Title
Pooled screening to reveal the primary effectors of miR-142

Author
Blau, James Alan

Publication Date
2015

Peer reviewed|Thesis/dissertation

Pooled screening to reveal the primary effectors of miR-142

by

Jamès Alan Blau

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Pharmaceutical Sciences and Pharmacogenomics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Dedication

To Jessica, who can probably write her own thesis on showing patience, love, and support for a chronically anxious husband in grad school.

Acknowledgements

The work that I present here was made possible by the creative mind of Dr. Michael McManus. His passion for using the latest technology to explore the dark matter of the genome has allowed me to gain a valuable skillset that will propel me to great heights in my young career. Another one of Michael's great talents is one of a general manager, bringing onto his team a strong roster that has great chemistry together. I've had the privilege of being in one of the most supportive labs anyone can ask for. I kindly acknowledge Robert Lebbink, Theresa Chan, and Weronika Patena, whose efforts toward starting this project allowed me to hit the ground running when I joined the lab. I thank Nikki Shariat and Gregory Ku, who all helped guide the way in my first years here. I particularly thank Ian Vaughn, Matthew Hangauer, Michael Boettcher, Sergio Covarrubias, and Susan Carpenter, for the countless screen discussions and for being invaluable helping hands during some of my more tedious experiments. I thank members of my thesis committee, David Erle and Nadav Ahituv, for the many great discussions related to my thesis project, and committing me to think of the bigger picture. Finally, I thank the several technicians, past and present, in the UCSF Viracore who must have processed thousands of different lentivirus preps for me over the years – you made my life so much easier.

Pooled screening to reveal the primary effectors of miR-142

By James Alan Blau

Abstract

Although microRNAs are key regulators of gene expression, few studies have thoroughly evaluated the upstream regulators of microRNA activity. We seek to understand the upstream regulators of miR-142, a microRNA thought to be important for B cell lymphogenesis and lymphocyte proliferation. We performed sequential whole-genome RNAi and focused CRISPR screens in a human B cell line to identify genes affecting miR-142 activity on a reporter construct bearing several perfect binding sites. Top hits include known core microRNA pathway genes, but also other genes such as *PSME4*, *FXR1*, *SKIV2L*, and *FAM208A*. The observation of *FXR1* as a top screen hit is intriguing given that it has been observed to post-transcriptionally promote the expression of several neuronal microRNAs, but has not been demonstrated for other microRNAs in general. The results gathered thus far demonstrate the power of using pooled high-throughput screening and genome-editing approaches to discover regulators of microRNA biogenesis and activity.

Table of Contents

Dedication..... iii

Acknowledgements..... iv

Abstract..... v

List of Tables..... vii

List of Figures..... viii

Chapter 1 – General Introduction..... 1

Chapter 2 – Development and Execution of microRNA Sensor Screens..... 11

Introduction..... 11

Results..... 13

Discussion..... 22

Methods..... 28

Chapter 3 – General Summary..... 67

References..... 71

Publishing Agreement..... 90

List of Tables

Table 1: Vectors used throughout study..... 55

Table 2: Top 50 genes by significance of sgRNA enrichment (Mann-Whitney rank sum test)... 56

Table 3: List of genes and enriched sgRNAs selected from sgRNA screen for validation..... 58

Table 4: List of primers used to amplify and deep-sequence samples from shRNA and CRISPR screens..... 66

List of Figures

Figure 1: *AGO2* shRNA-dependent shift of miR-142 sensor expression..... 37

Figure 2: Schematic of genome-wide shRNA screen in sensor cells..... 38

Figure 3: Performance of genome-wide shRNA screen..... 39

Figure 4: Analysis of top genes associated with enriched shRNAs..... 40

Figure 5: GFP validation of enriched shRNAs in sensor cells..... 41

Figure 6: *AGO2* and *MIR142* sgRNA-dependent shift of miR-142 sensor expression..... 42

Figure 7: Efficacy of CRISPR-Cas9 system on miR-142 locus..... 43

Figure 8: SgRNAs targeting hit genes from shRNA screen do not recapitulate GFP phenotypes of shRNAs..... 44

Figure 9: Rescue experiment for miR-142-3p activity after *FAM172A* knockdown..... 45

Figure 10: Cloning strategy and design for nuclease-active CRISPR-Cas9 sgRNA library..... 46

Figure 11: Schematic of nuclease-active CRISPR-Cas9 sgRNA screen..... 47

Figure 12: Performance of nuclease-active CRISPR-Cas9 sgRNA screen..... 48

Figure 13: Representation of sgRNAs across pooled library..... 49

Figure 14: Analysis of top genes associated with enriched sgRNAs..... 50

Figure 15: Validation of CRISPR-Cas9 sgRNA screen hits by Zeocin..... 51

Figure 16: Changes in sensor mRNA in sensor cells with sgRNAs against screen hit genes..... 52

Figure 17: Changes in sensor mRNA in sensor cells with dCas9-KRAB sgRNAs against screen hit genes..... 53

Figure 18: Changes in miR-142-3p expression in cells with sgRNAs against hit genes from CRISPR-Cas9 screen..... 54

Chapter 1 – General Introduction

MicroRNAs (miRNAs) have been fervently studied across almost all fields of basic biology and disease since their discovery in animals (1) and humans (2). The reasons for such heavy interest are plentiful. MiRNAs established a novel and simple means for translational inhibition of genes. This translational inhibition, dubbed miRNA-mediated repression, was found to affect genetic networks important to nearly all diseases and development. Also, the mechanism by which miRNAs function could be exploited through short interfering RNAs (siRNAs) to directly degrade messenger RNAs (mRNAs) in a process called RNA interference (RNAi); this enables researchers to inhibit proteins that are not druggable, and expands the potential therapeutic options to explore for diseases. Given the power for miRNAs and siRNAs to manipulate gene expression, it has become imperative to understand which gene networks are fine-tuned through miRNAs, and which gene networks can be exploited by siRNAs to elicit a desired phenotype. As such, a considerable number of studies have been centered on identifying the important mRNA targets for miRNAs. However, few studies have focused on the regulation of miRNA biogenesis itself. It is imperative to have a clear understanding of the network of proteins responsible for processing miRNAs transcribed from the genome and for mediating miRNA/siRNA inhibition. To put it simply: what regulates the regulators?

The biogenesis of miRNAs and shRNAs

MiRNAs are small, ~21-25-nucleotide (nt) non-coding RNAs that predominantly recruit the RNA-induced silencing complex (RISC) to target mRNAs for silencing (3). The core proteins required for the recognition and processing of most miRNAs have been identified and characterized (4).

Genes that miRNAs are derived from are diverse. Most miRNAs originate from mRNA transcripts (5) expressed from Pol-II promoters (6). While the majority of miRNAs in mammals may derive from noncoding transcripts, it appears that most conserved miRNAs derive from the same transcripts as those for protein coding genes (7–9). MiRNAs from these transcripts are often 'intronic,' as they originate from the introns of the transcripts and do not necessarily affect the stability of the mature mRNA transcript (10), while others are 'exonic.' MiRNAs deriving from non-coding Pol-II transcripts may be considered 'intergenic' in relation to traditional protein-coding genes if the transcript is not a component of or overlapped by any protein-coding gene. MiRNAs can also derive from Pol-III transcripts (11) and from more exotic sources like mirtrons, snoRNAs, and tRNAs (12). In a similar fashion, short hairpin RNAs (shRNAs), which are transgenically-expressed to produce siRNAs (13), are typically expressed from either Pol-II or Pol-III transcripts, often using the context sequence and hairpin shape of a endogenous pri-miRNA.

For the vast majority of Pol-II transcripts destined to yield miRNAs (primary precursor miRNAs, or 'pri-miRNAs'), the first catalytic step involves the Microprocessor, minimally composed of DGCR8 and the RNase III enzyme Drosha, which cleaves hairpins in pri-miRNA to form ~70nt precursor miRNA (pre-miRNA) (14,15). The pre-miRNA retains the hairpin shape, with a loop at one end and a 2nt 3'-overhang on the other (16). More rarely, miRNAs can be directly spliced from short introns and trimmed to a pre-miRNA without requiring Microprocessor; these species are called mirtrons (17).

shRNAs from Pol-II transcripts are recognized and cropped by Microprocessor via the same mechanism. However, shRNAs from Pol-III transcripts are usually designed to not require any sort of cropping; since the transcription start and end of a Pol-III transcript are more

predictable, the transcript can be designed to directly produce an shRNA with a 2nt 3'-overhang, akin to the shape of a pre-miRNA.

The pre-miRNA or shRNA is exported to the cytoplasm via Exportin-5 (XPO5) with Ran-GTP (18–20). The oligonucleotide is further processed into a ~21-25nt miRNA/miRNA* duplex or siRNA duplex by a second RNase III enzyme, DICER (21). DICER requires the 2nt 3'-overhang on the pre-miRNA/shRNA to work efficiently at cutting off the loop from the other side, creating another 2nt 3'-overhang in the process. The dsRNA-binding protein TRBP aids the next process as it helps recruit DICER to Argonaute (22). One strand of the miRNA/miRNA* or siRNA duplex is preferentially selected and loaded into Argonaute, forming the RISC. Of the four Argonautes expressed in cells (AGO1-4), only AGO2 is competent to perform RNAi since it is the only one to retain functional Slicer activity (23). RNAi is the intended mechanism through which siRNA functions, as siRNAs are designed to be perfectly complementary to their mRNA targets.

Gene networks controlling specific miRNA biogenesis and activity

Analogous to any protein-coding gene, the transcripts that become miRNAs are subject to a wide range of basal and dynamic gene signaling networks to either maintain their levels in the cell, or to raise or lower expression in response to a stimulus. The number of factors demonstrating control on the biogenesis or activity of individual or subsets of miRNAs continues to grow as we come to understand miRNAs as vital cogs on the many gears of gene pathways.

One seemingly obvious route to gene regulation is at the transcriptional level. Indeed, as many miRNAs are expressed from Pol-II transcripts, they require promoters that are in turn controlled by various transcription factors, enhancers, and other elements. The first

demonstration of transcriptional control was that of let-7 miRNA in *C. elegans* under the regulation of a temporal regulatory element in its promoter (24). Several transcription factors have been found to regulate miRNA transcription. One of the first transcription factors found to control miRNA expression is c-Myc, which was shown to induce expression of the miR-17~92 cluster (25). Another example is p53, which positively regulates the transcription of miR-34a in response to DNA-damage (26–28). Since then, a plethora of other transcription factors have been found to control miRNAs as part of their regulatory networks.

Many of the effects on miRNA biogenesis have been found to occur post-transcriptionally, as well. The Microprocessor recognition and cleavage of pri-miRNAs seems to be particularly susceptible to regulation. Several transcription factors have been found to regulate Microprocessor efficiency through association with p68/p72 helicases (29). For example, p53 not only transcriptionally promotes miR-34a, but it also positively regulates the efficiency of Microprocessor cropping activity (30). In the same study, p53 was found to regulate a whole subset of miRNAs, including the miR-15a/16-1 cluster, miR-143, and miR-145. SMAD proteins, upon activation of BMP signaling, also positively regulate the process for several cardiac miRNAs like miR-21 (31). The SMAD proteins, in particular, recognize a sequence motif (CAGAC) on the pri-miRNA stem to define the subset of miRNAs affected (32). Recently, SRp20 has been found to bind a small motif (CNNC) 3' to the stem-loop of the pri-miRNA and aid in Microprocessor efficiency for at least the subset of pri-miRNAs with the motif (33).

Several 'loop-binding' proteins associate with the loop of the hairpin to affect both Microprocessor activity and DICER activity. One of the earliest miRNAs discovered, let-7, featured an antagonistic relationship with lin-28 (34). Lin-28 recognizes a loop sequence in let-7 and promotes degradation of let-7 through oligouridylation (35,36). Examples of positive loop-

binding regulators include KSRP, which can promote the efficiency of Microprocessor in response to DNA-damage. HnrnpA1 can promote the expression of specific miRNAs such as miR-18a, which is intriguing given that miR-18a is part of the much larger miR-17~92 cluster (37). Also, FXR1P, the protein product of *FXR1*, has been found to promote DICER-mediated maturation of miR-9 and miR-124 in neuronal cells (38).

No peripheral factors have been found so far to regulate pre-miRNA export to the cytoplasm via XPO5. However, given that Ran-GTP is an important cofactor for XPO5 function, perhaps a protein or gene network that would regulate or disrupt the Ran cycle could affect miRNA expression.

Once the mature miRNA is loaded into RISC, the activity of the miRNA can still be highly regulated. For example, AGO2 requires the association of Hsp90 to load miRNA and siRNA duplexes, and to recruit the complex to processing bodies (P-bodies) (39,40). This association can be altered in hypoxic conditions (41). Additionally, AGO2 itself can be ubiquitinated and degraded in scenarios such as T-cell activation (42).

RISC can also be guided to target certain transcripts in response to associated factors. FMRP, the protein product of *FMR1*, has been shown to associate with RISC and facilitate the targeting of certain FMRP targets (43). Oddly, in quiescent cells, AGO2 associated with FXR1P can actually promote translational *activation* (44).

miRNAs are relevant to disease

Upon their discovery in humans, miRNAs were immediately investigated for their relevance to diseases such as cancer. Indeed, in cancer, and in almost every disease studied, the loss or gain of miRNA expression has been shown to play at least a contributing role to the

pathology. One of the earliest demonstrations of miRNA aberration in cancer was with the miR-15a/16-1 cluster. The locus coding for these miRNAs was found to be deleted in the majority of B cell chronic lymphocytic leukemia (45,46). It was later found that the miR-15a/16-1 cluster was tumor suppressive in part due to its targeting of BCL2 to promote apoptosis (47). Another example in cancer is the miR-142 locus, found to be part of the t(8;17) translocation with *MYC* in B cell leukemia (BCL) (48). Some aberrations to miRNA expression in cancer have been found to not be the result of transcriptional regulation, but rather the result of post-transcriptional regulation. A survey of miRNA expression across several primary tumor samples revealed that mature miRNA species were often down-regulated in relation to their primary transcript counterparts, highlighting extensive regulation of miRNA levels at the Microprocessor step or later (49).

miR-142

Some of the earliest efforts to identify and characterize miRNAs in mammals focused on the hematopoietic system, and one survey identified miR-142 as one of the highest expressed miRNAs (50). Intriguingly, it is almost exclusively expressed in hematopoietic tissues (50), at least in part due to its dependence on PU.1 transcription factor (51). The miR-142 locus has been found to play a role in more cancers aside from BCL, as aberrant expression of miR-142 has been found in several other cancer contexts, such as adult T-cell leukemia (52), acute lymphoblastic leukemia (53,54), acute myeloid leukemia (55,56), and B-cell lymphomas (57–59).

This miRNA is an attractive subject to study in terms of its regulation due to the many roles it plays in cells that require either its dynamic or sustained expression levels. There are

several examples of miR-142 controlling hematopoietic cell differentiation. It mediates part of the miR-223-CEBP-β-LMO2 regulatory circuit controlling myeloid cell differentiation by attenuating cell proliferation (60). It controls the maturation of megakaryocytes through regulation of the actin cytoskeleton network (61). This miRNA also controls several responses by virtue of its dynamic expression. Its control of the actin cytoskeleton network extends to the regulation of phagocytosis (62,63), CD4+ T cell migration (64), and cancer cell migration and invasion (65,66). miR-142 potentiates the innate immune response to endotoxin in dendritic cells by suppressing IL-6 until the miRNA itself is suppressed by TLR-4 (67). It also helps to control the function of T_{REG} cells by maintaining a cyclic AMP (cAMP) gradient via targeting adenylyl cylase 9 mRNA, where it is suppressed in CD4+CD25+ T_{REG} cells and expressed in CD4+CD25- T cells (68). Perhaps the most dynamic expression of miR-142 is observed in the context of circadian rhythm, where it controls circadian rhythm genes like BMAL1 and is in turn regulated by the CLOCK gene (69–72).

There have been several indications that miR-142 is post-transcriptionally regulated. It was found that a ~20nt conserved sequence 3' of the hairpin is lost in its translocations with *MYC*, driving speculation that it may be required for recognition of the hairpin (73). In addition, it is notably one miRNA to be targeted by adenosine deaminases acting on RNA (ADARs), leading to A-to-I RNA editing of the pri-miRNA, which inhibits its processing by Drosha (74). MiR-142 displays 5'-end heterogeneity, which results in two distinct 'isomiRs' that can target different mRNAs (75). Identifying and understanding novel genes that may regulate miR-142 may shed light on these and other post-transcriptional mechanisms.

Utility of pooled screening to discover novel gene functions

Since the discovery that RNAi is functional in mammals (76), efforts were made to exploit the process to systematically change the expression of any desired gene in the genome in the context of a high-throughput screen (77). This was first demonstrated in a study looking for modulators of TRAIL-induced apoptosis (78). These siRNAs had to be transfected into cells, which limit their applicability to highly transfectable cells. A more widely applicable strategy was developed to get around this limitation by using shRNAs, first in retroviral systems (79), then in improved lentiviral systems (80). Genome-wide lentiviral shRNA libraries were developed shortly thereafter (81).

The advent of retroviral and lentiviral systems to stably introduce shRNA libraries allowed the first pooled screens to be performed. These screens have several advantages over arrayed libraries listed here below.

One advantage is that pooled screening vastly reduces variability in experimental conditions. Arrayed screens typically have to be executed in 96- or 384-well plates, which are subject to well-to-well variables like pipetting error and evaporative edge effects. In pooled screens, several million cells can be cultured in large tissue-culture plates or flasks, or even cultured in bioreactors, with each individual cell stably expressing an individual shRNA.

Another advantage is that pooled screening is often less expensive to execute than arrayed screening, and producing and maintaining pooled libraries are much easier. If 96-well plates were used, arrayed screening would require the use of scores or even hundreds of plates to reach the same throughput as one 15cm dish of cells in a pooled screen, cutting down on the need for almost all materials.

Finally, comprehensive pooled libraries are much easier to produce and maintain than arrayed libraries. Pooled libraries can be synthesized on high-content array chips and chemically cleaved to produce one pool of oligonucleotides to clone en masse in a vector of choice. Cloning procedures often take only 2 or 3 days, and pooled libraries can be quality controlled using deep-sequencing. Since pooled libraries have to be amplified in competent bacteria, the library can be easily renewed when stocks become low. Arrayed libraries practically require the use of robotics to produce, and are difficult to quality control and replenish. As such, much higher throughput libraries can be produced using a pooled approach, allowing for more library elements to be made against individual genes, ensuring the ability to target all genes with several potent shRNAs (82).

Given the ability to produce and use massive, genome-wide shRNA libraries with relative ease, one was used to assay the miRNA/RNAi activity of miR-142-3p, which is detailed in the next chapter.

Advent of CRISPR-Cas9 gene editing technology

The latest breakthrough in biotechnology to control gene expression comes in the form of the innate immune system of many bacterial strains: the clustered regularly interspaced short palindromic repeats (CRISPR) – CRISPR-associated 9 (Cas9) system, or CRISPR-Cas9 for short. As opposed to the mRNA targeting that is characteristic of miRNAs and siRNAs, the CRISPR-Cas9 system targets double-stranded DNA, with the Cas9 inducing double-stranded breaks (DSBs) and subsequent destruction of their targets, primarily viruses like bacteriophage (83). The system from *S. pyogenes* was simplified to a single Cas9 protein and a single-guide RNA (sgRNA) composed of a chimera of a CRISPR (crRNA) and tracr (tracrRNA) (84). This

system was shown to be functional in human cells, targeting specific genomic loci, and creating indels through error-prone non-homologous end-joining (NHEJ) (85,86).

Unlike miRNA-mediated repression or RNAi, a nuclease-active Cas9 can edit the genome and cause frameshift mutations, resulting in functional knockouts of genes. The orthogonal mechanism of CRISPR-Cas9 for genetic manipulation serves as an appropriate counterpoint to RNAi, and could be used as an effective tool to validate the observations from studies using RNAi. Importantly, this orthogonal mechanism is completely independent of any miRNA/RNAi activity, which makes it uniquely suited to validate observations from miRNA/RNAi studies that used RNAi to manipulate genes. One longstanding concern for using RNAi to study miRNA/RNAi is the phenomenon of 'recursive RNAi.' Recursive RNAi paradoxically leads to ineffective targeting of RNAi components like DICER and AGO2 when RNAi is used to potently inhibit these components; when a protein like AGO2 is suppressed, then its own low levels result in a loss of RNAi efficacy and its rescue of expression. Though not often observed, it may explain occasional negative results in using RNAi to target core components like DICER (87), and has been modeled to demonstrate that potent RNAi of core components would be relatively ineffective (88). As such, using the CRISPR-Cas9 system to study miRNA biogenesis and activity would circumvent this issue.

For reasons of using an orthogonal mechanism to confirm observations from an RNAi screen, and avoiding recursive RNAi, a CRISPR-Cas9 library is also used in the next chapter in the efforts to find novel modulators of miRNA biogenesis and activity.

Chapter 2 – Development and Execution of miRNA Sensor Screens

Introduction

The discovery of RNAi in animals (1), and more specifically mammals (76), came around the time the human genome was successfully mapped (89–91). This allowed for any known or predicted mRNA transcript could be quickly inhibited using RNAi. The fact that most of the over 20,000 protein-coding genes across the human genome were mapped allowed RNAi to be utilized at a genome-wide scale. High-throughput RNAi technologies have developed into several arrayed and pooled libraries with robust screening and analysis methods (77,92). A ~600,000-element lentiviral shRNA was developed, making use of 30 shRNAs per gene to guarantee several effective shRNAs for each gene, and to safeguard against off-targeting effects (82). The highly complex 'EXPAND' shRNA libraries have been successfully used in several screens to date (93–96).

In this study, we use an EXPAND genome-wide shRNA library to screen for factors governing miR-142-3p biogenesis and RNAi activity. Since the levels of miRNA cannot be directly measured in live cells, only their activity on a reporter construct can be examined and selected for. This reporter construct, called a 'miRNA sensor,' has been constructed and validated for miR-142-3p, demonstrating that the miRNA sensor is sensitive to miR-142-3p levels (97). We utilize a variant of this miRNA sensor, using perfect binding sites to potently suppress the expression of the reporter; this version of the sensor will allow us to assay modulators of miR-142-3p expression and RNAi activity.

As outlined in Chapter 1, much of the core machinery of miRNA biogenesis has already been determined. In addition, some successful high-throughput RNAi arrayed screens have been executed in worms (98,99), fly (87,100–102), and human cells (103) to identify some of these

genes. However, most of these screens focused either on synthetic or non-human miRNAs, and the one human screen was centered on miR-21 biogenesis and activity. There still lays ahead a comprehensive assessment of the various signaling pathways and more subtle components of the miRNA machinery that act as regulators of individual or specific subsets of miRNAs. As the several examples of post-transcriptional modulators of miRNA processing, summarized in Chapter 1, indicate, future studies in miRNA biology should focus not on all miRNAs in general, but on individual miRNAs or their families.

During the course of this study, the use of a CRISPR-Cas9 system to manipulate gene expression was published (85,86). As this happened after the execution of our genome-wide shRNA library screen, we took advantage of this technology to further validate the results of the screen using a focused sgRNA library in a secondary screen. The results of the two screens show known miRNA pathway genes, in addition to several unknown candidates, highlighting the effectiveness of using high-throughput screens to study miRNA biogenesis.

