the same two host miRNAs, encoded within the miR17/92 cluster, are down-regulated in PBMCs from HIV-infected donors resulting in similarly enhanced viral replication (see section *'HIV-mediated perturbations of host miRNA profiles'*). Furthermore, miRNAs encoded within the mouse CMV genome have been linked

to replication in salivary glands [206], providing additional data for the role of CMV miRNAs in lytic but not latent infection.

γ -Herpesviruses

The first viral-encoded miRNAs were identified in B cells latently infected with Epstein-Barr Virus (EBV), a member of the γ –herpesvirus sub-family [207]. EBV maintains a life-long latent persistence in humans and is linked with numerous lymphocyte proliferative disorders including Burkitt's lymphoma, Hodgkin's lymphoma and NK/T cell lymphoma [208]. Another γ –herpesvirus that infects B cells and causes similar proliferative disorders is Kaposi's sarcoma-associated herpesvirus (KSHV), including its namesake Kaposi's sarcoma. Both EBV and KSHV are directly linked to human cancers [209] and while the complex interactions between host and viral proteins have been largely characterized, a blossoming appreciation of EBV/KSHV-encoded miRNAs has developed. Both of these viruses encode distinct miRNAs in terms of sequence, but share a similar genomic organization with a clustering of miRNAs in latency associated regions. Specifically, EBV encodes 25 miRNAs organized into the BART (22 miRNAs) and BHRF1 (3 miRNAs) clusters, while all 12 KSHV miRNAs are located within one latency cluster [210].

Both EBV and KSHV utilize their miRNAs in immune evasion, by disrupting signaling between infected B cells and surrounding immune cells, as well as to regulate apoptosis and the host cell cycle. EBV-encoded miRNAs target the IFN-inducible T-cell attracting chemokine CXCL-11 which is an NK cell and Th1 lymphocyte ligand, thereby disrupting NK-mediated cytotoxicity of infected B cells [211]. Similarly, NK cells and CD8+ T lymphocytes generate a cytotoxic response through binding of their NKG2D receptors and the major histocompatibility complex class I-related chain B protein (MICB) expressed on infected B cells. EBV miRNAs BART1, BART3, BART5 and BART9 as well as KSHV miRNA-K12-7 all target MICB transcripts resulting in decreased MICB expression and subsequent NKG2D-mediated cell death [212]. The observation that MICB mRNA is targeted by multiple miRNAs from different herpesvirus members suggests that MICB is a key cellular obstacle during infection. Several additional EBV miRNA targets have been identified but functional studies to elucidate the mechanisms of immune evasion are lacking [208]. For example, host PDE7A, which is required for proliferation of NK cells [213], is a target of EBV miRNA-BART1 and miRNA-BART3-3p, but the effects on the viral life cycle remain unknown. Similarly, SP100, elements of promyelocytic leukemia-nuclear

bodies that function as part of the cellular antiviral response, is silenced by EBV miRNA-BART1-5p but the mechanisms have not been defined [213].

Viral pathogens that initiate latent infections have evolved mechanisms to control cellular apoptosis and cell cycling as a means to ensure persistence. Both EBV and KHSV utilize encoded miRNAs to target pro- and anti-apoptotic host proteins as well as cell cycle arrest factors to promote long-term viral infection. The pro-apoptotic protein Bcl-2-associated factor (BCLAF1) is targeted by KSHV miRNAs-K12-5, -9, -10b and -3 [214], while the pro-apoptotic modulator BCL2L11 (Bim) is down-regulated by numerous KSHV and EBV-encoded miRNAs [215, 216]. Caspase 3, a pivotal regulator of cellular apoptosis is a direct target of KSHV miRNAs [217] and various tumor suppressors, including p53 and PUMA (BBC3), are targeted by EBV and KSHV miRNAs [218, 219]. Interestingly, KSHV miRNA-K12-1 down-regulates the cyclin-dependent kinase inhibitor protein p21 to halt cell-cycle arrest and ensure availability of host factors in latently infected cells [220]. Furthermore, the KSHV LANA protein has been shown to inhibit the TGFβ-type II receptor, a key component of the anti-proliferation pathway, by mediating epigenetic silencing as a means to regulate host cell cycling [221]. In an analogous manner, EBV controls cell cycling by upregulating host miR-34a, which is required for cellular proliferation [222]. Overall, the use of herpesvirus-encoded miRNAs as opposed to immunogenic viral proteins provides these pathogens with a highly effective mechanism to control apoptosis during latent infection while minimizing exposure to innate immunity.

While herpesviruses generally maintain a latent state of infection, they do switch to a lytic phase at some point in their lifecycle. The control of the latent to lytic state is critical and seems to be mediated by viral-encoded miRNAs that target both host and herpesvirus mRNAs. EBV miRNA-BART6 targets Dicer and fine-tunes the levels of this protein to control infection through all the latent (I, II, III) and lytic stages [223]. During latency, BART6-mediated reduction in Dicer results in suppression of genes that facilitate lytic replication and BART6-specific antagomirs shift EBV back into a lytic phase [224]. EBV also maintains latency by expressing miRNA-BHRF1 that targets host p53 [218], as well as BART2 that has complete sequence complementarity to BALF5 mRNA encoding the viral DNA polymerase [225]. Upon lytic reactivation, this suppression is relieved and EBV shifts out of a latent phase. KSHV miRNA-K1 targets the NFκB inhibitor IκBα to ensure activated NFκB and viral latency [226]. Similarly, KSHV miRNA-K12-11 targets IKKε resulting in decreased interferon signaling and ongoing

latency [227]. Intriguingly, KSHV miRNA-K12-4-5 represses host Rbl2, a DNA methyltransferase (DNMT) inhibitor, to control the latency to lytic phase shift [228]. Suppression of Rbl2 caused activation of DNMTs, loss of DNA CpG methylation and subsequent silencing of various cellular genes.

Influenza

Influenza A viruses are a worldwide cause of acute respiratory disease and can cause substantial morbidity and mortality each year [229-231] especially in the young and elderly. Treatment or prophylaxis with licensed antivirals curtail influenza morbidity by >80% efficacy during inter-pandemic influenza periods [232]. Influenza is an Orthomyxovirus with a genome that contains 8 negative-sense, single-stranded RNA segments that encode up to 11 proteins [233]. Owing to their segmented genome, influenza virus segments can reassert in species such as pigs that are susceptible to both human and avian influenza virus strains.

Host miRNA expression (*in vitro* and *in vivo*) is significantly deregulated following challenge with a wide variety of influenza viruses from human, avian and swine origin. Early computational studies predicted that host-encoded miRNAs interact with genome segments of both swine and human origin influenza A viruses [234]. miRNA seed sequences incorporated into influenza A virus to generate live attenuated vaccines showed complete attenuation, a significant (>2 log) reduction in mortality and a diversified antibody response to viral challenge, suggesting that miRNA-based attenuation of influenza viruses can be an effective strategy for vaccine development [235]. Chicken lungs and trachea differentially express 377 and 149 miRNAs relative to non-infected animals following low pathogenic H5N3 avian influenza virus infection. Gene ontology analysis of induced miRNAs showed that multiple categories involved in virus regulation and immune response were significantly enriched, suggesting that these miRNAs potentially regulate the host response to infection [236].

Whole genome miRNA expression profiles of mice infected with recombinant pandemic 1918 pandemic influenza virus (r1918) and a non-lethal seasonal influenza strain (A/Texas/36/91) identified miR-200a and miR-223 as significantly deregulated miRNAs that were predicted to regulate pathways associated with immune response and cell death. Expression profiles of numerous genes predicted to be regulated by these miRNAs showed significant reciprocal expression in mice infected with r1918 and A/Texas/36/91 suggesting that these genes are likely targets for the deregulated miRNAs [237]. Not only are host miRNAs deregulated by influenza infection, but host miRNAs miR-323, -491 and

-654 can bind to influenza H1N1 PB1 transcript and cause mRNA degradation to inhibit viral replication [238]. The leader sequences at the 5' end of the influenza genome segments are also processed by nuclear Drosha and cytosolic Dicer to generate a pool of small viral leader sequences (leRNAs) that potentially regulate the polymerase switch from transcription to viral genome synthesis [239].

Infection of human lung epithelial cells by influenza A/WSN/33 and A/ Udorn/72 induces expression of miR-7, -132, -146a, -187, -200c and miR-1275 and 26 innate immunity-associated genes (including IRAK1 and MAPK3) are targeted by these miRNAs [240]. A recent study showed that miR-29a induced in A549 lung epithelial cells negatively regulates the expression of DNA methyl transferase 3a and 3b to induce the expression of cyclooxygenase 2 (COX2) and prostaglandin E2 (PGE$_2$) [241]. miR-29c has also been shown to regulate BCL2L2 RNA that induces cellular apoptosis during infection [242]. In a recent study, we identified host proteases that are essential for influenza virus replication and showed that miR-106b and miR-124* regulate expression of ADAMTS7,DPP3, MST1, while miR-106b* was shown to regulate PRSS12 expression [243]. miR-451 inhibition was recently shown to induce the expression of pro-inflammatory cytokines IL-6, TNF, CCL5/RANTES and CCL3/MIP1α in primary murine dendritic cells upon influenza infection by inhibiting transcription factor YWHAZ/14-3-3ζ as well as two negative regulators of cytokine production FOXO3 and ZFP36/Tristetraprolin [244]. Deep sequencing approaches to identify small RNAs differentially expressed in broiler chickens upon influenza A challenge identified gga-miR-206 and gga-miR-451 and gga-miR-146a which were subsequently shown to regulate expression of ARL11, CHMP2B, POU1F1, PDHB and HIF1AN [245].

Infection of A549 cells with highly pathogenic avian influenza A/H5N1 identified miR-21*, miR-100*, miR-141, -574-3p, -1274a and -1274b as highly induced miRNAs. miR-141 was found to regulate the expression of pro-inflammatory cytokine TGF-β [246]. Swine challenged with influenza H1N2 significantly induced the expression of miR-15a, -21, -146, -206, -223 and -451 in addition to several key cytokines [247]. miR-146a induced in A549 cells also regulates miR-146a promoter activity and knockdown of this non-coding RNA increases viral replication [248]. We recently showed that miR-149* regulates NEK8, miR-548d regulates MAP3K1 and miR-1228 and miR-138 regulate CDK13, all of which are kinases that are crucial for influenza virus replication [249]. Overall these studies suggest that influenza virus infection in human, avian and swine deregulates expression of multiple miRNA species to regulate the host anti-viral response.

CONCLUSIONS

Clearly host-encoded miRNAs have a vital function in regulating the immune response to infection, and tissue-specific miRnomes that change temporally provide cells with an expansive repertoire of RNA-based 'tools' to rapidly and specifically fine-tune gene expression. Similarly, pathogens (and viruses in particular) have evolved both their own encoded miRNAs as well as mechanisms to subvert host miRNA-mediated immunity. While there is a clear need to improve the identification of miRNA targets, especially when considering seed sequence alignment alone, the overall number of potential binding sites has been drastically downscaled with an improved understanding of RNA secondary structure and its impact on miRNA target accessibility. Furthermore, as deep sequencing methods capable of detecting extremely low copies of short sequences continue to improve, and more studies are conducted in physiologically relevant primary cells, the true nature of the complex interplay occurring in the cellular noncoding RNA space during infection will be revealed.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors confirm that this chapter contents have no conflict of interest.

REFERENCES

[1]   Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998;391(6669):806-11.

[2]   Yamada K, Lim J, Dale JM, Chen H, Shinn P, Palm CJ, *et al*. Empirical analysis of transcriptional activity in the Arabidopsis genome. Science. 2003;302(5646):842-6.

[3]   Ponting CP, Belgard TG. Transcribed dark matter: meaning or myth? Human molecular genetics. 2010;19(R2):R162-8.

[4]   Saxena A, Carninci P. Long non-coding RNA modifies chromatin: epigenetic silencing by long non-coding RNAs. BioEssays : news and reviews in molecular, cellular and developmental biology. 2011;33(11):830-9.

[5]   Liu C, Bai B, Skogerbo G, Cai L, Deng W, Zhang Y, *et al*. NONCODE: an integrated knowledge database of non-coding RNAs. Nucleic acids research. 2005;33(Database issue):D112-5. Epub 2004/12/21.

[6]   Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, *et al*. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860-921.

[7]   Berezikov E. Evolution of microRNA diversity and regulation in animals. Nature reviews Genetics. 2011;12(12):846-60. Epub 2011/11/19.

[8]   Grimson A, Srivastava M, Fahey B, Woodcroft BJ, Chiang HR, King N, *et al*. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature. 2008;455(7217):1193-7. Epub 2008/10/03.

[9]   Campo-Paysaa F, Semon M, Cameron RA, Peterson KJ, Schubert M. microRNA complements in deuterostomes: origin and evolution of microRNAs. Evolution & development. 2011;13(1):15-27. Epub 2011/01/08.

[10]   Pederson T. Regulatory RNAs derived from transfer RNA? RNA. 2010;16(10):1865-9. Epub 2010/08/20.

[11]   Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, *et al*. A human snoRNA with microRNA-like functions. Molecular cell. 2008;32(4):519-28. Epub 2008/11/26.

[12]   Piriyapongsa J, Marino-Ramirez L, Jordan IK. Origin and evolution of human microRNAs from transposable elements. Genetics. 2007;176(2):1323-37. Epub 2007/04/17.

[13]   Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, *et al*. MicroRNA genes are transcribed by RNA polymerase II. The EMBO journal. 2004;23(20):4051-60.

[14]   Lee Y, Jeon K, Lee JT, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. The EMBO journal. 2002;21(17):4663-70.

[15]   Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, *et al*. The Microprocessor complex mediates the genesis of microRNAs. Nature. 2004;432(7014):235-40.

[16]   Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, *et al*. The nuclear RNase III Drosha initiates microRNA processing. Nature. 2003;425(6956):415-9.

[17]   Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ. Processing of primary microRNAs by the Microprocessor complex. Nature. 2004;432(7014):231-5.

[18]   Zeng Y, Yi R, Cullen BR. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. The EMBO journal. 2005;24(1):138-48.

[19]   Morlando M, Ballarino M, Gromak N, Pagano F, Bozzoni I, Proudfoot NJ. Primary microRNA transcripts are processed co-transcriptionally. Nature structural & molecular biology. 2008;15(9):902-9.

[20]   Janssen R, Pennings J, Hodemaekers H, Buisman A, van Oosten M, de Rond L, *et al*. Host transcription profiles upon primary respiratory syncytial virus infection. Journal of virology. 2007;81(11):5958-67. Epub 2007/03/23.

[21]   Han J, Pedersen JS, Kwon SC, Belair CD, Kim YK, Yeom KH, *et al*. Posttranscriptional crossregulation between Drosha and DGCR8. Cell. 2009;136(1):75-84.

[22]   Berezikov E, Chung WJ, Willis J, Cuppen E, Lai EC. Mammalian mirtron genes. Molecular cell. 2007;28(2):328-36.

[23]   Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. Nature. 2007;448(7149):83-6.

[24]   Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC. The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. Cell. 2007;130(1):89-100.

[25]   Lund E, Dahlberg JE. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. Cold Spring Harbor symposia on quantitative biology. 2006;71:59-66.

[26]   Jaskiewicz L, Filipowicz W. Role of Dicer in posttranscriptional RNA silencing. Current topics in microbiology and immunology. 2008;320:77-97.

[27]   Jinek M, Doudna JA. A three-dimensional view of the molecular machinery of RNA interference. Nature. 2009;457(7228):405-12.

[28]   Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature. 2001;411(6836):494-8.

[29]   Elbashir SM, Lendeckel W, Tuschl T. RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes & development. 2001;15(2):188-200.

[30]   Daniels SM, Melendez-Pena CE, Scarborough RJ, Daher A, Christensen HS, El Far M, *et al*. Characterization of the TRBP domain required for dicer interaction and function in RNA interference. BMC molecular biology. 2009;10:38.

[31]   Lee Y, Hur I, Park SY, Kim YK, Suh MR, Kim VN. The role of PACT in the RNA silencing pathway. The EMBO journal. 2006;25(3):522-32.

[32]   Hammond SM, Bernstein E, Beach D, Hannon GJ. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature. 2000;404(6775):293-6.

[33]   Zamore PD, Tuschl T, Sharp PA, Bartel DP. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell. 2000;101(1):25-33.

[34]   Martinez J, Patkaniowska A, Urlaub H, Luhrmann R, Tuschl T. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell. 2002;110(5):563-74.

[35]   Preall JB, Sontheimer EJ. RNAi: RISC gets loaded. Cell. 2005;123(4):543-5.

[36]   Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, *et al*. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature. 2005;436(7051):740-4.

[37]   Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. Cell. 2003;115(2):199-208.

[38]   Azuma-Mukai A, Oguri H, Mituyama T, Qian ZR, Asai K, Siomi H, *et al*. Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(23):7964-9.

[39]   Eulalio A, Huntzinger E, Izaurralde E. Getting to the root of miRNA-mediated gene silencing. Cell. 2008;132(1):9-14.

[40]   Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, *et al*. Argonaute2 is the catalytic engine of mammalian RNAi. Science. 2004;305(5689):1437-41.

[41]   Yekta S, Shih IH, Bartel DP. MicroRNA-directed cleavage of HOXB8 mRNA. Science. 2004;304(5670):594-6.

[42]   Martinez J, Tuschl T. RISC is a 5' phosphomonoester-producing RNA endonuclease. Genes & development. 2004;18(9):975-80.

[43]   Nilsen TW. Mechanisms of microRNA-mediated gene regulation in animal cells. Trends in genetics : TIG. 2007;23(5):243-9.

[44]   Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Molecular and cellular biology. 2007;27(11):3970-81.

[45]   Hock J, Meister G. The Argonaute protein family. Genome biology. 2008;9(2):210.

[46]   Braun JE, Huntzinger E, Izaurralde E. The role of GW182 proteins in miRNA-mediated gene silencing. Advances in experimental medicine and biology. 2013;768:147-63.

[47]   Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, *et al*. The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. Science. 2005;310(5755):1817-21.

[48]   Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005;120(1):15-20.

[49]   Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, *et al*. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature. 2005;433(7027):769-73.

[50]   Barbato C, Arisi I, Frizzo ME, Brandi R, Da Sacco L, Masotti A. Computational challenges in miRNA target predictions: to be or not to be a true target? Journal of biomedicine & biotechnology. 2009;2009:803069.

[51]   Sashital DG, Doudna JA. Structural insights into RNA interference. Current opinion in structural biology. 2010;20(1):90-7.

[52]   Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP. The impact of microRNAs on protein output. Nature. 2008;455(7209):64-71.

[53]   Ebert MS, Sharp PA. Roles for microRNAs in conferring robustness to biological processes. Cell. 2012;149(3):515-24.

[54]   Chen SJ, Chen GH, Chen YH, Liu CY, Chang KP, Chang YS, *et al*. Characterization of Epstein-Barr virus miRNAome in nasopharyngeal carcinoma by deep sequencing. PloS one. 2010;5(9).

[55]   Li Z, Zhang S, Huang C, Zhang W, Hu Y, Wei B. MicroRNAome of splenic macrophages in hypersplenism due to portal hypertension in hepatitis B virus-related cirrhosis. Experimental biology and medicine. 2008;233(11):1454-61.

[56]   Marcinowski L, Tanguy M, Krmpotic A, Radle B, Lisnic VJ, Tuddenham L, *et al*. Degradation of cellular mir-27 by a novel, highly abundant viral transcript is important for efficient virus replication *in vivo*. PLoS pathogens. 2012;8(2):e1002510.

[57]   Meshesha MK, Veksler-Lublinsky I, Isakov O, Reichenstein I, Shomron N, Kedem K, *et al*. The microRNA Transcriptome of Human Cytomegalovirus (HCMV). The open virology journal. 2012;6:38-48.

[58]   Ostermann E, Tuddenham L, Macquin C, Alsaleh G, Schreiber-Becker J, Tanguy M, *et al*. Deregulation of type I IFN-dependent genes correlates with increased susceptibility to cytomegalovirus acute infection of dicer mutant mice. PloS one. 2012;7(8):e43744.

[59]   Yao Y, Zhao Y, Smith LP, Lawrie CH, Saunders NJ, Watson M, *et al*. Differential expression of microRNAs in Marek's disease virus-transformed T-lymphoma cell lines. J Gen Virol. 2009;90(Pt 7):1551-9.

[60]   Yao Y, Zhao Y, Xu H, Smith LP, Lawrie CH, Watson M, *et al*. MicroRNA profile of Marek's disease virus-transformed T-cell line MSB-1: predominance of virus-encoded microRNAs. J Virol. 2008;82(8):4007-15.

[61]   Kuehbacher A, Urbich C, Zeiher AM, Dimmeler S. Role of Dicer and Drosha for endothelial microRNA expression and angiogenesis. Circulation research. 2007;101(1):59-68.

[62]   Dalzell JJ, Warnock ND, Stevenson MA, Mousley A, Fleming CC, Maule AG. Short interfering RNA-mediated knockdown of drosha and pasha in undifferentiated Meloidogyne incognita eggs leads to irregular growth and embryonic lethality. International journal for parasitology. 2010;40(11):1303-10.

[63]   Rehwinkel J, Natalin P, Stark A, Brennecke J, Cohen SM, Izaurralde E. Genome-wide analysis of mRNAs regulated by Drosha and Argonaute proteins in Drosophila melanogaster. Molecular and cellular biology. 2006;26(8):2965-75.

[64]   Fan P, Chen Z, Tian P, Liu W, Jiao Y, Xue Y, *et al*. miRNA biogenesis enzyme Drosha is required for vascular smooth muscle cell survival. PloS one. 2013;8(4):e60888.

[65]   Francia S, Michelini F, Saxena A, Tang D, de Hoon M, Anelli V, *et al*. Site-specific DICER and DROSHA RNA products control the DNA-damage response. Nature. 2012;488(7410):231-5.

[66]   Cobb BS, Hertweck A, Smith J, O'Connor E, Graf D, Cook T, *et al*. A role for Dicer in immune regulation. The Journal of experimental medicine. 2006;203(11):2519-27.

[67]   Liston A, Lu LF, O'Carroll D, Tarakhovsky A, Rudensky AY. Dicer-dependent microRNA pathway safeguards regulatory T cell function. The Journal of experimental medicine. 2008;205(9):1993-2004.

[68]   Fedeli M, Napolitano A, Wong MP, Marcais A, de Lalla C, Colucci F, *et al*. Dicer-dependent microRNA pathway controls invariant NKT cell development. Journal of immunology. 2009;183(4):2506-12.

[69]   Ruggiero T, Trabucchi M, De Santa F, Zupo S, Harfe BD, McManus MT, *et al*. LPS induces KH-type splicing regulatory protein-dependent processing of microRNA-155 precursors in macrophages. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2009;23(9):2898-908.

[70]   Lim DH, Lee L, Oh CT, Kim NH, Hwang S, Han SJ, *et al*. Microarray analysis of Drosophila dicer-2 mutants reveals potential regulation of mitochondrial metabolism by endogenous siRNAs. Journal of cellular biochemistry. 2013;114(2):418-27.

[71]   Inchley CS, Sonerud T, Fjaerli HO, Nakstad B. Reduced Dicer expression in the cord blood of infants admitted with severe respiratory syncytial virus disease. BMC infectious diseases. 2011;11:59.

[72]   Zhang C, Huys A, Thibault PA, Wilson JA. Requirements for human Dicer and TRBP in microRNA-122 regulation of HCV translation and RNA abundance. Virology. 2012;433(2):479-88.

[73]   Matskevich AA, Moelling K. Dicer is involved in protection against influenza A virus infection. J Gen Virol. 2007;88(Pt 10):2627-35.

[74]   Ghosh J, Bose M, Roy S, Bhattacharyya SN. Leishmania donovani targets Dicer1 to downregulate miR-122, lower serum cholesterol, and facilitate murine liver infection. Cell host & microbe. 2013;13(3):277-88.

[75]   Ma F, Xu S, Liu X, Zhang Q, Xu X, Liu M, *et al*. The microRNA miR-29 controls innate and adaptive immune responses to intracellular bacterial infection by targeting interferon-gamma. Nature immunology. 2011;12(9):861-9.

[76]   Schulte LN, Eulalio A, Mollenkopf HJ, Reinhardt R, Vogel J. Analysis of the host microRNA response to Salmonella uncovers the control of major cytokines by the let-7 family. The EMBO journal. 2011;30(10):1977-89.

[77]   Sheedy FJ, Palsson-McDermott E, Hennessy EJ, Martin C, O'Leary JJ, Ruan Q, *et al*. Negative regulation of TLR4 *via* targeting of the proinflammatory tumor suppressor PDCD4 by the microRNA miR-21. Nature immunology. 2010;11(2):141-7.

[78]   Taganov KD, Boldin MP, Chang KJ, Baltimore D. NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(33):12481-6.

[79]   Cullen BR. MicroRNAs as mediators of viral evasion of the immune system. Nature immunology. 2013;14(3):205-10.

[80]   Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, *et al*. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell. 2006;126(6):1203-17.

[81]   Bentwich I, Avniel A, Karov Y, Aharonov R, Gilad S, Barad O, *et al*. Identification of hundreds of conserved and nonconserved human microRNAs. Nature genetics. 2005;37(7):766-70.

[82]   Brameier M, Wiuf C. Ab initio identification of human microRNAs based on structure motifs. BMC bioinformatics. 2007;8:478.

[83]   Nam JW, Shin KR, Han J, Lee Y, Kim VN, Zhang BT. Human microRNA prediction through a probabilistic co-learning model of sequence and structure. Nucleic Acids Res. 2005;33(11):3570-81.

[84]   Helvik SA, Snove O, Jr., Saetrom P. Reliable prediction of Drosha processing sites improves microRNA gene prediction. Bioinformatics. 2007;23(2):142-9.

[85]   Lim LP, Lau NC, Weinstein EG, Abdelhakim A, Yekta S, Rhoades MW, *et al*. The microRNAs of Caenorhabditis elegans. Genes & development. 2003;17(8):991-1008.

[86]   Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP. Vertebrate microRNA genes. Science. 2003;299(5612):1540.

[87]   Terai G, Okida H, Asai K, Mituyama T. Prediction of conserved precursors of miRNAs and their mature forms by integrating position-specific structural features. PloS one. 2012;7(9):e44314.

[88]   Bonnet E, Wuyts J, Rouze P, Van de Peer Y. Detection of 91 potential conserved plant microRNAs in Arabidopsis thaliana and Oryza sativa identifies important target genes. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(31):11511-6.

[89]   Oulas A, Boutla A, Gkirtzou K, Reczko M, Kalantidis K, Poirazi P. Prediction of novel microRNA genes in cancer-associated genomic regions--a combined computational and experimental approach. Nucleic acids research. 2009;37(10):3276-87.

[90]   Liu CG, Calin GA, Volinia S, Croce CM. MicroRNA expression profiling using microarrays. Nature protocols. 2008;3(4):563-78.

[91]   Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nature methods. 2007;4(9):721-6.

[92]   Vermeulen A, Robertson B, Dalby AB, Marshall WS, Karpilow J, Leake D, *et al*. Double-stranded regions are essential design components of potent inhibitors of RISC function. RNA. 2007;13(5):723-30.

[93]   Robertson B, Dalby AB, Karpilow J, Khvorova A, Leake D, Vermeulen A. Specificity and functionality of microRNA inhibitors. Silence. 2010;1(1):10.

[94]   Santhakumar D, Forster T, Laqtom NN, Fragkoudis R, Dickinson P, Abreu-Goodger C, *et al*. Combined agonist-antagonist genome-wide functional screening identifies broadly active antiviral microRNAs. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(31):13830-5.

[95]   Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell. 2009;136(2):215-33.

[96]   Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Molecular cell. 2007;27(1):91-105.

[97]   Long D, Lee R, Williams P, Chan CY, Ambros V, Ding Y. Potent effect of target structure on microRNA function. Nature structural & molecular biology. 2007;14(4):287-94.

[98]   Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. Nucleic acids research. 2008;36(Database issue):D154-8.

[99]   John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human MicroRNA targets. PLoS Biol. 2004;2(11):e363.

[100] Alexiou P, Maragkakis M, Papadopoulos GL, Reczko M, Hatzigeorgiou AG. Lost in translation: an assessment and perspective for computational microRNA target identification. Bioinformatics. 2009;25(23):3049-55.

[101] Thomas M, Lieberman J, Lal A. Desperately seeking microRNA targets. Nat Struct Mol Biol. 2010;17(10):1169-74.

[102] Hendrickson DG, Hogan DJ, Herschlag D, Ferrell JE, Brown PO. Systematic identification of mRNAs recruited to argonaute 2 by specific microRNAs and corresponding changes in transcript abundance. PloS one. 2008;3(5):e2126.

[103] Aldred SF, Collins P, Trinklein N. Identifying targets of human micrornas with the LightSwitch Luciferase Assay System using 3'UTR-reporter constructs and a microRNA mimic in adherent cells. Journal of visualized experiments : JoVE. 2011(55).

[104] Smith BD, Morgan RL, Beckett GA, Falck-Ytter Y, Holtzman D, Teo CG, *et al*. Recommendations for the identification of chronic hepatitis C virus infection among persons born during 1945-1965. MMWR Recommendations and reports : Morbidity and mortality weekly report Recommendations and reports / Centers for Disease Control. 2012;61(RR-4):1-32.

[105] El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology. 2007;132(7):2557-76.

[106] Venook AP, Papandreou C, Furuse J, de Guevara LL. The incidence and epidemiology of hepatocellular carcinoma: a global and regional perspective. The oncologist. 2010;15 Suppl 4:5-13.

[107] Bruni R, Marcantonio C, Tritarelli E, Tataseo P, Stellacci E, Costantino A, *et al*. An integrated approach identifies IFN-regulated microRNAs and targeted mRNAs modulated by different HCV replicon clones. BMC genomics. 2011;12:485.

[108] Bala S, Tilahun Y, Taha O, Alao H, Kodys K, Catalano D, *et al*. Increased microRNA-155 expression in the serum and peripheral monocytes in chronic HCV infection. Journal of translational medicine. 2012;10(1):151 %U http://www.translational-medicine.com/content/10/1/.

[109] Bhanja Chowdhury J, Shrivastava S, Steele R, Di Bisceglie AM, Ray R, Ray RB. Hepatitis C virus infection modulates expression of interferon stimulatory gene IFITM1 by upregulating miR-130A. Journal of virology. 2012;86(18):10221-5 %U http://www.ncbi.nlm.nih.gov/pubmed/22787204.

[110] Raychoudhuri A, Shrivastava S, Steele R, Kim H, Ray R, Ray RB. ISG56 and IFITM1 proteins inhibit hepatitis C virus replication. J Virol. 2011;85(24):12881-9.

[111] Wilkins C, Woodward J, Lau DT, Barnes A, Joyce M, McFarlane N, *et al*. IFITM1 is a tight junction protein that inhibits hepatitis C virus entry. Hepatology. 2013;57(2):461-9.

[112] Li S, Xing X, Yang Q, Xu H, He J, Chen Z, *et al*. The effects of hepatitis C virus core protein on the expression of miR-122 *in vitro*. Virology journal. 2013;10:98- %U http://www.ncbi.nlm.nih.gov /pmc/articles/PMC3623887/.

[113] Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. Science. 2005;309(5740):1577-81.

[114] Conrad KD, Giering F, Erfurth C, Neumann A, Fehr C, Meister G, *et al*. microRNA-122 Dependent Binding of Ago2 Protein to Hepatitis C Virus RNA Is Associated with Enhanced RNA Stability and Translation Stimulation. PloS one. 2013;8(2 %U http://www.ncbi.nlm.nih.gov/pmc/articles/ PMC3566042/).

[115] Li Y, Masaki T, Lemon SM. miR-122 and the Hepatitis C RNA genome: More than just stability. RNA biology. 2013;10(6 %U http://www.ncbi.nlm.nih.gov/pubmed/23770926).

[116] Mortimer SA, Doudna JA. Unconventional miR-122 binding stabilizes the HCV genome by forming a trimolecular RNA structure. Nucleic Acids Research. 2013;41(7):4230-40 %U http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3627571/.

[117] Sagan SM, Sarnow P, Wilson JA. Modulation of GB virus B (GBV-B) RNA abundance by microRNA-122: Dependence on and escape from microRNA-122 restriction. %U http://www.ncbi.nlm.nih.gov/pubmed/23616647. Journal of virology. 2013.

[118] Fukuhara T, Tani H, Shiokawa M, Goto Y, Abe T, Taketomi A, *et al*. Intracellular delivery of serum-derived hepatitis C virus. Microbes and infection / Institut Pasteur. 2011;13(4):405-12.

[119] Kambara H, Fukuhara T, Shiokawa M, Ono C, Ohara Y, Kamitani W, *et al*. Establishment of a novel permissive cell line for the propagation of hepatitis C virus by expression of microRNA miR122. J Virol. 2012;86(3):1382-93.

[120] Narbus CM, Israelow B, Sourisseau M, Michta ML, Hopcraft SE, Zeiner GM, *et al*. HepG2 cells expressing microRNA miR-122 support the entire hepatitis C virus life cycle. J Virol. 2011;85(22):12087-92.

[121] Sainz B, Jr., Barretto N, Yu X, Corcoran P, Uprichard SL. Permissiveness of human hepatoma cell lines for HCV infection. Virol J. 2012;9:30.

[122] Jopling CL, Schütz S, Sarnow P. Position-dependent function for a tandem microRNA miR-122-binding site located in the hepatitis C virus RNA genome. Cell host & microbe. 2012;4(1):77-85 %U http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3519368/.

[123]   Elmen J, Lindow M, Schutz S, Lawrence M, Petri A, Obad S, *et al*. LNA-mediated microRNA silencing in non-human primates. Nature. 2008;452(7189):896-9.

[124]   Lanford RE, Hildebrandt-Eriksen ES, Petri A, Persson R, Lindow M, Munk ME, *et al*. Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. Science (New York, NY). 2012;327(5962):198-201 %U http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3436126/.

[125]   Janssen HLA, Reesink HW, Lawitz EJ, Zeuzem S, Rodriguez-Torres M, Patel K, *et al*. Treatment of HCV Infection by Targeting MicroRNA. %U http://www.ncbi.nlm.nih.gov/pubmed/23534542. The New England journal of medicine. 2013.

[126]   Li A, Song W, Qian J, Li Y, He J, Zhang Q, *et al*. MiR-122 modulates type I interferon expression through blocking suppressor of cytokine signaling 1. The international journal of biochemistry & cell biology. 2013;45(4):858-65 %U http://www.ncbi.nlm.nih.gov/pubmed/23348614.

[127]   Pager CT, Schütz S, Abraham TM, Luo G, Sarnow P. Modulation of hepatitis C virus RNA abundance and virus release by dispersion of processing bodies and enrichment of stress granules. Virology. 2013;435(2):472-84 %U http://www.ncbi.nlm.nih.gov/pubmed/23141719.

[128]   Chen Y, Chen J, Wang H, Shi J, Wu K, Liu S, *et al*. HCV-Induced miR-21 Contributes to Evasion of Host Immune System by Targeting MyD88 and IRAK1. PLoS pathogens. 2013;9(4 %U http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3635988/).

[129]   Shirasaki T, Honda M, Shimakami T, Horii R, Yamashita T, Sakai Y, *et al*. MicroRNA-27a regulates lipid metabolism and inhibits hepatitis C virus replication in human hepatoma cells. %U http://www.ncbi.nlm.nih.gov/pubmed/23449803. Journal of virology. 2013.

[130]   Shiu T-Y, Huang S-M, Shih Y-L, Chu H-C, Chang W-K, Hsieh T-Y. Hepatitis C Virus Core Protein Down-Regulates p21(Waf1/Cip1) and Inhibits Curcumin-Induced Apoptosis through MicroRNA-345 Targeting in Human Hepatoma Cells. PloS one. 2013;8(4):e61089 %U http://www.ncbi.nlm.nih.gov/pubmed/23577194.

[131]   UNAIDS. UNAIDS global report. 2012.

[132]   Gupta A, Nagilla P, Le HS, Bunney C, Zych C, Thalamuthu A, *et al*. Comparative expression profile of miRNA and mRNA in primary peripheral blood mononuclear cells infected with human immunodeficiency virus (HIV-1). PloS one. 2011;6(7):e22730.

[133]   Strebel K, Luban J, Jeang KT. Human cellular restriction factors that target HIV-1 replication. BMC medicine. 2009;7:48.

[134]   Mallory AC, Reinhart BJ, Bartel D, Vance VB, Bowman LH. A viral suppressor of RNA silencing differentially regulates the accumulation of short interfering RNAs and micro-RNAs in tobacco. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(23):15228-33.

[135]   Zhou Z, Dell'Orco M, Saldarelli P, Turturo C, Minafra A, Martelli GP. Identification of an RNA-silencing suppressor in the genome of Grapevine virus A. The Journal of general virology. 2006;87(Pt 8):2387-95.

[136]   Zhang X, Yuan YR, Pei Y, Lin SS, Tuschl T, Patel DJ, *et al*. Cucumber mosaic virus-encoded 2b suppressor inhibits Arabidopsis Argonaute1 cleavage activity to counter plant defense. Genes & development. 2006;20(23):3255-68.

[137]   Hemmes H, Lakatos L, Goldbach R, Burgyan J, Prins M. The NS3 protein of Rice hoja blanca tenuivirus suppresses RNA silencing in plant and insect hosts by efficiently binding both siRNAs and miRNAs. RNA. 2007;13(7):1079-89.

[138]   Abe M, Suzuki H, Nishitsuji H, Shida H, Takaku H. Interaction of human T-cell lymphotropic virus type I Rex protein with Dicer suppresses RNAi silencing. FEBS letters. 2010;584(20):4313-8.

[139]   Haasnoot J, de Vries W, Geutjes EJ, Prins M, de Haan P, Berkhout B. The Ebola virus VP35 protein is a suppressor of RNA silencing. PLoS pathogens. 2007;3(6):e86.

[140]   Bennasser Y, Jeang KT. HIV-1 Tat interaction with Dicer: requirement for RNA. Retrovirology. 2006;3:95.

[141]   Bennasser Y, Le SY, Benkirane M, Jeang KT. Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing. Immunity. 2005;22(5):607-19.

[142]   Lin J, Cullen BR. Analysis of the interaction of primate retroviruses with the human RNA interference machinery. J Virol. 2007;81(22):12218-26.

[143]   Bennasser Y, Yeung ML, Jeang KT. HIV-1 TAR RNA subverts RNA interference in transfected cells through sequestration of TAR RNA-binding protein, TRBP. The Journal of biological chemistry. 2006;281(38):27674-8.

[144]   Sanghvi VR, Steel LF. A re-examination of global suppression of RNA interference by HIV-1. PloS one. 2011;6(2):e17246.

[145]   Coley W, Van Duyne R, Carpio L, Guendel I, Kehn-Hall K, Chevalier S, *et al*. Absence of DICER in monocytes and its regulation by HIV-1. The Journal of biological chemistry. 2010;285(42):31930-43.

[146]   Nathans R, Chu CY, Serquina AK, Lu CC, Cao H, Rana TM. Cellular microRNA and P bodies modulate host-HIV-1 interactions. Molecular cell. 2009;34(6):696-709.

[147]   Triboulet R, Mari B, Lin YL, Chable-Bessia C, Bennasser Y, Lebrigand K, *et al*. Suppression of microRNA-silencing pathway by HIV-1 during virus replication. Science. 2007;315(5818):1579-82.

[148]   Yeung ML, Bennasser Y, Myers TG, Jiang G, Benkirane M, Jeang KT. Changes in microRNA expression profiles in HIV-1-transfected human cells. Retrovirology. 2005;2:81.

[149]   Houzet L, Yeung ML, de Lame V, Desai D, Smith SM, Jeang KT. MicroRNA profile changes in human immunodeficiency virus type 1 (HIV-1) seropositive individuals. Retrovirology. 2008;5:118.

[150]   Hayes AM, Qian S, Yu L, Boris-Lawrie K. Tat RNA silencing suppressor activity contributes to perturbation of lymphocyte miRNA by HIV-1. Retrovirology. 2011;8:36.

[151]   Huang J, Wang F, Argyris E, Chen K, Liang Z, Tian H, *et al*. Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes. Nat Med. 2007;13(10):1241-7.

[152]   Ahluwalia JK, Khan SZ, Soni K, Rawat P, Gupta A, Hariharan M, *et al*. Human cellular microRNA hsa-miR-29a interferes with viral nef protein expression and HIV-1 replication. Retrovirology. 2008;5:117.

[153]   Sung TL, Rice AP. miR-198 inhibits HIV-1 gene expression and replication in monocytes and its mechanism of action appears to involve repression of cyclin T1. PLoS pathogens. 2009;5(1):e1000263.

[154]   Baur AS. HIV-Nef and AIDS pathogenesis: are we barking up the wrong tree? Trends in microbiology. 2011;19(9):435-40.

[155]   Deacon NJ, Tsykin A, Solomon A, Smith K, Ludford-Menting M, Hooker DJ, *et al*. Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. Science. 1995;270(5238):988-91.

[156]   Eulalio A, Behm-Ansmant I, Izaurralde E. P bodies: at the crossroads of post-transcriptional pathways. Nature reviews Molecular cell biology. 2007;8(1):9-22.

[157]   Chable-Bessia C, Meziane O, Latreille D, Triboulet R, Zamborlini A, Wagschal A, *et al*. Suppression of HIV-1 replication by microRNA effectors. Retrovirology. 2009;6:26.

[158]   Baulcombe D. RNA silencing in plants. Nature. 2004;431(7006):356-63.

[159]   Cullen BR. Viral and cellular messenger RNA targets of viral microRNAs. Nature. 2009;457(7228):421-5.

[160]   Kaul D, Ahlawat A, Gupta SD. HIV-1 genome-encoded hiv1-mir-H1 impairs cellular responses to infection. Molecular and cellular biochemistry. 2009;323(1-2):143-8.

[161]   Omoto S, Fujii YR. Regulation of human immunodeficiency virus 1 transcription by nef microRNA. The Journal of general virology. 2005;86(Pt 3):751-5.

[162]   Klase Z, Kale P, Winograd R, Gupta MV, Heydarian M, Berro R, *et al*. HIV-1 TAR element is processed by Dicer to yield a viral micro-RNA involved in chromatin remodeling of the viral LTR. BMC molecular biology. 2007;8:63.

[163]   Ouellet DL, Plante I, Landry P, Barat C, Janelle ME, Flamand L, *et al*. Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element. Nucleic acids research. 2008;36(7):2353-65.

[164]   Purzycka KJ, Adamiak RW. The HIV-2 TAR RNA domain as a potential source of viral-encoded miRNA. A reconnaissance study. Nucleic acids symposium series. 2008(52):511-2.

[165]   Klase Z, Winograd R, Davis J, Carpio L, Hildreth R, Heydarian M, *et al*. HIV-1 TAR miRNA protects against apoptosis by altering cellular gene expression. Retrovirology. 2009;6:18.

[166]   Yeung ML, Bennasser Y, Watashi K, Le SY, Houzet L, Jeang KT. Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence for the processing of a viral-cellular double-stranded RNA hybrid. Nucleic acids research. 2009;37(19):6575-86.

[167]   Schopman NC, Willemsen M, Liu YP, Bradley T, van Kampen A, Baas F, *et al*. Deep sequencing of virus-infected cells reveals HIV-encoded small RNAs. Nucleic acids research. 2012;40(1):414-27.

[168]   CDC. Respiratory syncytial virus activity-- United States, July 2007-December 2008. MMWR Morb Mortal Wkly Rep. 2008;57(50):1355-8. Epub 2008/12/19.

[169] Li Q, Brass AL, Ng A, Hu Z, Xavier RJ, Liang TJ, *et al*. A genome-wide genetic screen for host factors required for hepatitis C virus propagation. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(38):16410-5. Epub 2009/09/01.

[170] Meliopoulos VA, Andersen LE, Birrer KF, Simpson KJ, Lowenthal JW, Bean AG, *et al*. Host gene targets for novel influenza therapies elucidated by high-throughput RNA interference screens. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2012;26(4):1372-86. Epub 2012/01/17.

[171] Panda D, Das A, Dinh PX, Subramaniam S, Nayak D, Barrows NJ, *et al*. RNAi screening reveals requirement for host cell secretory pathway in infection by diverse families of negative-strand RNA viruses. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(47):19036-41. Epub 2011/11/09.

[172] Alvarez R, Elbashir S, Borland T, Toudjarska I, Hadwiger P, John M, *et al*. RNA interference-mediated silencing of the respiratory syncytial virus nucleocapsid defines a potent antiviral strategy. Antimicrob Agents Chemother. 2009;53(9):3952-62. Epub 2009/06/10.

[173] Spann KM, Tran KC, Collins PL. Effects of nonstructural proteins NS1 and NS2 of human respiratory syncytial virus on interferon regulatory factor 3, NF-kappaB, and proinflammatory cytokines. J Virol. 2005;79(9):5353-62. Epub 2005/04/14.

[174] Bossert B, Marozin S, Conzelmann KK. Nonstructural proteins NS1 and NS2 of bovine respiratory syncytial virus block activation of interferon regulatory factor 3. J Virol. 2003;77(16):8661-8. Epub 2003/07/30.

[175] Lo MS, Brazas RM, Holtzman MJ. Respiratory syncytial virus nonstructural proteins NS1 and NS2 mediate inhibition of Stat2 expression and alpha/beta interferon responsiveness. J Virol. 2005;79(14):9315-9. Epub 2005/07/05.

[176] Elliott J, Lynch OT, Suessmuth Y, Qian P, Boyd CR, Burrows JF, *et al*. Respiratory syncytial virus NS1 protein degrades STAT2 by using the Elongin-Cullin E3 ligase. J Virol. 2007;81(7):3428-36. Epub 2007/01/26.

[177] Tayyari F, Marchant D, Moraes TJ, Duan W, Mastrangelo P, Hegele RG. Identification of nucleolin as a cellular receptor for human respiratory syncytial virus. Nat Med. 2011. Epub 2011/08/16.

[178] Oshansky CM, Krunkosky TM, Barber J, Jones LP, Tripp RA. Respiratory syncytial virus proteins modulate suppressors of cytokine signaling 1 and 3 and the type I interferon response to infection by a toll-like receptor pathway. Viral Immunol. 2009;22(3):147-61. Epub 2009/05/14.

[179] Murawski MR, Bowen GN, Cerny AM, Anderson LJ, Haynes LM, Tripp RA, *et al*. Respiratory syncytial virus activates innate immunity through Toll-like receptor 2. J Virol. 2009;83(3):1492-500. Epub 2008/11/21.

[180] Kurt-Jones EA, Popova L, Kwinn L, Haynes LM, Jones LP, Tripp RA, *et al*. Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. Nature immunology. 2000;1(5):398-401. Epub 2001/03/23.

[181] Moore EC, Barber J, Tripp RA. Respiratory syncytial virus (RSV) attachment and nonstructural proteins modify the type I interferon response associated with suppressor of cytokine signaling (SOCS) proteins and IFN-stimulated gene-15 (ISG15). Virol J. 2008;5:116. Epub 2008/10/15.

[182] Tripp RA, Moore D, Jones L, Sullender W, Winter J, Anderson LJ. Respiratory syncytial virus G and/or SH protein alters Th1 cytokines, natural killer cells, and neutrophils responding to pulmonary infection in BALB/c mice. J Virol. 1999;73(9):7099-107. Epub 1999/08/10.

[183] Tripp RA, Jones L, Anderson LJ. Respiratory syncytial virus G and/or SH glycoproteins modify CC and CXC chemokine mRNA expression in the BALB/c mouse. J Virol. 2000;74(13):6227-9. Epub 2000/06/14.

[184] Tripp RA, Jones LP, Haynes LM, Zheng H, Murphy PM, Anderson LJ. CX3C chemokine mimicry by respiratory syncytial virus G glycoprotein. Nature immunology. 2001;2(8):732-8. Epub 2001/07/31.

[185] Martinez I, Lombardia L, Garcia-Barreno B, Dominguez O, Melero JA. Distinct gene subsets are induced at different time points after human respiratory syncytial virus infection of A549 cells. The Journal of general virology. 2007;88(Pt 2):570-81. Epub 2007/01/26.

[186] Munday DC, Emmott E, Surtees R, Lardeau CH, Wu W, Duprex WP, *et al*. Quantitative proteomic analysis of A549 cells infected with human respiratory syncytial virus. Mol Cell Proteomics. 2010;9(11):2438-59. Epub 2010/07/22.

[187]   Gibbs JD, Ornoff DM, Igo HA, Zeng JY, Imani F. Cell cycle arrest by transforming growth factor beta1 enhances replication of respiratory syncytial virus in lung epithelial cells. Journal of virology. 2009;83(23):12424-31. Epub 2009/09/18.

[188]   Mohapatra S, Park SJ, Boyapalle S, Pastey MK, Graham BS, Blanck G. Human respiratory syncytial virus reduces the number of cells in S-phase and increases GADD153 expression in HEp-2 cells. Acta Virol. 2009;53(3):207-11. Epub 2009/12/04.

