

Stony Brook University



OFFICIAL COPY

The official electronic file of this thesis or dissertation is maintained by the University Libraries on behalf of The Graduate School at Stony Brook University.

© All Rights Reserved by Author.

RNA interference Screens as a Tool for Discovering Gene Function

A Dissertation Presented

by

Despina Siolas

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

In

Genetics

Stony Brook University

August 2008

Stony Brook University

The Graduate School

Despina Siolas

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Greg Hannon-Dissertation Advisor

Professor, Cold Spring Harbor Laboratory, Genetics Program, Stony Brook University

Scott Lowe- Chairperson of Defense

Professor, Cold Spring Harbor Laboratory, Genetics Program, Stony Brook University

Scott Powers

Associate Professor, Cold Spring Harbor Laboratory,

Genetics Program, Stony Brook University

Vivek Mittal

Assistant Professor, Cold Spring Harbor Laboratory,

Genetics Program, Stony Brook University

Mike Hemann

Assistant Professor, Massachusetts Institute of Technology

This dissertation is accepted by the Graduate School

Lawrence Martin

Dean of the Graduate School

Abstract of the Dissertation
RNA interference Screens as a Tool for Discovering Gene Function
by
Despina Siolas
Doctor of Philosophy
in
Genetics
Stony Brook University
2008

Recent advances in the field of RNA interference (RNAi) have enabled researchers to conduct in depth investigations of gene function. High throughput screens in cultured mammalian cells can now be performed using RNAi libraries such as our own Hannon-Elledge library. Our second-generation shRNA libraries consist of over 200,000 short hairpin RNA constructs targeting over 45,000 human and mouse genes modeled after primary miRNA transcripts. Each hairpin is linked to a unique 60 nucleotide identification sequence, which serves as a barcode and allows us to virtually count the number of cells that contain a specific hairpin in a cell population. Small changes in barcode copy number can be monitored through the use of microarray technology. The barcode can be amplified from a cell's genomic DNA and fluorescently labeled to produce a probe that is hybridized to a microarray. We have optimized our probe labeling methods, probe size and hybridization conditions using library plasmid DNA in a Nimblegen platform. This optimized protocol allows us to detect 78.6% of probes within 1 standard deviation above the mean background from a complex mixture of approximately 1500 hairpins. By using two color hybridizations, we can detect control subsets of hairpins known to be depleted from a sample population. I applied this RNAi barcode screening method in a screen using a complex mixture of 7500 library hairpins in a p53 isogenic HCT 116 colon cancer cells to identify genes that modify sensitivity to a common chemotherapeutic, doxorubicin. Knockdown of Chk1, a kinase that mediates the G2/M checkpoint, increases sensitivity to doxorubicin in the p53 deficient HCT116 cells but not to the p53 wt HCT116 cell line. This work illustrates the powerful use of RNAi screens to search for genes that synergize with existing therapeutics, and suggests strategies for genetically informed combination therapies.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS.....	vi
LIST OF FIGURES.....	vii
LIST OF TABLES.....	ix
CHAPTER 1	
INTRODUCTION.....	1
1.1 Cancer.....	2
1.1.1 Cancer Chemotherapy as a cure.....	2
1.1.2 Targeted Therapeutics.....	3
1.1.3 Personalized cancer treatment	4
1.1.4 Challenges to current cancer research and treatment.....	5
1.2 RNA interference.....	5
1.2.1 RNA interference mechanism.....	6
1.2.2 Challenges to RNA interference in mammalian cells.....	8
1.2.3 Rational Design of RNAi triggers.....	10
1.2.4 Development of small RNA libraries for functional genomics.....	11
1.3 Functional Genomics.....	12
1.3.1 RNAi Genetic Screens.....	12
1.4 RNAi Therapeutics.....	19
1.5 Questions addressed in this Thesis.....	22
CHAPTER 2	
SYNTHETIC shRNAs AS HIGHLY POTENT RNAI TRIGGERS.....	30
2.1 Introduction.....	32
2.2 Results.....	33
2.2.1 Dicer cleaves a predominant small RNA from the end of each shRNA.....	33
2.2.2 shRNAs are generally more effective than siRNAs.....	36
2.2.3 siRNAs and shRNAs give similar profiles of off-target effects at saturation.....	38
2.3 Discussion.....	39
2.4 Methods.....	40
2.5 Figures.....	44
CHAPTER 3	
SECOND-GENERATION shRNA LIBRARIES COVERING THE MOUSE AND HUMAN GENOMES.....	53
3. 1 Introduction.....	55
3. 2 Results	57
3.2.1 Design and construction of second generation shRNA libraries.....	57
3.2.2 Validation of the second-generation shRNA libraries.....	63
3.3 Discussion.....	65
3.4 Methods.....	67
3.5 Figures.....	76
CHAPTER 4	
HIGH THROUGHPUT RNA INTERFERENCE BARCODE SCREENS AS A TOOL FOR DISCOVERING GENE FUNCTION	94
4.1 Introduction.....	95

4.2 Experimental System.....	97
4.3 Results.....	100
4.3.1 Different labeling methods result in different frequencies of incorporation of labeled nucleotides.....	100
4.3.2 Probe size and quantity affect microarray intensities.....	103
4.3.3 Frequency of incorporation affects hybridization intensity.....	105
4.3.4 Random primer labeling with smaller probes is the best method for doing negative selection screens.....	106
4.3.5 The biggest source of variability for our hybridizations is the day of hybridization.....	108
4.3.6 Correlation data can indicate hybridization quality.....	110
4.3.7 Specialized data analysis for hairpins with variable knockdown.....	112
4.4 Discussion.....	115
CHAPTER 5	
SYNTHETIC LETHALITY SCREENS FOR THE DISCOVERY OF CANCER CHEMOTHERAPEUTIC TARGETS.....	120
5.1 Introduction.....	121
5.2 Experimental System.....	123
5.3 Results.....	126
5.3.1 Preprocessing of data reveals variability within the screen.....	126
5.3.2 Different statistical analyses reveals different candidates.....	128
5.3.3 Candidates list includes DNA damage response genes, novel genes and checkpoint related genes.....	138
5.3.4 Chk1 knockdown sensitizes p53 null cells to Doxorubicin.....	141
5.3.5 Additional Validation Assays.....	143
5.4 Discussion.....	148
5.4.1 Synthetic lethal screen execution, data analysis and validation....	148
5.4.2 Chk1 inhibitors for combination therapy with doxorubicin in p53 null cells.....	151
CHAPTER 6	
CONCLUSIONS AND PERSPECTIVES.....	167
6.1 Experimental Conclusions and their Context in the Field.....	168
6.1.1 Synthetic shRNAs are more potent silencing triggers than siRNAs.....	168
6.1.2 Establishment of second-generation shRNA libraries covering the mouse and human genomes.....	170
6.1.3 Barcode shRNA screens are tools for discovering gene function...171	
6.1.4. A synthetic lethal microarray screen identifies Chk1 as a modifier of doxorubicin sensitivity to p53 null cells.....	172
6.2 Future Directions.....	173
6.2.1 Expansion of the Hannon Elledge Library and our screening capabilities.....	173
6.2.2 Screens across multiple drugs and genotypes.....	175
6.2.3 RNAi therapeutics.....	177
6.3 Summary.....	178
BIBLIOGRAPHY.....	192

List of Abbreviations

CML Chronic Myeloid Leukemia
CSHL Cold Spring Harbor Laboratory
Cy3 Arylsulphonate Cyanine Fluorescent Dye 3
Cy5 Arylsulphonate Cyanine Fluorescent Dye 5
DMEM Dulbecco's Modified Eagle Medium
DLBCL Diffuse Large B-cell Lymphoma
DNA Deoxyribonucleic Acid
dsRNA Double stranded RNA
DXR Doxorubicin
EDGE Extraction of Differential Gene Expression
FACS Fluorescence Activated Cell Sorting
GFP Green Fluorescent Protein
IRES Internal Ribosome Entry Site
FOI Frequency of Incorporation
LB Luria Broth
LTR Long Terminal Repeat
MAGIC Mating-Assisted Genetically Integrated Cloning
MEF Mouse Embryonic Fibroblast
miRNA micro-RNA
MK2 Mapkapk2
MLP MSCV LTR MiR-30 Puro-IRES-GFP Vector
MSCV Murine Stem Cell Virus Retroviral Expression System
PCA Principal Component Analysis
PCR Polymerase Chain Reaction
PIG Puro-IRES-GFP Cassette
Pri-miRNA Primary microRNA
P-bodies Processing bodies
PKR double-stranded RNA activated *protein kinase*
pSM2 plasmid SHAG-MAGIC 2 Vector
RISC RNA Induced Silencing Complex
RLC RISC Loading Complex
RNA Ribonucleic Acid
RNAi RNA Interference
ROMA Representational Oligonucleotide Microarray Analysis
SAFE Significance Analysis of Function and Expression
SAM Significance Analysis of Microarray
shRNA short-hairpin RNA
shRNA^{mir} second generation short-hairpin library RNA
siRNA small-interfering RNA
UTR Untranslated Region

List of Figures

Chapter 1

1.1: Acquired capabilities of Cancer.....	2
1.2: RNA interference.....	7
1.3: Example of siRNA design algorithm.....	11
1.4: Features of the Hannon Elledge Library.....	15
1.5: RNAi screening approaches.....	18
1.6: Using RNAi screens to identify targets controlling the hallmarks of cancer.....	23

Chapter 2

2.1: <i>In vitro</i> processing of 29 bp. shRNAs by Dicer generates a predominant siRNA from the end of each short hairpin.....	44
2.2: Primer extension analysis reveals similar small RNAs are generated by Dicer processing <i>in vitro</i> or <i>in vivo</i>	44
2.3: Gene suppression by shRNAs is comparable to or more effective than that achieved by siRNAs targeting the same sequences.....	45
2.4: Microarray profiling reveals gene expression profiles are more similar between 29mer shRNAs and corresponding siRNAs than between siRNAs and 19mer shRNAs.....	46
Supplementary Figure 2.1 RNase III digestion of shRNA.....	46
Supplementary Figure 2.2 Northern blotting indicates that siRNAs and 19mer and 29mer shRNAs all give rise to RISC.....	47
Supplementary Figure 2.3 Structures of synthetic RNAs used for comparing siRNA and shRNA.....	47

Chapter 3

3.1: Design and structure of shRNA ^{mir} cassettes.....	76
3.2: Construction of the second-generation library.....	77
3.3: Validation of the second-generation library.....	78
3.4: Performance of the second-generation library in a small-scale high-throughput screen.....	79
Supplementary Figure 3.1 Drosha and Dicer processing sites.....	79
Supplementary Figure 3.2 shRNA mir hairpin sequences.....	81
Supplementary Figure 3.3 Stable suppression by pSMC2.....	82

Chapter 4

4.1: A typical RNAi screen in a stably selected population.....	98
4.2: Diagram of a PCR product from the MLP shRNA vector.....	100
4.3: Labelling methods.....	102
4.4: Smaller probes have higher intensity signal on a microarray hybridization.....	104
4.5: Frequency of Incorporation affects hybridization intensity.....	106
4.6: Frequency of incorporation affects hybridization intensity.....	107
4.7: Random primer labeling with smaller probes is the best method for doing negative selection screens.....	108
4.8: The biggest source of variability for our hybridizations is the day of hybridization.....	110

4.9: Correlation coefficients can indicate hybridization quality.....	112
4.10 Multiple hairpins for the same gene can be analyzed by SAFE.....	114

Chapter 5:

5.1: Synthetic lethality and Cancer.....	123
5.2: Different screen analyses answer different questions.....	126
5.3: Correlation coefficients for synthetic lethal screen reveal variables.....	127
5.4: PCA Analysis shows p53 status as a component of variability.....	128
5.5: Two graphical interpretations of Edge results.....	129
5.6: Chk1 hairpins knock down at varying efficiency.....	143
5.7: Chk1 knockdown increases sensitivity to doxorubicin in p53 null cells....	143
5.8: PCA analysis of Agilent screen.....	146
5.9: Agilent re-screen samples have strong correlation coefficients between biological replicates in both the experimental and reference channel.....	147

List of Tables

Chapter 2

Supplementary Table 2.1 Sequences of the siRNAs used in this study.....	48
---	----

Chapter 3

Table 3.1 Library coverage by functional group.....	75
Supplementary Table 3.1: shRNA sequences used in Figure 3.3b	82
Supplementary Table 3.2: shRNA sequences used in Figure 4.....	82
Supplementary Table 3.3: Oligonucleotides used in construction of the library vectors	84

Chapter 5

Table 5.1 Probes depleted in the p53 null cell line but not the p53 wt cell line (DMSO only).....	130
Table 5.2 Probes depleted in HCT116 p53 null cells after treatment with Doxorubicin.....	132
Table 5.3 Probes depleted in HCT116 p53 null cells with paclitaxel treatment.....	136
Table 5.4 Candidate gene list resulting from screen for validation experiments.....	139
Table 5.5 Probes depleted in doxorubicin samples regardless of p53 status.....	145
Table 5.6 Probes depleted in doxorubicin samples for the p53 null cell line with a p value less than 0.1.....	145

Acknowledgements

As the daughter of a cancer survivor, I wanted to do a PhD because of my love of research and my passion to eradicate cancer. Throughout the duration of this thesis, my wits were tested beyond anything I could have imagined. I would have surely given up if it weren't for the support and encouragement of my friends and family.

I want to recognize my committee members, Scott Lowe, Scott Powers, Vivek Mittal and Mike Hemann for their help and guidance. Ken Chang taught me everything I know about microarrays. Through Scott Hammond, I met Joel Parker, whose long standing collaboration taught me much about bioinformatics, though I have problems explaining it verbally. Asking someone to edit a thesis is like asking for a ride to the airport or help moving out of an apartment. No one wants to do it but they will because they are your friends. I need to thank Adam Rosebrock, Emily Hodges, Kata Fejes-Toth and Claudio Scuoppo (who's still offering me western blots and papers). I really wish Jonathan Kui, Darren Burgess, Faby Rivas, Jack Zilfou, Ross Dickins, Monica Dus, and Ahmet Denhli were able to see me graduate because they helped me tremendously in the beginning and throughout of my graduate pursuits. I want to thank McClintock Tissue Culture Finer Things Society members, particularly Danny Khalil, Dave Simpson, Simone Salghetti, Ingrid Ibarra and John Kurland for their patience with my hours of ipod music and political podcasts.

I want to thank my friends for so many good times and lots of comfort for the bad times: Kasandra Burgos, Sabrina Boettcher, Jo Leonardo, Sihem Cheloufi, Astrid Haase, Ingrid Ibarra, Kata Fejes-Toth, Hana Mizuno and Adam Rosebrock. I would like to thank the AACR for providing me with support to attend and give presentations at their meetings. The Dept of Defense Breast Cancer Research Program made this research possible through their grants.

Finally, I want to thank my parents, John and Catherine Siolas for sacrificing so much to see me progress. They are truly role models. I need to thank my thesis advisor, Greg Hannon. If I didn't meet him in a random application process as an undergraduate, the course of my life would be radically changed. I look up to Greg and I hope to some day be 1/10 of the person he is.

The author hereby certifies that the use of any copyrighted material in the manuscript beyond brief excerpts is with the permission of the copyright owner.

The following authors have given permission to use their figures in this manuscript:

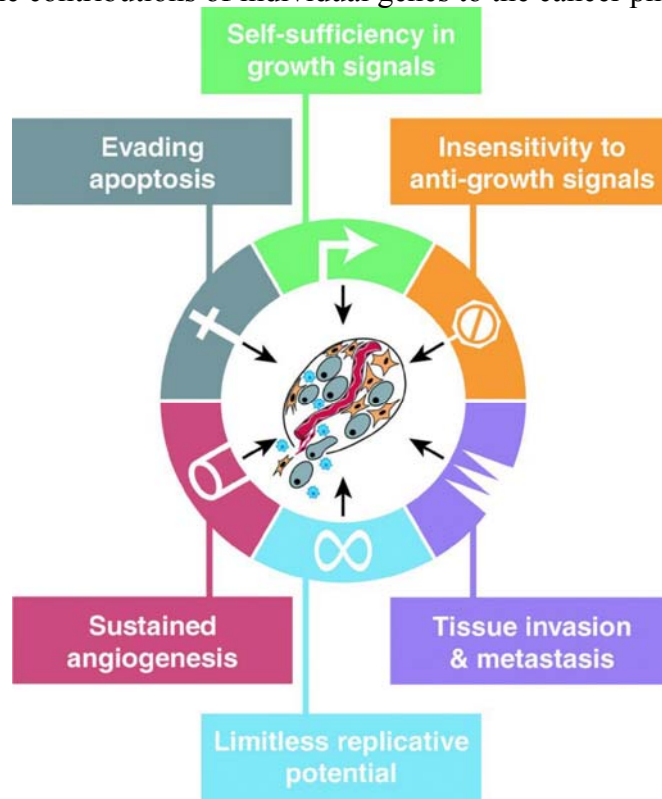
Alan Ashworth
Tony De Fougérolles
Doug Hanahan
William G Kaelin
Judy Lieberman
Jeff MacKeigan

Chapter 1

Introduction

Section 1.1: Cancer

Cancer is a genetic disease resulting in uncontrolled cellular division, aberrant apoptosis and unregulated homeostasis¹ (see Figure 1.1). While no single genetic defect is causal, cancer arises from multiple mutations in tumor suppressors, oncogenes, and repair genes.² The last 30 years of biomedical research have provided great insight into how these mutations occur and their resulting phenotype. Cancer is mostly studied in cultured cells and mouse models of human tumors, systems which have shed light onto the contributions of individual genes to the cancer phenotype.¹ Researchers face a



number of major challenges in eradicating cancer. These include discovering new genes that play a causal role in neoplasia, understanding how multiple genetic mutations interact to produce cancer, and finally, developing new therapeutics that will target cancer by exploiting the genetic changes unique to cancer cells.^{1,2}

Figure 1.1: Acquired capabilities of cancer. Most cancer cells have acquired the same set of functional capabilities during their development as a result of multiple genetic mutations¹. (Figure taken from Hanahan and Weinberg, 2000).

1.1.1 Cancer chemotherapy as a cure

Before 1950, cancer therapy was mainly enacted through the surgeon's scalpel. Since then, our cancer treatment methods have grown to include radiation therapy and drugs.

Early cancer chemotherapy focused on using toxins developed for chemical warfare, such as nitrogen mustard, to combat cancer.³ While these single agents caused a brief remission, they were largely not curative. Further biological studies established the utility of drug therapy. Antifolates, such as methotrexate, were effective in leukemia patients, who also underwent a brief but clear remission. The beginnings of modern chemotherapy were not established until studies by James Holland, Emil Freireich and Emil Frei in the 1960's showed that a combination of methotrexate, vincristine, 6-MP and prednisone could cause long term remission in patients with acute lymphoblastic leukemia.⁴ This approach was the first attempt at combination therapy. It is now known that combination therapy is often essential for tumor eradication because of the resistance of tumors to single agents. Human tumors are genetically diverse and the selective pressure of a single agent can create an environment for drug resistant cells to emerge.⁵ In order to make the correct combination of drugs work in each cancer, we must have a deepened understanding of how our current chemotherapeutics work and find new drugs that will enhance our existing arsenal.

1.1.2 Targeted therapies

Research from the 1980's and 1990's have illuminated our understanding of the cell signaling pathways that are altered in cancer. This new knowledge of molecular pathways combined with innovations in combinatorial chemistry brought forth the era of targeted therapies. Targeted therapies are designed to interfere with a specific molecular target that is believed to have a critical role in tumor growth or progression.⁶ Pharmacological achievements include Imatinib (targeting Bcr-Abl, c-Kit and PDGFR),

Bevacizumab (targeting VEGF ligand) and Tarceva (targeting EGFR). While these drugs hold strong promise, there is still much research needed to use them effectively. For example, some patients with Chronic Myeloid Leukemia (CML) have resistant cells even before exposure to Imatinib.⁷ Other patients with the acute leukemic phase of CML have a brief remission in response to Gefitinib, but end up developing drug resistant cells.⁸

Targeted therapeutics can be made more effective by knowing what subset of patients will best respond to them. For example, studies of non small cell lung cancer patients identified molecular changes in EGFR that are associated with increased response to Gefitinib therapy.⁹ Distinct mutations that cause varying degrees of resistance to Imatinib have been found by studying patients whose CML relapsed after initial treatment.⁸ Some of these mutations were altered amino acids that directly contacted Imatinib, and others prevented BCR-ABL from achieving the conformational state required for drug binding⁸. Thus, our increasingly sophisticated knowledge of the molecular profiles of human tumors is altering our treatment of cancer as we are beginning to understand why certain tumors respond to chemotherapeutics and why others seem resistant.

1.1.3 Personalized cancer treatment

Cancer treatment options for patients have largely been decided by the histological classification of tumors. Because of the genetic origin of cancer, this can be ineffective as a treatment method as only a subset of patients with a solid tumor may prove responsive to any given agent.⁵ While clinical trials can help decipher what combination of drugs will work for most patients, genomic and molecular assays are

being developed to identify subsets of patients that will respond to certain drugs. This will help avoid the needless cost and toxicity associated with ineffective treatment. A recent example of this is the FDA approved test MammaPrint that uses the Amsterdam 70-gene breast cancer gene signature¹⁰ to classify breast cancer as “high risk” or “low risk” clarifying the risk of breast cancer recurrence in an individual patient. The intention of this test is to use the genomic signature of breast tumors to help physicians plan appropriate follow-up for patients.¹¹ While time will assess the efficacy and utility of this test, it is a forward step towards using our knowledge of the molecular basis of tumors to plan patient treatment.

1.1.4 Challenges to current cancer research and treatment

Our understanding of cancer and the appropriate drug regimen to treat cancer has developed extensively in the last 100 years. Researchers have made great depths to understanding cancer as a genetic disease and we are using this knowledge to make more effective chemotherapeutic treatments. In addition, cancer’s genetic origins mean that molecular profile of a tumor will help indicate what therapy will be most effective. However, much research will be needed to develop the genetic models that will allow us to discover new therapy and use them in appropriate combinations with our current therapy. Additional studies are needed to decide what therapy should be used on different cancer types.

Section 1.2: RNA interference

RNA interference (RNAi) has now become an almost universally used tool for studying gene function. Since the initial studies by Andrew Fire and Craig Mello showed they could induce sequence specific gene silencing in *C. elegans* through the expression of double stranded RNA (dsRNA)¹², RNAi related pathways have been found in numerous organisms including plants, flies, yeast, humans and planaria.¹³ Since this landmark discovery 10 years ago, small RNAs have been implicated in regulation of gene expression, transposon silencing, and heterochromatin formation.¹⁴

1.2.1 RNA interference mechanism

The initial studies of long dsRNA mediated gene silencing transformed the fields of fly and worm genetics as it allowed scientists to decrease expression of specific genes both *in vivo* and *in vitro*. This eventually allowed both model organisms to be exploited for subsequent genome-wide RNAi screens. In the RNAi pathway, double stranded RNA (dsRNA) is processed by an RNaseIII enzyme named Dicer into 21 to 23 nucleotide small interfering RNAs (siRNAs).¹⁵ These products have 5' monophosphate ends and two-nucleotide 3' overhangs.¹⁶ One strand of each siRNA duplex is incorporated into the RNA induced silencing complex (RISC).¹³ It is then used as a guide to target and destroy the cell's messenger RNA in a sequence specific manner.¹⁷

A class of endogenous small RNA molecules called microRNAs (miRNAs) have emerged as playing an important role in regulating gene expression. These small 21-25 base pair noncoding RNAs were initially discovered in worms^{18, 19} but are present in numerous species including mammals. Class II RNA polymerases make long single stranded primary transcripts (pri-miRNA) that form a hairpin structure with significant

loops and bulges.^{20, 21} These pri-miRNA are processed in the nucleus by the RNase III enzyme, Drosha, into 70 nucleotide hairpin RNAs (pre-miRNA) which are transported into the cytoplasm by exportin-5.^{22, 23} Dicer cleaves the pre-miRNAs into 21 nucleotide small RNAs (miRNAs) that are loaded into RISC.²⁴ (see Figure 1.2).