Results

A 'sensor' system to measure miRNA activity

In order to capture miRNA activity in live cells, we generated pSensor.miR142, which is a polycistronic reporter construct to express GFP and Sh ble, an antibiotic resistance-marker for Zeocin (hereafter referred to as 'ZeoR'), under the regulation of four perfect binding sites for miR-142-3p in the 3'UTR (Table 1). pSensor.miR142 was cloned into a lentiviral backbone and stably introduced into Raji B cells, a human Burkitt lymphoma cell line. Raji cells highly express miR-142-3p (104). To demonstrate that the cell lines established were sensitive to miRNA activity from miR-142-3p, the two clones established (Raji-miR142-clA and Raji-miR142-clB) were infected with an shRNA against *AGO2* (Figure 1). Rescue of GFP expression was observed compared to a negative control shRNA against *LacZ*. The perturbation of *AGO2* expression established that the cell lines are sensitive to changes in RISC activity.

Whole-genome shRNA screen to enrich for genes that affect the miRNA sensor

To find novel genes that can affect miR-142-3p activity, we utilized a highly-complex EXPAND (82) human genome-wide shRNA library composed of ~600,000 elements (93). This library was infected into the Raji-miR142-clA clone, and the infected pools of cells were separated into one of two screening arms: cells that were treated with 400ug/mL Zeocin, or cells that were cultured without drug (Figure 2). After 13 days, the two sets of cells had their genomic DNA extracted to amplify the shRNA libraries. As a quality control measure of the performance of the Zeocin selection during the screen, cell samples were measured by flow cytometry to measure enrichment of GFP in library-infected cells (Figure 3). Indeed, the levels of GFP in library-infected cells increased more with Zeocin selection arm compared to the untreated arm.

Both libraries were subjected to massively parallel sequencing and compared to one another. Two core components of the miRNA biogenesis pathway had multiple shRNAs highly enrich: *DICER1* and *AGO2* (Figure 4). A total of 456 genes had at least two shRNAs enrich above a 10-fold threshold, which is more than expected by random chance ($p < 0.01$).

Validation of hit genes from whole-genome shRNA screen

From the top enriched screen hits, several genes were chosen for post-screen validation. The 2 or 3 most enriched shRNAs for each gene were individually cloned into MP177 (Table 1) and infected into Raji-miR142-clB cells to measure rescue of GFP expression, an indication of sensor rescue from RNAi (Figure 5). Several shRNAs validated by this method, with many genes such as *AKAP8*, *MYEF2*, *SET* (data not shown), *EIF5*, *SNRPD1*, *ZCRB1*, *P4HA1*, *NUDT16*, *GJC1*, *RBM12B*, *HNRPA2B1*, *PRPF18*, *TRIM55*, *KIAA1524*, *ST8SIA4*, *PKP4*, *SFRP4*, *RGS5*, *NUTF2*, *PSME4*, *LRRK2,* and *TRUB1* with shRNAs that enrich GFP levels more than negative control shRNAs. The rescue in GFP expression using these shRNAs that also confer Zeocin resistance in the genome-wide shRNA screen suggests that the shRNAs are working specifically on sensor expression, consistent with affecting miRNA activity.

Development and validation of nuclease-active CRISPR-Cas9 system

We sought an orthogonal method for validating the observations of gene knockdown leading to sensor rescue, so we established a nuclease-active CRISPR-Cas9 system to use for shRNA screen validation. To test the activity of the nuclease-active CRISPR-Cas9 system, we cloned a human-codon-optimized *S. pyogenes* Cas9 sequence and a U6-promoter-driven cassette driving expression of mosaic crRNA-tracrRNA (85) (sgRNA) into the combined vector

pSgRNA.Cas9 (Table 1). 10 days after lentiviral transduction of pSgRNA.Cas9, with an sgRNA against GFP (using the 'T1' guide (85)) and a scrambled sgRNA, into GFP-positive 293T cells, the cells expressing the 'T1' sgRNA have massive reductions in GFP levels compared to cells with the scrambled sgRNA as measured by flow cytometry (Figure 6a). 82.9% of cells infected with the vector with GFP-targeting sgRNA show ablation of GFP expression, compared to only 7.13% in cells expressing the scrambled sgRNA. The nuclease-active CRISPR-Cas9 system we cloned works efficiently in cells.

Next, we sought to validate the system in Raji-miR142-clB cells against endogenous gene targets. The nuclease-active CRISPR-Cas9 system was separated into two vectors, pLibrary.Cas9 and pSgRNA.1 (Table 1). SgRNAs targeting *AGO2* and miR-142 were cloned into pSgRNA.1, and both this vector and pLibrary.Cas9 were transduced by lentivirus into the sensor cells. After 15 days of Zeocin selection, the infected cells were analyzed by flow cytometry to measure sensor activity by GFP (Figure 6b). Cells expressing the *AGO2* and miR-142 sgRNA vectors had on the order of 10 to 100-fold higher GFP expression. It is worth noting that the miR-142 sgRNAs appear to be more affective than the *AGO2* sgRNAs; miR-142 has no family members, and so loss of the miRNA leads to a complete rescue, but the targeting of AGO2 may lead to compensatory activity by other Argonaute proteins. Overall, the results indicate that not only does the nuclease-active CRISPR-Cas9 system work for endogenous genes, but also it can ablate genes that result in a miRNA sensor rescue.

We also established that most sgRNAs have sufficient efficacy on their cognate targets. A total of 57 sgRNAs were cloned into pSgRNA.1 to target the hairpin and context of miR-142, a ~325nt region, and transduced into pLibrary.Cas9-expressing Raji-miR142-clB cells. 11 days after infection, cells were analyzed by flow cytometry for GFP rescue (Figure 7). The most

potent sgRNAs target the stem of the hairpin directly, with 6 of the 9 sgRNAs in this region showing efficacy. No other targeting of regions outside the hairpin effectively worked to reduce miR-142 activity, suggesting that the indels produced by the nuclease-active Cas9 are mostly small, and that the only potent sites on the locus affecting miR-142 expression are in the hairpin, and not in the context. Nonetheless, most sgRNAs targeting the hairpin showed efficacy.

Confirmation of validated hit genes from shRNA screen using CRISPR-Cas9

We introduced a validated CRISPR-Cas9 system (pCas9.HP, Table 1) into Raji-miR142-clB cells. Next, 3 sgRNAs targeting each of most of the validated hits were cloned into pSgRNA.2 and transduced into the pCas9.HP-expressing Raji-miR142-clB cells. As expected, targeting of *AGO2* rescued GFP, demonstrating that in this experiment, the system works on the RNAi regulation of the sensor. However, none of the sgRNAs against any of the other hit genes from the screen rescued GFP above scrambled control (Figure 8).

The different mechanisms by which each system works may explain the discrepancy between shRNA and CRISPR-Cas9 experimental results. As the miR-142 sensor is sensitive to miRNA levels, it is reasonable to suspect that shRNAs that are highly present in the RISC complex, for one reason or another, flood the RNAi compartment. This leads to a lowering of the miR-142-loaded RISC complexes available, therefore leading to reduced RNAi on the sensor transcript and consequent rescue of Zeocin resistance and GFP expression. Due to this possibility, even if the shRNAs are effective at knocking down its target, it does not necessarily follow that this knockdown causes a reduction in miR-142 expression or activity. A rescue experiment for one of the top-enriched genes, *FAM172A*, was attempted, introducing cDNAs for all 3 isoforms of FAM172A (Figure 9). The shRNA-targeting of *FAM172A* rescues GFP activity

of the sensor, validating the gene. However, the cDNAs failed to rescue the effect of the shRNA knockdown. In this experiment, it is possible that the cDNAs fail to properly express the FAM172A proteins, though a Western blot of cells with transcript variant 1 does show overexpression (data not shown). It is also possible that the overexpressed proteins may not be properly modified post-translationally. Nonetheless, the most plausible explanation is that the shRNA activity on the miRNA sensor is not through *FAM172A* knockdown.

Design of focused nuclease-active Cas9 library

In light of the discrepancy between shRNA validation and CRISPR-Cas9 validation results, it is imperative to rescreen the top genes from the shRNA screen using the CRISPR-Cas9 format to rule out genes that enriched simply due to nonspecific shRNA activities on the miRNA machinery. Validation of the nuclease-active CRISPR-Cas9 system in Raji-miR142-clB cells allowed us to utilize the system for high-throughput pooled screening in the sensor cell line. The layout of library oligonucleotides is shown in Figure 10a. The library oligonucleotides ranged in size from 95nt to 103nt, with the variance in size due to differences in primer binding sites (PBS1 and PBS2), which allows for amplification of a sub-library of oligonucleotides. The NheI site and proximal U6 promoter sequence allow the oligonucleotides to be ligated into the AvrII site near the end of the U6 promoter on pSgRNA.1. The AarI site allows for scarless cloning of the spacer just upstream of the tracrRNA. Within the proximal U6 promoter region is a 6nt randomized barcode sequence; this tags each library element with a unique barcode, and allows for barcode-sgRNA combinations to be analyzed at the end of the screen to measure enrichment/depletion trends for library elements at a single-cell resolution. The barcodes can also be used to create pseudoreplicates to increase statistical power.

Spacers were designed to target the most 5' coding exons of every Refseq annotated transcript for the gene set. To check for off-targeting, a candidate spacer was checked across all coding exons across the genome; the schematic for the off-target checking protocol is shown in Figure 10b. If there was no perfect alignment in the 11nt 3' seed region of the spacer, the spacer was not considered as off-target, in accordance to published off-targeting observations (86). If a perfect alignment did exist, then the 9nt 5' region comprising the rest of the spacer must have had 7 or more mismatches to be considered as not off-target (84). Since editing at any off-target site with more than 4 mismatches looks to be negligible (105), we find this off-targeting check protocol to be sufficiently conservative. From the list of spacers not thought to off-target, 12 spacers were selected for each Refseq transcript. About 12,500 spacers were designed for the entire library. The genes that had at least two shRNAs enrich from the whole-genome screen were selected to design targeted spacers (457 genes). Spacers targeting these genes made up roughly half of the library, with the remaining half targeting RNA binding proteins, genes known to affect the miRNA pathway to some degree, and a robust set of negative control (non-targeting) spacers. In total, 893 genes were targeted by these spacers to make the sgRNA library.

Focused nuclease-active CRISPR-Cas9 screen for genes that affect the miRNA sensor

The spacers were cloned into pSgRNA.1, creating the 12,500-element sgRNA library. Sequence analysis of a small sample of 24 colonies from the cloning indicates that the library has ~71% perfect sequences (data not shown). To express Cas9 in the sensor cells, Raji-miR-142-clB cells were transduced with pLibrary.Cas9-packaged lentivirus. Positive cells were selected gradually using low levels of puromycin (0.2-0.5ug/mL). These pLibrary.Cas9-expressing Raji-miR142-clB cells were then used in the screen (Figure 11). Cells were transduced with the

sgRNA library. After 6 days of Hygromycin B selection (100ug/mL) followed by 2 days recovery, the cells were split into two separate screening arms in technical duplicate; one arm treated with Zeocin for 30 days, and one untreated control. As a demonstration that Zeocin-selected cells were enriched in a sensor-specific manner, GFP was measured by FACS at days 3, 15, 24, and 30 (Figure 12). Over time, particularly in days 24 and 30, a tail and small population of GFP-high cells was observed.

The cells that survived the screen were subjected to the same library-amplification and massively parallel sequencing procedure as for the shRNA screen. As a control for library representation in the sample cells, a sample of cells taken just prior to the start of Zeocin selection ("Time 0") was processed and subjected to massively parallel sequencing (Figure 13). The library was fairly even, with ~85% of the sgRNAs within one order of magnitude of representation. No sgRNAs were particularly over-represented, and fairly few sgRNAs (~15%) were under-represented.

The sequencing results from sgRNA screen were more striking compared to the shRNA screen; *AGO2* was the top screen hit, with *DICER1* ranked 11th (Figure 14). *MIR142,* the miR-142 gene, also ranked 3rd. Other genes described as part of the miRNA/RNAi pathway also ranked highly, such as *FXR1*, *SKIV2L*, and *CNOT7*. Overall, the results of the screen indicate that the CRISPR/Cas9 screen was effective at selecting for genes important to the miRNA/RNAi pathway.

Validation of hit genes from CRISPR-Cas9 screen

To validate some of the strongest-enriched genes from the screen, the three top-enriched sgRNAs for selected hit genes from the main Fisher's exact analysis, in addition to the Mann-

Whitney analysis (Table 2), were compiled (Table 3) and individually re-cloned into pSgRNA.2. These sgRNAs were transduced into Cas9-expressing Raji-miR-142-clB cells, and then diluted to a low infection frequency (~4-8%) with non-sgRNA-transduced cells. By subjecting these lowly-transduced cell populations to Zeocin selection, many of strongest-enriched sgRNAs validated in terms of Zeocin-resistance, including those targeting *AGO2*, *DICER1*, *MIR142*, *PSME4*, *FXR1*, and *SKIV2L* (Figure 15). *FAM208A* targeting trends toward enrichment with Zeocin treatment, and also shows enrichment in other experiments (data not shown), demonstrating that the editing of this gene likely leads to Zeocin resistance in Raji-miR142-clB cells.

Next, we wanted to observe increased sensor mRNA expression for sgRNAs that conferred a Zeocin resistance phenotype. From edited cells treated with Zeocin for 30 days, total RNA was extracted and subjected to reverse-transcription followed by quantitative polymerase chain reaction (RT-qPCR) (Figure 16). Only editing of *AGO2*, *DICER1*, and *MIR142* resulted in increased expression of the sensor by more than 3-fold, while sensor levels for other edited cells were around the levels found in the negative control. However, since these RNA samples came from Zeocin-treated cells, we may be seeing a background signal of nonspecific sensor rescue in the negative control. To address this, we used the dCas9-KRAB system, since the activity of the system across all cells appears to be more uniform (data not shown), allowing us to not use Zeocin to amplify the targeted cell population. Total RNA from cells targeted with the dCas9-KRAB system was subjected to reverse transcription followed by droplet digital PCR (RT-ddPCR), measuring sensor mRNA (Figure 17). The changes are modest, but significant. The changes may be modest due to several reasons. One reason is that Zeocin was not used to select for cells with more exacerbated phenotypes. In addition, finding the best sgRNA to target the

promoter of a gene is more difficult than simply editing an open reading frame; a gene may have several transcription start sites that cannot be simultaneously targeted by one sgRNA. This is likely the case with *MIR142*, which has many possible transcription start sites. Nonetheless, the targeting of *PSME4*, *FXR1*, and *FAM208A*, as well as *AGO2* and *MIR142*, rescues sensor mRNA expression, suggesting the potential for these genes to affect sensor expression, possibly through miRNA/RNAi mechanisms.

To directly examine regulation of miR-142 expression in Zeocin-resistant cells, total RNA from edited cells selected for 30 days with Zeocin was subjected to RT-ddPCR, using a probe specific to miR-142-3p. The Zeocin-selected cells transduced with a *MIR142* sgRNA had miR-142-3p completely ablated (Figure 18a). Since the nuclease-active CRISPR-Cas9 system has the unique ability to target non-coding RNAs like miRNAs (compared to shRNAs), this result confirms that miRNA sensor expression is sensitive to the levels of miR-142-3p. Transduction with *AGO2* sgRNA reduced levels of miR-142-3p by about 50%. Interestingly, two of the three sgRNAs against *FXR1* (FXR1_b and FXR1_c) reduced levels of miR-142-3p modestly but significantly, about 25-30%. Unfortunately, this observation of reduction in miR-142-3p levels could not be reproduced in a repeat of the experiment using similar conditions (Figure 18b). This repeated experiment also shows that editing of *FAM208A* also fails to change levels of miR-142-3p. Overall, the only hits that we have shown to reliably affect the expression levels of miR-142-3p are core miRNA components.

Discussion

We have executed the first pooled high-throughput screens in human cells examining the factors modulating miRNA expression and RNAi activity. This is the first genome-wide RNAi screen examining miR-142, and the CRISPR-Cas9 screen is the first for the activity of any miRNA. These screens will hopefully be a springboard to further understanding the genetic networks affecting miR-142, and a springboard to promoting similar screens for other miRNAs highly relevant to various pathologies.

Performing a pooled shRNA screen for miRNA/RNAi factors is somewhat more difficult than a similar screen for other phenotypes like drug resistance, apoptosis, or proliferation, to name a few. Several other phenotypes like these can be assayed directly, or the screening pressure is directly relevant to the biology of interest. However, miRNA biology cannot be directly assayed in live cells; we rely on reporter constructs to act as a proxy for miRNA mediated repression and RNAi. As such, screening procedures do not directly measure miRNA or siRNA activity. Instead, they measure phenotypes likes fluorescence, surface marker expression, or antibiotic resistance. These phenotypes have their own factors that affect them directly, which increases the false-positive rate of hit genes toward the indirect biology of miRNAs. Nevertheless, the shRNA screen still yielded some of the core components of the miRNA biogenesis pathway (Figure 4), giving promise to the hit list produced yielding some novel data relevant to miR-142.

The validation of many of the top enriched genes from the shRNA screen by GFP (Figure 5) shows that while there are several theoretical ways to produce false-positive genes in the screen related to Zeocin resistance or cell proliferation, many of the top enriched shRNAs still rescue the activity of the sensor, and make the possibility of these genes being true hits ever

greater. However, the utilization of the CRISPR-Cas9 system to confirm these genes as being true hits (Figure 8) revealed a nefarious possibility that the screen primarily selected for shRNAs that commandeered the RISC pool and pushed out native miRNAs like miR-142, resulting in a non-specific rescue of the sensor. While it is possible that some the sgRNAs designed for the CRISPR-Cas9 may not have worked at knocking down their targets, the targeting of *AGO2* and *MIR142* (Figure 6b, Figure 7, Figure 8) clearly works to rescue the sensor activity. Given the rather robust rescue of sensor expression when genes like *PKP4* and *NUTF2* are targeted by shRNAs, we expect to see some perceivable sensor rescue with the CRISPR-Cas9 system, even if the genes are not being targeted in the ideal locations of their open reading frames. It is my opinion that the enriched shRNAs for these genes are artifacts that non-specifically reduce the activity of native miRNAs. This observation highlights the need to use the CRISPR-Cas9 system to assay miRNA/RNAi biology.

The performance of the CRISPR-Cas9 screen appears to be successful, given the GFP rescue observed during the screen (Figure 12), the even representation across the library (Figure 13), the enrichment of several of the core components of the miRNA machinery (Figure 14), and high validation rate by Zeocin of most of the top enriched genes (Figure 15).

Despite the great performance of the CRISPR-Cas9 screen, and the good validation rate of the top enriched genes, the validation of these genes by rescue of the sensor appears to be slight, at best (Figure 16, Figure 17). In addition, only manipulation of the core genes *AGO2* and *MIR142* appears to have any robust impact on the expression levels of miR-142 itself (Figure 18). This calls into question the ability of any of the novel candidate genes to affect the levels of miR-142. However, we cannot yet rule out any impacts the editing of these genes have on the activity of miR-142. Though the rescues of sensor expression appear to be slight, Zeocin

resistance may only require this slight rescue, as we have seen that Zeocin resistance can occur at much lower levels of reporter expression than, for example, the expression of GFP required to observe by flow cytometry (unpublished observations). Nevertheless, the fact that the top genes enriched in the screen validate by Zeocin resistance demonstrates that these hits are not a result of random noise, even if the circumstances by which they confer Zeocin resistance may be under contention.

Among the candidate genes that enriched outside of *AGO2*, *MIR142*, and *DICER1*, the #2 gene is *GFP*. This may have occurred due to GFP having to separate from ZeoR by 2A-element self-cleavage. Even though T2A cleavage is efficient (106), this efficiency may vary between cell types. If we presume that fused ZeoR-T2A-GFP fails to work on Zeocin, then edits to *GFP* that result in GFP truncations may rescue the activity of the fused proteins, resulting in higher functional ZeoR concentration, thus explaining the enrichment of most *GFP* sgRNAs in the screen. Western blots showing GFP fusions with the ZeoR protein, in addition to testing the functionality of ZeoR protein fused with various GFP ORF truncations should adequately test for this hypothesis.

PSME4, the #3 ranked gene, may have enriched in the screen due to its functional knockout conferring Zeocin resistance independent of the sensor. Its ortholog in yeast is *BLM10*, the knockout of which sensitizes yeast to bleomycin and phleomycin antibiotics, of which Zeocin is a family member (107–109). Additionally, *PSME4* is crucial for genomic stability in the face of DNA-damaging ionizing radiation (110). So why would ablation of this gene cause a *resistance* phenotype in the screen? This may be due to cells without PSME4 protein only exhibiting sensitivity when glutamine is limiting, yet appear somewhat resistant when glutamine is supplemented (111). Cell culture conditions during the screen would provide glutamine

replenishment every 3 days, explaining the observed resistance phenotype. Testing *PSME4*-edited cells vs. wild-type cells for resistance and growth under Zeocin treatment, with differing glutamine concentrations, should indicate whether the edited cells are only resistant with excess glutamine present.

Another strongly enriched gene, *SKIV2L*, may have also acted more directly on the sensor through its protein's role as part of the SKI complex that helps in degrading sliced RNA products of RNAi (112). However, I have been unable to observe any rescue of miR-142 sensor mRNA expression in quantitative PCR experiments. Future experiments testing a role for SKIV2L in the DNA damage response or Zeocin kinetics may reveal novel functions for the protein.

FAM208A represents one of the more interesting screen hits. It has been relatively uncharacterized except in two studies to date: one showing that a knockout of the mouse *Fam208a* rescued an epigenetically-repressed *GFP* transgene (113), and one showing that the human protein, renamed TASOR, is part of an epigenetic silencing complex dubbed the human silencing hub (HUSH) complex (114). One hypothesis for how *FAM208A* came up as a hit in my screen is that the sensor was at least partially repressed directly or indirectly by the HUSH complex, and that editing *FAM208A* released this repression. *FAM208A* editing does not appear to change miR-142-3p levels, though it does seem to slightly rescue sensor mRNA levels, making this model plausible. Alternatively, the sensor-expressing cells may have a second copy of the sensor integrated in the genome and epigenetically repressed by the HUSH complex. In this model, the editing of *FAM208A* leads to an activation of the second copy of the sensor. Fluorescence in-situ hybridization would reveal the sensor copy number in these cells.

Another hypothesis we explored for why *FAM208A* enriched in the screen involves its paralog, *FAM208B* (also known as *C10orf18*). TargetScan (115) predictions for miRNA targets show *FAM208B* is strongly predicted to be targeted by miR-142-3p. I hypothesized that editing of *FAM208A*, with subsequent loss of protein, triggered a feedback that resulted in greatly increased expression of *FAM208B* mRNA, which could have sponged away miR-142-3p from the sensor enough for a slight sensor rescue. However, *FAM208B* expression levels in *FAM208A*-edited cells do not appear to be changed, nor does *FAM208B* appear to be rescued in *MIR142*-edited cells (data not shown), leaving a possible epigenetic role on the sensor transgene itself the most likely reason for its enrichment in the screen.