[189]   Lindquist ME, Lifland AW, Utley TJ, Santangelo PJ, Crowe JE, Jr. Respiratory syncytial virus induces host RNA stress granules to facilitate viral replication. Journal of virology. 2010;84(23):12274-84. Epub 2010/09/17.

[190]   Lindquist ME, Mainou BA, Dermody TS, Crowe JE, Jr. Activation of protein kinase R is required for induction of stress granules by respiratory syncytial virus but dispensable for viral replication. Virology. 2011;413(1):103-10. Epub 2011/03/08.

[191]   Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem. 2010;79:351-79.

[192]   Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. Widespread changes in protein synthesis induced by microRNAs. Nature. 2008;455(7209):58-63.

[193]   Othumpangat S, Walton C, Piedimonte G. MicroRNA-221 Modulates RSV Replication in Human Bronchial Epithelium by Targeting NGF Expression. PloS one. 2012;7(1):e30030. Epub 2012/01/25.

[194]   Zhang J, Han L, Ge Y, Zhou X, Zhang A, Zhang C, *et al*. miR-221/222 promote malignant progression of glioma through activation of the Akt pathway. International journal of oncology. 2010;36(4):913-20. Epub 2010/03/04.

[195]   Terasawa K, Ichimura A, Sato F, Shimizu K, Tsujimoto G. Sustained activation of ERK1/2 by NGF induces microRNA-221 and 222 in PC12 cells. The FEBS journal. 2009;276(12):3269-76. Epub 2009/05/15.

[196]   Zhang C, Zhang J, Zhang A, Wang Y, Han L, You Y, *et al*. PUMA is a novel target of miR-221/222 in human epithelial cancers. International journal of oncology. 2010;37(6):1621-6. Epub 2010/11/03.

[197]   Hamada N, Fujita Y, Kojima T, Kitamoto A, Akao Y, Nozawa Y, *et al*. MicroRNA expression profiling of NGF-treated PC12 cells revealed a critical role for miR-221 in neuronal differentiation. Neurochemistry international. 2012;60(8):743-50. Epub 2012/04/03.

[198]   Bakre A, Mitchell P, Coleman JK, Jones LP, Saavedra G, Teng M, *et al*. Respiratory syncytial virus modifies microRNAs regulating host genes that affect virus replication. J Gen Virol. 2012;93(Pt 11):2346-56. Epub 2012/08/17.

[199]   Skalsky RL, Cullen BR. Viruses, microRNAs, and host interactions. Annual review of microbiology. 2010;64:123-41.

[200]   Jurak I, Kramer MF, Mellor JC, van Lint AL, Roth FP, Knipe DM, *et al*. Numerous conserved and divergent microRNAs expressed by herpes simplex viruses 1 and 2. J Virol. 2010;84(9):4659-72.

[201]   Umbach JL, Kramer MF, Jurak I, Karnowski HW, Coen DM, Cullen BR. MicroRNAs expressed by herpes simplex virus 1 during latent infection regulate viral mRNAs. Nature. 2008;454(7205):780-3.

[202]   Umbach JL, Wang K, Tang S, Krause PR, Mont EK, Cohen JI, *et al*. Identification of viral microRNAs expressed in human sacral ganglia latently infected with herpes simplex virus 2. J Virol. 2010;84(2):1189-92.

[203]   Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grasser FA, *et al*. Identification of microRNAs of the herpesvirus family. Nature methods. 2005;2(4):269-76.

[204]   Stark TJ, Arnold JD, Spector DH, Yeo GW. High-resolution profiling and analysis of viral and host small RNAs during human cytomegalovirus infection. J Virol. 2012;86(1):226-35.

[205]   Lee S, Song J, Kim S, Kim J, Hong Y, Kim Y, *et al*. Selective Degradation of Host MicroRNAs by an Intergenic HCMV Noncoding RNA Accelerates Virus Production. Cell host & microbe. 2013;13(6):678-90.

[206]   Dolken L, Krmpotic A, Kothe S, Tuddenham L, Tanguy M, Marcinowski L, *et al*. Cytomegalovirus microRNAs facilitate persistent virus infection in salivary glands. PLoS pathogens. 2010;6(10):e1001150.

[207]   Pfeffer S, Zavolan M, Grasser FA, Chien M, Russo JJ, Ju J, *et al*. Identification of virus-encoded microRNAs. Science. 2004;304(5671):734-6.

[208]   Ramalingam D, Kieffer-Kwon P, Ziegelbauer JM. Emerging themes from EBV and KSHV microRNA targets. Viruses. 2012;4(9):1687-710.

[209] Zur Hausen H. The search for infectious causes of human cancers: where and why. Virology. 2009;392(1):1-10.

[210] Walz N, Christalla T, Tessmer U, Grundhoff A. A global analysis of evolutionary conservation among known and predicted gammaherpesvirus microRNAs. J Virol. 2010;84(2):716-28.

[211] Xia T, O'Hara A, Araujo I, Barreto J, Carvalho E, Sapucaia JB, *et al*. EBV microRNAs in primary lymphomas and targeting of CXCL-11 by ebv-mir-BHRF1-3. Cancer research. 2008;68(5):1436-42.

[212] Nachmani D, Stern-Ginossar N, Sarid R, Mandelboim O. Diverse herpesvirus microRNAs target the stress-induced immune ligand MICB to escape recognition by natural killer cells. Cell host & microbe. 2009;5(4):376-85.

[213] Skalsky RL, Corcoran DL, Gottwein E, Frank CL, Kang D, Hafner M, *et al*. The viral and cellular microRNA targetome in lymphoblastoid cell lines. PLoS pathogens. 2012;8(1):e1002484.

[214] Ziegelbauer JM, Sullivan CS, Ganem D. Tandem array-based expression screens identify host mRNA targets of virus-encoded microRNAs. Nature genetics. 2009;41(1):130-4.

[215] Skalsky RL, Samols MA, Plaisance KB, Boss IW, Riva A, Lopez MC, *et al*. Kaposi's sarcoma-associated herpesvirus encodes an ortholog of miR-155. J Virol. 2007;81(23):12836-45.

[216] Marquitz AR, Mathur A, Nam CS, Raab-Traub N. The Epstein-Barr Virus BART microRNAs target the pro-apoptotic protein Bim. Virology. 2011;412(2):392-400.

[217] Suffert G, Malterer G, Hausser J, Viiliainen J, Fender A, Contrant M, *et al*. Kaposi's sarcoma herpesvirus microRNAs target caspase 3 and regulate apoptosis. PLoS pathogens. 2011;7(12):e1002405.

[218] Li Z, Chen X, Li L, Liu S, Yang L, Ma X, *et al*. EBV encoded miR-BHRF1-1 potentiates viral lytic replication by downregulating host p53 in nasopharyngeal carcinoma. Int J Biochem Cell Biol. 2012;44(2):275-9.

[219] Choy EY, Siu KL, Kok KH, Lung RW, Tsang CM, To KF, *et al*. An Epstein-Barr virus-encoded microRNA targets PUMA to promote host cell survival. The Journal of experimental medicine. 2008;205(11):2551-60.

[220] Gottwein E, Cullen BR. A human herpesvirus microRNA inhibits p21 expression and attenuates p21-mediated cell cycle arrest. J Virol. 2010;84(10):5229-37.

[221] Di Bartolo DL, Cannon M, Liu YF, Renne R, Chadburn A, Boshoff C, *et al*. KSHV LANA inhibits TGF-beta signaling through epigenetic silencing of the TGF-beta type II receptor. Blood. 2008;111(9):4731-40.

[222] Forte E, Salinas RE, Chang C, Zhou T, Linnstaedt SD, Gottwein E, *et al*. The Epstein-Barr virus (EBV)-induced tumor suppressor microRNA MiR-34a is growth promoting in EBV-infected B cells. J Virol. 2012;86(12):6889-98.

[223] Forte E, Luftig MA. The role of microRNAs in Epstein-Barr virus latency and lytic reactivation. Microbes and infection / Institut Pasteur. 2011;13(14-15):1156-67.

[224] Iizasa H, Wulff BE, Alla NR, Maragkakis M, Megraw M, Hatzigeorgiou A, *et al*. Editing of Epstein-Barr virus-encoded BART6 microRNAs controls their dicer targeting and consequently affects viral latency. The Journal of biological chemistry. 2010;285(43):33358-70.

[225] Barth S, Pfuhl T, Mamiani A, Ehses C, Roemer K, Kremmer E, *et al*. Epstein-Barr virus-encoded microRNA miR-BART2 down-regulates the viral DNA polymerase BALF5. Nucleic acids research. 2008;36(2):666-75.

[226] Lei X, Bai Z, Ye F, Xie J, Kim CG, Huang Y, *et al*. Regulation of NF-kappaB inhibitor IkappaBalpha and viral replication by a KSHV microRNA. Nature cell biology. 2010;12(2):193-9.

[227] Liang D, Gao Y, Lin X, He Z, Zhao Q, Deng Q, *et al*. A human herpesvirus miRNA attenuates interferon signaling and contributes to maintenance of viral latency by targeting IKKepsilon. Cell research. 2011;21(5):793-806.

[228] Lu F, Stedman W, Yousef M, Renne R, Lieberman PM. Epigenetic regulation of Kaposi's sarcoma-associated herpesvirus latency by virus-encoded microRNAs that target Rta and the cellular Rbl2-DNMT pathway. J Virol. 2010;84(6):2697-706.

[229] Thompson WW, Shay DK, Weintraub E, Brammer L, Bridges CB, Cox NJ, *et al*. Influenza-associated hospitalizations in the United States. JAMA. 2004;292(11):1333-40. Epub 2004/09/16.

[230] Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, *et al*. Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA. 2003;289(2):179-86. Epub 2003/01/09.

[231]   MMWR. Estimates of Deaths Associated with Seasonal Influenza - United States, 1976–2007. Atlanta, GA: Centers for Disease Control and Prevention, 2010.

[232]   Regoes RR, Bonhoeffer S. Emergence of drug-resistant influenza virus: population dynamical considerations. Science. 2006;312(5772):389-91. Epub 2006/04/22.

[233]   Palese P SM. Fields Virology. 5th ed. Knipe DM HP, editor. Philadelphia: Raven; 2007.

[234]   He T, Feng G, Chen H, Wang L, Wang Y. Identification of host encoded microRNAs interacting with novel swine-origin influenza A (H1N1) virus and swine influenza virus. Bioinformation. 2009;4(3):112-8 %U http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2828889/.

[235]   Perez JT, Pham AM, Lorini MH, Chua MA, Steel J, tenOever BR. MicroRNA-mediated species-specific attenuation of influenza A virus. Nature biotechnology. 2009;27(6):572-6 %U http://www.ncbi.nlm.nih.gov/pubmed/19483680.

[236]   Wang Y, Brahmakshatriya V, Zhu H, Lupiani B, Reddy SM, Yoon B-J, *et al*. Identification of differentially expressed miRNAs in chicken lung and trachea with avian influenza virus infection by a deep sequencing approach. BMC genomics. 2009;10:512- %U http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2777939/.

[237]   Li Y, Chan EY, Li J, Ni C, Peng X, Rosenzweig E, *et al*. MicroRNA expression and virulence in pandemic influenza virus-infected mice. Journal of virology. 2010;84(6):3023-32 %U http://www.ncbi.nlm.nih.gov/pubmed/20071585.

[238]   Song L, Liu H, Gao S, Jiang W, Huang W. Cellular microRNAs inhibit replication of the H1N1 influenza A virus in infected cells. Journal of virology. 2010;84(17):8849-60 %U http://www.ncbi.nlm.nih.gov/pubmed/20554777.

[239]   Umbach JL, Yen H-L, Poon LLM, Cullen BR. Influenza A virus expresses high levels of an unusual class of small viral leader RNAs in infected cells. mBio. 2010;1(4 %U http://www.ncbi.nlm.nih.gov/pubmed/20842206).

[240]   Buggele WA, Johnson KE, Horvath CM. Influenza A virus infection of human respiratory cells induces primary microRNA expression. The Journal of biological chemistry. 2012;287(37):31027-40 %U http://www.ncbi.nlm.nih.gov/pubmed/22822053.

[241]   Fang J, Hao Q, Liu L, Li Y, Wu J, Huo X, *et al*. Epigenetic changes mediated by microRNA miR29 activate cyclooxygenase 2 and lambda-1 interferon production during viral infection. Journal of virology. 2012;86(2):1010-20 %U http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3255816/.

[242]   Guan Z, Shi N, Song Y, Zhang X, Zhang M, Duan M. Induction of the cellular microRNA-29c by influenza virus contributes to virus-mediated apoptosis through repression of antiapoptotic factors BCL2L2. Biochemical and biophysical research communications. 2012;425(3):662-7 %U http://www.ncbi.nlm.nih.gov/pubmed/22850539.

[243]   Meliopoulos VA, Andersen LE, Brooks P, Yan X, Bakre A, Coleman JK, *et al*. MicroRNA regulation of human protease genes essential for influenza virus replication. PloS one. 2012;7(5 %U http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3351457/).

[244]   Rosenberger CM, Podyminogin RL, Navarro G, Zhao G-W, Askovich PS, Weiss MJ, *et al*. miR-451 regulates dendritic cell cytokine responses to influenza infection. Journal of immunology (Baltimore, Md : 1950). 2012;189(12):5965-75 %U http://www.ncbi.nlm.nih.gov/pubmed/23169590.

[245]   Wang Y, Brahmakshatriya V, Lupiani B, Reddy SM, Soibam B, Benham AL, *et al*. Integrated analysis of microRNA expression and mRNA transcriptome in lungs of avian influenza virus infected broilers. BMC genomics. 2012;13(1):278 %U http://www.biomedcentral.com/1471-2164/13/278.

[246]   Lam W-Y, Yeung AC-M, Ngai KL-K, Li M-S, To K-F, Tsui SK-W, *et al*. Effect of avian influenza A H5N1 infection on the expression of microRNA-141 in human respiratory epithelial cells. BMC microbiology. 2013;13:104- %U http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3663648/.

[247]   Skovgaard K, Cirera S, Vasby D, Podolska A, Breum SØ, Dürrwald R, *et al*. Expression of innate immune genes, proteins and microRNAs in lung tissue of pigs infected experimentally with influenza virus (H1N2). %U http://www.ncbi.nlm.nih.gov/pubmed/23405029. Innate immunity. 2013.

[248]   Terrier O, Textoris J, Carron C, Marcel V, Bourdon J-C, Rosa-Calatrava M. Host microRNA molecular signatures associated with human H1N1 and H3N2 influenza A viruses reveal an unanticipated antiviral activity for miR-146a. %U http://www.ncbi.nlm.nih.gov/pubmed/23343627. The Journal of general virology. 2013.

[249]   Bakre A, Andersen LE, Meliopoulos V, Coleman K, Yan X, Brooks P, *et al*. Identification of Host Kinase Genes Required for Influenza Virus Replication and the Regulatory Role of MicroRNAs. PloS one. 2013;8(6):e66796.

CHAPTER 7

RNAi Screening in Cells of the Immune System: Challenges and Opportunities

Sinu P. John[1], Michael Freeley[2], Aideen Long[2] and Iain D.C. Fraser[1],*

[1]*Signaling Systems Unit, Laboratory of Systems Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA and* [2]*Department of Clinical Medicine, Trinity College Dublin, Dublin, Ireland*

Abstract: RNAi screening and the use of small silencing RNAs for specific gene knockdown has revolutionized basic science and translational medicine, both through the discovery of novel gene function and as a means to perturb disease-causing genes for therapeutic intervention. Availability of genome-wide RNAi libraries has made it possible to screen, in an unbiased manner, for all genes involved in any cellular process. This promises a more comprehensive understanding of complex cellular response networks, a fundamental goal in the emerging field of systems biology. Despite the obvious potential of this technology, cells of the immune system pose certain challenges for application of large scale RNAi screening, particularly in balancing the efficient delivery of silencing RNA while avoiding non-specific immune responses to the introduced nucleic acid. However, recent advancements in RNAi technology, improvements in delivery methods and the development of robust screening assays have made this technology more accessible to immunologists. Consequently, several examples of successful application of RNAi screening at both genome and sub-genome scales in immune cells are emerging, and are significantly advancing our knowledge of immune cell function. In this chapter, we outline the major challenges of using large scale RNAi screening in hematopoietic cells and describe different methodologies and assays that have been adopted for screening, with an emphasis on how these published studies have advanced our understanding of the immune system in health and disease. We conclude with a discussion of future opportunities and screening approaches that will realize the potential of RNAi screening in immune cells.

Keywords: Assay design, electroporation, Hematopoietic cells, immune cells, innate immune response, nucleofection, RNAi screening, siRNA delivery, viral shRNA.

INTRODUCTION

Increased understanding of small RNA biology following the discovery of RNA

***Corresponding to Iain D.C. Fraser:** Signaling Systems Unit, Laboratory of Systems Biology, Laboratory of Systems Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 9000 Rockville Pike, Bldg. 4, Rm. 109A, MSC-0421, Bethesda, MD 20892, USA; Tel: 301-443-5998; Fax:: 301-480-5725; E mail: fraseri@niaid.nih.gov

Interference (RNAi) has led to the development of large-scale libraries that allow researchers to rapidly screen for novel targets relevant to their area of interest. RNAi takes advantage of the endogenous microRNA pathway, permitting the silencing of mRNA transcripts by introduction of a short interfering RNA (siRNA) complementary to the target gene messenger RNA. Large-scale application of this technology using the robotic platforms originally developed for small molecule screens now permits genome-wide forward genetic screens in a wide range of mammalian cell types [1-3]. Accordingly, RNAi screening and the use of siRNAs for specific gene perturbation are revolutionizing both basic science and translational medicine, as they permit rapid and unbiased discovery of novel gene function and the manipulation of disease-causing genes for therapeutic intervention.

A wide range of cell types derived from hematopoietic stem cell precursors orchestrates the mammalian immune system. Hematopoietic development generates two branches of immune cells, a myeloid and a lymphoid branch (Fig. **1**). The lymphoid branch gives rise to B and T lymphocytes and natural killer (NK) cells, while the myeloid branch differentiates to all other blood cells including neutrophils and macrophages [4, 5]. Dendritic cells (DCs), another immune cell type, have been reported to develop from both myeloid and lymphoid lineages [6]. The application of genome-wide siRNA screening in hematopoietic cells has the potential to provide significant insight into mechanisms of immune cell activation, host-pathogen interactions, inflammation and tumor biology. Uncovering new components in immune cell function may also reveal novel targets for immuno-suppression, anti-viral and anti-cancer therapies.

However, since the key to successful siRNA screening is the efficient delivery of siRNA into target cells, the subsequent uptake by the RNA Induced Silencing Complex (RISC), and the degradation of the target gene mRNA, hematopoietic cells present some significant technical challenges in terms of siRNA delivery. Moreover, since certain immune cell types are primed to recognize and respond to pathogen-derived nucleic acids, researchers planning RNAi screening studies in immune cells must be particularly vigilant to ensure that RNA delivery protocols do not induce a non-specific interferon (IFN) response to the RNAi reagents.

In this chapter, we discuss the challenges of siRNA screening in cells of the immune system, outline available methodology for RNAi effector delivery and assay design, highlight examples where this technology has been used successfully for immune cell studies, and discuss future opportunities for the use of RNA screening to address immunological questions.
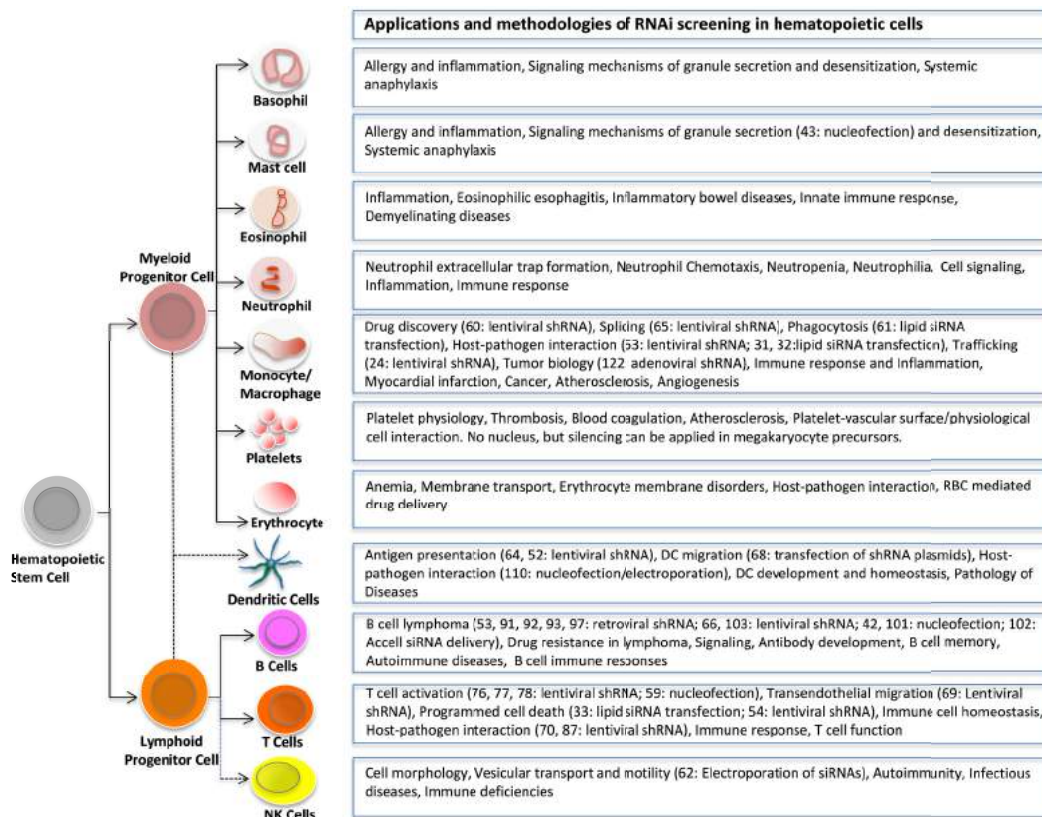
Figure 1: Applications and methodologies of RNAi screening in hematopoietic cells. Differentiated cell types of hematopoietic origin are shown diagrammatically (on left) with potential applications of RNAi screening for various cellular functions (on right). Already reported RNAi screens for specific functions are given in parentheses, with the respective references si/shRNA delivery method used.

CHALLENGES FOR RNAi SCREENING IN IMMUNE SYSTEM CELLS

Successful perturbation of target genes in RNAi screening depends on the efficient delivery of siRNA into the cell type under study. Transfection of siRNA using lipid-based reagents is the most popular delivery method and has been used successfully in a wide range of non-immune adherent mammalian cell types such as HeLa, U2OS and HEK293 cells [7-9]. However, many hematopoietic cell types are non-adherent in culture, and are often less tractable due to the low percentage of transfection achieved in these cells with currently available protocols. While lipid-based transfection of siRNA remains the most convenient approach to siRNA delivery whenever possible, viral vector mediated delivery of short hairpin

RNAs (shRNA) offers an alternative to lipid-mediated methods [10, 11]. However, compared to siRNAs, shRNA have higher variability in terms of silencing efficiency [1], and viral-mediated transduction may also increase the off-target effects due to higher cellular concentrations of shRNA [12, 13].

The possibility of a non-specific cellular response to the silencing RNA must also be considered, particularly in cells of the innate immune system. While the IFN response to double stranded RNA (dsRNA) was long established as an anti-viral defense strategy, the seminal studies with siRNA in mammalian cells suggested a size threshold of 19-21 nt under which dsRNA could promote gene silencing while avoiding an IFN-induced global inhibition of protein expression [14]. While this rule also appeared initially to hold true in immune cells, it was later discovered that certain immunostimulatory motifs in the small dsRNA could stimulate IFN responses that are particularly potent in certain innate immune cell types [15-19]. While all mammalian cells express cytosolic sensors responsive to dsRNA, such as retinoic acid-inducible gene-I (RIG-I), phagocytic immune cells, such as DCs and macrophages, are also primed to respond to pathogen-derived nucleic acids in endosomal compartments through a number of different toll-like receptors (TLR3, TLR7, TLR8 and TLR9) [20, 21]. The higher sensitivity of certain immune cells to dsRNA can therefore present an additional challenge in attempting to specifically perturb individual genes with siRNA while avoiding IFN induction. For example, TLR7/8 in human monocytes and PBMC have been shown to be activated by both single stranded as well as double stranded siRNAs, with single stranded RNA being comparatively more immuno-potent [22, 23]. This can result in the production of immune cytokines including IFNs, which in turn induce the production of antiviral proteins to inhibit protein synthesis [24]. Recent studies have also suggested that certain cationic lipids used for nucleic acid transfection can non-specifically activate cells and that certain immune cell types may be particularly sensitive to these effects [25]. It is thus vital for any RNAi screening endeavor in immune cells, particularly in cells expressing high levels of the described nucleic acid sensors, to carefully evaluate delivery protocols to identify conditions that avoid any significant non-specific responses to either the transfection reagents or the nucleic acid.

As with any siRNA/shRNA mediated gene knockdown, off target effects are also a major challenge in siRNA/shRNA based high throughput screens in hematopoietic cells, particularly through microRNA (miRNA)-like targeting of the 3'UTR of unintended target genes by the seed sequence of the siRNAs [26, 27]. Strategies to evaluate seed-based off target effects will be covered in detail

elsewhere in this volume, so we will not discuss them in detail here. However we would emphasize that while statistical and experimental validation of individual hits from the screen may help to minimize the off target effect, more retrospective quality control on commercially available siRNAs may be needed to remove library siRNAs with particularly promiscuous seed sequences [28, 29].

APPROACHES TO RNAi SCREENING IN IMMUNE SYSTEM CELLS

Successful siRNA screening depends not only on the efficient delivery of siRNA into the cells of choice, but also on the development of a specific and sensitive assay to reliably identify an altered phenotype in response to the silencing of genes in a high throughput format. The sections below first discuss the advantages and disadvantages of various dsRNA delivery methods (Fig. **2**), and then highlight considerations that should be taken into account for assay design.

Silencing Reagent Delivery

Common practice for siRNA delivery into cells uses either lipid-based transfection or viral mediated transduction. However, in 'hard to transfect' cells, such as primary cells and many hematopoietic cell lines, other alternatives have also been adopted to improve the delivery of siRNA, and these are also discussed below.

Lipid-Based Transfection

The most widely used method of lipid-based siRNA transfection is often not easily transferrable to primary hematopoietic cells and cell lines because of the lower efficiency of siRNA delivery. The advantage of this approach is that in a reverse-transfection format, where plates can be prepared and stored with a pre-arrayed genome-wide library, it is the most straightforward approach to conduct and interpret a large-scale screen. Although not practical with suspension cell types, it can be used with adherent innate immune cells such as macrophages or DCs if considerable effort is dedicated to delivery optimization in the plate format being used for the primary screening assay. Accordingly, lipid-mediated transfection has been reported for a screen in the mouse macrophage cell line, Raw 264.7 cells [30]. Among the various commercially available lipids tested, only Lipofectamine LTX, Lipofectamine RNAi Max and HiPerFect were able to deliver sufficient siRNA to induce significant silencing in this study [30]. Another lipid-based transfection screen has been reported in the mouse macrophage cell line J774.1 for host factors regulating mycobacterium tuberculosis infection [31].
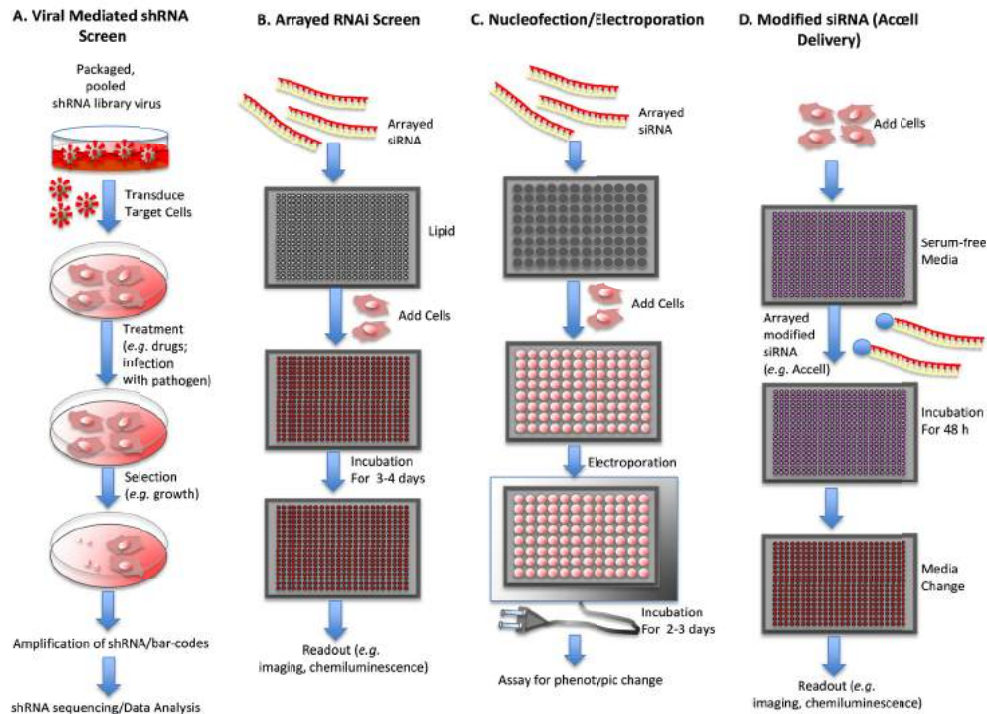
Figure 2: Available methods for RNA interference in immune cells. Schematic representation of commonly used methods for shRNA and siRNA delivery and screening workflows are shown. **A**. Viral mediated shRNA knockdown can utilize several different types of viral vector for shRNA expression, such as lentiviral, retroviral or adenoviral vectors. **B**. Arrayed RNAi screening can be used for siRNA (lipid-mediated) or shRNA (lipid-mediated or viral-mediated) delivery. **C**. Nucleofection/Electroporation is used to deliver silencing RNA (most commonly siRNA but can also be plasmid-expressed shRNA) through application of an electric field, which can be performed currently in 96-well format. **D**. Modified siRNAs or Accell delivery use modified siRNAs which can be delivered without a transfection reagent into otherwise hard to transfect cells**.**

In order to obtain prolonged knockdown, this group used two successive rounds of transfection into the same cells [31]. A siRNA screen with lipid-mediated siRNA delivery has also been conducted in the human macrophage cell line THP1, again for identification of host factors involved in the response to mycobacterium tuberculosis [32], while lipid-based siRNA transfection has also been used for screening in the Jurkat T cell line [33]. Despite these few published examples, achieving robust transfection efficiency of siRNA into hematopoietic cells using any commercially available lipid remains a significant challenge. However, it should be noted that if the target cells used in a screen stably express a reporter such as luciferase or green fluorescent protein (GFP) (see 'Screen assay

design' section below), this can provide a convenient target for optimizing siRNA delivery protocols. The ability to assay such reporters directly in the screening assay plate format provides two advantages. First, it allows measurement of protein rather than mRNA knockdown (the most common validation method for siRNA delivery), providing a more direct assessment of the required endpoint needed for an effective screen. Second, the ability to run the delivery optimization assays in 96 or 384-well format allows for a more extensive matrix of experimental conditions that improves the chance of identifying an optimal siRNA delivery protocol.

Electroporation/Nucleofection

An alternative to lipid-mediated transfection for siRNA delivery is electroporation or nucleofection, which has been shown to provide higher nucleic acid transfection efficiencies in many primary cell types and cell lines of hematopoietic origin [34-39]. Electroporation involves applying an electric field pulse to induce transient cell permeability *via* the formation of microscopic pores in the plasma membrane through which nucleic acids can traverse. If the electric field pulse is optimized carefully, most electroporated cells can recover to function normally [40]. Nucleofection® is an electroporation-based procedure consisting of a proprietary device, cell type-specific solutions and pre-set programs for the delivery of nucleotides into a wide variety of cell types, including 'hard-to-transfect' cells. A disadvantage of siRNA delivery *via* electroporation/nucleofection is its low throughput, although advances have been made to electroporate/nucleofect hematopoietic-derived cells in 96-well format [41-43]. It should be noted that using this approach, electroporation of primary T cells required a comparatively higher concentration of siRNA (10x more than for HeLa cells) for efficient gene silencing [44]. Requirement of a relatively higher concentration of siRNA, and often also a need for proprietary delivery buffers for optimal nucleic acid uptake and cell survival, can make this method considerably more expensive on a genome-wide scale. However, a potential advantage of this approach that can reduce the cost per assay is that larger numbers of cells (typically up to 1×10^6 cells per sample, depending on cell type) can be electroporated with each siRNA and then distributed into multiple plates to be used for numerous different downstream assays. This could potentially allow for more comprehensive biological readouts from a single screen. Although the instrumentation for high throughput screening in 384-well format is yet to be developed, the well established reliability of this technique in immune cells suggest this is an area for promising future screening applications.

Viral Mediated Transduction

Retroviral, adenoviral and lentiviral-transduction are an alternative method of exogenous interfering RNA expression in hematopoietic cell models where transient transfection is inefficient. Initial shRNA designs employed plasmid or viral-based expression RNA polymerase III-driven stem-loop (hairpin) structures of approximately 50-60 nucleotides resembling pre-miRNA, the endogenous substrate of the Dicer ribonuclease [45]. More recent studies have demonstrated that shRNA can be more efficiently processed through the host miRNA processing pathway if they are transcribed by RNA polymerase II in the context of the larger 200-400 nucleotide endogenous primary miRNA transcripts that are first processed by the Drosha ribonuclease [46]. Although transient expression of the above types of shRNA designs from plasmid templates can be employed in some easily tractable cell types, this shRNA expression approach is more widely used for screens adopting viral delivery systems such as adenoviral, retroviral and lentiviral infection that can be applied to almost any cell type [8, 47].

However, a caveat with this approach is that the shRNA may compete with endogenous miRNAs for the small RNA processing machinery described above, and there have also been reports that viral driven shRNA expression can trigger an IFN response [48-50]. In this regard, designing the shRNA in the context of the primary miRNA transcript as described above and expressing the transcripts at lower levels from weaker pol II promoters appears to be the best approach, showing reduced toxicity in mouse studies and providing a promising platform for the stable knockdown of genes with minimum cytotoxicity [51].

Another practical issue is that the generation of individual viruses for all genes in the genome could be too expensive and labor-intensive for many research groups. While there are examples of arrayed high throughput lentiviral screens conducted in hematopoietic cells [52], many virus mediated shRNA screening studies address the large-scale virus production challenge by adopting a pooling strategy using mixed viral preparations containing shRNAs targeted against many genes. Cells are infected to achieve on average, one stably integrated shRNA per cell, followed by an assay that permits selection for loss or gain of a particular phenotype. To then identify a specific shRNA that causes a phenotypic change, RNA is isolated and over or under-representation of particular shRNA hairpin sequences (or unique shRNA construct-embedded bar codes) is determined by deep sequencing or microarray [53, 54].

A notable advantage of shRNA-based screening is that viral vectors with inducible promoters can be used to tightly regulate gene knockdown [53, 55]. This is particularly useful in circumstances where genes essential for cell viability need to be depleted acutely over a short time frame. Another useful application of this platform is the ability to use cell-type specific promoters to selectively express the required inducible response elements and deplete a target gene or genes only in a particular immune cell subset. Selective promoters for many immune subsets are well established, such as the myeloid promoters lysM, csf1r, CD11c, CD68, macrophage SRA, and CD11b, lck for T cell expression, CD19 for B cells, Mcpt5 for mast cells and Ncr1 for NK cells. Although retroviral and adenoviral vectors are appropriate for most cell lines and some primary cells, such as activated lymphocytes, efficient delivery of shRNA to non-dividing cells, including immature DCs, is better achieved with lentiviral systems [56].

Modified siRNA (Accell)

Accell siRNA delivery is a relatively recent technology developed by Dharmacon (part of GE Healthcare) for use in less tractable cells. This technology uses modified siRNA for efficient uptake by cells without a transfection reagent. In a recently published study, transfection efficiencies of siRNA in primary cells with Accell delivery was 85% (range, 71–97%) compared to 38% (23–65%) using nucleofection [57]. This technique has been optimized for use in 96-well format and has the potential for development in 384-well format to use in high throughput transfection of hematopoietic cells. However, no high throughput siRNA screen has been reported so far with this method to the best of our knowledge. A caveat with this technique is the requirement to incubate cells in serum-free media for a minimum of 48 h during the siRNA delivery and, depending upon the cell type, tolerance to the serum-free conditions may vary. While there are few reports testing the tolerance of immune cells to serum-free conditions in culture, Accell siRNA delivery can also be achieved in low serum conditions down to 2.5%, which could broaden the applicability of this delivery method. It should also be noted that the required concentration of siRNA for efficient delivery using Accell technology is higher than conventional approaches using lipid-based delivery.

Screen Assay Design

Various types of assays have been adapted for RNAi screening in hematopoietic cells. Selection of an optimal assay depends on the nature of the cells, whether they are suspension or adherent and also on the biological question under

consideration. For genome-wide screens, the amenability of an assay to high throughput microwell formats is a primary consideration. Here we describe several types of assays researchers have adopted for RNAi screening in immune cells.

Viability and Cell Survival Assays

Due to their routine use in evaluating cancer cell survival, cell viability assays are one of the most highly validated and straightforward screen readouts. Most widely used cell viability assays measure the energy metabolite ATP [58], and these assays also have important applications as a secondary readout for assays that require cell number-based normalization. Viability assays have immunological applications as a primary readout outside of the cancer field, for example in assaying pathogen infections that typically lead to cell death or necrosis, or in the study of apoptotic signaling pathways that lead to the negative selection of self-reactive lymphocytes. As an example, a FADD-deficient Jurkat cell line, I42, that undergoes programmed necrosis in response to tumor necrosis factor-alpha (TNFα) stimulation, has been used to screen for kinases important in this pathway [59].

Endpoint Reporter Assays

Genetic reporter systems provide a convenient and well-established approach to studying signaling and subsequent gene expression by RNAi screening. Most applications of such assays adopt a sensitive endpoint readout of an easily assayed reporter such as luciferase, driven by a signal-responsive promoter derived sequence. In Jurkat T cells, for example, a stably integrated reporter was created with Gaussia luciferase expression under the control of 8xNFκB sites, and used to screen for signaling components involved in activation of NF-κB by either TNFα or anti-CD3 [43]. In another novel reporter assay designed specifically for pooled shRNA screens, a macrophage cell line was created with the TNFα promoter driving a lethal diphtheria toxin A (DT-A) cassette, which promoted cell death in response to TLR stimulation. Thus, upon infection with a pooled shRNA library, cells transduced with shRNAs inhibiting genes necessary for TLR-dependent pro-inflammatory signals would be positively selected and the causative shRNAs could be identified by deep sequencing [60].

High Content Imaging and Cell-Based Fluorescence Assays

Advances in high throughput microscopy platforms have enabled the development of increasingly complex cellular imaging and fluorescence-based screening

assays. In addition to the simultaneous detection of multiple readouts using spectrally distinct fluorescent proteins or conjugated antibodies, changes in cell shape and morphology can also be rapidly profiled using advanced image analysis software. Such 'high content' imaging can be especially advantageous in RNAi screening as it can allow the detection of multiple assay readouts in a single screen, significantly increasing the amount of data that can be collected to gain a more comprehensive understanding of the effects of gene perturbation.

The use of multiple fluorescent channels can also permit differential staining to screen for effectors of key immune processes. For example, an image-based phagocytosis assay was developed in mouse macrophage J774A.1 cells using different fluorescent dyes to distinguish external and internalized opsonized particles. This assay was used to screen for key phagocytosis effectors among a panel of 20 GTPases [61].

Another image-based fluorescence assay adapted for RNAi screening in immune cells is the Forster resonance energy transfer (FRET) assay. Using a combination of fluorescent proteins, these assays can effectively monitor changes in protein-protein interactions or in the activities of signaling enzymes and effectors that drive immune responses. For example, the small G protein Cdc42 regulates changes in cell morphology, motility and vesicular transport, and its activity can be monitored with the FRET biosensor Raichu-Cdc42, which changes conformation in response to Cdc42 activating stimuli. This has been successfully used for RNAi screening in YTS NK cells, where Cdc42 activity was used as a readout downstream of activating and inhibitory NK receptors that drive the morphological changes essential for immune surveillance and cytotoxicity [62].

Although not microscopy-based, flow cytometry assays are another fluorescence-based readout that are often adaptable to high throughput RNAi screening applications, and these assays can be particularly useful in non-adherent immune cell types which are less amenable to microscopy-based high content assays. This long-established technology provides the advantage of rapid collection of data from a large number of single cells and is particularly well suited to assays that can provide a robust change in a fluorescent readout. It has been widely adopted for RNAi screening in immune cells to study host-pathogen interactions, immune activation and cell signaling. In one such study, GFP-expressing *F. tularensis subsp. holoarctica* was used in a genome-wide siRNA screen for genes mediating resistance to *F. tularensis* infection in the THP1 macrophage-like cell line. Hits were scored by decreased levels of GFP in cells 24 hr after infection with the labeled bacteria [63].

Secreted Effector Assays

There are many well-established methods to assay for effectors secreted from immune cells or released by degranulation. Enzyme linked immunosorbent assays (ELISA) are widely used for secreted cytokine quantification, and more recently developed bead-based assays can permit the simultaneous evaluation of multiple cytokines in a single sample of cell supernatant. The advantage of adopting these assays as RNAi screen readouts are that they are already well validated in microplate formats, and the collection of the secreted factors from cell supernatants can permit additional cell-based assays to be run from the cells left behind in the same well. A potential drawback with these assays on a genome-wide scale is the expense of the antibodies and reagents to run a large-scale screen. Examples of ELISA-based assays used in RNAi screens include interleukin (IL)-2 production in response to antigen presentation by DCs [64], and IL-1β production in response to *E. coli* infection of THP-1 macrophages [65]. A bead-based antibody detection assay for IFN-γ was used to screen for genes that promote target cell resistance to human NK cell mediated lysis [66]. Enzymatic activities of secreted proteins can also be used for developing RNAi screening assays in immune cells. For example, mast cell degranulation releases various enzymes that can be assayed by colorimetric detection. One such study screened for phosphatases involved in IgE-mediated mast cell degranulation using β-hexosaminidase activity as a readout [67].

Migration Assays

Transwell migration assays have been developed for studying immune cell migration and chemotaxis. While these assays do not lend themselves well to a high-throughput arrayed assay format, they are better suited to pooled shRNA screens as cells from a transduced population that show an aberrant migration phenotype can be easily isolated. A screen of this type has been run in RAW264.7 cells to screen for genes influencing chemotaxis towards the complement component C5a [68], and also in SupT1 lymphoma T cells to identify genes that influenced migration towards CXCL12 [69].

Colony Forming Assays

While studying host cell infection by pathogens such as bacteria or viruses, growth or replication can be measured by colony forming or focuses forming assays (CFU). Although not normally considered practical in a high throughput format, this type of assay has been used for a genome-wide RNAi screen to

identify host factors involved in the response to *Mycobacterium tuberculosis* infection [31].

Microarray and Next Generation Sequencing Assays

Although they are not widely used as a primary assay for evaluating cell responses in RNAi screens, microarrays and next generation sequencing have important applications in identifying enriched or depleted shRNAs (and hence the genes) responsible for phenotypic effects in pooled shRNA screens. For example, microarray analysis has been used to identify enriched shRNAs after human immunodeficiency virus (HIV) infection of Jurkat cells [70].

INVESTIGATION OF IMMUNE CELL FUNCTION USING RNAi SCREENING

Despite the challenges facing researchers wishing to adopt RNAi screening approaches in hematopoietic cells, this technology has been used successfully in numerous studies in a variety of immune cell types. While we do not have the scope to reference all of these studies in this review, here we highlight several examples organized by the target cells in which the studies were performed.

T Cells

T cells are a vital arm of the adaptive immune response and provide surveillance against micro-organisms, pathogens and tumors [71]. T cells arise from common lymphoid progenitors in the bone marrow (Fig. **1**) and are transported to the thymus where some will eventually develop into CD4$^+$ T cells and CD8$^+$ T cells. In response to activation by antigen-presenting cells (APCs), CD4$^+$ T cells proliferate and differentiate into helper T cell subtypes (*e.g.* Th1, Th2 and Th17) that secrete distinct cytokines that influence both innate and adaptive immunity, while CD8$^+$ T cells proliferate and differentiate into cytotoxic CD8$^+$ T cells that mediate direct target-cell killing. A lineage of CD4$^+$ T cells known as regulatory T cells (Tregs) inhibit the function of T cells and other immune cell types. T cell differentiation also results in the formation of long-lived memory T cells that are rapidly mobilized should the host encounter these pathogens in the future. T cell activation and differentiation takes place in secondary lymphoid organs and is followed by trafficking of these cells from the bloodstream into inflamed target tissues such as the skin or gastrointestinal tract where they mediate effector function. The capacity of T cells to migrate and infiltrate into tissues, however, is also a major contributing factor in the development of autoimmunity [72], allergy [73] and graft rejection [74]. Understanding the genes and proteins that regulate T cell activation and trafficking

into tissues may therefore reveal novel immunosuppressive or anti-inflammatory targets. T cells (and monocytes/macrophages) are also the primary site of infection of the HIV virus [75]. Since HIV is dependent on host factors for both viral entry and replication inside T cells, elucidating such host factors may provide novel therapies in the continuing fight against this virus. In addition, targeting malignant T cells *via* silencing of cell survival genes may provide novel therapeutic targets for T cell leukemias and lymphomas. Here we outline siRNA screening studies that have led to the elucidation of new components that regulate T cell signaling, cytokine production, migration and viral-host interactions.

T Cell Activation and Function

A number of RNAi screens performed in human T cell lines or primary T cells have revealed new molecular components that influence T cell activation and cytokine production. For example, the heterogeneous ribonucleoprotein hnRNPLL was identified as a key splicing factor that regulated alternative splicing of CD45, a transmembrane protein tyrosine phosphatase, in the JSL1 T cell line following T cell activation [76]. In primary CD4[+] T cells, a lentiviral screen targeting ~ 1000 kinase and phosphatases identified 55 genes (such as the tyrosine kinases FLT-3 and ZAP-70) that positively or negatively regulated cytokine production in Th1 T cells (IFN-γ), Th2 T cells (IL-13) and Treg cells (IL-10) [77]. In a separate study, the same group recently used this lentiviral shRNA library to identify factors that regulate the expression of CD46, a type 1 membrane protein required for T cell activation and function, in primary human CD4[+] T cells [78]. Two members of the serine/threonine kinase G protein-coupled receptor kinase (GRK) family, GRK2 and GRK3, influenced the expression of CD46, as knockdown of these kinases inhibited the expression of this membrane receptor [78]. Brechmann *et al.* [59] recently screened a siRNA library targeting 298 known or putative phosphatase genes to identify novel regulators of NF-κB signaling in the Jurkat leukemic T cell line using nucleofection in combination with a 96-well shuttle system. The protein phosphatase 4 regulatory subunit 1 (PP4R1) was identified as a negative regulator of NF-κB activation in this study and characterized further [43].

Screening of siRNA libraries have also led to the identification of new components that regulate T cell migration and trafficking. A nucleofection-based small-scale siRNA screen of Rho guanosine triphosphatases in CCRF-CEM T cells identified RhoA as a crucial mediator of trans-endothelial migration (the process by which T cells migrate through endothelial cells that line blood vessel walls into peripheral tissues or secondary lymphoid organs) [79]. In a separate

study, a lentiviral shRNA screen targeting 300 genes encoding predominantly kinases and phosphatases was carried out in SupT1 lymphoma T cells to identify genes that influenced migration towards the CXCL12 chemokine in Transwell chambers [69]. Eleven genes, including ZAP-70 and three members of the synaptotagmin family of proteins that regulate vesicle fusion (synaptotagmins 2, 5 and 7), were confirmed as being positive or negative regulators of CXCL12-mediated chemotaxis following a secondary screening assay [69]. The identification of synaptotagmins from this screen led the authors to examine in more detail the role of synaptotagmin-7 in leukocyte migration. Leukocytes derived from synaptotagmin-7 knockout mice migrated less efficiently *in vitro* in comparison to wild-type leukocytes. Furthermore, synaptotagmin-7 deficient leukocytes were perturbed in their ability to migrate *in vivo* using an inflammatory mouse model of gout [69].