While miRNAs and siRNAs utilize similar processing enzymes for their biogenesis, they induce gene silencing in different manners. miRNAs guide the RISC complex to the 3' untranslated regions (UTRs) of mRNA transcripts, but do not share

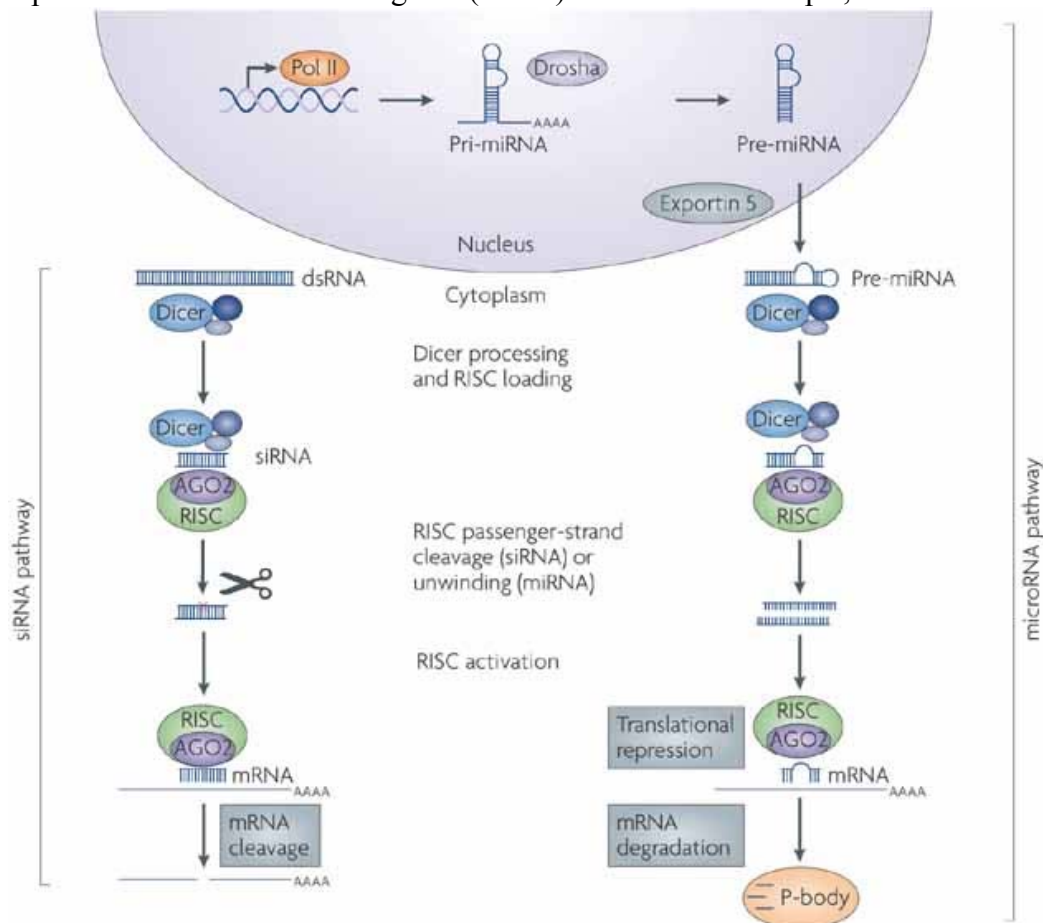


Figure 1.2 RNA interference. (RNAi) pathways are guided by small RNAs that include small interfering RNA (siRNA) and microRNAs (miRNAs). Both of these pathways share common enzymatic machinery, specifically, Dicer and Ago2. siRNAs mediate sequence specific mRNA cleavage and miRNAs direct translational repression and in some cases mRNA degradation. (Figure from de Fougères, A, et al, 2007).

exact complementarity with their targets.²⁵ Consequently, this leads to the repression of protein expression through mRNA destabilization or translational repression.²⁵ miRNAs

may mediate mRNA degradation in cytoplasmic compartments known as processing bodies (P-bodies).²⁶ In the siRNA pathway, the guide strand of each siRNA duplex is used as a template for cleavage of its exact complementary mRNA sequence by the RISC complex.²⁷

Argonaute 2 has been identified as the catalytic engine of the RISC complex.²⁷ Dicer and Argonaute proteins both contain a PAZ domain that they use to bind the single-stranded 3' ends of RNA in a sequence-independent manner.²⁸ This is particularly interesting as both Drosha and Dicer are RNase III enzymes which characteristically leave two nucleotide 3'-overhangs on their processed products.¹⁶ It is thought that Argonaute and Dicer distinguish small regulatory RNAs from degraded RNAs that are derived from non-related pathways by binding to their characteristic 3'-overhangs.²⁹ Argonaute also has a PIWI domain which contains the slicing site of the enzyme.³⁰ The PIWI domain shares structural similarities with the RNase H family of enzymes.³¹ The 5' end of the guide strand is likely to be anchored in a highly conserved binding pocket in the PIWI domain with the seed nucleotides (2–8) facing towards the target mRNA.³²

1.2.2 Challenges to RNA interference in mammalian cells

While dsRNA induced RNA interference was used extensively as a tool for knocking down genes in lower model organisms, dsRNA delivery could not be used as a method for gene silencing in mammalian cells for some time because of the innate interferon response. The RNA-dependent protein kinase (PKR) pathway of mammalian cells responds to dsRNA of greater than 30 nucleotides by phosphorylating EIF-2 α and nonspecifically arresting translation.^{33,34} This mechanism is thought to be a component of

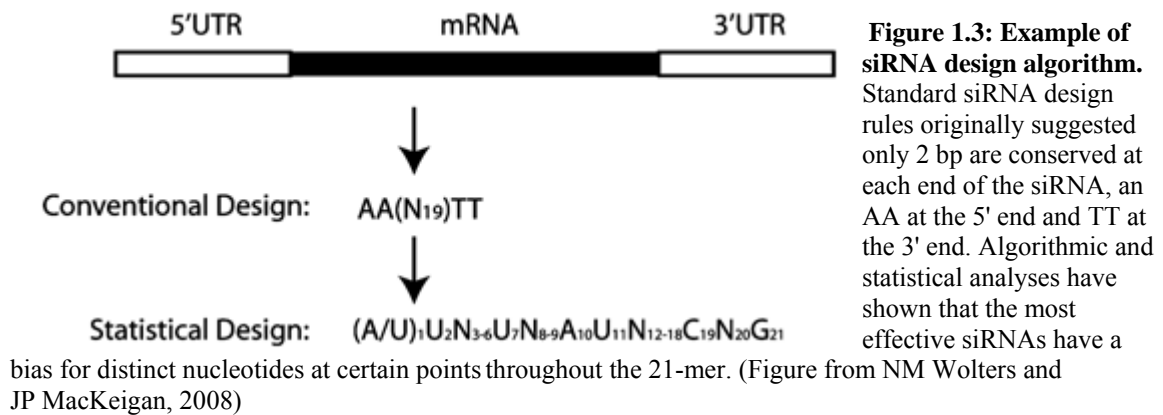
the mammalian immune system's defense against dsRNA viruses. Studies from the lab of Tom Tuschl showed that siRNAs themselves could be used to induce effective silencing in many mammalian cells by transient transfection.³⁵ These small RNAs are chemically synthesized mimics of Dicer products that are incorporated into RISC resulting in gene specific knockdown. Concurrently, work by Patrick Paddison indicated that exogenously produced short hairpin RNAs (shRNAs) could be used to knockdown a specific gene of interest in mammalian cells.³⁶ However, these hairpins are processed by Dicer before they are incorporated into the RISC complex. These methods worked because siRNAs and shRNAs are too small to trigger an interferon response. This landmark discovery gave scientists the ability to knockdown a gene of choice and study its function in mammalian cells. While previously, a knockout mouse could take months or years to breed, scientists could now study gene function in a matter of days. It is important to take into account though that mRNA levels are significantly depleted but not abolished. RNAi loss-of-function applications are termed 'knock-down' to distinguish them from 'knockouts' achievable by mutagenesis at the DNA level.³⁷ When a gene is silenced by RNAi, production of the target mRNA continues regardless of mRNA degradation and a certain level of translation is expected. Transfection of chemically synthesized RNAs is transient and mammals lack the molecular mechanisms that amplify silencing in worms and plants. However, RNAi has been made stable and heritable by enforced expression of the silencing trigger through the infection of retroviral expression cassettes.³⁸

1.2.3 Rational Design of RNAi triggers

Once siRNA duplexes could be used as a tool to specifically silence mammalian genes, it was thought that almost any siRNA could work effectively. In actuality, the efficacy of silencing depends on the small RNA sequence resulting in a spectrum of potency with only a subset of small RNAs being highly effective.³⁹ Because of this, concerns arose about identifying the siRNA or shRNA sequence that gave the most potent knockdown at the lowest possible concentration of antisense agent. Many different groups have pursued algorithms for predicting efficient sequences. It was discovered that small shifts in the position within the target mRNA were enough to change the degree of small RNA silencing in an unpredictable manner.³⁹ In addition, reports of the non-specific off-target effects of RNAi became apparent. It is thought these effects occur because of cross-hybridization of small RNAs containing partial identity to non-targeted transcripts causing unwanted silencing effects.⁴⁰ In addition, RNAi has been associated with stimulating nonspecific RNA degradation by RNase L, activating toll-like receptors and the retinoic acid-inducible gene-1 protein, both of which promote adverse immune reactions.⁴¹

Advancements in understanding the RNAi mechanism have shown that thermodynamic properties of the siRNA sequences and sequence asymmetry can help in identifying what sequence is loaded into the RISC complex and subsequently used for silencing.⁴² Our increased knowledge of the biology of RNAi together with the statistical analyses of libraries of siRNAs with experimentally determined efficiency has led to computer-based approaches that increase the likelihood of identifying effective and specific siRNAs.⁴⁰ Artificial neural networks have been utilized to train algorithms based on the analysis of randomly selected siRNAs and have deciphered that certain base pair

positions have a tendency to possess specific nucleotides perhaps due to the loading and unwinding of the small RNA duplex into RISC.⁴² While no algorithm is perfect, these rules have helped to ensure efficient and potent gene silencing while minimizing off-target effects (see Figure 1.3). It is important to recapitulate phenotypes from an RNAi experiment through the use of multiple small RNAs or through a rescue experiment where the phenotype is restored by re-expression of the gene of interest.



1.2.4 Development of small RNA libraries for functional genomics

The ability to perform genetic screens has been popularized in model organisms such as bacteria, yeast, worms and flies because of the molecular tools available to study gene function combined with advances in genomics, which has allowed the sequences of these organisms to become available to the public. The development of RNAi as a silencing tool for mammalian cells combined with the publication of the human genome has made it possible to produce large-scale RNAi libraries with the intention of assigning function to genes and outlining molecular pathways through the use of genetic screens. Multiple scientific groups and private companies have developed siRNA and shRNA libraries that have covered the human and mouse genome.⁴³⁻⁴⁵ The Hannon-Elledge library of Cold Spring Harbor Laboratory consists of over 100,000 shRNA expression

cassettes targeting over 30,000 human and mouse genes with approximately 3 shRNAs per gene (see Figure 1.4).⁴⁴ The expression cassettes are contained within multifunctional vectors that can be packaged into retroviruses and shuttled into customized vectors through bacterial mating.⁴⁶ These hairpins can be tracked in a mixed cell population by means of a random 60-mer DNA barcode. Retroviral and lentiviral based systems are available which can work well with cells that are difficult to transfect. Lentiviruses are capable of infecting both dividing and nondividing cells, which is beneficial for cell types that are extremely slow-growing or nondividing. While it is difficult to determine how much virus is necessary for substantial knockdown, care can be taken to achieve a multiplicity of infection that is normalized among the infected cells. shRNAs in the design of a primary miRNA have been shown to effectively silence genes even when present in the genome as a single copy.⁴⁷ In addition library hairpins can be moved into inducible vectors that give scientists greater freedom in developing complex assays.⁴⁸ Given our improved understanding of the biology behind RNA interference and the recent availability of the human genome sequence, we now have a formidable tool for studying functional genomics and attacking the molecular vulnerabilities of cancer.

Section 1.3: Functional Genomics

Scientists have been challenged to describe gene functions and interactions from the vast wealth of data produced by the Human Genome Project. The published human genome sequence has made it possible to expose molecular pathways relevant to development and disease. The genome can be an investigative tool for scientists to functionalize genetic events and study their relevance on a biochemical level.⁴² Efficient

interpretation of the functions of human genes requires resources and strategies to be developed to enable large-scale investigations across entire genomes. Recent advances in the field of RNA interference have allowed for large-scale loss-of-function studies in mammals. This is especially helpful to the study of cancer, which is caused by multiple genetic events. We can now create realistic and accurate models of cancer and examine tumor response to chemotherapy.

1.3.1 RNAi Genetic Screens

RNAi genetic screens hold much promise because they allow the use of reverse genetics in mammalian cells. Previously, forward genetics screens consisted of mutating genes at random and selecting candidates based on the display of a phenotype. The mutated gene would have to be mapped, which can be a time consuming process, and is not easily applicable to mammalian systems. In a reverse screen, the gene of interest is already known and is mutated or deleted. The function of the gene is inferred from the resulting phenotype. An example of classical reverse genetics is the creation of knockout cell lines. However, generating knockout cell lines to conduct a genome-wide screen in mammals is impractical. RNAi screens allow a fast and simple method of screening genes by repressing their expression.⁴⁹ Because of the ease of knocking down gene expression, RNAi screens can be automated or high-throughput. In addition, RNAi does not generate complete null alleles, but allows the creation of hypomorphs. This allows the study of essential genes, whose biology could not be investigated using true knockout technology.⁵⁰ Most interestingly, intrinsic differences between individual shRNA expression vectors targeting the same gene allows the generation of an 'epi-allelic series'

for dissecting gene function.⁵¹ This allows scientists to study gene dosage effects rather than complete abrogation of gene function.

RNAi screens can be performed in numerous ways (see Figure 1.5). Cells can be screened gene-by-gene in a multiwell format for activation or repression of a reporter in a cell-based or biochemical assay. These assays can range from simple readouts of immunofluorescence or colorimetric reporter assays to complex high-content screens involving microscopic image analysis. These usually involve the transient transfection of chemically synthesized individual siRNAs or shRNAs. Analysis of the phenotype is usually carried out by a robot in a high throughput manner. While this method reflects the ease of RNAi and allows rapid identification of the target gene, it can be quite costly to purchase synthetic small RNAs and lipid based reagents. In addition, multiple siRNAs for each gene must be used to decrease the probability of off-target effects, which can increase costs. It is important to normalize the transfection efficacy and titrate the siRNA concentration to the lowest nanomolar range, yet still allow sufficient knockdown. Determining the most potent siRNA concentration can help limit potential off-target effects.

Screens where cells are infected with pools of RNAi constructs are not as costly as screens involving individual transfection of synthetic RNAs. In addition, knockdown can be monitored over extended periods of time. Cells can be infected with pools of RNAi constructs followed by selection of individual colonies that can be scored phenotypically for easily observable alterations such as morphological changes.⁴⁶

Library vector

Vector description. Second generation: pSM2 (see Fig. 2; ref. 30).

Viral vector backbone: self-inactivating retroviral vector. Plasmid pSM2 is related to pSM1 (first-generation library vector; Fig. 2; ref. 27).

Major modifications: (i) in pSM2 the shRNA is embedded in an miR-30 cassette, (ii) pSM2 contains elements to swap the miR-30 expression cassette (with the 60-nt bar code and microbial selection gene but without the U6 promoter) into recipient plasmids containing other promoters by mating-assisted bacterial recombinational cloning.

Additional features: in both pSM1 and pSM2, the shRNA expression cassette is under the control of the U6 pol III promoter. The vectors also contain unique 60-nt bar-code sequences.

Library stability. Under the recommended handling and growth conditions, the second-generation library clones are very stable and more stable than first-generation clones. Upon sequence verification, second-generation library plasmids are recovered and retransformed into a low recombination bacteria host. We evaluated the recombination rate in a random sampling of clones by monitoring DNA quality by electrophoresis and observed recombination in ~5% of clones (see <http://www.openbiosystems.com>).

Infectivity. A variety of human and mouse cell lines have been tested with pSM2 virus: 293T, 293, MCF10A, NIH3T3, IMR90 and derivatives, as well as a variety of cancer cell lines.

Silencing. No data available.

Library production

Vector DNA production yield. Both first and second-generation vectors contain the π protein-dependent origin of replication (*RR67*). Yield is typically less than with pUC origin plasmids. Yield can be increased by using *pir1* strains and/or growing in rich medium. There are other variables that influence DNA yield, of note is the method of DNA extraction and reagent quality. If grown in normal medium (LB) yield is 2–3 $\mu\text{g}/\text{ml}$ (after 16–18 h at 37 °C), whereas in rich medium (GS96 or 2XYT), the yield is 5–7 $\mu\text{g}/\text{ml}$.

Viral titer. Production conditions: Typically, 12 μg of library DNA along with helper plasmids (4 μg of packaging helper plasmid, and 8 μg of pVSV.G expressing the pseudotyping envelope protein) are transfected into packaging cells (LinX) plated on a 10-cm dish (cell confluence of 50–70%), using lipid carriers. Two virus harvests (filtration-clarified culture supernatant) are collected at 48 h and 72 h after transfection. The two harvests can be pooled if necessary but for most applications, virus from a single harvest is sufficient. Viral titer is determined by infection of human (HEK293) and mouse (NIH3T3) cells by counting of antibiotic-resistant transduced colonies. These cells are highly infectable and thus serve as a standard for comparing infectivity with other cells. Typical virus titers for the second-generation library are $\sim 1\text{--}1.3 \times 10^6$ I.U./ml for both HEK293 and NIH3T3 cells.

Library coverage

As of June 2006, the second-generation library contains shRNAs directed against:

32,216 human genes (87,283 shRNAs)
30,629 mouse genes (76,896 shRNAs)

Gene coverage distribution. Average: 2–3 shRNAs per target gene. For human genes:

18% of target genes (5,800 genes) covered by exactly 1 shRNA
26% of target genes (8,362 genes) covered by 2 shRNAs
34% of target genes (11,056 genes) covered by 3 shRNAs
22% of target genes (6,998 genes) covered by ≥ 4 shRNAs

For mouse genes:

19% of target genes (5,740 genes) covered by exactly 1 shRNA
31% of target genes (9,359 genes) covered by 2 shRNAs
36% of target genes (10,983 genes) covered by 3 shRNAs
15% of target genes (4,547 genes) covered by ≥ 4 shRNAs

Region of transcripts targeted: open reading frame.

List of target genes: available online (<http://codex.cshl.edu/> and <http://www.openbiosystems.com>)

shRNA sequences. The shRNA were designed using a proprietary algorithm of Rosetta Inpharmatics.

A list of shRNA sequences is available online (<http://codex.cshl.edu/> and <http://www.openbiosystems.com>).

Functional tests

Sequence verification. Sequencing success rate is 35–55% (sequencing of one strand). In the second-generation library, 87,283 and 76,896 shRNA clones against human and mouse genes, respectively, were sequence-verified. In the first-generation library, 28,659 and 9,119 shRNA clones against human and mouse genes, respectively, were sequence-verified.

Validation data. Validation is only available for a small proportion of clones as a large-scale effort to systematically evaluate each individual shRNA for either collection is not currently planned. Rather, validation of candidates from functional screens have been and will continue to be the means in which library clones will be validated for the time being.

1. First- and second-generation library clones were validated in a proteasome function screen^{28,30}.
2. Genetic screen for candidate tumor suppressor identified REST (first-generation library)³¹.
3. Probing tumor phenotypes using stable and regulated synthetic miRNA precursors (second-generation library)³⁷.
4. Identification and validation of oncogenes in liver cancer³⁹.
5. Validation in context of other vector^{36,37}.

Nonspecific off-target effect. In a test of several second-generation vectors, no induction of the interferon response was observed, as monitored by EIF2- α .

Distribution

Glycerol stocks (arrayed in 96-well format) and plasmid DNA of individual clones and pools of clones are available from Open Biosystems.

Sets of individual clones in glycerol stock are also available grouped by functions of targets (see Table 1).

Figure 1.4: Features of the Hannon Elledge Library. The library can be accessed at

<http://codex.cshl.edu>⁴⁴ (Figure taken from Chang, K, 2006).

Infected pools of RNAi cells can also be monitored via sequencing based or microarray based deconvolution. These methods are becoming increasingly attractive as the rise of “deep sequencing” techniques make it cheaper and easier to sequence many short sequences. A method developed in our lab has used the availability of DNA 'barcodes' contained within the shRNAs to identify target genes after phenotypic selection.⁴⁴ These barcodes can then be monitored to reflect the relative abundance of individual shRNA vectors in a large population. This property allows us to identify shRNA targets that are both selected for (positive selection) or selected against (negative selection) (See Chapter 4).

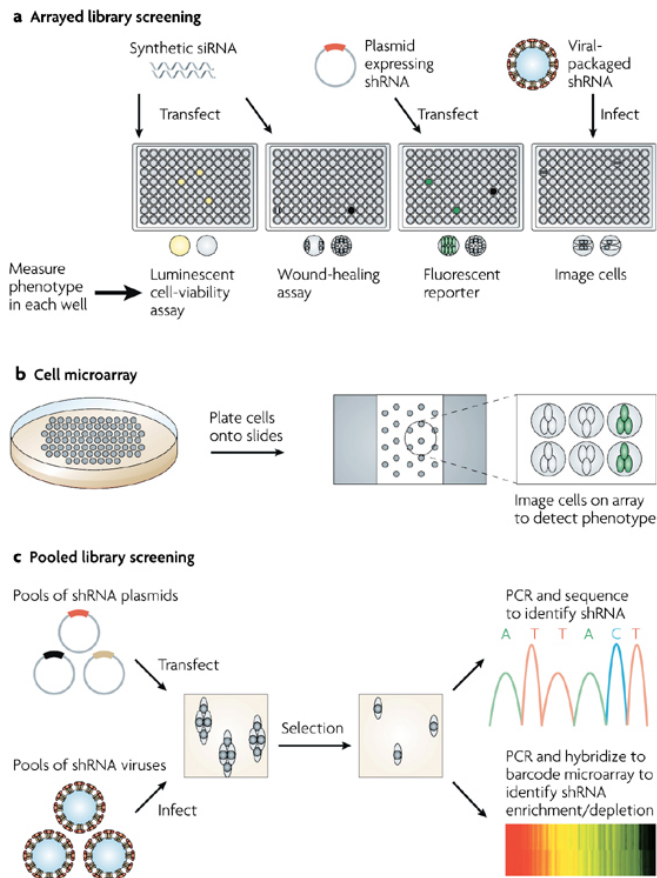
Much success with RNAi screens have come in the form of positive selection screens, where the relevant phenotype results from being resistant to an applied stress. The surviving population would be greatly enriched for relevant hairpins relative to the starting population.⁵⁰ Negative selection screens are more complicated as they require identifying cells that are lost from a population. Screens like these can illuminate genes that cause lethality to a cell when placed in a stressful situation. They also have the potential to identify genes that are lethal only in combination with a drug or deletion of another gene (synthetic lethality).⁵² (See Chapter 5).

Currently available libraries can be organized into subsets of gene families based on gene ontology.⁴² This allows researchers to choose gene subsets from an RNAi library that will reflect the goals of the screen. Care should be taken so that gene numbers are not limited to a level that diminishes potential novelty of the results. It is important to survey relevant genes and genes of unknown function. RNAi screens will deliver a number of positive hits and the quality of these hits will depend on the cut-off stringency and

statistical relevance. While there is no set cut-off, some screens will accept a hit that is more than three standard deviations above or below the mean value. However, there is no standard data analysis method or candidate selection criteria for RNAi screens. In addition, most groups validate results with a secondary screen or follow up cell-based assay. It is important to show that two or three sequences for each candidate show adequate knockdown and recapitulate the same phenotype.

A number of RNAi screens have been published in recent years that have allowed scientists to associate genes with molecular pathways and cell death processes. This new information has greatly increased our knowledge of the inner workings of the cell. Many groups have focused on issues regarding human disease such as cancer. Most of the early RNAi screens in mammalian cells examined a small subset of genes in a multiwell format. Despite the relative small scale of these screens, they still yielded a wealth of information. Aza-Blanc et al used an siRNA library of a 500 hairpins to identify regulators of TRAIL-induced apoptosis.⁵³ Since then, RNAi based screens have identified multiple tumor suppressors^{54, 55} as well as kinases associated with chemoresistance.⁵⁶ Scientists were able to scale up the size and complexity of screens by using pooled hairpins and tracking them with DNA barcodes.^{57, 58} These screens illustrate the potential for genome-scale examinations of mammalian cells. Positive selection screens have been very successful at probing cancer relevant pathways. Some have relied upon shRNA sets targeting the 'druggable' genome such as kinases and phosphatases with the hope that these would make good chemotherapeutic targets.⁵⁹ Negative selection screens, while more complicated, have also shed new insight into the vulnerabilities of human cells. The first successful study was conducted by Louis Staudt who showed that knockdown of

genes in the NF-kappaB pathway differentially affected two subtypes of B-cell lymphoma.⁶⁰



Nature Reviews | Drug Discovery

Figure 1.5:RNAi screening approaches. a. Arrayed screens target each gene separately in a multi-well format. b. Arrayed screening by RNAi cell microarrays. c. Pooled library screens target many genes. (Figure from Iorns, et al 2007)

RNAi screens are also illuminating genes whose knockdown synergize with current cancer therapeutics. For example, a recent screen by Whitehurst et al looked for enhancers of paclitaxel, a popular chemotherapeutic that destabilizes microtubules.⁶¹ By using a one well/one gene approach, they found a panel of 87 genes that when knocked down, increase sensitivity of lung cancer cells to paclitaxel.⁶¹ These types of screens for modifiers of a particular compound have also been shown to work with shRNA libraries.

A recent shRNA library based barcode screen for resistance genes to Herceptin, an antibody that is antagonistic to the VEGF receptor, identified components of the PTEN pathway.⁶² These types of successful screens bring forth information that can be used to select the proper chemotherapy for cancer patients with the hope of eventually leading to personalized treatment.

siRNA/shRNA libraries are being expanded to entire genome sets in a variety of mammalian organisms. While we don't have the ability to absolutely guarantee each silencing trigger will knockdown to the utmost efficacy, our knowledge of the biological pathways governing RNAi are allowing us to make better silencing tools that have less side effects. These libraries can be shuttled to a variety of stable and inducible vectors that are available in both a lentiviral and retroviral format. Our screening capabilities have improved where both negative and positive selection screens can be performed targeting thousands of genes at a time. These have been aided by advancements in both deep sequencing and microarray technology. Indeed, RNAi screens have the ability to reveal the secrets of the genome.