FXR1 is another one of the more interesting candidates from the screen, as there is already some precedent for the role of FXR1P in the biogenesis of at least some miRNAs. It is part of a family of three closely related proteins comprising the Fragile X mental retardation gene family (FMRP, FXR1P, *FXR2* product FXR2P). It was shown that FXR1P associates with DICER and promote its processing efficiency of pre-miR-9 and pre-miR-124 (38). The Drosophila dFXR, a version of FMRP, has been shown to associate with RISC to promote mRNA degradation (116,117). Indeed, all three human proteins can associate with ~22nt small RNAs (118). Fragile X proteins also contribute to some of the same genetic networks as miR-142, making the potential for an interaction between the genes more plausible; they contribute to the regulation of the actin cytoskeleton through RAC1 (119) and circadian rhythm (120,121). Nonetheless, I have not established FXR1P as contributing to the expression of miR-142 (Figure 18). However, I have not ruled out the possibility that FXR1P may be affecting miR-142-3p activity independent of miR-142 expression. Luciferase assays using 3'UTRs for validated miR-142 targets may indicate if general miR-142 activity is affected in *FXR1*-edited cells.

It is also possible that *FXR1* could have enriched in the screens through a DNA-damage-dependent mechanism. FMRP was recently established as a factor in the DNA-damage response to genomic stress and single-strand DNA breaks (122), so one plausible explanation is that FXR1P acts in a similar manner on the double-strand DNA breaks induced by Zeocin. Preliminary Western blots examining the phosphorylation of H2A.X, an indicator for activation of the DNA-damage response, suggest this is not the case, but still remains a current focus.

Methods

Lentiviral vectors

Several vectors were used in this study, and fully annotated sequences are available upon request. The parent vector for all vectors used is the pSicoR lentiviral vector (Jacks Lab, MIT). Schematics for the relevant cassettes are detailed in Table 1.

Vector pSensor.miR142 expresses a reporter cassette for miR-142-3p activity. It contains a CMV promoter driving expression of the *Sh ble* gene for resistance to Zeocin, a T2A ribosomal skipping peptide, and *EGFP*. In the 3'UTR of the cassette are four perfect binding sites for miR-142-3p, separated by four nucleotides each, cloned using *BsrGI* and *EcoRI* sites.

MP177 was used for the human whole-genome shRNA library. The library was cloned as previously described (82,93).

Several Cas9-expressing vectors were used in this study. The vector used in the library screen is pLibrary.Cas9, which has a short EF1α promoter driving a puromycin resistance gene (Puro), T2A peptide, mCherry, a second T2A peptide, and a human codon optimized nuclease-active Cas9 from *S. pyogenes*, derived from the AddGene vector #41815, and mutated at one base to ablate an AarI site. One vector used in post-screen validation assays is pCas9.Blast, which is similar to pLibrary.Cas9, except it has a Blasticidin resistance marker in place of Puro, and blue fluorescent protein (BFP) in place of mCherry. A second vector used in the post-screen validation assays is pCas9.HP, which contains an SFFV promoter driving expression of nuclease-active Cas9, a T2A peptide, and BFP. pCas9.HP was a kind gift from Haopeng Wang and Art Weiss at UCSF. A third vector, pDCas9.KRAB, contains a nuclease-dead Cas9 (dCas9) fused to BFP and a KRAB domain. pDCas9.KRAB was a kind gift from Luke Gilbert and Jonathan Weissman at UCSF.

Several single-guide RNA (sgRNA) expressing vectors were used in this study. The vector used in the library screen is pSgRNA.1, which has a U6 promoter driving the expression of the sgRNA, with the sequence of the tracrRNA portion of the sgRNA derived elsewhere (84). pSgRNA.1 also has a reporter, with short EF1α promoter driving expression of a Hygromycin B resistance gene, a T2A peptide, and a tailless mouse CD4. The U6 promoter is derived from MP177 (93), with one base mutated to introduce an AvrII site. The sgRNA cassette contained a ~1kb 'stuffer' sequence in place of the CRISPR sequence to facilitate cloning of the sgRNA library and individual sgRNAs upon digestion with AarI (see *Generation of focused CRISPR-Cas9 library and vectors for post-screen validation*). pSgRNA.2 is similar to pSgRNA.1, except it uses a full-length EF1α promoter, an mCherry fluorescent marker, and a modified tracrRNA (123).

Lentivirus production

Lentivirus production was performed by the UCSF Viracore <http://viracore.ucsf.edu> using standard transfection procedures.

Cell culture

Raji B cells were cultured in RPMI-1640 media supplemented with 10% FBS, glutamine, and penicillin/streptomycin.

Generation and validation of Raji miRNA sensor cell lines

Two miR-142 sensor cell lines, Raji-miR142-clA and Raji-miR142-clB, were established by packaging pSensor.miR142 into lentivirus and transducing into Raji cells. Single-cell sorts

were made to generate clones. The cell lines were validated by lentiviral transduction of an shRNA against *AGO2*, and observing an increase in GFP after 7 days by flow cytometry.

Genome-wide shRNA screen in Raji miRNA sensor cells

The genome-wide shRNA library, cloned into MP177, was transduced into Raji-miR142-clA cells at ~44-fold coverage. Cells were selected 3 days post-transduction with 0.5µg/mL for 4 days. 7 days post-transduction, cells were either left not treated, or treated with 400ug/mL Zeocin. Each sample was passaged every 2-3 days in Zeocin-containing media, or drug-free media, in 4 15cm plates each with volumes between 160-200mL. Samples were measured for GFP expression by flow cytometry after each passage. 13 days into the Zeocin-treatment, ~50 million cells from each sample were lysed. Phenol-chloroform extraction was used to purify the DNA. The incorporated libraries for each sample were amplified with Phusion polymerase using primers that incorporated Illumina 5' and 3' adaptors, and the libraries were deep-sequenced for 50 nucleotides on an Illumina flow cell using the sequencing primer used before (82).

Validation of genes from shRNA library

Selected shRNAs were individually recloned in array format and packaged in lentivirus. Raji-miR142-clB cells were transduced, selected for shRNA expression by puromycin, and measured for GFP expression 14 days post-transduction.

Validation of CRISPR-Cas9 system

A version of pLibrary.Cas9 with an upstream U6 cassette from pSgRNA.1 and a long EF1α promoter was cloned with spacer sequence T1GFP or NT for a negative control.

Constructs were packaged in lentivirus and transduced into 293T cells that highly express GFP. The cells were selected with puromycin to enrich for transduced cells. Cells were measured for GFP expression 9 days post-transduction. To validate the system's function in Raji-miR142-clB cells, pLibrary.Cas9 was transduced by lentivirus into cells and selected for with puromycin. pSgRNA.1 with sgRNAs targeting *AGO2* or *MIR142* were then transduced into the cells, and after 3 days were selected with 100ug/mL Hygromycin B for 6 days. 11 days post-transduction, the cells were either treated with 200ug/mL Zeocin for 15 days or untreated. Cells were measured for GFP by flow cytometry after the 15-day treatment.

Tiling miR-142 locus with CRISPR-Cas9 system

All possible sgRNAs targeting miR-142 ~120bp upstream of the hairpin to ~200bp downstream of the hairpin (57 sgRNAs) were cloned into pSgRNA.1, packaged in lentivirus, transduced into pLibrary.Cas9-expressing Raji-miR142-clB cells in a 96-well plate, and selected with Hygromycin B. 11 days post-transduction, cells were measured for GFP expression by flow cytometry.

Validation of shRNA hit genes using CRISPR-Cas9 system

3 sgRNAs per gene were individually cloned into pSgRNA.2, packaged in lentivirus, and transduced into pCas9.HP-expressing Raji-miR-142-clB cells. Cells were selected with puromycin at 3 days post-transduction. Cells were measured for GFP expression by flow cytometry 22 days post-transduction.

FAM172A rescue

Three different ORFs (without UTRs to make them immune to shRNA targeting) for *FAM172A* were cloned into lentiviral vectors (MP633, MP678, and MP679, sequences available upon request), and transduced into Raji-miR142-clB cells. Cells were selected with Hygromycin B for 11 days at 5 days post-transduction. 4 shRNAs against *FAM172A* (#2851, #2868, #3880, #4059, sequences available upon request), 1 shRNA against *AGO2*, and 1 shRNA against *LacZ* were then transduced in the cDNA-expressing cells. Cells were measured for GFP expression by flow cytometry 5 days after shRNA transduction.

Generation of focused CRISPR-Cas9 library

A ~12,500 element sgRNA library was designed and ordered for synthesis on a CustomArray 12K chip. Lyophilized oligonucleotides were received, amplified with Phusion polymerase, digested with AvrII and AarI, PAGE-purified, and cloned into AvrII/AarI-cut pSgRNA.1. 4.5ng of the library was ligated to ~100ng of the vector with T4 DNA ligase, transformed into ElectroMAX competent cells. The cells were recovered for 1 hour in 1mL SOC media and plated on large ~500cm^2 LB-agar plates with 100ug/mL ampicillin. The plates were incubated overnight, and all colonies were scraped from the agar surface. An estimated 3 million colonies were isolated for a theoretical 243-fold coverage of the library. The colonies were maxiprepped en masse using Nucleobond Plasmid Maxi Plus Kit (Macherey Nagel). Sample colonies were sequenced, and the library was determined to have a 71% perfect-sequence rate.

Focused sgRNA screen in Raji miRNA sensor cells

pLibrary.Cas9 was transduced by lentivirus into cells and selected for with puromycin. pSgRNA.1 with the sgRNA library was then transduced into the cells at an estimated ~200-fold coverage. 3 days post-transduction, cells were selected with 100ug/mL Hygromycin B for 6 days. 11 days post-transduction, the cells were divided into four replicates, and two of each were either treated with 200ug/mL Zeocin for 15 days or untreated. The selection pressure was increased to 400ug/mL Zeocin for another 15 days, for a total of 30 days Zeocin treatment. At least 3000-fold coverage was maintained for each replicate sample at all times. Cells were measured for GFP by flow cytometry after each passage. An excess amount of cells (at least 2-fold in excess of 3000-fold coverage) were saved just prior to each passage to be collected and frozen down for genomic DNA. Frozen cell samples from the day Zeocin-selection started (day 0) and the end of the screen for all samples (day 30), were processed using the Qiagen Blood and Cell Culture DNA Maxi Kit. DNA was digested overnight with PstI-HF to liberate the sgRNA library from the genomic DNA, and the genomic DNA was run on a 1% agarose gel. Bands were cut in the expected range for the liberated library bands, and DNA was gel extracted. Each library sample was amplified using one of the 7 sets of indexed Illumina adaptor primers (Table 4). Samples were deep-sequenced for 50 nucleotides using an Illumina HiSeq 2000 platform, using the sequencing and Truseq index primers also listed in Table 4. Reads were segregated by index and downloaded.

Deep-sequencing analysis

For the shRNA screen, the sequencing reads were aligned to the shRNA library index using Novoalign, and assigned to their cognate genes using Python scripts. Genes were ranked

by number of shRNAs enriched in Zeocin treated cells, followed by the magnitude of the fold-enrichment.

For the sgRNA screen, the sequencing reads were aligned to the sgRNA library index using Novoalign, and assigned to their cognate genes using Python scripts. Fold-changes for sgRNAs between Zeocin-treated cells after 30 days and untreated cells at the same timepoint was determined using DESeq (124). Genes were ranked off of unadjusted P-value determined by Fisher's exact test for sgRNAs above or below the 95th percentile of enrichment. An alternative analysis used to identify some of the genes selected for validation was a Mann-Whitney rank sum test, where enriched genes were ranked by unadjusted P-value.

Zeocin validation of top enriched genes from sgRNA screen

Genes and sgRNAs selected for validation are listed in Table 3. The selected sgRNAs were re-ordered and cloned in arrayed format into pSgRNA.2, and then packaged in lentivirus. The sgRNAs were transduced in pCas9.Blast-expressing Raji-miR142-clB cells in 96-well plates, and briefly selected with low puromycin (0.1ug/mL) to help maintain some infection, but infection rate was intentionally kept low (~5-7% infection rate). Plates were expanded to 4 replicates, and 10 days post-transduction, two replicates each were either treated with 200ug/mL Zeocin, or untreated. Selection was increased to 400ug/mL after 18 days. Plates were measured for infection rate by mCherry fluorescence on a BD FACSarray 96-well flow cytometer at days 31 and 32 of Zeocin treatment. Afterwards, cell pellets were flash frozen. The pellets were later lysed with TRIzol, and RNA was extracted using the Qiagen miRNeasy Mini Kit.

A repeat of this experiment was performed using similar parameters with pCas9.HP-expressing Raji-miR142-clB cells. Afterwards, cell pellets were flash frozen. The pellets were later lysed with TRIzol, and RNA was extracted using the Qiagen miRNeasy Mini Kit.

dCas9-KRAB knockdown of selected genes

Selected dCas9-KRAB compatible sgRNAs were cloned into pSgRNA.2 and packaged into lentivirus. These sgRNAs were transduced into pDCas9.KRAB-expressing Raji-miR142-clB cells, and selected with puromycin. 14 days post-transduction, cell pellets were flash frozen. The pellets were later lysed with TRIzol, and RNA was extracted using the Qiagen miRNeasy Mini Kit.

qPCR and ddPCR

For droplet digital PCR (ddPCR) experiments measuring miR-142-3p expression, ~1-10ng of RNA sample was subjected to reverse transcription using miR-142-3p and HRPT1 RT primers. Reverse transcription was performed using the Multiscribe RT Kit, following manufacturer's protocol. The cDNA was then subjected to ddPCR on the Bio-Rad QX100 using using 20X FAM-labeled miR-142-3p and VIC-labeled RNU48 commercial Taqman assays from Applied Biosystems, following the Bio-Rad protocol for droplet generation, PCR, and droplet reading on the QX100.

For ddPCR experiments measuring sensor mRNA expression, ~300-500ng of RNA sample was subjected to reverse transcription using oligodT(20) primer and Superscript III. The cDNA was then diluted 1:4 and subjected to ddPCR on the Bio-Rad QX100 using using 20X FAM-labeled custom GFP probe and primer set (5' primer: ACGACGGCAACTACAAGACC,

3’ primer: GTCCTCCTTGAAGTCGATGC, Probe: /56AM/CTGGTGAACCGCATCGAGCTGA/3IABkFQ/) and VIC-labeled HPRT1 commercial Taqman assays from Applied Biosystems, following the Bio-Rad protocol for droplet generation, PCR, and droplet reading on the QX100.

For qPCR experiments measuring sensor mRNA expression, ~60-90ng of RNA was subjected to RT-qPCR using the Taqman RNA-to-Ct kit, following manufacturer’s specifications. The custom FAM-labeled GFP primer and probe set was used to measure the sensor, and the VIC-labeled HPRT1 commercial kit was used as control.

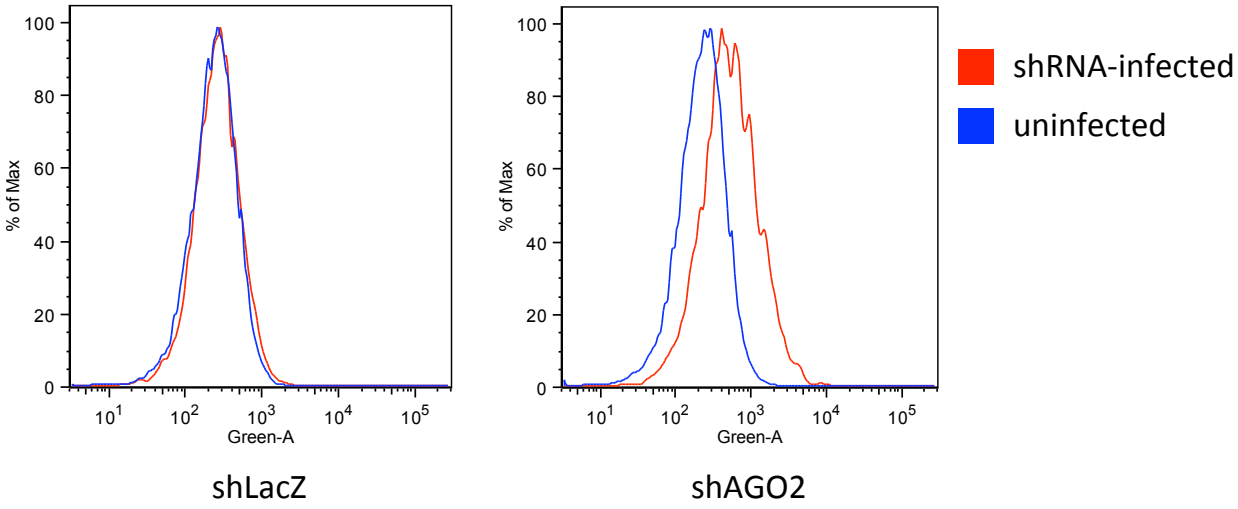


Figure 1. *AGO2* shRNA-dependent shift of miR-142 sensor expression.

shRNAs targeting either *LacZ* or *AGO2* were infected in Raji cells expressing pSensor.miR142.

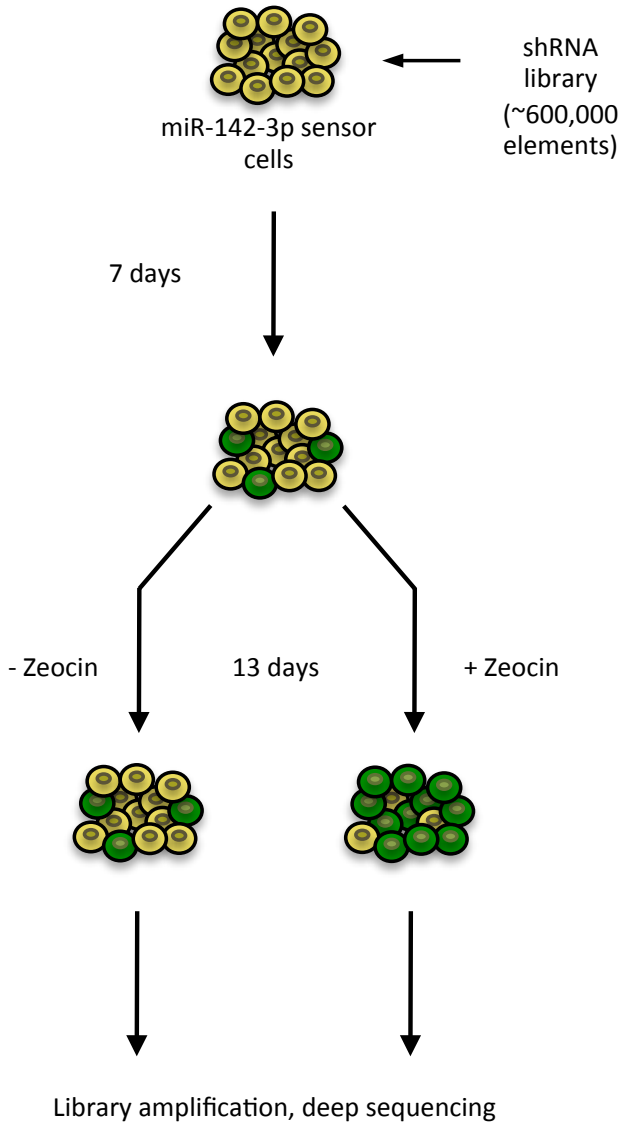


Figure 2. Schematic of genome-wide shRNA screen in sensor cells.

Raji-miR142-clA cells were infected with an EXPAND whole-genome shRNA library with a complexity of ~600,000 shRNAs. After puro selection and recovery, the library-infected cells were either treated with 400ug/mL Zeocin, or not treated. Cells were under culture for 13 days during Zeocin treatment, and then frozen at -80C for subsequent PCR amplification of the library from the genomic DNA and deep-sequencing.

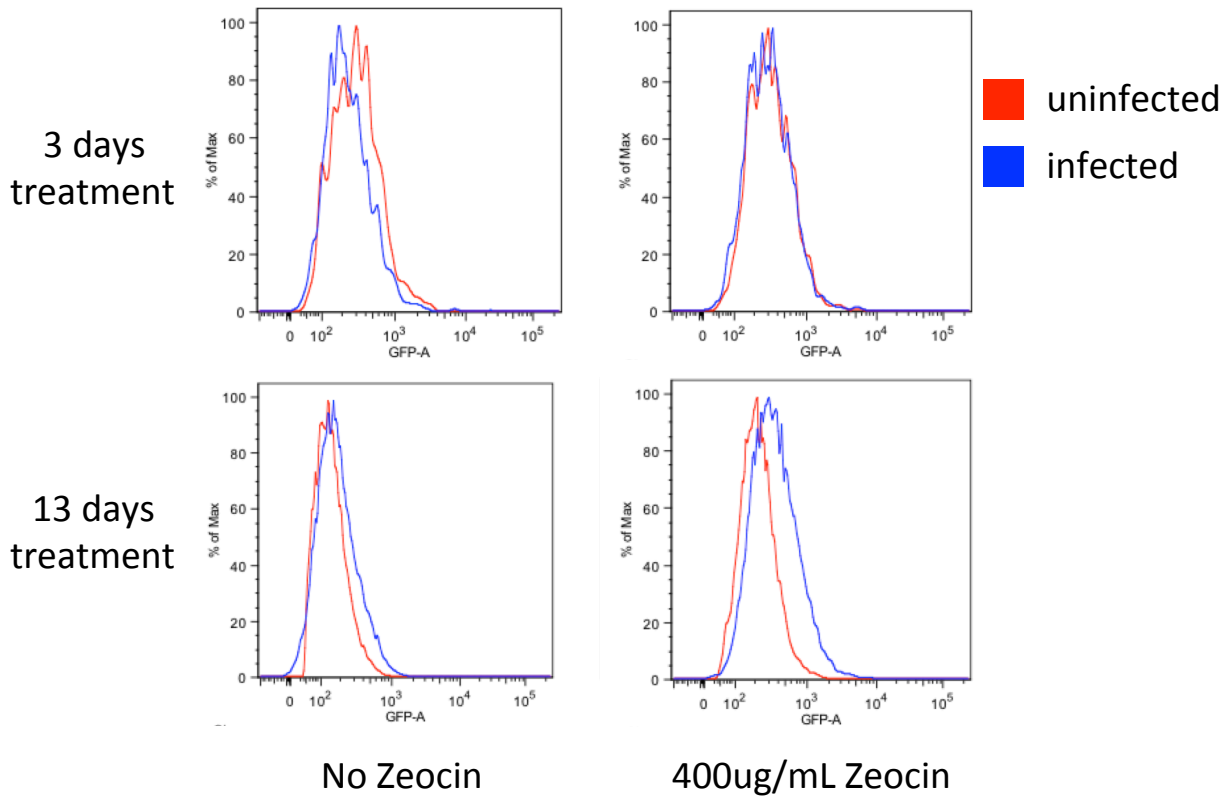


Figure 3. Performance of genome-wide shRNA screen.

Flow cytometry analysis of GFP levels during the screen. Note the increase in GFP levels across the Zeocin-selected population between day 3 and day 13.