Screening of siRNA libraries in T cell lines have also identified genes that regulate distinct forms of cell death and pathways implicated in T cell malignancies. A variant of the Jurkat T cell line that undergoes programmed necrosis in response to stimulation with the cytokine TNF was used to screen a siRNA library targeting 691 human kinase genes [59]. Ten distinct genes that regulated TNF-induced necrosis were identified in this screen and one of these, RIP3, was characterized further [59]. In another study, a lentiviral shRNA screen targeting ~9500 genes was carried out in Jurkat T cells to identify genes whose inhibition rendered T-cells resistant to Fas-induced apoptosis, a pathway that functions in immune cell homeostasis [54]. As well as identifying genes with well-established roles in this pathway, this study identified ARD1A (a SWI/SNF chromatin remodeling complex component) and CBX1 (a chromatin silencing protein) as critical regulators of Fas-mediated apoptosis [54]. Further siRNA screens carried out in this study revealed genes that were essential for the proliferation of Jurkat T cells, SupT1 T cells and ten other cancer cell lines [59]. The molecular events that contribute to T cell acute lymphoblastic leukemia (T-ALL) [80] and to the survival of T-ALL cells [81] have also been elucidated further using siRNA screening.

T Cells and HIV Host-Pathogen Interaction

HIV encodes nine viral genes, is the causative factor of AIDS and infects T cells, monocytes/macrophages and other non-immune cell types including cells of the central nervous system [75]. The CD4 receptor that is expressed on the surface of these host cells serves as the primary site of viral attachment and entry, while chemokine receptors such as CXCR4 and CCR5 also serve as co-receptors for

viral entry at later stages [75]. Once inside the cell, viral replication and integration into the host genome is dependent on both viral and host cell genes. In 2008, three genome-wide siRNA screens identified hundreds of host cell genes that were crucial for HIV replication and infectivity [82-84]. Surprisingly, very little overlap in the 'hits' or host cell genes required for viral replication were observed between the three screens. This was attributed to differences in the experimental procedures and assay read-outs of the individual studies [85]. A fourth study that utilized a whole genome siRNA library to identify intrinsic resistance factors to HIV replication was recently reported [86]. A notable caveat with all of these studies was that HeLa-derived cervical cancer cells expressing exogenous CD4 were used as the host cells for HIV infection in three of the studies [82, 84, 86], while an HIV isolate that lacked the viral envelope protein required for infection was pseudotyped with the Vesicular Stomatitis Virus Glycoprotein to allow viral entry into CD4-negative HEK293T cells in the other study [83]. Hence these cells were not representative of the physiological cell types infected by the virus. In contrast, a genome-wide siRNA screen has recently been performed in a T cell line to identify host cell factors required for HIV infection. Yeung *et al.* [70] used the Jurkat leukemic T cell line that expresses endogenous CD4 to screen an shRNA library targeting 54,509 human transcripts to identify host cell proteins contributing to HIV replication. Interestingly, 252 genes were identified from this screen and although most of them did not overlap with the genes identified from previous screens, many of the genes mapped to common cellular pathways [70]. An shRNA screen targeting 622 human kinases and 180 human phosphatases was also recently performed in Jurkat T cells to identify genes required for HIV infection and replication [87]. Fourteen new genes were identified in this screen and implicated in steps of the HIV life cycle that preceded viral integration into the host genome (*i.e.* viral entry, viral uncoating or viral transcription) [75]. In theory, the identification of kinases and phosphatases that influence the life cycle of HIV may provide novel anti-viral therapeutic targets, since kinases and phosphatases are generally 'druggable' with small molecule inhibitors and as such may provide alternative strategies to block viral replication in T cells with agents other than siRNAs which are known as 'hard-to-transfect' into these cells. Although current reports on HIV screens in Jurkat T cells and other cell lines sheds valuable insight into the cellular pathways involved in the complex interplay between the virus and host proteins, identification of host factors in non-transformed primary T cells will provide a better understanding of the host-virus interaction that occurs *in vivo* and may provide better therapeutic opportunities to target HIV.

B Cells

B cells, like T cells, develop from common lymphoid progenitors in the bone marrow (Fig. **1**). B cells recognize soluble antigen *via* the B cell receptor (BCR) and differentiate into antibody-producing plasma cells and memory cells [88]. Antibodies promote antigen clearance *via* neutralization, opsonization or phagocytosis by other immune cells. Most B cell immune responses require the 'help' of T cells in the form of co-stimulatory receptors or cytokines for optimal proliferation and antibody production. In the presence of T cell 'help', B cells genetically modify the type and affinity of antibody produced by processes known as class switching and somatic hypermutation, respectively. Class switching involves structural changes and permits some antibodies to cross certain tissues and allows the antibody to be recognized by specific cell types. Somatic hypermutation enables more potent antibodies to be developed against the antigen during the immune response. Class switching and somatic hypermutation occur in distinct areas of secondary lymphoid organs known as germinal centers.

The vast majority of siRNA screens performed in B cells have focused on the identification of genes required for survival and pathogenesis of malignant B cell leukemias and lymphomas. Hence, the discovery that silencing the expression of these genes results in cell death reveals exciting new opportunities to target these malignant cells. B cells are prone to malignant transformation because the genetic re-arrangements that lead to BCR diversification and antibody production may promote constitutive activation of signaling pathways leading to cellular proliferation and oncogenesis [89]. The number of B cell malignancies are broad and are categorized according to clinical, histological and genetic features (The World Health Organization Classification of Tumours of Hematopoietic and Lymphoid Tissues, Fourth Edition; 2008). Regulators of cell survival in malignant B cells have been revealed by siRNA screens in cell lines or primary cells derived from diffuse large cell B cell lymphoma [53, 90-93], Burkitt's lymphoma [94, 95], follicular lymphoma [96], multiple myeloma [66, 97-100], chronic lymphocytic leukemia [101], mantle cell lymphoma [101] and acute lymphoblastic leukemia [42]. An siRNA screen in chronic lymphocytic leukemia cells also revealed molecules that perturbed T cell/APC interaction and thus potential anti-tumor surveillance [102].

While the majority of all siRNA screens to date have been performed *in vitro*, advances have been made to adopt the siRNA/shRNA screen format to *in vivo* systems. The ***Eµ-Myc;Arf***$^{/-}$ B cell lymphoma mouse model is such a system whereby transfer of B cell lymphoma cells into mice results in malignancy. B cell

lymphoma cells were transduced with a lentiviral shRNA library targeting 1000 genes with known or suspected roles in cancer and injected into mice to identify shRNAs that were depleted or enriched *in vivo* during tumorigenesis [103]. shRNAs targeting dynamic actin reorganization and cell adhesion genes such as Rac2, CrkL and Twinfilin were depleted *in vivo*, with subsequent studies demonstrating that these genes influenced tumor cell mobilization and metastasis [103]. A similar *ex vivo* study was conducted using *Eµ-Myc* hematopoietic stem and progenitor cell (HSPC) to identify shRNAs capable of accelerating lymphoma. This study identified a number of tumor suppressors (*Sfrp1, Numb, Mek1, Angiopoietin*) and components of the DNA damage response machinery (such as Rad17) as being important for lymphoma development [104].

Dendritic Cells

DCs can be derived from both lymphoid and myeloid progenitors during hematopoiesis (Fig. **1**) [105], and they perform pivotal functions in host defense and antigen-presentation in innate and adaptive immune responses [106]. Application of siRNA screening technology in these cells has been primarily in the areas of antigen presentation, DC migration and host-pathogen interaction.

Antigen Presentation

DCs collect antigens and present them to CD4 T cells on major histocompatibility complex (MHC) class II [107]. Researchers have used alternative cell types with DC characteristics to study proteins and networks involved in MHC-II expression and peptide loading. One such study performed a genome-wide siRNA screen in the MelJuso cell line followed by validation in human primary monocyte-derived DCs [108]. Although MelJuso cells are not hematopoietic, they express many immune-specific genes involved in regulation ofMHC-II transport in a similar manner to DCs. This study identified nine proteins involved in transcriptional regulation of MHC-II; five of these regulate the expression levels of the class II major histocompatibility complex transactivator (CIITA), and the remaining genes modulate MHC-II expression without affecting CIITA levels [108]. In addition, they were able to identify six genes involved in MHC-II distribution in DCs. Further analysis identified one of the candidates, ARL14/ARF7, as a MYO1E receptor on the MHC-II compartment required for actin-based transport control [108]

Another property of DCs is their ability to present exogenous antigens on MHC class I, a function termed cross-presentation, and this is crucial for the generation of cytotoxic T cell immunity in response to antigens associated with viral

infection, tumorigenesis and DNA vaccination [109]. The role of Rab GTPases, key regulators of membrane trafficking, in cross-presentation has been studied using lentiviral shRNA mediated knockdown of 57 Rab GTPases in the mouse DC2.4 cell line [64]. Upon establishing stable shRNA expression, the DC2.4 cells were infected with BL21 *E.coli* and antigen presentation to the B3Z T-cell hybridoma was measured by IL-2 ELISA. This identified 12 GTPases involved in antigen cross-presentation, and further experiments showed that internalized MHC class I molecules accumulated in Rab3b/3c-positive recycling endosomes, implicating these vesicles in the process of cross-presentation of exogenous antigens [64].

Another lentiviral shRNA screen was done in mouse DCs with the goal of identifying signaling molecules that regulate cross-presentation [52]. This screen used an arrayed shRNA library in 96-well format containing 3450 shRNAs against 691 kinases and phosphatases. The shRNA-expressing DCs were challenged with ovalbumin (OVA) peptide-expressing yeast to induce antigen presentation to $CD8^+$ T cells from OT-1 mice, and T cell proliferation was assayed by a thymidine incorporation assay. This screen unveiled several new regulators of antigen cross-presentation. For example, Acvr1c, also known as Alk7, a type I receptor for the TGF-β family of signaling molecules, was identified as a novel negative regulator [52].

DC Migration

Upon encountering antigens in peripheral tissues, DCs must localize to secondary lymphoid organs such as the spleen and lymph nodes to present the antigens to T cells. An shRNA screen has been conducted to identify the Rho guanine nucleotide exchange factors involved in the actin cytoskeleton signaling mechanisms controlling DC migration [68]. This screen used the macrophage-like RAW264.7 cell line, but then validated hits in primary DCs. Delivery was achieved through lipid-based transfection of plasmid-expressed shRNAs targeting 38 Rho guanine nucleotide exchange factors (GEFs). The shRNA vector was modified by incorporating a GFP-luciferase fusion protein expression cassette. At 48 h post-transfection, transmigration of the cells was assayed in 24-well transwell plates and quantitated by luciferase expression levels. From an initial set of 6 hits showing more than 50% inhibition of cell migration, Arhgef5 was found to be a key protein essential for the transmigration of DCs. Further *in vitro* studies showed a direct interaction of Arhgef5 with Gβ1γ2, a key effector in G-protein mediated chemotactic signaling to the actin-cytoskeletal regulator RhoA. MIP1-α induced chemotaxis of immature DCs and migration of DCs from skin to lymph

node were shown to be impaired in Arhgef5 knockout mice [68], thus identifying Arhgef5 as a significant regulator of RhoA-dependent immature DC migration.

DCs and Host-Pathogen Interactions

DCs also have important innate immune functions in the host response to pathogen infection. The 96-well format nucleofection approach described above has been employed to study the role of the RIG-I RNA-sensing pathway during Newcastle disease virus (NDV) infection of DCs. While not run as a screen for novel pathway components, this study confirmed a role for RIG-I in NDV-driven IFNβ induction, and provides a basis for the application of this siRNA delivery approach for large-scale siRNA screens in primary DCs [110].

RNAi screening has also been employed in DCs to study the cellular response network mediating differential responses to TLR stimuli. Using gene expression arrays to implicate genes involved in regulating the DC transcriptional network, candidate regulators were tested by lentiviral-based shRNA perturbation in primary BMDC [111]. From a screen of 125 genes, it was concluded that on average, about 14 regulators activate a target gene and about 5 regulators repress it. This study also identified novel regulators of transcription in response to TLR stimuli [111]. A similar approach was undertaken to study the signaling components involved in the TLR response in BMDCs [112]. In combination with transcriptional profiling, small molecule-based perturbation and phosphoproteomics, lentiviral mediated shRNA knockdown identified novel signaling regulators involved in TLR signaling in DCs. Of particular note, Polo-like kinases 2 and 4 activated a novel signaling branch including several proteins that included Tnfaip2, a gene previously associated with autoimmune diseases [112].

Monocytes and Macrophages

Monocytes play a vital role in immune system function linking inflammatory responses and the defense against pathogens to the adaptive immune response. Circulating monocytes serve as a reservoir for the renewal of both macrophages and antigen presenting DCs [113]. Macrophages are derived from parent monocytes in the blood and are highly phagocytic cells that play a pivotal role as first responders to invading pathogens and also in maintaining tissue homeostasis through scavenging and clearance of dead and damaged host cells [114]. In response to various stimuli, macrophages are activated and acquire specialized functional phenotypes. Activated macrophages are generally classified into the following two types: classically activated macrophages (M1), which promote

inflammation, and alternatively activated macrophages (M2), which inhibit inflammation, but promote wound repair and tissue remodeling [113]. Although, macrophages are involved in host defense against microorganisms, these cells are often a primary site for replication of bacteria, protozoans and viruses [115, 116]. Recognizing the broad range of functions performed by monocytes and macrophages, RNAi screens have been used to study their immune system function in several areas.

Immune Activation and Cytokine Response

Interleukin-1β is a key inflammatory mediator produced by monocytes and macrophages in response to bacterial infection. A lentiviral shRNA screen has been conducted in THP-1 monocytes to identify components of the splicing machinery, among a set of 425 genes, regulating the expression and secretion of IL-1β in response to *E. coli* infection [65]. THP-1 cells were infected and selected for stable shRNA expression, challenged with bacteria, and the level of secreted IL-1β protein was measured by ELISA. This study identified 10 factors whose perturbation decreased the level of IL1-β, and 20 factors whose depletion led to increased in cytokine secretion, the latter set including SFRS3, a member of the SR protein family required for constitutive pre-mRNA splicing and regulation of alternate splice site selection [65, 117].

TNFα is another major inflammatory cytokine released by macrophages in response to a broad range of infectious stimuli. Using a novel reporter assay (described in 'Endpoint reporter assays' sub-section above), RAW264.7 macrophage cells were screened with a randomly generated lentiviral library of shRNA sequences to implicate genes required for TLR pathway induced TNFα expression. Using a diphtheria toxin reporter, this assay positively selected for cells expressing shRNAs that blocked the TLR2 and TLR4-mediated pro-inflammatory TNFα response [60]. While the authors reported identification of enriched shRNAs by this method, they did not identify putative gene targets with a complete match to the shRNA sequences, possibly due to miRNA-like targeting through partial complementarity in the shRNA seed sequence.

Phagocytosis

Phagocytosis by macrophages is orchestrated by coordinated movement of the actin cytoskeleton that involves Rho GTPase signaling [118]. A small-scale siRNA screen in the J774.1 mouse macrophage cell line attempted to identify the relevant GTPases (among a target set of 20) controlling phagocytosis through the Fcγ receptor R (FcγR) and complement receptor 3 (CR3) [61]. After lipid-

mediated transfection of siRNA in 24-well plates, cells were challenged with either IgG-opsonized (for FCγR) or C3bi-opsonized (for CR3) sheep red blood cells (RBCs), and phagocytosis levels were assayed by immunofluorescence microscopy. Cdc42 and Rac2 were shown to be the primary effectors of FcγR-mediated phagocytosis, RhoA primarily mediated CR3 mediated phagocytosis, while RhoG was required for both [61].

Host-Pathogen Interaction

During innate and adaptive immune responses to viral and intracellular bacterial infection, macrophages are activated by IFN-γ produced by NK cells and by CD4 Th1 and CD8 cytotoxic T cells. IFN-γ can inhibit the intracellular replication of *Francisella tularensis* in both human-derived macrophages and in mice. A lentiviral shRNA screen has been conducted to identify host factors conferring resistance to *F. tularensis*. In this screen, THP-1 cells were transduced with a genome-wide lentiviral library containing 50,000 shRNAs. At 9-10 days after infection and selection with puromycin, the cells were differentiated to a macrophage-like state with PMA and treated with IFN-γ prior to infection with fluorescently labeled *F. tularensis* for 2 h. At 24 h post-infection, the cells with higher fluorescence expression were collected by FACS and the shRNAs expressed in these cells were detected by genomic DNA sequencing. From five independent experiments, those genes that appeared multiple times in independent screens were studied further. Eight of the validated proteins (TNFRSF9, SERPINI1, SERPINA7, HLA-DRB1, ATG5, ATG16L1, PLEK2 and PLS1) were also required for *Listeria monocytogenes* resistance, implicating these genes in broader host mechanisms of IFN-γ restriction of bacterial replication [63].

Another focus of siRNA screening in monocytes and macrophages has been to identify host factors influencing *Mycobacterium tuberculosis* (Mtb) infection. These studies have great potential in drug development efforts to alleviate the continuing global health burden of tuberculosis. A focused screen of 744 kinases and 288 phosphatases for their influence on Mtb replication was carried out in murine J774.1 macrophages [31]. Two rounds of lipid-based transfection 48 h apart were implemented to obtain prolonged silencing, and primary screen hits were validated with different siRNAs from alternative vendors. Among a validated group of 41 hits, the TGF-β receptor isoforms TGFβRI and TGFβRII showed particular promise, as in addition to their siRNA perturbation phenotype, they showed marked elevation in expression in response to Mtb infection in microarray experiments. The TGF-β pathway inhibitor, D4476, showed marked inhibition of intracellular bacterial replication and also eliminated Mtb from

infected mice [31]. This study emphasizes the potential of siRNA screens in identifying drug targets for pathogen control in hematopoietic cells.

A comparable Mtb screen was conducted on a genome-wide scale in THP-1 cells using the same virulent H37Rv strain [32]. This screen used pooled siRNA targeting 18,174 human genes, and after several rounds of hit validation, 270 genes were identified whose knockdown led to a reduction in intracellular Mtb load, while 5 genes showed an elevated bacterial load. Since the Mtb bacterium exhibits a broad spectrum of genotypic and phenotypic variation, the authors re-screened the validated hits against different field isolates of Mtb and noted limited overlap in the host factor requirement across Mtb strains [32]. However, pathway and protein interaction network analysis of the specific hits for each isolate suggested a greater degree of overlap at the pathway level, a pattern also observed in influenza and HIV screens [119] [120].

Tumor-Associated Macrophage Biology

Tumor-associated macrophages (TAMs) have an important role in facilitating tumor outgrowth through anti-inflammatory, immunosuppressive, pro-angiogenic and pro-metastatic properties [121]. It would thus be of great clinical importance to identify the pathways that lead to generation of TAM-like macrophages, and this has been recently addressed with an adenoviral shRNA screen in a human macrophage precursor, the CD14+ peripheral blood mononuclear cell [122]. Using 8495 adenoviral shRNA constructs targeting 2825 genes, shRNA-expressing cells were stimulated to form TAM-like cells by co-culture with MCF-7 breast cancer cells, and assayed for IL10 secretion (a key cytokine produced by TAM-like macrophages). Five proteins, IL-4 receptor alpha (IL-4RA, IL-4R), cannabinoid receptor 2 (CNR2, CB2), 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductase (HMGCR), and bruton's tyrosine kinase (BTK), were identified as regulators of IL-10 and/or IL-16 production and possible targets for control of TAM function [122].

NK Cells

NK cells are large granular leukocytes that play a pivotal role in host defense against viruses and tumors [51]. Although large-scale RNAi screens in NK cells have not yet been reported, smaller scale screens have been carried out. As described in above (sub-section 'High content imaging and cell-based fluorescence assays'), a FRET-based assay was developed in YTS NK cells to measure activity of the Rho GTPase Cdc42, which plays important roles in NK cell morphology, vesicular transport and motility [123, 124]. Known binding

partners of Cdc42 were screened by siRNA electroporation for their effect on Cdc42 activity upon challenge with 721.221 NK activating cells [62]. This identified the guanine nucleotide exchange factors RhoGEF6 and RhoGEF7, the kinase Akt, and the p85α subunit of phosphoinositide 3-kinase (PI3K) as essential for Cdc42 activation [62].

An shRNA screen was also conducted on IM9 cells, a multiple myeloma cell line, to identify molecular pathways which modulate tumor cell susceptibility to NK cell-mediated killing [66]. IM9 cells were transduced with a lentiviral shRNA library targeting 476 kinases, 180 phosphatases, and 372 genes representing tumor suppressors, DNA binding proteins, and modification enzymes. The transduced cells were incubated with NKL effector cells, and release of IFN-γ from NKL cells was assayed as a measure of productive NKL-IM9 cell interaction. Of 83 hits identified in the screen, 66 were kinases, 4 were phosphatases and 12 were from the 372 gene set with non-kinase function. Genes in the MAPK pathway were highly represented with 15 genes, and two members of the JAK kinase family, JAK1 and JAK2, were shown to increase the susceptibility of tumor cells to NK-mediated lysis [66].

Mast Cells

Mast cells trigger allergic reactions and IgE-associated immune responses. Degranulation and release of inflammatory mediators from mast cells are initiated upon aggregation of FcεRI on these cells [125, 126]. To understand the role of phosphatases in IgE-mediated mast cell activation, a siRNA screen has been reported using a library targeting all the 198 mouse phosphatase genes in the mouse mast cell line, MMC-1[43]. Using IgE-Ag-induced mast cell degranulation as a functional readout, 10 phosphatases enhanced and 7 inhibited FcεRI-induced degranulation. Among the top hits identified were subunits of calcineurin, where siRNA knockdown of both the Ppp3r1 regulatory subunit and the Ppp3cc enzymatic subunit inhibited degranulation [43].

FUTURE OPPORTUNITIES AND CLOSING REMARKS

The studies we have described emphasize that, while immune cells often present an experimentally challenging cell system, RNAi screening can be successfully implemented in hematopoietic cells to discover important new biology and provide insight to cellular mechanisms of immunological disease. Technological advances in si/shRNA delivery and screening assay formats now permit large-scale screens to be conducted in primary immune cells, which were once

considered intractable to genetic screening. The screens we have described here have generated a wealth of new knowledge on genes required for numerous aspects of immune cell function, including innate and adaptive immune cell activation, migration and trafficking, cell death and clonal selection, hematopoietic cell malignancies and host-pathogen interactions. Furthermore, these reported studies only scratch the surface of potential applications and opportunities for RNAi screens in other cell types of hematopoietic origin (Fig. **2**). Also, with further improvements in screening and delivery methods, we expect continued application of this technology will lead to many further insights to immune cell function in health and disease. Areas that could benefit especially from this technology are studies of immune cells *in vivo* in adoptive transfer models that were not previously amenable to large-scale unbiased screening approaches. For example, the ability to deliver immune cell subsets transduced with pooled shRNA libraries, and to leverage the increasing capacity of next-generation sequencing technology to identify the shRNAs expressed by small numbers of cells that show a specific phenotype *in vivo* holds great promise. Another area of potential technical development and new opportunities will be in the targeted delivery of siRNA to specific cell types *in vivo*. For example, targeted delivery of siRNAs into T cells *in vivo* is now feasible with peptide/protein-antibody fusion proteins [127, 128], integrin-targeted and stabilized nanoparticles [129, 130] and aptamers [131, 132], and it is anticipated that these novel reagents may also be useful for *in vivo* delivery of siRNAs into other immune cell types.

While the described screens have produced valuable insight, a major advantage of siRNA screening that has not yet realized its full potential in the immunology field is its ability to screen the entire genome in an unbiased manner. The majority of immune cell screens published to date have targeted subsets of genes, in many cases focusing specifically on a gene family already known to have a central role in a particular process. While this can be useful in highlighting selective usage of functionally related proteins in certain processes (such as the differential usage of GTPases in FcγR-mediated and CR3-mediated phagocytosis in macrophages), these studies ultimately miss the opportunity to identify novel regulators of the process under study, and the potential to reach a comprehensive understanding of all genes required in a given cellular response. While admittedly faster and cheaper, these selective screens remain biased towards known 'canonical' gene families, and the field will only begin to realize the full potential of RNAi with an increased commitment to genome-wide screening. At the same time, research groups planning on undertaking siRNA screens should be fully aware of the potential pitfalls and difficulties that may arise when performing and interpreting

large-scale or genome-wide screens [133]. This may include issues of screen reproducibility, as demonstrated by two whole genome siRNA screens performed by the same research group within six months of each other where notable intra and inter-screen variability was reported [134], and the lack of overlapping hits in screens performed by different research groups studying HIV and influenza. Given recent concerns regarding whole genome RNAi screening studies as a method to identify potential drug targets [135] and reports from industry citing lack of reproducibility of primary data from the literature [136, 137], it is imperative that all steps of the screening process, including siRNA delivery, assay design, identification/confirmation of 'hits' and secondary screening assays are robust and validated if results from siRNA screens can be exploited for therapeutic purposes.

Another area where screening data is potentially compromised is in the use of sub-optimal target cells for screening. A prime example has been in several large-scale studies to identify host factors involved in the response to HIV and influenza infection, where use of a variety of 'easily transfectable' cell types has led to the identification of highly non-overlapping hit lists [70, 119, 120]. While siRNA delivery to more physiological target cells can be challenging, it is feasible with a commitment to rigorous protocol development, and we would argue is significantly more valuable in identifying physiologically relevant targets. While the above-mentioned screens have succeeded in identifying overlapping classes of host cellular pathways and processes involved in the response to important viral infections, they have been less useful in identifying specific candidates for therapeutic development. This question also applies to screens conducted in malignant B cell lines, where further work will be required to determine what proportion of hits has a significant role in primary tumor B cells, which are genetically heterogenous in nature and may have varied oncogene dependence

This brings up the challenge of translating primary data from siRNA screens into potential drug targets, and the question of whether small molecule inhibitors of screen hits, or siRNAs themselves, hold the most promise as drug candidates. The complexity of many diseases and the mutation frequency of pathogens suggest a combinatory approach may be most productive, and thus the identification of multiple potential gene targets in large-scale siRNA screens may be particularly attractive in guiding approaches to such combined therapies. For example, elucidation of the mechanism of HIV viral entry and replication using several approaches, including RNAi, has led to trials of a combinatory anti-viral therapy using hematopoietic stem cells transduced with a lentiviral vector encoding an

shRNA targeting HIV tat/rev, in combination with a ribozyme targeting CCR5 (a co-receptor for HIV entry) and an RNA decoy to the HIV TAR sequence [138].

We predict there will be numerous future opportunities to exploit RNAi screening and the new knowledge it provides for the benefit of human health in the immunological field, particularly through identification of anti-viral and anti-inflammatory therapies, targeting of malignant blood cancers, identification of host factors regulating bacterial diseases, elucidation of regulatory steps in protective immunity and identification of molecular mechanisms underlying immune disorders. It is anticipated that increased understanding of RNAi in mammalian cells and the continual improvements made in siRNA screening will see this technology realizing its full potential for the discovery of novel gene function and for therapeutic purposes.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors confirm that this chapter contents have no conflict of interest.

REFERENCES

[1]   Sachse C, Echeverri CJ. Oncology studies using siRNA libraries: the dawn of RNAi-based genomics. Oncogene. 2004;23(51):8384-91. Epub 2004/11/02.
[2]   Ren YG, Wagner KW, Knee DA, Aza-Blanc P, Nasoff M, Deveraux QL. Differential regulation of the TRAIL death receptors DR4 and DR5 by the signal recognition particle. Molecular biology of the cell. 2004;15(11):5064-74. Epub 2004/09/10.
[3]   Silva J, Chang K, Hannon GJ, Rivas FV. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. Oncogene. 2004;23(51):8401-9. Epub 2004/11/02.
[4]   Akashi K, Traver D, Miyamoto T, Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature. 2000;404(6774):193-7. Epub 2000/03/21.
[5]   Kondo M, Weissman IL, Akashi K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell. 1997;91(5):661-72. Epub 1997/12/11.
[6]   Steinman RM, Inaba K. Myeloid dendritic cells. Journal of leukocyte biology. 1999;66(2):205-8. Epub 1999/08/17.
[7]   Cullen LM, Arndt GM. Genome-wide screening for gene function using RNAi in mammalian cells. Immunology and cell biology. 2005;83(3):217-23. Epub 2005/05/10.
[8]   Echeverri CJ, Perrimon N. High-throughput RNAi screening in cultured cells: a user's guide. Nature reviews Genetics. 2006;7(5):373-84. Epub 2006/04/12.
[9]   Zhang S, Zhao B, Jiang H, Wang B, Ma B. Cationic lipids and polymers mediated vectors for delivery of siRNA. Journal of controlled release : official journal of the Controlled Release Society. 2007;123(1):1-10. Epub 2007/08/25.

[10]   Arts GJ, Langemeijer E, Tissingh R, Ma L, Pavliska H, Dokic K, *et al*. Adenoviral vectors expressing siRNAs for discovery and validation of gene function. Genome research. 2003;13(10):2325-32. Epub 2003/09/17.

[11]   Rubinson DA, Dillon CP, Kwiatkowski AV, Sievers C, Yang L, Kopinja J, *et al*. A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. Nature genetics. 2003;33(3):401-6. Epub 2003/02/19.

[12]   Persengiev SP, Zhu X, Green MR. Nonspecific, concentration-dependent stimulation and repression of mammalian gene expression by small interfering RNAs (siRNAs). RNA. 2004;10(1):12-8. Epub 2003/12/19.

[13]   Sledz CA, Holko M, de Veer MJ, Silverman RH, Williams BR. Activation of the interferon system by short-interfering RNAs. Nature cell biology. 2003;5(9):834-9. Epub 2003/08/28.

[14]   Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature. 2001;411(6836):494-8. Epub 2001/05/25.

[15]   Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. Nature. 2001;413(6857):732-8. Epub 2001/10/19.

[16]   Heil F, Hemmi H, Hochrein H, Ampenberger F, Kirschning C, Akira S, *et al*. Species-specific recognition of single-stranded RNA *via* toll-like receptor 7 and 8. Science. 2004;303(5663):1526-9. Epub 2004/02/21.

[17]   Lund JM, Alexopoulou L, Sato A, Karow M, Adams NC, Gale NW, *et al*. Recognition of single-stranded RNA viruses by Toll-like receptor 7. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(15):5598-603. Epub 2004/03/23.

[18]   Judge AD, Sood V, Shaw JR, Fang D, McClintock K, MacLachlan I. Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. Nature biotechnology. 2005;23(4):457-62. Epub 2005/03/22.

[19]   Cekaite L, Furset G, Hovig E, Sioud M. Gene expression analysis in blood cells in response to unmodified and 2'-modified siRNAs reveals TLR-dependent and independent effects. Journal of molecular biology. 2007;365(1):90-108. Epub 2006/10/24.

[20]   Takeda K, Akira S. Toll-like receptors in innate immunity. International immunology. 2005;17(1):1-14. Epub 2004/12/09.

[21]   Sioud M. Innate sensing of self and non-self RNAs by Toll-like receptors. Trends in molecular medicine. 2006;12(4):167-76. Epub 2006/03/15.

[22]   Sioud M. Induction of inflammatory cytokines and interferon responses by double-stranded and single-stranded siRNAs is sequence-dependent and requires endosomal localization. Journal of molecular biology. 2005;348(5):1079-90. Epub 2005/04/28.

[23]   Sioud M. Single-stranded small interfering RNA are more immunostimulatory than their double-stranded counterparts: a central role for 2'-hydroxyl uridines in immune responses. European journal of immunology. 2006;36(5):1222-30. Epub 2006/04/13.

[24]   Garg S, Sharma M, Ung C, Tuli A, Barral DC, Hava DL, *et al*. Lysosomal trafficking, antigen presentation, and microbial killing are controlled by the Arf-like GTPase Arl8b. Immunity. 2011;35(2):182-93.

[25]   Lonez C, Vandenbranden M, Ruysschaert JM. Cationic lipids activate intracellular signaling pathways. Advanced drug delivery reviews. 2012;64(15):1749-58.

[26]   Bracken AP, Helin K. Polycomb group proteins: navigators of lineage pathways led astray in cancer. Nature reviews Cancer. 2009;9(11):773-84.

[27]   Sauvageau M, Sauvageau G. Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. Cell stem cell. 2010;7(3):299-313.

[28]   Birmingham A, Selfors LM, Forster T, Wrobel D, Kennedy CJ, Shanks E, *et al*. Statistical methods for analysis of high-throughput RNA interference screens. Nature methods. 2009;6(8):569-75. Epub 2009/08/01.

[29]   Echeverri CJ, Beachy PA, Baum B, Boutros M, Buchholz F, Chanda SK, *et al*. Minimizing the risk of reporting false positives in large-scale RNAi screens. Nature methods. 2006;3(10):777-9. Epub 2006/09/23.

[30]   Carralot JP, Kim TK, Lenseigne B, Boese AS, Sommer P, Genovesio A, *et al*. Automated high-throughput siRNA transfection in raw 264.7 macrophages: a case study for optimization procedure. Journal of biomolecular screening. 2009;14(2):151-60. Epub 2009/02/07.

[31]   Jayaswal S, Kamal MA, Dua R, Gupta S, Majumdar T, Das G, *et al*. Identification of host-dependent survival factors for intracellular Mycobacterium tuberculosis through an siRNA screen. PLoS pathogens. 2010;6(4):e1000839. Epub 2010/04/27.

[32]   Kumar D, Nath L, Kamal MA, Varshney A, Jain A, Singh S, *et al*. Genome-wide analysis of the host intracellular network that regulates survival of Mycobacterium tuberculosis. Cell. 2010;140(5):731-43. Epub 2010/03/10.

[33]   Cho YS, Challa S, Moquin D, Genga R, Ray TD, Guildford M, *et al*. Phosphorylation-driven assembly of the RIP1-RIP3 complex regulates programmed necrosis and virus-induced inflammation. Cell. 2009;137(6):1112-23.

[34]   Oliveira DM, Goodell MA. Transient RNA interference in hematopoietic progenitors with functional consequences. Genesis. 2003;36(4):203-8. Epub 2003/08/21.

[35]   Walters DK, Jelinek DF. The effectiveness of double-stranded short inhibitory RNAs (siRNAs) may depend on the method of transfection. Antisense & nucleic acid drug development. 2002;12(6):411-8. Epub 2003/02/06.

[36]   Weil D, Garcon L, Harper M, Dumenil D, Dautry F, Kress M. Targeting the kinesin Eg5 to monitor siRNA transfection in mammalian cells. BioTechniques. 2002;33(6):1244-8. Epub 2002/12/31.

[37]   Dunne J, Drescher B, Riehle H, Hadwiger P, Young BD, Krauter J, *et al*. The apparent uptake of fluorescently labeled siRNAs by electroporated cells depends on the fluorochrome. Oligonucleotides. 2003;13(5):375-80. Epub 2004/03/06.

[38]   Jiang ZY, Zhou QL, Coleman KA, Chouinard M, Boese Q, Czech MP. Insulin signaling through Akt/protein kinase B analyzed by small interfering RNA-mediated gene silencing. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(13):7569-74. Epub 2003/06/17.

[39]   Wilson JA, Jayasena S, Khvorova A, Sabatinos S, Rodrigue-Gervais IG, Arya S, *et al*. RNA interference blocks gene expression and RNA synthesis from hepatitis C replicons propagated in human liver cells. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(5):2783-8. Epub 2003/02/21.

[40]   Herweijer H, Wolff JA. Progress and prospects: naked DNA gene transfer and therapy. Gene therapy. 2003;10(6):453-8. Epub 2003/03/07.

[41]   Ovcharenko D, Jarvis R, Hunicke-Smith S, Kelnar K, Brown D. High-throughput RNAi screening *in vitro*: from cell lines to primary cells. RNA. 2005;11(6):985-93. Epub 2005/06/01.

[42]   Tyner JW, Deininger MW, Loriaux MM, Chang BH, Gotlib JR, Willis SG, *et al*. RNAi screen for rapid therapeutic target identification in leukemia patients. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(21):8695-700. Epub 2009/05/13.

[43]   Zhang J, Mendoza M, Guiraldelli MF, Barbu EA, Siraganian RP. Small interfering RNA screen for phosphatases involved in IgE-mediated mast cell degranulation. J Immunol. 2010;184(12):7178-85.

[44]   McManus MT, Haines BB, Dillon CP, Whitehurst CE, van Parijs L, Chen J, *et al*. Small interfering RNA-mediated gene silencing in T lymphocytes. J Immunol. 2002;169(10):5754-60. Epub 2002/11/08.

[45]   Brummelkamp TR, Bernards R, Agami R. A system for stable expression of short interfering RNAs in mammalian cells. Science. 2002;296(5567):550-3. Epub 2002/03/23.

[46]   Silva JM, Li MZ, Chang K, Ge W, Golding MC, Rickles RJ, *et al*. Second-generation shRNA libraries covering the mouse and human genomes. Nature genetics. 2005;37(11):1281-8. Epub 2005/10/04.

[47]   Moffat J, Sabatini DM. Building mammalian signalling pathways with RNAi screens. Nature reviews Molecular cell biology. 2006;7(3):177-87. Epub 2006/02/24.

[48]   Bridge AJ, Pebernard S, Ducraux A, Nicoulaz AL, Iggo R. Induction of an interferon response by RNAi vectors in mammalian cells. Nature genetics. 2003;34(3):263-4. Epub 2003/06/11.

[49]   Fish RJ, Kruithof EK. Short-term cytotoxic effects and long-term instability of RNAi delivered using lentiviral vectors. BMC molecular biology. 2004;5:9. Epub 2004/08/05.

[50]   Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, *et al*. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. Nature. 2006;441(7092):537-41. Epub 2006/05/26.

[51]   McBride JL, Boudreau RL, Harper SQ, Staber PD, Monteys AM, Martins I, *et al*. Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(15):5868-73. Epub 2008/04/10.

[52]   Moita CF, Chora A, Hacohen N, Moita LF. RNAi screen for kinases and phosphatases that play a role in antigen presentation by dendritic cells. European journal of immunology. 2012;42(7):1843-9. Epub 2012/05/16.

[53]   Ngo VN, Davis RE, Lamy L, Yu X, Zhao H, Lenz G, *et al*. A loss-of-function RNA interference screen for molecular targets in cancer. Nature. 2006;441(7089):106-10. Epub 2006/03/31.

[54]   Luo B, Cheung HW, Subramanian A, Sharifnia T, Okamoto M, Yang X, *et al*. Highly parallel identification of essential genes in cancer cells. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(51):20380-5. Epub 2008/12/19.

[55]   Shin KJ, Wall EA, Zavzavadjian JR, Santat LA, Liu J, Hwang JI, *et al*. A single lentiviral vector platform for microRNA-based conditional RNA interference and coordinated transgene expression. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(37):13759-64.

[56]   Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H, Yu EY, An DS, *et al*. Lentivirus-delivered stable gene silencing by RNAi in primary cells. RNA. 2003;9(4):493-501. Epub 2003/03/22.

[57]   Larsen HO, Roug AS, Nielsen K, Sondergaard CS, Hokland P. Nonviral transfection of leukemic primary cells and cells lines by siRNA-a direct comparison between Nucleofection and Accell delivery. Experimental hematology. 2011;39(11):1081-9. Epub 2011/08/23.

[58]   Kanemura Y, Mori H, Kobayashi S, Islam O, Kodama E, Yamamoto A, *et al*. Evaluation of *in vitro* proliferative activity of human fetal neural stem/progenitor cells using indirect measurements of viable cells based on cellular metabolic activity. Journal of neuroscience research. 2002;69(6):869-79.

[59]   Brechmann M, Mock T, Nickles D, Kiessling M, Weit N, Breuer R, *et al*. A PP4 holoenzyme balances physiological and oncogenic nuclear factor-kappa B signaling in T lymphocytes. Immunity. 2012;37(4):697-708.

[60]   Cho JS, Kim YC, Morrison SL. Inhibitors of MyD88-dependent proinflammatory cytokine production identified utilizing a novel RNA interference screening approach. PloS one. 2009;4(9):e7029. Epub 2009/09/16.

[61]   Tzircotis G, Braga VM, Caron E. RhoG is required for both FcgammaR- and CR3-mediated phagocytosis. Journal of cell science. 2011;124(Pt 17):2897-902. Epub 2011/09/01.

[62]   Carlin LM, Evans R, Milewicz H, Fernandes L, Matthews DR, Perani M, *et al*. A targeted siRNA screen identifies regulators of Cdc42 activity at the natural killer cell immunological synapse. Science signaling. 2011;4(201):ra81. Epub 2011/12/01.

[63]   Zhou H, DeLoid G, Browning E, Gregory DJ, Tan F, Bedugnis AS, *et al*. Genome-wide RNAi screen in IFN-gamma-treated human macrophages identifies genes mediating resistance to the intracellular pathogen Francisella tularensis. PloS one. 2012;7(2):e31752. Epub 2012/02/24.

[64]   Zou L, Zhou J, Zhang J, Li J, Liu N, Chai L, *et al*. The GTPase Rab3b/3c-positive recycling vesicles are involved in cross-presentation in dendritic cells. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(37):15801-6. Epub 2009/09/01.

[65]   Moura-Alves P, Neves-Costa A, Raquel H, Pacheco TR, D'Almeida B, Rodrigues R, *et al*. An shRNA-based screen of splicing regulators identifies SFRS3 as a negative regulator of IL-1beta secretion. PloS one. 2011;6(5):e19829. Epub 2011/05/26.

[66]   Bellucci R, Nguyen HN, Martin A, Heinrichs S, Schinzel AC, Hahn WC, *et al*. Tyrosine kinase pathways modulate tumor susceptibility to natural killer cells. The Journal of clinical investigation. 2012;122(7):2369-83. Epub 2012/06/12.

[67]   Hohman RJ, Dreskin SC. Measuring degranulation of mast cells. Current protocols in immunology / edited by John E Coligan [et al]. 2001;Chapter 7:Unit 7 26. Epub 2008/04/25.

[68]   Wang Z, Kumamoto Y, Wang P, Gan X, Lehmann D, Smrcka AV, *et al*. Regulation of immature dendritic cell migration by RhoA guanine nucleotide exchange factor Arhgef5. The Journal of biological chemistry. 2009;284(42):28599-606. Epub 2009/08/29.

[69]   Colvin RA, Means TK, Diefenbach TJ, Moita LF, Friday RP, Sever S, *et al*. Synaptotagmin-mediated vesicle fusion regulates cell migration. Nature immunology. 2010;11(6):495-502. Epub 2010/05/18.

[70]   Yeung ML, Houzet L, Yedavalli VS, Jeang KT. A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication. The Journal of biological chemistry. 2009;284(29):19463-73. Epub 2009/05/23.

[71]   Bonilla FA, Oettgen HC. Adaptive immunity. The Journal of allergy and clinical immunology. 2010;125(2 Suppl 2):S33-40. Epub 2010/01/12.

[72]   Norman MU, Hickey MJ. Mechanisms of lymphocyte migration in autoimmune disease. Tissue antigens. 2005;66(3):163-72. Epub 2005/08/17.

[73]   Islam SA, Luster AD. T cell homing to epithelial barriers in allergic disease. Nature medicine. 2012;18(5):705-15. Epub 2012/05/09.

[74]   Wood KJ, Goto R. Mechanisms of rejection: current perspectives. Transplantation. 2012;93(1):1-10. Epub 2011/12/06.

[75]   Singh SK. RNA interference and its therapeutic potential against HIV infection. Expert opinion on biological therapy. 2008;8(4):449-61. Epub 2008/03/21.

[76]   Oberdoerffer S, Moita LF, Neems D, Freitas RP, Hacohen N, Rao A. Regulation of CD45 alternative splicing by heterogeneous ribonucleoprotein, hnRNPLL. Science. 2008;321(5889):686-91. Epub 2008/08/02.

[77]   Astier AL, Beriou G, Eisenhaure TM, Anderton SM, Hafler DA, Hacohen N. RNA interference screen in primary human T cells reveals FLT3 as a modulator of IL-10 levels. J Immunol. 2010;184(2):685-93. Epub 2009/12/19.

[78]   Kickler K, Maltby K, Ni Choileain S, Stephen J, Wright S, Hafler DA, *et al*. Prostaglandin E2 affects T cell responses through modulation of CD46 expression. J Immunol. 2012;188(11):5303-10. Epub 2012/05/01.

[79]   Heasman SJ, Carlin LM, Cox S, Ng T, Ridley AJ. Coordinated RhoA signaling at the leading edge and uropod is required for T cell transendothelial migration. The Journal of cell biology. 2010;190(4):553-63. Epub 2010/08/25.

[80]   Mavrakis KJ, Wolfe AL, Oricchio E, Palomero T, de Keersmaecker K, McJunkin K, *et al*. Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia. Nature cell biology. 2010;12(4):372-9. Epub 2010/03/02.

[81]   Sanda T, Tyner JW, Gutierrez A, Ngo VN, Glover J, Chang BH, *et al*. TYK2-STAT1-BCL2 Pathway Dependence in T-Cell Acute Lymphoblastic Leukemia. Cancer Discov. 2013. Epub 2013/03/09.

[82]   Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, Xavier RJ, *et al*. Identification of host proteins required for HIV infection through a functional genomic screen. Science. 2008;319(5865):921-6. Epub 2008/01/12.

[83]   Konig R, Zhou Y, Elleder D, Diamond TL, Bonamy GM, Irelan JT, *et al*. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Cell. 2008;135(1):49-60. Epub 2008/10/16.

[84]   Zhou H, Xu M, Huang Q, Gates AT, Zhang XD, Castle JC, *et al*. Genome-scale RNAi screen for host factors required for HIV replication. Cell host & microbe. 2008;4(5):495-504. Epub 2008/11/04.

[85]   Pache L, Konig R, Chanda SK. Identifying HIV-1 host cell factors by genome-scale RNAi screening. Methods. 2011;53(1):3-12. Epub 2010/07/27.

[86]   Liu L, Oliveira NM, Cheney KM, Pade C, Dreja H, Bergin AM, *et al*. A whole genome screen for HIV restriction factors. Retrovirology. 2011;8:94. Epub 2011/11/16.

[87]   Rato S, Maia S, Brito PM, Resende L, Pereira CF, Moita C, *et al*. Novel HIV-1 knockdown targets identified by an enriched kinases/phosphatases shRNA library using a long-term iterative screen in Jurkat T-cells. PloS one. 2010;5(2):e9276. Epub 2010/02/23.

[88]   Chaplin DD. Overview of the immune response. The Journal of allergy and clinical immunology. 2010;125(2 Suppl 2):S3-23. Epub 2010/03/05.

[89]   Shaffer AL, 3rd, Young RM, Staudt LM. Pathogenesis of human B cell lymphomas. Annual review of immunology. 2012;30:565-610. Epub 2012/01/10.

[90]   Bidere N, Ngo VN, Lee J, Collins C, Zheng L, Wan F, *et al*. Casein kinase 1alpha governs antigen-receptor-induced NF-kappaB activation and human lymphoma cell survival. Nature. 2009;458(7234):92-6. Epub 2009/01/02.

[91] Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, *et al*. Oncogenically active MYD88 mutations in human lymphoma. Nature. 2011;470(7332):115-9. Epub 2010/12/24.

[92] Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PB, *et al*. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. Nature. 2010;463(7277):88-92. Epub 2010/01/08.

[93] Lam LT, Davis RE, Ngo VN, Lenz G, Wright G, Xu W, *et al*. Compensatory IKKalpha activation of classical NF-kappaB signaling during IKKbeta inhibition identified by an RNA interference sensitization screen. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(52):20798-803. Epub 2008/12/24.

[94] Schmitz R, Young RM, Ceribelli M, Jhavar S, Xiao W, Zhang M, *et al*. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. Nature. 2012;490(7418):116-20. Epub 2012/08/14.

[95] Ruiz-Vela A, Aggarwal M, de la Cueva P, Treda C, Herreros B, Martin-Perez D, *et al*. Lentiviral (HIV)-based RNA interference screen in human B-cell receptor regulatory networks reveals MCL1-induced oncogenic pathways. Blood. 2008;111(3):1665-76. Epub 2007/11/23.

[96] Oricchio E, Nanjangud G, Wolfe AL, Schatz JH, Mavrakis KJ, Jiang M, *et al*. The Eph-receptor A7 is a soluble tumor suppressor for follicular lymphoma. Cell. 2011;147(3):554-64.

[97] Shaffer AL, Emre NC, Lamy L, Ngo VN, Wright G, Xiao W, *et al*. IRF4 addiction in multiple myeloma. Nature. 2008;454(7201):226-31. Epub 2008/06/24.

[98] Tiedemann RE, Zhu YX, Schmidt J, Shi CX, Sereduk C, Yin H, *et al*. Identification of molecular vulnerabilities in human multiple myeloma cells by RNA interference lethality screening of the druggable genome. Cancer research. 2012;72(3):757-68. Epub 2011/12/08.

[99] Tiedemann RE, Zhu YX, Schmidt J, Yin H, Shi CX, Que Q, *et al*. Kinome-wide RNAi studies in human multiple myeloma identify vulnerable kinase targets, including a lymphoid-restricted kinase, GRK6. Blood. 2010;115(8):1594-604. Epub 2009/12/10.

[100] Zhu YX, Tiedemann R, Shi CX, Yin H, Schmidt JE, Bruins LA, *et al*. RNAi screen of the druggable genome identifies modulators of proteasome inhibitor sensitivity in myeloma including CDK5. Blood. 2011;117(14):3847-57. Epub 2011/02/04.