Section 1.4: RNAi Therapeutics

RNAi screens have enabled scientists to identify genes that would make good chemotherapeutic targets, either by the creation of a small molecule inhibitor or antibody mediated therapy. However, there is potential to use RNA mediated gene silencing as a therapy for patients. A number of groups have been able to express gene knockdown in mice in both a stable and transient way.⁴⁸ *In vivo* studies using chemically modified short interfering RNAs (chol-siRNAs) can systemically silence an endogenous gene after

intravenous injection in mice.⁶³ siRNAs have even been shown to inhibit viruses after administration in mice intranasally with or without transfection reagents.⁶⁴ It is possible to harness RNA interference as a therapy by using a viral vector to express shRNA or by introducing siRNAs into a human. While viral delivery of shRNA is still being developed, synthetic siRNAs are currently in clinical trials. Though naked siRNAs are degraded in human plasma in a couple of minutes, chemical modifications have been made to protect siRNAs from endo- and exonucleases. Introduction of a phosphorothioate backbone linkage at the 3' end and 2' sugar modifications (such as 2'-O-methyl and 2'-fluoro) stabilize the siRNA duplex while maintaining silencing activity.⁶⁵ These RNAs can be delivered either systemically or locally as cholesterol conjugates, liposomes, or complexed with peptides, polymers or antibodies.⁶⁵ Direct delivery of small RNAs has the advantage of needing a relatively lower dose to elicit potent gene silencing responses without undesired systemic side effects. This is one of the reasons RNAi therapeutics have been successful in the human eye and lung.⁶⁶ In addition to the easily accessible local delivery by physicians, the eye and lung are composed of mostly nondividing cells.⁶⁷ Indeed, three different siRNAs have reached different phase level clinical trials. Two of these, bevasiranib and Allergan's AGN211745 are used in the treatment of wet-age related macular degeneration and the third, ALN-RSV01 is used in the treatment of Respiratory Syncytial Virus infections. Bevasiranib is currently in a Phase III trial with results expected in mid-2009. There is one siRNA based drug targeting solid and metastatic tumors currently in preclinical/phase I clinical trials. CALAA-01 is an siRNA packaged nanoparticle that targets the M2 subunit of ribonucleotide reductase.⁶⁷

Pharmaceutical companies are also interested in making use of the therapeutic value of miRNAs. AntagomiRs, modified antisense oligonucleotides, have been developed to block miRNA expression directly and specifically.⁶⁸ It is unclear precisely how they work, though they may work by blocking mature miRNAs from active pools or binding to miRNA precursors and inhibiting the biogenesis of mature miRNAs. While the details of antagomiR delivery and enzymatics of cellular processing are still being developed in mice, they do show great therapeutic promise. miRNAs have been shown to act as tumor suppressors⁶⁹ or oncogenes⁷⁰ by regulating cancer related genes. For example, the mir-17-92 cluster has been shown to be overexpressed in human B cell lymphomas and also cooperate with Myc to drive lymphoma formation in mice.⁷⁰ This may be a reasonable target of antagomiR-based therapy.⁶⁷

The ability to use RNAi as a tool has not stopped in the laboratory, as clearly it is making great strides in the clinic as a therapeutic tool. We have yet to see the direct handoff from RNAi genetic screen to RNAi therapeutics as both fields are in their infancy. However, there are examples of RNAi screens that have dramatic clinical implications. For example, Brummelkamp et al screened 50 de-ubiquitinating enzymes involved in tumor formation and found that suppression of CYLD, (the cylindromatosis tumor suppressor gene) enhanced activation of the NF-kB pathway.⁵⁹ CYLD is a gene that is mutated in familial cylindromatosis causing tumors to develop principally on the head and neck.⁷¹ CYLD was found to have a similar effect on NF-kB levels as aspirin, even though they work on different parts of the pathway. It was thought that aspirin may compensate for CYLD dysfunction and thus, clinical trials of topical application of aspirin to tumors promptly began.⁷² This is an example of how the results of a simple

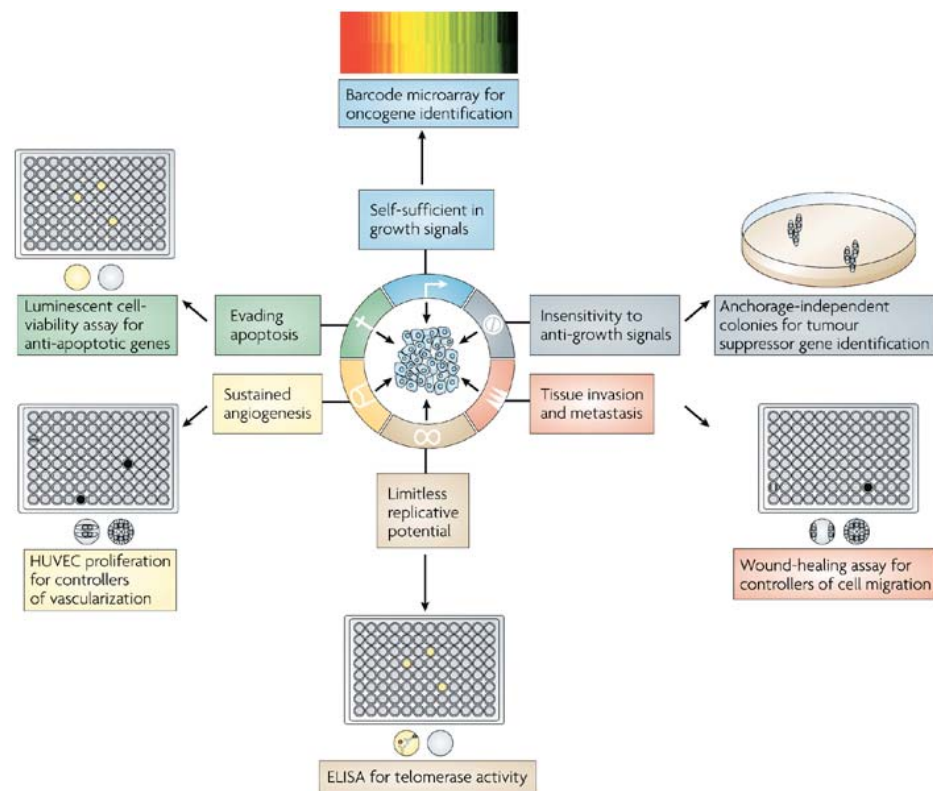
RNAi genetic screen influenced patient treatment. Surely, as our siRNA therapeutic technology advances we will be able to translate our laboratory findings directly into patient cures.

Section 1.5: Questions addressed in this Thesis

This thesis work illustrates how technological advances of RNAi tools have evolved in lock-step with our increased knowledge of the biology of RNA interference. Our deeper understanding of Dicer processing and RISC silencing has empowered us to make more potent and specific gene silencing triggers (Chapter 2). It seems fortuitous for scientists to have stumbled upon a biological phenomenon that can be exploited for exploring gene function at a time when the human genome sequence reached its completion. This technology has allowed us to build a massive library of shRNAs that have the potential to address questions relating to a multitude of human diseases (Chapter 3). The hairpins in this library are based on an endogenous primary miRNA transcript and have silencing algorithms incorporated in their design. However, a proper technological platform needed to be developed for large scale genomic studies using pools of hairpins (Chapter 4). The optimization of our microarray technology made it possible for both negative and positive selection screens to be accomplished. I used these methods to conduct a synthetic lethal screen for discovering chemotherapeutic targets that will cause p53 null cells to die in the presence of doxorubicin (Chapter 5).

As a genetic tool, RNAi holds much promise for both understanding and treating genetic diseases, such as cancer. This tool enables scientists to manipulate specific genes resulting in better cancer models. The development of RNAi screening

technology allows the identification of genes that cause the well-defined phenotypic hallmarks of cancer (see Figure 1.6). We can expose the genetic vulnerabilities of cancer through RNAi. This work has spanned great advances and discoveries in both RNA biology and genomics. The complex world of small RNAs as regulators of the genome has warranted the Nobel prize. The sequencing of the genome and our better understanding of the cell's biological processes are empowering scientists to conquer the most formidable of human diseases.



Nature Reviews | Drug Discovery

Figure 1.6: Using RNAi screens to identify targets controlling the hallmarks of cancer.
(from Iorns et al, 2007)

Literature Cited

1. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* 100, 57-70 (2000).
2. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nat Med* 10, 789-99 (2004).
3. Papac, R. J. Origins of cancer therapy. *Yale J Biol Med* 74, 391-8 (2001).
4. Frei, E., 3rd et al. The effectiveness of combinations of antileukemic agents in inducing and maintaining remission in children with acute leukemia. *Blood* 26, 642-56 (1965).
5. Chabner, B. A. & Roberts, T. G., Jr. Timeline: Chemotherapy and the war on cancer. *Nat Rev Cancer* 5, 65-72 (2005).
6. Sawyers, C. Targeted cancer therapy. *Nature* 432, 294-7 (2004).
7. Roche-Lestienne, C. et al. Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. *Blood* 100, 1014-1018 (2002).
8. Shah, N. P. et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* 2, 117-25 (2002).
9. Sordella, R., Bell, D. W., Haber, D. A. & Settleman, J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* 305, 1163-7 (2004).
10. van 't Veer, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530-6 (2002).
11. Agendia, Inc. Mammaprint Brochure
http://usa.agendia.com/en/for_healthcare_professionals_2.html
12. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-11 (1998).
13. Hannon, G. J. RNA interference. *Nature* 418, 244-51 (2002).
14. Birchler, J. A. & Kavi, H. H. Molecular biology. Slicing and dicing for small RNAs. *Science* 320, 1023-4 (2008).
15. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363-6 (2001).

16. Carmell, M. A. & Hannon, G. J. RNase III enzymes and the initiation of gene silencing. *Nat Struct Mol Biol* 11, 214-8 (2004).
17. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293-6 (2000).
18. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843-54 (1993).
19. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858-62 (2001).
20. Lee, Y. et al. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* 23, 4051-60 (2004).
21. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* 431, 343-9 (2004).
22. Han, J. et al. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18, 3016-27 (2004).
23. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415-9 (2003).
24. Song, J. J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* 10, 1026-32 (2003).
25. He, L. & Hannon, G. J. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 5, 522-31 (2004).
26. Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* 7, 719-23 (2005).
27. Liu, J. et al. Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science* 305, 1437-1441 (2004).
28. Tolia, N. H. & Joshua-Tor, L. Slicer and the Argonautes. *Nat Chem Biol* 3, 36-43 (2007).
29. Song, J.-J. & Joshua-Tor, L. Argonaute and RNA -- getting into the groove. *Current Opinion in Structural Biology* 16, 5-11 (2006).

30. Faehnle, C. R. & Joshua-Tor, L. Argonautes confront new small RNAs. *Current Opinion in Chemical Biology* 11, 569-577 (2007).
31. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 9, 22-32 (2008).
32. Parker, J. S. & Barford, D. Argonaute: a scaffold for the function of short regulatory RNAs. *Trends in Biochemical Sciences* 31, 622-630 (2006).
33. Davis, S. & Watson, J. C. In vitro activation of the interferon-induced, double-stranded RNA-dependent protein kinase PKR by RNA from the 3' untranslated regions of human alpha-tropomyosin. *Proc Natl Acad Sci U S A* 93, 508-13 (1996).
34. Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H. & Williams, B. R. Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* 5, 834-9 (2003).
35. Elbashir, S. M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411, 494-8 (2001).
36. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. & Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 16, 948-58 (2002).
37. Sarov, M. & Stewart, A. F. The best control for the specificity of RNAi. *Trends Biotechnol* 23, 446-8 (2005).
38. Paddison, P. J., Caudy, A. A. & Hannon, G. J. Stable suppression of gene expression by RNAi in mammalian cells. *Proc Natl Acad Sci U S A* 99, 1443-8 (2002).
39. Pei, Y. & Tuschl, T. On the art of identifying effective and specific siRNAs. *Nat Meth* 3, 670-676 (2006).
40. Jackson, A. L. et al. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotech* 21, 635-637 (2003).
41. Barik, S. RNAi in moderation. *Nat Biotech* 24, 796-797 (2006).
42. Wolters, N. M. & MacKeigan, J. P. From sequence to function: using RNAi to elucidate mechanisms of human disease. *Cell Death Differ* 15, 809-19 (2008).
43. Bernards, R., Brummelkamp, T. R. & Beijersbergen, R. L. shRNA libraries and their use in cancer genetics. *Nat Methods* 3, 701-6 (2006).

44. Chang, K., Elledge, S. J. & Hannon, G. J. Lessons from Nature: microRNA-based shRNA libraries. *Nat Methods* 3, 707-14 (2006).
45. Root, D. E., Hacohen, N., Hahn, W. C., Lander, E. S. & Sabatini, D. M. Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Meth* 3, 715-719 (2006).
46. Silva J, C. K., Hannon GJ, Rivas FV. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. *Oncogene* 51, 8401-9. (2004).
47. Dickins, R. A. et al. Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nat Genet* 37, 1289-95 (2005).
48. Dickins, R. A. et al. Tissue-specific and reversible RNA interference in transgenic mice. *Nat Genet* 39, 914-21 (2007).
49. Silva, J., Chang, K., Hannon, G. J. & Rivas, F. V. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. *Oncogene* 23, 8401-9 (2004).
50. Burgess, D. J. RNAi Genetic Screens for Chemotherapy Response Modifier Genes. (PhD dissertation, Watson School of Biological Sciences) (2007).
51. Hemann, M. T. et al. An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nat Genet* 33, 396-400 (2003).
52. Kaelin, W. G., Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer* 5, 689-98 (2005).
53. Aza-Blanc, P. et al. Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. *Mol Cell* 12, 627-37 (2003).
54. Westbrook, T. F. et al. A genetic screen for candidate tumor suppressors identifies REST. *Cell* 121, 837-48 (2005).
55. Kolfschoten, I. G. et al. A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. *Cell* 121, 849-58 (2005).
56. MacKeigan, J. P., Murphy, L. O. & Blenis, J. Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. *Nat Cell Biol* 7, 591-600 (2005).
57. Berns, K. et al. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* 428, 431-7 (2004).

58. Paddison, P. J. et al. A resource for large-scale RNA-interference-based screens in mammals. *Nature* 428, 427-31 (2004).
59. Brummelkamp, T. R., Nijman, S. M., Dirac, A. M. & Bernards, R. Loss of the cylindromatosis tumour suppressor inhibits apoptosis by activating NF-kappaB. *Nature* 424, 797-801 (2003).
60. Ngo, V. N. et al. A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* 441, 106-110 (2006).
61. Whitehurst, A. W. et al. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature* 446, 815-9 (2007).
62. Berns, K. et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 12, 395-402 (2007).
63. Soutschek, J. et al. Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* 432, 173-8 (2004).
64. Bitko, V., Musiyenko, A., Shulyayeva, O. & Barik, S. Inhibition of respiratory viruses by nasally administered siRNA. *Nat Med* 11, 50-5 (2005).
65. de Fougerolles, A., Vornlocher, H. P., Maraganore, J. & Lieberman, J. Interfering with disease: a progress report on siRNA-based therapeutics. *Nat Rev Drug Discov* 6, 443-53 (2007).
66. Tara, M. L., Howell, F. M. & Carl, D. N. Not miR-ly small RNAs: Big potential for microRNAs in therapy. *The Journal of allergy and clinical immunology* 121, 309-319 (2008).
67. Krutzfeldt, J. et al. Silencing of microRNAs in vivo with 'antagomirs'. *Nature* 438, 685-689 (2005).
68. He, L. et al. A microRNA component of the p53 tumour suppressor network. *Nature* 447, 1130-1134 (2007).
69. He, L. et al. A microRNA polycistron as a potential human oncogene. *Nature* 435, 828-833 (2005).
70. Bignell, G. R. et al. Identification of the familial cylindromatosis tumour-suppressor gene. *Nat Genet* 25, 160-165 (2000).
71. Lakhani, S. R. Putting the Brakes on Cylindromatosis? *N Engl J Med* 350, 187-188 (2004).

72. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs Exhibit Strand Bias. *Cell* 115, 209-216 (2003).
73. Carmell, M. A., Xuan, Z., Zhang, M. Q. & Hannon, G. J. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* 16, 2733-42 (2002).
74. Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear export of microRNA precursors. *Science* 303, 95-8 (2004).

Chapter 2

Synthetic shRNAs as highly potent RNAi triggers

Despina Siolas^{1,2}, Cara Lerner³, Julja Burchard³, Wei Ge³, Peter Linsley³, Patrick Paddison², Gregory J. Hannon^{2*} and Michele A. Cleary^{3*}

The biochemistry experiments in this chapter were performed by Despina Siolas. The microarray profiling and gene suppression studies were conducted by the scientists above at Rosetta Inpharmatics under the direction of Michele Cleary.

¹ Program in Genetics
Stony Brook University
Stony Brook, NY 11794

² Cold Spring Harbor Laboratory
Watson School of Biological Sciences
1 Bungtown Road
Cold Spring Harbor, NY 11724

³Rosetta Inpharmatics LLC,
A Wholly Owned Subsidiary of Merck & Co., Inc.
401 Terry Ave North
Seattle, Washington 98109, USA

* to whom correspondence should be addressed
hannon@cshl.edu
Michele_cleary@merck.com

Designing highly potent silencing triggers is key to successful application of RNAi in mammals. Recent studies suggested that assembly of RNAi effector complexes (RISC) is coupled to Dicer cleavage. We therefore examined whether transfection of optimized Dicer substrates might give an improved RNAi response. Dicer cleavage of chemically synthesized short hairpin RNAs (shRNAs) with 29 bp. stems and 2 nt. 3' overhangs produced predictable homogeneous small RNAs comprising the 22 bases at the 3' end of the stem. Consequently, direct comparisons of synthetic siRNAs and shRNAs that yield the same small RNA became possible. We found synthetic, 29mer shRNAs to be more potent inducers of RNAi than siRNAs. Maximal inhibition of target genes can be achieved at lower concentrations and silencing at 24 hours is often greater. These studies provide an improved methodology for triggering experimental silencing via the RNAi pathway.

2.1 Introduction

Many eukaryotic organisms respond to double-stranded RNA (dsRNA) by activating a sequence-specific silencing pathway, known as RNA interference or RNAi. Initiation of RNAi occurs upon processing of double-stranded RNAs into ~22nt fragments, known as siRNAs¹⁻³, by an RNaseIII family nuclease, Dicer⁴. These small RNAs are used as guides for selection and cleavage of complementary mRNAs through their incorporation into an effector complex, RISC,^{2,3,5,6} whose catalytic subunit, Argonaute 2 has recently been identified^{7,8}. These mechanistic insights led to the development of a methodology for experimentally programming the RNAi machinery in mammalian cells by directly transfecting chemically synthesized siRNA duplexes of ~21 nt. consisting of 19 paired bases with 2 nucleotide 3' overhangs to produce a transient silencing response⁹⁻¹¹.

In many organisms, the RNAi machinery also serves an effector function for endogenous, non-coding RNAs, known as microRNAs (miRNAs) (reviewed in¹²). MiRNAs are initially generated as long primary transcripts (pri-miRNA), which are cleaved in the nucleus by another RNaseIII family nuclease, Drosha¹³. The liberated pre-miRNAs are exported to the cytoplasm where Dicer performs a second cleavage to produce small RNAs that are loaded into RISC¹⁴⁻¹⁷. In the case of miRNAs, the cleavage sites are specific, and most often a single, discrete sequence is liberated from the precursor (reviewed in¹²). These discoveries prompted the development of a second approach for triggering RNAi in mammalian cells using DNA vectors encoding short hairpin RNAs (shRNAs), modeled roughly after endogenous microRNAs¹⁸⁻²³.

Remarkably, for both miRNAs and siRNAs, the two strands of the processed dsRNA are treated unequally. Cloning efforts in a variety of organisms yielded overwhelmingly one strand for each miRNA²⁴⁻²⁶. A potential explanation for this outcome came from biochemical studies of siRNAs in *Drosophila* suggesting that relative thermodynamic instability at the 5' end of a given strand of the Dicer product favors its loading into RISC²⁷. This is in accord with analysis of predicted Dicer cleavage products of endogenous miRNAs^{28,29} and with studies of the efficacy of large numbers of siRNAs, which indicate greater suppression if the antisense strand (relative to the target mRNA) has an unstable 5' end⁷³. Recent reports have suggested that this loading might occur in a complex and be coordinated with Dicer cleavage³⁰⁻³². A possibility suggested by these mechanistic insights is that Dicer substrates might be more efficiently incorporated into RISC than siRNAs. We therefore sought to understand how Dicer processes shRNAs in order to permit comparison of the efficiency of silencing triggers that are predicted to produce equivalent RISC enzymes.

2.2 Results

2.2.1 Dicer cleaves a predominant small RNA from the end of each shRNA

We began by producing ~70 chemically synthesized shRNAs, targeting various endogenous genes and reporters. We focused on a detailed analysis of one set of four shRNAs that target firefly luciferase (Fig. 2.1a). The individual species differed in two distinct ways. First, the stems of the shRNAs were either 19 or 29 bp in length. These sizes were chosen to reflect the two stem sizes most commonly used for vector-expressed shRNAs. Second, each shRNA either contained or lacked a 2 nt. 3' overhang, identical

to that produced by processing of pri-miRNAs by Drosha. Each species was end-labeled by enzymatic phosphorylation and incubated with recombinant human Dicer. The 29mer shRNA bearing the 3' overhang was converted almost quantitatively into a 22 nt. product by Dicer (Fig. 2.1b). In contrast, the 29mer shRNA that lacked the overhang generated very little discrete 22 nt. labeled product, despite a Dicer-dependent depletion of the starting material. Neither 19mer shRNA was cleaved to a significant extent by the Dicer enzyme. This result was not due to the lack of double stranded structure in the 19mer shRNAs, as all shRNA substrates were efficiently cleaved by bacterial RNaseIII (Supplementary Figure 2.1). These results suggested that the shRNAs bearing a 3' overhang produced predominantly one specific and unique small RNA product, while a blunt ended hairpin was processed into a range of products. This hypothesis was consistent with parallel analysis of identical shRNA substrates that were produced by *in vitro* transcription with T7 polymerase and uniformly labeled (Fig. 2.1c). Cleavage of the uniformly labeled hairpin with an overhang resulted in products accumulating from 29mer shRNAs both with and without overhangs. In the latter case, the products did not accumulate to the same degree as seen with the overhang-containing material. Additionally, shRNAs with overhangs yielded products of two discrete sizes (21 and 22 nt.). Considered together, our results suggest that Dicer requires a minimum stem length for efficient cleavage. Furthermore, they are consistent with a hypothesis that the presence of a correct 3' overhang enhances the efficiency and specificity of cleavage, directing Dicer to cut ~22 nt. from the end of the substrate.

A number of previous studies have suggested that Dicer might function as an end-recognizing endonuclease, without positing a role for the 3' overhang. Blocking of the

ends of dsRNAs using either fold-back structures or chimeric RNA-DNA hybrids attenuated, but did not abolish, the ability of human Dicer to generate siRNAs⁷⁴. Lund and colleagues suggested that Dicer cleaved ~22 nt from the blunt end of an extended pre-miRNA, designed in part to mimic a pri-miRNA (see⁷⁵). Structural analysis of the Argonaute 2 PAZ domain suggested that it engages very short (~2-3 nt.) stretches of the 3' ends of single-stranded RNAs³⁵⁻³⁹. This led Song and colleagues to propose a model in which the 3' overhangs of pre-miRNAs, generated by Drosha cleavage, would serve as an important recognition and specificity determinant for subsequent processing by Dicer³⁸. The results presented here are consistent with this model and suggest further that the 3' overhang aids in determining the specificity of cleavage, directing processing to a site 22 nt. from the 3' end of the substrate. These findings are in full accord with a recently published model for Dicer action⁴⁰.

To validate our biochemical analysis, we also mapped the position of Dicer cleavage *in vivo* using primer extension. Precursors were transfected into cells, and the processed form of each was isolated by virtue of its co-immunoprecipitation with co-expressed myc-tagged human Argonaute proteins, Ago1 and Ago2. The 29mer shRNA with an overhang gave rise to a relatively discrete product of 20 nt. as predicted for a cleavage 22 nt. from the 3' end of the substrate. Primer extension suggested identical cleavage specificities upon exposure of shRNAs to Dicer *in vitro* and in living cells (Fig. 2.2a). Control experiments using a luciferase 29mer shRNA alone (without myc-tagged Ago1 or Ago2 expression) or cells transfected with myc-tagged Ago1 or Ago2 alone (no shRNA) did not yield extension products (Fig. 2.2b).

Although the inability of Dicer to effectively cleave shRNAs with 19 bp. stems may seem at odds with the effective use of such structures for triggering RNAi using vector-based expression, there is presently no evidence that these RNAs require Dicer for their action. Indeed, our results using RNAi to deplete Dicer from cells suggests a strong dependence of silencing on Dicer for shRNAs with 29 bp. stems with a reduced or lack of dependency on Dicer for shRNAs with 19 bp. stems (not shown). However, 19mer shRNA do enter RISC. Human 293 cells that constitutively express Ago 1 were transfected with siRNAs, 29mer shRNAs or 19mer shRNAs. RISC was recovered by immunoprecipitation and associated RNAs were examined by Northern blotting. (Supplementary Figure 2.2) The 29mer shRNA with an overhang and the 22mer siRNA both entered RISC, giving 22 nt. small RNAs. The 19mer shRNA also entered RISC but gave 2 distinct small RNAs of 21 and 23 nt. While we do not understand the mechanistic basis for this observation, it may reflect Dicer-independent cleavage of the 19mer shRNA in the loop by a single-strand specific ribonuclease.

2.2.2 shRNAs are generally more effective than siRNAs

Since we could predict which single, specific 22 nt. sequence would be incorporated into RISC from a given shRNA, we could directly compare the silencing efficiency of shRNAs and siRNAs. Toward this goal, we selected 43 sequences targeting a total of 5 genes (3-9 sequences per gene). For each sequence, we synthesized a 21mer siRNA (19 bp stem) and shRNAs with 19 or 29 bp. stems that were predicted to give Dicer products that were either identical to their corresponding siRNAs or that differed by the addition of one 3' nucleotide homologous to the target. Importantly, each

was predicted to give precisely the same 5' end following cleavage of a 22mer RNA from the shRNA (N in Supplementary Figure 2.3). Sequences for siRNAs are provided in Supplemental Table 2.1. Each RNA species was transfected into HeLa cells at a relatively high concentration (100 nM). The level of suppression was determined by semi-quantitative RT-PCR of RNA from HeLa cells 24 hours after transfection and the performance of each shRNA compared with the performance of the corresponding siRNA. Studies assessing siRNAs and 19mer shRNAs revealed that there was little difference in silencing at 24 hours with these species (Fig. 2.3a). A comparison of siRNAs with shRNAs having 29 bp. stems gave a different result. Clustering of the data points above the diagonal indicated consistently better inhibition with the 29mer shRNAs (Fig. 2.3b). As predicted by the aforementioned results, direct comparisons of shRNAs with 19 and 29 basepair stems indicated a greater overall effect with latter structure (Fig. 2.3c).

The generally better inhibition with 29mer shRNAs at the high dose used for these studies led us to investigate the potency of these silencing triggers as compared with siRNAs. Seventeen complete sets comprising an siRNA, a 19mer shRNA and a 29mer shRNA were examined for suppression in titration experiments. In all cases, the 19mer shRNAs performed as well as or worse than the corresponding siRNAs. In contrast, 29mer shRNAs exceeded the performance of siRNAs in the majority of cases. In most cases, the 29mer shRNAs showed greater inhibition at the maximal dose; however even when this inhibition at the maximal dose did not differ much from the siRNA or 19mer shRNA the efficacy of the 29mer at lower concentrations was

significantly better. The dose response curves for four representative sets of RNAs are shown in Figure 2.3d.