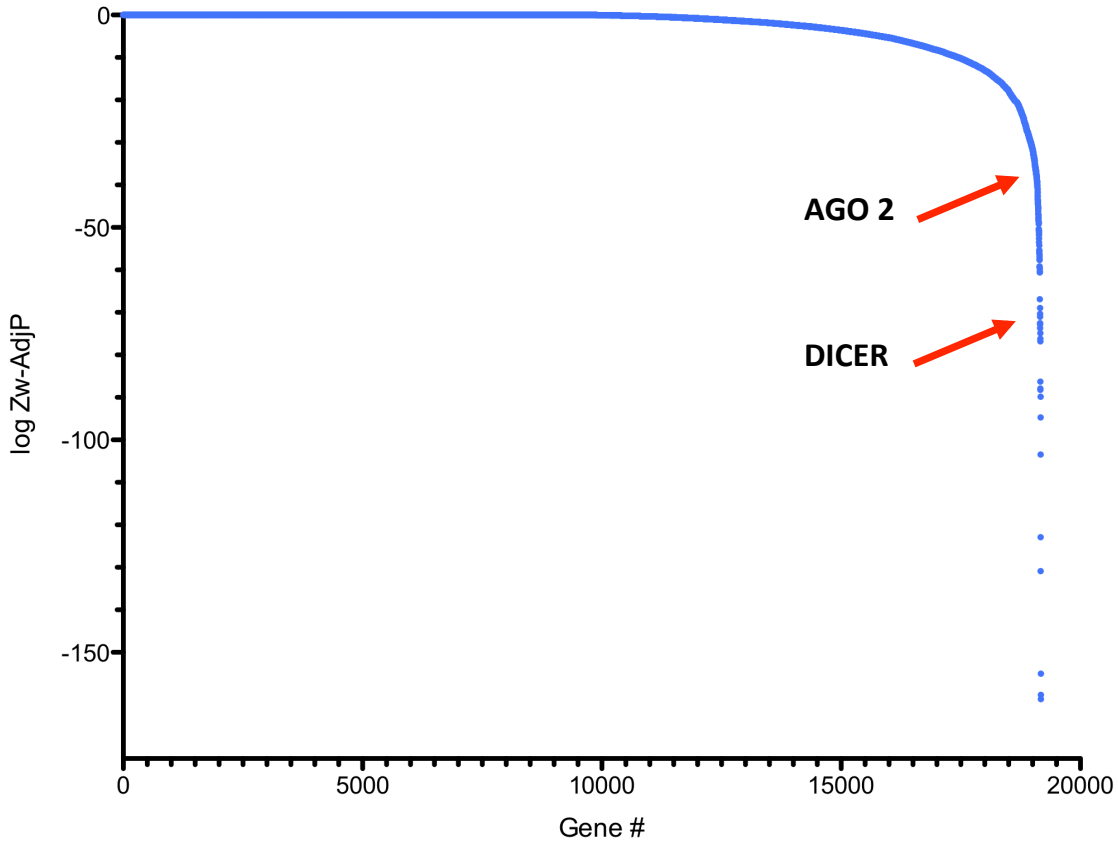


Figure 4. Analysis of top genes associated with enriched shRNAs.

Deep-sequencing results of the library samples were analyzed by the weighted Z-method (125) and ranked according to an adjusted P-value.

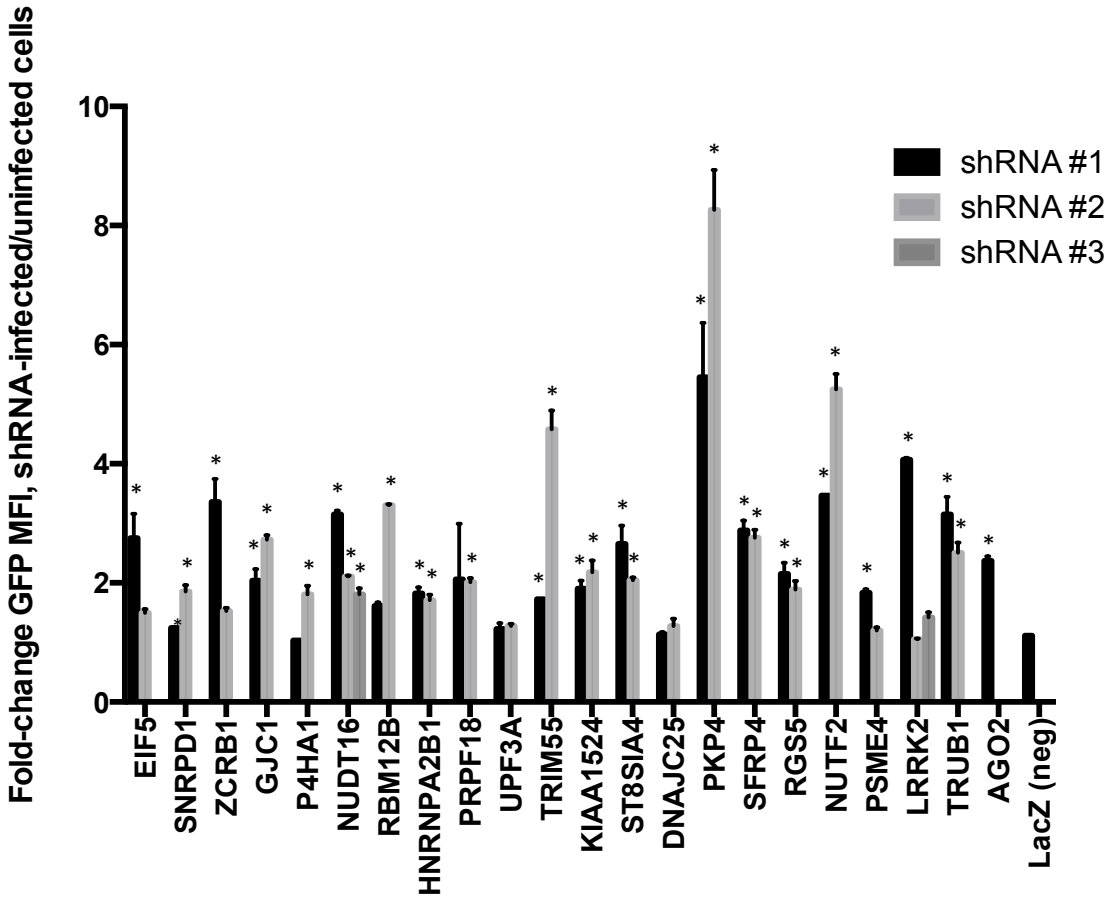


Figure 5. GFP validation of enriched shRNAs in sensor cells.

Raji-miR142-clB cells were infected with enriched shRNAs against a variety of hit genes. 14 days post-infection, the cells were analyzed by flow cytometry for GFP. GFP enrichments are calculated as fold-enrichments of GFP geometric mean fluorescence intensity in shRNA-infected cells vs. uninfected cells in the same wells. Error bars are standard deviation of two independent experiments. MFI = geometric mean fluorescence intensity. * = $p<0.05$ Fisher's Least Significant Difference (LSD).

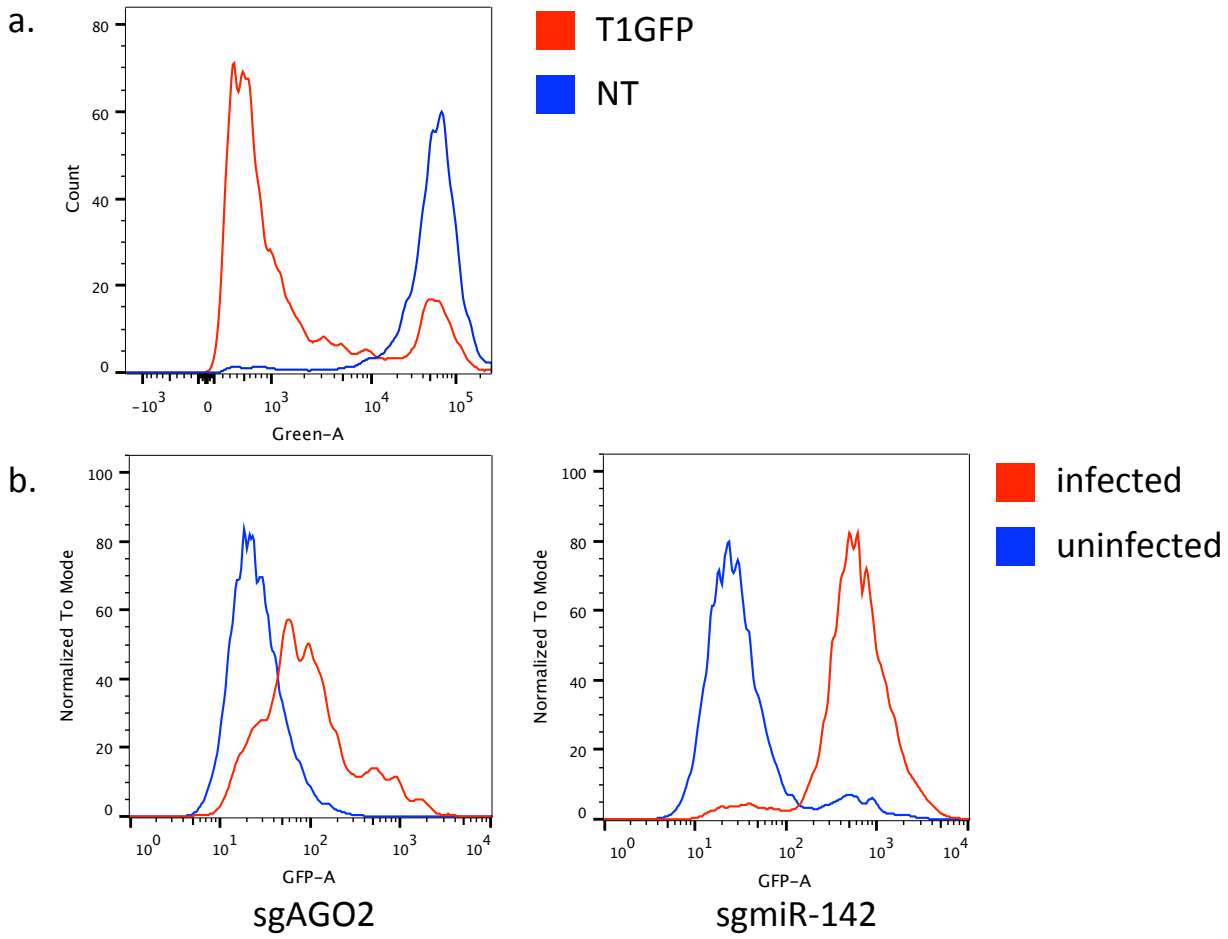


Figure 6. *AGO2* and *MIR142* sgRNA-dependent shift of miR-142 sensor expression.

(a) A combined vector expressing both nuclease-active Cas9 and sgRNAs targeting (T1GFP) or not targeting (NT) *GFP* was infected into GFP+ 293T cells, demonstrating the function of Cas9. **(b)** Using pLibrary.Cas9 in combination with pSgRNA.1 with sgRNAs against *AGO2* and *MIR142* to demonstrate both the CRISPR-Cas9 system functionality in Raji cells expressing pSensor.miR142 and the sensor functionality at specifically measuring miR-142-3p activity.

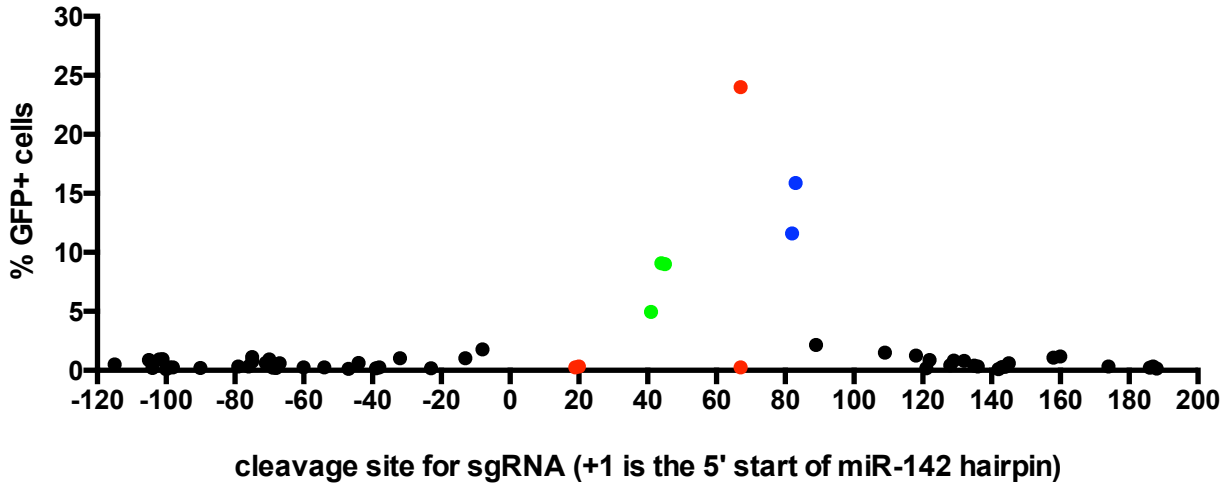


Figure 7. Efficacy of CRISPR-Cas9 system on miR-142 locus.

All possible sgRNAs targeting the stem-loop and surrounding context of *MIR142* on Chromosome 17 were infected in pLibrary.Cas9-expressing Raji-miR142-clB cells, then analyzed by flow cytometry for GFP rescue. Points representing the sgRNAs are shown at the predicted Cas9 nuclease sites. Points in blue directly target the basal portion of the hairpin stem; points in red directly target either miR-142-5p or miR-142-3p sequence; points in green target the loop of the hairpin. Y-axis is the % of cells infected with the sgRNAs that reached into the GFP+ gate, indicating functional editing of *MIR142*.

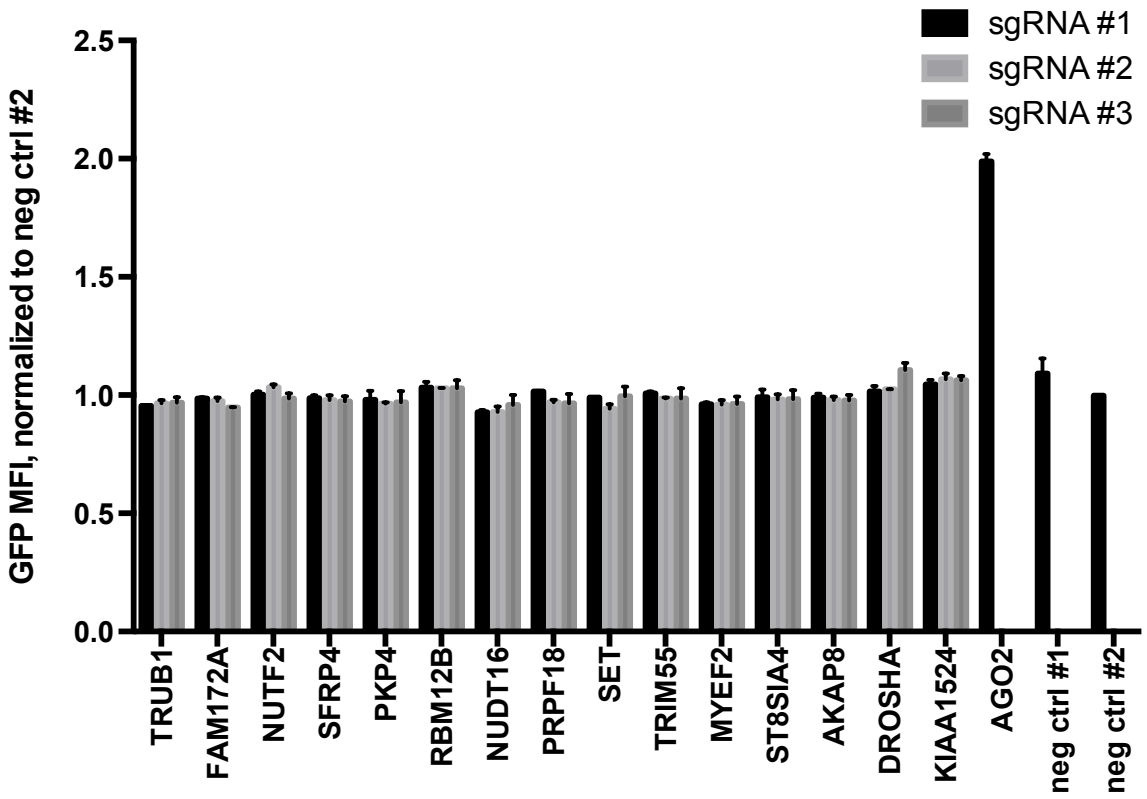


Figure 8. SgRNAs targeting hit genes from shRNA screen do not recapitulate GFP phenotypes of shRNAs. GFP expression of pCas9.HP-expressing Raji-miR142-clB cells when targeted by sgRNAs against the specified genes. Error bars are standard deviation across two biological replicates. MFI = geometric mean fluorescence intensity.

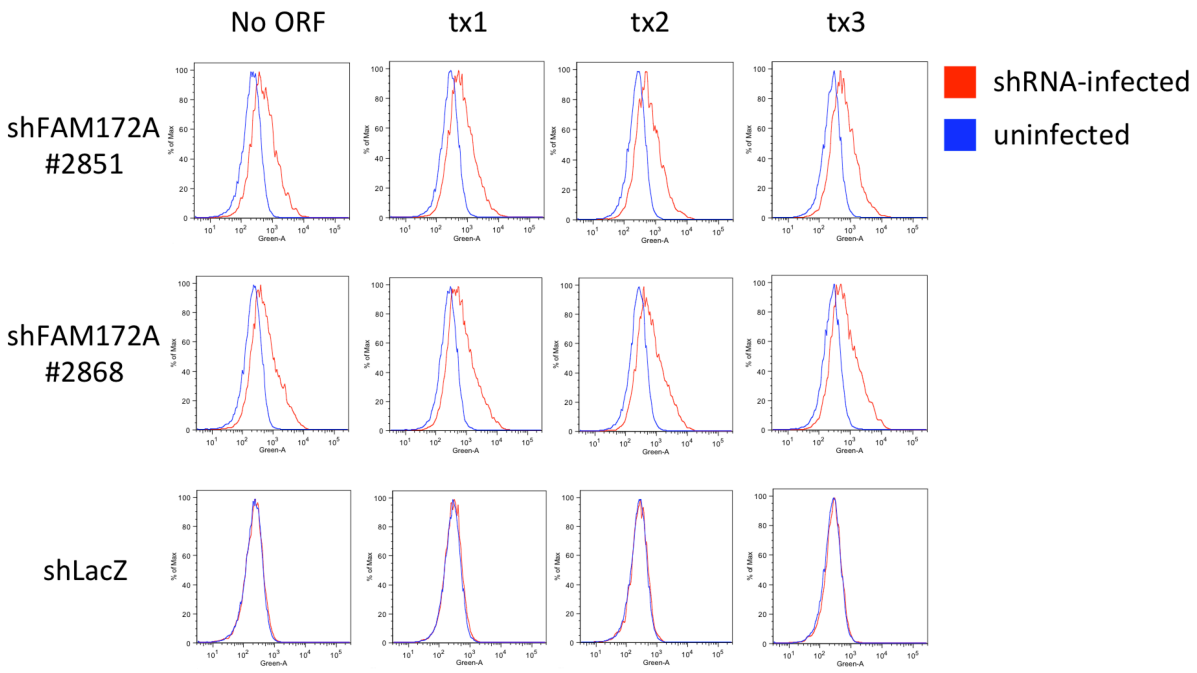


Figure 9. Rescue experiment for miR-142-3p activity after *FAM172A* knockdown.

cDNAs for three predicted isoforms of FAM172A were transduced by lentivirus to Raji-miR142-clB cells. The cDNAs lacked the target sites for shRNAs against *FAM172A*, making them resistanct to knockdown. Afterwards, 2 shRNAs against *FAM172A* and 1 shRNA against *LacZ* were transduced into cells. Flow cytometry analysis of GFP expression was performed 5 days post-infection of shRNAs.

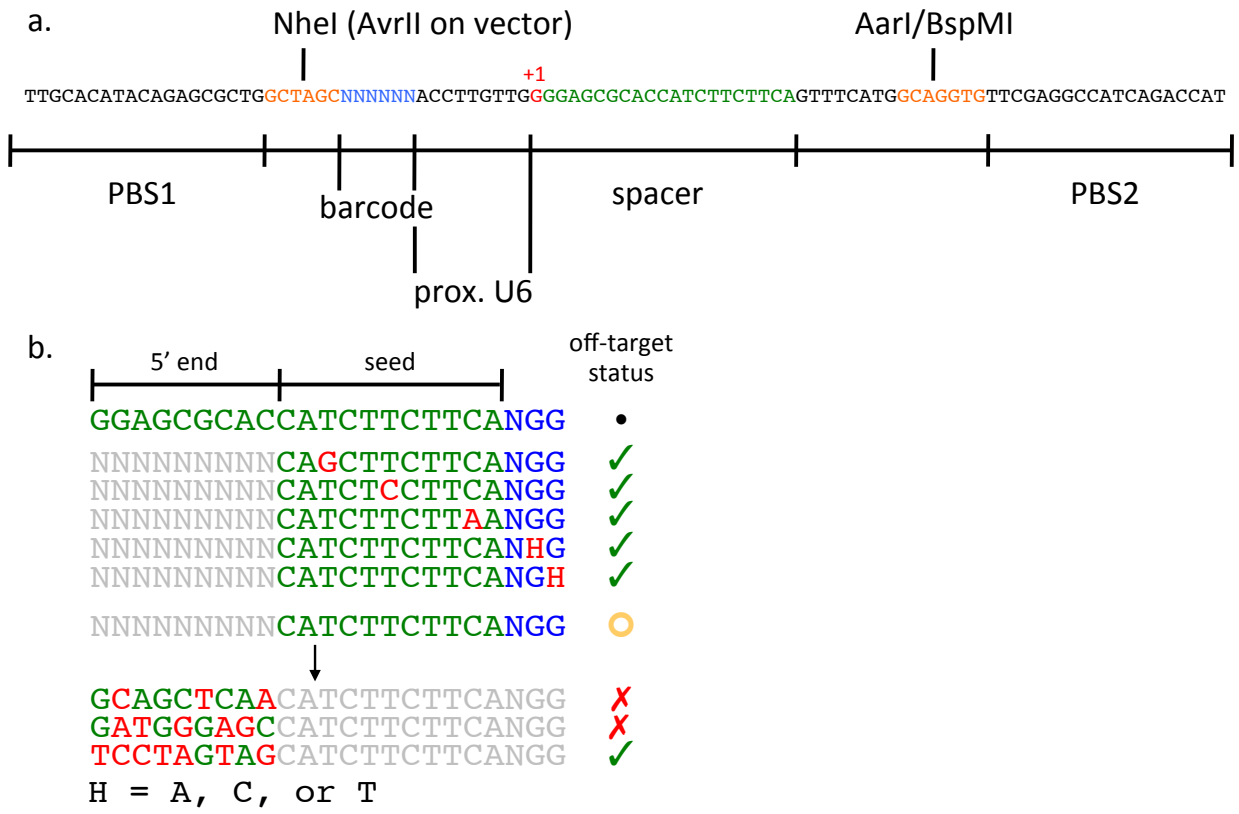


Figure 10. Cloning strategy and design for nuclease-active CRISPR-Cas9 sgRNA library.

(a) Schematic for DNA oligonucleotides used to clone sgRNA library into pSgRNA.1. Library oligonucleotides were PCR amplified from the primer binding sites (PBS1 and PBS2), then digested with AarI and NheI (which leaves the same sticky end as the AvrII site on pSgRNA.1) and ligated to the vector. The 6nt barcode lies on the end of the U6 promoter. **(b)** Design of 20nt protospacers. Sites targeting the coding regions of Refseq transcripts, as determined by the presence of the PAM (NGG) are checked against off-targeting against all other potential sites. If the 11nt seed region adjacent to the PAM is not a perfect match, the site is not considered an off-target. If the seed is perfect, and there are less than 7 mismatches in the remaining 9nt region, then the site is considered an off-target and the protospacer is rejected. Of all protospacers that pass QC, the 12 earliest protospacers on the ORFs of each transcript are selected for the library.