[101] Farfsing A, Engel F, Seiffert M, Hartmann E, Ott G, Rosenwald A, *et al*. Gene knockdown studies revealed CCDC50 as a candidate gene in mantle cell lymphoma and chronic lymphocytic leukemia. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK. 2009;23(11):2018-26. Epub 2009/07/31.

[102] Ramsay AG, Clear AJ, Fatah R, Gribben JG. Multiple inhibitory ligands induce impaired T-cell immunologic synapse function in chronic lymphocytic leukemia that can be blocked with lenalidomide: establishing a reversible immune evasion mechanism in human cancer. Blood. 2012;120(7):1412-21. Epub 2012/05/02.

[103] Meacham CE, Ho EE, Dubrovsky E, Gertler FB, Hemann MT. *In vivo* RNAi screening identifies regulators of actin dynamics as key determinants of lymphoma progression. Nature genetics. 2009;41(10):1133-7. Epub 2009/09/29.

[104] Bric A, Miething C, Bialucha CU, Scuoppo C, Zender L, Krasnitz A, *et al*. Functional identification of tumor-suppressor genes through an *in vivo* RNA interference screen in a mouse lymphoma model. Cancer cell. 2009;16(4):324-35. Epub 2009/10/06.

[105] Traver D, Akashi K, Manz M, Merad M, Miyamoto T, Engleman EG, *et al*. Development of CD8alpha-positive dendritic cells from a common myeloid progenitor. Science. 2000;290(5499):2152-4. Epub 2000/12/16.

[106] Kushwah R, Hu J. Complexity of dendritic cell subsets and their function in the host immune system. Immunology. 2011;133(4):409-19. Epub 2011/06/02.

[107] Steinman RM. The dendritic cell system and its role in immunogenicity. Annual review of immunology. 1991;9:271-96. Epub 1991/01/01.

[108] Paul P, van den Hoorn T, Jongsma ML, Bakker MJ, Hengeveld R, Janssen L, *et al*. A Genome-wide multidimensional RNAi screen reveals pathways controlling MHC class II antigen presentation. Cell. 2011;145(2):268-83. Epub 2011/04/05.

[109] Heath WR, Belz GT, Behrens GM, Smith CM, Forehan SP, Parish IA, *et al*. Cross-presentation, dendritic cell subsets, and the generation of immunity to cellular antigens. Immunological reviews. 2004;199:9-26. Epub 2004/07/06.

[110] Bowles R, Patil S, Pincas H, Sealfon SC. Validation of efficient high-throughput plasmid and siRNA transfection of human monocyte-derived dendritic cells without cell maturation. Journal of immunological methods. 2010;363(1):21-8. Epub 2010/09/30.

[111] Amit I, Garber M, Chevrier N, Leite AP, Donner Y, Eisenhaure T, *et al*. Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science. 2009;326(5950):257-63.

[112] Chevrier N, Mertins P, Artyomov MN, Shalek AK, Iannacone M, Ciaccio MF, *et al*. Systematic discovery of TLR signaling components delineates viral-sensing circuits. Cell. 2011;147(4):853-67.

[113] Auffray C, Sieweke MH, Geissmann F. Blood monocytes: development, heterogeneity, and relationship with dendritic cells. Annual review of immunology. 2009;27:669-92. Epub 2009/01/10.

[114] Dale DC, Boxer L, Liles WC. The phagocytes: neutrophils and monocytes. Blood. 2008;112(4):935-45. Epub 2008/08/08.

[115] Verani A, Gras G, Pancino G. Macrophages and HIV-1: dangerous liaisons. Molecular immunology. 2005;42(2):195-212. Epub 2004/10/19.

[116] Halstead SB, Mahalingam S, Marovich MA, Ubol S, Mosser DM. Intrinsic antibody-dependent enhancement of microbial infection in macrophages: disease regulation by immune complexes. The Lancet infectious diseases. 2010;10(10):712-22. Epub 2010/10/05.

[117] Zhong XY, Wang P, Han J, Rosenfeld MG, Fu XD. SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. Molecular cell. 2009;35(1):1-10. Epub 2009/07/15.

[118] Groves E, Dart AE, Covarelli V, Caron E. Molecular mechanisms of phagocytic uptake in mammalian cells. Cellular and molecular life sciences : CMLS. 2008;65(13):1957-76. Epub 2008/03/07.

[119] Meliopoulos VA, Andersen LE, Birrer KF, Simpson KJ, Lowenthal JW, Bean AG, *et al*. Host gene targets for novel influenza therapies elucidated by high-throughput RNA interference screens. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2012;26(4):1372-86. Epub 2012/01/17.

[120] Bushman FD, Malani N, Fernandes J, D'Orso I, Cagney G, Diamond TL, *et al*. Host cell factors in HIV replication: meta-analysis of genome-wide studies. PLoS pathogens. 2009;5(5):e1000437. Epub 2009/05/30.

[121] Qian BZ, Pollard JW. Macrophage diversity enhances tumor progression and metastasis. Cell. 2010;141(1):39-51. Epub 2010/04/08.

[122] Ley S, Weigert A, Heriche JK, Mille-Baker B, Janssen RA, Brune B. RNAi screen in apoptotic cancer cell-stimulated human macrophages reveals co-regulation of IL-6/IL-10 expression. Immunobiology. 2012. Epub 2012/03/27.

[123] Nobes CD, Hall A. Rho GTPases control polarity, protrusion, and adhesion during cell movement. The Journal of cell biology. 1999;144(6):1235-44.

[124] Weber KS, Klickstein LB, Weber PC, Weber C. Chemokine-induced monocyte transmigration requires cdc42-mediated cytoskeletal changes. European journal of immunology. 1998;28(7):2245-51.

[125] Kalesnikoff J, Galli SJ. New developments in mast cell biology. Nature immunology. 2008;9(11):1215-23.

[126] Wilson BS, Pfeiffer JR, Oliver JM. FcepsilonRI signaling observed from the inside of the mast cell membrane. Molecular immunology. 2002;38(16-18):1259-68.

[127] Kumar P, Ban HS, Kim SS, Wu H, Pearson T, Greiner DL, *et al*. T cell-specific siRNA delivery suppresses HIV-1 infection in humanized mice. Cell. 2008;134(4):577-86. Epub 2008/08/12.

[128] Peer D, Zhu P, Carman CV, Lieberman J, Shimaoka M. Selective gene silencing in activated leukocytes by targeting siRNAs to the integrin lymphocyte function-associated antigen-1. Proc Natl Acad Sci U S A. 2007;104(10):4095-100. Epub 2007/03/16.

[129] Peer D, Park EJ, Morishita Y, Carman CV, Shimaoka M. Systemic leukocyte-directed siRNA delivery revealing cyclin D1 as an anti-inflammatory target. Science. 2008;319(5863):627-30. Epub 2008/02/02.

[130] Kim SS, Peer D, Kumar P, Subramanya S, Wu H, Asthana D, *et al*. RNAi-mediated CCR5 silencing by LFA-1-targeted nanoparticles prevents HIV infection in BLT mice. Mol Ther. 2010;18(2):370-6. Epub 2009/12/10.

[131] Neff CP, Zhou J, Remling L, Kuruvilla J, Zhang J, Li H, *et al*. An aptamer-siRNA chimera suppresses HIV-1 viral loads and protects from helper CD4(+) T cell decline in humanized mice. Sci Transl Med. 2011;3(66):66ra6. Epub 2011/01/21.

[132] Wheeler LA, Trifonova R, Vrbanac V, Basar E, McKernan S, Xu Z, *et al*. Inhibition of HIV transmission in human cervicovaginal explants and humanized mice using CD4 aptamer-siRNA chimeras. J Clin Invest. 2011;121(6):2401-12. Epub 2011/05/18.

[133] Jiang M, Instrell R, Saunders B, Berven H, Howell M. Tales from an academic RNAi screening facility; FAQs. Brief Funct Genomics. 2011;10(4):227-37. Epub 2011/04/30.

[134] Barrows NJ, Le Sommer C, Garcia-Blanco MA, Pearson JL. Factors affecting reproducibility between genome-scale siRNA-based screens. J Biomol Screen. 2010;15(7):735-47. Epub 2010/07/14.

[135] Kaelin WG, Jr. Molecular biology. Use and abuse of RNAi to study mammalian gene function. Science. 2012;337(6093):421-2. Epub 2012/07/28.

[136] Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. Nature. 2012;483(7391):531-3. Epub 2012/03/31.

[137] Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? Nat Rev Drug Discov. 2011;10(9):712. Epub 2011/09/06.

[138] DiGiusto DL, Krishnan A, Li L, Li H, Li S, Rao A, *et al*. RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. Science translational medicine. 2010;2(36):36ra43. Epub 2010/06/18.

siRNA Microarray-Based Genomic Screening

Yong-Jun Kwon, HiChul Kim, Jin Y. Kim, Namyoul Kim, Jinyeoung Heo, TaeKyu Lee, Michael A.E. Hansen and Veronica Soloveva*

Institute Pasteur Korea, 696 Sampyeong-dong, Bundang-gu, Gyeonggi-do, 463-400, South Korea

Abstract: The following chapter describes a miniaturized array-based platform for RNAi screening. The PhenomicID™ platform combines high density spotted microarray technology, high content imaging of the cells, and computational algorithms for image and data analysis. This platform provides an efficient and cost effective way to use reverse genetic tools for analysis of mammalian gene functions on a genome-wide scale. The miniaturization of this process allows for experimental complexity and applications not previously feasible when performed in a well-based format. Our team has employed this technology to identify functional genes involved in the progression of infectious diseases, in cell-based infectious models for HIV, Chagas, Dengue virus, Chikungunya virus, and Influenza virus. Importantly, the power of miniaturized technologies is not limited to screening of RNAi libraries, but can allow performance of complex experiments - combining drug and siRNA treatments to identify drug targets. In addition the flexibility of the describe platform allows researchers to profile multiple primary cell lines in search for essential genes. In this chapter we will discuss practical guidance for developing microarray-based genome-wide library siRNA screening and its applications.

Keywords: Assay optimization, data analysis, image analysis, infectious disease, microarray printing, reverse transfection, siRNA screening, ultra-high throughput.

INTRODUCTION

Reverse Genetic Approach

Roughly a decade ago, two important findings in the areas of genetics and molecular biology were published by multiple groups: the first was a draft sequence of the human genome [1]. The second was the description of RNA interference (RNAi), an endogenous pathway that used small double-stranded oligonucleotides (referred to as small interfering RNAs, siRNA) to mediate post-

#**Corresponding author Veronica Soloveva:** Institute Pasteur Korea, 696 Sampyeong-dong, Bundang-gu, Gyeonggi-do, 463-400, South Korea; E-mail: VSoloveva@yahoo.com

transcriptional gene silencing in a broad variety of organisms [2-4]. These key discoveries brought forth a broad spectrum of new ideas and resulted in development of tools not only for basic research but diagnostic applications and also drug discovery.

As a tool, RNAi can now be applied to address a variety of biological questions [5]. The ability of siRNA to reduce or "knock down" levels of a targeted mRNA (and thus protein) in model biological systems has enabled the technology to become a reliable method for therapeutic target identification in screens aimed at identifying regulatory components of various biological pathways [6]. Depending upon the question being addressed, screening can be divided into two approaches. In the first, a defined set of target genes known to be operating in a specific biological pathway are tested. This type of study is typically limited to testing a well-defined collection of genes or gene families such as kinases, phosphatases, or GPCRs [7-8]. The second approach, which is the focus of this review, is represented by genome scale RNAi screening. This is typically performed using a phenotypic readout with the goal of comprehensively screening to identify genes that were previously not known to participate in the phenotype of under investigation.

With the publication of various drafts of genome sequences, libraries comprised of different sets of silencing reagents (*e.g.*, dsRNA, siRNA, shRNA) were developed to target the open reading frames in both model biological systems as well as human genomes [9-11].

To date, the most conventional uses of RNAi screens in drug discovery have been to identify genes that encode new druggable candidates [12]. The clear benefit of RNAi-based technology in drug discovery is the ability of siRNAs to mimic the inhibitory effect of a small molecule compound of interest on the target protein. Such strategies have been employed in screening campaigns for a variety of disease models including cancer [13, 14], inflammation [15, 16], and neurological (CNS) diseases [17]. RNAi screens have also been used to identify cellular host factors implicated in viral infectious diseases [18] including HIV [19-23], HCV [24-26], Influenza [27-29], WNV and Dengue [29] as well as non-viral parasitic pathogens [30].

Interestingly, while the initial RNAi screens identified new sets of genes related to the phenotypes under investigation, they also broadened our understanding of issues associated with the technology. For instance, screens performed in independent labs and designed to identify critical host factors for HIV [31] and

Influenza [32, 33] revealed substantial variability in the data and only a small overlap in the lists of essential genes. This analysis drew attention to the importance of the assay details and screen design, the selection of efficient and specific silencing reagents, and the incorporation of rigorous validation procedures to confirm primary screen hits and reduce false positives [5, 12, 34, and 35].

Detailed studies of RNAi reagents facilitated our understanding of the complexities associated with performing reverse genetics with siRNAs. One of the major issues is tied to off-target effects (OTEs), where the introduced siRNA modulates the expression of unintended gene targets [36-38]. Studies have shown that the most general mechanism associated with OTEs is siRNA "seed" sequences (nucleotides 2-7) induce a natural regulatory effect similar to endogenous microRNA [37]. Due to the tolerance of single nucleotide mismatches in this process, a single siRNA can potentially target hundreds of unrelated sequences [39, 40]. Separately, researchers have observed siRNA-mediated activation of Toll-Like Receptors (TLRs) and other cellular signaling pathways being facililtated by the presence of GU-rich motifs in introduced oligonucleotides [41, 42]. Due to these and related issues, siRNAs design improvements became a major goal of reagent-producing groups.

The transient nature of siRNA-mediated effects and their ease of delivery into cells allows flexibility in assay design and facilitates the targeting of genes in a range of cell lines. However, potential cell toxicity effects associated with the lipid transfection reagents used to introduce siRNAs into the intracellular space, and the high copy-number of siRNAs required for target silencing can produce false results [43-45]. To address these issues, academic and commercial organizations have optimized both the sequence and chemical modifications of siRNA, thereby minimizing micro-RNA-like off-target effects and increasing overall molecular stability. Simultaneously, chemical modifications have been identified that enable lipid-independent delivery of siRNAs [44, 46, 47].

Despite these efforts, RNAi screens appear to be inherently more variable and produce a broader range of hits than small molecule screens. Birmingham and colleagues [48] performed a statistical analysis comparing results from 13 RNAi screens and 19 small molecule HTS, and found that RNAi screens have, on average, a two-fold smaller assay window and twice the median coefficient of variance (26.5% for RNAi screens *vs* 13.4% for small molecules HTS). This observation emphasizes the importance in screen optimization to minimize variability and false positive results. One way to enhance the reproducibility of

screening results is to run a statistically significant number of replicates in conjunction with multiple relevant experimental controls [5, 47, 34]. To this end, we have developed stringent approaches for designing and performing screens on and analyzing data from high-density live cell microarrays.

High Density Live Cell Microarrays

High density spot microarrays carrying cDNAs, beads, oligonucletides, protein, tissues samples, antibodies and chemical compounds have been used by several groups in both research and diagnostic settings (see, for example, Schena *et al.* [49]). Microarrays allow thousands of individual experiments to be performed on multiple samples under well-controlled, identical conditions (as opposed to the limited number associated with 94, 384, and 1536 well formats). The uniformity of the platform conditions enhances the reproducibility and statistical reliability of individual observation.

There are many techniques described in literature for producing arrays with the desirable characteristics specific for the materials involved in study [50]. The first application of microarrays for introducing genetic material into cells was reported by Ziauddin J. and Sabatini M. in 2001 [51]. One hundred ninety-two plasmids expressing different cDNAs were combined with gelatin for cell adhesion and arrayed as 120-150μm spots on glass slides. When cells were mixed with transfection reagent and seeded on the top of slides, they adhered as clusters in the area of printed spots and showed good uptake of plasmid material and expression of plasmid-encoded proteins in the cells localized on the printed spots. This work demonstrated the throughput and flexibility of microarray formats for phenotypic cell-based assays, expanding the platform from its traditional use in DNA/cDNA hybridization experiments [52].

Despite the advantages mentioned above, microarray technology has its caveats. Although the reagents, including the siRNA libraries and microarray printers of all kinds are available commercially, the optimization of all processes associated with library management, array fabrication, and data acquisition and analysis need to be planned thoroughly and require additional investments. Much effort is needed for tuning each step of the screening process before it becomes efficient and reliable. This is probably one of the strongest reasons why there are only few groups who have established the cell-based arrays fabrication and reported successful results in this format using cDNA [53, 54] or siRNA [55-59, and our group 23, 30].

DEVELOPMENT OF SIRNA MICROARRAY TECHNOLOGY FOR CELLULAR TRANSFECTION

Because the validation of candidate genes takes longer than the screening (similar to screening of small molecules) it is essential for the screening process to be rapid and to provide reliable data for analysis.

The Work Flow Process for Experiments with siRNA Microarrays

The overall flow of the process for microarray preparation and handling is similar to protocols describes by Erfle H. *et al.* 2007 [55] and illustrated in the Fig. **1**.

The initial step involves preparation of the siRNA library for printing by reformatting the siRNA material from original 96-well plate format into 384-well plates. This first miniaturization step is the main step for saving the reagents because only small fractions of total volume will be used from each well to print hundreds of spot replicates or copies. The reagent mixture includes siRNA, transfection reagents and a labeled siRNA-red tracer for downstream visualization of the printed spots. We use the Genomic Solution Omnigrid 100 high-throughput pin-tool based printer to deliver the mixture onto slides in well-defined spots. In all, 3,888 spots (including positive and negative controls) are positioned on one slide. This allows the complete siRNA library targeting 18,000 human genes to be printed on five slides. Slides are dried at 23°C at 50% humidity and can be stored in desiccated conditions for at least 1 year without significant loss of gene knockdown efficiency.

Once experiments commence, slides are placed in 4 chamber culture well plates and seeded with cells contained in ~5~8 ml of appropriate media (per slide). It has previously been shown that synthetic siRNA printed/ or plated in this fashion can be taken up by cells and produce a knock-down phenotype within 48-72 hours post plating [55, 56]. Subsequent phenotypic effects of transfection can be visualized using immuno-staining or marker expression (*e.g.*, GFP) and depend upon the half-life of the protein and its regulation in the cell line of choice. Following fixation (4% PFA) and staining, slides are placed in an MDS ImageXpress Ultra confocal plate reader for image acquisition. In our hands, spot localization is achieved by detecting a labeled siRNA (*e.g.* siGLO, Dharmacon, part of GE Healthcare) at 561nm. Another channel (635 nm) is used in conjunction with DRAQ5 staining to identify the position of nuclei. Finally, a third channel (488 nm) is available to detect fluorophores (such as FITC or GFP) introduced by immuno-staining, cellular dyes or expression of recombinant markers (Fig. **2**).
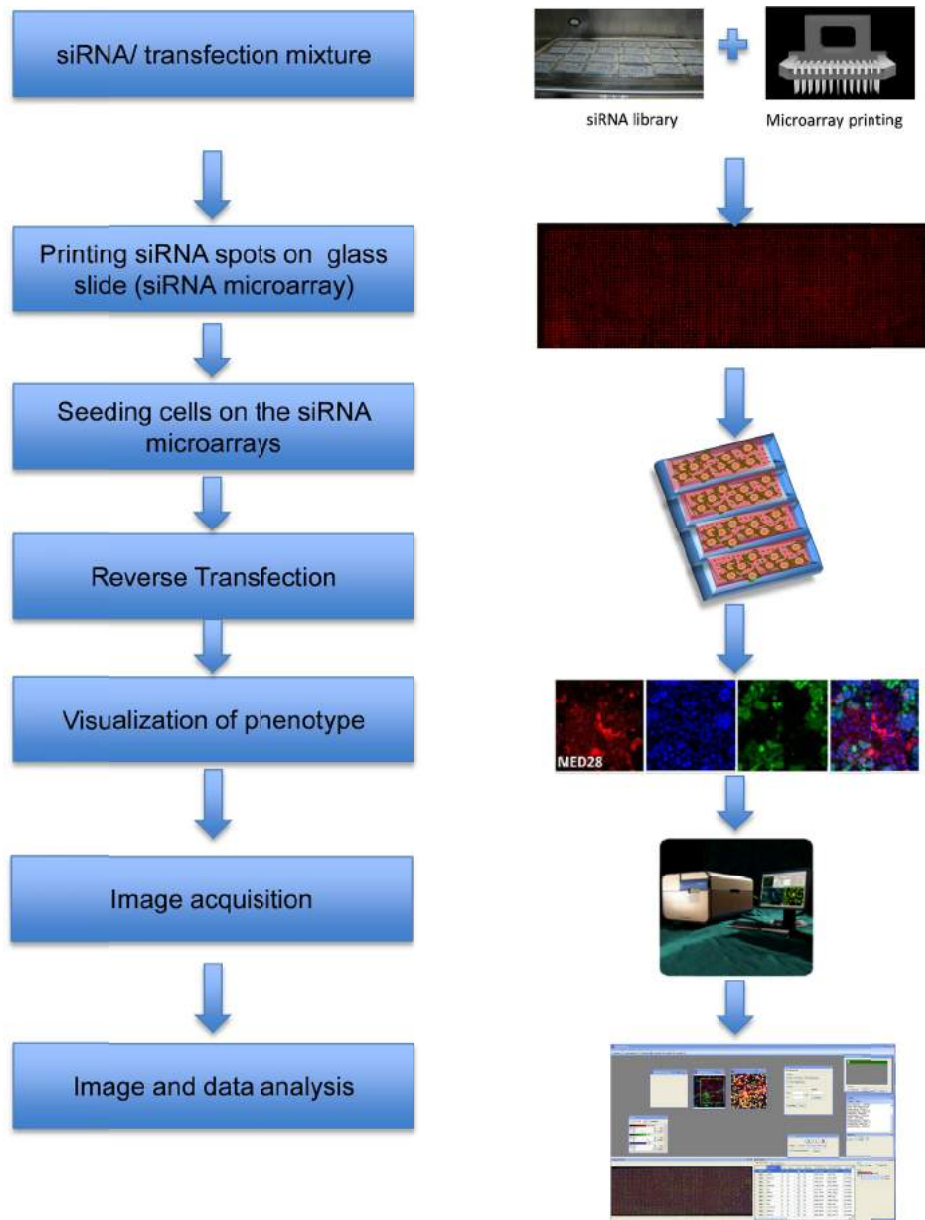
Figure 1: Process flow for cellular microarray-based siRNA experiments. Our work-flow comprises the following steps: 1) the siRNAs are mixed with the transfection reagent and Red-siRNA tracer, 2) the mixture is printed on the glass slide, 3) cells are seeded on the surface of the slide for reverse transfection, 4) cells are incubated, 5) the phenotype is detected by immunoassay, 6) images of the whole slide are acquired using a plate reader (MDS ImageXpress Ultra), and 7) the images and data are analyzed using proprietary and commercial software.
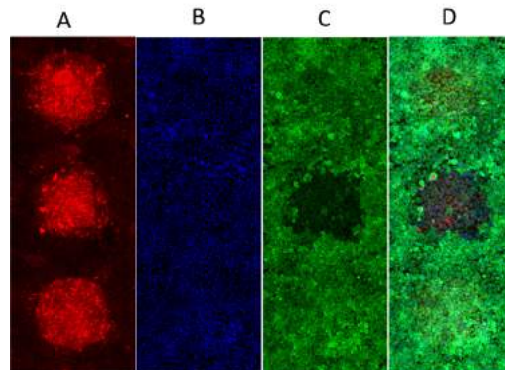
Figure 2: Images of microarray spots. Images were taken in 3 channels: (A) 561 nm for detection of siGLO-Red tracer mixed with transfection reagents and siRNA to visualize spot. (B) Deep red at 635 nm for detection DRAQ5 staining of nuclei material. (C) In this example at 488 nm for detection of Alexa 488 signal associated with immuno-staining of endogenous p65 (NFkB)-RelA. (D) integration of all 3 images. Scale bar= 100µm.

The image acquisition with a 10x objective generally takes an hour for each slide. During this time, the entire slide is scanned collecting total of 408 individual images. An image montage is then created using MetaXpress software and transferred to a server where a proprietary IPK algorithm (IM) is used to analyze the images. These images can also be analyzed by MDS MetaXpress or other alternative software packages. The corresponding meta-data generated by the algorithm is then organized by a custom-designed IM database created for IPK for data analysis. Alternatively, data analysis can also be performed using commercial software such as Excel, Graph Pad Prism or Matlab for more complex multi-parametrical analysis like PCA and multivariate analysis..

When beginning a new project or assay there are several critical factors that need to be thoroughly tested to assess the compatibility of microarray technology with screening goals and optimize conditions for microarray-based screening. The predominant steps that need to be validated include testing cell line compatibility with the microarray surface, the efficiency at which siRNAs are reverse-transfected into cells, optimization array printing and confirmation of the reliability and reproducibility of data. Below, the discussion is focused on the specific details involved in those four steps of screening preparation.

Cellular Adherence and Uniformity of Culture on Microarray Surface

The properties of the surface upon which the cells have been seeded plays a role in both cell growth and the development of signaling networks. Glass, which is

usually selected for microscopic experiments due to its excellent optical properties, combines a hydrophobic surface with a very hard tension (10,000 times harder than in any mammalian tissues), unsuitable for the direct seeding and culturing of mammalian cells. However there are several approaches that have been adopted that to facilitate cell culture on glass surfaces. These include: coating surfaces with defined synthetic molecules free of any animal products or with extracellular matrix proteins (ECM). The Epoxide and Gamma Amino Propyl Silane (GAPS) and more recently, MAS, MAS-GP, and aminopropylsilane (APS) are synthetic coatings particularly valuable for use under the harsh conditions employed during *in situ* hybridization experiments but was tested for cell culture as well [For more information check websites of Corning; BD; IBIDI; Matsunami Glass Ind.].

Materials that include collagen I, collagen IV, fibronectin, gelatin, laminin, matrigel matrix, poly-D-Lysine (PDL), and Poly-L-Lysine (PLL) mimic the *in vivo* cellular environment and can be readily applied in typical laboratories to any plastic or glass surface. However, for optimal consistency and quality control, it is best to purchase pre-coated cell culture ware from vendors who can ensure optimal surface quality and coating homogeneity.

We have successfully used PDL, PLL, collagen and MAS coatings to promote cell seeding for reverse transfection experiments on glass slides.

With regard to the slides employed in screening, we typically use traditional tissue culture slides (of 26x76 mm and 0.9-1.2 mm thickness) and N1 glass cover slips (24x60 mm and 0.16-0.19 μm thickness). The optimal thickness of the cover slips allows positioning on the reader in a direct orientation. In contrast, the 0.9-1.2 mm thick slides need to be mounted with cover glass and inverted for scanning by ImageXpress Ultra. However, working with the thin cover slips is complicated by their fragility, so in cases where this platform is employed, we use a custom-designed slide holder for printing the arrays and a slide chamber to perform immunostaining and image acquisition. Regular 1mm thick microscope slides pre-treated with different coating reagents do not required any special holders and readily available from a variety of vendors.

As the printed spots containing the siRNAs are arrayed across the complete surface of the slide, it is critical that cells uniformly cover the entire surface. Any gaps or empty spaces will result in an underrepresentation of gene replicates and potentially misleading results. Assessment of the quality of the cell layer can be

done by visual inspection (Fig. **3**) or by image analysis of selected picture-frames (discussed later, Fig. **14**).

After cells are confirmed to be compatible with culture conditions on the slide, the next step is to test how well the cells can be transfected. Because of the usual variation in cellular susceptibility to transfection, this test is often performed prior to testing it in the cells cultured on the slide. However the efficiency of transfection and mixture reagents used for siRNA printing also very much depends on the slide coating. If reagents applied to the spot will not adhere to the slide coating, the material may easily defuse from spot to spot across the slide.
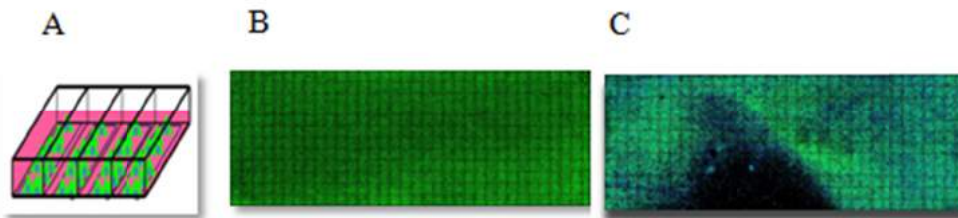
Figure 3: Visual assessment of uniformity for the cell monolayer on arrays. (**A**) The cells were seeded on glass slides in 4 well culture dishes and cultured for 48 h. Immuno-staining of endogenous p65 was used as a marker for cells (green color). Quality assessment of the cellular monolayer can be done by visual inspection. (**B**) Uniform cell layer (green color) on commercially available PLL– coated slide and (**C**) distorted cell layer due to the poor quality of coating.

Reverse Transfection of Cells on a Microarray

Transfection is the process by which nucleic acids (NAs) are introduced into cells. There are a wide range of methods by which transfection can be achieved, the most popular for RNAi screening involves treating cells with siRNAs encapsulated in cationic lipids [57, 58].

Lipid-mediated delivery of NAs can be achieved by one of two protocols; forward-transfection (FT) and reverse-transfection (RT) (Fig. **4**). For microarray-based RNAi screening, reverse transfection, where siRNA-lipid formulations spotted on slides at high densities are then overlaid with cells, is the method of choice [59]. Regardless of whether RT takes place in a well or microarray format, critical parameters include testing of the amenability of the cell line to reverse transfection, identifying preferred lipid transfection reagents and optimal siRNA-to-lipid ratio. At the same time, transfection on the slides has its own unique set of complications including the immobilization of nano-liter amounts of reagents in a

tightly defined area and averting migration of siRNAs from one area (*i.e.*, spot) to another during submerging of arrays in cell culture media.
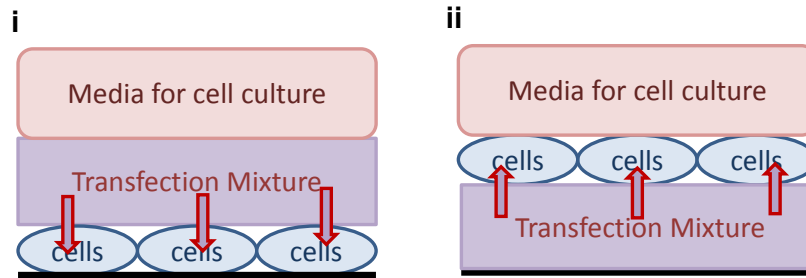
Figure 4: Schematic illustration of forward and reverse transfection. (i) Forward transfection and (ii) reverse transfection

Forward and Reverse Transfection in Plate Wells

To select the best method for our experiments, we usually first compare forward and reverse transfection in the wells of 96 or 384 well plates. Fig. **5** (**A** and **B**) depicts such an experiment using four different transfection reagents, DharmaFECT 1-4 (DF1-4; Dharmacon, part of GE Healthcare) for transfection of siRNA against the p65 protein into Hela cells. Several wells were assessed in parallel, using QPCR to measure the decrease in amount of mRNA and immunostaining to determine how the decrease of mRNA expression affects protein detection within the cells. Our results for p65 and other genes showed that the decrease of mRNA detectable by the level of the protein at 48-72 hours of incubation. Interestingly, both forward and reverse transfection give comparable results in terms of mRNA and protein level reduction, about 60%-80% (Fig. **5A** and **B**).

For each microarray-based screening project, we first test in wells the efficiency of reverse transfection of siRNA targeting cellular genes like p65, as well as "control" genes related to pathways or cellular mechanisms involved in the assay. This test demonstrates that the reverse transfection efficiency is acceptable in the selected cell line, and shows the percentage knockdown of protein expression that can be achieved for already known control genes. In well-based tests, we generally use three concentrations of siRNA ranging from 1-20 nM in 100μl of the final media dispensed in wells of 96 well plates. To achieve this, volumes of siRNA and transfection reagent are mixed and incubated for 20 min prior to dilution in culture media (see Table **2**). If these results are acceptable, yielding >50% knockdown of the

signal (*e.g.*, transcript levels) with no significant cytotoxic effect determined by the cell/nuclei count during image analysis, and observed in 2-3 independent experiments, then tests with microarrays can proceed.

A- Direct Transfection   B- Reverse Transfection
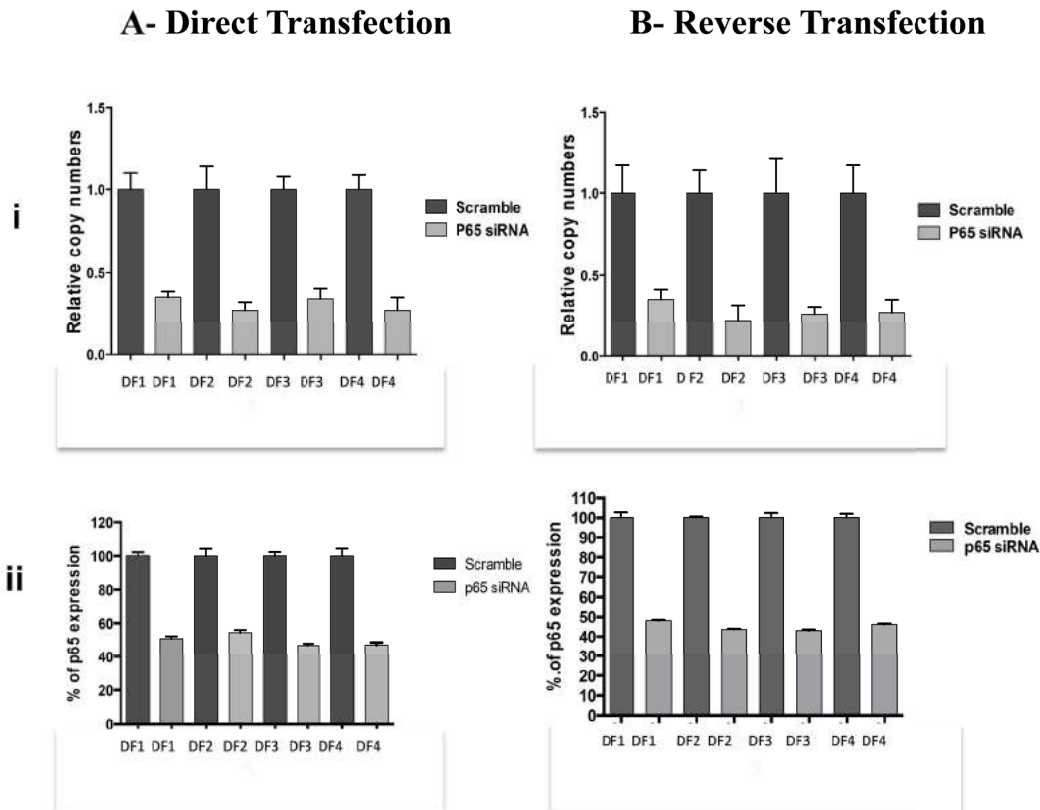
Figure 5: (A) Forward transfection and (B) Reverse transfection of HeLa cells with p65 siRNA. (i) QPCR for p65 mRNA in transfected cells, with parallel wells containing cells treated for immune-staining and imaging; (ii) quantification of the immunostaining signal analyzed using IM algorithm, normalized and plotted as % of cells expressing p65. DF 1- 4 are DharmaFECT 1-4.

MICROARRAY PRINTING

Before discussing the details of assay optimization for transfection on microarrays, we will briefly describe the array printing itself.

Approaches to Printing

An ideal printing system should be able to create uniform, dense arrays of nano-quantities of probes, while maintaining integrity and purity of reagents by

preventing contamination. Methods for printing or spotting nano- and pico-liter-quantities of different biological materials on solid substrate can be roughly divided into two major categories: contact and non-contact printing. Contact printing includes pin-printing, micro-stamping and nano-tip printing. Contact printing instruments have solid or split pins that come in direct contact with the surface of the slide. Even though this technique has not progressed much over the earlier printers devised by Brown and colleagues, it has proved reliable over many years for oligonucleotide microarray production for a variety of experiments including work with siRNA molecules for transfection into live cells, as shown in Table **1**. The printers with pin-tools are manufactured by Bio-Rad Labs; Genetix; Invatis, and other companies.

The non-contact printing approach includes: photochemistry-based printing, electro- or electrochemistry printing, laser writing and droplet-dispensing Inkjet technology. Inkjet technology has become popular during recent years with the development of piezo-based systems that allow the precise, quantifiable and reproducible dispensing of small volumes containing fragile biological samples such as protein molecules [60]. Such printers are produced by companies such as Abbis, Arrayjet; Biofluidix; Gesim; Perkin Elemer, Scienion, and Shimadzu.

We use the Genomic Solution Omnigrid 100 high-throughput pin-tool based printer designed to simultaneously hold 100 1'x3'microscopic slides. Our system employs stealth solid pins (SMP9, Telechem International, USA) to create spots having a 250μm diameter separated by a distance of 150 -200 μm (Fig. **6**). Using images of red spots we determined the size of the spots arrayed on PLL coated slides to be around 215±10 μm.

Many factors can influence the size and quality of the spot, including:

- Size of pin

- Viscosity of printing solution

- Depth of submerging of the pin in the well of 384 well plate

- Total volume of the mixture in the well

- Time that pin is in contact with slide surface

- How many spots are made using one dip of the pin in the reagent

Table 1: List of arrayers used for printing of cDNA or siRNA for transfection of live cells

	Publications	Arrayer	Pins	Spot size (D)	Distance between spots
1 - 2	Ziauddin, J. &Sabatini, D.M. 2001 [51] Lindquist RA *et al.* 2007 [61]	PixSys 5500 PixScientific, NH	Genomic Solution (Telechem International) : SMP7 stealth pins	220μm	450 μm
3	Yoshikawa T, *et al.* 2004 [62]	MacroSys 400 Cartesian Technology			
4,5,6	Erfle H *et al.* 2004 [63] Simpson J.C. *et al.* 2006 [64] Erfle H *et al.* 2007, 2008 [55,65]	Chip Writer Compact Robot and Pro (Bio-Rad, USA)	TeleChem Int: SMP10 Eurogentec, Belgium: quill pins	400 μm	900 μm 1125 μm
			Point technologies: Solid pins PTS 600		1500 μm
			PTS 400	400 μm 270 μm	2250 μm
		Prosys 5210DNA microarrayer (Genomic Solution, USA)	Telecehm Int: SMP10B stealth pins	350μm	550 μm
7	Mannherz, O, *et al.* 2006 [53]	Omnigrid (GeneMachines, USA)	Telechem Int: SMP15B	600 μm	1000 μm
8	Rantala, J.K. *et al.* 2011 [56]	Genetix Qarray2 (Genetix Ltd)	Point Technology: solid pins 200 μm tip	200 μm 400 μm	500 μm
9, 10, 11	Genovesio, A. *et al.* 2011 [23,30] Kwon, Y.J. *et al.* 2013 [66]	Omnigrid 100 (Genomic solution, USA)	Telechem Int: SMP9 stealth pins	250 μm	500 μm

Due to the fluctuations of those factors during printing of the siRNA transfection material, spots can be missed, although on average less than 1% of the spots are missing, can have irregular shapes, or contain less material than desired. The use of siGLO tracer helps to assess the presence of printed spot and its quality.

Another challenge associated with high-density spot arrays is the relatively small number of cells associated with each spot area [67]. The reduction of the size of the spots can become a critical issue due to the smaller number of cells attached to the area. One study reported 50 ±3 and 150±8 PC-3 cells in spot sizes of 200 μm and 400 μm, respectively, using a technique that allowed cells to adhere only to the spot area [56]. In work that utilized HEK293 cells, other study reported 30-80 cells were counted on a spot of 120-150 μm [51], while in a separate study, 100-400 cells were associated with a much larger spot of 600 μm [53].
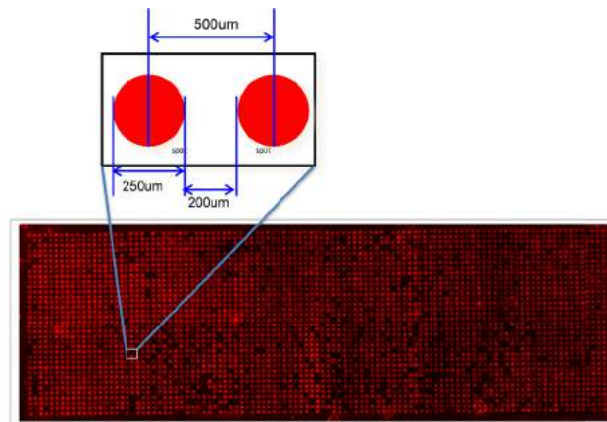
Figure 6: Dimensions of transfection mixture spots printed on a microarray. The image of a microarray was taken on ImageXpress Ultra at 10x magnification in the Red channel (to detect Red-siGLO), the dimensions are summarized in drawing. Fluorescently labeled siRNAs were printed on a PLL-coated slide. The size of spots and their variability was measured by MetaXpress software, average diameter of spots ∼ 215 μm ±10 (error- is standard deviation, calculated in Graphpad Prism, n=3.
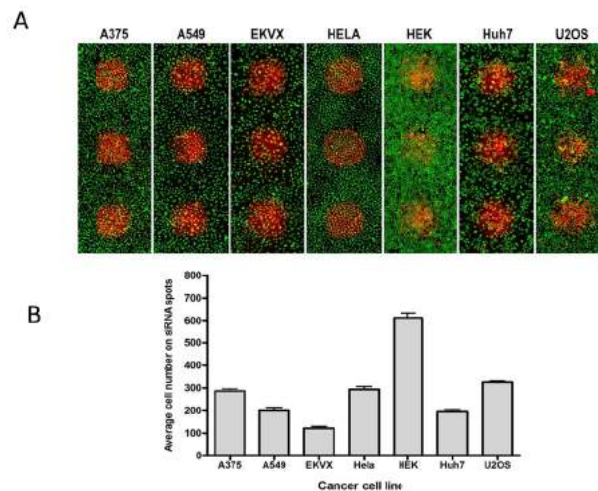
Figure 7: Number of cells adhered on the spot. (**A**) Composite image of a PLL-coated microarrays with different cell lines seeded at 10^6 cells/slide and cultured for 48h before fixation and immunostaining of p65 and nuclei staining. Green: p65 antibody-Alexa 488 labeling, Red: siGLO-siRNA tracer spots, Blue: DRAQ5 staining of nuclei. Scale bar is 100μm (**B**) Quantification of cell number per spot based on a nuclei count by MetaMorph algorithm. Average number for 3 spots is plotted and error bar represents standard deviation.

Cell number will also vary depending on confluence of cell cultures. We counted the number of cells per spot for several cell lines typically used in the laboratory, and found that average number was about 300 cells/spot (after plating about 10^6 cells/slide) for A375, HeLa and U2OS (Fig. **7**). EKVX and Huh7 showed fewer cells on the spot, with HEK293 being the most aggressively growing and reaching as many as 600 cells per spot.

The most important issue associated with microarrays is the efficiency of transfection on the spot, which varies with the type of cell line. Some cell lines such as PC-9, TC-7, some primary cells or even immortalized cell lines known to have reasonable efficiencies in well-based assays were transfected at unexpectedly low efficiency on microarrays. In some cases transfection could be optimized by switching the transfection reagent but we also found that different coatings increase transfection efficiency for some cell lines. For instance A375 shows better transfection on MAS-coated slides with a slightly increased concentration of sucrose (100 mM) and 0.09% gelatin compared with the PLL coating with 25 mM sucrose and 0.06% gelatin.

Optimization of Reverse Transfection on Microarrays

Table 2: Comparison of transfection mixture components for reverse transfection in wells or on arrays

Mixture for Reverse Transfection					
in wells			**on array**		
1	lipofectamine	0.1μl	**1**	lipofectamine	3μl
	DMEM	9.9μl			
2	siRNA (1uM)	5μl (1μM)	**2**	siRNA (20μM)	2μl (20μM)
	DMEM	5μl	**3**	0.3M Sucrose in OptiMEM	2μl (0.3M)
				RNAse free water	2μl
mixed and incubated for 20 min			mixed and incubate for 20 min		
3	OptiMEM	80μl	**4**	siGLO (20μM)	2μl
			5	0.2% gelatin	5μl
Final Step:					
add 100μl to cells in 1 well/96 well plate or 20μl in 1 well 384 well plate			print 3-6 nL/ 300 um spot		

A large variety of lipid-based transfection reagents can be used in conjunction with test/control siRNAs (*e.g.*, targeting p65) to create test slides for determining the best reagents for any cell line of interest. We typically print slides comprised of control siRNA formulated with Effectene (Qiagen), MetafectenePro (Biontex), Lipofectamine 2000 (Invitrogen), Lipofectamine RNAiMax (Invitrogen) and DharmaFECTs (Dharmacon, part of GE Healthcare) on a single test microarray. In general, these reagents show little toxicity under conditions that provide highly efficient transfection. In our analysis, which regularly employs p65-targeting siRNA and uses the average p65 signal intensity to assess transfection efficiency, this panel array (with few exceptions) confirms the optimal reagents that provide acceptable levels of transfection ($\geq$50%).

As the most expensive components of microarray-based screening, the transfection reagent and siRNA need to be titrated to select the minimal concentrations required for efficient knockdown. Fig. **8** shows a typical titration optimization results for HeLa cells using Lipofectamine (**8A**) and siRNA anti-p65 (Fig. **8B**). The 2.5-3.0 μM of siRNA and 2.5-3μl of Lipofectamine (Table **2**) were selected as optimal for this type of experimental conditions.
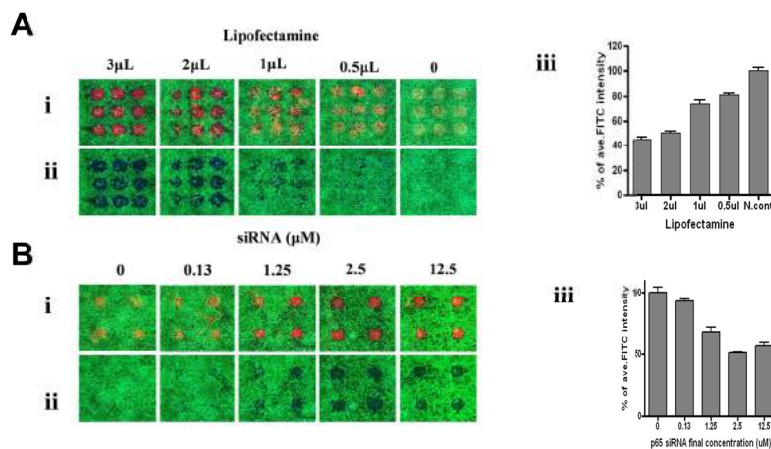
Figure 8: Titration of (A) Lipofectamine and (B) anti-p65 siRNA for reverse transfection on microarrays. Composite images of cells on microarrays. Green represents p65 antibody labeling. Red identifies siGLO-siRNA tracer spots. Blue represents DRAQ5 staining of nuclei. (A) Lipofectamine was tested with different volumes of 0.5-3 μl using protocol described in Table **1**. (B) anti-p65 siRNA was titered at concentrations of 0, 0.13, 1.25, 2.5 and 12.5 μM. For both A and B: (i) Overlay of images from 3 channels to show the spots, 9 (A) and 4 (B) replicates (spots) for each condition; (ii) the overlay of images from two channels to show presence of the blue nuclei (cells) in area of spots to assess the absence (or presence) of possible toxicity; (iii) quantification of the average intensity signal for the cells on the spots (n=9 (A) and n=4 (B) replicates)

It is worth noting the differences in siRNA concentrations required for knockdown in the two (well-based and microarray-based) formats (Table **2**). For reverse transfection of anti-p65 siRNA, we found a 10-20nM concentration of siRNA in wells gave a similar knockdown effect as 2.5-3 μM of siRNA in the mixture printed on the slide. Clearly, the miniscule amount of mixture used for printing, the total amount of siRNA stock is significantly smaller than that used in well-based experiments. That said, the amounts and kinetics of siRNA delivery in microarray platforms is difficult to quantify. In addition to smaller amounts of material, another advantage of the microarray-based approach is that unused portions of the mixture can be stored at -20°C and reused weeks later without loss of efficiency in transfection experiments (data not shown).

The timeline for siRNA-mediated gene knockdown on was found to be very similar to well-based reverse-transfection. On average, an incubation period of 48h-72h induces the best knockdown results (Fig. **9**).
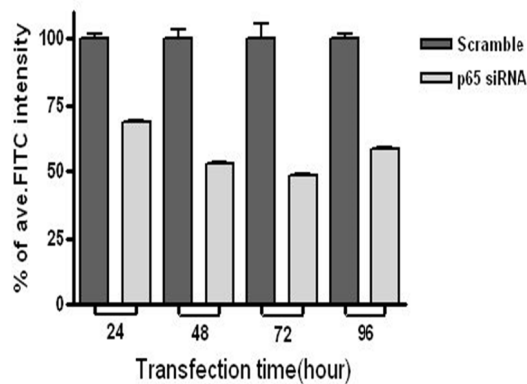
Figure 9: Time window for incubations on microarray. HeLa cells were plated on siRNA microarrays (anti-p65 siRNA and scrambled siRNA printed) and cultured for 24- 96 hours. The average signal for p65 immuno-staining in the cells on the spots is analyzed and presented as % of Intensity compared to scrambled siRNA negative control. The error-bar is SD, N=4 spots.