Consistent with our results for most of the RNA sets tested, in the case of MAPK14, KIF14 and KIF11, the maximal level of suppression for the 29mer shRNA was approximately two-fold greater than the maximal level of suppression for the corresponding siRNA (Fig. 2.3e,g,h). More importantly, in some cases, the amount of RNA required to achieve maximal inhibition was up to 20-fold lower with 29mer shRNA than with a similar 21mer siRNA. The increase in potency for 29mer shRNA versus the other two RNA species may reflect the entry of these RNAs into the RNAi pathway as natural intermediates and may explain their greater efficacy when delivered from vectors

19

2.2.3 siRNAs and shRNAs give similar profiles of off-target effects at saturation

Microarray analysis has revealed down-regulation of many non-targeted transcripts following transfection of siRNAs into HeLa cells⁴⁰. Notably, these gene expression signatures differed between different siRNAs targeting the same gene. Many of the “off target” transcripts contained sites of partial identity to the individual siRNA, possibly explaining the source of the effects. To examine potential off-target effects of synthetic shRNAs, we compared shRNA signatures with those of siRNAs derived from the same target sequence. Using microarray gene expression profiling, we obtained a genome-wide view of transcript suppression. Figure 2.4 shows a two-dimensional clustering analysis of the signatures produced in HeLa cells 24 hours after transfection of 19mer and 29mer shRNAs compared with those generated by corresponding siRNAs. As

indicated by the dendrogram on the vertical axis, each set of three RNAs derived from the same core sequence was accurately clustered. Furthermore, in all but 2 of 7 cases, although the 19mer shRNAs produced signatures similar to those of the corresponding siRNAs, the signatures of the 29mer shRNAs and the siRNAs were more closely related. Note, that in one of the 2 cases in which the 19mer shRNA and the siRNA clustered more closely (MAPK14-1), these two RNA species did not significantly silence the target gene, whereas the 29mer shRNA did. The agreement between the signatures of 29mer shRNAs and siRNAs is consistent with precise processing of the shRNA to generate a single siRNA rather than a random sampling of the hairpin stem by Dicer. The overall smaller signature sizes of the 19mer shRNA and the basis of their divergence from the signature of the corresponding siRNA is presently unclear. However, our goal was not to extensively analyze off target effects potentially associated with these shRNAs.

2.3 Discussion

Considered together, our results suggest that chemically synthesized, 29mer shRNAs are often substantially more effective triggers of RNAi than are siRNAs. A mechanistic explanation for this finding may lie in the fact that 29mer shRNAs are substrates for Dicer processing both *in vitro* and *in vivo*. We originally suggested that siRNAs might be passed from Dicer to RISC in a solid-state reaction on the basis of an interaction between Dicer and Argonaute2 in *Drosophila* S2 cell extracts¹⁵. More recently, results from several laboratories have strongly suggested a model for assembly of the RNAi effector complex in which a multi-protein assembly containing Dicer and

accessory proteins interacts with an Argonaute protein and actively loads one strand of the siRNA or miRNA into RISC³⁰⁻³². Such a model implies that Dicer substrates, derived from nuclear processing of pri-miRNAs or cytoplasmic delivery of pre-miRNA mimetics, might be loaded into RISC more effectively than siRNAs. Our data support such a prediction, since it is not the hairpin structure of the synthetic RNA that determines its increased efficacy but the fact that the shRNA is a Dicer substrate that correlates with enhanced potency (see also accompanying paper by Rossi and colleagues). In *Drosophila*, Dicer is also required for siRNAs to enter RISC, and similar data has been obtained in mammalian cells^{30, 42}. Thus, it is possible that even siRNAs enter RISC via a Dicer-mediated assembly pathway and that our data simply reflect an increased affinity of Dicer for longer duplex substrates. Alternatively, hairpin RNAs, such as miRNA precursors, might interact with specific cellular proteins that facilitate delivery of these substrates to Dicer, whereas siRNAs might not benefit from such chaperones. Overall, our results suggest an improved method for triggering RNAi in mammalian cells that uses higher potency RNAi triggers. This remains a critical issue for both cell culture studies and for potential therapeutic use *in vivo*. Mapping the predominant 22 nt. sequence that appears in RISC from each of these shRNAs now permits the combination of this more effective triggering method with rules for effective siRNA design.

2.4 Methods

RNA sequence design

Each set of RNAs began with the choice of a single 19mer sequence. These 19mers were used directly to create siRNAs. To create shRNAs with 19mer stems, we appended a 4-base loop (either CCAA or UUGG) to the end of the 19mer sense strand target sequence followed by the 19mer complementary sequence and a UU overhang. We tested a variety of loop sequences and have noted no significant influence of the sequences examined on the performance of triggers. To create 29mer stems, we increased the length of the 19mer target sequence by adding 1 base upstream and 9 bases downstream from the target region and used the same loop sequence and UU overhang. All synthetic RNA molecules used in this study were purchased from Dharmacon.

Dicer processing

RNA hairpins corresponding to luciferase were end-labeled with [α -³²P] ATP and T4 Polynucleotide kinase. 0.1 pmoles of RNA were then processed with 2 units of Dicer (Stratagene) at 37 °C for 2 hours. Reaction products were trizol extracted, isopropanol precipitated, run on an 18% polyacrylamide, 8M urea denaturing gel. For RNaseIII digestion, 0.1 pmoles were digested with 1 unit of E. coli RNase III (NEB) for 30 minutes at 37 °C and analyzed as described above. Uniformly labeled hairpins were produced using a T7 Megashortscript kit (Ambion) with [γ -³²P] UTP and then incubated with dicer as indicated above.

For primer extension analysis, hairpins were processed with Dicer at 37 °C for 2 hours, followed by heat inactivation of the enzyme. The sequence of the primer is the first 16 nucleotide sequence from the 5' end of the hairpin:

5' AGTTGCGCCCGCGAAC 3'. DNA primers were 5' labeled with PNK and annealed to 0.05 pmole of RNA as follows : 95 °C for one minute, 10 minutes at 50 °C and then 1 min on ice. Extensions were carried out at 42 °C for 1 hour using MoMLV reverse transcriptase. Products were analyzed by electrophoresis on a 8M Urea/20% polyacrylamide gel. For analysis of *in vivo* processing, LinxA cells were transfected in 10 cm plates using Mirus TKO (10 ug hairpin RNA) or Mirus LT4 reagent for DNA transfection (12 ug of tagged Ago1/Ago 2 DNA; J. Liu, unpublished). 293 cells constitutively expressing Ago1 were utilized for Northern blot experiments. Cells were lysed and immunoprecipitated after 48 hours using with myc Antibody (9E14) Antibody. IPs were washed 3x in lysis buffer and treated with DNase for 15 minutes. Immunoprecipitates were then primer extended as described above.

siRNA and shRNA Transfections and mRNA Quantitation

HeLa cells were transfected in 96-well plates by use of Oligofectamine (Invitrogen) with the final nanomolar concentrations of each synthetic RNA indicated in the graphs. RNA quantitation was performed by Real-time PCR, using appropriate Applied Biosystems TaqMan™ primer probe sets 24 hours after RNA transfection, and the percent mRNA remaining was compared with cells treated with transfection reagent alone.

Microarray Gene Expression Profiling

HeLa cells were transfected in 6-well plates with 100 nM final concentration of the appropriate RNA by use of Oligofectamine (according to the manufacturer's instructions). RNA from transfected cells was hybridized competitively with RNA from mock-transfected cells (treated with transfection reagent in the absence of synthetic

RNA). Total RNA was purified by the Qiagen RNeasy kit, and processed as described previously⁴³ for hybridization to microarrays containing oligonucleotides corresponding to approximately 21,000 human genes. Ratio hybridizations were performed with fluorescent label reversal to eliminate dye bias. Microarrays were purchased from Agilent Technologies. Error models have been described previously⁴³. Data were analyzed using Rosetta Resolver™ software.

2.5 Figures

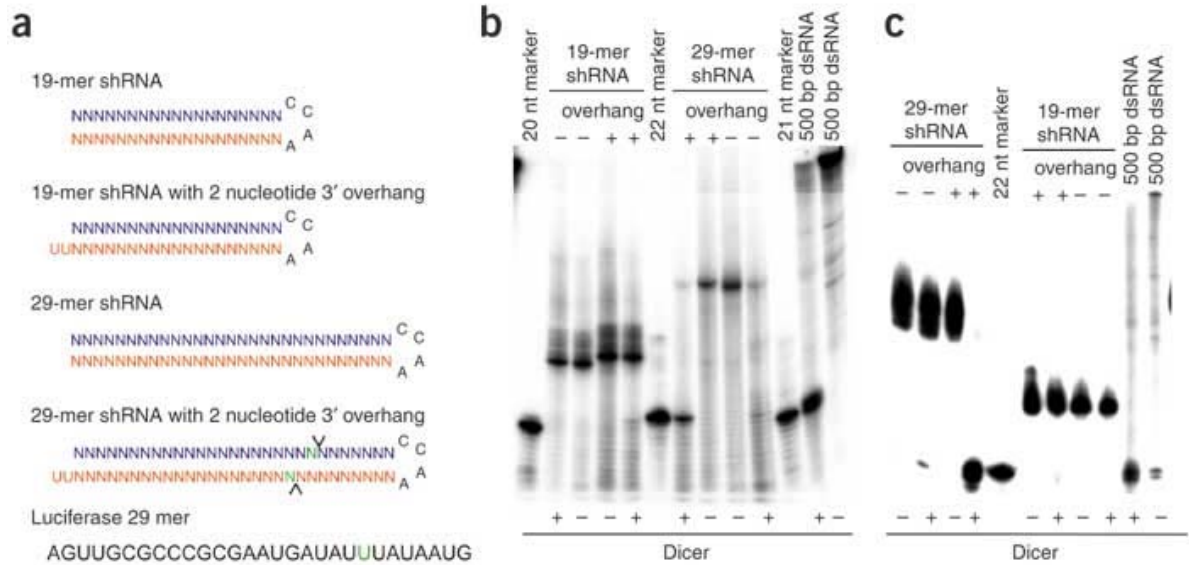


Figure 2.1. *In vitro* processing of 29 bp. shRNAs by Dicer generates a predominant siRNA from the end of each short hairpin. A. The set of shRNAs containing 19 or 29 bp. stems and either bearing or lacking a 2 nt. 3' overhang is depicted schematically. For reference the 29 nt sequence from luciferase (top, blue) strand is given. The presumed cleavage sites (as predicted by analysis of Dicer processing products) are indicated in green and by the arrows. **B.** *In vitro* Dicer processing of shRNAs. 5' end-labeled substrates as depicted in **A** were incubated either in the presence or absence of recombinant human Dicer (as indicated). Processing of a 500 bp. blunt-ended dsRNA is shown for comparison. Markers are end-labeled, single-stranded, synthetic RNA oligonucleotides. **C.** Uniformly labeled shRNAs with structures as indicated in Fig. 1A were processed by Dicer to produce a small RNA product (as indicated). A long 500 bp. blunt-ended dsRNA is processed and shown for comparison.

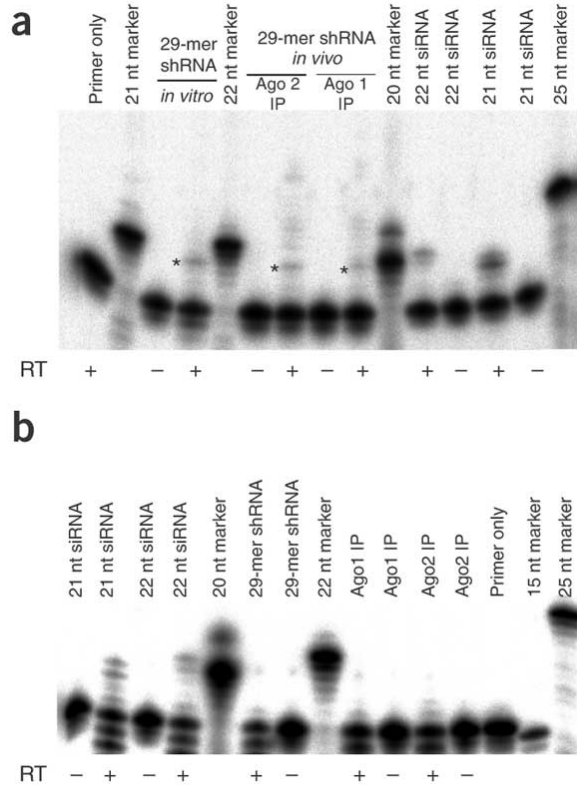


Figure 2.2 Primer extension analysis reveals similar small RNAs are generated by Dicer processing *in vitro* or *in vivo*. **A.** Primer extension was used to analyze products from processing of overhung 29 mer shRNAs *in vivo*. Total RNAs were extended with a specific primer that yields a 20 base product if cleavage occurs 22 bases from the 3' end of the overhang-containing RNA (see Fig 1A). For comparison, extensions of *in vitro* processed material are also shown. Lanes labeled siRNA are extensions of synthetic RNAs corresponding to predicted siRNAs that would be released by cleavage 21 or 22 nt. from the 3' end of the overhung precursor. Observation of extension products depends entirely on the inclusion of RT (indicated). The * indicates the specific extension product. Markers are phosphorylated, synthetic DNA oligonucleotides. **B.** Total RNA from control transfections, as indicated, did not show a primer extension product. The same primer was used for all extensions and is compatible with all RNAs. Controls with each RNA, as indicated, lacked a co-expressed tagged Ago protein, making it impossible to recover small RNAs in the

immunoprecipitates. Controls labeled Ago1 or Ago2 lacked co-transfected target RNAs.

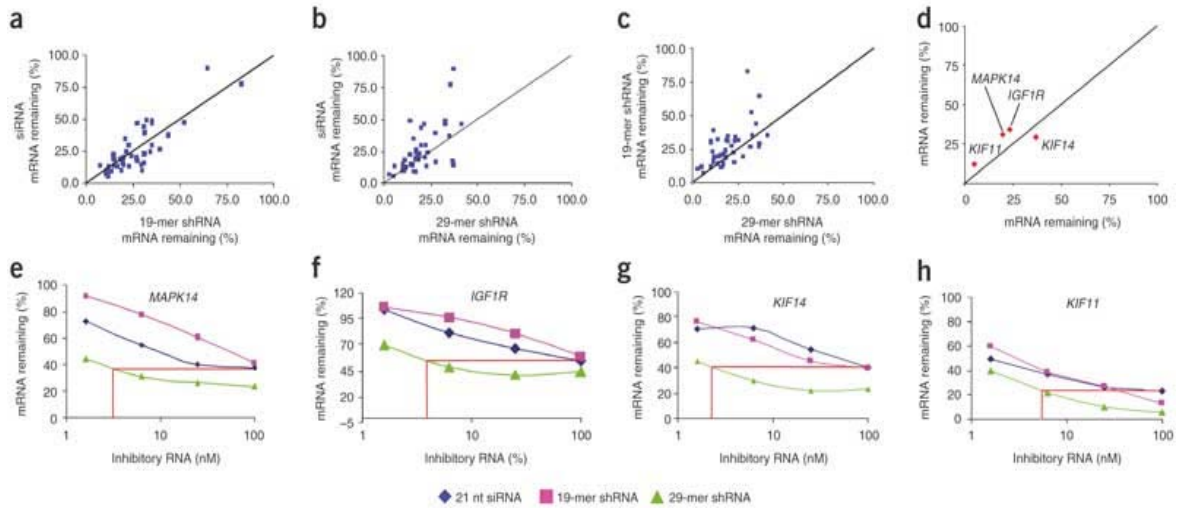


Figure 2.3. Gene suppression by shRNAs is comparable to or more effective than that achieved by siRNAs targeting the same sequences. **A,B, C.** mRNA suppression levels achieved by 43 siRNAs targeting 6 different genes were compared with levels achieved by 19mer or 29mer shRNAs derived from the same target sequences. 19mer and 29mer shRNAs were also directly compared. All RNAs were transfected at a final concentration of 100 nM. Values indicated on the X and Y axes reflect the percentage of mRNA remaining in HeLa cells 24 hours after RNA transfection compared with cells treated with transfection reagent alone. **D-H.** Four representative sets of siRNA and 19mer and 29mer shRNAs were used in dose-response analysis to compare the potency of representative RNAi triggers targeting 4 genes. Comparisons of relative suppression at the maximal dose are shown for reference in panel **D**. Titration curves were also performed reporting the percent of target mRNA remaining (Y axis) graphed from data derived from transfections at 1.56, 6.25, 25, and 100 nM final concentrations of RNA (X axis). Percent remaining RNA was determined by semi-quantitative RT-PCR. Gene targets were MAPK14, KIF 11, IGF1R and KIF14. (Sequences used were MAPK14-4, KIF11-6, IGF1R-1, KIF14-1 as in Suppl. Table 1.) (Blue diamonds: 21mer siRNAs; pink squares: 19mer shRNAs; green triangles: 29mer shRNAs). Red lines indicate the concentration of 29mer shRNA that gives the level of inhibition achieved by 100nM siRNA.

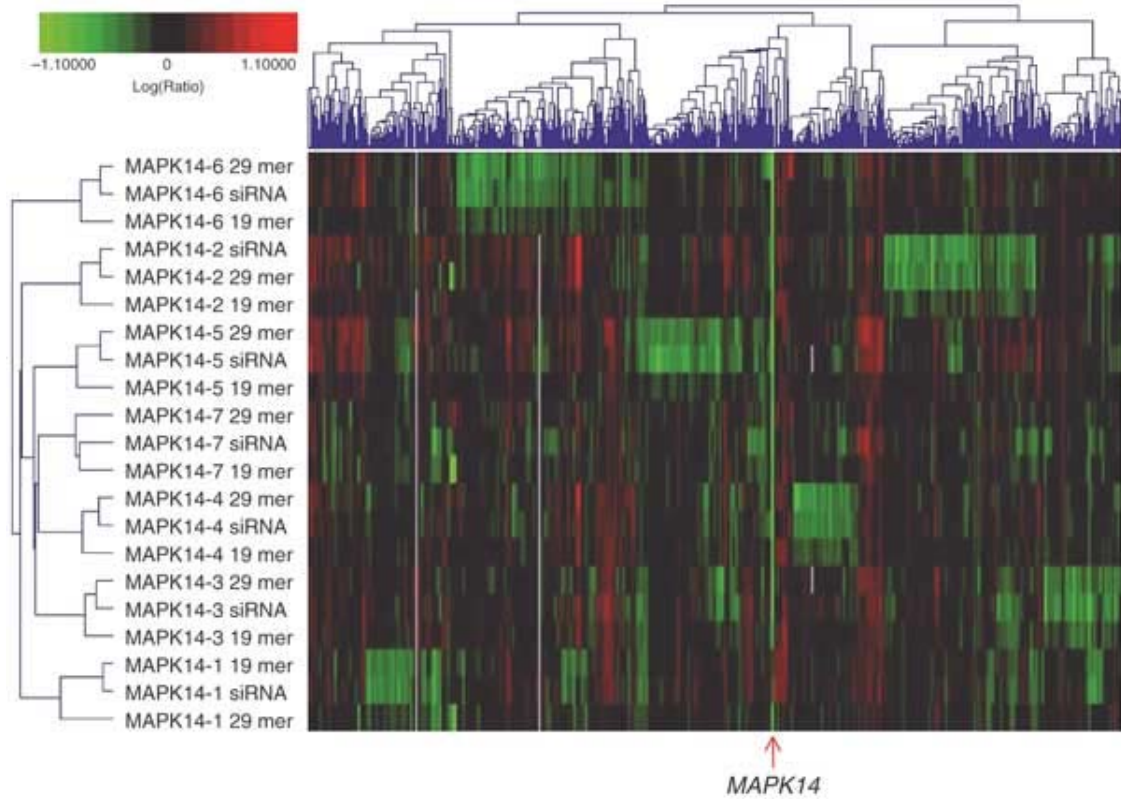
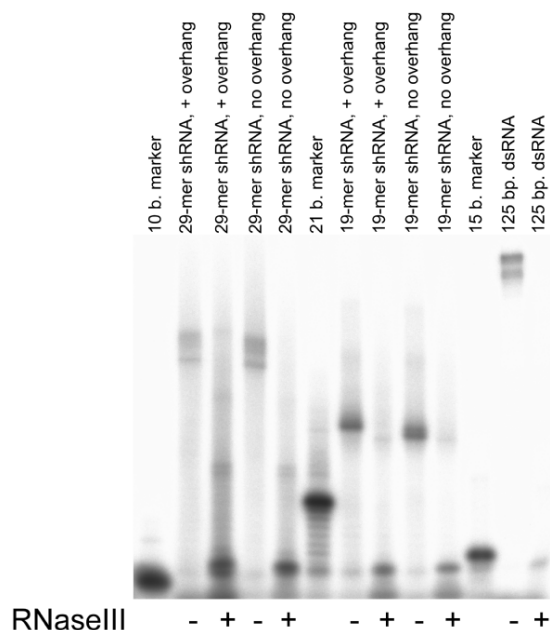
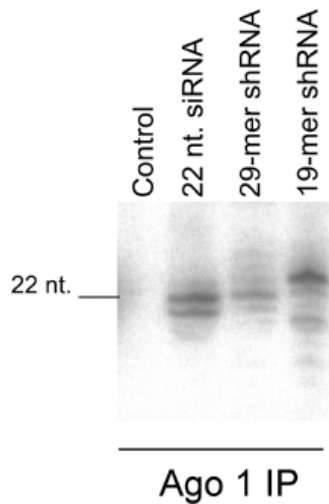


Figure 2.4. Microarray profiling reveals gene expression profiles are more similar between 29mer shRNAs and corresponding siRNAs than between siRNAs and 19mer shRNAs. 19mer and 29mer shRNAs and siRNAs designed for seven different target sequences within the coding region of MAPK14 were tested for gene silencing 24 hours after transfection into HeLa cells. Each row of the heat map reports the gene expression signature resulting from transfection of an individual RNA. Two-dimensional clustering of the data groups RNAs (vertical axis dendrogram) and regulated genes (horizontal axis dendrogram) according to signature similarities. Data shown represent genes that display at least a 2-fold change in expression level (p value < 0.01 and \log_{10} intensity > 1) relative to mock-transfected cells. Green indicates decreased expression relative to mock transfection whereas red indicates elevated expression. Black indicates no change and gray indicates data with a p value > 0.01 . The red arrow indicates MAPK14.



Green indicates decreased expression relative to mock transfection whereas red indicates elevated expression. Black indicates no change and gray indicates data with a p value > 0.01 . The red arrow indicates MAPK14.

Supplementary Figure 2.1. The set of shRNAs containing 19 or 29 bp. stems coding a luciferase sequence and either bearing or lacking a 2 nt. 3' overhang were incubated with bacterial RNase III to verify their double-stranded nature.



Supplementary Figure 2.2 Northern blotting indicates that siRNAs and 19mer and 29mer shRNAs all give rise to RISC. Each RNA species was transfected into HeLa cells constitutively expressing tagged Ago1 protein. Small RNAs in RISC were detected by Northern blotting of Ago immunoprecipitates.

siRNA



Synthetic 19-mer shRNA



Synthetic 29-mer shRNA



Supplementary Figure 2.3. Structures of synthetic RNAs used for comparing siRNA and shRNA. A total of 43 sequences were used with a matching set containing a 21mer siRNA (19 bp stem) and shRNAs with 19 or 29 bp. stems that were predicted to give Dicer products that were either identical to their corresponding siRNAs or that differed by the addition of one 3' nucleotide homologous to the target. Each was predicted to give precisely the same 5' end following cleavage of a 22mer RNA from the shRNA.

Supplementary Table 2.1. Sequences of the siRNAs used in this study

Gene	Accession number	Target sequence ID	Target sequence
IGF1R	NM_000875	IGF1R-1	GGAUGCACCAUCUUAAGG
IGF1R	NM_000875	IGF1R-2	GACAAAUCCCCAUCAGGA
IGF1R	NM_000875	IGF1R-3	ACCGCAAAGUCUUUGAGAA
IGF1R	NM_000875	IGF1R-4	GUCCUGACAUGCUGUUUGA
IGF1R	NM_000875	IGF1R-5	GACCACCAUCAACAAUGAG
IGF1R	NM_000875	IGF1R-6	CAAAUUAUGUGUUUCCGAA
IGF1R	NM_000875	IGF1R-7	CGCAUGUGCUGGCAGUAUA
IGF1R	NM_000875	IGF1R-8	CCGAAGAUUUCACAGUCA
IGF1R	NM_000875	IGF1R-9	ACCAUUGAUUCUGUUACUU
KIF11	NM_004523	KIF11-1	CUGACAAGAGCUCUAGGAA
KIF11	NM_004523	KIF11-2	CGUUCUGGAGCUGUUGAUA
KIF11	NM_004523	KIF11-3	GAGCCCAGAUCAACCUUUA
KIF11	NM_004523	KIF11-4	GGCAUUAACACACUGGAGA
KIF11	NM_004523	KIF11-5	GAUGGCAGCUCUAAAGCAA
KIF11	NM_004523	KIF11-6	CAGCAGAAAUCUAAAGGAUA
KIF11	NM_004523	KIF11-7	GACCUGUGCCUUUAGAGA
KIF11	NM_004523	KIF11-8	AAAGGACAACUGCAGCUAC
KIF11	NM_004523	KIF11-9	GACUUCAUUGACAGUGGCC
KIF14	NM_014875	KIF14-1	CAGGGAUGCUGUUUGGAUA
KIF14	NM_014875	KIF14-2	ACUGACAACAAAGUGCAGC
KIF14	NM_014875	KIF14-3	AAACUGGGAGGCUACUUAC
KIF14	NM_014875	KIF14-4	CACUGAAUGUGGGAGGUGA
KIF14	NM_014875	KIF14-5	GUCUGGGUGGAAAUUCAAA
KIF14	NM_014875	KIF14-6	CAUCUUUGCUGAAUCGAAA
KIF14	NM_014875	KIF14-7	GGGAUUGACGGCAGUAAGA
KIF14	NM_014875	KIF14-8	CAGGUAAGUCAGAGACAU
KIF14	NM_014875	KIF14-9	CUCACAUUGUCCACCAGGA
MAPK14	NM_139012	MAPK14-1	AAUAUCCUCAGGGGUGGAG
MAPK14	NM_139012	MAPK14-2	GUGCCUCUUGUUGCAGAGA
MAPK14	NM_139012	MAPK14-3	GAAGCUCUCCAGACCAUUU
MAPK14	NM_001315	MAPK14-4	CUCCUGAGAUCAUGCUGAA
MAPK14	NM_001315	MAPK14-5	GCUGUUGACUGGAAGAACA
MAPK14	NM_001315	MAPK14-6	GGAAUUCAAUGAUGUGUAU
MAPK14	NM_001315	MAPK14-7	CCAUUUCAGUCCAUCAUUC
PLK	NM_005030	PLK-1	CCUGUGUGGGACUCCUAA
PLK	NM_005030	PLK-2	CCGAGUUAUUCUUCGAGAC
PLK	NM_005030	PLK-3	GUUCUUUACUUCUGGCUAU
PLK	NM_005030	PLK-4	CGCCUCAUCCUCUACAUG
PLK	NM_005030	PLK-5	AAGAGACCUACCUCGGAU
PLK	NM_005030	PLK-6	GGUGUUCGCGGGCAAGAUU
PLK	NM_005030	PLK-7	CUCCUUAAAUAUUUCCGCA
PLK	NM_005030	PLK-8	AAGAAGAACCAGUGGUUCG
PLK	NM_005030	PLK-9	CUGAGCCUGAGGCCCGAUA

Literature Cited

1. Hamilton, A.J. & Baulcombe, D.C. A species of small antisense RNA in posttranscriptional gene silencing in plants [see comments]. *Science* **286**, 950-952 (1999).
2. Zamore, P.D., Tuschl, T., Sharp, P.A. & Bartel, D.P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25-33 (2000).
3. Hammond, S.M., Bernstein, E., Beach, D. & Hannon, G.J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293-296 (2000).
4. Bernstein, E., Caudy, A.A., Hammond, S.M. & Hannon, G.J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363-366. (2001).
5. Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R. & Hannon, G.J. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**, 1146-1150. (2001).
6. Tuschl, T., Zamore, P.D., Lehmann, R., Bartel, D.P. & Sharp, P.A. Targeted mRNA degradation by double-stranded RNA *in vitro*. *Genes Dev.* **13**, 3191-3197 (1999).
7. Song, J.J., Smith, S.K., Hannon, G.J. & Joshua-Tor, L. Crystal Structure of Argonaute and its implications for RISC slicer activity. *Science* **Published Online 7/29/04** (2004).
8. Liu, J. *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **Published online 7/29/04** (2004).
9. Caplen, N.J., Parrish, S., Imani, F., Fire, A. & Morgan, R.A. Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc. Natl. Acad. Sci. U S A* **98**, 9742-9747. (2001).
10. Elbashir, S.M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-498. (2001).
11. Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W. & Tuschl, T. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *Embo. J.* **20**, 6877-6888. (2001).

12. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297 (2004).
13. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419 (2003).
14. Hutvagner, G. et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834-838. (2001).
15. Ketting, R.F. et al. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes. Dev.* **15**, 2654-2659. (2001).
16. Grishok, A. et al. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**, 23-34. (2001).
17. Knight, S.W. & Bass, B.L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**, 2269-2271. (2001).
18. Brummelkamp, T.R., Bernards, R. & Agami, R. A System for Stable Expression of Short Interfering RNAs in Mammalian Cells. *Science* **21**, 21 (2002).
19. Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J. & Conklin, D.S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes. Dev.* **16**, 948-958. (2002).
20. Zeng, Y., Wagner, E.J. & Cullen, B.R. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol. Cell* **9**, 1327-1333. (2002).
21. Sui, G. et al. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5515-5520. (2002).
22. Lee, N.S. et al. Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat. Biotechnol.* **20**, 500-505. (2002).
23. Paul, C.P., Good, P.D., Winer, I. & Engelke, D.R. Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.* **20**, 505-508. (2002).
24. Lee, R.C. & Ambros, V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**, 862-864. (2001).
25. Lau, N.C., Lim, L.P., Weinstein, E.G. & Bartel, D.P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858-862. (2001).

26. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853-858. (2001).
27. Schwarz, D.S. et al. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199-208 (2003).
28. Silva, J.M., Sachidanandam, R. & Hannon, G.J. Free energy lights the path toward more effective RNAi. *Nat. Genet.* **35**, 303-305 (2003).
29. Khvorovova, A., Reynolds, A. & Jayasena, S.D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209-216 (2003).
30. Lee, Y.S. et al. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**, 69-81 (2004).
31. Pham, J.W., Pellino, J.L., Lee, Y.S., Carthew, R.W. & Sontheimer, E.J. A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. *Cell* **117**, 83-94 (2004).
32. Tomari, Y. et al. RISC assembly defects in the Drosophila RNAi mutant armitage. *Cell* **116**, 831-841 (2004).
33. Zhang, H., Kolb, F.A., Brondani, V., Billy, E. & Filipowicz, W. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *Embo. J.* **21**, 5875-5885. (2002).
34. Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E. & Kutay, U. Nuclear export of microRNA precursors. *Science* **303**, 95-98 (2004).
35. Ma, J.B., Ye, K. & Patel, D.J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**, 318-322 (2004).
36. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat. Struct. Mol. Biol.* **11**, 576-577 (2004).
37. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* **426**, 465-469 (2003).
38. Song, J.J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat. Struct. Biol.* **10**, 1026-1032 (2003).
39. Yan, K.S. et al. Structure and conserved RNA binding of the PAZ domain. *Nature* **426**, 468-474 (2003).

40. Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E. & Filipowicz, W. Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**, 57-68 (2004).
41. Jackson, A.L. *et al.* Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol.* **21**, 635-637 (2003).
42. Doi, N. *et al.* Short-Interfering-RNA-Mediated Gene Silencing in Mammalian Cells Requires Dicer and eIF2C Translation Initiation Factors. *Curr. Biol.* **13**, 41-46. (2003).
43. Hughes, T.R. *et al.* Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.* **19**, 342-347 (2001).

Chapter 3

Second-generation shRNA libraries covering the mouse and human genomes

Jose M. Silva ^{1#}, Mamie Z. Li ^{2#}, Ken Chang ^{1#}, Wei Ge ³, Michael C. Golding ¹, Richard J. Rickles², Despina Siolas ¹, Guang Hu², Patrick J. Paddison ¹, Michael R. Schlabach², Nihar Sheth ¹, Jeff Bradshaw³, Julia Burchard³, Amit Kulkarni³, Guy Cavet³, Ravi Sachidanandam¹, W. Richard McCombie ¹, Michele A. Cleary ³, Stephen J. Elledge ^{2*}, and Gregory J. Hannon^{1*}

The authors from Cold Spring Harbor Laboratory conducted library construction and validation experiments. Despina Siolas and Patrick Paddison performed the northern blot in figure 3B. The authors from Harvard Medical School developed the bacterial mating procedures. The authors from Rosetta Inpharmatics developed the bioinformatics algorithms applied to the library design.

¹ Cold Spring Harbor Laboratory
Watson School of Biological Sciences
1 Bungtown Road
Cold Spring Harbor, NY 11724

² Department of Genetics
Center for Genetics and Genomics
Harvard Medical School
Room 158D, NRB
77 Avenue Louis Pasteur
Boston, MA 02115

³ Rosetta Inpharmatics LLC,
A Wholly Owned Subsidiary of Merck & Co., Inc.
401 Terry Ave North
Seattle, Washington 98109, USA
#These authors contributed equally

Loss-of-function phenotypes often hold the key to understanding the connectivity and biological functions of biochemical pathways. We and others have previously constructed libraries of short hairpin RNAs (shRNAs) that allow systematic analysis of RNAi-induced phenotypes in mammalian cells^{57,58}. Here we report the construction and validation of second-generation shRNA (shRNA^{mir}) expression libraries that have been designed based on an increased knowledge of RNAi biochemistry. In these constructs, silencing triggers have been designed to mimic a natural microRNA primary transcript, and each target sequence has been selected based on thermodynamic criteria for optimal small RNA performance. Biochemical and phenotypic assays have indicated that the new libraries are substantially improved compared to first-generation reagents. We have generated large-scale arrayed, sequence-verified libraries comprising more than 140,000 shRNA^{mir} expression plasmids, covering a substantial fraction of all predicted genes in the human and mouse genomes. These libraries are presently available to the scientific community.

3. 1 Introduction

RNA interference (RNAi) has been exploited in organisms ranging from plants to fungi to animals for deciphering gene function through suppression of gene expression. Particularly in systems where targeted genetic manipulation is difficult or time consuming, RNAi has transformed the way in which gene function can be approached on a single gene or genome-wide level ^{49, 76-79}.

The RNAi machinery can be programmed by exogenous or endogenous sources of double-stranded RNA. The most well characterized source of endogenous triggers are the microRNA genes ^{25, 80}. It was initially assumed that microRNAs were transcribed from the genome as short, hairpin RNAs ⁸¹ that were directly processed by Dicer to yield the mature small RNAs that enter RISC ⁸²⁻⁸⁵. Over the past year, however, a different picture has emerged. Numerous studies have demonstrated that, in animals, miRNAs are transcribed by RNA polymerase II to generate long primary polyadenylated RNAs (pri-miRNAs) ^{20, 69}. Through mechanisms not yet fully understood, the pri-microRNA is recognized and cleaved at a specific site by the nuclear Microprocessor complex ^{22, 23, 86-88}. This contains an RNase III family enzyme, Drosha, that cleaves the hairpin to produce a ~70-90 nucleotide microRNA precursor (pre-miRNA) with a 2 nucleotide 3' overhang ²³. This distinctive structure signals transport of the pre-miRNA to the cytoplasm by a mechanism mediated by Exportin-5 ^{23, 75}. Only then is the pre-miRNA recognized by Dicer and cleaved to produce a mature microRNA. This probably involves recognition the 2 nucleotide 3' overhang created by Drosha to focus Dicer cleavage at a single site ~22 nucleotides from the end of the hairpin ^{24, 89}.

Mature miRNAs are superficially symmetrical, with 2 nucleotide 3' overhangs at each end having been generated by Drosha and Dicer, respectively. However, the individual strands of the mature miRNA enter RISC in an unequal manner. As with siRNAs, the thermodynamic asymmetry of the Dicer product is sensed such that the strand with the less stable 5' end has a greater propensity to enter RISC and guide substrate selection^{73,90}. This observation of thermodynamic asymmetry within small RNAs led to the development of rules for predicting effective siRNA sequences that have greatly improved the efficiency of those RNAs as genetic tools.

Previously, several groups, including our own, described the design and construction of arrayed short hairpin RNA (shRNA) libraries that covered a fraction (~1/3) of human genes^{57,58}. At the time when these tools were developed, our knowledge of microRNA maturation was relatively incomplete. This led most groups to the notion of expressing a simple hairpin RNA that mimicked the pre-miRNA.

Here we report the construction of a new generation of shRNA libraries (shRNA^{mir}) that takes into consideration our advancing understanding of microRNA biogenesis. As has previously proven successful in plants and in animals, second-generation shRNA^{mir}s are modeled after endogenous microRNAs, specifically being harbored within the backbone of the primary mir-30 microRNA⁹¹. This natural configuration proved to be up to 12 times more efficient in the production of the mature synthetic miRNAs than simple hairpin designs. Additionally, we have biochemically characterized processing of these synthetic microRNAs, allowing us to predict the mature small RNA product(s) that will be generated from each vector. This has allowed selection of target sequences that maximize efficiency by directing preferential

incorporation of the correct strand into RISC. Using these criteria, we have produced and sequence-verified more than 140,000 shRNAs covering a substantial fraction of the predicted genes in the mouse and human genomes. We have assayed a selected subset of shRNAs from the library for their ability to knock-down the expression of targeted genes by quantitative RT-PCR. We have also tested this set in a phenotypic assay and compared the performance of the first- and second-generation library designs. Overall, the shRNA^{mir} libraries that we describe here provide a convenient, flexible and effective tool for studying gene function in human cells. Additionally, they, for the first time, extend the possibility of large-scale RNAi screens to mouse systems.

3.2 Results

3.2.1 Design and construction of second generation shRNA libraries

We have previously shown that expression of a simple, 29 bp hairpin from a U6 snRNA promoter can induce effective suppression of target genes when delivered either transiently or stably from integrated constructs^{54, 58, 92}. We also found that longer hairpin structures were more effective inhibitors of gene expression than were shorter structures with stems of 19-21 nucleotides. All of these constructs, however, were designed to express a pre-miRNA hairpin, an intermediate in microRNA biogenesis, rather than a transcript that closely resembles a primary microRNA. Cullen and colleagues had previously shown that effective suppression could be achieved by redesigning an endogenous microRNA, miR-30, such that its targeting sequence was directed against a reporter gene⁹¹. We sought to compare directly the abundance of small RNAs produced from vectors with simple hairpin structures to those that more closely resemble a natural

microRNA. Since it had been previously shown that the efficient ectopic expression of endogenous microRNAs requires substantial flanking sequence⁹³, we developed a vector in which sequences from a remodeled miR30 are flanked by ~125 bases of 5' and 3' sequence derived from the primary transcript. Incorporation of appropriate cloning sites into this vector required altering only 3 positions in the precursor. This cassette was inserted into a vector equivalent to that in which we constructed our first-generation shRNA library (pSM1), with the new shRNA vector being designated, pSM2. To distinguish the second-generation shRNAs from those in our first-generation library, we have dubbed these shRNA^{mir}.

In order to enable the use of small RNA design rules to potentially enhance the efficacy of our shRNAs, it was necessary to understand how the shRNA^{mir} was processed *in vivo*. To address this issue, we took advantage of existing studies of miR30 biogenesis that mapped its processing sites⁹⁴. Using this information as a guide, we designed a series of constructs predicted to generate small RNAs targeting mouse p53, human PTEN and luciferase. We transfected human 293 cells with pSM2 carrying each of these inserts and mapped the mature 3' ends of the guide and passenger strands of p53 and PTEN shRNAs and the guide strand of the luciferase shRNA by RACE-PCR (Fig. 3.1a; Supplementary Fig. 3.1). Since Northern Blotting indicated that maturation of shRNA^{mir}s produced 22 nucleotide species (Fig. 1b), we were able to infer the 5' end of each small RNA species. We consider the possibility of two processing sites at each end of the shRNA, since our analysis in the cases of p53 and PTEN shRNAs could not distinguish between processing at either of two terminal bases (Fig. 3.1a). However, in the case of luc1309, the answer was relatively clear that the guide strand was cleaved in most cases

(8/10) at the most 3' indicated site. Two out of ten sequences indicated cleavage 1 base 5' of that site, perhaps reflecting a genuine heterogeneity in Drosha cleavage (RACE1, RACE3; Supplementary Fig. 3.1). We therefore feel it most likely that cleavage is most prevalent at the sites indicated by the heavy red (drosha) and blue (dicer) lines, but our data is consistent with the possibility of some cleavage also occurring at the sites indicated by the lighter lines.

To test the performance of pSM2 in comparison to pSM1, we used both vectors to express a sequence targeting firefly luciferase. The sequence was inserted such that an identical mature small RNA would be generated from each construct following processing *in vivo* (Supplementary Fig. 3.2). Of primary concern was the overall amount of mature small RNA that would be generated from each construct. This was critical as dose-response experiments for shRNAs indicate that suppression correlates very well with the amount of RNA delivered⁸⁹, particularly at the relatively low doses that are expected to be achieved by expression from transfected or integrated constructs as compared to directly transfected synthetic RNAs. We transfected pSM1-luc and pSM2-luc into 293 cells, prepared RNA and assayed the processed small RNA by northern blotting. Cells transfected with pSM2-luc contained roughly 12-fold more of the small RNA than did cells transfected with pSM1-luc (Fig 3.1b).

As it is now clear that primary microRNAs are transcribed mainly by RNA polymerase II^{20, 69}, we wished to compare the performance of shRNA^{mir}s driven by a variety of different promoters. We therefore cloned two different shRNA^{mir} cassettes targeting firefly luciferase downstream of three different RNA polymerase III promoters (tRNA-val⁹⁵, U6⁹² and H1⁹⁶) and two different RNA polymerase II promoters (MSCV-

LTR and CMV⁹⁷). These constructs were each prepared in a plasmid backbone that carried no other mammalian promoter. Each was transfected in combination with a homologous target expression plasmid encoding firefly luciferase and with a non-targeted reporter plasmid, encoding *Renilla* luciferase, as a means of normalization. We compared the performance of these plasmids in a four different cell lines including two from human, one from mouse and one from dog. When the ability of these constructs to suppress the luciferase target was compared using a very efficient shRNA^{mir} (luc1309), we saw virtually no difference in the performance of the various promoters (Fig. 3.1c). However, when a less efficient shRNA^{mir} (luc311) was used, differences became apparent (Fig. 3.1c). In this, and numerous experiments with other shRNAs (not shown), the U6 snRNA and CMV promoters gave the best and most consistent repression. Based upon these studies, we chose to retain the U6 snRNA promoter in our base library vector, pSM2. It is important to note that all of our studies have been carried out in transient assays. In situations in which constructs are stably integrated into the genome at single copy, different configurations of promoters and flanking sequences perform more efficiently than U6^{47,98}. However, we can also suppress gene expression by stable integration of pSM2 directly (Supplementary Figure 3.3).

Based upon these tests we constructed our second-generation shRNA library vector, pSM2, as shown in figure 2a. The shRNA^{mir} expression cassette is carried within a self-inactivating murine stem cell virus. Expression of the small RNA is driven by the U6 snRNA promoter. As with the first-generation shRNAs, a U6 snRNA leader sequence lies between the promoter and the 5' end of the miR-30 flanking region. Synthetic oligonucleotides encoding shRNAs are inserted into XhoI and EcoRI sites that

lie within the miR-30 primary microRNA sequences. Immediately following the miR-30 cassette in each vector is a RNA polymerase III termination signal and a randomly generated 60 nucleotide barcode region to facilitate tracking of individual hairpin RNAs in complex populations. This feature is similar to that described for our first-generation RNAi library^{54, 58}. The pSM2 vector is also designed such that inserts can be moved by an *in vivo* recombination strategy (MAGIC)⁹⁹. One key difference between the first and second-generation shRNA libraries is that the shRNA^{mir} cassette can be transferred without the constitutive U6 snRNA promoter. This allows the construction of mating recipients that contain inducible or tissue specific promoters^{47, 98}.

Six different shRNA^{mir} sequences were designed for each of 34,711 different known and predicted human genes and 32,628 mouse genes. In each case, shRNAs were designed such that the mature small RNA generated from each construct followed thermodynamic asymmetry rules that have been successfully applied for the design of siRNAs. Based upon the approaches used to map the termini of the mature small RNAs generated from our vectors, we could not definitively distinguish between processing at two possible sites. Additionally, Dicer has been shown to generate some 3' end heterogeneity in processing its substrates. Therefore we chose sequences that gave similar thermodynamic profiles even if cleavage sites were shifted by a base at either end (the cleavage positions indicated in Fig. 3.1).

Construction of the library proceeded stochastically using a highly parallel *in situ* synthesis approach for oligonucleotide production (Fig. 3.2b). Groups of ~22,000 oligonucleotides, each containing a different shRNA^{mir} cassette were synthesized on glass-slide microarrays⁵⁸. Populations were eluted from the arrays and amplified by

PCR. In order to insure efficient cloning, the pSM2 backbone was inserted into a lambda phage backbone such that it was flanked by loxP sites. λ -pSM2 contains unique XhoI, EcoRI for subcloning amplified hairpins and unique FseI and AvrII sites for insertion of bar code 60mers. λ -pSM2 was first barcoded with a mixed library of random 60 nucleotide sequences amplified with a primer set which included one primer with an FseI site and one primer with a T7 promoter followed by the AvrII site. Amplified barcoded λ -pSM2 libraries were lysogenized into a strain we constructed for this purpose, DH10 β λ_{KP} , which has a wild-type *pir1* gene and the lambda repressor, *cI*, to allow λ -pSM2 to replicate as a 42 kb plasmid. Approximately 10^8 Cm^RKm^R lysogens were selected and served as a bar coded library pool. Bar coded λ -pSM2 was CsCl purified, then cleaved with EcoRI and XhoI before being ligated to gel purified EcoRI-XhoI cleaved pooled shRNA^{mir} inserts from an individual chip and packaged. Average library sizes were $\sim 5 \times 10^7$ recombinants per pool. To generate pSM2 library plasmid pools, the phage were used to infect an *E. coli* strain we constructed, BUN25, that expresses both Cre recombinase and the *pir1-116* gene, needed for high copy RK6 γ replication. Pooled plasmid libraries were then transformed into a mating competent host strain (BW F'DOT) and individual clones were sequenced at random. Clones with perfect inserts represented between 25 and 50% of the population, and these were selected and saved as an arrayed set. Accumulation of new clones from each pool was monitored dynamically and once a pool began to yield fewer unique clones per sequencing run, sequencing was halted and the pool was resynthesized without those sequences that had already been obtained. Approximately 70 chips were reiteratively synthesized to maximize unique sequencing. Also, once 3 or more verified shRNAs were obtained for any given gene, the remaining

shRNAs targeting that gene were also withdrawn from population selected for resynthesis.

To date, we have sequence verified 79,805 shRNAs targeting 30,728 human genes and 67,676 shRNAs targeting 28,801 mouse genes. A tabulation of coverage within selected functional groups can be found in Table 1 for the mouse and human libraries. The ultimate goal is to generate 3 shRNAs for each target locus. The full collection of mouse and human shRNAs, updated dynamically, can be accessed at <http://codex.cshl.edu>.

3.2.2 Validation of the second-generation shRNA libraries

To test the efficiency of the second-generation shRNA libraries, we took an approach that we had previously used to assess the performance of the first generation reagents⁵⁸. A green fluorescent protein (ZsGreen) reporter harboring the PEST domain of the mouse ornithine decarboxylase is normally degraded by the proteasome¹⁰⁰. Thus, cells harboring a destabilized ZsGreen expression plasmid show very low levels of fluorescence. Interference with proteasome function, for example using a synthetic proteasome inhibitor, causes accumulation of the protein and a corresponding increase in fluorescence. The protein can also be stabilized by suppression of any gene required for proteasome function. Thus, co-transfection of the reporters with an shRNA^{mir} expression plasmid can reveal whether a target protein is involved in the proteasome pathway (Fig. 3a). Using this assay as a primary test we compared a series of shRNAs targeting proteasomal subunits that were obtained from either the first- or second-generation libraries (Fig. 3b; Supplementary Table 1).

We chose a total of 53 shRNAs targeting 13 different genes that were known to be involved in proteasome function (Fig. 3b). 24 were from the first-generation library and 29 were from the second-generation library. These were co-transfected with the reporter in combination with a dsRED-encoding plasmid that allowed normalization of the transfections. It was immediately apparent that the second-generation shRNAs performed substantially better than the first-generation shRNAs. We noted that most of the plasmids derived from the second-generation library were as potent as the best shRNAs that had been selected from a screen of the first-generation library.

To gain a more detailed picture of the performance of the second-generation libraries, we compared results from the proteasome assay for 36 shRNA^{mir} expression plasmids to suppression of target RNAs as measured by semi-quantitative RT-PCR. Plasmids were transfected into HeLa cells with approximately 80% efficiency, as measured by reference to a co-transfected reporter plasmid. Despite this incomplete transfection, all but 6 of the shRNA^{mir}s reduced the levels of their target RNAs by ~60% or more with 13/36 of the shRNA^{mir}s suppressing their targets by the theoretical maximum of ~80% (Fig. 3.3c, upper panel; Supplementary Table 3.1). Similar results were seen with an additional 12 shRNAs that did not target proteasome subunits (not shown). These studies were also illuminating, as they revealed that the functional assay in some cases, e.g., pSMB3, did not show a large activation of the reporter despite substantial suppression of the targeted mRNA (Fig. 3.3c, lower panel). Thus, the functional assay underestimated slightly the efficacy of the library.

To test the performance of the library on a larger scale, we assayed a set of 515 kinase shRNAs that contained within it 47 hairpins directed to proteasome subunits using

the phenotypic assay for proteasome function via a high-throughput protocol in 96-well plates (Fig. 3.4). In this context, 34/47 shRNAs targeting the proteasome scored as positives (72%) as compared to 10 shRNAs that had not previously been linked to proteasome function (1.9%). A secondary screen of those 44 potential positives from the primary screen again revealed positive signals from all 34 proteasomal shRNAs. However, only 5/10 of the non-proteasomal RNAs continued to activate the reporter, and none of these scored with more than one shRNA in the library (Supplementary Table 3.2).

3.3 Discussion

Since the discovery that an RNAi pathway was conserved in mammals, the exploitation of this silencing response as a genetic tool has evolved in concert with our deeper understanding of its biochemical mechanism. The initial applications of siRNAs as triggers of the silencing response required comprehension of the way in which Dicer processes long dsRNA substrates in *Drosophila*¹⁶. Similarly, studies of dicer-mutant *C. elegans* demonstrated that endogenous loci could encode triggers of the RNAi machinery, and this led to the notion that such loci could be altered to target genes for experimental silencing⁸²⁻⁸⁵.

Many strategies have been developed for producing miRNA-like triggers of the RNAi pathway. As we have mapped the processing sites on precursor shRNA^{mir}s, we can predict what small RNA is generated from each shRNA^{mir} expression vector. This enables us to apply siRNA design rules to shRNA^{mir} expression cassettes. A combination

of increased small RNA production with better shRNA design yielded a pronounced increase in the performance of these silencing tools.

Guided by these design strategies, we have constructed large libraries of sequence-verified shRNAs targeting the majority of the known and predicted genes in the human and mouse genomes. On average, each locus is covered by 2 shRNA^{mir}s presently; however, the ultimate goal is to have 3 sequence-verified shRNA^{mir}s for each gene. The second generation libraries resemble those that we have previously reported in that they reside in flexible vectors that permit shuttling of shRNA^{mir} expression cassettes into virtually any desired expression vector using a bacterial mating strategy⁹⁹. A unique feature of the second generation library is that the expression cassette can be moved without the need to move also the constitutive U6 snRNA promoter. This permits large scale construction of secondary libraries under the control of tissue specific and inducible promoters. Indeed, regulated expression of our library cassettes from RNA polymerase II promoters has been shown to effectively suppress gene expression both in cultured cells and in animals^{47,98}. These recipient vectors can be directly used with any shRNA^{mir} encoded by the library described herein.

The use of large-scale resources for suppressing gene expression via RNAi promises to revolutionize genetic approaches to biological problems in numerous model systems. The libraries described here should prove useful for assessing the functions of individual genes and for taking genome-wide approaches. Strategies reported in accompanying papers^{47,98} will permit large-scale application of these tools for screens which require long-term suppression of gene expression using single-copy integrants or

inducible repression. Thus, we have produced coherent system of RNAi reagents with utility in both mouse and human experimental systems.