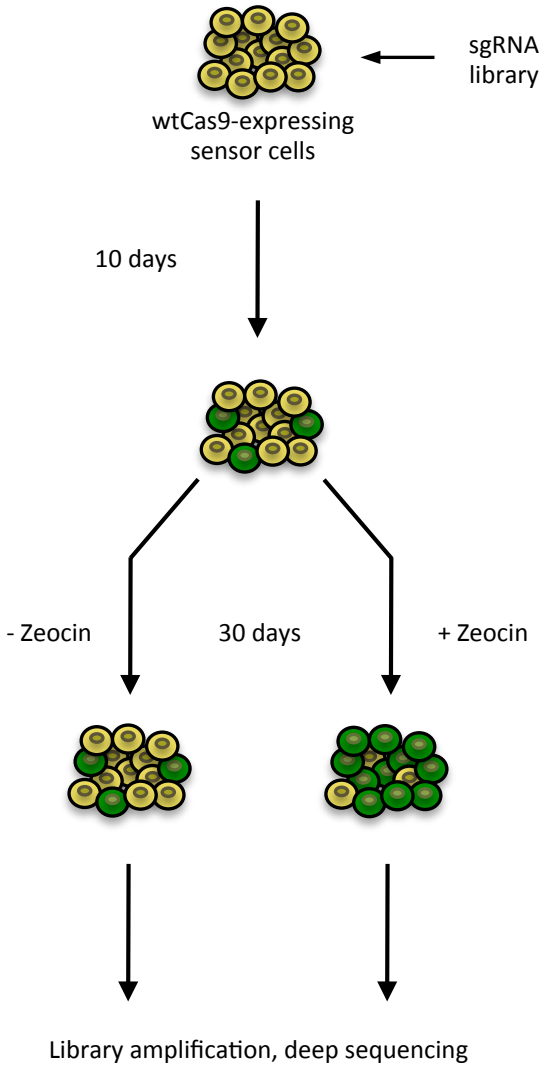


Figure 11. Schematic of nuclease-active CRISPR-Cas9 sgRNA screen.

pLibrary.Cas9-expressing Raji-miR142-clB cells were transduced with a library with a complexity of ~12,500 sgRNAs targeting a focused set of 893 genes. After Hygro selection, the library-infected cells were treated with 200ug/mL Zeocin for 15 days, followed by 400ug/mL Zeocin for another 15 days, or not treated. Cells were under culture for a total of 30 days during Zeocin treatment, and then frozen at -80C for subsequent PCR amplification of the library from the genomic DNA and deep-sequencing.

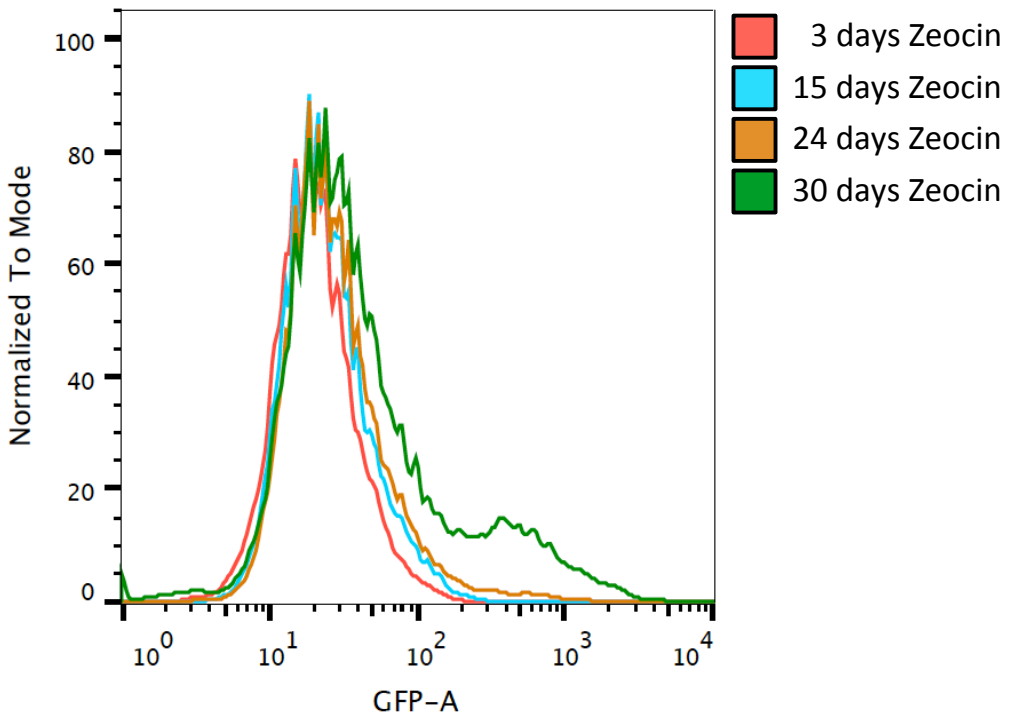


Figure 12. Performance of nuclease-active CRISPR-Cas9 sgRNA screen.

Samples from the Zeocin-selected cells during the sgRNA screen were analyzed by flow cytometry for GFP expression as a measure of sensor rescue.

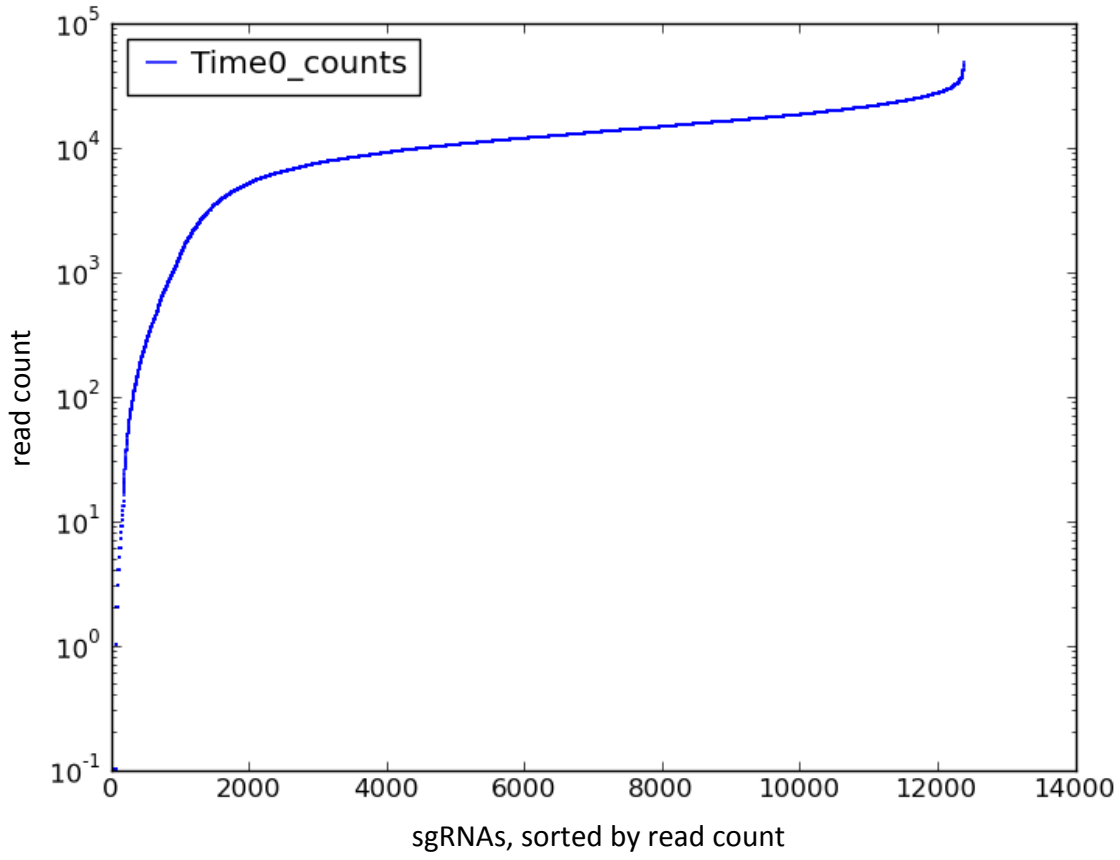


Figure 13. Representation of sgRNAs across pooled library.

Perfectly-aligned read counts for the "Time 0" sample transduced with the pooled sgRNA library. The X-axis represents each sgRNA in the library, sorted by read count.

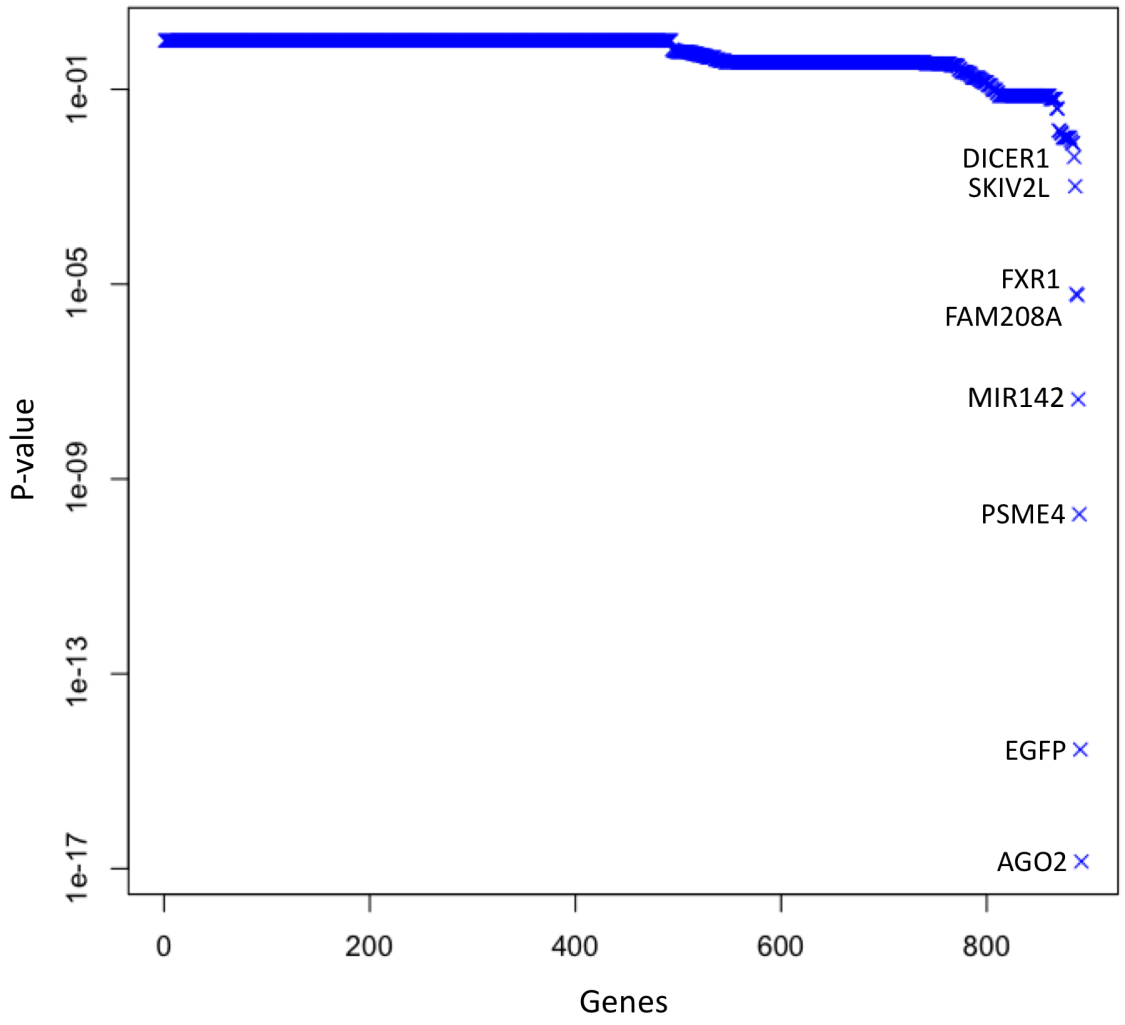


Figure 14. Analysis of top genes associated with enriched sgRNAs.

Massively parallel sequencing results of the library samples were analyzed by Fisher's exact test and ranked according to unadjusted P-value.

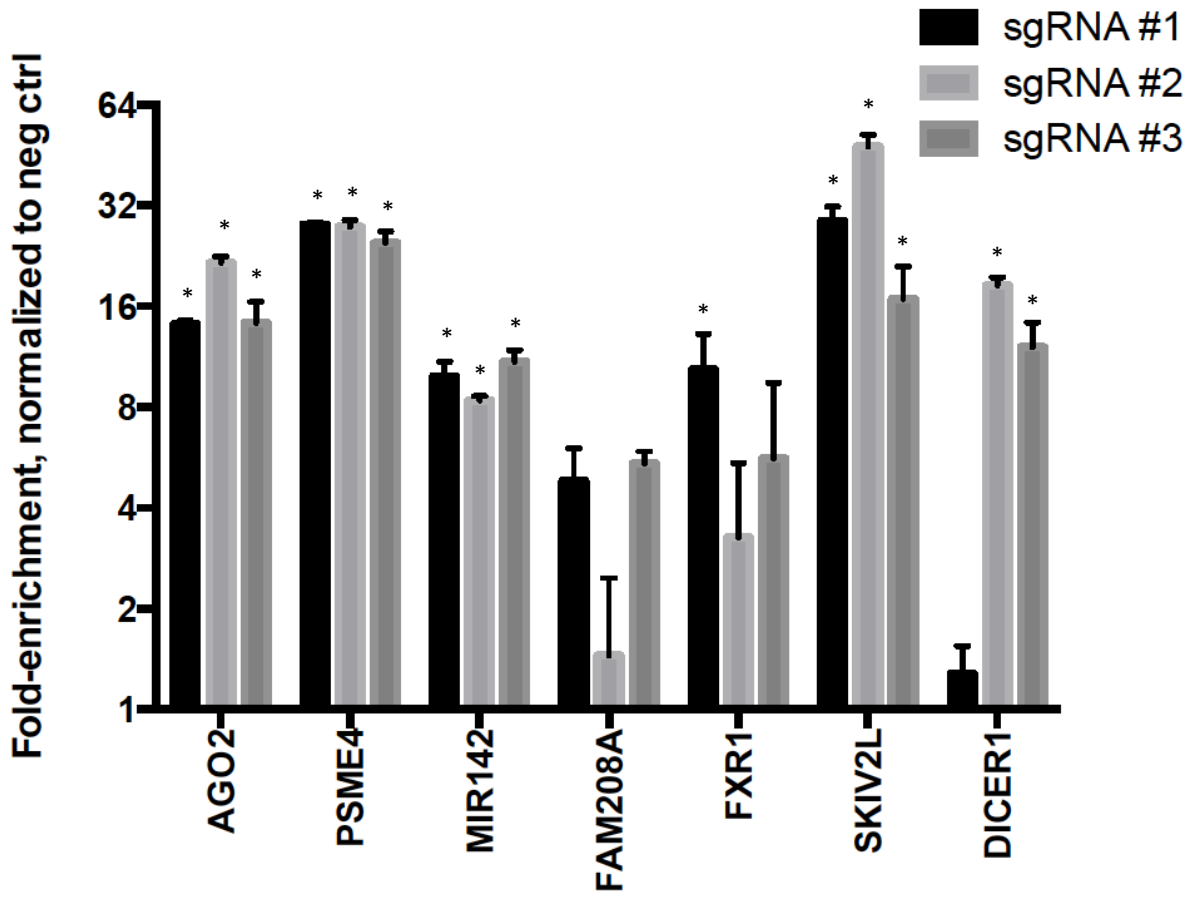


Figure 15. Validation of CRISPR-Cas9 sgRNA screen hits by Zeocin.

pCas9.Blast-expressing Raji-miR142-clB cells were infected with sgRNAs cloned in pSgRNA.2 at a low level of infection (less than 10%) and selected with Zeocin over a course of 31 days. Fold-enrichment is the proportion of mCherry+ cells from Zeocin-treated plates to untreated plates. Error bars are standard error of the mean for two biological replicates. Error bars are standard error of the mean of two biological replicates. * = $p<0.05$ Fisher's LSD.

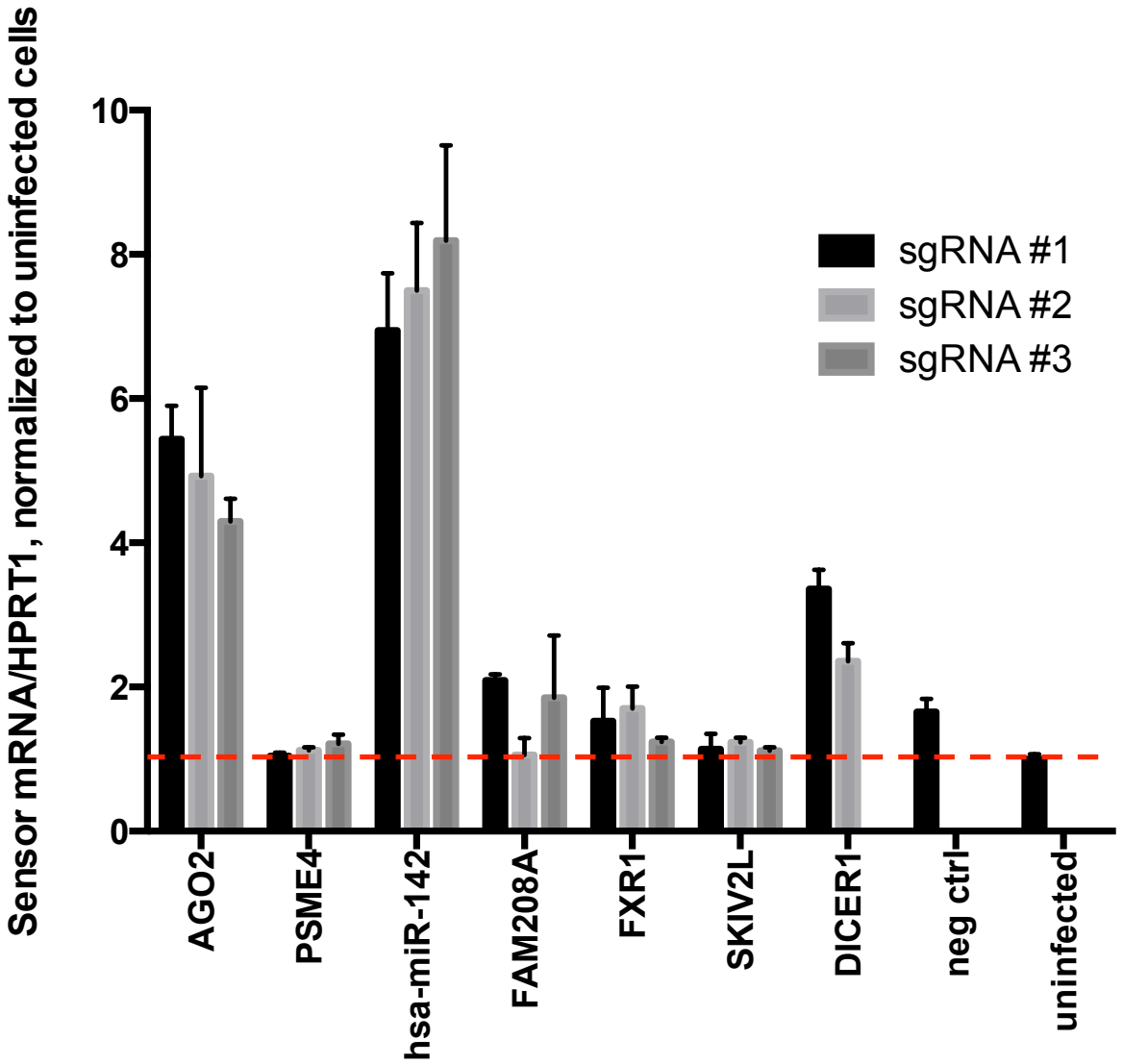


Figure 16. Changes in sensor mRNA in sensor cells with sgRNAs against screen hit genes. ddPCR to measure miR-142 sensor mRNA in pCas9.HP-expressing, sgRNA-expressing Raji-miR142-clB cells with 30 days Zeocin treatment. Sensor mRNA (using a GFP primer and probe set) levels normalized to *HPRT1* are compared to levels in uninfected cells, which are set at 1 and marked with a red dashed line. Error bars are standard error of the mean of three biological replicates.

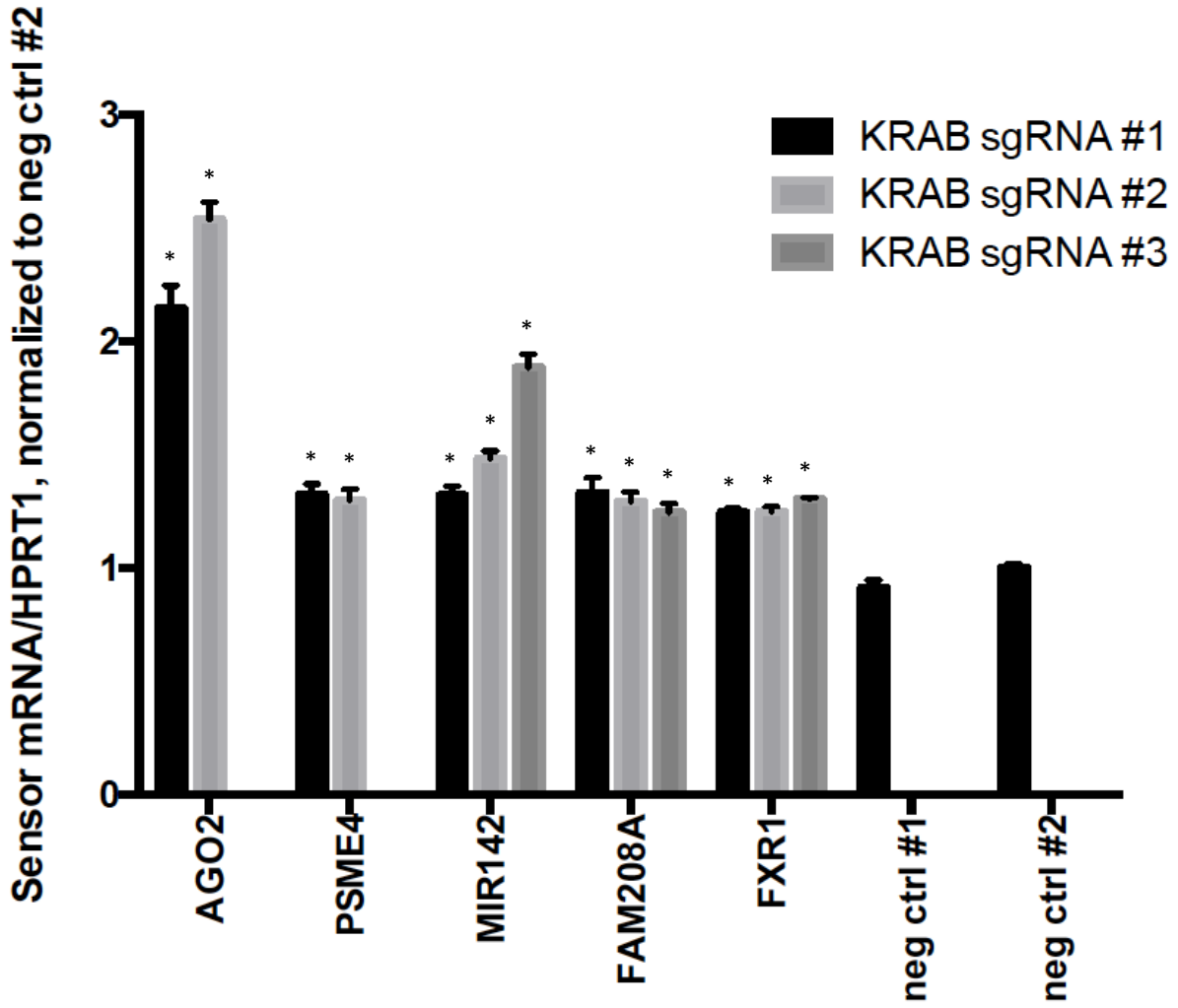


Figure 17. Changes in sensor mRNA in sensor cells with dCas9-KRAB sgRNAs against screen hit genes. Rescue of GFP expression in pDCas9.KRAB-expressing, sgRNA-expressing Raji-miR142-clB cells. Sensor mRNA (using a GFP primer and probe set) levels normalized to *HPRT1* are compared to levels in negative-control sgRNA-infected cells. Error bars are standard deviation of two biological replicates. * = $p<0.05$ Fisher's LSD.