Protocols for microarray printing have mentioned two other essential components of the transfection mixture: gelatin and sucrose [51, 55, 67, 56, 69, and 70]. Fjeldbo *et al.* 2008 [69] reported that gelatin helped the transfection mixture to transfer from the surface to cells, and sucrose stabilizes the transfection complexes and acts as a critical preservative for storage of the printed slides. This work also reported testing of different concentrations of sucrose and gelatin for transfection of GFP-expressing cDNA using the X-teremeGene (Roche) transfection reagent on Ultra-GAPS coated slides (Corning). As we were

developing a system to use lipofectamine or RNAiMax, in conjunction with PLL- or MAS-coated slides, we tested the effect of different concentrations of sucrose and gelatin under our conditions (Fig. **10**)
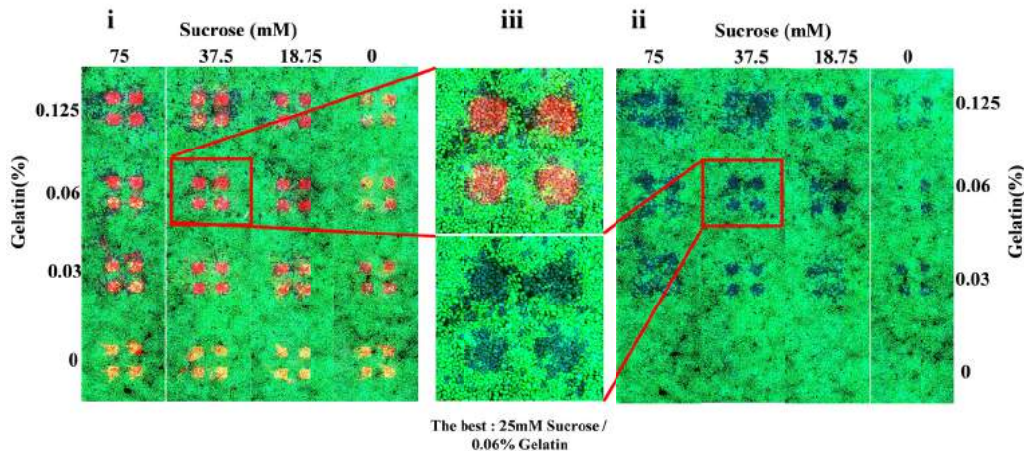
Figure 10: Combined titration of gelatin and sucrose. HeLa cells (1×10^6 cells per slide) were plated on microarray with p65 siRNA and cultured for 48 hours. Cells were fixed and stained with anti-p65 antibodies and DRAQ5. Green: p65 antibody-Alexa 488 labeling, Red: siGLO-siRNA tracer spots, Blue: DRAQ5 staining of nuclei. (i) composite image from 3 channels, 4 replicates (spots) for each condition; (ii) composite image from two channels to show presence of the blue nuclei (cells) in area of spots to assess for possible toxicity; (iii) zoom-in view of the section with optimal conditions of 37.5 mM sucrose in combination with 0.06% gelatin

Fig. **10** shows the combined titration of sucrose and gelatin for transfection of anti-Rel-A (p65) siRNA into HeLa cells on PLL-coated slides. As was expected, both gelatin and sucrose proved essential for successful transfection of siRNA. However, the highest concentrations of sucrose (75 mM) and gelatin 0.125% caused a significant change in the viscosity of the mixture resulting in diffusion of reagents, washing-off the clear borders of the spots, and reduction in adherence of the reagent mixture to the surface during the process of air-drying of the printed slides. The combination of 37.5 mM sucrose and 0.06% gelatin was optimal for our experimental conditions. These results are quite similar to previous observations and indicate the importance of correctly calibrating the range of concentrations for gelatin as well as sucrose [68,71].

Once the transfection reagent mixture is optimized and ready for printing, the management of the printing process for the large genomic siRNA library (18,000 to 20,000 samples) is the next critical step.

Printing the Human Genome siRNA Library on Microarrays

To print the library for easy to transfect cell lines, the siRNA transfection solution is prepared as described in Table **2**. The library of Dharmacon "ON-TARGETplus*"* siRNA "SMARTpool" format consists of 4 different sequences pooled together for each gene. That helps to reduce amount of off-target hits and increases the chance of achieving efficient knockdown effect for the average gene in the genome. 2 μL of a 20 μM solution of each siRNA, from library is transferred into each well of a 384 well plate. 6 μL of 20 μM RED siGLO is then combined with 2 μL of 0.3 M sucrose dissolved in OptiMEM media, and 2 μL of RNAse free water. 3 μl of RNAiMax is added to each well, mixed thoroughly, and then incubated for 20 min at RT before addition of 5 μL of 0.2% (w/v) gelatin

The optimized spot dimensions (Fig. **6**) allow us to position 3,888 spots (108 columns by 36 rows) on a single glass slide. Since cross-contamination could occur if the transfection reagent or cells containing the transfection reagent cross from one spot to another, multiple tests were done to ensure that does not occur. One of those tests is presented in the Fig. **11**, where the spots with the scrambled (non-targeted) siRNA were printed around the spot with anti-p65 siRNA. When the image of fixed and stained HeLa cells was assessed for p65 expression, signal reduction was only observed in spots containing anti-p65 siRNA (observed silencing ~ 20-40% of control signal). Other regions of the slide (*e.g.*, other spots and areas between spots) exhibited normal p65 expression suggesting spot-to-spot cross contamination was not occurring (Fig. **11B**).

With such a high density of spots, the 18,000 pooled siRNAs library can be distributed on as few as five slides. We also added several controls on each slide. These include 52 equally distributed negative control spots containing scrambled siRNA on each slide, along with an equivalent number of positive control spots containing pools of siRNA targeting p65 (Fig. **15**). We also printed anti-GFP siRNA as a positive control in cases where a recombinant cell line containing a GFP expression marker is used. Printing was done at 22-25°C, 55-65% humidity in an enclosed, custom-built HEPA-filtered clean chamber.

One of the advantages of an array-based format of an siRNA library is that it eliminates the problems associated with multiple freezing and thawing of plates with aqueous siRNA solutions. After preparation of the siRNA plates in the desired format (we use 384 well plates), plates are used for direct printing of as many as 200-400 copies of the siRNA genome library. The remaining reagents on a plate can be sealed and stored at -20°C (or -80°C) for several days and re-used

for QC arrays or additional test runs. The arrays can be stored under desiccated conditions for up to 1 year.
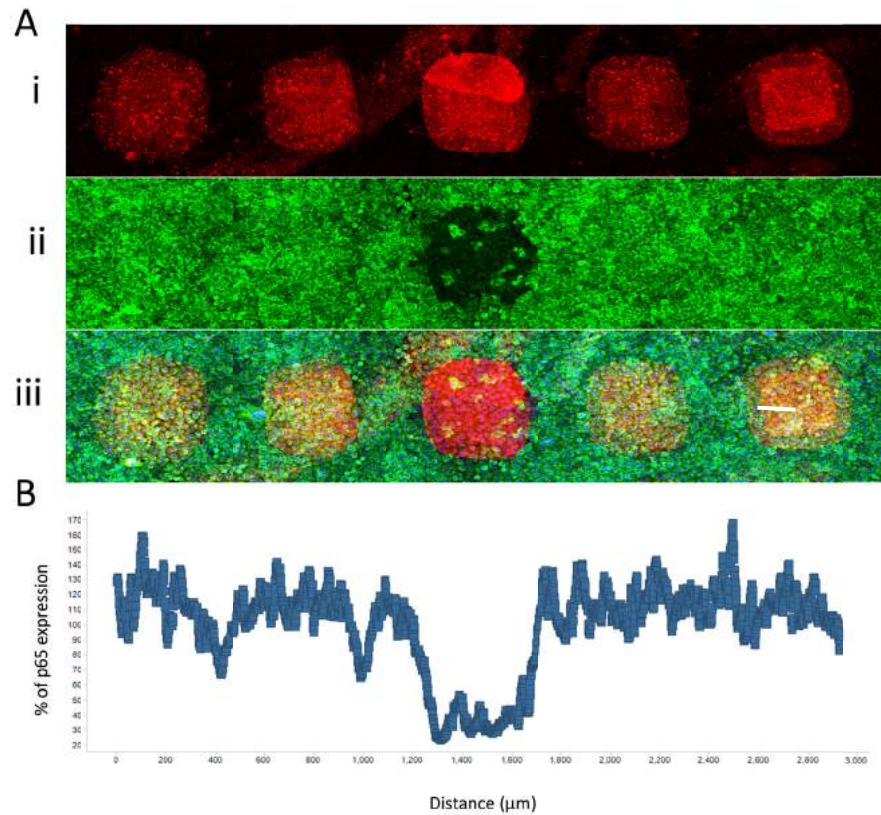
Figure 11: Test for cross contamination between spots. (**A**) To test for cross-spot contamination, anti-p65 siRNAs (center) were flanked by spots containing scrambled siRNA. After a 48h exposure, HeLa cells were fixed and stained with anti-p65 antibody and DRAQ5. Color Indications: Green: anti-p65antibody labeling, Red: siGLO-siRNA tracer spots, Blue: DRAQ5 staining of nuclei. (i) Spot visualization in red channel only; (ii) green channel for p65 signal; (iii) composite image from 3 channels. Scale bar = 100μm (**B**) The presence of cross contamination was assessed by scanning the slide, measuring intensity of p65 signal, and presenting it as the percentage of p65 expression across the images. There is a significant decrease in p65 signal in spots that contain p65 siRNA down to 20-40% of total signal while spots containing scrambled (non-targeting) siRNAs exhibited normal levels of p65 expression.

Microarray Image Analysis and Library Annotation

There are various considerations for analyzing visually addressable microarrays designed for cultured cellular monolayers. The main difference between the high-

density spot microarrays and the same experiments done in wells is that the spots first have to be localized precisely enough so that image analysis can be applied to cells positioned exactly above the spot. Therefore, the first and the most important tasks in image analysis are spot fitting and spot localization.

Typically, the process of spot localization is based on prior knowledge of the spot "map", which is usually the number of spots in x and y axis on the slide. More accurate information, such as the size and actual distance between spots (scale information), removes further degrees of freedom during spot detection, and algorithms have been proposed for resolving this issue [72-75].

A predecessor of the microarray technology described here, cDNA hybridization microarrays, used visually-addressable, precise spot localization, in conjunction with many approaches for spot detection and grid-fitting [72, 76]. Our computational group modified those approaches to address critical problems associated with spot detection and sub-spot fitting for our microarrays [60]. Like all the algorithms used in the analysis, the spot fitting should be performed with efficient speed. It also should be precise because the arrayed spots do not always have the expected shape and intensity of the signal, or may be missing entirely. Precision is critical for statistical analysis of data obtained from the cells growing on the spots. For our group, spot detection is achieved by visualizing a fluorescent labeled tracer which has been added into each sample. Specifically, we use a double-stranded siRNA-tracer, covalently labeled with the DY547 fluorophore (siGLO-Red with excitation maximum of 557 nm and emission maximum of 570 nm, Dharmacon, part of GE Healthcare). Images acquired at 560 nm (red channel) are then used for spot identification and correction of grid position. Fig. **12** shows an example of how addition of the fluorescent tracer allows both the localization of spots and the identification of missing spots. A total of 408 individual images, 34 columns x 12 rows, of cells on microarrays usually is taken by MDS ImageXpress Ultra at 10x magnification by scanning the entire surface of a single slide. Those images are stored in a database and stitched into a single montage by our Image Mining (IM) program for spot fitting and image analysis. The resulting array containing the genomic library comprises 108 columns by 36 rows of spots. This type of accurate spot fitting allows not only image analysis of relevant cells but also correct annotation of spots with information regarding targeted genes. The annotation file generated by the Genomic Solution Omnigrid printer then transfers the information from the plate-based format into the array format after the spots have been found (Fig. **12B**).
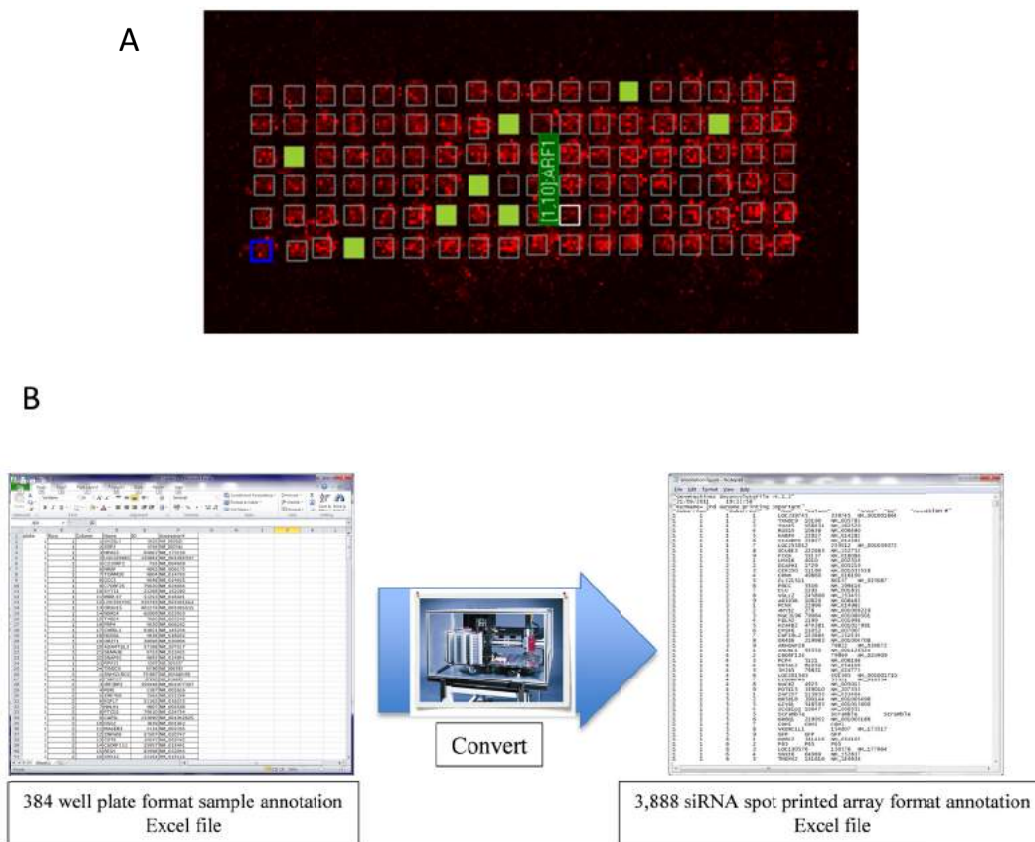
Figure 12: Spot grid fitting and annotation. (**A**) The images of spots on a microarray acquired on red channel only. The grid is fitted to a portion of the array and the missing spots are identified as solid green squares. The grid is associated with an annotation file that assigns information about a gene to the corresponding spot. (In this case, a spot associated with the gene ARF-1 is identified). (**B**) The Genomic Solution Omnigrid printer converts the plate-based format of the genomic library into an array-based format.

Another channel is used to detect the cell nuclei and select only those localized in the "spot" area. The area of image defined by the so called "nuclei mask" defines the cell spaces associated with those nuclei and is used for further expansion and analysis of the signals in other channels (Fig. **13**). Because of the limited number of channels we can use for microarrays imaging, we use DRAQ5 (excitation 598 nm and emission 646 nm) for nuclei staining. The third channel, as mentioned before, is typically used to image signals from immunostaining or detection of fluorescent proteins expressed in cells. This signal is collected in the area designed around the defined nuclei or masked area. The specific design of those algorithms depends on the type of cellular response interrogated in this particular

assay. The intensity of the signal collected within the masked area in the green channel generates data with which to calibrate the siRNA effect on cells in a particular spot. The other two channels (blue and green) collect data for multi-parametric analysis of complex phenotypic events.
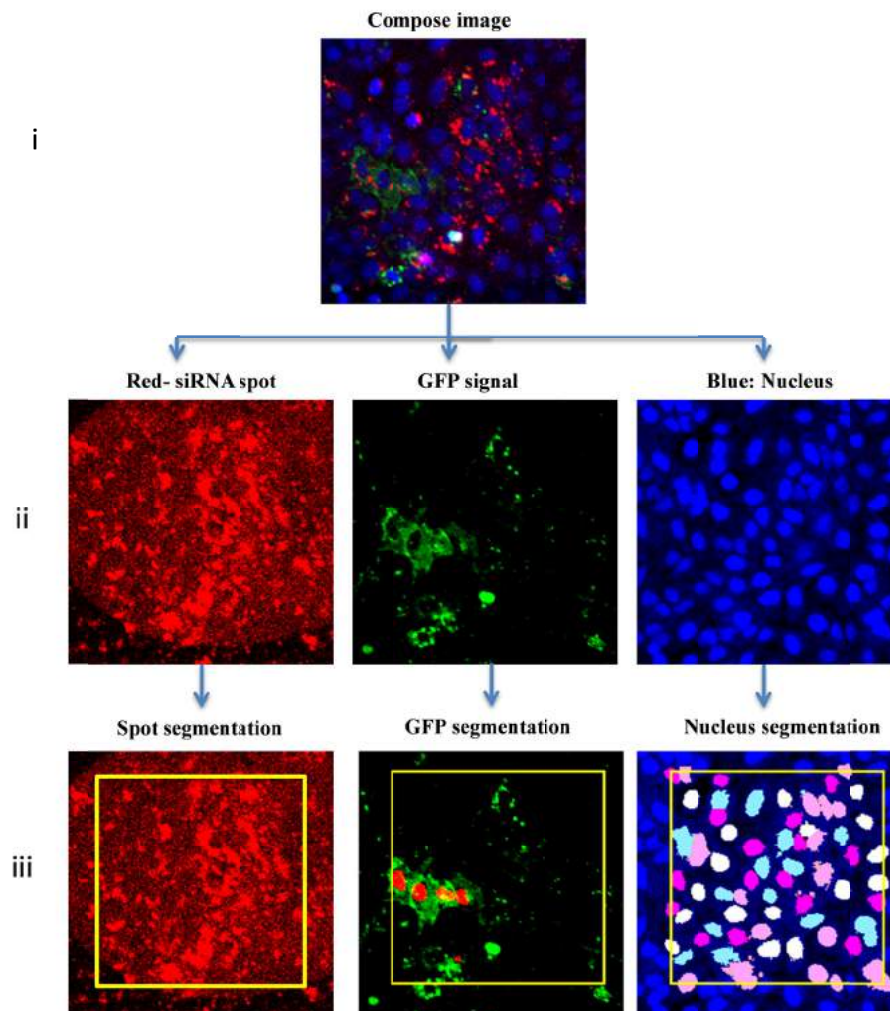
Figure 13: Image analysis of cells on a microarray. (i) Composite image of 3 channels (ii) Individual channels: red- siRNA tracer signal, blue- DRAQ5 nuclear signal, green- GFP cell/pathway signal; (iii) Segmentation, or selection and localization of spot area(red), nuclei (blue) and GFP signal associated with nuclei (green). Multicolored nuclei represent masks for cells located in the defined spot area and thus selected for analysis. The population of cells transfected by siRNA but not expressing the marker protein can be clearly seen by the absence of green color but presence of blue nuclei in locations within the spot area.

Typically, each new assay requires specific design of "plug-in" software for analyzing images. In laboratories with no access to support groups designing custom algorithms, commercially available software such as Metamorph (MDS), Columbus (PE), and open-sourced software such as CellProfiles [77], DetectTiff [78], and ImageJ [79] can be used for automated identification and quantification of cellular phenotypes, and image analysis.

Validation of Microarray–Based Assays

Assay validation is a critical step in any type of high-throughput screening. In the case of microarray-based format screening it is essential due to the variability associated with transient effects of siRNA, the small number of cells analyzed in each spot, and the significant number of samples (3,888 in our case) positioned on each array.

We have developed a multi-step validation process to confirm the feasibility of assays for screening. One of them is the test for uniformity of the cell layer on the microarray and uniformity of immunostaining signal through the slide. We addressed the uniformity of cell distribution on the microarrays earlier in Fig. **3**. Now we describe the image-based approach to analyze the quality of the cell layer. This can be done by analyzing images taken by a plate reader for the cross-section of the slide. Fig. **14** illustrates the general design of the test. Data from the reader is collected for several consecutive independent experiments (or slides), and analyzed for statistical deviations. This analysis can help to optimize the cell seeding density that in 48 hour provides a homogeneous monolayer of cells. The analysis of images also confirms optimal conditions for immunostaining, including volumes for antibody solutions that provide homogeneous staining without edge effects and other artifacts. Also this test is used to confirm uniformity of viral or parasitic infection throughout the whole slide when the assay involves infectious agents. As an example, Fig. **14** shows the titration of GFP-expressing Chikungunya virus detected in cells after 48h infection. Such optimizations were used in functional genome screening to identify host factors involved in HIV infection of human HeLa cells [23], *Trypanosoma cruzi* parasitic infection of U2OS cells [30] and H1N1 influenza virus.

It is helpful for assay validation to be able to test positive and negative controls positioned throughout the whole slide. The difference between positive and negative controls should be statistically significant and reproducible in 3 independent experiments.
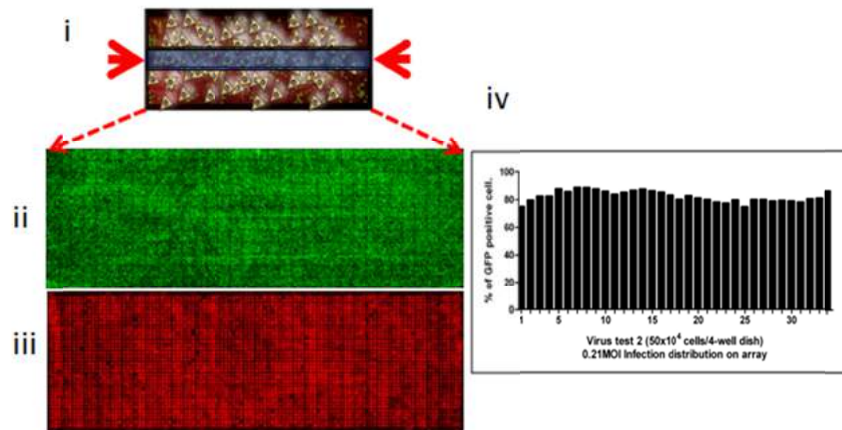
Figure 14: Assessment of uniformity for cell monolayer on arrays. The endogenous p65 used as a marker for the immuno-staining procedure (green color). The assessment of quality of cellular monolayer was done by quantification of the images of the cells on the selected fields of the slide. (i) the slide scanned by plate reader and raw of images of immuno-staining signal (ii) and red spot tracer (iii) could be used for quantification of the signal throughout the slide. (i) only one set of field were selected through the length of the slide to generate the data, represented by bar graph (iv).
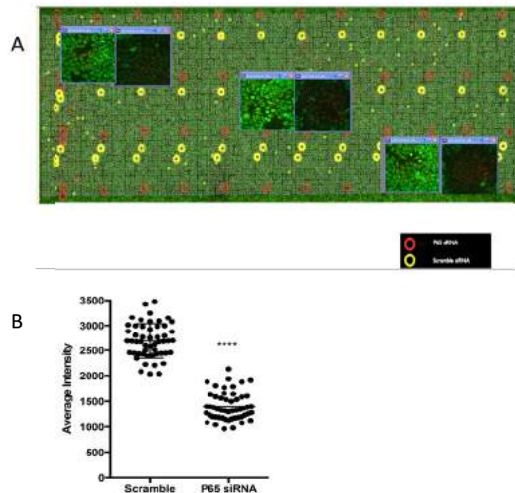
Figure 15: Positive and negative controls on slides printed with portions of the library. (**A**) Scrambled negative controls (yellow circles) and p65 siRNA (red circles) are printed throughout the slides to assess variability on different portions of the microarray. Shown is a composite image of the 3 channels for the mosaic composition of all the images for the whole scanned slide. The zoomed-in views of several individual images illustrate the control spots in different locations on the slide. (**B**) Scatter plot of average intensity for each control spot. The statistically significant difference with P<0.001 in Student t-test for the average values are indicated by asterisks (n = 52 spots for each p65 and for scrambled control spots)

We arrayed at least 52 scrambled siRNA (negative controls) throughout the length of each slide containing the genomic library (Fig. **15**). This figure also illustrates the statistical significance of the difference between two control populations of data points on the array (Fig. **15B**). As mentioned previously we use p65 as a positive control and this standard tests allows us to validate steps for complex assays without wasting precious reagents or viruses.

Unfortunately, it is impossible to prepare specific positive controls on genomic arrays for all possible types of assays. For projects that use GFP as a reporter fluorescent protein, we added 52 spots of anti-GFP siRNA. For other projects, we create separate new control arrays with specific positive controls among large populations of scrambled siRNAs. These arrays are added to the five genome arrays for quality controls during screening. If there are no known genes that could work as positive controls in a given project, we focus our attention on the quality of signal and noise for the scrambled non-targeting siRNA.

Data Analysis

Having a large population of scrambled or non-targeted siRNA samples provides a powerful tool to work with the data from the screening campaign. The data derived from the scrambled siRNA population is used in computational methods for normalization of the data across different conditions, such as between slides and different batches of slides or even different experiments. The genome library set has 5 slides covering 18,000 individual genes and total of 300 scramble siRNA samples. Usually, to ensure statistical significance of the observed events, we conduct screening with six or more copies of the genome library, which results in a total of more than 1800 negative controls. This is a large enough population to differentiate specific change of the signal from noise and reveal any artifacts. The assumption is that the negative controls give results similar to any genes that are not involved in the modification of the observed cellular phenotype. For example, the knock down of p65 did not play a role in early steps of viral infection and replication that was measured by immunostaining for NA protein after cells were infected with influenza (H1N1). Fig. **16** shows the distribution of results of the influenza assay done against scrambled spots compared to anti-p65 siRNA, 300 copies per one genome. The observations labeled in black show the difference (distance) between the multivariate phenotypic readouts of p65 and a control (scrambled) reference population in blue. We see that there is no significant difference between the two sets, which means that the gene knockdown shows no significant difference from scrambled results.
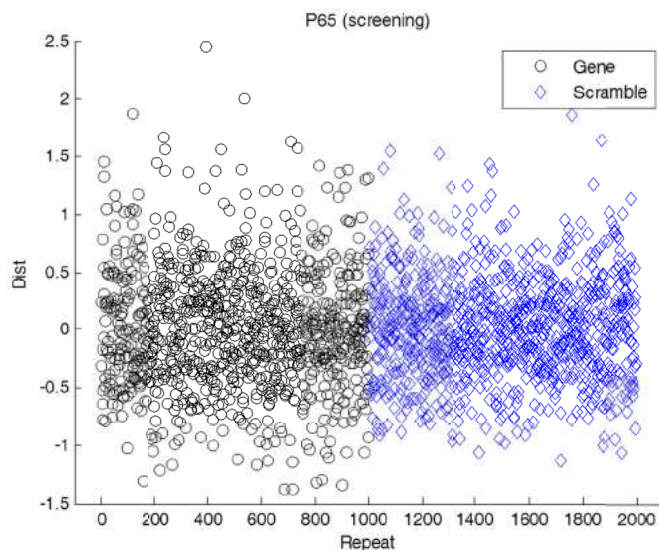
Figure 16: Scatter plot of negative control data. Metadata collected from cell image analysis for scrambled siRNA spots (Blue) and p-65 siRNA (Black) gene.

For typical functional genome screening, we calculate the distance of each gene from the scrambled population using Matlab applications. Fig. **17A** shows the total population of 17,753 (in this case) P values for each gene from a human genome screen using an influenza assay. Scrambled samples represents a compact population with P=0.999. The simple cutoff of P<0.01 was used to select 209 hits that significantly differed from noise in the sample population (Fig. **17B**).

To ensure reproducibility of the obtained results, we run at least five but usually more than six replicates of the library, as was mentioned previously [5, 11]. To make hit selection more robust we also used an additional score to reflect how many copies of the same gene give data with P value below the cutoff. If we have total of six replicates then at least four and more copies of the gene should be present in analysis and satisfy the cut-off criteria to be selected as a hits. In addition, the advantage of spot-microarrays is that we know if the spot was present on the array (see Fig. **12A**) and can determine how many copies were missing. The total number of hits is usually variable and depends on the quality of screening. In some cases, such as the HIV screen by Genovesio *et al.* 2011 [23], screening requires more complex data analysis.

Methods of computational data analysis are now commonly used for normalization of the screening data and selecting hits. A variety of software

programs can be used for statistical analysis of microarray data, such as the open source Bioconductor [80], CellHTS (which is designed for analysis of cell based siRNA screens) and more general applications such as SpotFire (IDBS). For more multi-variant and multi-parametric analysis, the R-open source language is a good solution for statistical computing and graphics (http://www.r-project.org/). There are several good reviews of statistical approaches developed for the analysis of RNAi screening data [81-83]. Having the support of a statistician experienced in screening routines could be essential in case of complicated assays.
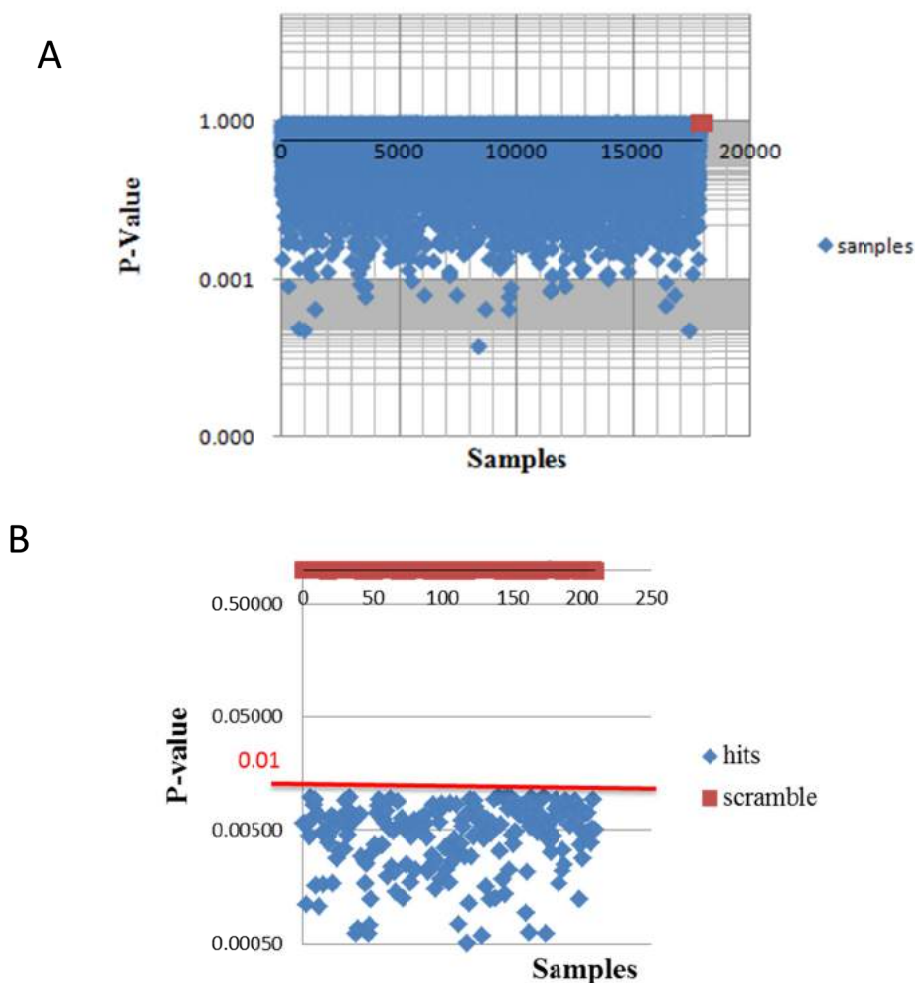
Figure 17: Scatter plot of screening data (**A**) P-values for all samples in screen (blue). (**B**) Selected hits (blue), P<0.01 for the scrambled controls (re**d**).

APPLICATION OF siRNA MICROARRAYS IN SCREENING OF INFECTIOUS DISEASE MODELS

We have used siRNA microarray technology to identify genes involved in specific disease progression and host gene-pathogen interactions for viral infections such as HIV [23], Influenza, Chikungunya, and Dengue fever; as well as parasitic diseases like Chagas [30]. Here we will describe details of the screening process with the siRNA microarray technology to identify genes involved in HIV infection and replication [23]. Study of these critical host factors is important for improving the fundamental understanding of HIV-host interactions, and for developing novel anti-HIV therapeutics. The work-flow process for microarray-based whole genome siRNA library screening includes the following steps: establishing and validating the cellular model; assay development in the micro-titer plates format; assay adaptation to the microarray format; screening; hit selection and confirmation; and deconvolution of selected hits.

Assay Development and Genome Wide Screening for HIV Infectious Model

First we established the most robust cellular model for the phenotypic assay. Among the available human cell lines and types of assays capable of detecting HIV, we chose an assay which provides the highest level of infection in cells that can be efficiently transfected on the microarrays. HeLa CD4+ LTR-GFP cells, developed in Peter Somers laboratory at IPK, have these characteristics and recapitulate early steps in HIV infection, which enables TAT-driven trans-activation of stably integrated GFP.

Based on previous studies of HIV host factors [19, 20, 22] we chose the CD4 receptor, which is essential for viral entry into the cell, as a positive control for viral infection. We optimized conditions for HIV infection and siRNA transfection in the plate format first to achieve the best assay window which ranged from >90% of HIV infectivity to < 20% of HIV expression when cells were transfected with CD4 siRNA. (Fig. **18**)

Using these optimized conditions, we validated the assay on a small-scale siRNA microarray containing printed spots of scrambled and CD4 siRNAs. After 24 h of reverse transfection, cells on the array were infected with HIV-1 at a multiplicity of infection (MOI) of 0.14 for an additional 48hr. As observed in the well-based assay, the HIV infection was significantly repressed in cells transfected with a CD4 siRNA (Fig. **19**). The uniform distribution of infected cells throughout the entire slide confirmed that the HIV assay adapts well to the siRNA microarray system.
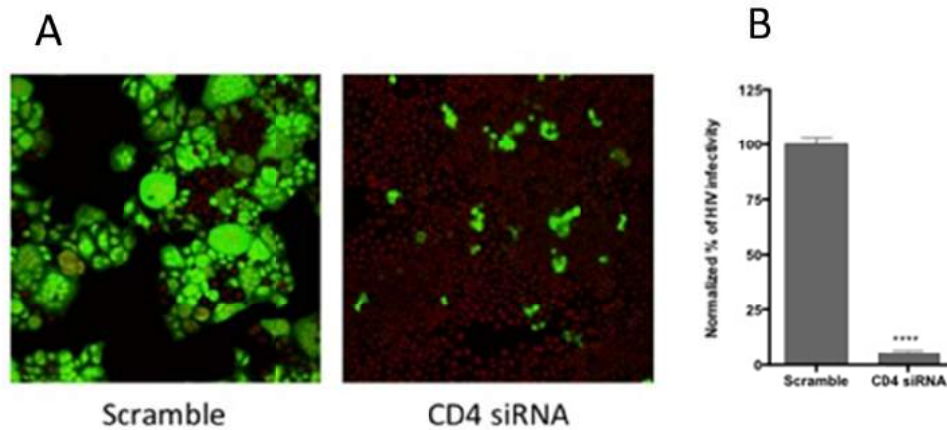
Figure 18: The cellular HIV assay in plate format. (**A**) Hela CD4+ LTR-GFP cells were seeded in 96 well plates and incubated overnight. The next day, cells were transfected with siRNA (scrambled control and anti-CD4) for 24 hours, then infected with HIV-IIIB. After an additional 48 hour incubation, cells were fixed and stained for nuclei detection with DRAQ5. The GFP expression associated with virus decreased in the presence of anti-CD4 siRNA. Scale bar is 50 μm. (**B**) GFP signal in all cells was analyzed using an Image analyzing algorithm and data was normalized as % of infected cells, **** significance of difference was assessed using student t-test (n=6, p<0.001)
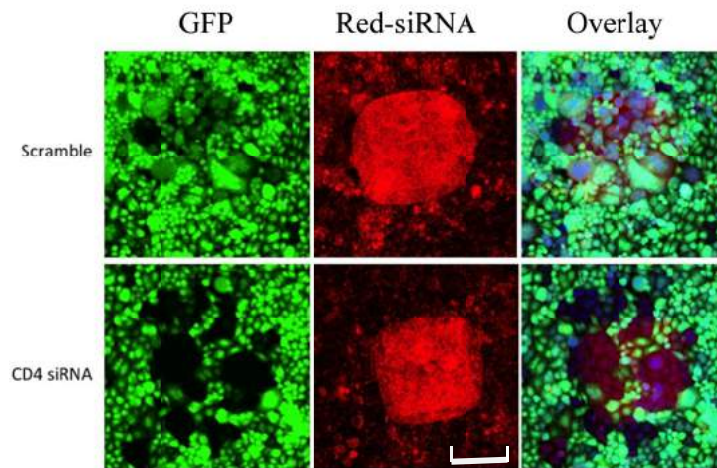
Figure 19: HIV infection assay was adapted to siRNA microarrays. HeLa LTR GFP cells were seeded on a test microarray containing scrambled and anti-CD4 siRNA and incubated for 24 hours. Cells were infected with HIV-1 (MOI 0.14) and after 48 hours treated same way as in well-based assay. (Green: HIV infection, Red: siRNA for spot detection, Blue: nuclei). Scale bar = 100μm

Five replicates of the siGENOME whole genome library, 20,000 siRNAs SMARTpool (Dharmacon, part of GE Healthcare) printed on seven slides, were screened in one experiment with a total of 35 slides. If this level of screening was performed in 384 well plates, it would require that 360 plates be screened in one experiment. Once siRNA microarray images were acquired and each spot localized and identified, HIV infection was independently analyzed in cells growing on each siRNA spot. Fifteen imaging parameters were used in data analysis, including measurements such as cell number, area of relative cell distribution, size of individual cells, syncytium formation, and intensity of GFP signal. The multi-parametric data analysis was performed such that it allowed selection of hits with parameters similar to those collected using the CD4 siRNA (control). This work demonstrated the advantage of using different approaches for hit selection, based on the hits position relative to controls. In this case, we selected 56 genes with Wilks' Lambda = 0.66 (a value under 0.9 indicates that this distribution is different form distribution of scrambled controls [23])

Confirmation and Deconvolution of Screening Results

The 56 hits selected from the primary screening were confirmed using plate-based assays that in addition to detecting GFP expression activated by HIV, also directly detected the presence of p24 HIV proteins [84]. To verify the significance of the screening results, it was essential that we demonstrate that depletion of the candidate genes in LTR-GFP Hela cell blocks HIV replication in a similar way to the replication block induced by CD4 knockdown. To test this, we selected RNASEH2A, MED28 and JMY from the hit list and used CD4 as a control. Cells were transfected with siRNA for 24 hours and infected with HIV for 48 hours. Viral replication was measured by the appearance of GFP expression as well as p24 protein expression (Fig. **20 A**, **B**, **C**). Depletion of RNASEH2A, MED28 and JMY in Hela cells totally blocked HIV infection as detected by quantification of GFP expression and immunostaining of p24 protein expression (Fig. **20A**). Both biomarkers gave comparable results in cells transfected with RNASEH2A siRNA, which was an ~ 80% reduction of HIV infection. Cells transfected with siRNA against MED28 and JMY had somewhat higher levels of HIV infection at ~50% (Fig. **20B**, **C**).

To demonstrate that the observed effect is mediated by the reduction of the targeted genes and not due to off-target effects, the level of mRNA of the selected genes was measured using RT-PCR (Fig. **21**). The validity of the effect produced by RNASEH2A was confirmed in Jurkat (T lymphocyte) cells, a more physiologically relevant model of HIV infection. This model was also used for de-

convolution of pooled siRNA. Together, these results confirmed that hits selected in microarray-based screening campaigns can identify host genes, such as RNASEH2A, MED28 and JMY that are essential for HIV infection in Hela cells.
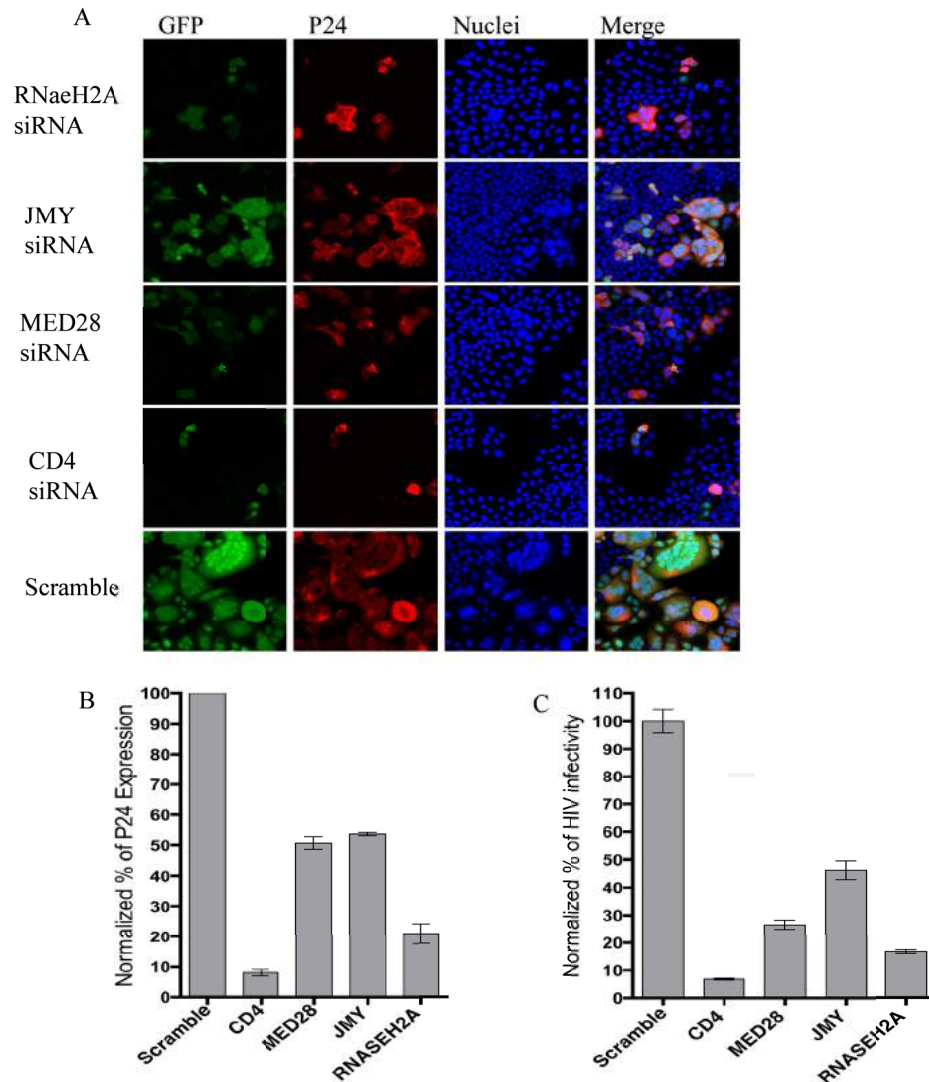
Figure 20: RNASEH2A, MED28 and JMY are required for HIV infection. (**A**) 3 channel images of cells transfected with siRNA against RNASEH2A, JMY, MED28 and CD4 (positive control). Green: LTR-GFP expression, Red: p24 expression, Blue: nuclei stained with DRAQ5. Composite image shows the overlay of 3 images. Scale bar is 40μm. (**B**) Quantification of P24 expression; (**C**) -Quantification of LTR-GFP expression. These values are expressed as intensity/pixel/cell and plotted as a bar-graph, n=3, standard deviation shown as error bar.
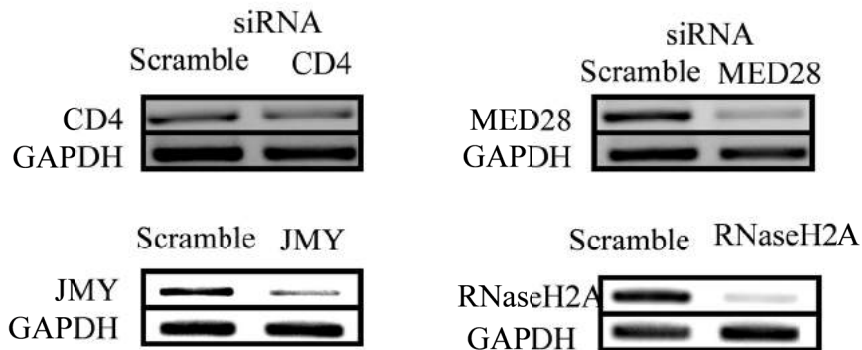
Figure 21: RT-PCR of RNASEH2A, MED28 and JMY in control samples and depleted cells. RT-PCR results show the level of RNASEH2A, MED28 and JMY mRNA reduction by individual siRNA. Cells were transfected with each indicated siRNA for 72hours and cDNA was prepared, and tested by RT-PCR.

SUMMARY

In this chapter, we discussed the technical details associated with the screening of phenotypic effects produced by the transfection of human siRNA whole genome libraries into cells plated as a monolayer on microarrays. This technology is complex, but enables a variety of applications that are otherwise unfeasible or prohibitively labor-intensive if using commercially available plate-based approaches. The technology becomes even more useful for phenotypic target-free applications in drug discovery that require drug target deconvolution, as well as to search for new targets and pathways involved in drug responses. By virtue of the high throughput and economy of scale, future applications of such microarrays will support complex parallel experiments with multiple samples such as primary cells from patients or multiple cell lines, and to test chemical compounds in search of new mechanisms of action and cellular pathways by screening for siRNA-drug interactions.

ACKNOWLEDGEMENTS

This work was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) and the National Research foundation of Korea (NRF) grant funded by the Korea government (MEST) (No.2012-00011 and No. 2012M3A9B6055466), Gyeonggi-do and KISTI.

CONFLICT OF INTEREST

The authors confirm that this chapter contents have no conflict of interest.

REFERENCES

[1]   Venter CJ, Adams MD, Myers EW. The sequence of human genome. Science 2001; 291: 1304.
[2]   Fire A, *et al.* Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 1998; 391:806-8811.
[3]   Sharp PA. RNA interference -2001. Genes Dev 2001; 15: 485-490.
[4]   Elbashir SM, Harborth J, Lendeckel W. *et al.* Duplex of 21-nucleolide RNAs mediate RNA interference in cultured mammalian cells. Nature 2001; 411: 494-498.
[5]   Sharma S & Rao A. RNAi screening: tips and techniques. Nat Immunol. 2009; 10(8):799-804.
[6]   SiRNA screening trends 2010. HTStec 2010.
[7]   Aza-Blanc P, Cooper CL, Wagner K, *et al.* Identification of modulators of TRIAL-induced apoptosis *via* RNAi-based phenotypic screening. MolCell 2003; 12:627-37.
[8]   Murrow LM, Garimella SV, Jones TL, *et al.* Identification of Wee1 as a potential molecular target in cancer cells by RNAi screening of the human tyrosine kinome. Breast Cancer Res.Treat. 2010; doi10.1007/s10549-009-0571-2.
[9]   Bantounas I, Phylatou LA, & Uney JB,. RNA interference and the use of small interfering RNA to study gene function in mammalian systems. J. of Mol. Endocrinology 2004; 33:545-557.
[10]   Berns K, Hijmans EM, Mullenders J, *et al.* A large-scale RNAi screen in human cells identifies new components of the p53 pathway. Nature 2004; 428:4431-437.
[11]   Echeverri CJ, & Perrimon N, High-throughput RNAi screening in cultured cells: a user's guide. Nature rev. genetics 2006; 7:373-384.
[12]   Seyhan AA, RNAi screens for the identification and validation of novel targets: current status and challenges. Eur. Pharmaceutical review.2010; issue 6.
[13]   Iorns E, lord C.J, Turner N, & Ashworth A, Utilizing RNA interference to enhance cancer drug discovery. Nat.Rev. Drug Discovery 2007: 6(7):556-568.
[14]   Ngo VN, Davis RE, Lamy L, *et al.* A loss-of-function RNA interference screen for molecular targets in cancer. Nature 2006; 441:106-110.
[15]   Choudhary S, Rosenblatt KP, Fang L, *et al.* High throughput siRNA screening of the human kinome identifies novel kinases controlling the canonical NF-kB activation pathway. JBC 2011; 286(43):37187-95.
[16]   Nickles D, Falschlehner C., Metzig M, & Boutros M. A genome-wide RNAi screen identifies CASP4 as a factor required for TNF-alpha signaling. Mol. Cell Biol. 2012doi:10.1128/MCB.06739-11.
[17]   Azorsa DO, Robeson RH, Frost D, *et al.* High-content siRNA screening of the kinome identifies kinases involved in Alzheimer's disease-related tau hyperphosphorylation. BMC Genomics 2010; 11(1):25.
[18]   Houzet L, & Jeang KT, Genome-wide screening using RNA interference to study host factors in viral replicaiton and pathogenesis. EMB 2010; 236:962-967.
[19]   Brass DM, Dykxhoorn DM, Benita Y, *et al.* Identification of host proteins required for HIV infection through a functional genomic screen, Science 2008; 319:921-926.
[20]   Konig R, Zhou Y, Elleder D, *et al.* Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Plos Pathog. 2009; 5:e1000437.