3.4 Methods

Construction of the lysogenic strain DH10 β λ_{KP} and excision strain BUN25

DH10 β λ_{KP} [*mcrA* Δ (*mrr-hsdRMS-mcrBC*) ϕ 80 *lacZ* Δ M15 Δ *lacX74* *deoR* *recA1* *endA1* *ara* Δ 139 Δ (*ara, leu*)7697 *galU* *galK* λ *rpsL* *nupG* *tonA* λ -*pir1-npt*] is a strain containing λ cI and the *pir1* gene that was constructed in order to lysogenize the λ SM2-barcode library prior to introduction of the hairpin fragments. To generate this strain, λ_{KP} containing the *pir1* and Km^R genes was constructed. To generate λ_{KP} , the *pir1* gene was amplified from BW23473 using primers MZL393 and MZL51, and cloned into the pCR2.1 TOPO TA cloning vector. The *pir1* gene was excised from the above clone on a BamHI fragment and ligated into BamHI cleaved pSE356, which contains an *npt* gene and a *Bam*HI restriction site flanked by two 1 kb λ DNA fragments¹⁰¹ to generate pSE356pirWT. The *pir1Km*^R fragment was recombined onto wild type λ by amplifying λ on LE392/pSE356pirWT and the resulting phage were collected and used to infect DH10 β . 100 μ l of DH10 β cells were infected with 10⁶ PFU at 30°C for 30 minutes in LB + 10 mM MgSO₄, diluted with 900 μ l of LB incubated at 30°C for 2 h with shaking, and plated on LB containing 50 μ g/ml kanamycin at 37°C overnight to select λ_{KP} lysogens. Lysogens were tested for the ability to lysogenize λ vectors containing R6K γ origins of replication as extrachromosomal elements. A strain capable of doing this was selected and named DH10 β λ_{KP} .

The BUN25 [F' *traD36 lacI^q Δ(lacZ)M15 proA⁺B⁺/e14(McrA⁻) Δ(lac-proAB) thi gyrA96 (Nal^r) endA1 hsdR17 (r_k⁻m_k⁺) relA1 glnV44 λ-cre-npt umuC::pir116-Frt sbcDC-Frt*] strain containing *pir1-116* and *cre* was constructed to allow the conversion of λSM2 shRNA libraries into pSM2 shRNA libraries. A PCR fragment containing *pir1-116* gene was generated using primers MZL393 and MZL51 (see Supplementary table 3), cleaved with BamHI and ligated into BamHI-cleaved pUC18 to generate pML284. A fragment containing *BstBI-Frt-cat-Frt-NdeI* (filled-in) was isolated from KD3¹⁰² and inserted into the *SmaI* site of pML284 to generate pML334. A *HpaI* fragment containing *UmuDC* was isolated from pSE117 and cloned into pBluescript *XhoI* (filled in)-*EcoRV* to generate pML236. We eliminated one of the *BamHI* site on pML236 by digesting it with *PstI-XbaI*, filling in with T4 DNA polymerase and ligating. The *Frt-cat-Frt-pir116* was isolated from pML334 as a *KpnI-SacI* (filled-in) fragment and ligated into *MluI/BamHI* (filled-in) cleaved pML236-Δ*BamHI* to generate pML346. The 3.8 kb *KpnI-SacI UmuDC-Frt-cat-Frt-pir116-UmuC* fragment from pML346 was integrated into BNN132/pML104 by homologous recombination using the λ recombinase expressed from pML104, and confirmed by colony PCR. The *cat* gene was removed by FLP-mediated excision *in vivo* using pCP20¹⁰³ which expresses the FLP recombinase to generate BUN24. A cassette that has *Frt-cat-Frt* flanked by 50 bp homology to *sbcD* and 50 bp homology to *sbcC* was amplified by primers MZL493/MZL494 and using KD3 as a template. This cassette was used to replace *sbcD* and part of *sbcC* on BNN132 by homologous recombination and the deletion were confirmed by colony PCR. The strain was named BNN132sbcDC-Frt-cat-Frt. We then used a pair of outside primers (MZL495/MZL496) that gave about 500 bp homology regions to the upstream of *sbcD*

and 500 bp homology regions to the *sbcC* to amplify a PCR product from BNN132sbcDC-Frt-cat-Frt to recombine onto the *sbcDC* region of BUN24. The resulting strain was named BUN25 and is used to stabilize inverted repeats in *E. coli*¹⁰⁴.

Library vector construction

A pair of loxP-*NotI*-loxP duplexed oligos (MZL524/MZL525) were inserted into the pSM2 *BstXI* site to generate pSM2c-loxP. A second pair of duplexed oligos (MZL541/MZL542), carrying the proper restriction sites for cloning barcodes into λ SM2, were inserted into the *BbsI*-*MluI* sites of pSM2c-loxP to create pML375. λ ACT2 was digested with *NotI*, and the λ arms were gel purified and ligated to *NotI* digested pML375 to generate λ SM2. The ligation mixture was packaged using MaxPlaxTM lambda packaging extracts from Epicentre. We selected a λ SM2 lysogen by infecting 200 μ l of BW23473 cell ($A_{600} = \sim 0.8$) with 100 μ l of λ SM2 packaging mix in the presence of 10 mM $MgSO_4$ and 0.2 % (w/v) maltose, incubated at 30°C for 30 minutes then added 900 μ l of LB and incubated at 30°C for 2 hrs with shaking to express the Cm^R marker, and plated on LB containing 17 μ g/ml of chloramphenicol (Cm) at 30°C overnight. The proper recombinants were confirmed by restriction analysis. See Supplementary table 3 for sequences of referenced oligonucleotides.

Barcode library construction

The 60 base pair barcode oligos, were amplified using barcode primer 1 and barcode primer 2 (Supplementary table 3). The PCR conditions were: 0.1 pmol of barcodes, 50 pmol of each primer, 25 nmol of each dNTP, and 2.5 U of Taq DNA polymerase; 94°C for 45 seconds, (94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 30 seconds) x 13, 72°C for 10 minutes, 4°C forever. Ten PCR reactions were pooled together, purified

using a QIAquick PCR purification kit, digested with *Bam*HI, *Eco*RI, *Xho*I and *Sal*I to remove these sites in the barcodes, and gel purified. The purified barcodes were digested with *Fse*I and *Avr*II and purified using the QIAquick gel extraction kit. Two micrograms of *Fse*I-*Avr*II digested λ SM2 ligated with 10 ng of *Fse*I-*Avr*II digested barcodes with 1 x ligation buffer and 0.5 μ l T4 DNA ligase in a 5 μ l final volume at 16°C overnight. The ligation mixture was packaged and amplified. The size of the λ SM2-barcode library was 4.2×10^7 . We used 20 ml of DH10 β λ_{KP} cells ($A_{600} = \sim 1$) to lysogenize 2×10^9 of λ SM2-barcode library as 42 kb plasmids. The cell and the phage were mixed in the presence of 10 mM MgSO₄ and 0.2 % (w/v) maltose and incubated at 30°C for 30 minutes, added 200 ml of LB to recover at 30°C for 1 h by shaking. The mixture were concentrated by centrifugation at 4000 rpm for 20 minutes, resuspended in 3 ml of LB, plated on 10 large LB/Cm 17 μ g/ml, and incubated at 30°C overnight. The cells were scraped from plates and grown in 3 L of TB containing 17 μ g/ml of Cm overnight. Supercoiled λ SM2-barcode library DNA was prepared by cesium chloride. The lysogenization efficiency was approximately 30%.

Oligonucleotide Cleavage and PCR Amplification

To harvest oligonucleotides, we treated microarrays for 2 h with 2-3 mL of 35% NH₄OH solution (Fisher Scientific) at room temperature. We transferred the solution to 1.5-mL microcentrifuge tubes and subjected it to speed vacuum drying at medium heat (~55 °C) overnight. We resuspended the dried material in 200 μ l of RNase/DNase free water and performed PCR amplification in 50 μ l reaction volumes using Invitrogen's Platinum® *Pfx* DNA Polymerase. To obtain a sufficient amount of PCR product, four 50 μ l reactions were required for each sample. Each reaction contained 2X *Pfx* PCR amplification buffer,

0.3 mM of each dNTP, 1 mM MgSO₄, 0.3 uM of each primer, 0.5X PCR enhancer solution, 0.5 units of Platinum® *Pfx* DNA Polymerase, and 10 µl of template DNA. The primers used for amplification were 5'-mir30-PCR-xhoI-F (5' CAGAGGCTCGAGAAGGTATATTGCTGTTGACAGTGAGCG 3') and 3'-mir30-PCR-*ecorI*-R (5' CGCGGCGAATTCCGAGGAGTAGGCA 3'). After an initial denaturation step of 94°C for 5 min, DNA amplification occurred through 25 cycles of denaturing at 94°C for 45 seconds and annealing and extension at 68° for 1 min and 15 sec followed by a final 7 minute extension at 68°. We then combined the four reactions into one tube, cleaned up the PCR product by use the QIAGEN® Minelute PCR Purification Kit and eluted in a total volume of 26 µl.

ShRNA library construction

The λSM2-barcode library and shRNA PCR products were digested with *EcoRI* and *XhoI* overnight and gel purified. Ligations with shDNA oligos were set up as following: 1.5 µg of *XhoI-EcoRI* cleaved vector, 8-10 ng of *XhoI-EcoRI* cleaved inserts generated from the PCR of shDNA oligos from the parallel microarray synthesis, 1 µl of 10 x ligation buffer, 0.5 µl of T4 DNA ligase, and water to 10 µl final. The ligation mixtures were incubated at 16°C for overnight and packaged. We typically observed 30- to 90-fold stimulation of plaque forming units (PFU) and 2 x 10⁷ to 8 x 10⁷ PFU total for each library pool. We typically amplify 2 x 10⁷ PFU for each pool to generate a stock. To verify the ligation efficiency, we excised 10 µl of package mix by infecting 100 µl of BUN25 (A = ~0.5) and selected colonies on LB/Cm 30 µg/ml 30°C overnight. Colony PCR was performed using forward and reverse primers (Supplementary table 3) and 85 to 95 % correct sized inserts were typically observed with some containing multiple inserts.

To generate plasmid DNA from these libraries, we typically excised 5×10^7 PFUs through infection of BUN25 cells as described earlier. The cells were scraped from plates and grown in 2 L of LB plus 13 g/l of circle growth 37°C for 7 to 8 h. Cesium chloride method was used to prepare DNA. DNAs were transformed into BW23474 F'DOT SbcC , and individual clones were sequenced using primer5' (TGTGGAAAGGACGAAACACC). Correct clones were individually rearrayed to form the final library.

Small RNA Northern Blots

293 cells were transfected in 10cm dishes at 60% confluency with 15 ug of shRNA plasmid DNA along with 5 ug of pDsRed-N1 (Clontech) using TransIT-LT1 (Mirus). 48hs post transfection, transfection efficiency was confirmed by estimating the percentage of cells expressing DsRed (~80%) and then total RNA was Trizol extracted and purified. Small RNA northern blots were carried out as described in ¹⁰⁵ using 30 ug RNA/lane. For hairpins targeting EGFP at starting position 481 (Fig 1b), northern probes were DNA oligos corresponding to the anti-sense strand (ccggcatcaaggtgaacttcaa) of the mature RNA.

Proteasome assays

Bacteria cultures were grown in 96 well plates for 36 h in GS96 media (Bio101). Plasmid DNA was extracted using Quiagen Ultrapure plasmids minipreps in a 96 well plate format. DNA concentrations were determined by mixing an aliquot of each sample with picogreen (Molecular Probes) and determining fluorescence on a Victor2 plate reader. HEK 293T cells were plated in 96 well optic plates (Corning) at 1×10^5 cells per ml. For the proteasome assay, 12.5 ng of the plasmid dsRed N-1 (Clontech), 5 ng of the

Zsprosensor (Clontech) and 75 ng of each individual shRNA construct were cotransfected per well using 0.3 μ l of LT-1 (Mirus) transfection reagent. After 24 h the transfection media was replaced. After 72 h, media was removed and replaced with PBS in order to read fluorescence. Fluorescence signals were read on a Victor2 plate reader. Signals in the green channel were normalized to transfection efficiency using customized scripts with fluorescence in the red channel serving as a normalization criterion. Cut-offs were assigned by using control shRNA transfections to determine the range for a negative outcome.

Plasmid Transfections and mRNA Quantitation

HeLa cells were seeded at 0.5×10^5 cells/well in 24-well plates and transfected 24 h later with 1 μ g/well of the appropriate plasmid. Each plasmid was delivered to 4 wells by use of Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocols. Transfection efficiency was determined by parallel transfection of a GFP-expressing plasmid and the percentage of fluorescent cells assayed by flow cytometry. For analysis of target gene mRNA knock down, cell lysates were collected 24 h after transfection, and total RNA was prepared by use of RNeasy columns (Qiagen) following the manufacturer's protocols. Messenger RNA quantitation was performed by Real-time PCR of reverse transcription products, using available Applied Biosystems TaqMan™ primer probe sets, and the percent mRNA remaining was determined by comparison with mRNA levels from cells treated with transfection reagent alone.

Table 3.1 Library coverage by functional group

Group	Human Genes	Human Hairpins	Mouse Genes	Mouse Hairpins
Cancer	859	3082	890	2524
Cell cycle	531	2166	482	1552
Checkpoint	123	541	116	367
DNA repair	118	512	130	355
DNA replication	238	961	248	719
Enzymes	2943	10456	2818	8302
GPCRs	669	2101	663	1795
Kinases	618	2648	575	2250
Dual Specificity Phosphatases	35	144	32	114
Tryosine Phosphatases	36	184	33	166
Phosphotases	206	765	187	628
Proteases	454	1431	441	1168
Proteolysis	302	1458	270	858
Signal Transduction	2650	9046	2541	7274
Protein Trafficking	476	1596	458	1309
Transcription	820	2865	767	2209
Apoptosis	581	2061	558	1538

3.5 Figures

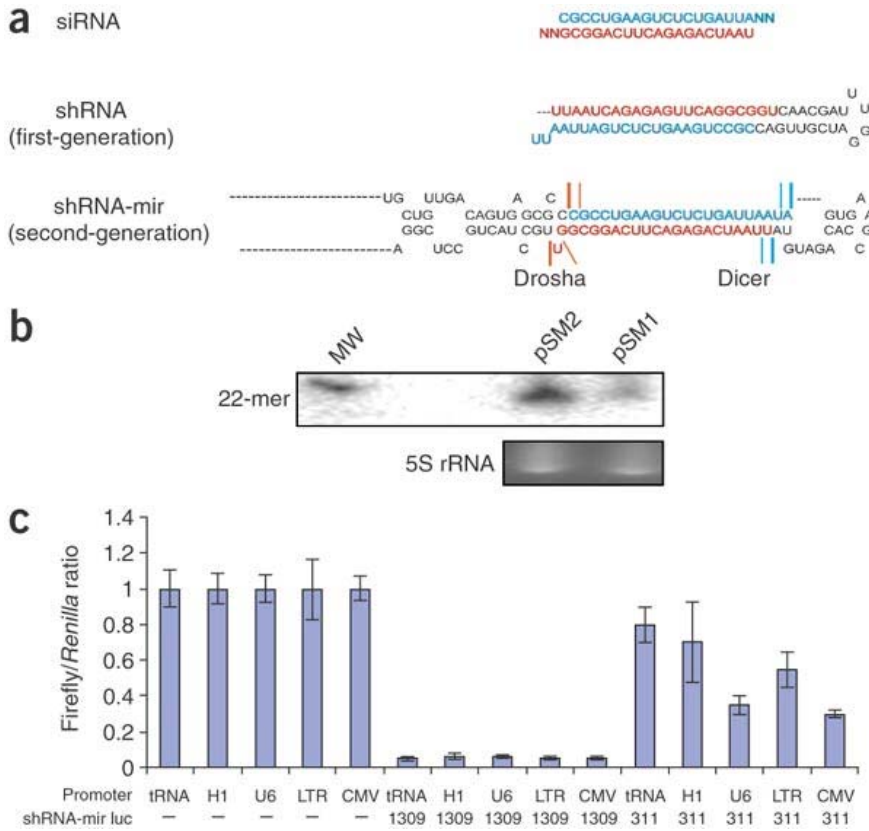


Figure 3.1. Design and structure of shRNA^{mir} cassettes.

(a) A comparison of the structures of several silencing triggers is shown. These include an siRNA, a portion of the shRNA precursor, as generated from our first-generation design in pSM1, and a segment of the shRNA^{mir} precursor produced by pSM2. The sequence of the target site (sense orientation) from firefly luciferase (luc1309, see c) is shown in blue (passenger strand) with the guide strand shown in red. For pSM2, mapped potential cleavage sites for Dicer and Drosha are indicated by blue and red lines respectively.

(b) Northern blotting was used to detect the mature small RNA produced after transfection of HEK-293T cells with shRNA and shRNA^{mir} cassettes expressed from pSM1 and pSM2, respectively, by the U6 snRNA promoter. In neither case was significant accumulation of pre-miRNA observed. Transfection rates were normalized using a co-delivered dsRED expression plasmid.

(c) Five different promoters (human tRNA^{Val}, Human H1 RNA, Human U6 snRNA, MSCV LTR and Human CMV IE, as indicated) were tested for their ability to drive shRNA^{mir} expression and silence luciferase in transient transfections. Two different shRNAs were used, a highly efficient shRNA (luc1309) and a less efficient shRNA (luc311). In each case, the level of firefly luciferase was normalized to a non-targeted *Renilla* luciferase. Controls with empty vectors lacking a hairpin insert are also shown.

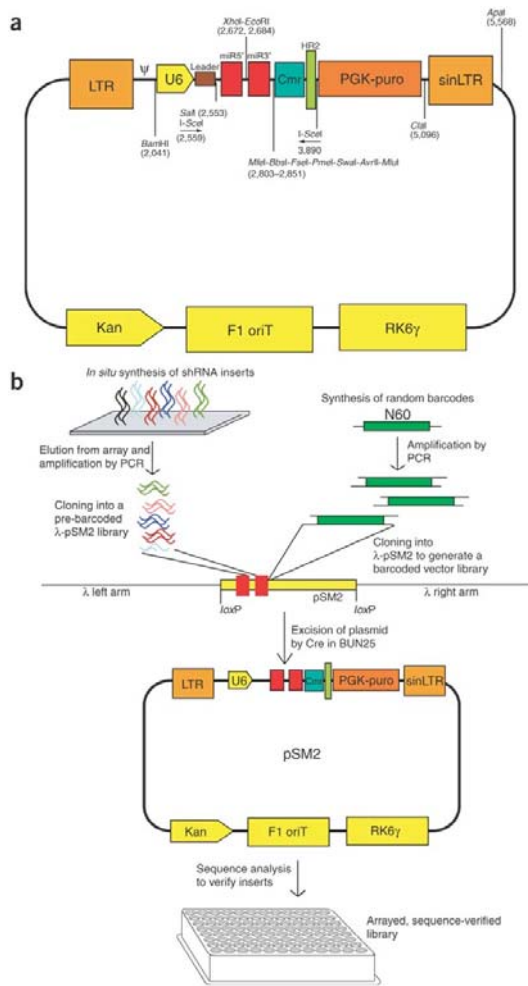


Figure 3.2. Construction of the second-generation library. (a) The pSM2c vector contains a U6 promoter, a U6 terminator following mir3', a self inactivating retroviral backbone; two bacterial antibiotic resistance markers kanamycin and chloramphenicol; a protein-dependent origin (RK6 γ); a mammalian selectable marker (puromycin) driven by the PGK promoter; a homology region (HR2) for use in bacterial recombination and a randomly generated 60 mer barcode sequence. The shRNA^{mir} inserts were cloned between the 5' and 3' flanking sequences derived from the mir-30 primary transcript. The nucleotide positions for sites in an excised version of an empty vector (no shRNA or barcode) are given. (b) Construction of the second-generation libraries began with the generation of a λ derivative of pSM2. A bar coded pre-library was generated by the ligation of PCR amplified random 60 mers into FseI-AvrII cleaved λ pSM2 to generate a bar-coded library pool (upper right). The bar-codes λ pSM2 was converted into a shRNA library by insertion of PCR amplified shRNA inserts prepared by *in situ* synthesis in pools of ~22,000 into the EcoR1-XhoI cleaved pre-library (upper right). Packaged phage were amplified and used to infect BUN25, which express Cre recombinase and *pir1-116* for pSM2 replication. Each excision event gave rise to a Kan^r+Cm^r resistant colony. These were pooled and used for preparation of library DNA. This, in turn, was transformed into BW F'DOT and individual colonies were selected for sequence analysis.

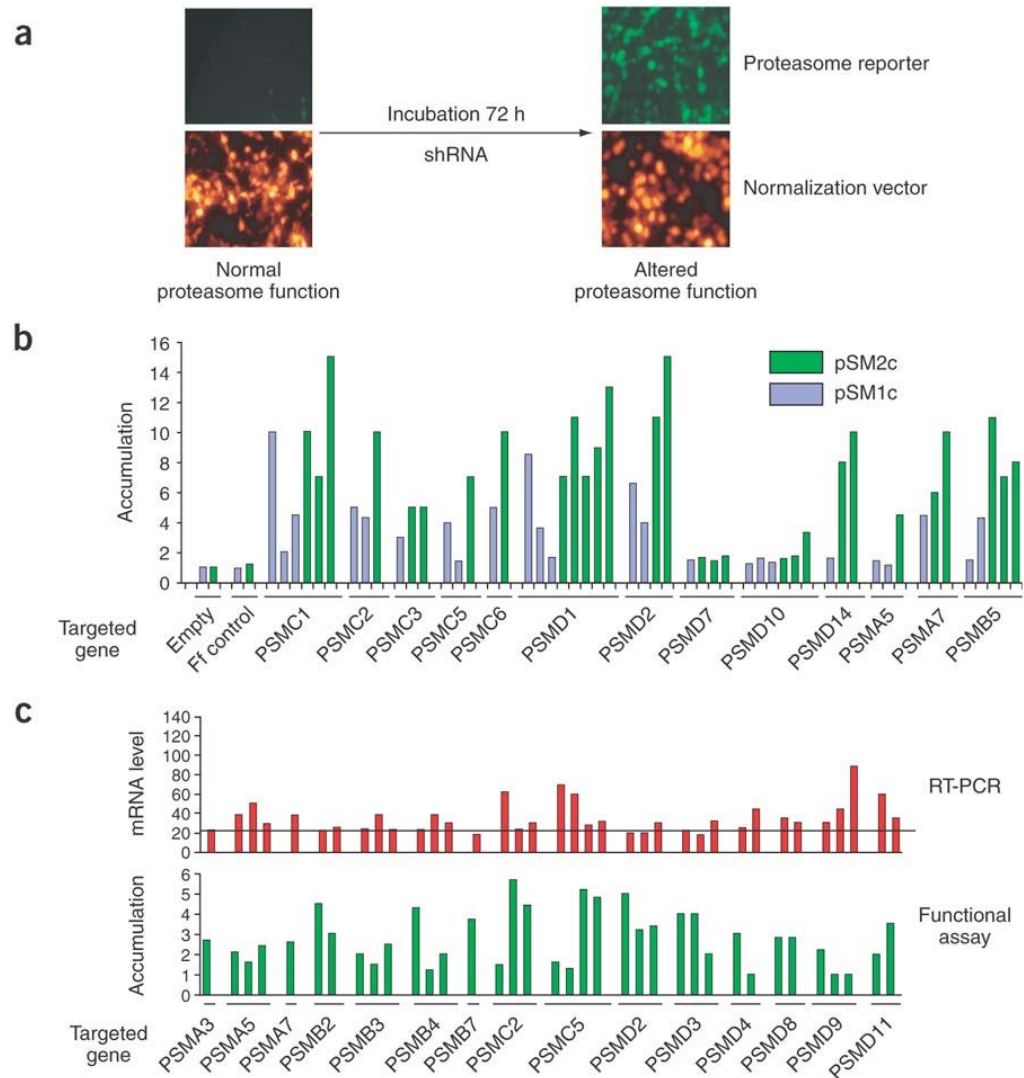


Figure 3.3 Validation of the second-generation library. (a) A schematic representation of the phenotypic assay for proteasome function (see text) is shown. (b) Thirteen proteasome subunits were chosen because of their representation in both the first- and second-generation libraries (for sequences see Supplementary table 1). ShRNA expression clones corresponding to each were assayed for activation of the proteasome reporter. Blue bars indicate first-generation clones while green indicate second-generation clones. In all cases, the activity of the proteasome reporter (green channel) was normalized for transfection using a dsRED expression plasmid. (c) In a separate study, 36 different proteasome shRNAs (for sequences see Supplementary table 1) were tested for their ability to suppress their target RNAs (upper panel). QRT-PCRs were performed 24 h after transfection of HeLa cells at an average efficiency of 80% as measured by a co-transfected normalization reporter (dsRED). The hypothetical maximum suppression, as calculated by transfection efficiency, is indicated by the black line. For comparison, functional assays for proteasome inhibition were performed in parallel (lower panel).

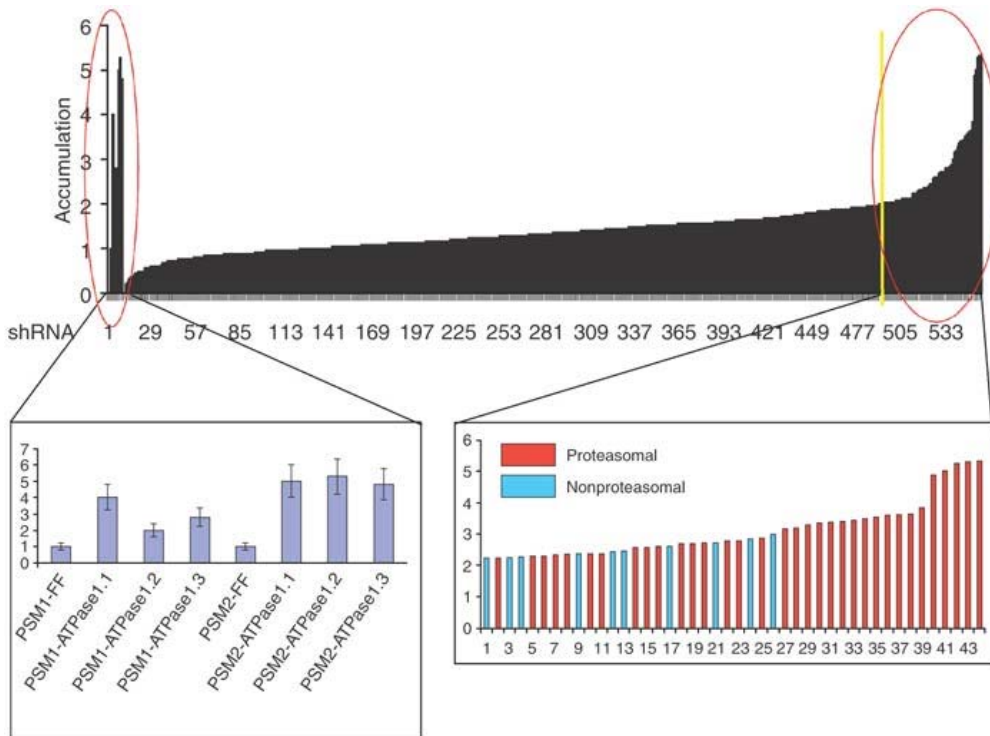


Figure 3.4. Performance of the second-generation library in a small-scale high-throughput screen. 47 shRNAs targeting proteasome subunits were distributed among a series of 562 hairpins targeting human kinases (upper panel). The lower left panel shows the negative (FF) and the positives controls (ATPase 1.1 to 1.3) from first (pSM1) and second (pSM2) libraries. The lower right panel shows the shRNAs that displayed accumulation of the proteasome reporter over the cut-off (2-fold or greater activation; yellow line). These are highly enriched for proteasome shRNAs (red). In blue are 10 additional non-proteasomal shRNAs that also scored positive in the screen. Of these, 5 were also positive on a retest of individual clones. The sequences of the shRNAs, in order from left to right, for the lower right panel are given in Supplementary Table 2.