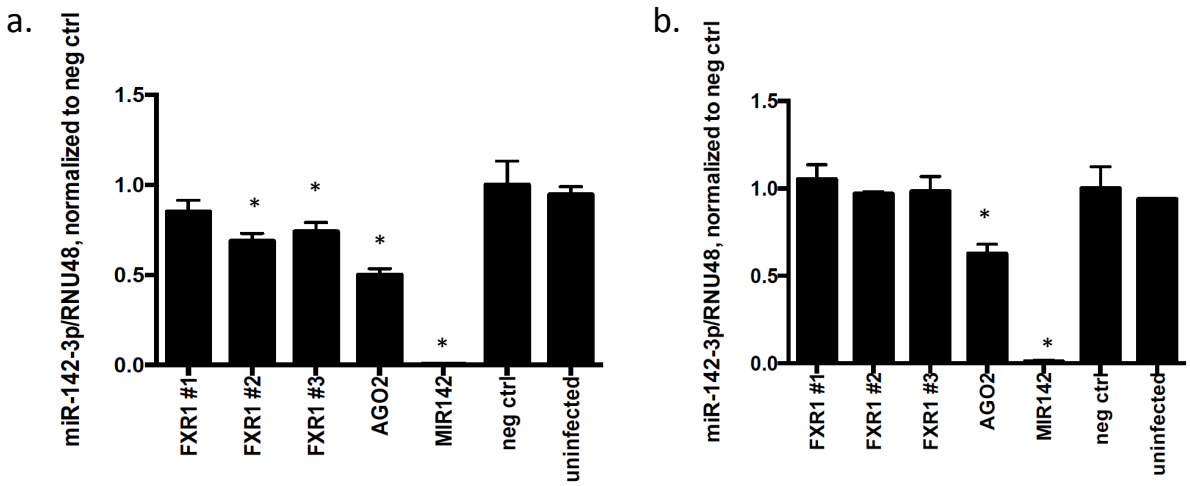


Figure 18. Changes in miR-142-3p expression in cells with sgRNAs against hit genes from CRISPR-Cas9 screen. (a) ddPCR to measure miR-142-3p in pCas9.Blast-expressing, sgRNA-expressing Raji-miR142-clB cells with Zeocin treatment. miR-142-3p levels normalized to RNU48 are compared to levels in negative-control sgRNA-infected cells. **(b)** The experiment was repeated under similar conditions (the Cas9 vector used was pCas9.HP) in an attempt to reproduce the apparent lower levels of miR-142-3p in *FXR1*-edited cells. Error bars are standard deviation of two biological replicates. * = $p<0.05$ Fisher's LSD.

Table 1. Vectors used throughout study.

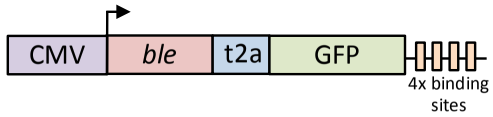
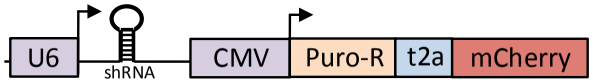
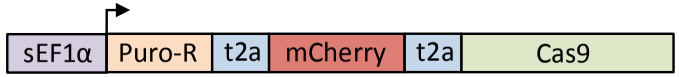
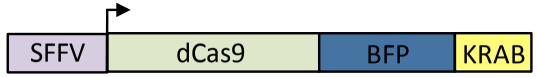
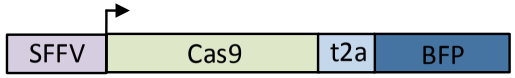
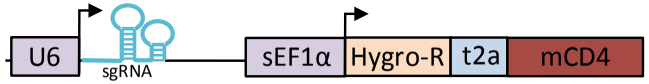
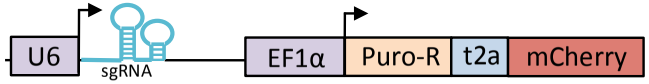
Name	Schematic	shRNA screen	shRNA validation	CRISPR screen	CRISPR validation
pSensor.miR142		✓	✓	✓	✓
MP177		✓	✓		
pLibrary.Cas9				✓	
pDCas9.KRAB					✓
pCas9.HP					✓
pSgRNA.1				✓	✓
pSgRNA.2					✓

Table 2. Top 50 genes by significance of sgRNA enrichment (Mann-Whitney rank sum test).

Gene	Z score	P-value (unadjusted)
AGO2	6.475210372	9.47E-11*
EGFP	5.883202697	4.02E-09*
PSME4	5.094672914	3.49E-07*
TP53	4.408566073	1.04E-05*
ANKRD28	4.067845669	4.74E-05*
ADARB1	3.221238855	0.001276377
FXR1	3.133906489	0.001724958
BAP1	3.04366171	0.002337178
SRSF9	2.915586238	0.00355021
MIR142	2.828539585	0.004676092
GRSF1	2.736569975	0.006208339
YAP1	2.73159316	0.006302892
UPF3A	2.708330394	0.006762267
HNRNPF	2.637524371	0.008351362
FAM208A	2.585100497	0.009735062
TNRC6A	2.584419853	0.009754297
SLC26A7	2.51097427	0.012039847
MYO1B	2.510465498	0.01205721
DDX23	2.465566885	0.013679663
PSPC1	2.455451739	0.014070764
ESR1	2.35935785	0.018306593
SIK2	2.354947433	0.018525326
SEPW1	2.306073434	0.021106529
NEGR1	2.251811049	0.024334215
MS4A2	2.22786095	0.025889789
HNRNPA0	2.210159444	0.027094099
TRIM71	2.200044298	0.027803752
DICER1	2.176999366	0.029480615
MYO6	2.168886675	0.030091289
NEDD4L	2.147120987	0.031783651
CNOT4	2.14338562	0.032082147
ZNF536	2.134295848	0.032818571
CBL	2.107791276	0.035049048
SKIV2L	2.098892836	0.035826349
DDX39A	2.098892836	0.035826349
PPRC1	2.095784212	0.036101338

HDC	2.091306476	0.0365006
LHX6	2.087054383	0.036883219
SLC30A1	2.068547397	0.038588579
CD80	2.06404994	0.039012982
ENDOD1	2.063489824	0.039066114
PSMA6	2.053374678	0.040036253
HNRNPL	2.036891355	0.041660931
CD3E	2.025710781	0.042794431
MPHOSPH8	2.012914093	0.044123668
ATXN7L3	2.010385307	0.044390424
CD19	1.998119449	0.045703712
SRSF2	1.997936873	0.045723504
LYPLA1	1.987424696	0.046875353

*- significant by FWER p<0.05

Table 3. List of genes and enriched sgRNAs selected from sgRNA screen for validation.

Gene	Sequence Name	Sequence
ABHD5	ABHD5_005__00026	AAAAGCATAGACAGGTCTGT
ABHD5	ABHD5_007__00028	AATAGGTCAAAAGCATAGAC
ADARB1	ADARB1_008__00192	GGCTCTCAGCTCTCCAATGG
ADARB1	ADARB1_009__00193	GGGCTCTCAGCTCTCCAATG
ADARB1	ADARB1_017__00201	GCCATGCAGAAATAATATCT
AGO1	AGO1_003__00259	CCACCTCAAAGTAATTGGCC
AGO1	AGO1_007__00263	GGCCTGGCATTGGCACTGTG
AGO2	AGO2_001__00269	AAGGCATATCCTTGGATGGG
AGO2	AGO2_013__03012	GTGCCTTGGATCCTCGCAGG
AGO2	AGO2_016__03015	GTACATGGTGGCGCCGCCGA
AGO2	AGO2_04_ctrl	AGGTCCCAAAGTCGGGTCT
ANKRD28	ANKRD28_011__00440	CTGGTGCAAGCTATATTTAA
ANKRD28	ANKRD28_021__00450	ATCCATTGGGCAGCATATAT
ANKRD28	ANKRD28_024__00453	TATCCATTGGGCAGCATATA
ATXN7L3	ATXN7L3_008__00668	CTACTTCTTCTTGGACGACA
ATXN7L3	ATXN7L3_009__00669	GATATACGCGGACCTGGTCG
ATXN7L3	ATXN7L3_011__00671	ATGGAGGAAATGTCTTTGTC
AVPR1A	AVPR1A_007__00679	CGTTGCGCACGTCCCTCGGT
AVPR1A	AVPR1A_009__00681	GGGCAACAGCAGCGTACTGC
BAP1	BAP1_002__00686	CAAGGTAGAGACCTTTCGCC
BAP1	BAP1_006__00690	TAGAGACCTTTCGCCGGGAC
BAP1	BAP1_009__00693	CGACCTTCAGAGCAAATGTC
CDKL4	CDKL4_010__01277	AGACTGGAGAAGGGTCTTAT
CDKL4	CDKL4_011__01278	GACTGGAGAAGGGTCTTATG
CTTNBP2	CTTNBP2_008__01845	CCCGCAGGGCCTCGATGACA
CTTNBP2	CTTNBP2_012__01849	TGCTCCTCAGCGTGATGGAA
DCP1B	DCP1B_010__01919	TCCACGATGCGGTTGATATA
DCP1B	DCP1B_012__01921	TGCGGTTGATATAGGGGTCG
DDX19B	DDX19B_002__02015	ACCGCCGCAGGGCGCGATCT
DDX19B	DDX19B_016__02029	CGCCAGGGCCCATGAGTCAG
DDX21	DDX21_008__02053	CAAAAAAGTGACAAAAAATG
DDX21	DDX21_013__02058	ACTCCGTAGTGACGCTGGTT
DDX23	DDX23_002__02063	AGGAGGAAAGGAAGCGATCA
DDX23	DDX23_005__02066	CCGCTTTCTATCTTTAGATG
DDX23	DDX23_006__02067	CCTTGGAAGGTGATGCATCA
DDX24	DDX24_004__02077	AGTAAAAATGTAGCAACTGA

DDX24	DDX24_006__02079	CTGTGGCAAATTTCAGACAA
DDX24	DDX24_008__02081	GGATGACTTGGTGTGCTTTG
DDX39A	DDX39A_001__02146	CCAGCATGAGTGCATTCCCC
DDX39A	DDX39A_007__02152	CCCTTGATGTCTTTCTTAGG
DDX39A	DDX39A_011__02156	TCCCCCTAAGAAAGACATCA
DDX54	DDX54_002__02354	CGTCCTCCGAGTCGCTGCCG
DDX54	DDX54_012__02364	TCCGAGGTGCATTCCGAGGT
DICER1	DICER1_003__02657	GACCAGGTTCCACGAAACGA
DICER1	DICER1_004__02658	GTGTTGAGGCCTGACGATGG
DICER1	DICER1_01_ctrl	AAATACCTACCTGAGGTAT
DICER1	DICER1_015__02669	ACCCCTGCTTCCTCACCAAT
DNM3	DNM3_001__02760	GGACTTTCTCCCTCGAGGGT
DNM3	DNM3_012__02771	TCCTGCAGACGGTTCACCAG
DROSHA	DROSHA_006__02813	GGTCGTGGAGGGAGAAAATT
DROSHA	DROSHA_010__02817	TCATATTGATATTGCACAGG
DROSHA	DROSHA_012__02819	TGTCGTTCCACCCGGGACGA
DUS4L	DUS4L_001__02832	AGCGTCGCTGTAGCTGGCCT
DUS4L	DUS4L_006__02837	GGGATTCTGCGTCACCACCT
DUS4L	DUS4L_012__02843	GCAGCAAACTGAACAATCAA
DYNLT3	DYNLT3_010__02873	TTAACACACCTGGTTAAGTT
DYNLT3	DYNLT3_011__02874	AGCCCACAATATTGTCAAAG
DYNLT3	DYNLT3_012__02875	GTACCATCGCCACTGCGACG
ESR1	ESR1_003__03290	CTTGAGCTGCGGACGGTTCA
ESR1	ESR1_004__03291	GCCCTCGGGGTAGTTGTACA
ESR1	ESR1_006__03293	TCAGATCCAAGGGAACGAGC
FAM208A	FAM208A_001__03483	AGCATGTAGGACTCGAGACT
FAM208A	FAM208A_010__03492	CCTTATAAAACAGTGCTGGA
FAM208A	FAM208A_018__03500	ATCTGAAAACTCCTCCTAAC
FAM35A	FAM35A_005__03535	CTTCCCAACATCGTATATAC
FAM35A	FAM35A_008__03538	ATTGGTTCTCCAGATCTTAG
FCHSD1	FCHSD1_006__03635	TCCGCCCTGTACCCCCTGCT
FCHSD1	FCHSD1_008__03637	GCCCATTCCTGAAGAGGGAA
FCHSD1	FCHSD1_012__03641	GCAGCCATTGAACGGGAGTA
FKTN	FKTN_002__03667	TTGGATTTGATAGCACACAG
FKTN	FKTN_009__03674	TTTCGGATAGCTGAGAATAT
FXR1	FXR1_010__03727	GAGCTGACGGTGGAGGTTCG
FXR1	FXR1_011__03728	GGAGGTTCGCGGCTCTAACG
FXR1	FXR1_018__03735	GGTGGTGGTTGGCTAAAGTT
GRSF1	GRSF1_008__03939	GCCGCTGGCTCTATCCCGTC

GRSF1	GRSF1_012__03943	TTCGTTTCCCATCTCTGTTT
GRSF1	GRSF1_019__03950	GTCTTTCTCATTCGAGCTCA
GTF2IRD2	GTF2IRD2_004__03977	GAAATGGAATTACCAATGGA
GTF2IRD2	GTF2IRD2_008__03981	CGTGTCTGCCCTCGAATCCA
HNRNPA0	HNRNPA0_004__04126	AGCAACGGGAGCGCTTGGTC
HNRNPA0	HNRNPA0_009__04131	CGCCTCCTCCACATTGGAGT
HNRNPA0	HNRNPA0_012__04134	GTGTTGCCGTCCACGGCATG
HNRNPF	HNRNPF_007__04233	CCATTTCATCTACACTAGAG
HNRNPF	HNRNPF_009__04235	CCTCTCTAGTGTAGATGAAA
HNRNPF	HNRNPF_010__04236	CTGACTGCACGATTCATGAT
IGF2BP3	IGF2BP3_005__04563	GAAGACTGGCTACGCGTTCG
IGF2BP3	IGF2BP3_012__04570	TTGAGGGCCCAGCTCTCGTC
IPCEF1	IPCEF1_003__04639	GCTCTTTCCACAGTGAAATC
IPCEF1	IPCEF1_008__04644	GAAGTTTCCTAAGCAACAAA
IPCEF1	IPCEF1_010__04646	GTACTGGTATAGCAATCAAA
KIAA1217	KIAA1217_002__04872	ACCAAGGAACGCCTTTCTAA
KIAA1217	KIAA1217_015__04885	TTCTGTCTGCTGAGTAAGGA
KIAA1217	KIAA1217_026__04896	TTGAGGACTGTGAGACAGTT
KIAA1407	KIAA1407_005__04902	TTTGAGAAGGACCTTCCCTG
KIAA1407	KIAA1407_011__04908	ACACTGTCAGGACCAAAAGT
KIAA1586	KIAA1586_006__04939	ATGTCCATGTGTCCAAGGAA
KIAA1586	KIAA1586_007__04940	CCTAAAATGCCAAAACGACA
KIAA1586	KIAA1586_011__04944	TGTTCAGCAGTTCGGCATTT
KRT73	KRT73_004__05047	CAGTCGGAGCCTTTACAGCC
KRT73	KRT73_006__05049	CGCCAATTCACCTACAAGTC
KRT73	KRT73_008__05051	CTCCCGACTTGTAGGTGAAT
LHX6	LHX6_001__05194	AGCGCTGAGAATCCCGACGC
LHX6	LHX6_012__05205	GGAACGGCCTCACGTTGGAG
LHX6	LHX6_018__05211	CATGATTGAGAACCTCAAGA
LYPLA1	LYPLA1_005__05351	CAGTTTCTGCTGTGTGGTAA
LYPLA1	LYPLA1_008__05354	GTGTCACTGCACTCAGTTGC
LYPLA1	LYPLA1_015__05361	CTTCCGGGCGGCGGGCACGA
MIR142	hsa-mir-142_002__04409	GCACTACTAACAGCACTGGA
MIR142	hsa-mir-142_003__04410	AGCACTACTAACAGCACTGG
MIR142	hsa-mir-142_005__04412	AGTACACTCATCCATAAAGT
MIR142	hsa-mir-142_005_6-ctrl	AGTACACTCATCCATAAAGT
MPHOSPH8	MPHOSPH8_002__05646	ATACATCGGATGATGATACC
MPHOSPH8	MPHOSPH8_003__05647	CAGTCCTCCAGGTGAATCTC
MPHOSPH8	MPHOSPH8_010__05654	GGAGGCCTTTGGCGACAGTG

MPP7	MPP7_005__05661	TCTCTGATCTCACTGTTTAA
MPP7	MPP7_008__05664	ACCTTCCTCTGGGATATGTT
MRPS5	MRPS5_007__05747	TGCACGGCTCAAGCTGGCGT
MRPS5	MRPS5_011__05751	TGCCGTCCCGCTACACAGCA
MS4A2	MS4A2_001__05753	ACTGTCAGCCATGTATGCAG
MS4A2	MS4A2_007__05759	GTCTGCCTGAAGATACTTCC
MS4A2	MS4A2_012__05764	TTCAGGCAGACTATTGAAGT
MYO1B	MYO1B_004__05924	CTGTATTCTTCCACTTTCTC
MYO1B	MYO1B_005__05925	GTGAATAAATGGGTAAAGAC
MYO1B	MYO1B_009__05929	TCTGTATGCTTCATCCGAAA
MYO6	MYO6_005__05937	CCATCTGAAATCCATCTGTA
MYO6	MYO6_008__05940	GTAGGGTGTGGCGCCCAAAC
MYO6	MYO6_010__05942	TGGAGGATGGAAAGCCCGTT
NCKAP5	NCKAP5_002__05970	TGAGAAGCTGATACATGAAC
NCKAP5	NCKAP5_009__05977	CTCAGCTTGAGGAGCAACAC
NCKAP5	NCKAP5_010__05978	AAAGAGACAGCTTGAGAAAA
NEDD4L	NEDD4L_003__06027	ATAGGAGTCTGTGATTAGAT
NEDD4L	NEDD4L_006__06030	GGAGCGACCCTATACATTTA
NEDD4L	NEDD4L_008__06032	GCCTATATGCCAAAAAATGG
NEGR1	NEGR1_001__06047	AATACTTGACCGGTTCAGCC
NEGR1	NEGR1_002__06048	AATGTTGAAATTGAAACTCG
NEGR1	NEGR1_003__06049	ATGGAGCTTCAAAGGGTGCC
nt (non-target)	nt_00_ctrl	CTCGCGTGGTAGAAGAAGT
NT5DC1	NT5DC1_002__06184	CTTTCTCCTTAACTAGGAAC
NT5DC1	NT5DC1_006__06188	TTTCTCCTTAACTAGGAACT
OR13C3	OR13C3_005__06268	CAAGGTGGCGTATGTATTGA
OR13C3	OR13C3_008__06271	GAAATTCTGACACAAGTGTC
OR13C3	OR13C3_009__06272	GAGAAAAAAATGACCATCAG
PABPC5	PABPC5_005__06412	CCCAGGGGGCTGCGGGTCAC
PABPC5	PABPC5_006__06413	CTCAAGGCCGCTCTGTACGT
PABPN1	PABPN1_003__06422	ATGAATATGAGTCCACCTCC
PABPN1L	PABPN1L_007__06438	AGGAGTGCAGCAACAGGCCG
PDP1	PDP1_005__06590	GGCAGGCTGGAACTTCTGAC
PDP1	PDP1_006__06591	GGTGGCAGTACACCCAAGGA
PHTF2	PHTF2_003__06663	TAACCACCACCGGAAGAAAA
PHTF2	PHTF2_008__06668	CCTGTACTTAGAGGAGGTTT
PIWIL2	PIWIL2_006__06726	CCTTTCCGACCATCGTTCAG
PIWIL2	PIWIL2_010__06730	TCCAAAGGTTTAGAAGCTTG
PKHD1L1	PKHD1L1_002__06758	CAACAAGGCTGACTATAAGA

PKHD1L1	PKHD1L1_011__06767	CCTGTGGCTCCTGGGTATTT
POF1B	POF1B_004__06821	ATGGTAGCAGTGGTAATGCT
POF1B	POF1B_005__06822	ATGTAGTGTATGAGCGAGTG
PPA2	PPA2_003__06951	ACCTGGAAGCTACTCTTAAT
PPA2	PPA2_007__06955	TGATGTTAAGAAGTTCAAAC
PPAPDC1B	PPAPDC1B_002__06971	ACACAATAAAACTGATCGTA
PPAPDC1B	PPAPDC1B_004__06973	GTCTGCCTTCTTGAGAAATT
PPRC1	PPRC1_005__07062	CTTTGTCAGTCTCTCTCGGC
PPRC1	PPRC1_010__07067	CCAAGCGCCGTATGGGACTT
PPRC1	PPRC1_012__07069	GTCGAAGCCAAGCGCCGTAT
PRC1	PRC1_001__07070	AAAAGATTTGCGCACCCAAG
PRC1	PRC1_002__07071	AGGCACTGTCAATATCATAG
PRC1	PRC1_006__07075	AAATCACCTTCGGGAAATAT
PRCC	PRCC_002__07085	AGGCGGTGGCTCCTACATCT
PRCC	PRCC_005__07088	CGGCTCGCTCTCATCGCTGC
PSME4	PSME4_003__07292	GAGCTGATTTGGAGTTACCC
PSME4	PSME4_010__07299	GGCCAGCTGCAAGTCGGACT
PSME4	PSME4_012__07301	GGCGTCTAGCCGCTCCGCGT
PSPC1	PSPC1_003__07304	AGTGCGCATTGAGAAAAACC
PSPC1	PSPC1_008__07309	CTCGCCCACCGCGGACTCCA
PSPC1	PSPC1_011__07312	TCCTCCTCCGTGATGTCGGT
PUF60	PUF60_003__07458	CTCCGTCACCATGAAGCACA
PUF60	PUF60_005__07460	GATGGGGCCAAAGGGGGCAA
PUF60	PUF60_013__07468	TCCTGCTGCTCGGGCGTCAG
RAB8B	RAB8B_005__07548	CGATGTCGAAAGAATGATCC
RAB8B	RAB8B_010__07553	CCCCCGAGTCGCCGATCAGC
random	random_251__07829	ATTGCAGGTGCTACAGCGCA
random	random_255__07833	TGGACAACCTATCGGGTGAT
random	random_304__07882	TCATCGAGATTAAGACTGCC
RBM10	RBM10_013__08104	CAGCCCCCGAAGGCCCTATC
RBM10	RBM10_024__08115	CGGGCTCCGAGACTCAGCGT
RBM10	RBM10_031__08122	TGGAGCCACTGACCGCTCGC
RNASEH2A	RNASEH2A_005__08896	CCAGGACGCAAGGCTCCTTG
RNASEH2A	RNASEH2A_007__08898	CCGCGGGCACAGGCGAACTC
RNASEH2A	RNASEH2A_011__08902	GGGCGTCGATGAGGCGGGCA
RPS2	RPS2_001__09016	TCTCTTCTCCCTGCCTATTA
RPS2	RPS2_005__09020	GGCCCTGGGATGGGGAACCG
RRAS2	RRAS2_001__09058	AAGCTTTCCATGAACTTGTC
RRAS2	RRAS2_006__09063	AAGACAGATTCTCAGAGTAA