[21]   Yeung ML, Houxet L, Venkat SRK, *et al.* A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication. J.Biol Chem. 2009; 284: 19463-19473.

[22]   Zhou H, Xu M, Huang Q, *et al.* Genome-scale RNAi screen for host factors required for HIV replication. Cell.Hot&Microbe 2008; 4:495-504.

[23]   Genovesio A, Kwon YJ, Windisch MP, *et al.* Automated genome-wide visual profiling of cellular proteins involved in HIV infection. J. Bimolecular screening. 2011; 16:945-958.

[24]   Ng TI, Mo H, Pilot-Matias T, *et al.* Identification of host genes involved in hepatitis C virus replication by small interfering RNA technology. Hepatology 2007; 45(6):1413-1421.

[25]   Li Q, Brass AL, Ng A, *et al.* A genome-wide genetic screen for host factors required for hepatitis C virus propagation. PNAS 2009; 106(38):16410-16415.

[26]   Tai AW, Benita Y, Peng LF, *et al.* A functional genomics screen identifies cellular cofactors of Hepatitis C virus replication. Ceel Host Microbe. 2009; 19(50):298-307.

[27]   Konig R, Stertz S., Zhou Y, *et al.* Human host factors required for influenza virus replicaiton. Nature 2010; 463:813-7.

[28]   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.

[29]   Brass AL, Huang IC, Benita Y, *et al.* The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus and Dengue virus. Cell 2009; 139:1243-54.

[30]   Genovesio A, Giardini MA, Kwon YJ, *et al.* Visual genome-wide RNAi screening to identify Human host factors requeres for tripanosoma cruzi infection. PLoS One 2011; 6(5):e19733.

[31]   Pache L, Konig R, Chanda SK, Identifying HIV host cell factors by genome-scale RNAi screening.YMETH 2010; doi:10.1016/j.ymeth.2010.07.009.

[32]   Min JY & Subbarao K. Cellular targets for influenza drugs. Nature biotech. 2010; 28:239-40.

[33]   Watanabe T, Watanabe S, Kawaoka Y., l. Cellular networks involved in the influenza virus life cycle. Cell Host & microbe 2010; 7:427-439.

[34]   Mohr S, Bakal C, Perrimon N. Genomic screening with RNAi: results and chellenges. Ann. Rev.biochem 2010; 79:37-64.

[35]   Mittal V, Improving the efficiency of RNA interference in mammals. Nature rev. Genetics 2004; 5:355-365.

[36]   Jackson AL,Bartx SR, Schelter J, *et al.* Expression profiling reveals off-target gene regulation by RNAi. Nature Biotechnol 2003; 21:635-636.

[37]   Jackson A.L, Buchard J, Schelter J, *et al.* Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity RNA. 2006; 12(7):1180-1187.

[38]   Ma Y, Creanga A., Lum L, & Beachy PA, Prevalence of off-target effects in Drosophila RNA interference screens. Nature 2006; 443:359-63.

[39]   Lewis BP, Burge CB, Bartel DP, *et al.* Conserved sequence pairing, often flanked by adenosine, indicates that thousands of human genes are mictroRNA targets. Cell 2005; 120:15-20.

[40]   Lim LP, Lau NC, Garrett-Engele P, *et al.* Microarray analysis shows that some microRNAs down regulate large number of target mRNA. Nature 2005; 433:769-773.

[41]   Agrawal S, & Kandimalla ER, Antisense and siRNA as agonists of Toll-like receptors Nature Biotechnol 2004; 22:1533-1537.

[42]   Agrawal S. & Kandimalla ER, Role of Toll-like receptors in antisense and siRNA [corrected] Nature Biotechnol 2004; 22:1533-1537.

[43]   Rao DD, Vorhies JS, Senzer N. *et al.* siRNA *vs* shRNA similarities and differences. Adv.Drug. deliv. Rev.2009; 61:746-59.

[44]   Caffrey DR, Zhao J, Song Z, *et al.* siRNA off-target effects 58can be reduced at concentrations that match their individual potency. PloSOne 2011; 6:e21503.

[45]   Bauer M, Kristensen BW, Meyer M, *et al.* Toxic effects of lipid-mediated gene transfer in ventral mesencephalic explant culture. BCPT 2006; 98:395-400.

[46]   Jackson AL, Burchard J, Leake D, *et al.* Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. RNA 2006; 12:1197-1205.

[47]   Echeverri CJ, Beachy PA, Baum B, *et al.* Minimizing the risk of reporting false positives in large-scale RNAi screens. Nature methods 2006; 3:777-779.

[48]   Birmingham A, Selfors LM, Forster T, *et al.* Statistical methods for analysis of high-throughput RNA interference screens.Nat. Methods 2009; 6(8):569-575.

[49]   Schena M, Shalon D, Davis RW & Brown PO, Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 1995; 270: 467-70.

[50]   Barbulovic-Nad I, Lucente M, Sun Y, *et al.* Bio-microarray fabrication techniques- a review. Critical Reviews in Biotechnology 2006; 26t:237-259.

[51]   Ziauddin J, & Sabatini D.M, Micro-array of cells expressing defined cDNAs. Nature 2001; 411:107.

[52]   Mayer TU, Kapoor TM, Haqqarty SJ, *et al.* Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. Science 1999; 286:971-974.

[53]   Mannherz O, Mertens D, Hahn M. & Lichter P. Functional screening for pro-apoptotic genes by reverse transfection cell array technology. Genomics 2006; 87:665-672.

[54]   Mishina YM, Chrsitopher JW, Bruett L, *et al.* Multiples GPCR assay in reverse transfection cell microarrays. Journal of Biomolecular Screening 2004; 9:196-207.

[55]   Erfle H, Naumann B, Liebel U, *et al.* Reverse transfection on cell arrays for high content screening microscopy. Nature Protocols 2007; doi:10.1038/npro.2006.483.

[56]   Rantala JK, Makela R, Aaltola AR, *et al.* A cell spot microarrays method for production of high density siRNA transfection microarrays. BMG Genomics 2011; 12:162.

[57]   Mortimer I.,Tam P., MacLachlan I., *et al.* Cationic lipod-mediated transfection of cells in culture requieres mitotic activity. Gene Therapy 1999; 6;403-41195.

[58]   Lonez C., Vandenbraden M., Ruysschaert J.-M., Cationic loposomal lipods: from gene carriers to cell signalling. Preogress in lipd research 2008; 47(5):340-347.

[59]   Carpenter AE, & Sabatinin DM, Systematic genome-wide screens of gene function. Nature Reviewa; Genetics 2004; 5:11-22.

[60]   Genovesio A & Emans N, Methods and apparatus for imaging of features ona substrate. WO 006727, 2010.

[61]   Lindquist RA, Ottina KA, Wheeler DB, *et al.* Genome-scale RNAi on living-cell microarrays identifies novel regulators of Drosophila melanogaster TORC1-S6K pathway signaling. Genome research 2011; 21:433-446.

[62]   Yoshikawa T, *et al.* Transfection microarrays of human mesenchymal stem cells and on-chip siRNA gene knockdown. Journal of Controled release. 2004; 96:227-232.

[63]   Erfle H, Simpson JC, Bastiaens PIH & Pepperkok R.siRNA cell arrays for high-content screening microscopy. Biotechniques 2004; 37:454-462.

[64]   Simpson JC, Cetin C, Erfle H, *et al.* An RNAi screening platform to identify secretion machinery in mammalian cells. J of Biotechnology 2007; 129:352-65.

[65]   Erfle H, Neumann B, Rogers P, *et al.* Work flow for multiplexing siRNA assays by solid-pahse reverse trasnfection in multiwell plates. J.of Biomolecular Screening 2008;13: 575-580.

[66]   Kwon YJ, Heo JY, Kim HC, *et al.* Phenotypic microRNA microarrays. Microarrays 2013;2 (2): 63-80.

[67]   Ovcharenko D, Jarvis R, Hunicke-Smith S, *et al.* High-throughput RNAi screening *in vitro*: from cell lines to primary cells. RNA 2005; 11:985-993.

[68]   Silva JM, Mizuno H, Brady A, *et al.* RNA interference microarrays: High-throughput loss-of-function genetics in mammalian cells. PNAS 2004; 101:6548-6552.

[69]   Fjeldbo CS, Misund K, Gunther CC, *et al.* Functional studies on transfected cell microarray analysed by linear regression modeling. Nuc.Acid Rerearch. 2008; 36:e97.

[70]   Conrad C, Erfle H., Warnat P, *et al.* Automatic identification of subcellular phenotypes on human cell arrays. Genome Res. 2004; 14:1130-36.

[71]   Fjeldbo CS, Misund K, Gunther CC, *et al.* Functional studies on RNA-transfected cell microarrays. Methods in Mol.Biol 2010; 629:173-189.

[72]   Ceccarelli M & Antoniol G A deformable grid-matching approach for microarray images, IEEE Transactions on Image Processing, 2006;15, (10):. 3178- 3188.

[73]   Rueda L, *et al.* A fully automatic gridding method for cDNA microarray images. BMC Bioinformatics 2011; 12:113.

[74]   Hartelius K, & Carstensen J, Bayesian grid matching, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003;25, (2):162 - 173.

[75] Ho-Youl J, & Hwan-Gue C, An automatic block and spot indexing with k-nearest neighbors graph for microarray image analysis. Bioinformatics 2002: 18: 141-151.

[76] Brandle N, Bishof H, Lapp H, Robust DNA microarray image analysis. Machine Visison and applications 2003: 115:11-28.

[77] Carpenter AE, Jones TR, *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biology 2006; 7:R100.

[78] Gilbert DF, Meinhof T., Pepperkok R, & Runz H, Detect Tiff; a Novel image analysis routine for high-content screening microscopy. J of Biomolecular screening 2009; 14:944-955.

[79] Abramoff MD, Magalhaes PJ, Ram SJ. Image processing with ImageJ. Biophotonics International 2004; 11(7): 36-42.

[80] Boutros M, Bras LP, & Huber W, Analysis ofc ell-based RNAi screens. Genome Biology 2006; 7:R66.

[81] Malo N, Hanley JA, Cerquozzi S, *et al.* Statistical practice in high-throughput screening data analysis. Nature Biotechnology 2006; 24(2):167-175.

[82] Konig R., Chiang CY, Tu BP, *et al.* A probability-based approach for the analysis of large-scale RNAi screens. Nature Methods 2007; 4(7):847-9.

[83] Zhang XD, Yang XC, Chung N, *et al.* Robust statistical methods for hit selection in RNA interference high-throughput screening experiments. Pharmacogenomics 2006; 793:299-309.

[84] Dorsky DI, Wells M., Harrington RD, Detection of HIV-1 infection with a green fluorescent protein reporter system. Journal of acquired immune deficiency syndromes and human retrovirology : official publication of the International Retrovirology Association 1996; 13: 308.

A Three-Dimensional Spheroid Cell Culture Model for Robust High-Throughput RNA Interference Screens

Geoffrey Bartholomeusz[1,*] and Arvind Rao[2]

[1]*Department of Experimental Therapeutics, Division of Cancer Medicine, The University of Texas M.D.Anderson Cancer Center, Houston, TX.77054, USA and* [2]*Department of Bioinformatics and Computational Biology, Division of Quantitative Sciences, The University of Texas M.D.Anderson Cancer Center, Houston, TX.77054, USA*

Abstract: The sequencing of the human genome and the discovery that synthetic siRNA between 19mer and 22mer could silence genes led to the development of siRNA libraries capable of targeting all known genes within a genome. The emergence of high throughput genetic screens represent a powerful unbiased approach for identifying new targets and may fundamentally change biological research by increasing the speed with which disease mechanisms and potential drug targets can be identified. High-throughput RNAi screens are typically performed using two-dimensional monolayer cell culture models due to ease, convenience, and high cell viability. Although conventional two dimensional cell culture systems have improved our understanding of basic cell biology, the morphology and physiology of cells grown as monolayers in dish cultures differ substantially from the morphology and physiology of cells grown *in vivo* within a complex three-dimensional microenvironment. There is now a growing realization that 3D cell culture models are superior in biological studies. Three dimensional cell culture models can boost the physiological relevance of cell-based assays and advance the quantitative modeling of biological systems, from cells to organisms. These models exhibit a high degree of structural complexity and homeostasis, analogous to the complexity and homeostasis of tissues and organs. In this chapter we discuss 3D cell culture models and describe a three dimensional spheroid cell culture system and the standard operating procedure for its successful use in high throughput RNAi screens.

Keywords: 3D cell culture, high content image analysis, high Throughput screening, matrix-free nanoculture plates, RNAi, spheroid cell culture.

INTRODUCTION

RNA interference (RNAi) emerged on the global stage in 1998 when Fire and colleagues demonstrated the ability of double-stranded RNA to reduce gene

*Corresponding author Geoffrey Bartholomeusz: Department of Experimental Therapeutics, Division of Cancer Medicine, The University of Texas M.D.Anderson Cancer Center, Houston, TX.77054, USA; Tel: 713-792-4158; Fax: 713-745-1710; E-mail: gbarthol@mdanderson.org

expression in *C. elegans* [1]. Soon after, it became apparent that the RNAi pathway is conserved in eukaryotic cells and can be exploited to specifically target and cleave the mRNA using a complementary, short double-stranded RNA molecule [2-4].

In the endogenous RNAi pathway, long precursor RNA sequences with hairpin structures are processed by a ribonuclease known as Dicer into small interfering RNAs (siRNAs). The RNA-induced silencing complex (RISC) binds to the siRNA, and then components of the RISC separate the two siRNA strands. The antisense siRNA strand functions as the template allowing the RISC to bind to and cleave a complementary mRNA that is then rapidly degraded.

The realization that siRNA can be produced synthetically and delivered into mammalian cells, thus circumventing Dicer mechanics [5], opened the way for use of RNAi in a wide range of biological studies. Theoretically, with appropriately designed siRNA, researchers can silence any gene, giving RNAi a broader therapeutic potential than that of small-molecule drugs. Now that the entire human genome has been sequenced, we can develop siRNA to target and silence most human genes, and we can develop loss-of-function genetic screens using distinct siRNA libraries. These genetic screens represent a powerful unbiased approach for identifying new targets and may accelerate biological research aimed at understanding disease mechanisms and discovering potential drug targets.

RNAi screens will be invaluable for unraveling signaling pathways and will permit us to delineate networks and map pathways of gene products within cells much more rapidly and in greater detail than was previously possible. RNAi screens will also allow us to identify functional genetic differences between cell lines of diverse origin, revealing how common signaling pathways are differentially regulated in different normal and cancerous tissues. To realize the full potential of RNAi technology, however, one must adhere to a stringent set of guidelines to ensure that the data generated from these screens are reliable.

2D MONOLAYER CELL CULTURE MODELS FOR RNAi SCREENS

Most high-throughput screens of RNAi libraries have been performed using two-dimensional monolayer cell culture models. The preferred platforms in these screens have been microwell plates, tissue culture flasks, and Petri dishes. The benefits of a two-dimensional cell culture model are ease, convenience, and high cell viability.

Conventional two-dimensional cell culture systems have improved our understanding of basic cell biology. However, the morphology and physiology of cells grown as monolayers in dish cultures differ substantially from the morphology and physiology of cells grown *in vivo* within a complex three-dimensional microenvironment [6]. Therefore, growing cells as monolayers compromises fundamental investigations not only in cell and developmental biology but also in clinical research [7-11].

THREE-DIMENSIONAL SPHEROID CELL CULTURE MODELS FOR RNAi SCREENS

The intrinsic limitations of two-dimensional monolayer cell culture models have prompted the development of three-dimensional cell culture models that more closely recapitulate the complex three-dimensional microenvironment associated with normal tissues and tumors. We are now beginning to realize that three-dimensional cell culture models can improve the physiological relevance of cell-based assays and advance the quantitative modeling of biological systems, from cells to organisms, because three-dimensional models exhibit a high degree of structural complexity and homeostasis, analogous to the complexity and homeostasis of tissues and organs [7, 12, 13]. Today, three-dimensional cell culture models are gaining in popularity and are being used in a broad range of cell biology studies, including studies of tumor biology [8].

Several platforms for three-dimensional cell culture models have been developed in an attempt to simulate the heterogeneous tumor microenvironment [14-17]. These platforms include scaffolds, hydrogels, hollow fiber cultures, hanging drop models, collagen gel models, microfluidic channel-based models, and multicellular spheroid cell culture models. Of these models, the multicellular spheroid model devised by Sutherland *et al*. [18] is one of the best characterized and most widely used.

Multicellular spheroid models take advantage of the natural tendency of cells to aggregate into microscale spherical clusters. These models closely simulate the pathophysiological milieu of solid tumors [16, 19-21] and are providing new insights into tumor biology as well as differentiation, tissue organization, and homeostasis [16]. Multicellular spheroids are of an intermediate complexity between *in vivo* tumors and two-dimensional monolayer cell cultures and close to that of experimental tumors in mice and natural tumors in humans. Multicellular spheroid systems exhibit oxygen, pH, and nutrient gradients resulting in a necrotic area in the center of the spheroid similar to that observed in tumors (Fig. **1**). Cells

located in the periphery of the spheroid are actively cycling, whereas cells located in the center of the spheroid become quiescent and eventually die *via* apoptosis or necrosis to form a secondary necrotic core [22]. Because of their concentric arrangement of heterogeneous cell populations and their growth pattern, spheroids are similar to solid tumors *in vivo* at their initial, avascular stages; not-yet-vascularized micrometastatic foci; or intercapillary tumor microregions with high proliferative activity close to the capillaries, quiescent cells at an intermediate distance, and necrotic areas farther from the capillaries [20, 21, 23]. Multicellular spheroid models are therefore considered valid *in vitro* tumor models, highlighting their potential in cancer research and treatment [19-21, 24-26]. Due to their structure and composition, spheroids are a useful model to study novel hypoxic markers, targeted therapy, multicellular-mediated drug resistance, and heavy ion irradiation. They have applications in biotechnology and are straightforward to employ in high-throughput screens [27, 28].
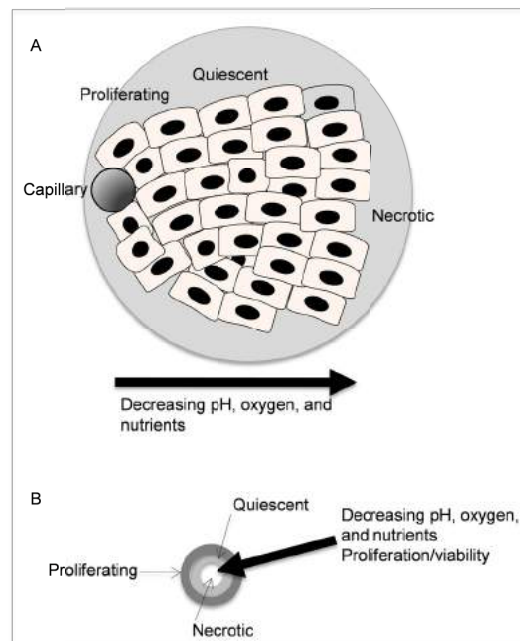
Figure 1: Similarity between (A) tumor and (B) multicellular spheroid.

There is great interest in the use of reliable and robust three-dimensional spheroid-based assays for high-throughput RNAi screening for target identification and cell signaling studies in cancer and other diseases. Kunz-Schughart and colleagues [11] have defined four important factors supporting the use of multicellular spheroids for

high-throughput screens. First, multicellular spheroids re-establish morphological, functional, and mass transport features of the corresponding tissue *in vivo*. In particular, tumor cells in multicellular spheroids restore a differentiation pattern similar to that observed *in vivo* and maintain this pattern for several weeks in culture. Second, as noted above, multicellular spheroids mimic many characteristics of avascular tumor nodules, micrometastases, or intervascular regions of large solid tumors with regard to both tumor growth kinetics and the pathophysiological micromilieu. Third, the well-defined, symmetry of multicellular spheroids allows comparison of structure to function by taking advantage of the microenvironmental gradients that affect spheroid morphology and are also spatially correlated with changes in cellular physiology. Fourth, multicellular spheroids are amenable to co-culture with different cell types, including both tumor cells and normal cells such as stromal fibroblasts, endothelial cells, or cells of the hematopoietic/immune system. In addition, cells within three-dimensional multicellular spheroids deposit extensive amounts of their own extracellular matrix [29] and form a complex three-dimensional network of cell-to-matrix and cell-to-cell interactions similar to that of tumors *in vivo* [30].

As is the case for RNAi screens performed utilizing two-dimensional monolayer cell culture models, achieving maximum benefits from an RNAi screen utilizing three-dimensional spheroid cell culture models depends on selection and optimization of an appropriate phenotypic assay and a relevant biological context. In designing RNAi screens utilizing three-dimensional spheroid cell culture models, it is important to consider the differences in growth characteristics between cells growing in a two-dimensional monolayer cell culture and cells growing in a three-dimensional spheroid cell culture. There are 3 main phases in the development of spheroids [31]. During the first phase, the cells migrate and aggregate into microspheroids. The links that bind these cells are weak, and the microspheroids are therefore fragile. In the second phase, the number of cells increases, and the general spheroid structure is reinforced. Intercellular links such as desmosomes or gap junctions appear and make the spheroid more compact. During this phase, 2 cell layers gradually appear. The external layer is proliferative because it is in contact with the nutritive medium. The internal layer is essentially nonproliferative because the farther cells are from the surface of the spheroid, the less oxygen and nutrients are available. Simultaneously, pH, osmolarity, and production of catabolites evolve as in tumors. These deficits induce a central necrosis that occurs when the spheroid diameter reaches 200 to 500 nm. The third phase in the development of a spheroid is characterized by a decline in growth rate until it reaches a plateau.

Another important consideration in designing RNAi screens utilizing three-dimensional spheroid cell culture models is the need to achieve a high efficiency of transfection. To achieve transfection efficiency in a three-dimensional spheroid cell culture model similar to that routinely achieved in two-dimensional monolayer cell culture models, it is imperative to select a matrix-free platform to eliminate any factors that might interfere with delivery of the siRNA/lipid complex into the cells.

MD ANDERSON 3D SPHEROID CELL CULTURE MODEL

Here, we describe a three-dimensional spheroid cell culture model that has been developed at The University of Texas MD Anderson Cancer Center. The platform used is a matrix-free nanoculture plate (SCIVAX Corp., Japan) comprising nanostructured scaffolds (Fig. **2**). When cells are cultured on nanoculture plates, they migrate and aggregate spontaneously into multicellular spheroids (Fig. **3B**). An important feature of this model is that once cells are dispensed into the plates, they remain as a monolayer for 24 to 48 hours before migrating and aggregating. This feature permits successful transfection of cells by the reverse transfection protocols routinely used for high-throughput RNAi screens utilizing two-dimensional monolayer cell culture models. One major difference between our three-dimensional spheroid model and two-dimensional monolayer models is that more cells are required in our three-dimensional model because a large number of cells are required to promote the formation of multicellular spheroids. When we perform screens using 96-well plates, we typically dispense 5000 to 10,000 cells per well for two-dimensional monolayer models and 20,000 to 30,000 cells per well for our three-dimensional spheroid model. Thus, reagents have to be adjusted accordingly.
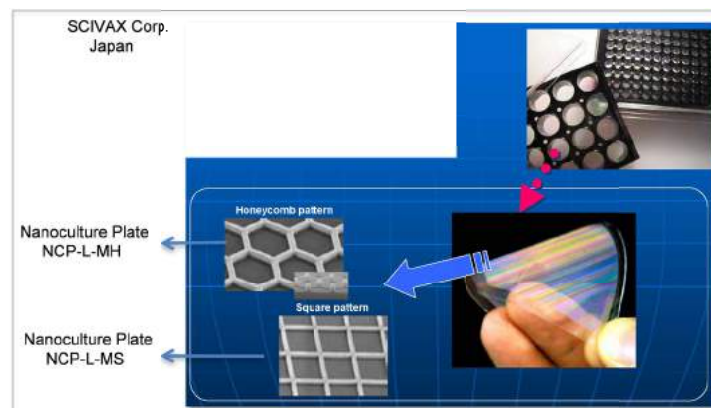
Figure 2: Matrix-free surface nanoculture NCP plate.
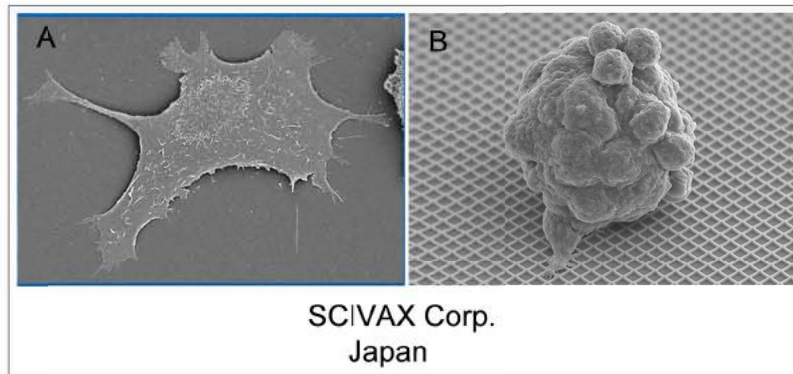
Figure 3: (A) Hela cells cultured in 2D monolayer, (B) HeLa cells cultured in NCP plate.

Assay Optimization

The most important element of a successful RNAi screen is a robust and specific assay. Development of the assay is usually the most time-consuming aspect, requiring repetition and careful attention to detail while parameters are fine-tuned and optimized. Failure to optimize all key parameters will result in a substandard screen.

One important parameter for a successful high-throughput RNAi screen is the selection of positive and negative controls to achieve high signal with the positive controls and low noise with the negative controls. The positive and negative controls yield important information about the reproducibility, robustness, and ease of the assay. In addition to selecting assay-specific controls, we have developed a standard approach to selecting the most appropriate positive and negative controls to identify the optimum transfection conditions with minimal lipid toxicity. Using a read-out of cell viability, we test a panel of positive and negative controls against a panel of eight lipid reagents (Fig. **4**). Our positive controls (siPLK1, siCOP2, siKIF11) (Dharmacon, part of GE Healthcare) lead to cell death when the corresponding targeted genes are silenced. To control sequence-independent off-target effects, we test a panel of nine non-targeting siRNA. These siRNAs used as our negative controls, are designed to avoid targeting any transcript expressed in a chosen sample and theoretically should have no effect on cell viability, which is determined quantitatively by ATP production, as measured by the activity of ATP-dependent luciferase [32, 33], or qualitatively by determining the morphology of the spheroids by imaging (Fig. **5**). An inherent problem with siRNA technology is the observation that in cases of low sequence complementarity between the siRNA and a gene the siRNA binds to the 3'-untranslated region of this secondary gene bringing about a microRNA–like effect. This phenomenon referred to as "off target effect" also

occurs even with the non-targeting siRNA as has been observed in our laboratory in the case of SCR1 silencing the EGFR gene and SCR2 silencing the luciferase gene. The inability to eliminate the "off target" effects by the use of non-targeting negative controls results in negative control-dependent reduction of viability in some cases. Therefore, for each new assay or cell combination, several candidate molecules should be tested to empirically evaluate whether the chosen negative control accurately reflect the cellular undisturbed baseline. We use a panel of nine non-targeting siRNAs (siSCR1, siSCR2, siSCR3, siSCR4, siSCR5, siOPT1, siOTP2, siOTP3, and siOTP4) (Dharmacon, part of GE Healthcare) and selected the negative controls that have negligible effects on viability. To optimize transfection and minimize variation it is beneficial to use statistical measures of data 'quality' such as the Z-factor. The Z-factor is a calculation that considers the dynamic range of the assay as well as data variability based on the measured values for internal positive and negative controls to produce a 'quality' score. Assays with a good Z-factor (0.5 - 1) provide a strong basis for a siRNA screen. The determination of the Z-factor will be based on differential toxicity (cell viability) induced by known toxic siRNAs (positive controls) – viability less than 30% and non-toxic non-targeting siRNA (negative controls) - viability greater than 80%.
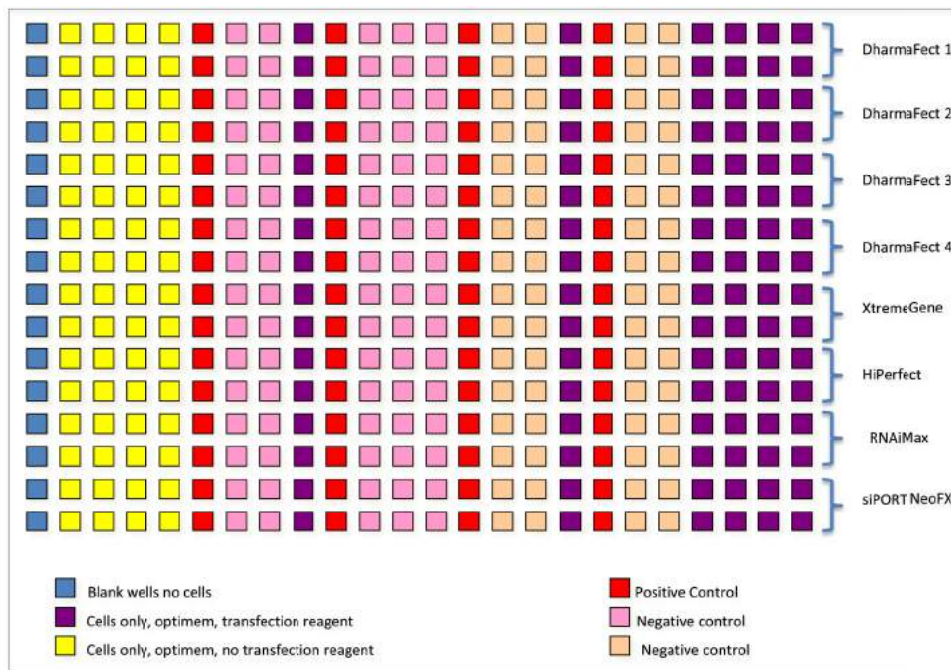
Figure 4: Assay development plate for the identification of optimum transfection conditions with minimum toxicity.
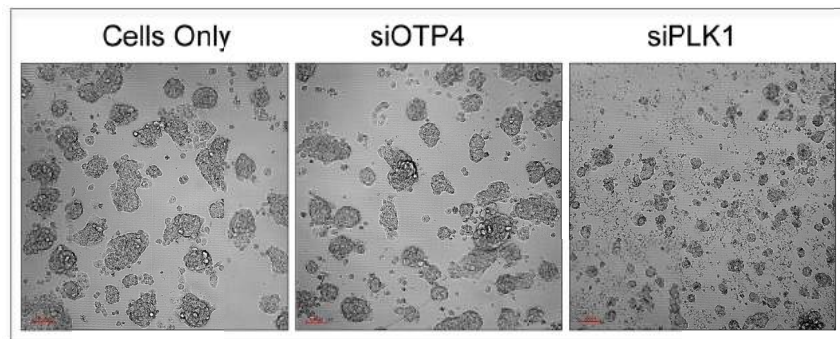
Figure 5: Cells transfected with nontargeting siRNA (OTP4 and targeting siRNA (siPLK1).

Applications of the 3D Spheroid Cell Culture Model

Three dimensional cell culture systems expand the utility of high-throughput screening and are especially suitable for interrogating basic cellular processes common to tumors.
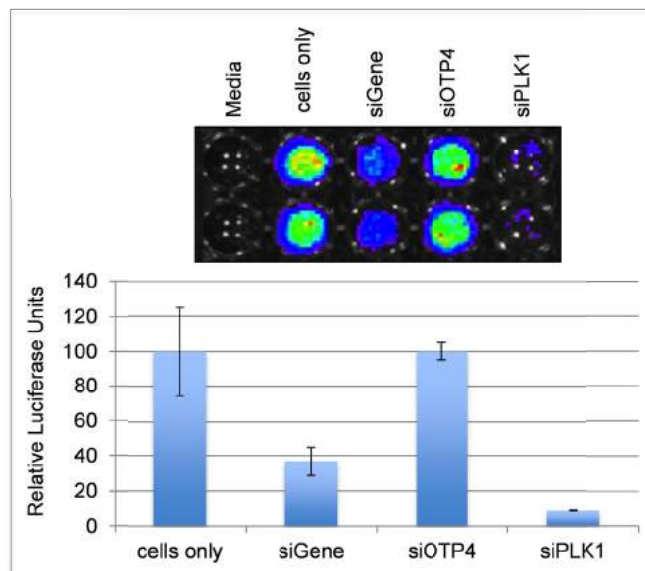
Figure 6: Reporter assay.

The simplest cell-based assay is one in which the phenotypes of a homogeneous population of cells are averaged across each well in a microtiter plate. For example, an ATP-dependent luciferase can be used to assess ATP production of the cells, which is correlated to cell viability [32, 33]. The ratio of ATP level to

ADP level, although routinely used to determine viability, can also serve as an indicator of the cells' metabolic rate. The heterogenous growth rate of cells within spheroids might have an effect on the interpretation of this assay: cells located deeper within the spheroid are quiescent and have reduced metabolic activity, and this assay might misidentify such cells as nonviable. An alternate approach is the use of more automated methods for assessment of spheroids morphology based on a microscope image analyzer.

Another cell-based assay that we routinely perform using our three dimensional spheroid cell culture model is the reporter assay. This assay is simple, robust, and reproducible. We use a cell line that stably expresses a fusion construct comprising the promoter of the gene of interest fused to luciferase. Promoter activation of the gene is determined by measuring the levels of luciferase expression. The plates are initially scanned to determine the average volume of spheroids per well. Next, luciferin is added to the medium in each well, and the plates are incubated for approximately 5 minutes to permit the luciferin to diffuse into the spheroids. The plates are then imaged to determine the intensity of luminescence per well, and this value quantitated (Fig. **6**).
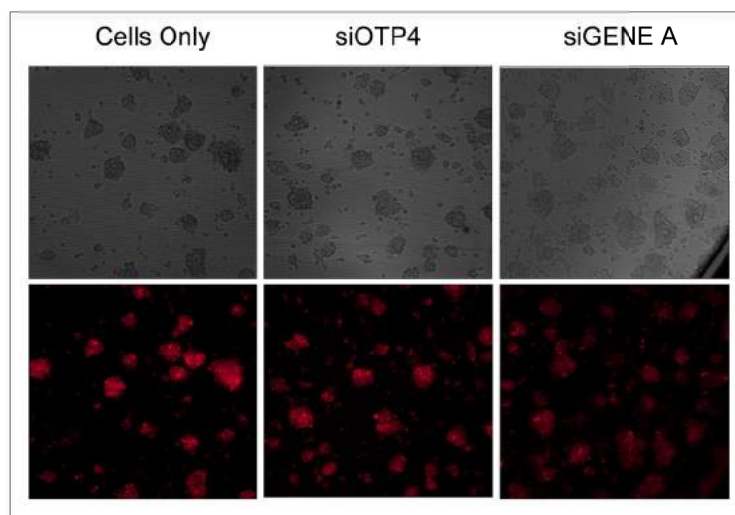
Figure 7: Hypoxia status of spheroid.

We have also used our three dimensional spheroid cell culture model for other assays with image-based readouts. For such assays, multiple images of cells are captured for each well, and the spheroids in each well are scored in terms of phenotypic descriptors, including degree of hypoxia, volume, and morphology. A

hypoxic probe, Lox-1 (SCIVAX. Inc. Japan) was used to detect the hypoxic environment of the spheroid. Lox-1 is a phosphorescence light emitting iridium complex that permeates cell membranes. Phosphorescence of Lox-1 is quenched by oxygen, but increased in response to low oxygen levels and detected as red fluorescence.

Spheroids are incubated with the probe for approximately 18 hours to permit sufficient diffusion of the probe into the spheroids. When the probe is exposed to light at its excitation wavelength, the probe emits red fluorescence. This reaction is quenched in the presence of oxygen. Spheroids with hypoxic inner cores emit intense red fluorescence, and any change in the level of oxygen in spheroids can be detected by a corresponding change in the fluorescence intensity (Fig. **7**). This assay has been successfully used in a high-throughput screen to identify proteins necessary to maintain the integrity of the spheroid architecture (personal communication)

Analysis of Images Generated Using the Three Dimensional Spheroid Cell Culture Model

The data generated from the assays described above can be quantified and analyzed by methods used to analyze data generated from two dimensional monolayer RNAi screens.
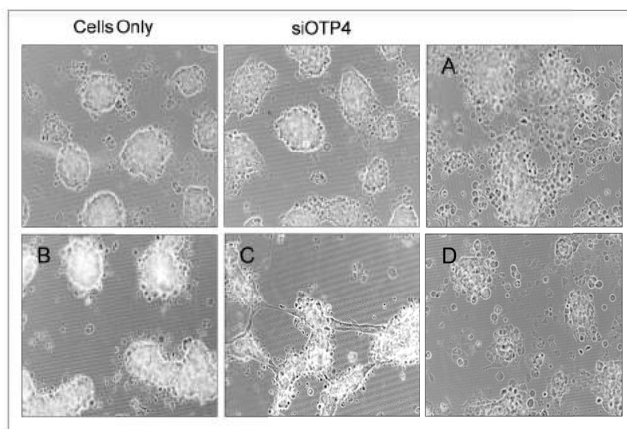
Figure 8: Changes in spheroid morphology.

One significant difference in data analysis between high-throughput screens performed using three dimensional spheroid cell culture models and those performed using two dimensional monolayer cell culture models is spheroid

morphology. In high-throughput screens using the three dimensional spheroid cell culture model, we observe a range of morphologies among the transfected cells (Fig. **8**). The methods used to analyze morphological data obtained from high-throughput primary screens using three dimensional spheroid cell culture models depend on the type of assay and the screening mode. These qualitative morphological data, which are more common in three dimensional spheroid screens, are analyzed differently from quantitative cell-based data. The analytical methods developed to analyze the morphological data are described below.

Automated analysis of the large volume of images acquired is perhaps the biggest bottleneck in the interpretation of high-throughput RNAi screens. The large number of images generated from systematic genome-wide knockdown necessitates the development of a reliable informatics infrastructure capable of handling large data volumes and facilitating integration of image data with other types of cross-modal genomic data (gene expression, methylation, copy number, protein abundance). Arguably, the cost of analysis (imaging and/or genomics analysis) is a significant if not major component of the economic costs associated with running large-scale phenotypic screens in the high-throughput RNAi setting.
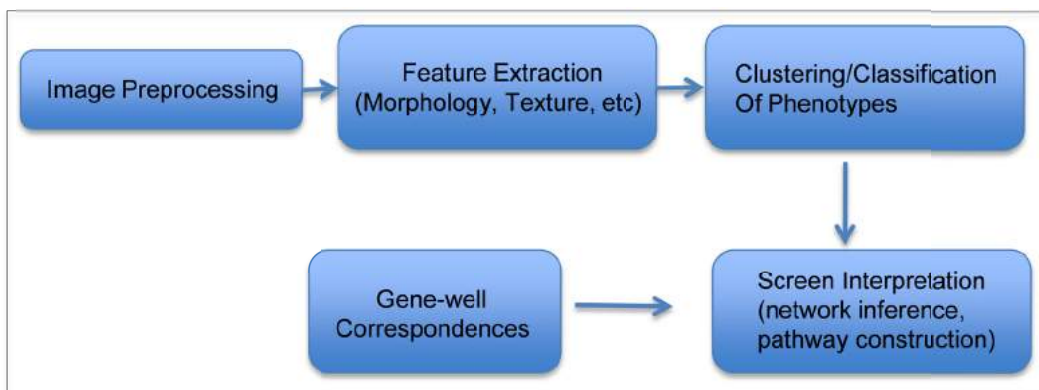
Figure 9: Workflow for high content analysis in a 3D setting.

Image analysis in the context of three dimensional culture screens can occur in either of 2 main scenarios: (i) analysis of two dimensional images of spheroids in three dimensional culture and (ii) analysis of three dimensional images of spheroids (z-stacks). Of course, higher-dimensional scenarios can also be envisioned (*e.g.*, two dimension + time, three dimension + time, three dimension + time + spectra). However, we will focus in this discussion on the 2 main cases above. In the RNAi setting, each well represents a different genetic perturbation,

so each well's image represents the phenotype resulting from that gene's knockdown. These cellular phenotypes can be measured for each cell in the image field or for the entire field (*i.e.*, the entire cell population in the well). Although field-level image processing is prevalent [34], most high content workflows pursue processing of single cells within each image field. A typical workflow for high content analysis in the three dimension setting is shown in Fig. **9**. The steps in this workflow are described below.

Image Preprocessing: Image data acquired from a microscope are initially preprocessed to correct for possible in homogeneities in illumination and contrast. The next important step, which influences almost all remaining downstream steps, is cell segmentation. Segmentation serves to segregate cellular regions from background. Cell segmentation is a field unto itself, and many sophisticated algorithms are available for 2D and 3D segmentation. However, because 3D imaging technologies have emerged only recently, there is still adequate space for the development of segmentation and feature extraction algorithms in the 3D domain [35, 36].

Feature Extraction: In the context of the screening scenarios listed above, some of the most common biological phenotypes require analyzing cells for a combination of morphological and texture features. Morphological features in general refer to measurements such as circularity, area, and shape, while texture features refer to measurements of gray-level intensity distribution of each cell image [37, 38]. The resulting feature-data matrix consists of genes along rows and cell phenotypic features along columns, *i.e.*, each row represents the feature measurements of a segmented blob (single cell or clumps of cells).

Clustering/Classification of Phenotypes: Clustering algorithms are routinely used to group image features into clusters. Each cluster defines a group of wells (genes from the RNAi screen) whose image measurements are most similar. A visual examination of the recovered clusters may reveal phenotypes of interest [39].

On the other hand, classification algorithms can be used when a human expert identifies what phenotypes are of interest a priori. These phenotypes are referred to as "labels". Such labels can be provided to a classification algorithm and used to find image features (measurements) that are associated with a phenotype of interest. Once a reliable classifier has been constructed, it can be used to classify previously unseen images into one of the desired phenotype labels [38].

Screen Interpretation: As mentioned above, clustering over image phenotypes identifies groups of genes whose knockdown results in similar phenotypes. Functional interpretation of these gene groups is necessary to relate these phenotypes to the biology of the underlying process. Similarly, classification algorithms can identify genes whose images correspond to a specified label. Genes identified using either the clustering or classification approach can then be mined for functional behavior *via* gene ontology [39], network construction, or pathway analysis (www.ingenuity.com, [40]).

Tools for 3D analysis: As mentioned earlier, there is significant research space for the development of 3D image analysis tools. There are several software packages with algorithms for 2D image analysis, and their 3D counterparts are gradually emerging. Multiple tools for 3D image analysis are now available, ranging from commercial tools like Volocity (Perkin Elmer-www.perkinelmer.com), Bitplane (www.bitplane.com), InCell Developer (www.ge.com), Pipeline Pilot (www.accelrys.com), and Definiens (www.definiens.com) to open-source tools like ImageJ (http://rsb.info.nih.gov/ij/),CellClassifier(http://acc.ethz.ch/), OMero (http://www.openmicroscopy.org/site/products/omero),CellProfiler (http://www.cellprofiler.org/), and V3D (http://vaa3d.org). There is continued interest in the development of algorithms that could simplify 3D image analysis.

Hit Validation

Even though the outcome of off-target effects can be reduced by carefully selecting target sequences and generating siRNA libraries in which each gene is targeted by a pool of siRNA the phenomenon of off target effect highlights the need to validate the selected target. Validation is performed utilizing the identical conditions selected for the primary screen. During the validation step the pooled siRNA targeting each selected hit is deconvoluted and each individual siRNA within this group tested separately. If two or more of the siRNA within each set leads to the desired outcome this candidate hit will be deemed validated. This will then be followed by Western blot analysis to determine expression levels of the targeted protein and qRTPCR to confirm gene knockdown.

The Multicellular spheroids generated in nanoculture plates are heterogeneous in size, and this feature has to be considered in the analysis of morphologies. A method was recently described for generating individual, uniformly sized spheroids in hanging drops [16, 41]. We modified this hanging drop technique to validate our targets. Utilizing the modified hanging drop technique, we tested each of the deconvoluted siRNAs directed against the same transcript to validate

the identified phenotype. Having two or more sequences elicit the same phenotype in each of the hanging drops is good proof of target specificity. An example of a hit validated for its ability to alter the hypoxic environment of the three dimensional spheroid is shown in Fig. **10**.
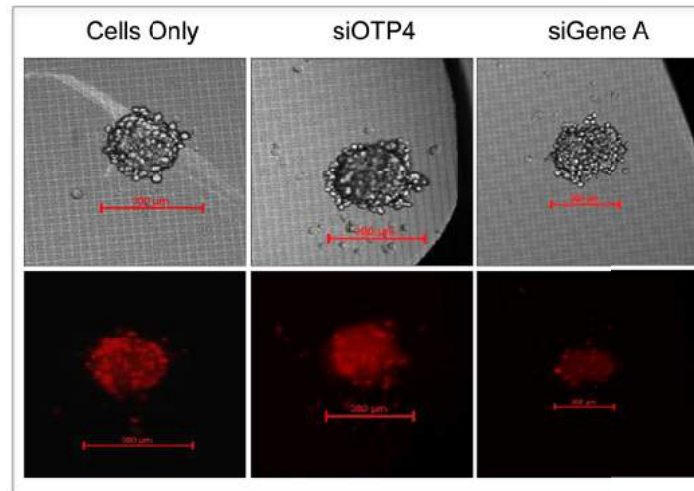
Figure 10: Hypoxia status of spheroid utilizing hanging drop model.

FUTURE DIRECTIONS

Because of their physiological relevance, three dimensional spheroid cell culture models have the potential to become an integral research tool in cell biology. Our biggest challenge in future studies will be to devise creative experiments that can capitalize on the benefits of working in three dimensions. The three dimensional spheroid cell culture model has been critical in furthering our understanding of the microenvironmental regulation of tumor cell physiology [20, 42-45]. In addition, this model has contributed considerably to our knowledge of cellular response to a variety of treatments. Our three dimensional spheroid cell culture model developed on the nanoculture plates allows us to use high-throughput RNAi screens to identify targets regulating each of the growth phases associated with this model. In some aspect, formation of spheroids is similar to the formation and growth of metastatic loci. Thus, three dimensional spheroid cell culture models have the potential to be more predictive and physiologically relevant than two dimensional monolayer models and will ultimately improve our understanding of tumor progression and improve the success rate of cancer drug candidates in preclinical studies and clinical trials.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors confirms that this chapter contents have no conflicts of interest.

REFERENCES

[1]   Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998;391(6669):806-11. Epub 1998/03/05.

[2]   Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature. 2001;411(6836):494-8.

[3]   Moffat J, Sabatini DM. Building mammalian signalling pathways with RNAi screens. Nat Rev Mol Cell Biol. 2006;7(3):177-87. Epub 2006/02/24.

[4]   Whitehead KA, Langer R, Anderson DG. Knocking down barriers: advances in siRNA delivery. Nat Rev Drug Discov. 2009;8(2):129-38. Epub 2009/01/31.

[5]   Tuschl T. RNA interference and small interfering RNAs. Chembiochem. 2001;2(4):239-45. Epub 2002/02/06.

[6]   Lee J, Cuddihy MJ, Kotov NA. Three-dimensional cell culture matrices: state of the art. Tissue Eng Part B Rev. 2008;14(1):61-86. Epub 2008/05/06.

[7]   Hess MW, Pfaller K, Ebner HL, Beer B, Hekl D, Seppi T. 3D *versus* 2D cell culture implications for electron microscopy. Methods Cell Biol. 2010;96:649-70. Epub 2010/09/28.

[8]   Pampaloni F, Reynaud EG, Stelzer EH. The third dimension bridges the gap between cell culture and live tissue. Nat Rev Mol Cell Biol. 2007;8(10):839-45. Epub 2007/08/09.

[9]   Simone NL, Paweletz CP, Charboneau L, Petricoin EF, 3rd, Liotta LA. Laser capture microdissection: beyond functional genomics to proteomics. Mol Diagn. 2000;5(4):301-7. Epub 2001/02/15.

[10]   Herrmann R, Fayad W, Schwarz S, Berndtsson M, Linder S. Screening for compounds that induce apoptosis of cancer cells grown as multicellular spheroids. J Biomol Screen. 2008;13(1):1-8. Epub 2007/11/28.