Supplementary Figure 3.1

P53_1224 Guide Strand (Drosha site)

3'-TGTTTCATGTACACATT-5' RACE primer
 3'-AGGUGAUGUUCAUGUACACAUU-5' – predicted small RNA
 5'-...TTTTTTTTTTTTTCCACTACAAGTACATGTGTAA... RACE1
 ...TTTTTTTTTTTTTCCACTACAAGTACATGTGTAA... RACE2
 ...TTTTTTTTTTTTTCCACTACAAGTACATGTGTAA... RACE3
 ...TTTTTTTTTTTTTCCACTACAAGTACATGTGTAA... RACE4
 ...TTTTTTTTTTTTTCCACTACAAGTACATGTGTAA... RACE5
 ...TTTTTTTTTTTTTCCACTACAAGTACATGTGTAA... RACE6
 ...TTTTTTTTTTTTTCCACTACAAGTACATGTGTAA... RACE7
 ...TTTTTTTTTTTTTCCACTACAAGTACATGTGTAA... RACE8
 ...TTTTTTTTTTTTTCCACTACAAGTACATGTGTAA... RACE9
 ...TTTTTTTTTTTTTCCACTACAAGTACATGTGTAA... RACE10

P53_1224 Passenger Strand (Dicer site)

3'-TGTACATGAACATCAC-5' RACE primer
 3'-AUAAGUGUGUACAUGAACAUAC-5' – predicted small RNA
 5'...TTTTTTTTTTTTTATTACACATGACTTGTAGTG...-3' RACE1
 5'...TTTTTTTTTTTTTATTACACATGACTTGTAGTG...-3' RACE2
 5'...TTTTTTTTTTTTTATTACACATGACTTGTAGTG...-3' RACE3
 5'...TTTTTTTTTTTTTATTACACATGACTTGTAGTG...-3' RACE4
 5'...TTTTTTTTTTTTTATTACACATGACTTGTAGTG...-3' RACE5
 5'...TTTTTTTTTTTTTATTACACATGACTTGTAGTG...-3' RACE6
 5'...TTTTTTTTTTTTTATTACACATGACTTGTAGTG...-3' RACE7
 5'...TTTTTTTTTTTTTATTACACATGACTTGTAGTG...-3' RACE8
 5'...TTTTTTTTTTTTTATTACACATGACTTGTAGTG...-3' RACE9
 5'...TTTTTTTTTTTTTATTACACATGACTTGTAGTG...-3' RACE10

Luc_1309 Guide Strand (Drosha site)

3'-ACTTCAGAGACTAATT-5' RACE primer
 3'-UGGCGGACUUCAGAGACUAAUU-5' – predicted small RNA
 5'...TTTTTTTTTTTTT.CCGCCTGAAGTCTCTGATTAA...-3' RACE1
 5'...TTTTTTTTTTTTTACCGCCTGAAGTCTCTGATTAA...-3' RACE2
 5'...TTTTTTTTTTTTT.CCGCCTGAAGTCTCTGATTAA...-3' RACE3
 5'...TTTTTTTTTTTTTACCGCCTGAAGTCTCTGATTAA...-3' RACE4
 5'...TTTTTTTTTTTTTACCGCCTGAAGTCTCTGATTAA...-3' RACE5
 5'...TTTTTTTTTTTTTACCGCCTGAAGTCTCTGATTAA...-3' RACE6
 5'...TTTTTTTTTTTTTACCGCCTGAAGTCTCTGATTAA...-3' RACE7
 5'...TTTTTTTTTTTTTACCGCCTGAAGTCTCTGATTAA...-3' RACE8
 5'...TTTTTTTTTTTTTACCGCCTGAAGTCTCTGATTAA...-3' RACE9
 5'...TTTTTTTTTTTTTACCGCCTGAAGTCTCTGATTAA...-3' RACE10

PTEN_1137 Guide Strand (Drosha site)

3'-AAGGACGTCTTTCTCA-5' RACE primer
 3'-ACCCUAAAGGACGUCUUUCUGA-5' – predicted small RNA
 5'...TTTTTTTTTTTTTGGGATTTCTGCAGAAAGACT-3' RACE1
 5'...TTTTTTTTTTTTTGGGATTTCTGCAGAAAGACT-3' RACE2
 5'...TTTTTTTTTTTTTGGGATTTCTGCAGAAAGACT-3' RACE3
 5'...TTTTTTTTTTTTTGGGATTTCTGCAGAAAGACT-3' RACE4
 5'...TTTTTTTTTTTTTGGGATTTCTGCAGAAAGACT-3' RACE5
 5'...TTTTTTTTTTTTTGGGATTTCTGCAGAAAGACT-3' RACE6
 5'...TTTTTTTTTTTTTGGGATTTCTGCAGAAAGACT-3' RACE7
 5'...TTTTTTTTTTTTTGGGATTTCTGCAGAAAGACT-3' RACE8
 5'...TTTTTTTTTTTTTGGGATTTCTGCAGAAAGACT-3' RACE9
 5'...TTTTTTTTTTTTTGGGATTTCTGCAGAAAGACT-3' RACE10

PETN_1137 Passenger Strand (Dicer site)

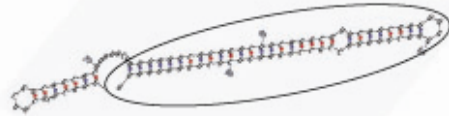
3'-AAAGACGTCCTTTAGG-5' RACE primer
 3'-AUUCAGAAAGACGUCCUUUAGG-5' – predicted small RNA
 5'...TTTTTTTTTTTTTAAGTCTTTCTGCAGGAAATCC-3' RACE1
 5'...TTTTTTTTTTTTTAAGTCTTTCTGCAGGAAATCC-3' RACE2
 5'...TTTTTTTTTTTTTAAGTCTTTCTGCAGGAAATCC-3' RACE3
 5'...TTTTTTTTTTTTTAAGTCTTTCTGCAGGAAATCC-3' RACE4
 5'...TTTTTTTTTTTTTAAGTCTTTCTGCAGGAAATCC-3' RACE5
 5'...TTTTTTTTTTTTTAAGTCTTTCTGCAGGAAATCC-3' RACE6
 5'...TTTTTTTTTTTTTAAGTCTTTCTGCAGGAAATCC-3' RACE7
 5'...TTTTTTTTTTTTTAAGTCTTTCTGCAGGAAATCC-3' RACE8

5'-...TTTTTTTTTTTAAAGTCTTTCTGCAGGAAATCC-3' RACE9
 5'-...TTTTTTTTTTTAAAGTCTTTCTGCAGGAAATCC-3' RACE10

Supplementary Figure 3.2:

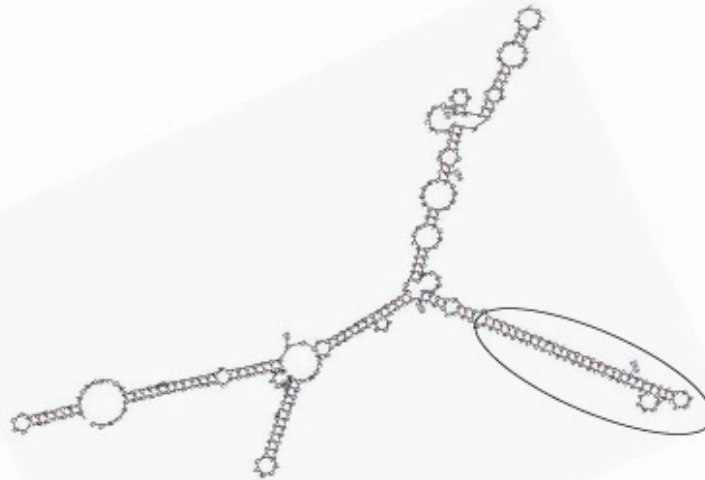
shRNA (first-generation)

gugcucgcuucgcagcacauauacuaUAAUCAGAGACUUCAGGCGUUCAACGAU*umgg*
 AUCGUUGACCGCCUGAAGUCUCUGAUUAAU



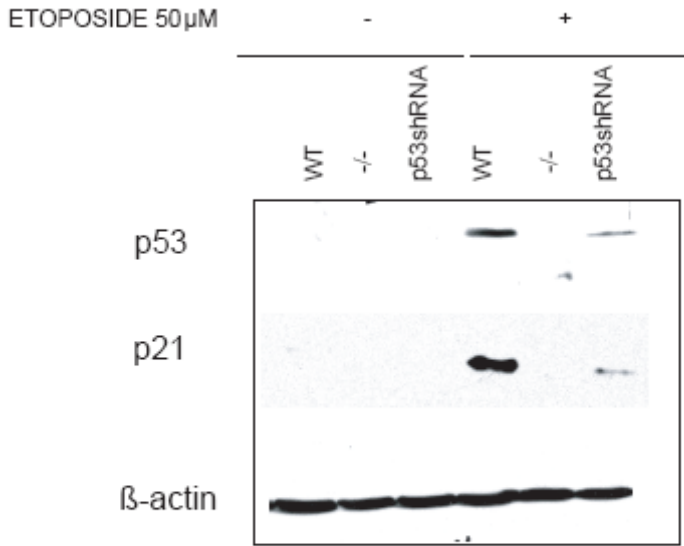
shRNA^{mir} (second-generation)

gugcucgcuucggcagcacauauacuagucgacuagggauaacaggguuuuuuuuugaauaggcucucaguacuu
uacagaauccguugccugcacaucuuggaaacacuugcugggaaucuuucagguuaaccaacagaaggcucga
gaagguaauuugcuguugacagugagcgc**CGCCUGAAGUCUCUGAUUAAU***Agugaagccaca*
*gangua*UAAUCAGAGACUUCAGGCGGUgccuacugccuggaaucaggggcacuuuagg
 agcauuuacuuguuuacuaaacugaaucuuucuaucucuugauacauu



Underline : Leader sequence
 CAPITAL: shRNA
CAPITAL BOLD: sense target sequence
Italic: shRNA loop

Supplementary Figure 3.3:



Supplementary Figure 3. Stable suppression by pSM2. HCT116 cells were infected with pSM2_hsP53_2120 (v2HS_93615) and selected as a population for resistance to puromycin. To induce p53, populations were treated with 50 microM etoposide for 24 prior to lysis for Western blotting. For comparison HCT116-p53null cells were also examined. Lysates were examined for levels of p53, a p53 target, p21, and -actin as a control. Titers were approximately 1x10⁶/ml.

Supplementary Table 3.1: ShRNAs used in Figure 3.

See website <http://www.nature.com/ng/journal/v37/n11/extref/ng1650-S4.pdf>

Supplementary Table 3.2:

Gene targeted	oligo ID	Acc. Number	sense/antisense shRNA sequence
CASK	v2MM_19396	NM_00368	CTGAAAGGATCACTGTTTA /
	5	8	TAAACAGTGATCCTTTTCAG
PSMA5	v2HS_170693	NM_00279	CTTTGCAGTTTGGAGAAGA /
	7	0	TCTTCTCCAAACTGCAAAG
LOC20004	v2HS_102066	XM_11409	GGAGAATTATTAAGAAAGA /
	8	9	TCTTTCTTAATAATTCTCC
CDC2L5	v2HS_24388	NM_00371	CTCGCTATGGTGAAACCAA /
		8	TGGTTTTACCATAGCGAG
PSMB4	v2HS_170730	NM_00279	CTGCGAGAAGTTCTGGAGA /
		6	TCTCCAGAACTTCTCGCAG
PSMC2	v2HS_170768	NM_00280	CCAGGGAGATTGGATAGAA /
		3	TTCTATCCAATCTCCCTGG
PSMD3	v2HS_170804	NM_00280	CAGGGCTTCTTCACTTCAA /
		9	TTGAAGTGAAGAAGCCCTG
PSMB5	v2HS_170736	NM_00279	CATCCTCCATCCTATTTAT /
		7	ATAAATAGGATGGAGGATG
PIK3C2G	v2HS_47794	NM_00457	GAGTGTTTCCCTGTAAAT /
		0	ATTTAACAGGGAAACACTC
PSMA3	v2HS_170678	NM_00278	CAGACAGAGTGGCCATGTA /
		8	TACATGGCCACTCTGTCTG
PSMC3	v2HS_170778	NM_00280	CCTCTATCATCTTCATTGA/TCAATGAAGATGATA
		4	GAGG
CLK2	v2HS_113094	NM_00129	CTCTCTATATACCTGTGAA /

		1_	TTCACAGGTATATAGAGAG
		NM_00503	CATATGAATTGTACAGAAT /
PLK1	v2HS_19709	0	ATTCTGTACAATTCATATG
		NM_00280	CACAGGGCCTGACACATTT /
PSMD2	v2HS_170801	8	AAATGTGTCAGGCCCTGTG
		NM_00281	GCCTCAAGAAGGAGAAAGT /
PSMD4	v2HS_170813	0	ACTTTCTCCTTCTTGAGGC
		NM_00280	GAGTTAAGGGAACAATTAA /
PSMC6	v2HS_170787	6	TTAATTGTTCCCTTAACTC
		NM_00728	CGCTACCACTTCTTCCTCT /
PTK9L	v2HS_91138	4	AGAGGAAGAAGTGGTAGCG
		NM_00280	GAAACATTATTCTAGGCAA /
PSMD2	v2HS_170803	8	TTGCCTAGAATAATGTTTC
		NM_00280	CACCGTCCATCGTGTTTAT /
PSMC1	v2HS_170763	2	ATAAACACGATGGACGGTG
		NM_00280	GCAAGTGTGCTGTGATCAA/TTGATCACAGCACAC
PSMC3	v2HS_170774	4	TTGC
		NM_00257	CCAAGAAAGAGCTGATTAT /
PAK1	v2HS_152621	6	ATAATCAGCTCTTTCTTGG
		NM_00281	CAAAGGTGGCCTGGGTTTA/TAAACCCAGGCCACC
PSMD10	v2HS_170838	4	TTTG
		NM_00280	CCGCAAAGTGTTACTATTA /
PSMD3	v2HS_170806	9	TAATAGTAACACTTTGCGG
		NM_00355	CTAGGTCTGCCTTCACTTT /
PIP5K1B	v2HS_33173	8	AAAGTGAAGGCAGACCTAG
		NM_00280	CCTGTGATGGGAGATTCAA /
PSMD2	v2HS_170798	8	TTGAATCTCCCATCACAGG
		NM_00662	GAGATGACAGATTTGACAA /
PLK2	v2HS_91307	2	TTGTCAAATCTGTATCTC
		NM_00281	GACATGCAGTCGGGTATTA /
PSMD11	v2HS_170842	5	TAATACCCGACTGCATGTC
		NM_00280	CAATCATGGTGGTGATATA/TATATCACCACCATG
PSMD1	v2HS_170794	7	ATTG
		NM_00281	CTGGATTAAGGCATTCTTA /
PSMD9	v2HS_170828	3	TAAGAATGCCTTAATCCAG
		NM_00279	CTCATCGTGGGTTTCGACT /
PSMA7	v2HS_170703	2	AGTCGAAACCCACGATGAG
		NM_00280	GCTCTGAACTGGTACAGAA /
PSMC5	v2HS_170785	5	TTCTGTACCAGTTCAGAGC
		NM_00279	CTCTATGTGTTACCAGTCA/TGACTGGTAACACAT
PSMB5	v2HS_170737	7	AGAG
		NM_00280	CTTAGTGGCATCTAAAGTA/TACTTTAGATGCCAC
PSMD1	v2HS_170792	7	TAAG
		NM_00281	CAGCTCAAATGCTACTACT /
PSMD8	v2HS_170823	2	AGTAGTAGCATTTGAGCTG
		NM_00281	CAGGGCACGAGTTATTTAA /
PSMD8	v2HS_170826	2	TTAAATAACTCGTGCCCTG
		NM_00280	CAGCAGTCTGCTTCCCAAT /
PSMC1	v2HS_170767	2	ATTGGGAAGCAGACTGCTG
		NM_00279	CTGATCTACATGAGAAGTA /
PSMB5	v2HS_170735	7	TACTTCTCATGTAGATCAG
		NM_00280	GGACCCACATATTTAAGAT /
PSMC2	v2HS_170770	3	ATCTTAAATATGTGGGTCC
		NM_00279	GAGATCAATCCCTCAAGAT /
PSMA7	v2HS_170706	2	ATCTTGAGGGATTGATCTC
PSMD1	v2HS_170797	NM_00280	CTCATATTGGGAATGCTTA/TAAGCATTCCCAATA

		7	TGAG
PSMC2	v2HS_170771	NM_00280 3	CTCGTTCAATGAGTGTTGA / TCAAACTCATTGAACGAG
PSMB4	v2HS_170726	NM_00279 6	CGCAACATCTCTCGCATT / TAATGCGAGAGATGTTGCG
PSMC5	v2HS_170784	NM_00280 5	CCAACAAGGTAGACCCATT / AATGGGTCTACCTTGTGG
PSMC1	v2HS_3	NM_00280 2	CCCAAACCTCGTACGGGAAT / ATTCCCGTACGAGTTTGGG

Supplementary table 3.3: Oligonucleotides used in construction of the library vectors

oligo names	oligo sequences
MZL393	AAATTTGGATCCGCGTTAAACATGAGTGGATAG
MZL51	TATATAGGATCCGGCGTTCGGCTCCTTGAGGG
MZL493	GCAAATGGCACATCTGTTTGGGTATAATCGCGCCCATGCTTTTTCGCCA GGTGTAGGCTGGAGCTGCTTCG
MZL494	CCGCGCTGTGTTCTGCGATGCGTTCCCAGTGTGGACGAAGATTTCTGTC CGGCATATGAATATCCTCCTTAGTT
MZL495	AATTAAGGATCCGCACGCGATCTTCTTCTTCC
MZL496	ATATATAAGCTTTGACATTCTGGATAGCGACC
MZL524	ATAACTTCGTATAGCATAACATTATACGAAGTTATGCGGCCGCATAA CTTCGTATAGCATAACATTATACGAAGTTATGCTG
MZL525	ATAACTTCGTATAATGTATGCTATACGAAGTTATGCGGCCGCATAACTTC GTATAATGTATGCTATACGAAGTTATCAGC
MZL541	ATGCGGCCGGCCGTTTAAACATTTAAATCCTAGGA
MZL542	CGCGTCCTAGGATTTAAATGTTTAAACGGCCGGCC
Barcodes	GAAGACTAATGCGGCCGGCCA (N)60GGGCCCTATAGTGAGTCGTATTA
Barcode primer 1	AAATTGCAATGAAGACTAATGCGGCCGGCCA
Barcode primer 2	ATATATGGACGCGTCCTAGGTAATACGACTCACTATAGGGCCC
Forward primer (for colony PCR)	GGACGAAACACCGTGCTCGC
Reverse primer (for colony PCR)	TTCTGCGAAGTGATCTTCCG

Literature Cited

1. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* 100, 57-70 (2000).
2. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nat Med* 10, 789-99 (2004).
3. Papac, R. J. Origins of cancer therapy. *Yale J Biol Med* 74, 391-8 (2001).
4. Frei, E., 3rd et al. The effectiveness of combinations of antileukemic agents in inducing and maintaining remission in children with acute leukemia. *Blood* 26, 642-56 (1965).
5. Chabner, B. A. & Roberts, T. G., Jr. Timeline: Chemotherapy and the war on cancer. *Nat Rev Cancer* 5, 65-72 (2005).
6. Sawyers, C. Targeted cancer therapy. *Nature* 432, 294-7 (2004).
7. Roche-Lestienne, C. et al. Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. *Blood* 100, 1014-1018 (2002).
8. Shah, N. P. et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* 2, 117-25 (2002).
9. Sordella, R., Bell, D. W., Haber, D. A. & Settleman, J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* 305, 1163-7 (2004).
10. van 't Veer, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530-6 (2002).
11. Agendia, Inc. Mammaprint Brochure
http://usa.agendia.com/en/for_healthcare_professionals_2.html
12. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-11 (1998).
13. Hannon, G. J. RNA interference. *Nature* 418, 244-51 (2002).
14. Birchler, J. A. & Kavi, H. H. Molecular biology. Slicing and dicing for small RNAs. *Science* 320, 1023-4 (2008).
15. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363-6 (2001).

16. Carmell, M. A. & Hannon, G. J. RNase III enzymes and the initiation of gene silencing. *Nat Struct Mol Biol* 11, 214-8 (2004).
17. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293-6 (2000).
18. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843-54 (1993).
19. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858-62 (2001).
20. Lee, Y. et al. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* 23, 4051-60 (2004).
21. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* 431, 343-9 (2004).
22. Han, J. et al. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18, 3016-27 (2004).
23. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415-9 (2003).
24. Song, J. J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* 10, 1026-32 (2003).
25. He, L. & Hannon, G. J. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 5, 522-31 (2004).
26. Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* 7, 719-23 (2005).
27. Liu, J. et al. Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science* 305, 1437-1441 (2004).
28. Tolia, N. H. & Joshua-Tor, L. Slicer and the Argonautes. *Nat Chem Biol* 3, 36-43 (2007).
29. Song, J.-J. & Joshua-Tor, L. Argonaute and RNA -- getting into the groove. *Current Opinion in Structural Biology* 16, 5-11 (2006).

30. Faehnle, C. R. & Joshua-Tor, L. Argonautes confront new small RNAs. *Current Opinion in Chemical Biology* 11, 569-577 (2007).
31. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 9, 22-32 (2008).
32. Parker, J. S. & Barford, D. Argonaute: a scaffold for the function of short regulatory RNAs. *Trends in Biochemical Sciences* 31, 622-630 (2006).
33. Davis, S. & Watson, J. C. In vitro activation of the interferon-induced, double-stranded RNA-dependent protein kinase PKR by RNA from the 3' untranslated regions of human alpha-tropomyosin. *Proc Natl Acad Sci U S A* 93, 508-13 (1996).
34. Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H. & Williams, B. R. Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* 5, 834-9 (2003).
35. Elbashir, S. M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411, 494-8 (2001).
36. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. & Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 16, 948-58 (2002).
37. Sarov, M. & Stewart, A. F. The best control for the specificity of RNAi. *Trends Biotechnol* 23, 446-8 (2005).
38. Paddison, P. J., Caudy, A. A. & Hannon, G. J. Stable suppression of gene expression by RNAi in mammalian cells. *Proc Natl Acad Sci U S A* 99, 1443-8 (2002).
39. Pei, Y. & Tuschl, T. On the art of identifying effective and specific siRNAs. *Nat Meth* 3, 670-676 (2006).
40. Jackson, A. L. et al. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotech* 21, 635-637 (2003).
41. Barik, S. RNAi in moderation. *Nat Biotech* 24, 796-797 (2006).
42. Wolters, N. M. & MacKeigan, J. P. From sequence to function: using RNAi to elucidate mechanisms of human disease. *Cell Death Differ* 15, 809-19 (2008).
43. Bernards, R., Brummelkamp, T. R. & Beijersbergen, R. L. shRNA libraries and their use in cancer genetics. *Nat Methods* 3, 701-6 (2006).
44. Chang, K., Elledge, S. J. & Hannon, G. J. Lessons from Nature: microRNA-based shRNA libraries. *Nat Methods* 3, 707-14 (2006).

45. Root, D. E., Hacohen, N., Hahn, W. C., Lander, E. S. & Sabatini, D. M. Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Meth* 3, 715-719 (2006).
46. Silva J, C. K., Hannon GJ, Rivas FV. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. *Oncogene* 51, 8401-9. (2004).
47. Dickins, R. A. et al. Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nat Genet* 37, 1289-95 (2005).
48. Dickins, R. A. et al. Tissue-specific and reversible RNA interference in transgenic mice. *Nat Genet* 39, 914-21 (2007).
49. Silva, J., Chang, K., Hannon, G. J. & Rivas, F. V. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. *Oncogene* 23, 8401-9 (2004).
50. Burgess, D. J. RNAi Genetic Screens for Chemotherapy Response Modifier Genes. (PhD dissertation, Watson School of Biological Sciences) (2007).
51. Hemann, M. T. et al. An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nat Genet* 33, 396-400 (2003).
52. Kaelin, W. G., Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer* 5, 689-98 (2005).
53. Aza-Blanc, P. et al. Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. *Mol Cell* 12, 627-37 (2003).
54. Westbrook, T. F. et al. A genetic screen for candidate tumor suppressors identifies REST. *Cell* 121, 837-48 (2005).
55. Kolfschoten, I. G. et al. A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. *Cell* 121, 849-58 (2005).
56. MacKeigan, J. P., Murphy, L. O. & Blenis, J. Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. *Nat Cell Biol* 7, 591-600 (2005).
57. Berns, K. et al. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* 428, 431-7 (2004).
58. Paddison, P. J. et al. A resource for large-scale RNA-interference-based screens in mammals. *Nature* 428, 427-31 (2004).