RRM2B	RRM2B_002__09073	CCCGTTAGATTGCAAGTTGC
RRM2B	RRM2B_003__09074	CCGGCCGCTTCCGGCCTTTC
RSPO1	RSPO1_005__09135	CTCCAGCAGGATGAACAGCT
RSPO1	RSPO1_010__09140	CATCAGCAGCCGGGGGATCA
SBNO1	SBNO1_011__09301	GCAAGCAGTAAATCTTGCCC
SBNO1	SBNO1_012__09302	GCCTACCCCGTCAGTTCAGC
SEPW1	SEPW1_001__09498	CATCACTTCAAAGAACCCGG
SEPW1	SEPW1_002__09499	CTTAGAGTGAATCAACTTCC
SEPW1	SEPW1_004__09501	TCTTTGAAGTGATGGTAGCC
SERPINB2	SERPINB2_005__09507	ATCCACAGGGAATTATTTAC
SERPINB2	SERPINB2_009__09511	CAGATGCTTGAATAAATTGA
SERPINB2	SERPINB2_010__09512	CCAGATGCTTGAATAAATTG
SESN3	SESN3_004__09518	GAGATTTCTGGATCTCTCTC
SESN3	SESN3_005__09519	TAATTGGTCTCTGCCTGAAC
SET	SET_008__09538	ATGTTGTTACCCAAAAATTT
SET	SET_016__09546	AGCTCAACTCCAACCACGAC
SF3B14	SF3B14_005__09591	GGGAACACACCTGAAACTAG
SF3B14	SF3B14_008__09594	TAGGTGATCACATGCATTCT
SFRP4	SFRP4_007__09629	CGGTGCGCATCCCTATGTGC
SFRP4	SFRP4_008__09630	GCCACAGGCACAGCGCCACT
SIK2	SIK2_001__09650	AAAGTTTGATTATGTGAGGG
SIK2	SIK2_003__09652	ACTATCTTGCTAATCATGGC
SIK2	SIK2_004__09653	GGTTAAATGAGTCTGAAGCC
SKIV2L	SKIV2L_005__09678	CCGTGGAGCTCGGATGCACG
SKIV2L	SKIV2L_008__09681	TCCGAGCTCCACGGCCCGAA
SKIV2L	SKIV2L_012__09685	TGGGAGCTGCTGAACTTGCC
SLC26A7	SLC26A7_005__09706	CACGGCGTTGGCTGATATTA
SLC26A7	SLC26A7_008__09709	ATTATCCCAGACACAGTGTC
SLC26A7	SLC26A7_010__09711	GTATAATGTCTTCACACTGG
SLC30A1	SLC30A1_003__09716	CCCAGAAGAACACGTTCGGC
SLC30A1	SLC30A1_005__09718	GAGCCCCCATTACCTCGGCT
SLC30A1	SLC30A1_007__09720	GCCACCCAGAAGAACACGTT
SNRPA	SNRPA_014__09987	GCGGGTCTCGGGAACTGCCA
SNRPA	SNRPA_015__09988	GGTTGTTGATATAAATAGTG
SPATA4	SPATA4_004__10074	AGGGTATTTGACACAGACTG
SPATA4	SPATA4_005__10075	CGGGAGCTCTTCGGCGCATG
SRSF9	SRSF9_001__10306	ATCGAGCTCAAGAACCGGCA
SRSF9	SRSF9_004__10309	GGTTCTTGAGCTCGATCTCG
SRSF9	SRSF9_010__10315	CCCGACCTCCATAAGTCCTG

STAU2	STAU2_001__10394	AGAGAAAACTGCAATGTGTC
STAU2	STAU2_003__10396	GGGGTTGGACTCTATTGAAA
SURF4	SURF4_003__10505	CACCTGGTCGGCGAAGTCCT
SURF4	SURF4_005__10507	CCGTGCCCATCAGGTCGTTC
SURF4	SURF4_007__10509	GGGCGCCTGCGATTGGACCC
tetR	tetR_001__10775	CGCATTAGAGCTGCTTAATG
tetR	tetR_003__10777	GCCAGCTTTCCCCTTCTAAA
tetR	tetR_010__10784	TAGAGCAGCCTACATTGTAT
TNRC6A	TNRC6A_006__11036	CTTGGCATTATTATTAGTGC
TNRC6A	TNRC6A_011__11041	TGTTCCTGGTGGCGAAATCG
TNRC6A	TNRC6A_012__11042	TTCACGAGGATACCGAGGCA
TP53	TP53_005__11087	CATCAAATCATCCATTGCTT
TP53	TP53_017__11099	AATCAACCCACAGCTGCACA
TP53	TP53_025__11107	GACCTGCCCTGTGCAGCTGT
TRIM71	TRIM71_004__11308	CCGGCGCGCCCGTTCTTGGG
TRIM71	TRIM71_005__11309	CGACGACGCGGACGAGTTGG
TRIM71	TRIM71_006__11310	CTAAGCAGGAAGGCGGACGA
TTC27	TTC27_004__11410	TCCTGAGGGTGTAAGTCAAC
TTC27	TTC27_006__11412	ATTCACCATCTAGAGTGAGC
TTC27	TTC27_015__11421	TTCCTACAATTGCTACTGGA
UPF3A	UPF3A_004__11613	CAACCGGAGGACGAAGTTGG
UPF3A	UPF3A_009__11618	GCGGTGGAACTGCACTTCTA
UPF3A	UPF3A_011__11620	GGCATGCGCTCGGAAAAGGA
VPS54	VPS54_007__11688	TAGGATCGTTTAATGCTGCT
VPS54	VPS54_011__11692	TTGATCTGTAACTAGAGATG
YAP1	YAP1_011__11864	GGTTGCCCGGGTCCGGACGG
YAP1	YAP1_020__11873	GGGCCAGAGACTACTCCAGT
YAP1	YAP1_032__11885	GCAGCAGAATATGATGAACT
ZMYM4	ZMYM4_006__12095	GAAACCTGATCAAATTGTTT
ZMYM4	ZMYM4_009__12098	TCCATAAGACATGCTGTCAA
ZMYM6	ZMYM6_008__12109	GAAAGCCTGGGTTCAACTGC
ZMYM6	ZMYM6_012__12113	CAAAGCAGTGGTACCACAGC
ZNF518A	ZNF518A_010__12170	TGCAGAAACACTTTCAAATG
ZNF518A	ZNF518A_015__12175	AATGGTGGTACCTAATACCT
ZNF518A	ZNF518A_016__12176	ACTGTTGTATGCATTTGGTT
ZNF536	ZNF536_003__12223	CTGGCCGCTCATGGGCACGT
ZNF536	ZNF536_007__12227	GCATACTGGCCGTTGAGGAC
ZNF536	ZNF536_008__12228	GCTCCGCCGAAGACACTCCA
ZNF638	ZNF638_002__12258	ATAGATCCAGGCCTCATAAA

ZNF638	ZNF638_011__12267	TCTCCTCGAGGATTAAACCT
ZNF800	ZNF800_002__12306	GGCTTCTAGGAGATCATTTA
ZNF800	ZNF800_009__12313	AGTGGTTGCTGTAACAAAGG
ZNF800	ZNF800_010__12314	ATAATTGAGTGCTTTCGATC

Table 4. List of primers used to amplify and deep-sequence samples from shRNA and CRISPR screens.

Screen	Primer name	Forward, reverse, or sequencing primer	Sequence
shRNA	5'BstX-2-illum	Forward	AATGATACGGCGACCACCGACAC TCTTTCCCACAAAAGGAAACTCAC CCTAAC
shRNA/CRISPR	3'XHO-2-illum	Reverse	CAAGCAGAAGACGGCATACGAGC GGTAATACGGTTATCCACG
shRNA	BstX Solexa 5' seq	Sequencing	GAGACTATAAGTATCCCTTGGAG AACCACCTTGTTGG
CRISPR	newpol3index1	Forward	AATGATACGGCGACCACCGAGAT CGGAAGAGCACACGTCTGAACTC CAGTCACACAGTGCACAAAAGGA AACTCACCCTAAC
CRISPR	newpol3index2	Forward	AATGATACGGCGACCACCGAGAT CGGAAGAGCACACGTCTGAACTC CAGTCACAGTCACCACAAAAGGA AACTCACCCTAAC
CRISPR	newpol3index3	Forward	AATGATACGGCGACCACCGAGAT CGGAAGAGCACACGTCTGAACTC CAGTCACTCTCAGCACAAAAGGA AACTCACCCTAAC
CRISPR	newpol3index4	Forward	AATGATACGGCGACCACCGAGAT CGGAAGAGCACACGTCTGAACTC CAGTCACTGAGTCCACAAAAGGA AACTCACCCTAAC
CRISPR	newpol3index5	Forward	AATGATACGGCGACCACCGAGAT CGGAAGAGCACACGTCTGAACTC CAGTCACCACTGTCACAAAAGGA AACTCACCCTAAC
CRISPR	newpol3index6	Forward	AATGATACGGCGACCACCGAGAT CGGAAGAGCACACGTCTGAACTC CAGTCACCTGACACACAAAAGGA AACTCACCCTAAC
CRISPR	newpol3index7	Forward	AATGATACGGCGACCACCGAGAT CGGAAGAGCACACGTCTGAACTC CAGTCACGAGACTCACAAAAGGA AACTCACCCTAAC
CRISPR	Truseq-index	Sequencing (Truseq index only)	GATCGGAAGAGCACACGTCTGAA CTCCAGTCAC
CRISPR	illum-g-lib-seq	Sequencing	ACTCACCCTAACTGTAAAGTAATT GTGTGTTTTGAGACTATAAGTATC CCTAGC

Chapter 3 – General Summary

The work presented here represents the first unbiased attempts to discover novel regulators of miR-142 and RNAi. In the study of miRNAs and RNAi in general, the screens performed are the first to be done in a pooled format, getting around several of the drawbacks of arrayed screening. In addition, the CRISPR-Cas9 screen was the first to be performed focusing on miRNA/RNAi activity, thus distancing the biology from the awkward use of shRNAs to study the very machinery required for their function.

Future experiments examining miR-142 activity

The hit lists from the shRNA and CRISPR-Cas9 screens yield promising inroads toward the understanding of how miR-142 functions, and they also give us some lessons on how the screening platform used, the transgene expression, and antibiotics can influence a screen. Manipulation of the core miRNA machinery easily influences the activity of miR-142, as expected. The challenge now is to identify which, if any, of the validated genes not in the canonical miRNA pathway alter the expression or activity of miR-142.

Despite our inability thus far to demonstrate any robust change in miR-142 basal expression upon *FXR1* editing, we have not ruled out effects on RISC activity, as *FXR1* editing does appear to increase levels of the miRNA sensor. If *FXR1* can be established to affect miR-142 activity, the next areas of research interest would include an assessment into the breadth of miRNAs regulated in this manner. A more general study into all miRNAs would touch on a broader question that has not yet been explored in the field: what are the individual contributions for each of the three Fragile X genes toward the expression or activity of miRNAs? Profiles of miRNAs regulated by the Fragile X family have not been performed in a comprehensive manner,

and so these individual and combinatorial contributions remain poorly characterized. We look to identify the set of miRNAs affected by *FXR1*, in addition to *FXR2* and *FMR1*, using combinatorial edits of each of the three genes.

Other than *FXR1*, other hits that have been assessed may have roles in transgene expression (*FAM208A*) or DNA-damage (*PSME4*), which would cause these genes to enrich independent of miR-142. These other phenotypes highlight a possible need to execute parallel screens from multiple angles, using unrelated antibiotics for selection pressure, and perhaps screening with multiple sensor clones and constructs to minimize transgene-specific effects.

Outside of the basal expression of miR-142 in B cells, several other genome-wide screens can be performed to assess the factors affecting the dynamic expression of miR-142, which occurs in several contexts, including development, the innate immune response, and the circadian cycle. Several viruses also rely on the miR-142 network to the degree that several viruses have developed miR-142 mimics and house crucial miR-142 binding sites on their genomes (75,126,127). Screens performed during viral infection could yield other important miR-142 regulating genes that would not otherwise be found in an assessment of basal expression or dynamic expression in other biological contexts.

The future of screening in the wake of CRISPR technology

During the execution of the CRISPR-Cas9 screen, the first two CRISPR-Cas9 screens were published, establishing the viability of the system to study various phenotypes at an unbiased, genome-wide scale (128,129). My screen lends further credence to the technology's potential. CRISPR-Cas9 technology has vastly increased the different types of low-throughput

and high-throughput experiments that can be performed, and a full review on these applications is written elsewhere (130).

Briefly, some of the applications for CRISPR-Cas9 technology include the ability to theoretically edit any locus in the genome, to knockdown coding and non-coding genes alike, and to activate the expression of genes. The ability to edit any region of the genome (131) in a high-throughput manner enables researchers to explore more than just the protein-coding genome and assay a large set of conserved genomic regions for phenotypic relevance. This will lead to screens editing putative enhancers, miRNA loci, regions with disease-relevant single nucleotide polymorphisms (SNPs), 5' and 3' UTR regulatory elements, and ultra-conserved non-coding elements (UCNEs). Utilizing dCas9-KRAB, or CRISPRi, will enable the first functional screens targeting a set of long non-coding RNAs (lncRNAs). Activating systems such as dCas9-VP64 (132), dCas9-SunTag (133), and dCas9-p300 Core (134) may serve as a high-throughput replacement for cDNA library screens, enabling the overexpression of proteins, miRNAs, and lncRNAs with the same simple set of tools. Each of these different applications widens the possibilities for almost any phenotype of interest. Indeed, future screens looking at miRNA activity may look at gene classes never examined before in the context of miRNA biology.

One paradigm shift in the field of high-throughput screening is the assessment of genetic interaction. Instead of looking at gene manipulations one at a time, looking at combinatorial gene manipulations allows researchers to see genes that are functionally similar, indicating that the genes are part of the same gene network, or even code for different parts of a larger protein complex. Screens of this nature have been used in yeast genetics for decades. With the recent construction of an epistasis map (or EMAP) of putative gene networks controlling ricin toxicity (94), genetic interaction screens have been shown to be viable in mammalian cells. However,

this EMAP was constructed from double-shRNA libraries, which knockdown their targeted genes, but do not fully mimic the full knockout screens performed in yeast. With the advent of CRISPR-Cas9, experiments in mammalian cells have been brought a step closer toward the yeast screens. The development of several different Cas9 platforms to edit, epigenetically inhibit, and activate genes allow for more types of genetic interaction screens to mimic double-knockouts, double-siRNA knockdowns, or introduction of two cDNAs. The Cas9 sgRNA has also been shown to be very customizable (135), allowing the possibility for one Cas9 protein to activate or inhibit genes depending on the sgRNA it has bound. A recent demonstration of the activity of another CRISPR system, Cpf1, in human cells gives another option toward using both Cas9 and Cpf1 together to produce gene manipulations going in opposite directions (136). Genetic interaction studies may reveal new networks and complexes governing the expression of certain miRNAs. The customizations of Cas9 sgRNAs can also localize protein fusions to specified genomic locations. This may provide possibilities to test in a high throughput manner how genomic structure regulates expression of specific genes, which has not really been explored for most genes, let alone miRNAs.

In short, there are seemingly endless possibilities for the types of high-throughput screens that can be performed with CRISPR systems. CRISPR screening technology will prove crucial toward understanding the genetic networks governing disease-relevant miRNAs.

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 Feb 19;391(6669):806–11.
2. Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, Hayward DC, Ball EE, Degnan B, Müller P, Spring J, Srinivasan A, Fishman M, Finnerty J, Corbo J, Levine M, Leahy P, Davidson E, Ruvkun G. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature. 2000 Nov 2;408(6808):86–9.
3. Bartel DP. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell. 2004 Jan 23;116(2):281–97.
4. Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol. 2009 Feb;10(2):126–39.
5. Cai X, Hagedorn CH, Cullen BR. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA. 2004 Dec 1;10(12):1957–66.
6. Lee Y, Kim M, Han J, Yeom K-H, Lee S, Baek SH, Kim VN. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 2004 Oct 13;23(20):4051–60.
7. Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A. Identification of Mammalian microRNA Host Genes and Transcription Units. Genome Res. 2004 Oct 1;14(10a):1902–10.
8. Chiang HR, Schoenfeld LW, Ruby JG, Auyeung VC, Spies N, Baek D, Johnston WK, Russ C, Luo S, Babiarz JE, Blelloch R, Schroth GP, Nusbaum C, Bartel DP. Mammalian

microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev. 2010 May 15;24(10):992–1009.

9. Chang T-C, Pertea M, Lee S, Salzberg SL, Mendell JT. Genome-wide annotation of microRNA primary transcript structures reveals novel regulatory mechanisms. Genome Res. 2015 Sep 1;25(9):1401–9.

10. Kim Y-K, Kim VN. Processing of intronic microRNAs. EMBO J. 2007 Feb 7;26(3):775–83.

11. Borchert GM, Lanier W, Davidson BL. RNA polymerase III transcribes human microRNAs. Nat Struct Mol Biol. 2006 Dec;13(12):1097–101.

12. Yang J-S, Lai EC. Alternative miRNA Biogenesis Pathways and the Interpretation of Core miRNA Pathway Mutants. Mol Cell. 2011 Sep;43(6):892–903.

13. Brummelkamp TR, Bernards R, Agami R. A System for Stable Expression of Short Interfering RNAs in Mammalian Cells. Science. 2002 Apr 19;296(5567):550–3.

14. Gregory RI, Yan K, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. The Microprocessor complex mediates the genesis of microRNAs. Nature. 2004 Nov 11;432(7014):235–40.

15. 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.

16. Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S, Kim VN. The nuclear RNase III Drosha initiates microRNA processing. Nature. 2003 Sep 25;425(6956):415–9.

17. Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. Nature. 2007 Jul 5;448(7149):83–6.

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

19. Lund E, Güttinger S, Calado A, Dahlberg JE, Kutay U. Nuclear Export of MicroRNA Precursors. Science. 2004 Jan 2;303(5654):95–8.

20. Bohnsack MT, Czaplinski K, Görlich D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA. 2004 Feb 1;10(2):185–91.

21. 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.

22. Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, Shiekhattar R. TRBP recruits the Dicer complex to AGO2 for microRNA processing and gene silencing. Nature. 2005 Aug 4;436(7051):740–4.

23. Schürmann N, Trabuco LG, Bender C, Russell RB, Grimm D. Molecular dissection of human Argonaute proteins by DNA shuffling. Nat Struct Mol Biol. 2013 Jul;20(7):818–26.

24. Johnson SM, Lin S-Y, Slack FJ. The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. Dev Biol. 2003 Jul 15;259(2):364–79.

25. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. Nature. 2005 Jun 9;435(7043):839–43.

26. Raver-Shapira N, Marciano E, Meiri E, Spector Y, Rosenfeld N, Moskovits N, Bentwich Z, Oren M. Transcriptional Activation of miR-34a Contributes to p53-Mediated Apoptosis. Mol Cell. 2007 Jun 8;26(5):731–43.

27. Chang T-C, Wentzel EA, Kent OA, Ramachandran K, Mullendore M, Lee KH, Feldmann G, Yamakuchi M, Ferlito M, Lowenstein CJ, Arking DE, Beer MA, Maitra A, Mendell JT. Transactivation of miR-34a by p53 Broadly Influences Gene Expression and Promotes Apoptosis. Mol Cell. 2007 Jun 8;26(5):745–52.

28. He L, He X, Lim LP, de Stanchina E, Xuan Z, Liang Y, Xue W, Zender L, Magnus J, Ridzon D, Jackson AL, Linsley PS, Chen C, Lowe SW, Cleary MA, Hannon GJ. A microRNA component of the p53 tumour suppressor network. Nature. 2007 Jun 28;447(7148):1130–4.

29. Davis-Dusenbery BN, Hata A. MicroRNA in Cancer. Genes Cancer. 2010 Nov;1(11):1100–14.

30. Suzuki HI, Yamagata K, Sugimoto K, Iwamoto T, Kato S, Miyazono K. Modulation of microRNA processing by p53. Nature. 2009 Jul 23;460(7254):529–33.

31. Davis BN, Hilyard AC, Lagna G, Hata A. SMAD proteins control DROSHA-mediated microRNA maturation. Nature. 2008 Jul 3;454(7200):56–61.

32. Davis BN, Hilyard AC, Nguyen PH, Lagna G, Hata A. Smad Proteins Bind a Conserved RNA Sequence to Promote MicroRNA Maturation by Drosha. Mol Cell. 2010 Aug 13;39(3):373–84.

33. Auyeung VC, Ulitsky I, McGeary SE, Bartel DP. Beyond Secondary Structure: Primary-Sequence Determinants License Pri-miRNA Hairpins for Processing. Cell. 2013 Feb 14;152(4):844–58.

34. Viswanathan SR, Daley GQ, Gregory RI. Selective Blockade of MicroRNA Processing by Lin28. Science. 2008 Apr 4;320(5872):97–100.

35. Heo I, Joo C, Cho J, Ha M, Han J, Kim VN. Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. Mol Cell. 2008 Oct 24;32(2):276–84.

36. Heo I, Joo C, Kim Y-K, Ha M, Yoon M-J, Cho J, Yeom K-H, Han J, Kim VN. TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. Cell. 2009 Aug 21;138(4):696–708.

37. Guil S, Cáceres JF. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. Nat Struct Mol Biol. 2007 Jul;14(7):591–6.

38. Xu X-L, Zong R, Li Z, Biswas MHU, Fang Z, Nelson DL, Gao F-B. FXR1P But Not FMRP Regulates the Levels of Mammalian Brain-Specific microRNA-9 and microRNA-124. J Neurosci. 2011 Sep 28;31(39):13705–9.

39. Pare JM, Tahbaz N, López-Orozco J, LaPointe P, Lasko P, Hobman TC. Hsp90 Regulates the Function of Argonaute 2 and Its Recruitment to Stress Granules and P-Bodies. Mol Biol Cell. 2009 Jul 15;20(14):3273–84.

40. Miyoshi T, Takeuchi A, Siomi H, Siomi MC. A direct role for Hsp90 in pre-RISC formation in Drosophila. Nat Struct Mol Biol. 2010 Aug;17(8):1024–6.

41. Wu C, So J, Davis-Dusenbery BN, Qi HH, Bloch DB, Shi Y, Lagna G, Hata A. Hypoxia Potentiates MicroRNA-Mediated Gene Silencing through Posttranslational Modification of Argonaute2. Mol Cell Biol. 2011 Dec 1;31(23):4760–74.