[11]   Kunz-Schughart LA, Freyer JP, Hofstaedter F, Ebner R. The use of 3-D cultures for high-throughput screening: the multicellular spheroid model. J Biomol Screen. 2004;9(4):273-85. Epub 2004/06/12.

[12]   Bissell MJ, Hall HG, Parry G. How does the extracellular matrix direct gene expression? J Theor Biol. 1982;99(1):31-68. Epub 1982/11/07.

[13]   Bissell MJ, Rizki A, Mian IS. Tissue architecture: the ultimate regulator of breast epithelial function. Curr Opin Cell Biol. 2003;15(6):753-62. Epub 2003/12/04.

[14]   Elliott NT, Yuan F. A review of three-dimensional *in vitro* tissue models for drug discovery and transport studies. J Pharm Sci. 2011;100(1):59-74. Epub 2010/06/10.

[15]   Toh YC, Zhang C, Zhang J, Khong YM, Chang S, Samper VD, *et al*. A novel 3D mammalian cell perfusion-culture system in microfluidic channels. Lab Chip. 2007;7(3):302-9. Epub 2007/03/03.

[16]   Kelm JM, Timmins NE, Brown CJ, Fussenegger M, Nielsen LK. Method for generation of homogeneous multicellular tumor spheroids applicable to a wide variety of cell types. Biotechnol Bioeng. 2003;83(2):173-80. Epub 2003/05/28.

[17]   Nichols GL, Raines MA, Vera JC, Lacomis L, Tempst P, Golde DW. Identification of CRKL as the constitutively phosphorylated 39-kD tyrosine phosphoprotein in chronic myelogenous leukemia cells. Blood. 1994;84(9):2912-8.

[18]   Sutherland RM, Durand RE. Radiation response of multicell spheroids--an *in vitro* tumour model. Curr Top Radiat Res Q. 1976;11(1):87-139. Epub 1976/01/01.

[19]   Friedrich J, Seidel C, Ebner R, Kunz-Schughart LA. Spheroid-based drug screen: considerations and practical approach. Nat Protoc. 2009;4(3):309-24. Epub 2009/02/14.

[20]   Mueller-Klieser W. Multicellular spheroids. A review on cellular aggregates in cancer research. J Cancer Res Clin Oncol. 1987;113(2):101-22. Epub 1987/01/01.

[21]   Desoize B. Contribution of three-dimensional culture to cancer research. Crit Rev Oncol Hematol. 2000;36(2-3):59-60. Epub 2000/10/18.

[22]   Desoize B, Gimonet D, Jardiller JC. Cell culture as spheroids: an approach to multicellular resistance. Anticancer Res. 1998;18(6A):4147-58. Epub 1999/01/19.

[23]   Sutherland RM. Cell and environment interactions in tumor microregions: the multicell spheroid model. Science. 1988;240(4849):177-84. Epub 1988/04/08.

[24]   Friedrich J, Ebner R, Kunz-Schughart LA. Experimental anti-tumor therapy in 3-D: spheroids--old hat or new challenge? Int J Radiat Biol. 2007;83(11-12):849-71. Epub 2007/12/07.

[25]   Jiang Y, Pjesivac-Grbovic J, Cantrell C, Freyer JP. A multiscale model for avascular tumor growth. Biophys J. 2005;89(6):3884-94. Epub 2005/10/04.

[26]   Stein AM, Demuth T, Mobley D, Berens M, Sander LM. A mathematical model of glioblastoma tumor spheroid invasion in a three-dimensional *in vitro* experiment. Biophys J. 2007;92(1):356-65. Epub 2006/10/17.

[27]   Ivascu A, Kubbies M. Rapid generation of single-tumor spheroids for high-throughput cell function and toxicity analysis. J Biomol Screen. 2006;11(8):922-32. Epub 2006/09/16.

[28]   Zhang X, Wang W, Yu W, Xie Y, Zhang Y, Ma X. Development of an *in vitro* multicellular tumor spheroid model using microencapsulation and its application in anticancer drug screening and testing. Biotechnol Prog. 2005;21(4):1289-96. Epub 2005/08/06.

[29]   Glimelius B, Norling B, Nederman T, Carlsson J. Extracellular matrices in multicellular spheroids of human glioma origin: increased incorporation of proteoglycans and fibronectin as compared to monolayer cultures. APMIS. 1988;96(5):433-44. Epub 1988/05/01.

[30]   Nederman T, Norling B, Glimelius B, Carlsson J, Brunk U. Demonstration of an extracellular matrix in multicellular tumor spheroids. Cancer Res. 1984;44(7):3090-7. Epub 1984/07/01.

[31]   Dubessy C, Merlin JM, Marchal C, Guillemin F. Spheroids in radiobiology and photodynamic therapy. Crit Rev Oncol Hematol. 2000;36(2-3):179-92. Epub 2000/10/18.

[32]   Boutros M, Kiger AA, Armknecht S, Kerr K, Hild M, Koch B, *et al*. Genome-wide RNAi analysis of growth and viability in Drosophila cells. Science. 2004;303(5659):832-5. Epub 2004/02/07.

[33]   Whitehurst AW, Bodemann BO, Cardenas J, Ferguson D, Girard L, Peyton M, *et al*. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. Nature. 2007;446:815-9.

[34]   Shariff A, Kangas J, Coelho LP, Quinn S, Murphy RF. Automated image analysis for high-content screening and analysis. J Biomol Screen. 2010;15(7):726-34. Epub 2010/05/22.

[35]   Li G, Liu T, Tarokh A, Nie J, Guo L, Mara A, *et al*. 3D cell nuclei segmentationbased on gradient flow tracking. BMC Cell Biology. 2007;4:8-40.

[36]   Kriston-Vizi J, Thong NW, Poh CL, Yee KC, Ling JS, Kraut R, *et al*. Gebiss: an ImageJ plugin for the specification of ground truth and the performance evaluation of 3D segmentation algorithms. BMC Bioinformatics. 2011;12:232. Epub 2011/06/15.

[37]   Jones TR, Carpenter AE, Lamprecht MR, Moffat J, Silver SJ, Grenier JK, *et al*. Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. Proc Natl Acad Sci U S A. 2009;106(6):1826-31. Epub 2009/02/04.

[38]   Zhang B, Pham TD. Phenotype recognition with combined features and random subspace classifier ensemble. BMC Bioinformatics. 2011;12:128. Epub 2011/05/03.

[39]   Fuchs F, Pau G, Kranz D, Sklyar O, Budjan C, Steinbrink S, *et al*. Clustering phenotype populations by genome-wide RNAi and multiparametric imaging. Mol Syst Biol. 2010;6:370. Epub 2010/06/10.

[40]   Krishnan MN, Ng A, Sukumaran B, Gilfoy FD, Uchil PD, Sultana H, *et al*. RNA interference screen for human genes associated with West Nile virus infection. Nature. 2008;455(7210):242-5. Epub 2008/08/12.

[41]   Timmins NE, Dietmair S, Nielsen LK. Hanging-drop multicellular spheroids as a model of tumour angiogenesis. Angiogenesis. 2004;7(2):97-103. Epub 2004/11/02.

[42]   Mueller-Klieser W. Tumor biology and experimental therapeutics. Crit Rev Oncol Hematol. 2000;36(2-3):123-39. Epub 2000/10/18.

[43]   Kunz-Schughart LA, Kreutz M, Knuechel R. Multicellular spheroids: a three-dimensional *in vitro* culture system to study tumour biology. Int J Exp Pathol. 1998;79(1):1-23. Epub 1998/06/06.

[44]   Mueller-Klieser W. Three-dimensional cell cultures: from molecular mechanisms to clinical applications. Am J Physiol. 1997;273(4 Pt 1):C1109-23. Epub 1997/11/14.

[45]   Santini MT, Rainaldi G. Three-dimensional spheroid model in tumor biology. Pathobiology. 1999;67(3):148-57. Epub 1999/07/08.

The Use of RNAi Technology in the Development of High Performance Bioproduction Cell Lines

Weilin Wu[1], Sabine van der Sanden[2], Paula Brooks[1], Jon M. Karpilow[3,*], Steven Oberste[2], Ralph A. Tripp[1]

[1]*University of Georgia, College of Veterinary Medicine, Dept. of Infectious Diseases, 111 Carlton Street, Athens, GA 30602, USA;* [2] *Centers For Disease Control and Prevention, 1600 Clifton Road, NE, Atlanta, GA, 30333, USA and* [3]*Thermo Fisher Scientific, 2650 Crescent Dr., Suite 202, Lafayette, CO 80026, USA*

Abstract: Vaccines have proven to be an effective means to protect communities from a range of human and agricultural pathogens. Unfortunately, costs associated with the development and manufacturing of vaccines often prevent some of the neediest populations from receiving and distributing these essential prophylactics. Advances in molecular and synthetic biology represent potential low cost solutions for enhancing bioproduction. In the following chapter, we describe a program in which RNA Interference (RNAi) has been successfully employed to identify gene modulation events that enhance poliovirus production in vaccine manufacturing cell lines. Transition of this technology into stable production lines promises to increase overall vaccine manufacturing capabilities – thereby making these essential, life-saving therapeutics available at an affordable cost.

Keywords: Cell line engineering, host-pathogen interactions, polio, RNA interference, vaccine.

INTRODUCTION

Since its discovery, applications of RNAi technology have primarily been in the areas of research and drug discovery. In these venues, small interfering RNAs (siRNA) delivered individually or as pools targeting discrete genes have successfully been used to down-regulate (silence) target transcripts. When combined with diligent validation, RNAi technology has enabled researchers to accurately assess the contributions that individual genes make to pathways, cellular phenotypes, and disease.

Early adopters initiated attempts to transition RNAi technology into applied fields. Primary among these pursuits was the application of RNAi to combat

*Corresponding author Jon Karpilow: GE Healthcare, Lafayette, CO., Thermo Fisher Scientific, 2650 Crescent Dr., Suite 202, Lafayette, CO 80026, USA; Tel: 720-890-5142; E-mail: jon.karpilow@thermofisher.com

human disease [1-5]. As was the case in previous attempts to develop nucleic acid-based therapies (*e.g.*, antisense, ribozymes) issues surrounding stability, specificity, and the single greatest hurdle, tissue-specific delivery, quickly challenged the application of this new technology in all but the most obvious of target tissues (*e.g.*, skin and ocular) [6-8]. While these hurdles have slowed the advancement of RNAi-based therapeutics, recent tangential movement of the technology into bioprocessing promises to greatly enhance the performance of current production platforms.

BIOPRODUCTION AND RNAi TECHNOLOGY

The bioproduction industry uses a wide range of prokaryotic and eukaryotic organisms to produce a variety of biomolecules including therapeutic antibodies, enzymes, hormones, and vaccines. The classic biomanufacturing workflow includes both upstream and downstream processes involved in creation and selection of cell lines expressing the biomolecule of interest, expansion of clones into small or large biofermentation reactors, removal of cells and cell debris, and purification of the biomolecule of interest from other constituents present in the media and/or cells. Upstream improvements in biomanufacturing include optimization of expression constructs, clonal selection, media formulation, and cell culture. Downstream optimization frequently focuses on improving processes tied to biomolecule purification.

A relatively small collection of well-characterized eukaryotic cell lines are currently employed in the biopharmaceutical industry. At present, nearly 80% of all biotherapeutic molecule production employs Chinese Hamster Ovary (CHO) cells with additional production platforms employing PerC6, and African Green Monkey Kidney (Vero) cells. Vaccine manufacturing utilizes a wider array of cell types including but not limited to human (MRC-5 for rubella, varicella, rabies, and diphtheria; Wi-38 for adenovirus), avian (primary chicken embryo fibroblasts for measles, mumps), primate (Vero cells for rotavirus, smallpox, influenza, and polio), and canine (MDCK cells for influenza) cell lines (see http://www.actip.org/pages/vaccinestable.html). In general, these platforms have been sufficiently characterized to accommodate large scale biofermentation. Still, a list of challenges still hinder production, including 1) suboptimal production levels of the desired biomolecule or vaccine, 2) absence of desired (or presence of undesired) post-translational modifications (PTMs), 3) aggregation or degradation (and subsequent loss) of the desired biomolecule, and 4) contamination of desired bioproducts with proteins derived from the host cell. As the costs of goods (COGS) associated with bioproduction can reach as high of 17% of total

expenditures, methods that facilitate upstream and downstream processes should be rapidly adopted to increase overall efficiency and minimize manufacturing costs.

RNAi technology can be applied to address many of the issues currently plaguing bioproduction. In the work described below, a genome-wide RNAi screen was performed as a first step in cell line engineering; enabling the identification of gene modulation events that greatly enhance the production of poliovirus in a vaccine production cell line. In the course of this screen, multiple genes with previously undisclosed anti-viral functions were identified. Elimination of individual gene functions greatly enhanced live poliovirus replication, in some cases increasing yields by greater than ten-fold.

The approach described herein has several advantages. Unlike other cell engineering efforts that build upon previous knowledge in the fields of immunology and virology, workflows that include broad up-front screening efforts have the potential of identifying novel host contributions that previously were unrecognized as "anti-viral". These discoveries can be rapidly transitioned into existing bioproduction cell lines through any number of mechanisms including gene editing by zinc finger nucleases, TALENs, and meganucleases or gene product inhibition by small molecules. Thus, unlike efforts to identify completely novel cell platforms which require lengthy regulatory approval, discoveries identified here can be rapidly transitioned into the industry. As the approach dramatically improves cellular platforms, and enables a reduction in overall production costs, the technologies hold the opportunity to greatly expand access to reagents that are essential to human and animal health.

A BRIEF HISTORY OF POLIO

The poliovirus is a small single stranded positive-sense RNA virus belonging to the picornavirus family. In a small fraction of cases, infection by the fecal-oral route leads to poliomyelitis, a debilitating paralysis resulting from infection, replication, and subsequent death of motor neurons in the CNS. Humans are the only known carriers of polio and while the annual number of paralytic cases previously ranged into the hundreds of thousands, widespread vaccination efforts have successfully reduced the number of cases by greater than 99% worldwide [9].

Despite the achievements in reducing the overall incidence of polio, it remains in a long list of vaccine-preventable diseases that would greatly benefit from the

development of a new high-production vaccine manufacturing cell line. Currently over a billion doses of polio vaccine are produced annually to prevent outbreaks across the globe. Two forms of the vaccine are currently distributed. These include the inactivated poliovirus vaccine (IPV) developed by Jonas Salk in 1952 as well as the oral vaccine (OPV) comprised of attenuated viruses developed by Albert Sabin in 1957 [10]. For the last 4 decades, the OPV vaccine has been the centerpiece in controlling poliomyelitis in developing countries. The vaccine is relatively inexpensive and provides strong immunity under ideal conditions. Unfortunately, the overall effectiveness of OPV is significantly diminished in settings where a collection of factors, including poor sanitation, the prevalence of diarrheal illnesses, immunosuppression, and the pervasiveness of rival (competing) enteric viruses, counter the overall immunogenicity of the OPV vaccine[11]. The fact that in rare cases the attenuated Sabin viruses used in OPV vaccines revert to a neurovirulent form, capable of triggering vaccine-associated paralytic poliomyelitis (VAPP), makes disease control and eradication (by OPV alone) even less attainable. Formalin-inactivated IPV vaccines are not subject to these limitations. Unfortunately, due to the costs and competing priorities in world health, expanded IPV use is not feasible. Identification of new technologies that can reduce per-dose IPV costs is viewed as one of the best mechanisms to eradicate polio.

Advances in synthetic biology and cell line engineering represent potential low cost solutions for enhancing bioproduction. Work by Jay Keasling and colleagues have recently accelerated the production of artemisinic acid, a precursor of the anti-malaria therapeutic artemisinin, by creating a novel biosynthetic pathway in a single organism [12]. Outside the field of therapeutics, collaborations between academic (U. Wisconsin) and industrial (DuPont) partners have led to the generation of synthetic organisms that greatly increase the efficiency of 1, 3, propanediol (1, 3, PDO) production, a chemical used in the manufacturing of commercial goods [13]. Similar improvements are achievable in the field of vaccine production. Large- and small-scale RNAi screens designed to identify host-pathogen interactions for human immunodeficiency virus (HIV), West Nile virus (WNV), respiratory syncytial virus (RSV), and influenza virus consistently identified rare gene knockdown events that enhance viral replication [14-17]. While these studies were not performed with the intent of enhancing vaccine manufacturing, a repeat of this work using viruses and cell lines that are relevant to the vaccine industry could significantly impact global public health.

RNAi SCREENING TO IDENTIFY GENE KD EVENTS THAT ENHANCE POLIOVIRUS PRODUCTION

The Workflow

The poliovirus vaccine cell line engineering program (Fig. **1**) includes two central workflows: target identification (Phase I) and stable cell line development (Phase II). Given the large number of hits typically identified in host-pathogen RNAi screens and the enormous cost burdens associated with stable cell line development, a critical goal of the Phase I workflow was to reduce the list of genes being considered for Phase II to a feasible cost-sensitive number. To accomplish this, Phase I incorporated multiple screening and validation filters. In addition to the primary screen, Phase I studies included 1) validation steps to ensure the identity of primary screen gene targets, 2) proof-of-principle in a vaccine production cell line, and 3) conversion of primary screen ELISA results into concrete increases in live virus titer. Furthermore, Phase I included testing with multiple poliovirus serotypes and a study to assess the effects of gene knockdown on viral immunogenicity. Additional work which included pathway analysis and multi-gene knockdown tests were performed to identify gene silencing combinations that further enhanced viral production. Phase II focuses on the creation of stable cell lines using one or more of the currently available gene editing technologies (*e.g.*, zinc finger nucleases, TALENs). Given the costs associated with this aspect of the program, pre-clinical studies that employ more cost-effective technologies such as recombinant adeno-associated virus gene editing (Horizon Discovery) are considered.
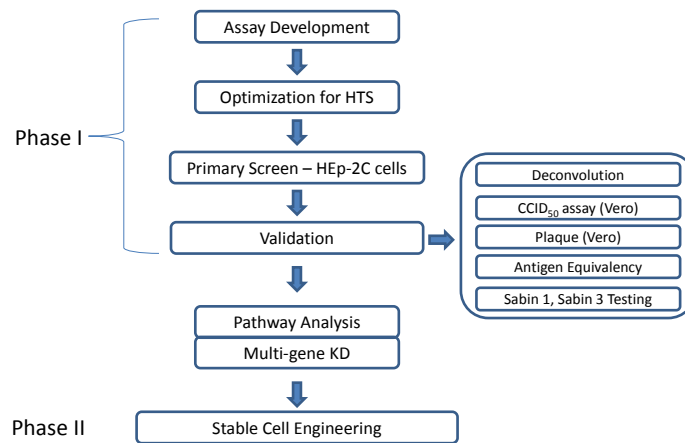
Figure 1: Schematic describing the overall workflow of the polio virus vaccine cell line engineering program. Primary screen was performed in HEp-2C cells. The hits identified from primary screen were then validated in Vero P cells.

Reagents and Assays

Cell Line and Virus Selection

While the prior art in the field pointed to clear opportunities for vaccine cell line engineering, these studies simultaneously highlighted the importance of cell line and virus selection. As has been noted in previous publications, RNAi screening results generated in three separate HIV host-pathogen screens identified distinctly disparate hit lists [18]. The discrepancies between these studies are attributed (in part) to differences in cell lines, virus genotypes, assay endpoints, and reagents. While it is true that meta-analysis of the existing data sets identified a more consistent collection of cellular pathways and networks, the take-home message in the context of this program was simple: to enhance the odds of identifying gene knockdown events that increased *vaccine production*, the virus and cell lines employed in current vaccine manufacturing had to be incorporated into the primary screen and validation processes. This proved challenging for while our group had access to both the viruses (Sabin 1, 2, 3) and cell line (Vero, African Green Monkey (AGM) kidney) currently employed in vaccine manufacturing, the silencing reagents available for a genome-wide RNAi screen were limited to the human, mouse, and rat genomes (*i.e.*, non-AGM). To further complicate matters, sequence information for Vero was not publically available, thus preventing a pre-screen bioinformatics study to assess the compatibility of human siRNA collections with the AGM genome. Given these circumstances, we were forced to decide between the lesser of two evils: screen in Vero cells and take the risk that (in some cases) the human siRNA collections would not adequately silence relevant AGM genes due to siRNA-target gene mismatches (false negatives), or screen in human cells and accept the consequences that some hits identified in human cells would not reproduce in the Vero platform (false positives). In the end, we chose the latter of the two alternatives (screen in human cells) for two reasons. First, performing the primary screen in human cells would allow us to carry out the first step of validation, pool deconvolution, with a high level of confidence. Secondly, we assumed that any hit that was 1) identified in the primary screen, and 2) confirmed by deconvolution, but 3) failed to reproduce in Vero cells, could be investigated downstream by cloning, sequencing, and (if necessary) redesign of siRNA to the Vero ortholog. Based on this reasoning, the HEp-2C cell line was selected for the primary screen. HEp-2C cells can be efficiently modulated with the commercially available RNAi reagents targeting the human genome and are effectively transfected using conventional lipid delivery reagents. Equally important, HEp-2C cells support poliovirus replication in time frames that are compatible with high-throughput screening [19].

The three poliovirus vaccine (Sabin) strains are LSc/2ab (serotype 1), P712 (serotype 2), and Leon (serotype 3). Sabin 1 has fifty-seven nucleotide substitutions that distinguish it from the parental virus. Similarly, Sabin 2 and Sabin 3 have two and ten nucleotide substitutions (respectively) that distinguish them from the parental strains. For our studies, the Sabin 2 virus was chosen for the primary screen while Sabin 1 and Sabin 3 were reserved for testing with hits that passed the initial validation studies, deconvolution and live virus assays in Vero cells. After selecting a virus and cell line for the primary screen, quantities that were sufficient for both the primary screen and downstream validations studies, were generated and banked so the same lots could be used for the entire screen.

Gene Silencing and Transfection Reagents

Previous RNAi screens have demonstrated that siRNAs can induce off-target effects through a seed-mediated process [20, 21]. Given the time and costs that can be lost to pursuing false positives, and the importance of rapidly transitioning our discoveries into a new high-value polio vaccine cell line, the ON-TARGETplus siRNA library (GE Healthcare) was chosen for the primary screen. The ON-TARGETplus siRNA collection is a chemically modified genome-wide siRNA library that contains over 72,000 siRNAs targeting greater than 18,200 protein-encoding genes of the human genome. Employed as pools (4 siRNAs per target gene), each reagent contains a specificity-enhancing chemical modification pattern that is based on the collaborative work between GE Healthcare and Rosetta Inpharmatics [22]. Gene expression profiling studies show that the ON-TARGETplus modifications greatly reduce the number of off-targeted genes. As such, application of this technology to the polio screen ensured a high degree of specificity without compromising the gene knockdown capabilities of the molecule.

In the context of the primary screen (Fig. **2A**), multiple controls were included in each 96-well screening plate to further enhance the quality and confidence in the results. These include 1) a non-targeting control having little or no effects on viral replication or cell viability (On TARGETplus Non-Targeting Control Pools, GE Healthcare, Cat. No. D-001810-10-50), and 2) a custom pool of siRNAs targeting the poliovirus VP1 & 3D genes to effectively decrease viral production (GE Healthcare). A third transfection control, TOX Transfection Control (GE Healthcare, Cat. # D-001500-01-05) which induced cell death was included as a method to evaluate transfection efficiency. Finally, a mock (untreated) control was included in each plate to provide a consistent baseline (Fig. **2B**).

Unfortunately, the described screen did not have a positive control *i.e.*, a gene KD event that increased poliovirus production.
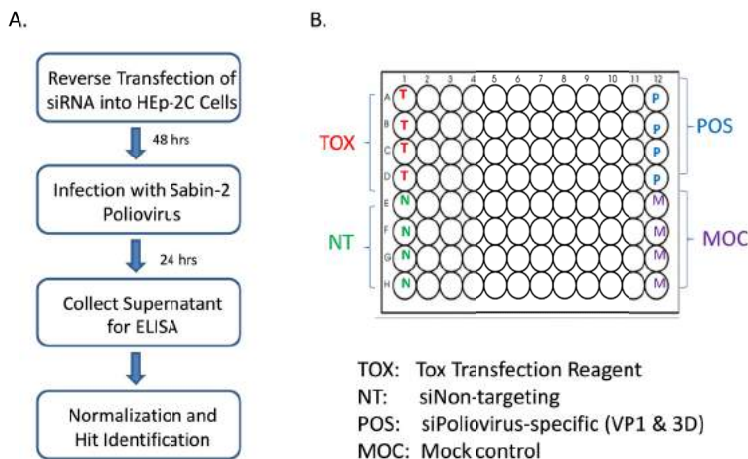
Figure 2: Schematic showing (**A**) overall workflow, and (**B**) plate format for the primary screen. Negative control (NT), positive control (POS) and mock control (MOC) were included. TOX Transfection Reagent was utilized to monitor the transfection efficiency for the entire screen.

To optimize delivery of the gene silencing reagents a batter of four lipid cocktails (GE Healthcare, DharmaFECTs 1 →4) were tested with both HEp-2C and Vero cells to identify a reagent that provided the greatest level of siRNA delivery with the negligible effects of cell viability. Based on these criteria, DharmaFECT 4 was employed in both the screening and validation studies.

Finally, in order to provide further consistency in our results, individual lots of media, supplements, and buffer were purchased for both the screen and validation studies (GE Healthcare, HyClone Reagents). The screen also employed the Nunc Edge Plate. The Edge plate design incorporates a well that encircles the perimeter of the plate and can be filled with sterile water or media. This design provides an evaporative buffer zone which minimizes edge effects frequently observed in high throughput screens. Importantly, the Nunc Edge Plate follows the American National Standards Institute (ANSI) footprint and is therefore amenable to the types of automation employed in genome-wide RNAi screening.

Assays

As stated previously, one of the goals of the program was to rapidly reduce the number of Phase I hits into a small, refined list of Phase II candidates. To achieve

this, no fewer than five separate assays were employed over the course of the primary screen and validation studies to enrich for hits that induced a narrow set of highly desired phenotypes. The primary screen (Fig. **3**) relied upon a polio-specific ELISA that detected the Sabin-2 poliovirus "D-antigen" and incorporated a mouse monoclonal antibody (HYB294-06, Thermo Scientific/Pierce). It is worth noting that special efforts were made to ensure that hits identified in the primary screen were relevant to the vaccine cell line program. Previous published host-pathogen RNAi screens have identified hundreds of host genes that play a role in viral replication. In most cases, these screens have utilized ELISA or reporter expression constructs as the primary tool for hit identification. Given the exhaustive and expensive work that is required to validate hits in our program and the absolute need for primary screen hits to convert into increases in live virus titers, a preliminary study was performed to understand how primary ELISA screen results related to live virus titers. Specifically, a small collection of ~40 hits 1) identified from screens of the kinase, protease, and ubiquitin libraries, and 2) having ELISA absorbance values ranging from 1.5-5.0 were tested in the $CCID_{50}$ (limiting dilution) live virus assay. From these studies we observed that most of the hits having absorbance values ranging from 1.5-2.7 showed consistent but modest increases in live virus titers (~2-3 fold). In contrast, hits having SD values of 3.0 (or greater) led to more significant increases in viral production (5-30x). Given the importance of identifying gene KD events that greatly increased live virus production we set the minimum absorbance cutoff for the primary screen at 3.0, thereby enhancing the likelihood of identifying hits that would provide value in vaccine manufacturing.

The polio-specific ELISA used in the primary screen was also used in the first step of our validation work, siRNA pool deconvolution. Deconvolution is a critical component of RNAi screening. As mentioned above, previous studies by multiple researchers have shown that siRNA-mediated knockdown can generate false-positive phenotypes. The mechanism behind these "off-target" effects is now well understood and one strategy for minimizing false-positives involves demonstrating that two or more siRNAs (targeting different sites on the same gene) induce the same phenotype (*e.g.* an increase in viral titer). As the primary ELISA screen was performed using pools of four separate siRNAs targeting different regions of each gene, the follow-up deconvolution study tested each of the four siRNAs individually along with as many as four additional siRNAs derived from a separate library, siGENOME (GE Healthcare).

The ELISA used in the primary screen and deconvolution studies detects the native "D-antigen" conformation of the viral proteins regardless of whether they are associated with a live virus, an inactive virus, or a suspended viral protein.

Since the IPV vaccine is derived from complete viral particles, two additional assays, a $CCID_{50}$ assay and a plaque assay, were added to the validation studies to quantitate the effects of gene knockdown on live virus particle production. Importantly, both of these assays incorporated a Vero cell line in the overall workflow. In doing so, this aspect of the validation provides a second essential piece of data that is required for the success of the program: the demonstration that the increased viral production (observed in HEp-2C cells) is similarly observed in a cell line employed in current poliovirus vaccine production.
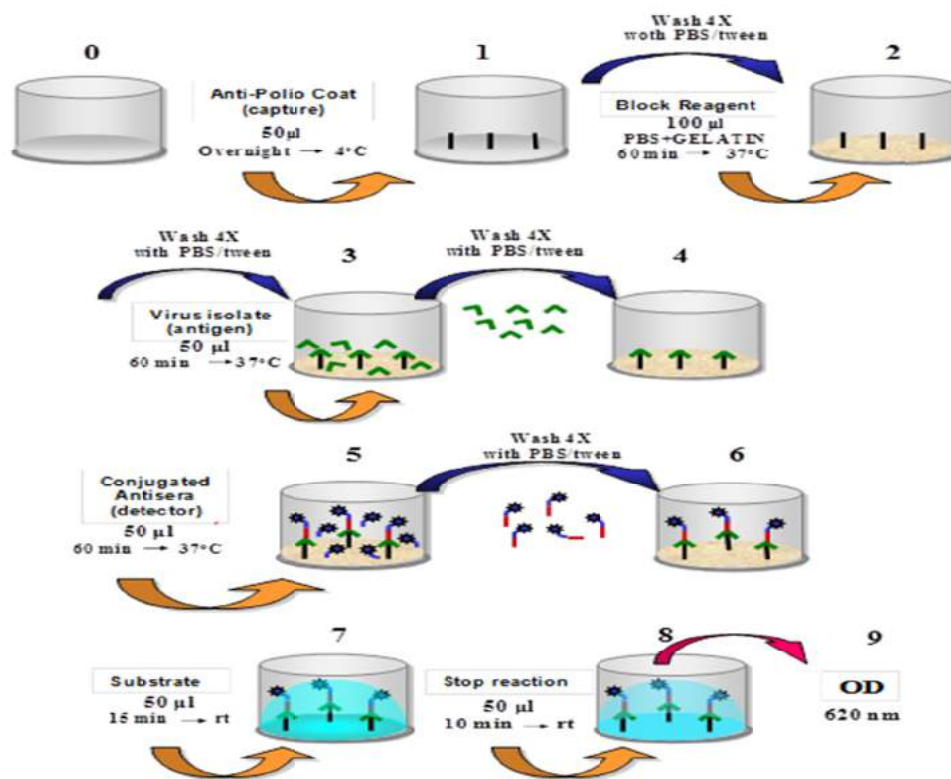
Figure 3: Diagram providing details of the polio ELISA used in the primary screen of the cell line engineering program. An in-house optimized ELISA protocol was utilized for primary screen in HEp-2C cells and validation in Vero cells.

A further step in validation involved testing how the hits identified in the primary (Sabin 2) screen affected other Sabin serotypes. As mentioned above, the current OPV vaccines incorporate three attenuated Sabin poliovirus strains (Sabin 1, 2, and 3). In contrast, current IPV vaccines utilize three distinct wildtype strains

(Note: several groups in the Netherlands, China and Japan are working to develop IPV vaccines based on the Sabin strains). Given the diverse makeup of the current vaccines, the highest priority hits would be those that positively influenced titers of a wide collection of poliovirus. To assess the effects of gene knockdown events on Sabin 1 and Sabin 3 viruses, ELISA and plaque assays were incorporated. ELISAs to detect Sabin 1 or Sabin 3 used identical assay configurations as those adopted in the primary screen but substituted serotype-specific antibodies unique to the virus under investigation (NBP1-05101 (Novus Biologicals) for Sabin poliovirus type 1; HYB 300-06 (Thermo Scientific, Pierce Products) for Sabin poliovirus type 3). Plaque assays incorporated Vero cells to assess the effect of gene knockdown on live virus production in a vaccine manufacturing cell line.

Lastly, an important aspect of our studies focused on antigen equivalency. Prior to these studies, it was unknown whether virus grown in siRNA-treated cells would be antigenically equivalent to those grown in unmodified cells. To assess the antigenic similarities between poliovirus grown in unmodified *vs.* siRNA-treated cells, a microneutralization assay was performed with Sabin 2 viruses from Vero cells transfected with siRNAs against selected genes and a pool of human sera collected from individuals previously exposed to poliovirus vaccine. In a 96-well format, 100 $CCID_{50}$ of Sabin 2 viruses from selected cell supernatants were combined with two-fold serial dilutions of the anti-polio serum, starting with a 1:8 dilution up to 1:1024. Sabin 2 viruses from cells not transfected with any siRNA were included as a control. Viruses and serum were incubated for 3 hours after which HEp-2C cells were added. After 5 days of incubation at 37°C, in 5% CO_2, cells were stained with crystal violet and endpoint serum neutralization titers calculated by the Kärber formula [23].

Summary of Results

As stated previously, the goal of the program was to rapidly identify gene knockdown events that significantly enhanced poliovirus production in a cell line currently employed in vaccine manufacturing. Fig. **4** shows the overall summary of results from our screening and validation studies. Using the relatively stringent absorbance cutoff for our primary ELISA, only 124 hits were identified in the primary screen. From this collection, 76 genes were found to have two or more siRNAs that induced the desired "increase-in-virus-production" phenotype. While the remaining 48 hits may represent true modulators of poliovirus replication, the observation that only a single siRNA increased viral protein expression amplified the risk that the observed changes were unrelated to target gene knockdown (*i.e.,* a false positive resulting from off-target effects). Overall, the primary screen and

deconvolution cutoffs allowed us to quickly reduce the number of genes being considered for further validation to a very manageable 0.4% of the whole genome.

Follow-up $CCID_{50}$ and plaque assay validation studies further reduced the list of genes that we would consider for Phase II. Using the cutoff of $\geq 5x$ increase in viral titer, fewer than 20 candidate genes out of the 76 under consideration performed to this level in both assays (Fig. **5**). The significant loss of potential candidates at this stage can be attributed to at least two factors. First, while the initial ELISA cutoff of Z-score ≥ 3.0 enhanced our chances of identifying relevant genes, this filter was not perfect. Secondly, the $CCID_{50}$ and plaque assays are the first instances where the Vero cell line was included in our validation studies. Though there is a high likelihood that siRNA designed to target genes of the human genome would also silence AGM genes, the frequency of siRNA-target mRNA mismatches (and therefore the relative efficiency of gene silencing) is unknown. These two factors in combination likely influenced the number of hits that were considered in subsequent rounds of validation.

Stage	Cell Type	Assay	Reagents	Hits
Primary Screen	Hep-2C	ELISA	OTP siRNA, pools	124 hits
Deconvolution	Hep-2C	ELISA	OTP & siGENOME siRNA, individual	76 hits
Live Virus Assays	Vero Cells	CCID50 & PA	OTP siRNA, Individual	20 hits
Ag Equivalency	Vero Cells	Micro-neutralization	OTP siRNA, Individual	20 hits
Sabin 1/Sabin3	Vero Cells	ELISA & PA	OTP siRNA, Individual	12 hits

Figure 4: Figure provides details of cell types, assays, and reagents used in Phase I of the polio vaccine cell line engineering project. "Hits" column describes number of genes that passed stage criteria. The 124 hits identified from primary screen in HEp-2 cells were first subjected to decovolution study; afterwards followed by validation in Vero cells examined by CCID-50, plaque assay and antigen equivalence test.

Surprisingly, the antigen equivalency assay which tested whether virus produced in siRNA-modified cells was recognized by sera taken from previous vaccine recipients (data not shown), did not provide a relevant filter for moving forward into Phase II studies. In all of the cases tested, the virus produced in cells where candidate genes had been silenced was recognized by patient sera (standard reference serum, CDC). In contrast, the Sabin1/Sabin 3 testing provided a strong

Phase II candidate filter. Only half of the "top 20" candidates that increased Sabin 2 production by 5x or more had similar effects on Sabin 1 and Sabin 3 serotypes. This finding supported previous work in the field that suggested small changes in cell lines, viral strains, and/or assays can lead to substantial differences in the effects of gene silencing. Overall, the combination of the primary screen ELISA together with multiple secondary validation assays allowed us to reduce the total number of Phase II candidates to approximately 10 genes or 0.005% of the original library.
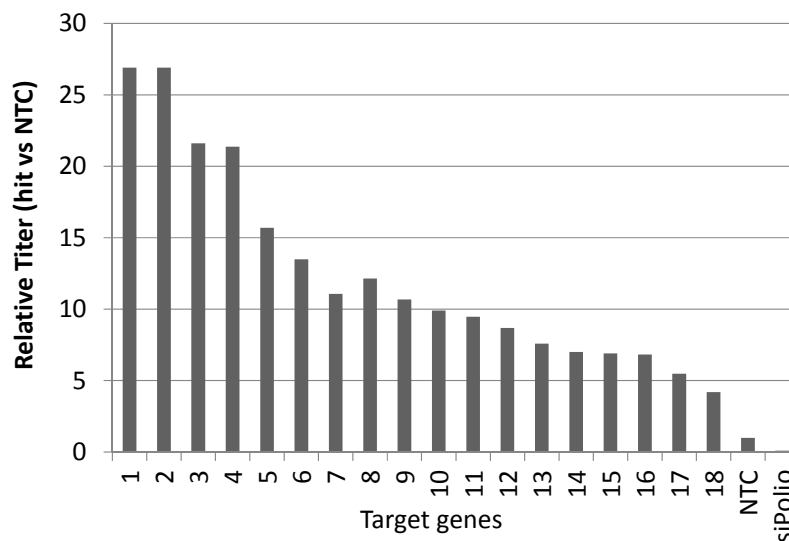
Figure 5: Graph shows the effect of single gene knockdown events on live virus titer ($CCID_{50}$ Assay) in Vero cell line for the top candidates. The hits shown here represent siRNA knowdown events which greatly enhanced poliovirus (Sabin 2) replication in Vero cells. X axis: gene number.

Moving Forward

While there is still much that needs to be done to understand the complexities and balances of cellular physiology, recent advances in gene sequencing, synthesis, and modulation have opened an array of new possibilities in synthetic biology and cell line engineering. The knowledge and techniques currently available to researchers should allow significant improvements in bioproduction. As shown here, the combination of high throughput screening and thorough validation can lead to the identification of multiple gene modulation events that significantly enhance polio virus production. These discoveries can be transitioned into novel engineered cell lines that greatly enhance vaccine production worldwide. While

the gene targets identified in this program are likely virus and cell line specific, expansion of this program to another virus-cell line combinations can identify additional host-encoded genes that (upon silencing) should enhance vaccine production. Adoption of such modified cell lines by the vaccine manufacturing industry could significantly reduce human disease and play a broad role in safeguarding agricultural stocks.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

J. Karpilow is employed by Thermo Fisher Scientific. Some of the materials described are products sold by Thermo Fisher Scientific.

REFERENCES

[1]   Blake SJ, Bokhari FF, McMillan NA. RNA interference for viral infections. Current drug targets. 2012;13(11):1411-20. Epub 2012/06/06.
[2]   Borkhardt A. Blocking oncogenes in malignant cells by RNA interference--new hope for a highly specific cancer treatment? Cancer cell. 2002;2(3):167-8. Epub 2002/09/21.
[3]   DeVincenzo JP. The promise, pitfalls and progress of RNA-interference-based antiviral therapy for respiratory viruses. Antiviral therapy. 2012;17(1 Pt B):213-25. Epub 2012/02/09.
[4]   Leung RK, Whittaker PA. RNA interference: from gene silencing to gene-specific therapeutics. Pharmacology & therapeutics. 2005;107(2):222-39. Epub 2005/05/24.
[5]   Meliopoulos VA, Andersen LE, Birrer KF, Simpson KJ, Lowenthal JW, Bean AG, *et al*. Host gene targets for novel influenza therapies elucidated by high-throughput RNA interference screens. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2012;26(4):1372-86. Epub 2012/01/17.
[6]   Paudel KS, Milewski M, Swadley CL, Brogden NK, Ghosh P, Stinchcomb AL. Challenges and opportunities in dermal/transdermal delivery. Therapeutic delivery. 2010;1(1):109-31. Epub 2010/12/07.
[7]   Thakur A, Fitzpatrick S, Zaman A, Kugathasan K, Muirhead B, Hortelano G, *et al*. Strategies for ocular siRNA delivery: Potential and limitations of non-viral nanocarriers. Journal of biological engineering. 2012;6(1):7. Epub 2012/06/13.
[8]   Wang J, Lu Z, Wientjes MG, Au JL. Delivery of siRNA therapeutics: barriers and carriers. The AAPS journal. 2010;12(4):492-503. Epub 2010/06/15.
[9]   Aylward B, Tangermann R. The global polio eradication initiative: lessons learned and prospects for success. Vaccine. 2011;29 Suppl 4:D80-5. Epub 2012/04/18.
[10]   Kalantzis G, Skiadas P, Lascaratos J. Constantin Levaditi (1874-1953): a pioneer in Immunology and Virology. Journal of medical biography. 2006;14(3):178-82.
[11]   Levine MM. Immunogenicity and efficacy of oral vaccines in developing countries: lessons from a live cholera vaccine. BMC biology. 2010;8:129. Epub 2010/10/06.

[12]   Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, *et al*. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature. 2006;440(7086):940-3. Epub 2006/04/14.

[13]   Biebl H, Menzel K, Zeng AP, Deckwer WD. Microbial production of 1,3-propanediol. Applied microbiology and biotechnology. 1999;52(3):289-97. Epub 1999/10/26.

[14]   Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, Xavier RJ, *et al*. Identification of host proteins required for HIV infection through a functional genomic screen. Science. 2008;319(5865):921-6. Epub 2008/01/12.

[15]   Kolokoltsov AA, Fleming EH, Davey RA. Venezuelan equine encephalitis virus entry mechanism requires late endosome formation and resists cell membrane cholesterol depletion. Virology. 2006;347(2):333-42. Epub 2006/01/24.

[16]   Krishnan MN, Ng A, Sukumaran B, Gilfoy FD, Uchil PD, Sultana H, *et al*. RNA interference screen for human genes associated with West Nile virus infection. Nature. 2008;455(7210):242-5. Epub 2008/08/12.

[17]   Zhou H, Xu M, Huang Q, Gates AT, Zhang XD, Castle JC, *et al*. Genome-scale RNAi screen for host factors required for HIV replication. Cell Host Microbe. 2008;4(5):495-504.

[18]   Bushman FD, Malani N, Fernandes J, D'Orso I, Cagney G, Diamond TL, *et al*. Host cell factors in HIV replication: meta-analysis of genome-wide studies. PLoS pathogens. 2009;5(5):e1000437. Epub 2009/05/30.

[19]   Albrecht P, Enterline JC, Boone EJ, Klutch MJ. Poliovirus and polio antibody assay in HEp-2 and Vero cell cultures. Journal of biological standardization. 1983;11(2):91-7. Epub 1983/04/01.

[20]   Anderson EM, Birmingham A, Baskerville S, Reynolds A, Maksimova E, Leake D, *et al*. Experimental validation of the importance of seed complement frequency to siRNA specificity. Rna. 2008;14(5):853-61. Epub 2008/03/28.

[21]   Fedorov Y, Anderson EM, Birmingham A, Reynolds A, Karpilow J, Robinson K, *et al*. Off-target effects by siRNA can induce toxic phenotype. Rna. 2006;12(7):1188-96. Epub 2006/05/10.

[22]   Jackson AL, Burchard J, Schelter J, Chau BN, Cleary M, Lim L, *et al*. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. Rna. 2006;12(7):1179-87. Epub 2006/05/10.

[23]   Flint SJE, W., Racaniello, V.R., and Skalka, A.M. Virological Methods. Principles of Virology: ASM Press; 2009.

RNAi Screening to Facilitate Drug Repurposing

Olivia Perwitasari and Ralph A. Tripp[*]

University of Georgia, College of Veterinary Medicine, Department of Infectious Diseases, 111 Carlton Street, Athens, GA 30302, USA

Abstract: Drug discovery is strangled by extraordinary time consuming and costly processes associated with high failure rates. In the United States, less than 5% of drug candidates that enter drug testing will be approved by the Food and Drug Administration (FDA) and offered for clinic use. An emerging solution to overcome this bottleneck in new drug development is to repurpose presently available drugs, a practice also known as drug repurposing. In this chapter, a general overview of drug repurposing is reviewed, along with screening methods that have yield successful outcomes. Emphasis is given on utilizing RNA interference (RNAi) screening to identify druggable genes that can be targeted by drug repurposing.

Keywords: Antiviral, drug repurposing, host-pathogen, proviral.

WHY REPURPOSE EXISTING DRUGS?

Drug development is an extremely time and resource consuming process. An average drug development program takes 10-15 years to complete and costs a billion dollars or more for the final product to reach the clinic [1-3]. These figures do not include the time and money vested in the basic research that precedes in-depth clinical studies. Despite the incredible investment associated with drug development, there is a >95% failure rate for compounds that enter the pipeline, *i.e.* one FDA-approved drug for every 10,000 compounds (illustrated in Fig. **1**). To compensate for the high failure rate there are currently 2,668 drugs in clinical trials, including new compounds, existing compounds for new indications, and new formulations (http://www.clinicaltrials.gov). Despite efforts by the National Institutes of Health (NIH) and FDA to promote new drug discovery, only 39 new drugs were approved in 2012, the second highest number of approvals by the FDA [4]. The high failure rate is especially disconcerting for development of therapeutics for emerging diseases, such as influenza virus where swift development is critical, and for rare or neglected diseases such as malaria where funding and demands are limited [5].

*Corresponding author Ralph A. Tripp: University of Georgia, College of Veterinary Medicine, Department of Infectious Diseases, 111 Carlton Street, Athens, GA 30302, USA; Tel: 706-542-1557; E-mail: ratripp@uga.edu

Figure 1
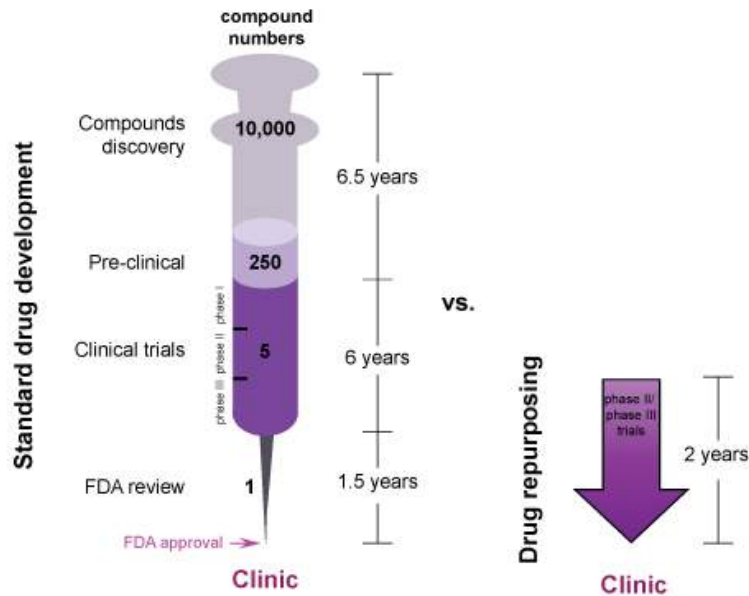
Figure 1: Drug development process. Traditional drug development process is shown on the left side of the figure (adapted from http://www.ncats.nih.gov/research/reengineering/process.html) and was compared to drug repurposing on the right side of the figure.

One alternative that could facilitate rapid drug development and accessibility to new therapeutics involves repurposing existing or previously approved drugs. This process, which is also termed drug repurposing, rescuing, repositioning, or reprofiling, makes use of existing drugs that have been successfully employed to treat unrelated diseases, or failed FDA approval in the latter phases of testing but passed the early safety trials [1, 6]. Drug repurposing takes advantage of the often non-specific and occasionally multifunctional nature of biology. By example, repurposing can result from the promiscuous (off-targeting) nature of small molecules *i.e.* a compound's ability to bind two or more different gene products that play a role in separate and unrelated pathways/diseases. Alternativelly, successful repurposing can be the consequence of a single gene product to which a molecule targets, participating in two or more biological/disease pathways. In this instance, a drug binding a single gene product can (fortuitously) be used to address afflictions that are associated with two unrelated diseases. By repurposing drugs, compounds that have already passed the initial Phase I clinical safety trial for one particular affliction can directly enter Phase II or III trials for a second, unrelated indication.

These drugs can be available for clinical use in approximately two years, as compared to the 10-15 year period required for standard drug development method (Fig. **1**). Importantly, this accelerated path, trims approximately 40% of the costs associated with standard drug development [1, 2].

Despite the stated advantages of drug repurposing there are several challenges to pursuing this approach (see Table **1**). One of these hurdles is associated with the limited access researchers have to the databases of compounds available for repurposing. Most pharmaceutical companies keep confidential lists of compounds that fail to transition into the clinic and these lists are not typically accessible to outside entities [7]. Furthermore, only a few of the publically accessible databases are searchable by genes or pathways. This shortcoming can thwart the scientific community's ability to identify relevant small molecules to gene targets identified in *e.g.*, a genome-wide screen.

Table 1: Advantages and disadvantages of drug repurposing

Advantages:

1. Cost savings - Drug repurposing reduces costs by >40% of total cost associated with traditional drug development.
    2. Rapid availability for clinic use -Safe drugs can immediately enter phase II clinical trials.
    3. Opportunity to develop therapies for emerging, rare, and neglected diseases.
    4. Predetermined pharmacokinetics and toxicology.

Disadvantages:

1. Limited database containing collection and organization of available compounds.
    2. Limited incentives for development of therapeutics against neglected diseases with negligible profitability.
    3. Difficulties to sort intellectual property issues.

Additional hurdles to adopting drug repurposing strategies come in the form of issues related to funding and intellectual property (IP). In the area of infectious disease, only 1% of new approved drugs are for treatments of neglected diseases, with the majority of these funds focused on therapeutics for malaria [8]. Given this overwhelming emphasis and the limited opportunity for profit, obtaining funding to repurpose drugs for treatments of additional rare and neglected diseases is an uphill battle. Roadblocks associated with ownership and freedom-to-operate represent a separate potential stumbling block for drug repurposing. To achieve government consent for new drugs in the United States applicants must file either a *New Drug Application* (NDA) or as an *Abbreviated New Drug Application* (ANDA). In general, therapeutic molecules are protected under patents submitted during the course of drug development and have a validity

period of twenty years plus sixty additional months of patent protection (Hatch-Waxman Act) to compensates drug manufacturers for the lengthy process associated with FDA approval. When a patent expires, a new application can be submitted to the FDA for generic versions as ANDA.

The compositions of matter claims for any particular drug are generally covered by an active pharmaceutical ingredient (API) patent and are complemented with additional patents/claims covering inventions related to formulations, methods of delivery, and treatment regimes. API patents are generally submitted very early in new drug development process, thus it is likely that these documents only cover a brief window of time following the drug's availability on the market. From the perspective of drug repurposing, new patents can cover a repurposed drug if they contain new ingredients, formulations, delivery methods, drug combinations, and (importantly from the context of this discussion) new methods of use, such as new indications. Such repurposed drugs can be developed from a previously-shelved API or an existing drug available in the market. That said, in many cases, previously-shelved therapeutics or therapeutic candidates have short (or no) patent protection what-so-ever. While this allows these entities available for immediate drug studies and repurposing, the ability to protect any new application can be limited.

DRUG REPURPOSING INITIATIVES AND RESOURCES

To combat the challenges described above, several initiatives and organizations including the Drugs for Neglected Diseases Initiative, Bill and Melinda Gates Foundation, USAID's Neglected Tropical Diseases program, and others have encouraged development of new therapeutics through public-private partnerships [9]. One noteworthy example of how these consortia can bridge the effects of basic research and therapeutic drug development is deomonstrated by the efforts to combat HIV infection. In 1987, AIDs was a full blown pandemic with no known treatments. Prior to the outbreak, the small molecule drug azidothymidine (AZT) had been developed as a treatment for cancer but was later abandoned from consideration due to the absence of efficacy during clinical trials. A partnership involving the National Cancer Institute (NCI), the Burroughs-Wellcome Company (now GlaxoSmithKline), and Duke University identified AZT as an effective therapeutic against HIV, launching the drug as the first FDA-approve treatment for the disease in 1987 [10, 11]. Remarkably, this effort took just over two years to move from the initial demonstration of AZT's anti-HIV property to its FDA approval.

Another notable initiative to promote partnerships involving industry and researchers in academic and government settings is the newly established, "Discovering New Therapeutic Uses for Existing Molecules" program launched in 2012 by the NIH's National Center for Advancing Translational Sciences (NCATS, [12, 13]. The main goal of the NCATS program is to bypass bottlenecks in drug development by allowing rapid testing and transition of existing drugs to target known diseases. Through this program, pharmaceutical companies (Abbott Laboratories,AstraZeneca, Bristol-Myers Squibb Company, Eli Lilly and Company, GlaxoSmithKline, Janssen Pharmaceutical Research & Development, L.L.C., Pfizer, and Sanofi) provide academic researchers with access to compound collections for repurposing. In addition, the program provides $20 million in funds for rapid access to multi-year research grants that focus on drug repurposing. To streamline legal and administrative processes and to guide the handling of intellectual property throughout the course of this partnership program, a memorandum of understanding was drafted between the NIH and industry partners. Templates for confidential disclosure and collaborative research agreements between are also provided and can be viewed on the NCATS website [13]. The UK's Medical Research Council (MRC) launched a similar program with AstraZeneca in late 2011, allowing access to 22 compounds for academic research [12].

An additional resource for drug repurposing is the NIH's Chemical Genomic Center (NCGC) Pharmaceutical Collection, a publicly accessible database of small molecules that have previously been accepted for human use [14]. This database contains ~2,400 of the approximately 2,750 drugs currently approved for use in the United States (FDA), European Union (EMA), Japan (NHI), and Canada (HC) (http://tripod.nih.gov/npc/). Other compound libraries that may aid drug repurposing include those developed by the National Institute of Neurological Disorders and Stroke (NINDS) which contains 1,040 compounds and Johns Hopkins Clinical Compound Library (JHCCL) containing 1,500 compounds [1, 15-17].

DRUG REPURPOSING THROUGH SERENDIPITY

AZT's repurposing resulted from an extensive collaborative effort involving governmental, academic, public, and private institutions. Several other successfully repurposed drugs were the product of serendipitous discoveries of the drug's off-target effects. One notable example is Sildenafil, or compound UK-92,480, which was developed by Pfizer in the late 1980s. Sildenafil targets phophodiesterase-5 (PDE-5) and was originally intended for treatment of

hypertension and angina [18]. In the course of clinical trials it was observed that treatment resulted in an unexpected "off-target" side effect (erectile-enhancing) in male volunteers [19, 20]. This led to a dramatic shift in Pfizer's marketing strategy for Sildenafil, targeting male patients afflicted with erectile dysfunction (ED). Under the trade name Viagra®, Sildenafil was approved as a medication for ED in 1998. This precedes its approval for the original intended indication, pulmonary arterial hypertension, in 2005 marketed under a different trade name Revatio® [19]. An additional example of serendipitous drug repurposing is minoxidil (Rogaine®) which, like Sildenafil, was developed to treat hypertension but displayed an off-target effect that promoted hair growth.

Despite these success stories, relying on serendipity for drug repurposing is ill-advised. In the following section, we provide details of how incorporation of high throughput screening, specifically genome wide RNAi screening, can facilitate drug repurposing.

RNAi SCREENING FOR THE IDENTIFICATION OF DRUG TARGETS AND DRUG REPURPOSING

RNAi is an endogenous post-transcriptional gene regulating system that utilizes small, non-coding RNAs to silence or knockdown gene expression. As reviewed in greater detail elsewhere in this book, the RNAi pathway has been utilized by researchers to investigate the contribution of a wide range of genes to cell and developmental biology. Individual or pools of synthetic RNAs, referred to as small interfering RNAs (siRNAs), can be designed to specifically target unique messenger RNAs (mRNA) and knockdown gene function for windows of time that are compatible with a range of tissue culture-based assay platforms. At the same time, RNAi libraries containing collections of siRNAs targeting thousands of genes have been developed. When combined with the appropriate mid-high throughput automation systems, research can quickly identify subsets of genes involved in a particular phenotypic responses including host-pathogen responses. An example of such technology is the siGENOME and ON-TARGETplus siRNA libraries developed by Dharmacon (part of GE Healthcare). Of a particular interest to drug repurposing efforts are the druggable libraries within these collections that contain reagents targeting approximately 7,500 genes. These "druggable genome" libraries are comprised of targets whose protein products contain functionally relevant secondary and/or tertiary structures that are predicted to be accessible by pharmacological inhibitors. By screening this subset of the genome a closer connection between the RNAi screening reagent and (eventual) small molecule discovery is achieved [21]. Importantly, these libraries are available in multiple

formats including pooled reagents (SMARTpool; mixture of four siRNA per target gene) or individual siRNAs targeting human, mouse, or rat.

A general workflow that describes how RNAi screening can facilitate drug repurposing is shown in Fig. **2**. In the first step, the cell is tranfected with an siRNA reagent(s) using conditions that are optimized for the cell line.
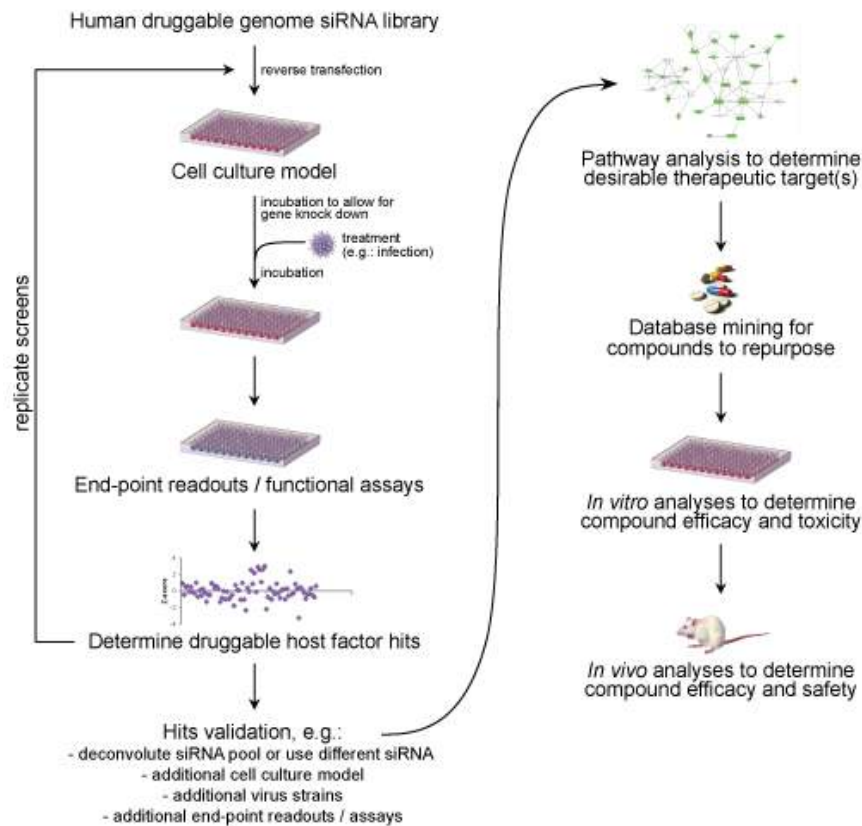
Figure 2: Workflow of siRNA screening to identify druggable target for drug repurposing.

Following a 48-72 hour period to allow for maximum RNAi-mediated gene knockdown, the cultures are treated as called for by the experimental protocol *i.e.* infection with pathogen of interest. Cultures are then assessed at desired time points post-treatment using an appropriate endpoint assay(s) to determine the contribution that each gene makes to a particular process or disease. Hits derived from the primary screen are frequently funneled through a series of validation assays designed to minimize false positives and provide additional details regarding the

contribution of the gene to the phenotype of interest. If pools of siRNAs were used in the primary screen, deconvolution, a process where each siRNA making up the pool is tested individually, often follows to demonstrate that multiple (single) gene targeting reagents induce the same phenotype. Additional validation often includes testing in complementary cell culture systems that incorporate different cell types, viruses, assay endpoints, and more. Gene targets that pass stringent validation studies are then investigated using various bioinformatic tools to gain insights into the gene's contribution in the context of the greater cellular physiology. These studies often include a gene ontology (GO), pathway mapping, and interaction databases (GeneGo MetaCore™, Ingenuity Pathway Analysis, Toppcluster) to further appreciate the contribution and pathway interaction of hits identified by the screen. [22, 23]. Potential therapeutic targets, including key pathway regulators or nodes can be identified from these *in silico* analyses and further evaluated for their druggability.

Once important regulators have been identified, researchers can mine public and/or private databases to identify drugs targeting hits identified during the screen. Publically available databases that contain both gene targets and small molecule effectors include PROMISCUOUS (http://bioinformatics.charite.de/promiscuous) [17]; ChemSpider (http://www.chemspider.com), [24]; DrugBank (http://www.drugbank.ca), [16] PubChem (http://pubchem.ncbi.nlm.nih.gov), ChEMBL (https://www.ebi.ac.uk/chembl), and the Clinician's Pocket Drug Reference (Scut manual). These platforms include anywhere from thousands to millions of compound references and in many cases provide interfaces that allow searches for targets, structures, and sources of each of the small molecules. Candidate small molecules identified through this sort of intensive database mining can be assessed (*in vitro* and *in vivo)* for relevance to any particular disease and (when appropriate) advanced toward clinical trials in an accelerated fashion.

It is worth noting that despite the extensive use of RNAi technology in biological research over the last decade [25] there are a limited number of examples of drugs being repurposed based on RNAi screens. Several factors appear to contribute to this. First, many of the hits identified by RNAi screens are not considered "druggable" in the classic sense. Certainly as screeners focus on RNAi collections that silence known drug targets, this hurdle will be overcome. Separately, at this time, the workflow that combines RNAi screening with drug repurposing has not been widely adopted by industrial or academic-industrial consortiums. This may be due to the additional efforts required for RNAi screening (*e.g.*, siRNA transfection, intensive validation studies) as compared to small molecule

screening. As the technology evolves to further simplify the workflow (*e.g.*, self-delivering siRNAs) we expect these challenges to be minimized. Other obstacles that are common to nearly all drug development investigations will need to be addressed to bridge the gap between gene target identification and therapeutically approved molecules. These include understanding dosing, delivery routes, bioavailability, and applicability to a range of cell types and/or pathogen strains. Despite these (common) hurdles, we believe the combination of RNAi screening and drug repurposing has the potential to greatly reduce the time required and financial burden needed to bring a drug to market.

CASE STUDY: USING RNAi SCREENING TO IDENTIFY PRO-INFLUENZA HOST FACTORS AND REPURPOSING THE PROTOTYPICAL OAT INHIBITOR, PROBENECID, FOR ANTI-INFLUENZA A THERAPEUTICS.

Influenza represents a global public health challenge due to its extraordinary pandemic potential. In the United States, over a quarter of a million hospitalizations and up to 49,000 fatalities are reported annually because of seasonal influenza infections [26-28]. Several influenza therapeutics are available, including the neuraminidase (NA) inhibitors zanamivir (Relenza®) and oseltamivir (Tamiflu®) and the M2 ion channel inhibitors amantadine and rimantidine [29-31]. These therapeutics are directed toward viral proteins and therefore set in motion a range of complex selective pressures for drug resistance. Unfortunately, despite the fact that drug resistance is increasing in circulating and pandemic influenza virus populations, only a small number of innovative drugs are advancing toward FDA approval[32], emphasizing the need to support programs that adopt new drug discovery approaches that have greater potential to identify novel classes of drug targets.

Both focused and genome-wide RNA interference screens have been performed to uncover cellular features required for influenza A virus replication [33-41]. In the course of infection, host factors can either suppress virus replication (anti-viral genes, such as factors associated with host immune response) or be hijacked by the virus to support its replication (proviral genes). Targeting pro-viral host factors represents a potentially innovative and refractory therapeutic approach to combat the growing issue of drug resistance. In contrast to the viral genome which is prone to a high mutation rate and therefore more capable of responding to selection pressures, normal host cells exhibit a relatively low mutation rate, thereby making them less capable of compensating to the presence of *e.g.*, a small molecule inhibitor targeting a host gene. As such, when a host-encoded proviral

function is targeted, either by RNAi reagents or small molecules, the lost function cannot be recovered by mutagenesis of the viral genome, thereby forcing the pathogen to undergo a much more dramatic evolution to compensate. While it is conceivable that viruses could adapt to become less dependent upon specific host functions (*e.g.*, by shifting its reliance onto a host gene that encodes a redundant function) these more radical changes in the viral lifecycle are considered more difficult to overcome than the single nucleotide changes often observed in viral genes that have circumvented the effects of small molecule therapeutics.

Adopting a therapeutic strategy that targets host genes is not without its own set of challenges. Targeting host-encoded proviral genes may lead to increased cellular toxicity, which would require candidate inhibitors to undergo extensive safety profile studies. Certainly, repurposing available inhibitors with known safety profile should minimize this concern. And while the majority of antiviral therapeutics are still focused on disrupting the function of viral targets, several inhibitors targeting host factors are currently being explored for anti-influenza A therapeutics. A list of current viral and host targets for influenza A virus therapeutics are included in Table **2**.

By performing an RNAi screen that employed the SMARTpool siGENOME drug target library, a recent study identified the organic anion transporter-3 (OAT3) as a pro-influenza A host factor [41]. Although OAT3 was not among the top pro-viral hits on this screen (z-score of -0.854), its potential as anti-influenza therapeutic target was evaluated due to the availability of a safe and reliable inhibitor. The classic OAT inhibitor probenecid is generally recommended as a medication for gout and other hyperuricemic disorders [42, 43]. In the study, probenecid is demonstrated to reduce growth of multiple influenza A virus subtypes *in vitro* ([41], also shown in Fig. **3**). Importantly, probenecid also limits influenza A virus infection *in vivo* when provided prophylactically or therapeutically (notably against the recent pandemic influenza A virus strain (A/California/04/09) and leads to a reduction of influenza A associated mortality and morbidity in infected mice [41]. Probenecid has been shown to maintain plasma levels of the active metabolite of oseltamivir. For this reason, it has been suggested that these two drugs be administered in combination [44-46]. This study also demonstrated the utility of oseltamivir-probenecid combinatorial treatments and ascribes the efficacy of this approach to probenecid's anti-influenza property and its role to prolong plasma oseltamivir levels. Overall, the program demonstrates how i) RNAi screening can identify a new collection of targets for anti-influenza A treatment, and ii) how probenecid and other approved medications can be repurposed for treatment of viral diseases.

Table 2: List of viral and host targets for anti-influenza A therapeutics

Target/inhibitor class	Available/known inhibitors	Mechanism of action	Reviewed in
Viral targets			
M2 ion channel inhibitor	Amantadine, Rimantadine	Block proton entry to virion and inhibit release of vRNP from endosome, Also induce premature HA maturation in trans-Golgi secretory network.	[29, 51-54]
NA inhibitor	Oseltamivir, Zanamivir, Peramivir	Block viral neuraminidase for virus release and infectivity.	[29, 55-57]
PA endonuclease inhibitor	diketo acid (L-742,001), Flutimide	Inhibit "cap-snatching" of host transcript mediated by viral acidic polymerase (PA).	[58, 59]
HA fusion inhibitor	Arbidol, HA mAb	Inhibit HA conformational change and block viral fusion	
NS1 inhibitor	NSC109834, NSC128164, NSC95676, NSC125044	Inhibit NS1 function to antagonize interferon response	[60]
NP inhibitor	Nucleozin	Prevent NP oligomerization and nuclear export	[61]
Host targets			
Proteases	7-amino-4-methylcoumarin	Host proteases are co-opted by virus for HA cleavage to release of infectious virus.	[37, 62, 63]
Sialic acid	Sialidase fusion protein	Sialidase removes surface sialic acid receptor of influenza virus.	[64]
MEK1/2	U0126	MEK inhibitors block MEK/ERK signaling important for virus replication.	[65, 66]
IKK and NFκB signaling	Bay 11-7082 (inhibits IκB phosphorylation)	Inhibition of NFκB signaling results in reduction of proinflammatory cytokines.	[66]
OAT3	Probenecid	OAT inhibitor reduces surface expression of sialic acid.	[41]
CRM1	Leptomycin B (*in vitro*), KPT335	CRM1 functions to aid vRNP nuclear export. Inhibition of CRM1 results in vRNP nuclear retention.	[67, 70]
CDC25B	NSC95397	NSC95397 inhibits CDK1 activity important for viral NS1 function.	[71]
Other			
Nucleoside analogs	Ribavirin, Favipiravir	Inhibit RNA synthesis	[68, 69]

Influenza virus proteins: M2, proton channel; NA, neuraminidase; PA, acid polymerase; HA, hemagglutinin; NS1, non-structural protein 1; NP, nucleoprotein. Host proteins: ERK, extracellular signal-regulated kinases; NFκB, nuclear factor kappa-light-chain-enhancer of activated B cells (κB); IκB, inhibitor of κB; IKK, IκB kinase; OAT3; organic anion transporter 3; CRM1, chromosome region maintenance 1; CDC25B, cell division cycle 25 homolog B.
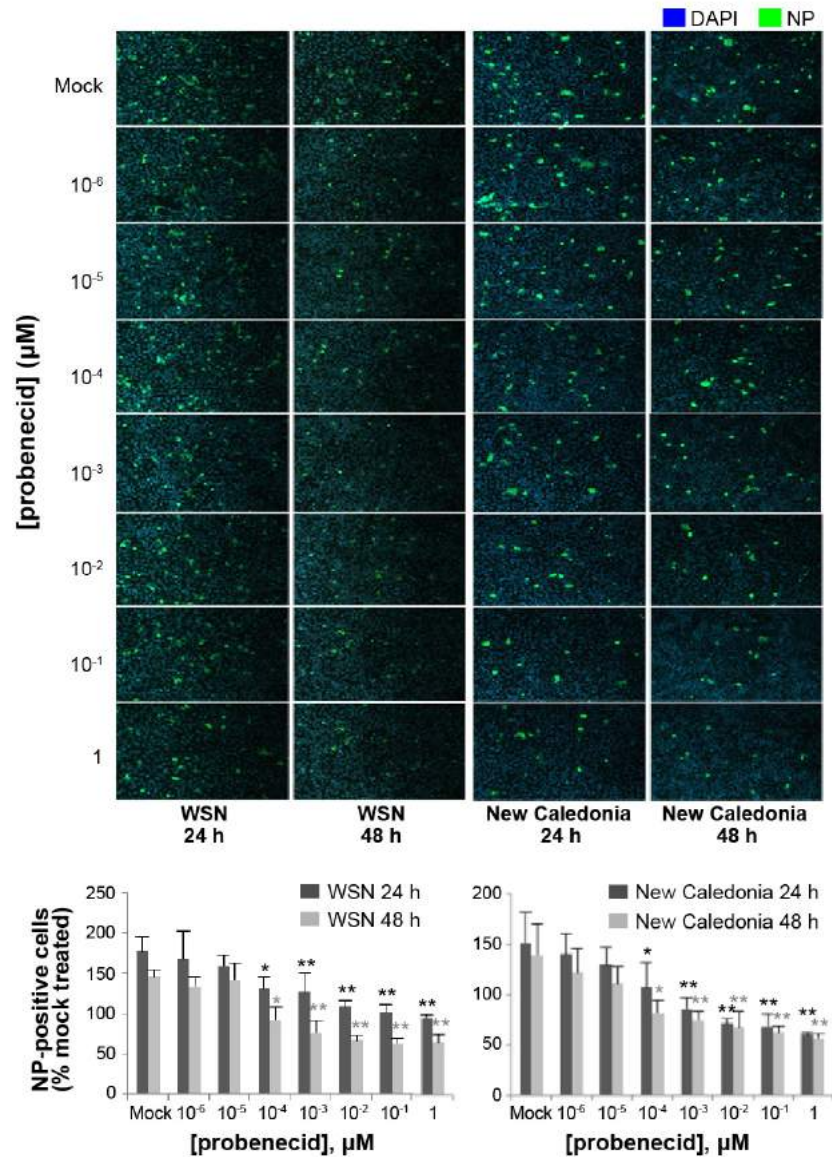
Figure 3: The OAT inhibitor, probenecid, limits influenza A virus infection *in vitro* in a dose dependent fashion. A549 human alveolar epithelial cells were exposed to increasing levels of probenecid for twenty-four hours before infection with influenza A/WSN/33(H1N1) (WSN) or A/New Caledonia/20/99(H1N1) (New Caledonia) virus at MOI = 0.001 or 0.05, respectively. At 24- or 48-hours post infection, samples were fixed, stained for influenza A virus nucleoprotein (NP; green) and nuclei (DAPI; blue), images were captured using the Cellomics array scan automated immunofluorescence microscopy system (Thermo Scientific). Number of NP-positive cells was quantified and graphed (bottom). Graphs represent results from six replicate experiments and error bars denote standard error of the mean. *,$p<0.05$; **,$p<0.005$.

OTHER APPROACHES TO DRUG REPURPOSING

Methods other than RNAi screening have been used to identify compounds for drug repurposing. As mentioned above, screens that employ small molecule repurposing libraries have been broadly used to identify compounds that possess the desired biological activities. However, as this approach does not automatically identify the gene(s) being targeted by the molecule, a significant drawback to this workflow is that it does not easily lend itself to downstream molecule optimization. Among the repurposed compounds listed in Table **3**, pyrvinium pamoate (tuberculosis and protozoal infections), tiagabine (Huntington's diseases), and ceftriaxone (Amyotrophic lateral schlerosis) were identified through NINDS library screens, while intraconazole (angiogenesis inhibitor), mycophenolic acid (lupus erythematosus), closantel (onchocerciasis), glefanine (anti-tumor), and digoxin (anti-cancer) were identified using screens of the JHCCL library (as reviewed in [5]). Separate and distinct from wet-lab screening, researchers have used *in silico* target predictions based on structural bioinformatics and drug target network analyses compiled from genomics, proteomics, metabolomics, and biomarker profiling [47-49] to repurpose drugs toward targets of interest.

Table 3: List of repurposed drugs

Drug	Original indication	New indication	Reviewed in
(±)-2-amino-3-phophonopropionic acid	Human metabolite	Malaria	[5]
Acrisorcin	Antifungal	Malaria	[5]
Amiodarone	Anti-arrhythmic	Chagas disease	[5]
Amphotericin	Antifungal	Leishmaniasis	[1]
Apomorphine	Parkinson's disease	Erectile dysfunction	[50]
Aprepitant	Nausea	Drug-resistant HIV	[5]
Arsenic	Tuberculosis and syphilis	Acute promyelocytic leukemia	[1]
Astemizole	Antihistamine	Malaria	[5]
Azidothiamidine	Cancer	HIV	[11]
Bupropion	Antidepressant	Smoking cessation	[50]
Cetriaxone	β-lactamase antibiotic	Amyotrophic lateral sclerosis	[1]
Closantel	Antihelmintic	Onchocerciasis	[5]
Dapoxetine	Antidepressant, analgesic	Premature ejaculation	[50]
Dapsone	Leprosy	Malaria	[1]

Table 1: contd…

DB289	Pneumocystis	Malaria and African trypanosomiasis	[1]
Digoxin, oubain, and proscillardin A	Congestive heart failure and arrhythmia	Cancer	[5]
Doxepin	Antidepressant	Insomnia, antipruritic	[50]
Duloxetine	Antidepressant	Urinary incontinence (stress-related)	[50]
Eflornithine	Cancer	African trypanosomiasis	[1]
Entacapone / tolcapone	Parkinson's disease	Tuberculosis	[5]
Finasteride	Benign prostatic hyperplasia	Male baldness	[50]
Fluoxetine	Antidepressant	Premenstrual dysphoria	[50]
Fosmidomycin	Urinary-tract infections	Malaria	[1]
Fumagillin	Antiamebic	Cancer (angiogenesis inhibitor)	[1]
Gemcitabine	Antiviral	Cancer	[50]
Glefenine	Analgesic	Chemotherapeutics for tumor resistance	[5]
Harmine	Cancer	Malaria	[5]
Hydroxychloroquine	Antiparasitic	Arthritis, systemic lupus erythematosus	[50]
Imidrapril	Hypertension	Cancer cachexia	[50]
Infliximab	Crohn's disease	Different arthritis form, Alzheimer's disease	[50]
Itraconazole	Antifungal	Angiogenesis inhibitor	[5]
Milnacipran	Antidepressant	Fibromyalgia	[50]
Miltefosine	Cancer	Visceral and cutaneous leishmaniasis	[1, 50]
Minicycline	Antibiotic	Amyotrophic lateral sclerosis	[1]
Minoxidil	Antihypertension	Hair loss	[50]
Mycophenolate mofetil	Immunosuppressive drug (transplant rejection)	Renal symptoms of systemic lupus erythematosus	[50]
Mycophenolic acid	Immunosuppressive drug	Angiogenesis inhibitor	[5]
Naltrexone	Opioid addiction	Alcohol withdrawal	[50]
Nitroxoline	Urinary-tract infections	Angiogenesis inhibitor	[5]
NSAID	Anti-inflammatory	Alzheimer	[1]
Paromomycin	Antiamebic	Visceral leishmaniasis	[1]
Pentamidine	Pneumonia (*Pneumocystis carnii*)	Trypanosomiasis and antimony-resistant leishmaniasis	[1]
Pioglitazone	Type-II diabetes	Nonalcoholic steatohepatitis	[50]

Table 1: contd…

Pyrvinium pamoate	Antihelmintic	Tuberculosis, protozoal infections (*Cryptosporidium parvum* and *Trypanosoma brucei*), Cancer (mitochondrial respiration)	[5]
Quiacrine	Malaria	Prion disease	[1]
Raloxifene	Breast cancer	Osteoporosis	[50]
Retinoic acid	Acne	Acute promyelocytic leukemia	[1]
Riluzole	Amyotrophic lateral sclerosis	Melanoma and other cancers	[5]
Ropinerole	Antihypertension	Parkinson's disease, restless legs syndrome	[50]
Serotonin receptor antagonists	Antipsychotic	Progressive multifocal leukoencephalopathy	[1]
Sertraline	Antidepressant	Neuroprotection for Huntington's disease	[5]
Sildenafil citrate	Antihypertension, antiangina	Erectile dysfunction	[50]
Tamoxifen	Antiestrogen	Protozoal infection	[5]
Thalidomide	Sedative / antiemetic	Cancer (angiogenesis inhibitor), erythema nodosum leprosum	[1, 50]
Tiagabine	Antiepileptic	Neuroprotection for Huntington's disease	[5]
Trimetrexate	Pneumonia (*Pneumocystis carnii*) in AIDS patients	Trypanosomiasis	[5]

AIDS, acquired immune deficiency syndrome; HIV, Human immunodeficiency virus; NSAID, Nonsteroidal anti-inflammatory drugs.

SUMMARY

The high cost, failure rate, and lengthy development time associated with standard drug development processes limit the number of new drugs available for clinic use. The NIH and other organizations have recently aimed to stimulate a more rapid drug development process by encouraging drug repurposing – a strategy that aims to find new uses for existing drugs which 1) are currently approved for treatment of unrelated diseases, or 2) failed to show adequate efficacy for the original application but passed the FDA's initial Phase I clinical trials. Repurposing drugs that have well-established pharmacokinetics and toxicity profiles allows for a more rapid deployment of these molecules against new diseases. This process can potentially reduce overall drug development times by as much as 80% and minimize costs by up to 40%. Thus, the drug repurposing strategy will facilitate rapid development and availability of desperately needed therapeutics, especially for emerging diseases, such as pandemic influenza, drug-resistant pathogens, or diseases with limited or no available therapeutics.

There are many approaches for repurposing available drugs including *in silico* prediction methods, high throughput small molecule screens, and RNAi screening. An example of an RNAi screen that led to a potentially valuable step in drug discovery is illustrated in a recent anti-influenza host-pathogen screen which identified probenecid, a drug currently available for treatment of gout and other uricemic disorders, as a potential anti-viral therapeutic. Thus, despite the workflow complexities associated with RNAi screening, this method represents a flexible target-oriented platform for drug repurposing.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors confirm that this chapter contents have no conflict of interest.

REFERENCES

[1]   Chong CR, Sullivan DJ, Jr. New uses for old drugs. Nature. 2007;448(7154):645-6. Epub 2007/08/10.
[2]   DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug development costs. Journal of health economics. 2003;22(2):151-85. Epub 2003/02/28.
[3]   DiMasi JA, Hansen RW, Grabowski HG, Lasagna L. Research and development costs for new drugs by therapeutic category. A study of the US pharmaceutical industry. PharmacoEconomics. 1995;7(2):152-69. Epub 1995/01/08.
[4]   FDA: 35 innovative new drugs approved in fiscal year 2011 Silver Spring, MD: U.S. Food and Drug Administration; 2011 [updated 11/3/11; cited 2012 10/26/12]; Available from: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm278383.htm.
[5]   Ekins S, Williams AJ, Krasowski MD, Freundlich JS. *In silico* repositioning of approved drugs for rare and neglected diseases. Drug Discov Today. 2011;16(7–8):298-310.
[6]   Collins FS. Mining for therapeutic gold. Nat Rev Drug Discov. 2011;10(6):397.
[7]   Mullard A. Could pharma open its drug freezers? Nat Rev Drug Discov. 2011;10(6):399-400.
[8]   Trouiller P, Olliaro P, Torreele E, Orbinski J, Laing R, Ford N. Drug development for neglected diseases: a deficient market and a public-health policy failure. The Lancet. 2002;359(9324):2188-94.
[9]   Butler D. Neglected diseases fund touted. Nature. 2010;465(7296):277. Epub 2010/05/21.
[10]   Fischl MA, Richman DD, Grieco MH, Gottlieb MS, Volberding PA, Laskin OL, *et al*. The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. A double-blind, placebo-controlled trial. The New England journal of medicine. 1987;317(4):185-91. Epub 1987/07/23.
[11]   Yarchoan R, Broder S. Development of antiretroviral therapy for the acquired immunodeficiency syndrome and related disorders. A progress report. The New England journal of medicine. 1987;316(9):557-64. Epub 1987/02/26.
[12]   Allison M. NCATS launches drug repurposing program. Nature biotechnology. 2012;30(7):571-2.
[13]   NCATS. Rescuing and Repurposing Drugs. Bethesda, MD: National Center for Advancing Translational Sciences (NCATS); [cited 2012 10/30/12]; Available from: http://www.ncats.nih.gov/research/reengineering/rescue-repurpose/rescue-repurpose.html.
[14]   Huang R, Southall N, Wang Y, Yasgar A, Shinn P, Jadhav A, *et al*. The NCGC pharmaceutical collection: a comprehensive resource of clinically approved drugs enabling repurposing and chemical genomics. Science translational medicine. 2011;3(80):80ps16. Epub 2011/04/29.

[15]   Grinnon ST, Miller K, Marler JR, Lu Y, Stout A, Odenkirchen J, *et al*. National Institute of Neurological Disorders and Stroke Common Data Element Project - approach and methods. Clinical trials (London, England). 2012;9(3):322-9. Epub 2012/03/01.

[16]   Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, *et al*. DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic acids research. 2008;36(Database issue):D901-6. Epub 2007/12/01.

[17]   von Eichborn J, Murgueitio MS, Dunkel M, Koerner S, Bourne PE, Preissner R. PROMISCUOUS: a database for network-based drug-repositioning. Nucleic acids research. 2011;39(Database issue):D1060-6. Epub 2010/11/13.

[18]   Terrett NK, Bell AS, Brown D, Ellis P. Sildenafil (VIAGRATM), a potent and selective inhibitor of type 5 cGMP phosphodiesterase with utility for the treatment of male erectile dysfunction. Bioorganic & Medicinal Chemistry Letters. 1996;6(15):1819-24.

[19]   Goldenberg MM. Safety and efficacy of sildenafil citrate in the treatment of male erectile dysfunction. Clinical Therapeutics. 1998;20(6):1033-48.

[20]   Raja SG, Nayak SH. Sildenafil: Emerging Cardiovascular Indications. The Annals of Thoracic Surgery. 2004;78(4):1496-506.

[21]   Russ AP, Lampel S. The druggable genome: an update. Drug Discov Today. 2005;10(23–24):1607-10.

[22]   Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. PLoS computational biology. 2012;8(2):e1002375. Epub 2012/03/03.

[23]   Kaimal V, Bardes EE, Tabar SC, Jegga AG, Aronow BJ. ToppCluster: a multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems. Nucleic acids research. 2010;38(suppl 2):W96-W102.

[24]   Williams AJ. Public chemical compound databases. Current opinion in drug discovery & development. 2008;11(3):393-404. Epub 2008/04/23.

[25]   Cejka D, Losert D, Wacheck V. Short interfering RNA (siRNA): tool or therapeutic? Clinical science (London, England : 1979). 2006;110(1):47-58. Epub 2005/12/13.

[26]   Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, *et al*. Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA : the journal of the American Medical Association. 2003;289(2):179-86. Epub 2003/01/09.

[27]   Clark NM, Lynch JP, 3rd. Influenza: epidemiology, clinical features, therapy, and prevention. Seminars in respiratory and critical care medicine. 2011;32(4):373-92. Epub 2011/08/23.

[28]   Taubenberger JK, Kash JC. Influenza Virus Evolution, Host Adaptation, and Pandemic Formation. Cell host & microbe. 2010;7(6):440-51.

[29]   Beigel J, Bray M. Current and future antiviral therapy of severe seasonal and avian influenza. Antiviral research. 2008;78(1):91-102. Epub 2008/03/11.

[30]   Dreitlein WB, Maratos J, Brocavich J. Zanamivir and oseltamivir: two new options for the treatment and prevention of influenza. 2001. p. 327-55.

[31]   Kandel R, Hartshorn KL. Novel strategies for prevention and treatment of influenza. Expert Opin Ther Targets. 2005;9(1):1-22. Epub 2005/03/11.

[32]   Beigel JH. ANTIVIRAL COMPOUNDS IN THE PIPELINE TO TACKLE H1N1 INFLUENZA INFECTION. Drugs Future.35(5):385-92. Epub 2010/05/01.

[33]   Brass AL, Huang IC, Benita Y, John SP, Krishnan MN, Feeley EM, *et al*. The IFITM Proteins Mediate Cellular Resistance to Influenza A H1N1 Virus, West Nile Virus, and Dengue Virus. Cell. 2009;139(7):1243-54.

[34]   Hao L, Sakurai A, Watanabe T, Sorensen E, Nidom CA, Newton MA, *et al*. Drosophila RNAi screen identifies host genes important for influenza virus replication. Nature. 2008;454(7206):890-3.

[35]   Karlas A, Machuy N, Shin Y, Pleissner K-P, Artarini A, Heuer D, *et al*. Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. Nature.463(7282):818-22.

[36]   Konig R, Stertz S, Zhou Y, Inoue A, Hoffmann HH, Bhattacharyya S, *et al*. Human host factors required for influenza virus replication. Nature.463(7282):813-7. Epub 2009/12/23.

[37]   Meliopoulos VA, Andersen LE, Brooks P, Yan X, Bakre A, Coleman JK, *et al*. MicroRNA regulation of human protease genes essential for influenza virus replication. PloS one. 2012;7(5):e37169. Epub 2012/05/19.

[38]   Shapira SD, Gat-Viks I, Shum BOV, Dricot A, de Grace MM, Wu L, *et al*. A Physical and Regulatory Map of Host-Influenza Interactions Reveals Pathways in H1N1 Infection. Cell. 2009;139(7):1255-67.

[39]   Sui B, Bamba D, Weng K, Ung H, Chang S, Van Dyke J, *et al*. The use of Random Homozygous Gene Perturbation to identify novel host-oriented targets for influenza. Virology. 2009;387(2):473-81.

[40]   Meliopoulos VA, Andersen LE, Birrer KF, Simpson KJ, Lowenthal JW, Bean AG, *et al*. Host gene targets for novel influenza therapies elucidated by high-throughput RNA interference screens. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2012;26(4):1372-86. Epub 2012/01/17.

[41]   Perwitasari O, Yan X, Johnson S, White C, Brooks P, Tompkins SM, *et al*. Targeting the Organic Anion Transporter-3 (OAT3) with Probenecid as a Novel Anti-Influenza A Virus Strategy. Antimicrobial agents and chemotherapy. 2012. Epub 2012/11/07.

[42]   Dantzler WH, Evans KK, Wright SH. Kinetics of interactions of para-aminohippurate, probenecid, cysteine conjugates and N-acetyl cysteine conjugates with basolateral organic anion transporter in isolated rabbit proximal renal tubules. J Pharmacol Exp Ther. 1995;272(2):663-72. Epub 1995/02/01.

[43]   Stamp LK, O'Donnell JL, Chapman PT. Emerging therapies in the long-term management of hyperuricaemia and gout. Intern Med J. 2007;37(4):258-66. Epub 2007/03/29.

[44]   Hill G, Cihlar T, Oo C, Ho ES, Prior K, Wiltshire H, *et al*. The Anti-Influenza Drug Oseltamivir Exhibits Low Potential to Induce Pharmacokinetic Drug Interactions *via* Renal Secretion€"Correlation of *in Vivo* and *in Vitro* Studies. Drug Metabolism and Disposition. 2002;30(1):13-9.

[45]   Holodniy M, Penzak SR, Straight TM, Davey RT, Lee KK, Goetz MB, *et al*. Pharmacokinetics and tolerability of oseltamivir combined with probenecid. Antimicrobial agents and chemotherapy. 2008;52(9):3013-21. Epub 2008/06/19.

[46]   Rayner CR, Chanu P, Gieschke R, Boak LM, Jonsson EN. Population pharmacokinetics of oseltamivir when coadministered with probenecid. J Clin Pharmacol. 2008;48(8):935-47. Epub 2008/06/06.

[47]   Ekins S, Mestres J, Testa B. *In silico* pharmacology for drug discovery: applications to targets and beyond. British journal of pharmacology. 2007;152(1):21-37. Epub 2007/06/06.

[48]   Iyengar R, Zhao S, Chung SW, Mager DE, Gallo JM. Merging systems biology with pharmacodynamics. Science translational medicine. 2012;4(126):126ps7. Epub 2012/03/24.

[49]   Cheng F, Liu C, Jiang J, Lu W, Li W, Liu G, *et al*. Prediction of drug-target interactions and drug repositioning *via* network-based inference. PLoS computational biology. 2012;8(5):e1002503. Epub 2012/05/17.

[50]   Henriksen K, Christiansen C, Karsdal MA. Serological biochemical markers of surrogate efficacy and safety as a novel approach to drug repositioning. Drug Discov Today. 2011;16(21-22):967-75. Epub 2011/07/13.

[51]   Davies WL, Grunert RR, Haff RF, McGahen JW, Neumayer EM, Paulshock M, *et al*. ANTIVIRAL ACTIVITY OF 1-ADAMANTANAMINE (AMANTADINE). Science (New York, NY). 1964;144(3620):862-3. Epub 1964/05/15.

[52]   Grunert RR, McGahen JW, Davies WL. THE *IN VIVO* ANTIVIRAL ACTIVITY OF 1-ADAMANTANAMINE (AMANTADINE). I. PROPHYLACTIC AND THERAPEUTIC ACTIVITY AGAINST INFLUENZA VIRUSES. Virology. 1965;26:262-9. Epub 1965/06/01.

[53]   Rabinovich S. Rimantadine therapy of influenza A infection in mice. Antimicrobial agents and chemotherapy. 1972;1(5):408-11. Epub 1972/05/01.

[54]   Rabinovich S, Baldini JT, Bannister R. Treatment of influenza. The therapeutic efficacy of rimantadine HC1 in a naturally occurring influenza A2 outbreak. The American journal of the medical sciences. 1969;257(5):328-35. Epub 1969/05/01.

[55]   Thomas GP, Forsyth M, Penn CR, McCauley JW. Inhibition of the growth of influenza viruses *in vitro* by 4-guanidino-2,4-dideoxy-N-acetylneuraminic acid. Antiviral research. 1994;24(4):351-6. Epub 1994/08/01.

[56]   Calfee DP, Hayden FG. New approaches to influenza chemotherapy. Neuraminidase inhibitors. Drugs. 1998;56(4):537-53. Epub 1998/11/07.

[57]   Li W, Escarpe PA, Eisenberg EJ, Cundy KC, Sweet C, Jakeman KJ, *et al*. Identification of GS 4104 as an orally bioavailable prodrug of the influenza virus neuraminidase inhibitor GS 4071. Antimicrobial agents and chemotherapy. 1998;42(3):647-53. Epub 1998/03/28.

[58]   DuBois RM, Slavish PJ, Baughman BM, Yun M-K, Bao J, Webby RJ, *et al*. Structural and Biochemical Basis for Development of Influenza Virus Inhibitors Targeting the PA Endonuclease. PLoS Pathog. 2012;8(8):e1002830.

[59]   Tomassini JE, Davies ME, Hastings JC, Lingham R, Mojena M, Raghoobar SL, *et al*. A novel antiviral agent which inhibits the endonuclease of influenza viruses. Antimicrobial agents and chemotherapy. 1996;40(5):1189-93. Epub 1996/05/01.

[60]   Basu D, Walkiewicz MP, Frieman M, Baric RS, Auble DT, Engel DA. Novel influenza virus NS1 antagonists block replication and restore innate immune function. J Virol. 2009;83(4):1881-91. Epub 2008/12/05.

[61]   Kao RY, Yang D, Lau LS, Tsui WH, Hu L, Dai J, *et al*. Identification of influenza A nucleoprotein as an antiviral target. Nature biotechnology. 2010;28(6):600-5. Epub 2010/06/01.

[62]   Baron J, Tarnow C, Mayoli-Nüssle D, Schilling E, Meyer D, Hammami M, *et al*. Matriptase, HAT and TMPRSS2 activate the hemagglutinin of H9N2 influenza A viruses. Journal of Virology. 2012.

[63]   Meyer D, Sielaff F, Hammami M, Bottcher-Friebertshauser E, Garten W, Steinmetzer T. Identification of the first synthetic inhibitors of the type II transmembrane serine protease TMPRSS2 suitable for inhibition of influenza virus activation. The Biochemical journal. 2013;452(2):331-43. Epub 2013/03/27.

[64]   Malakhov MP, Aschenbrenner LM, Smee DF, Wandersee MK, Sidwell RW, Gubareva LV, *et al*. Sialidase fusion protein as a novel broad-spectrum inhibitor of influenza virus infection. Antimicrobial agents and chemotherapy. 2006;50(4):1470-9. Epub 2006/03/30.

[65]   Pleschka S, Wolff T, Ehrhardt C, Hobom G, Planz O, Rapp UR, *et al*. Influenza virus propagation is impaired by inhibition of the Raf/MEK/ERK signalling cascade. Nat Cell Biol. 2001;3(3):301-5.

[66]   Pinto R, Herold S, Cakarova L, Hoegner K, Lohmeyer J, Planz O, *et al*. Inhibition of influenza virus-induced NF-kappaB and Raf/MEK/ERK activation can reduce both virus titers and cytokine expression simultaneously *in vitro* and *in vivo*. Antiviral research. 2011;92(1):45-56.

[67]   Watanabe K, Takizawa N, Katoh M, Hoshida K, Kobayashi N, Nagata K. Inhibition of nuclear export of ribonucleoprotein complexes of influenza virus by leptomycin B. Virus Res. 2001;77(1):31-42. Epub 2001/07/14.

[68]   Kiso M, Takahashi K, Sakai-Tagawa Y, Shinya K, Sakabe S, Le QM, *et al*. T-705 (favipiravir) activity against lethal H5N1 influenza A viruses. Proceedings of the National Academy of Sciences. 2010;107(2):882-7.

[69]   Oxford JS. Inhibition of the Replication of Influenza A and B Viruses by a Nucleoside Analogue (Ribavirin). Journal of General Virology. 1975;28(3):409-14.

[70]   Perwitasari O, Johnson S, Yan X, Howerth E, Shacham S, Landesman Y, *et al.* Verdinexor, a novel selective inhibitor of nuclear export, reduces influenza a virus replication in vitro and in vivo. J Virol. 2014;88(17):10228-43.

[71]   Perwitasari O, Torrecilhas AC, Yan X, Johnson S, White C, Tompkins SM, Tripp RA. Targeting cell division cycle 25 homolog B to regulate influenza virus replication. J Virol. 2013;87(24):13775-84.

Author Index

Subject Index

Ralph Tripp

Dr. Ralph Tripp is a professor in the College of Veterinary Medicine at the University of Georgia where he oversees the research activities at the Animal Health Research Center (AHRC). Dr. Tripp's broad interests include basic research designed to elucidate host-pathogen interactions, as well as translational studies in animal and human vaccines, small molecule antiviral therapeutics and pathogen biosensing technologies that employ nanotechnology. In addition to being an Eminent Scholar of the Georgia Research Foundation, Dr. Tripp is the Associate Director of the university's recently launched Center for Molecular Medicine.

Jon Karpilow

Dr. Jon Karpilow is a manager in the Biosciences Division of Thermo Fisher Scientific where he has worked on the development of synthetic and viral-based RNAi technologies for over a decade. During the past five years, Dr. Karpilow has worked closely with members of the Tripp lab as well as researchers at the Centers for Disease Control and Prevention to create a new generation of genetically engineered cell lines capable of increased polio vaccine production.