42. Bronevetsky Y, Villarino AV, Eisley CJ, Barbeau R, Barczak AJ, Heinz GA, Kremmer E, Heissmeyer V, McManus MT, Erle DJ, Rao A, Ansel KM. T cell activation induces proteasomal degradation of Argonaute and rapid remodeling of the microRNA repertoire. J Exp Med. 2013 Feb 11;210(2):417–32.

43. Jin P, Zarnescu DC, Ceman S, Nakamoto M, Mowrey J, Jongens TA, Nelson DL, Moses K, Warren ST. Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway. Nat Neurosci. 2004 Feb;7(2):113–7.

44. Vasudevan S, Steitz JA. AU-Rich-Element-Mediated Upregulation of Translation by FXR1 and Argonaute 2. Cell. 2007 Mar 23;128(6):1105–18.

45. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Aldler H, Rattan S, Keating M, Rai K, Rassenti L, Kipps T, Negrini M, Bullrich F, Croce CM. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci. 2002 Nov 26;99(24):15524–9.

46. Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, Iorio MV, Visone R, Sever NI, Fabbri M, Iuliano R, Palumbo T, Pichiorri F, Roldo C, Garzon R, Sevignani C, Rassenti L, Alder H, Volinia S, Liu C, Kipps TJ, Negrini M, Croce CM. A MicroRNA Signature Associated with Prognosis and Progression in Chronic Lymphocytic Leukemia. N Engl J Med. 2005 Oct 27;353(17):1793–801.

47. Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, Wojcik SE, Aqeilan RI, Zupo S, Dono M, Rassenti L, Alder H, Volinia S, Liu C, Kipps TJ, Negrini M, Croce CM. miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A. 2005 Sep 27;102(39):13944–9.

48. Gauwerky CE, Huebner K, Isobe M, Nowell PC, Croce CM. Activation of MYC in a masked t(8;17) translocation results in an aggressive B-cell leukemia. Proc Natl Acad Sci U S A. 1989 Nov;86(22):8867–71.

49. Thomson JM, Newman M, Parker JS, Morin-Kensicki EM, Wright T, Hammond SM. Extensive post-transcriptional regulation of microRNAs and its implications for cancer. Genes Dev. 2006 Aug 15;20(16):2202–7.

50. Chen C-Z, Li L, Lodish HF, Bartel DP. MicroRNAs Modulate Hematopoietic Lineage Differentiation. Science. 2004 Jan 2;303(5654):83–6.

51. Sun Y, Sun J, Tomomi T, Nieves E, Mathewson N, Tamaki H, Evers R, Reddy P. PU.1-Dependent Transcriptional Regulation of miR-142 Contributes to Its Hematopoietic Cell–Specific Expression and Modulation of IL-6. J Immunol. 2013 Apr 15;190(8):4005–13.

52. Bellon M, Lepelletier Y, Hermine O, Nicot C. Deregulation of microRNA involved in hematopoiesis and the immune response in HTLV-I adult T-cell leukemia. Blood. 2009 May 14;113(20):4914–7.

53. Ju X, Li D, Shi Q, Hou H, Sun N, Shen B. DIFFERENTIAL microRNA EXPRESSION IN CHILDHOOD B-CELL PRECURSOR ACUTE LYMPHOBLASTIC LEUKEMIA. Pediatr Hematol-Oncol. 2009 Jan 1;26(1):1–10.

54. Lv M, Zhang X, Jia H, Li D, Zhang B, Zhang H, Hong M, Jiang T, Jiang Q, Lu J, Huang X, Huang B. An oncogenic role of miR-142-3p in human T-cell acute lymphoblastic leukemia (T-ALL) by targeting glucocorticoid receptor-α and cAMP/PKA pathways. Leukemia. 2012 Apr;26(4):769–77.

55. Wang F, Wang X-S, Yang G-H, Zhai P-F, Xiao Z, Xia L-Y, Chen L-R, Wang Y, Wang X-Z, Bi L-X, Liu N, Yu Y, Gao D, Huang B-T, Wang J, Zhou D-B, Gong J-N, Zhao H-L, Bi X-H, Yu J, Zhang J-W. miR-29a and miR-142-3p downregulation and diagnostic implication in human acute myeloid leukemia. Mol Biol Rep. 2011 Jun 16;39(3):2713–22.

56. Wang X-S, Gong J-N, Yu J, Wang F, Zhang X-H, Yin X-L, Tan Z-Q, Luo Z-M, Yang G-H, Shen C, Zhang J-W. MicroRNA-29a and microRNA-142-3p are regulators of myeloid differentiation and acute myeloid leukemia. Blood. 2012 May 24;119(21):4992–5004.

57. Robbiani DF, Bunting S, Feldhahn N, Bothmer A, Camps J, Deroubaix S, McBride KM, Klein IA, Stone G, Eisenreich TR, Ried T, Nussenzweig A, Nussenzweig MC. AID Produces DNA Double-Strand Breaks in Non-Ig Genes and Mature B Cell Lymphomas with Reciprocal Chromosome Translocations. Mol Cell. 2009 Nov 25;36(4):631–41.

58. Zhao J-J, Lin J, Lwin T, Yang H, Guo J, Kong W, Dessureault S, Moscinski LC, Rezania D, Dalton WS, Sotomayor E, Tao J, Cheng JQ. microRNA expression profile and identification of miR-29 as a prognostic marker and pathogenetic factor by targeting CDK6 in mantle cell lymphoma. Blood. 2010 Apr 1;115(13):2630–9.

59. Kwanhian W, Lenze D, Alles J, Motsch N, Barth S, Döll C, Imig J, Hummel M, Tinguely M, Trivedi P, Lulitanond V, Meister G, Renner C, Grässer FA. MicroRNA-142 is mutated in about 20% of diffuse large B-cell lymphoma. Cancer Med. 2012 Oct 1;1(2):141–55.

60. Sun W, Shen W, Yang S, Hu F, Li H, Zhu T-H. miR-223 and miR-142 attenuate hematopoietic cell proliferation, and miR-223 positively regulates miR-142 through LMO2 isoforms and CEBP-β. Cell Res. 2010 Oct;20(10):1158–69.

61. Chapnik E, Rivkin N, Mildner A, Beck G, Pasvolsky R, Metzl-Raz E, Birger Y, Amir G, Tirosh I, Porat Z, Israel LL, Lellouche E, Michaeli S, Lellouche J-PM, Izraeli S, Jung S, Hornstein E. miR-142 orchestrates a network of actin cytoskeleton regulators during megakaryopoiesis. eLife. 2014 Jun 24;3:e01964.

62. Bettencourt P, Marion S, Pires D, Santos LF, Lastrucci C, Carmo N, Blake J, Benes V, Griffiths G, Neyrolles O, Lugo-Villarino G, Anes E. Actin-binding protein regulation by microRNAs as a novel microbial strategy to modulate phagocytosis by host cells: the case of N-Wasp and miR-142-3p. Front Cell Infect Microbiol [Internet]. 2013 Jun 5 [cited 2015 Nov 25];3. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3672780/

63. Naqvi AR, Fordham JB, Nares S. miR-24, miR-30b, and miR-142-3p Regulate Phagocytosis in Myeloid Inflammatory Cells. J Immunol. 2015 Feb 15;194(4):1916–27.

64. Liu J, Li W, Wang S, Wu Y, Li Z, Wang W, Liu R, Ou J, Zhang C, Wang S. MiR-142-3p Attenuates the Migration of CD4+ T Cells through Regulating Actin Cytoskeleton via RAC1 and ROCK2 in Arteriosclerosis Obliterans. PLoS ONE. 2014 Apr 17;9(4):e95514.

65. Wu L, Cai C, Wang X, Liu M, Li X, Tang H. MicroRNA-142-3p, a new regulator of RAC1, suppresses the migration and invasion of hepatocellular carcinoma cells. FEBS Lett. 2011 May 6;585(9):1322–30.

66. Zheng Z, Ding M, Ni J, Song D, Huang J, Wang J. miR-142 acts as a tumor suppressor in osteosarcoma cell lines by targeting Rac1. Oncol Rep. 2015 Mar;33(3):1291-9.

67. Sun Y, Varambally S, Maher CA, Cao Q, Chockley P, Toubai T, Malter C, Nieves E, Tawara I, Wang Y, Ward PA, Chinnaiyan A, Reddy P. Targeting of microRNA-142-3p in dendritic cells regulates endotoxin-induced mortality. Blood. 2011 Jun 9;117(23):6172–83.

68. Huang B, Zhao J, Lei Z, Shen S, Li D, Shen G-X, Zhang G-M, Feng Z-H. miR-142-3p restricts cAMP production in CD4+CD25− T cells and CD4+CD25+ TREG cells by targeting AC9 mRNA. EMBO Rep. 2009 Feb;10(2):180–5.

69. Shende VR, Goldrick MM, Ramani S, Earnest DJ. Expression and Rhythmic Modulation of Circulating MicroRNAs Targeting the Clock Gene Bmal1 in Mice. PLoS ONE. 2011 Jul 22;6(7):e22586.

70. Tan X, Zhang P, Zhou L, Yin B, Pan H, Peng X. Clock-controlled mir-142-3p can target its activator, Bmal1. BMC Mol Biol. 2012 Sep 7;13(1):27.

71. Shende VR, Neuendorff N, Earnest DJ. Role of miR-142-3p in the Post-Transcriptional Regulation of the Clock Gene Bmal1 in the Mouse SCN. PLoS ONE. 2013 Jun 5;8(6):e65300.

72. Shende VR, Kim S-M, Neuendorff N, Earnest DJ. MicroRNAs function as cis- and trans-acting modulators of peripheral circadian clocks. FEBS Lett. 2014 Aug 25;588(17):3015–22.

73. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T. Identification of Tissue-Specific MicroRNAs from Mouse. Curr Biol. 2002 Apr 30;12(9):735–9.

74. Yang W, Chendrimada TP, Wang Q, Higuchi M, Seeburg PH, Shiekhattar R, Nishikura K. Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. Nat Struct Mol Biol. 2006 Jan;13(1):13–21.

75. Manzano M, Forte E, Raja AN, Schipma MJ, Gottwein E. Divergent target recognition by coexpressed 5′-isomiRs of miR-142-3p and selective viral mimicry. RNA. 2015 Sep;21(9):1606-20.

76. 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.

77. Boettcher M, Hoheisel JD. Pooled RNAi Screens - Technical and Biological Aspects. Curr Genomics. 2010 May;11(3):162–7.

78. Aza-Blanc P, Cooper CL, Wagner K, Batalov S, Deveraux QL, Cooke MP. Identification of Modulators of TRAIL-Induced Apoptosis via RNAi-Based Phenotypic Screening. Mol Cell. 2003 Jan 9;12(3):627–37.

79. Brummelkamp TR, Bernards R, Agami R. Stable suppression of tumorigenicity by virus-mediated RNA interference. Cancer Cell. 2002 Sep;2(3):243–7.

80. Rubinson DA, Dillon CP, Kwiatkowski AV, Sievers C, Yang L, Kopinja J, Rooney DL, Zhang M, Ihrig MM, McManus MT, Gertler FB, Scott ML, Van Parijs L. A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. Nat Genet. 2003 Mar;33(3):401–6.

81. Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepfer AM, Hinkle G, Piqani B, Eisenhaure TM, Luo B, Grenier JK, Carpenter AE, Foo SY, Stewart SA, Stockwell BR, Hacohen N, Hahn WC, Lander ES, Sabatini DM, Root DE. A Lentiviral RNAi Library for Human and Mouse Genes Applied to an Arrayed Viral High-Content Screen. Cell. 2006 Mar 24;124(6):1283–98.

82. Bassik MC, Lebbink RJ, Churchman LS, Ingolia NT, Patena W, LeProust EM, Schuldiner M, Weissman JS, McManus MT. Rapid creation and quantitative monitoring of high coverage shRNA libraries. Nat Methods. 2009 Jun;6(6):443–5.

83. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature. 2012 Feb 16;482(7385):331–8.

84. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 2012 Aug 17;337(6096):816–21.

85. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-Guided Human Genome Engineering via Cas9. Science. 2013 Feb 15;339(6121):823–6.

86. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science. 2013 Feb 15;339(6121):819–23.

87. Dorner S, Lum L, Kim M, Paro R, Beachy PA, Green R. A genomewide screen for components of the RNAi pathway in Drosophila cultured cells. Proc Natl Acad Sci. 2006 Aug 8;103(32):11880–5.

88. Marshall WF. Modeling Recursive RNA Interference. PLoS Comput Biol. 2008 Sep 19;4(9):e1000183.

89. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001 Feb 15;409(6822):860–921.

90. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The Sequence of the Human Genome. Science. 2001 Feb 16;291(5507):1304–51.

91. Consortium IHGS. Finishing the euchromatic sequence of the human genome. Nature. 2004 Oct 21;431(7011):931–45.

92. Boutros M, Ahringer J. The art and design of genetic screens: RNA interference. Nat Rev Genet. 2008 Jul;9(7):554–66.

93. van de Weijer ML, Bassik MC, Luteijn RD, Voorburg CM, Lohuis MAM, Kremmer E, Hoeben RC, LeProust EM, Chen S, Hoelen H, Ressing ME, Patena W, Weissman JS, McManus MT, Wiertz EJHJ, Lebbink RJ. A high-coverage shRNA screen identifies TMEM129 as an E3 ligase involved in ER-associated protein degradation. Nat Commun. 2014 May 8;5:3832.

94. Bassik MC, Kampmann M, Lebbink RJ, Wang S, Hein MY, Poser I, Weibezahn J, Horlbeck MA, Chen S, Mann M, Hyman AA, LeProust EM, McManus MT, Weissman JS. A Systematic Mammalian Genetic Interaction Map Reveals Pathways Underlying Ricin Susceptibility. Cell. 2013 Feb 14;152(4):909–22.

95. Matheny CJ, Wei MC, Bassik MC, Donnelly AJ, Kampmann M, Iwasaki M, Piloto O, Solow-Cordero DE, Bouley DM, Rau R, Brown P, McManus MT, Weissman JS, Cleary ML. Next-Generation NAMPT Inhibitors Identified by Sequential High-Throughput Phenotypic Chemical and Functional Genomic Screens. Chem Biol. 2013 Nov 21;20(11):1352–63.

96. Qin H, Diaz A, Blouin L, Lebbink RJ, Patena W, Tanbun P, LeProust EM, McManus MT, Song JS, Ramalho-Santos M. Systematic Identification of Barriers to Human iPSC Generation. Cell. 2014 Jul 17;158(2):449–61.

97. Brown BD, Venneri MA, Zingale A, Sergi LS, Naldini L. Endogenous microRNA regulation suppresses transgene expression in hematopoietic lineages and enables stable gene transfer. Nat Med. 2006 May;12(5):585–91.

98. Kim JK, Gabel HW, Kamath RS, Tewari M, Pasquinelli A, Rual J-F, Kennedy S, Dybbs M, Bertin N, Kaplan JM, Vidal M, Ruvkun G. Functional Genomic Analysis of RNA Interference in C. elegans. Science. 2005 May 20;308(5725):1164–7.

99. Parry DH, Xu J, Ruvkun G. A Whole-Genome RNAi Screen for C. elegans miRNA Pathway Genes. Curr Biol. 2007 Dec 4;17(23):2013–22.

100. Saleh M-C, van Rij RP, Hekele A, Gillis A, Foley E, O'Farrell PH, Andino R. The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. Nat Cell Biol. 2006 Aug;8(8):793–802.

101. Ulvila J, Parikka M, Kleino A, Sormunen R, Ezekowitz RA, Kocks C, Rämet M. Double-stranded RNA Is Internalized by Scavenger Receptor-mediated Endocytosis in Drosophila S2 Cells. J Biol Chem. 2006 May 19;281(20):14370–5.

102. Zhou R, Hotta I, Denli AM, Hong P, Perrimon N, Hannon GJ. Comparative Analysis of Argonaute-Dependent Small RNA Pathways in Drosophila. Mol Cell. 2008 Nov 21;32(4):592–9.

103. Janas MM, Wang E, Love T, Harris AS, Stevenson K, Semmelmann K, Shaffer JM, Chen P-H, Doench JG, Yerramilli SVBK, Neuberg DS, Iliopoulos D, Housman DE, Burge CB, Novina CD. Reduced Expression of Ribosomal Proteins Relieves MicroRNA-Mediated Repression. Mol Cell. 2012 Apr 27;46(2):171–86.

104. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M, Lin C, Socci ND, Hermida L, Fulci V, Chiaretti S, Foà R, Schliwka J, Fuchs U, Novosel A, Müller R-U, Schermer B, Bissels U, Inman J, Phan Q, Chien M, Weir DB, Choksi R, De Vita G, Frezzetti D, Trompeter H-I, Hornung V, Teng G, Hartmann G, Palkovits M, Di Lauro R, Wernet P, Macino G, Rogler CE, Nagle JW, Ju J,

Papavasiliou FN, Benzing T, Lichter P, Tam W, Brownstein MJ, Bosio A, Borkhardt A, Russo JJ, Sander C, Zavolan M, Tuschl T. A Mammalian microRNA Expression Atlas Based on Small RNA Library Sequencing. Cell. 2007 Jun 29;129(7):1401–14.

105. Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol. 2013 Sep;31(9):839–43.

106. Kim JH, Lee S-R, Li L-H, Park H-J, Park J-H, Lee KY, Kim M-K, Shin BA, Choi S-Y. High Cleavage Efficiency of a 2A Peptide Derived from Porcine Teschovirus-1 in Human Cell Lines, Zebrafish and Mice. PLoS ONE. 2011 Apr 29;6(4):e18556.

107. Moore CW. Isolation and partial characterization of mutants of Saccharomyces cerevisiae altered in sensitivities to lethal effects of bleomycins. J Antibiot (Tokyo). 1980 Nov;33(11):1369–75.

108. Moore CW. Further characterizations of bleomycin-sensitive (blm) mutants of Saccharomyces cerevisiae with implications for a radiomimetic model. J Bacteriol. 1991 Jun 1;173(11):3605–8.

109. Febres DE, Pramanik A, Caton M, Doherty K, McKoy J, Garcia E, Alejo W, Moore CW. The novel BLM3 gene encodes a protein that protects against lethal effects of oxidative damage. Cell Mol Biol Noisy--Gd Fr. 2001 Nov;47(7):1149–62.

110. Blickwedehl J, Agarwal M, Seong C, Pandita RK, Melendy T, Sung P, Pandita TK, Bangia N. Role for proteasome activator PA200 and postglutamyl proteasome activity in genomic stability. Proc Natl Acad Sci. 2008 Oct 21;105(42):16165–70.

111. Blickwedehl J, Olejniczak S, Cummings R, Sarvaiya N, Mantilla A, Chanan-Khan A, Pandita TK, Schmidt M, Thompson CB, Bangia N. The Proteasome Activator PA200

Regulates Tumor Cell Responsiveness to Glutamine and Resistance to Ionizing Radiation. Mol Cancer Res. 2012 Jul 1;10(7):937–44.

112. Orban TI, Izaurralde E. Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. RNA. 2005 Apr 1;11(4):459–69.

113. Harten SK, Bruxner TJ, Bharti V, Blewitt M, Nguyen T-M-T, Whitelaw E, Epp T. The first mouse mutants of D14Abb1e (Fam208a) show that it is critical for early development. Mamm Genome. 2014 Apr 30;25(7-8):293–303.

114. Tchasovnikarova IA, Timms RT, Matheson NJ, Wals K, Antrobus R, Göttgens B, Dougan G, Dawson MA, Lehner PJ. Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. Science. 2015 Jun 26;348(6242):1481–5.

115. 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.

116. Caudy AA, Myers M, Hannon GJ, Hammond SM. Fragile X-related protein and VIG associate with the RNA interference machinery. Genes Dev. 2002 Oct 1;16(19):2491–6.

117. Ishizuka A, Siomi MC, Siomi H. A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev. 2002 Oct 1;16(19):2497–508.

118. Jin P, Zarnescu DC, Ceman S, Nakamoto M, Mowrey J, Jongens TA, Nelson DL, Moses K, Warren ST. Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway. Nat Neurosci. 2004 Feb;7(2):113–7.

119. Lee A, Li W, Xu K, Bogert BA, Su K, Gao F-B. Control of dendritic development by the Drosophila fragile X-related gene involves the small GTPase Rac1. Development. 2003 Nov 15;130(22):5543–52.

120. Zhang J, Fang Z, Jud C, Vansteensel MJ, Kaasik K, Lee CC, Albrecht U, Tamanini F, Meijer JH, Oostra BA, Nelson DL. Fragile X-Related Proteins Regulate Mammalian Circadian Behavioral Rhythms. Am J Hum Genet. 2008 Jul 11;83(1):43–52.

121. Xu S, Poidevin M, Han E, Bi J, Jin P. Circadian Rhythm-Dependent Alterations of Gene Expression in Drosophila Brain Lacking Fragile X Mental Retardation Protein. PLoS ONE. 2012 May 24;7(5):e37937.

122. Alpatov R, Lesch BJ, Nakamoto-Kinoshita M, Blanco A, Chen S, Stützer A, Armache KJ, Simon MD, Xu C, Ali M, Murn J, Prisic S, Kutateladze TG, Vakoc CR, Min J, Kingston RE, Fischle W, Warren ST, Page DC, Shi Y. A Chromatin-Dependent Role of the Fragile X Mental Retardation Protein FMRP in the DNA Damage Response. Cell. 2014 May 8;157(4):869–81.

123. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. Cell. 2013 Dec 19;155(7):1479–91.

124. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010 Oct 27;11(10):R106.

125. Whitlock MC. Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. J Evol Biol. 2005 Sep 1;18(5):1368–73.

126. Trobaugh DW, Gardner CL, Sun C, Haddow AD, Wang E, Chapnik E, Mildner A, Weaver SC, Ryman KD, Klimstra WB. RNA viruses can hijack vertebrate microRNAs to suppress innate immunity. Nature. 2014 Feb 13;506(7487):245–8.

127. Forte E, Raja AN, Shamulailatpam P, Manzano M, Schipma MJ, Casey JL, Gottwein E. MicroRNA-Mediated Transformation by the Kaposi's Sarcoma-Associated Herpesvirus Kaposin Locus. J Virol. 2015 Feb 15;89(4):2333–41.

128. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. Science. 2014 Jan 3;343(6166):80–4.

129. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F. Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science. 2014 Jan 3;343(6166):84–7.

130. Boettcher M, McManus MT. Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. Mol Cell. 2015 May 21;58(4):575–85.

131. Knight SC, Xie L, Deng W, Guglielmi B, Witkowsky LB, Bosanac L, Zhang ET, Beheiry ME, Masson J-B, Dahan M, Liu Z, Doudna JA, Tjian R. Dynamics of CRISPR-Cas9 genome interrogation in living cells. Science. 2015 Nov 13;350(6262):823–6.

132. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA-guided activation of endogenous human genes. Nat Methods. 2013 Oct;10(10):977–9.

133. Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. Cell. 2014 Oct 23;159(3):635–46.

134. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, Gersbach CA. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol. 2015 May;33(5):510–7.

135. Shechner DM, Hacisuleyman E, Younger ST, Rinn JL. Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. Nat Methods. 2015 Jul;12(7):664–70.

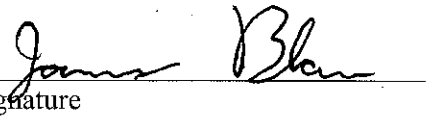
136. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell. 2015 Oct 22;163(3):759-71.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature

12/17/2015

Date