59. Brummelkamp, T. R., Nijman, S. M., Dirac, A. M. & Bernards, R. Loss of the cylindromatosis tumour suppressor inhibits apoptosis by activating NF-kappaB. *Nature* 424, 797-801 (2003).
60. Ngo, V. N. et al. A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* 441, 106-110 (2006).
61. Whitehurst, A. W. et al. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature* 446, 815-9 (2007).
62. Berns, K. et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 12, 395-402 (2007).
63. Soutschek, J. et al. Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* 432, 173-8 (2004).
64. Bitko, V., Musiyenko, A., Shulyayeva, O. & Barik, S. Inhibition of respiratory viruses by nasally administered siRNA. *Nat Med* 11, 50-5 (2005).
65. de Fougerolles, A., Vornlocher, H. P., Maraganore, J. & Lieberman, J. Interfering with disease: a progress report on siRNA-based therapeutics. *Nat Rev Drug Discov* 6, 443-53 (2007).
66. Tara, M. L., Howell, F. M. & Carl, D. N. Not miR-ly small RNAs: Big potential for microRNAs in therapy. *The Journal of allergy and clinical immunology* 121, 309-319 (2008).
67. Krutzfeldt, J. et al. Silencing of microRNAs in vivo with 'antagomirs'. *Nature* 438, 685-689 (2005).
68. He, L. et al. A microRNA component of the p53 tumour suppressor network. *Nature* 447, 1130-1134 (2007).
69. He, L. et al. A microRNA polycistron as a potential human oncogene. *Nature* 435, 828-833 (2005).
70. Bignell, G. R. et al. Identification of the familial cylindromatosis tumour-suppressor gene. *Nat Genet* 25, 160-165 (2000).
71. Lakhani, S. R. Putting the Brakes on Cylindromatosis? *N Engl J Med* 350, 187-188 (2004).
72. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs Exhibit Strand Bias. *Cell* 115, 209-216 (2003).

73. Carmell, M. A., Xuan, Z., Zhang, M. Q. & Hannon, G. J. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* 16, 2733-42 (2002).
74. Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear export of microRNA precursors. *Science* 303, 95-8 (2004).
75. Nakayashiki, H. et al. RNA silencing as a tool for exploring gene function in ascomycete fungi. *Fungal Genet Biol* 42, 275-83 (2005).
76. Tang, J., Erikson, R. L. & Liu, X. Checkpoint kinase 1 (Chk1) is required for mitotic progression through negative regulation of polo-like kinase 1 (Plk1). *Proc Natl Acad Sci U S A* 103, 11964-9 (2006).
77. Dasgupta, R. & Perrimon, N. Using RNAi to catch Drosophila genes in a web of interactions: insights into cancer research. *Oncogene* 23, 8359-65 (2004).
78. Fraser, A. Towards full employment: using RNAi to find roles for the redundant. *Oncogene* 23, 8346-52 (2004).
79. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281-97 (2004).
80. Reinhart, B. J. et al. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901-6. (2000).
81. Ketting, R. F. et al. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* 15, 2654-9. (2001).
82. Grishok, A. et al. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106, 23-34. (2001).
83. Knight, S. W. & Bass, B. L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* 293, 2269-71. (2001).
84. Hutvagner, G. et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834-8. (2001).
85. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231-5 (2004).
86. Landthaler, M., Yalcin, A. & Tuschl, T. The Human DiGeorge Syndrome Critical Region Gene 8 and Its *D. melanogaster* Homolog Are Required for miRNA Biogenesis. *Curr Biol* 14, 2162-7 (2004).

87. Gregory, R. I. et al. The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235-40 (2004).
88. Siolas, D. et al. Synthetic shRNAs as potent RNAi triggers. *Nat Biotechnol* 23, 227-31 (2005).
89. Schwarz, D. S. et al. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199-208 (2003).
90. Zeng, Y., Wagner, E. J. & Cullen, B. R. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* 9, 1327-33. (2002).
91. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. & Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 16, 948-58. (2002).
92. Chen, C. Z., Li, L., Lodish, H. F. & Bartel, D. P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* 303, 83-6 (2004).
93. Zeng, Y. & Cullen, B. R. Sequence requirements for micro RNA processing and function in human cells. *Rna* 9, 112-23 (2003).
94. Kawasaki, H. & Taira, K. Short hairpin type of dsRNAs that are controlled by tRNA(Val) promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells. *Nucleic Acids Res* 31, 700-7. (2003).
95. Brummelkamp, T. R., Bernards, R. & Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296, 550-553 (2002).
96. Zheng, L. et al. An approach to genomewide screens of expressed small interfering RNAs in mammalian cells. *Proc Natl Acad Sci U S A* 101, 135-40 (2004).
97. Stegmeier, F., Hu, G., Rickles, R. J., Hannon, G. J. & Elledge, S. J. A Lentiviral microRNA-based System for Single Copy Pol II Regulated RNAi in mammalian Cells. *Proc Natl Acad Sci U S A*, in press (2005).
98. Li, M. Z. & Elledge, S. J. MAGIC, an in vivo genetic method for the rapid construction of recombinant DNA molecules. *Nat Genet* 37, 311-9 (2005).
99. Li, X. et al. Generation of destabilized green fluorescent protein as a transcription reporter. *J Biol Chem* 273, 34970-5 (1998).
100. Elledge, S. J. & Walker, G. C. Phasmid vectors for identification of genes by complementation of *Escherichia coli* mutants. *J Bacteriol* 162, 777-83 (1985).

101. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97, 6640-5 (2000).
102. Cherepanov, P. P. & Wackernagel, W. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* 158, 9-14 (1995).
103. Chalker, A. F., Leach, D. R. & Lloyd, R. G. *Escherichia coli* sbcC mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* 71, 201-5 (1988).
104. Caudy, A. A., Myers, M., Hannon, G. J. & Hammond, S. M. Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev* 16, 2491-6. (2002).
105. Silva, J. M. et al. Profiling essential genes in human mammary cells by multiplex RNAi screening. *Science* 319, 617-20 (2008).
106. Lucito, R. a. W., M. Preparation of Slides and Hybridization. (ed. Sambrook), e. D. B. a. J.) (Cold Spring Harbor Press, Cold Spring Harbor, NY., 2003).
107. Ambion website <http://www.ambion.com/techlib/resources/arrays/labeling.html> viewed July 20,2008. (2008).
108. Ramdas, L. et al. Sources of nonlinearity in cDNA microarray expression measurements. *Genome Biol* 2, RESEARCH0047 (2001).
109. GE Healthcare-Amersham. Cy3-dUTP. http://www1.gelifesciences.com/aptrix/upp00919.nsf/content/71985FA39EE4F72CC1256EB40044AD3A?OpenDocument&Path=Catalog&Hometitle=Catalog&entry=3&newrel&LinkParent=C1256FC4003AED40-2359E91E6518BB7EC1257019004907EE_RelatedLinksNew-C821BEC677D8448BC1256EAE002E3030&newrel&hidesearchbox=yes&moduleid=164367
110. Ambion, Inc. in Amino Alkyl MessageAmp Instruction Manual. http://www.ambion.com/techlib/prot/fm_1752.pdf
111. Wall, M. E., Andreas Rechtsteiner, Rocha, LM. Singular value decomposition and principal component analysis". (Kluwer, Norwell, MA 2003).
112. Alter, O., Brown, P. O. & Botstein, D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A* 97, 10101-6 (2000).
113. Haney, S. A. Increasing the robustness and validity of RNAi screens. *Pharmacogenomics* 8, 1037-49 (2007).

114. Moore, D. The Basic Practice of Statistics (Paper) & Student CD. (W. H. Freeman, San Francisco, 2006).
115. Roberts, F. a. R., D. (2008).
116. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98, 5116-21 (2001).
117. Barry, W. T., Nobel, A. B. & Wright, F. A. Significance analysis of functional categories in gene expression studies: a structured permutation approach. Bioinformatics 21, 1943-9 (2005).
118. Elledge, S. J. & Hannon, G. J. An open letter to cancer researchers. Science 310, 439-41 (2005).
119. Lucito, R. et al. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. Genome Res 13, 2291-305 (2003).

Chapter 4

High Throughput RNA interference barcode screens as a tool for
discovering gene function

Abstract:

High throughput screens in cultured mammalian cells can now be performed using RNAi libraries such as our own Hannon-Elledge library. Each hairpin is linked to a unique 60 nucleotide identification sequence, which serves as a barcode and allows us to virtually count the number of cells that contain a specific hairpin in a cell population. Small changes in barcode copy number can be monitored through the use of microarray technology. The barcode can be amplified from a cell's genomic DNA and fluorescently labeled to produce a probe that is hybridized to a microarray. This chapter focuses on developing procedures for hybridizing barcode microarrays and analyzing data from these hybridizations. We have optimized our probe labeling methods, probe size and hybridization conditions using library plasmid DNA in a Nimblegen platform. This optimized protocol allows us to detect 78.6% of probes within 1 standard deviation above the mean background from a complex mixture of approximately 1500 hairpins. In addition, by using two color hybridization (Cy3 and Cy5) we can detect control subsets of hairpins known to be depleted from a sample population.

4.1 Introduction

High throughput screens in cultured mammalian cells can now be performed due to the creation of extensive RNAi libraries such as the Hannon-Elledge shRNA library. Our current library consists of over 100,000 shRNA expression cassettes targeting 32,216 human genes and 30,629 mouse genes.⁴⁴ This library provides a formidable functional genomics tool for studying the molecular vulnerabilities of cancer. A distinguishing feature of this library is presence of DNA barcodes which allow quantitation of individual hairpins within a population of cells. These barcodes make it easier to conduct

high-throughput assays as many hairpins can be screened in parallel rather than using a serial (one by one) approach. Single hairpin approaches require both considerable infrastructure and substantial investment for each cell line and hairpin to be screened¹⁰⁶; parallel approaches can be performed by a single individual using standard tissue culture tools. shRNA expression can be normalized through retroviral knockdown. Retroviral infection of small RNA triggers enables us to monitor the effects of gene knockdown over long periods of time and broadens the scope of our assays. Perhaps most importantly, cells can be examined in competitive assays where thousands of different hairpins are integrated into different cells in a single plate. This type of screening allows both positive and negative selection of thousands of hairpins at a time. This screening method is ideal for investigation of cancer cell lines as it could identify genes important to cancer. A variation of this technique is the ability to conduct synthetic genetic screens. We can search for genes that modify sensitivity to chemotherapeutic drugs that are currently used in the clinic in an effort to help determine the appropriate treatment for cancer patients. Since this is an emerging technology, development of a standardized barcode screening strategy and improved protocols for microarray hybridization is required.

Barcode screens are inherently complex because even the simplest screens involve the use of highly technological methods. This complexity can be increased further if we examine multiple treatment groups or attempt to survey thousands of genes. Barcode screens do provide us with the ability of generate a lot of new information over a relatively short period of time. Because this is a new technology, optimized protocols that allow extraction of reliable and accurate data are needed. As of 2005, only positive

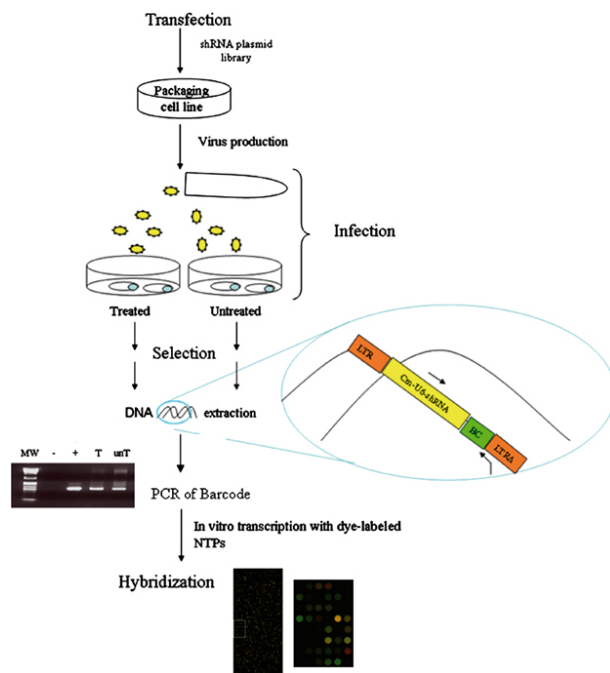
selection screens using pooled hairpins had been published.^{57, 58} Positive selection screens involve identifying genes that give a robust signal in an assay or on a microarray. Negative selection screens involve identifying genes that are lost from a population, which is harder to visualize on a microarray (see below).

Methods must be standardized if we are to compare screens between multiple people within a lab or wish to collaborate with other labs on a screening project. The smallest variable can change the data on a microarray and so it is important to use methods that are as consistent as possible. In addition, different types of data analysis can generate different data from a microarray. This can result in multiple lists of candidates for validation. We are interested in developing data analysis methods that could be used for RNAi screens. This chapter focuses on developing procedures for hybridizing barcode microarrays and analyzing data from these hybes. While our previous methods for hybridization were adequate for detecting robust positive signals, they were inadequate for detection of probes that are depleted from a population or for examining more than several thousand genes. Our microarray intensities were low and our data analysis was rudimentary. Thus, considerable time and effort was spent to develop hybridization methods for detection of depleted probes and quality control methods to decrease variability.

4.2 Experimental System

The overall scheme for a barcode screen is shown in figure 4.1. These screens begin with transfection of a shRNA plasmid library into a retroviral packaging cell line to

produce viruses containing our short hairpin constructs. Use of a viral vector allows increased efficiency of shRNA construct delivery, and generates a stable, selectable integrant. Infected cells are selected for by resistance to puromycin and then placed into control and experimental groups. Following further selection (chemotherapeutic or otherwise), cells are lysed and genomic DNA is extracted. Integrated shRNA constructs containing a 60 nucleotide DNA barcode are used to track the fate of individual shRNAs in a complex population by monitoring relative barcode levels on a microarray. The barcode is amplified from the genomic DNA using a primer containing a T7 RNA polymerase promoter, allowing *in vitro* transcription of fluorescently labeled RNA. These RNA probes are hybridized to a microarray containing complementary sequences. This technology allows us to monitor both enrichment and depletion of hairpins by monitoring their barcode proxy. Two color microarrays are used in which pre-transfection pooled *E.coli* plasmid DNA serves as a reference channel and barcodes amplified from extracted



cellular genomic DNA serve as the experimental channel. Hairpin-harboring cells that increase in number relative to the rest of the population will be reflected by an increase of barcode signal on the array. Conversely, cells carrying hairpins that decrease in the population will be reflected by a reduction of barcode signal.

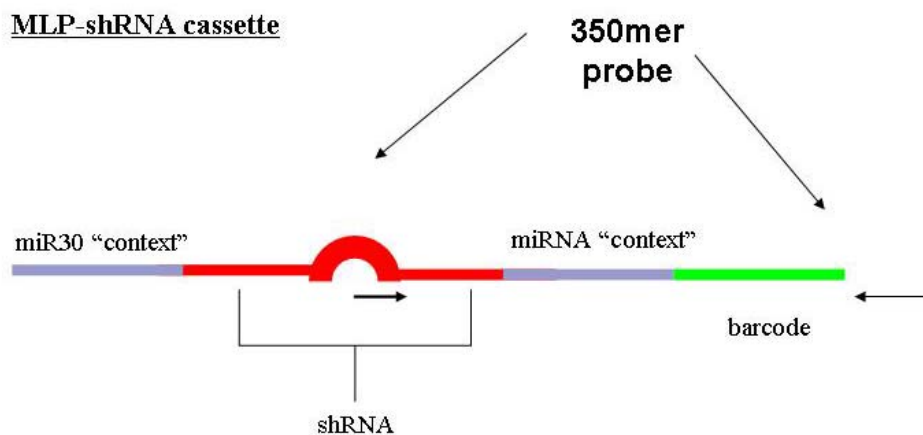
Figure 4.1: A typical RNAi screen in a stably selected population. (taken from Silva et al 2004).

The Hannon Elledge library design allows the use of DNA barcodes or hairpin PCR products as microarray probes (see figure 4.2), which can be used as independent identifiers during microarray hybridization. Work by Paddison et al showed that probes that contain the entire hairpin gave substantially lower hybridization signals relative to the corresponding barcode probes, indicating that using entire hairpin probes is not an optimal strategy for monitoring shRNA representation in a population.⁵⁸ We have been able to use the sense strand sequence of the hairpin as probes by using primers that begin in the loop of the hairpin (see figure 4.2). This is especially useful as the barcodes of all library hairpins have not yet been sequenced. Half-hairpin oligos on the array are designed to contain a tandem repeat of the 22 nucleotide sense strand rather than the both the sense and antisense strands of the hairpin, eliminating any potential secondary structure that could impede array hybridization. The array platform used for our experiments is a Nimblegen custom design 12-plex microarray. These arrays contain twelve identical subarrays consisting of 13,000 features of custom sequence. All hybridizations are two color (cy5/cy3) competitive hybridizations against a common reference. For the experiments in this chapter, amplified PCR products from *E. coli* plasmid DNA were used as reference.

Enrichment or depletion of hairpins can result in small fold changes on a microarray. In order to make sure we could detect even the smallest changes in abundance, we began to reevaluate our hybridization protocol. We had concerns that our probe intensities were poor and efforts were made to improve the overall quality of our hybridizations. Our hybridizations and labeling methods were lengthy and prone to error. The lab's current protocol is based on methods originally designed by Patrick Paddison.

Paddison's method amplified DNA barcodes and then used them as templates for T7 *in-vitro* transcription reactions (T7 Mega Short Script, Ambion) labeled with either Cy-3 or Cy-5 UTP (Amersham) to make fluorescent probes. Reactions were carried out at 37 degrees for 8 hours in the dark using 1.2 mM unlabelled UTP and 0.2uM labeled UTP. Probes were purified from unincorporated nucleotides using RNAeasy spin columns (Qiagen). These 950 nucleotide probes were hybridized to microarrays of complementary oligos. While previous experiments in the lab were using arrays with only 24mers printed on the chips, we saw increased probe performance using arrays with 50mer oligos (Ken Chang, unpublished).

Figure 4.2: Diagram of a PCR product from the MLP shRNA vector. This product can be used to hybridize both half-hairpin probes and barcode probes. The arrows represent the location of the PCR primers.



4.3 Results

4.3.1 Different labeling methods result in different frequencies of incorporation of labeled nucleotides

Our main goals were to improve the quality of our microarray hybridizations. To this end, we re-examined each step of our microarray hybridization protocol, beginning

with the quality of our labeled probes. We had concerns that the fluorescent intensity of our probes was poor. Three different labeling methods were examined:

- Direct incorporation of fluorescent dye through T7 *in-vitro* transcription.

This method involves a simple T7 RNA polymerase reaction using a fluorescent labeled UTP nucleotide. Unfortunately, T7 polymerase has difficulty incorporating bulky modified nucleotides (see Figure 4a for an example of a cyanine dye modified nucleotide). To obtain a sufficient yield of fluorescent RNA, the reaction must proceed for 6 to 8 hours. Specific activity of probes generated in this manner is quite low; total yield (in mg RNA) is high, however.

- Random Primer Labeling

This procedure uses random primers to make fluorescent DNA probes. Klenow DNA polymerase is used to incorporate Cy labeled dCTP into high specific activity probes in a 2 hour reaction.¹⁰⁷ A disadvantage of this procedure is the requirement for more starting material in comparison to the other methods presented.

- Amino Allyl labeling

This is a modification of the direct T7 labeling method above, in which amino-modified UTP (amino allyl-UTP) is incorporated during transcription. Amino allyl UTP contains a reactive amino group on a 2-carbon spacer attached to the methyl group on the base portion of UTP (see figure 4.3b). Due to this relatively small modification, these amino-modified nucleotides are incorporated at efficiencies similar to that of unmodified NTPs. NHS-activated cyanine dyes are coupled to purified amino-allyl modified RNAs

resulting in probe populations with equivalent labeling efficiencies¹⁰⁸. This method is appealing as amino allyl RNA is quite stable and can be frozen until needed.

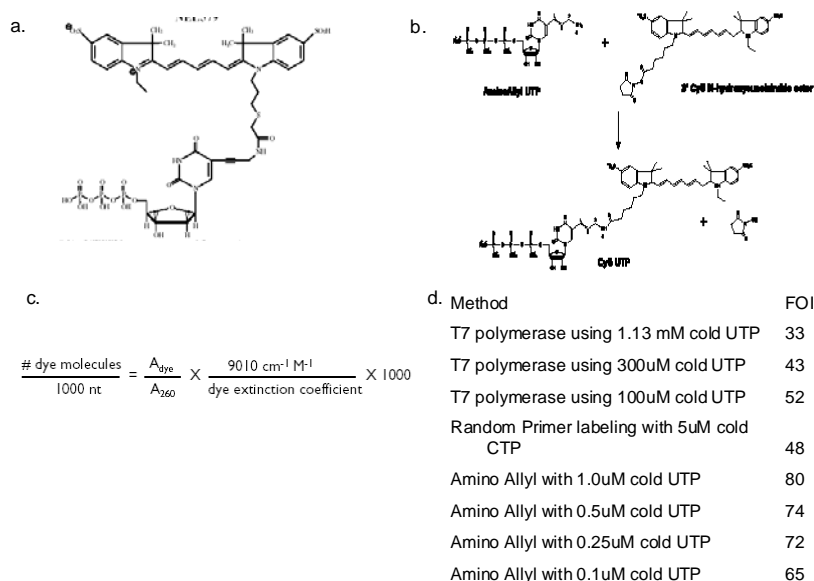


Figure 4.3 Labeling methods

a. Diagram of a cyanine labeled nucleotide.¹⁰⁹ b. Amino Allyl coupling reaction with an NHS modified dye to generate a Cy labeled nucleotide.¹¹⁰ c. Formula for calculating the frequency of incorporation of Cy dyes. d. Table indicating the different frequency of incorporations that resulted from the different labeling methods.

Independent reactions of each method were performed using plasmid DNA as template. The absorbance measurements of the final products were taken at 260nm (DNA) and 550 nm (for the Cy dye) using the nanodrop. The frequency of incorporation (FOI) was calculated using the formula in figure 4.3c. The FOI is the number of dye molecules incorporated per 1000 nucleotides. Results are listed in Figure 4.3d. The lab-standard method of using 1.13mM unlabelled UTP resulted in a relatively low FOI of 33. We titrated the amounts of unlabeled UTP in the T7 and amino allyl labeling reactions to see how high an FOI could be obtained. As the amount of unlabeled UTP was decreased, the frequency of incorporation increased. The maximum FOI obtained using the T7 polymerase reaction was 52. For the amino allyl labeled samples, the FOI increased to a

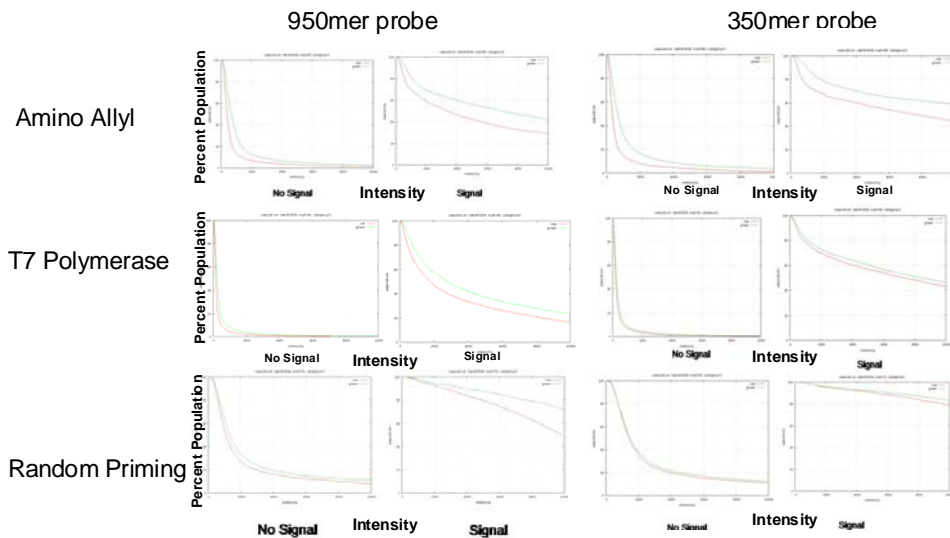
maximum of 83.9. The FOI for the random primer labeling method was 52.1. We decided to compare the hybridization efficiencies of probes generated by the different labeling methods to our current labeling method. We were concerned that too high an FOI could result in quenching, a reduction in fluorescence which occurs when multiple dye moieties are in close proximity.¹¹¹ We decided to hybridize samples with an FOI of approximately 50 in comparison to our current method, which had an FOI of 33. We wanted to explore how these differently labeled probes would perform on a microarray.

4.3.2 Probe size and quantity affect microarray intensities
(Sample preparation by Despina Siolas, Hybridization by Ken Chang, Analysis by Ravi Sachidanadam)

We were also interested in optimizing the size and amount of labeled probe used in our microarray hybridization, in an attempt to improve our ability to see changes in barcode abundance. While previous experiments in the lab were using arrays with only 24mers printed on the chips, we saw increased probe performance using arrays with 50mer oligos (Ken Chang, unpublished). Using Nimblegen specified conditions, we performed test hybridizations of 1,489 barcodes from *E. coli* plasmid DNA. These barcodes came from three different pooled functional sets of hairpins: the kinase and phosphatase set contain 1300 hairpins total and the c600 set, a small pool of miscellaneous cancer and proteasome related hairpins, contains 189 hairpins. We used these functional sets as samples because they would be the actual pools of hairpins that would be used in our subsequent screens. To see how well we could detect a reduction in barcode abundance, we spiked in c600 DNA at 1/10 the amount of the

kinase/phosphatase set for the Cy5 channel. Equal amounts of all hairpins were hybridized in the Cy3 (reference) channel.

Figure 4.4: Smaller probes have higher intensity signal on a microarray hybridization. Analysis by Ravi Sachidanadam. The graphs with no signal show the intensity level of unused spots on the microarray. The graphs with signal represent the intensity levels of probes that are expected to be hybridizing to the array.



The graphs in Figure 4.4 illustrate the signal intensities of the probe population in both the red and green channels. Plots were made using a program designed by Ravi Sachidanadam. The graphs in the right columns of each pair show the intensity levels of the probes present in our experiment. Graphs in the left column show the intensities of spots whose probes are not present. The right column shows the intensities of spots that should have signal and the left column are the unused spots that should have no signal. These graphs allow us to determine the signal to noise ratio by using the turning point of the graph as a cutoff between real and background signals. Larger probes reduce signal intensity. The 350mer probes yield a higher intensity signal compared to background than the 950mer probes across all different labeling methods (see figure 4.4). Quantity of

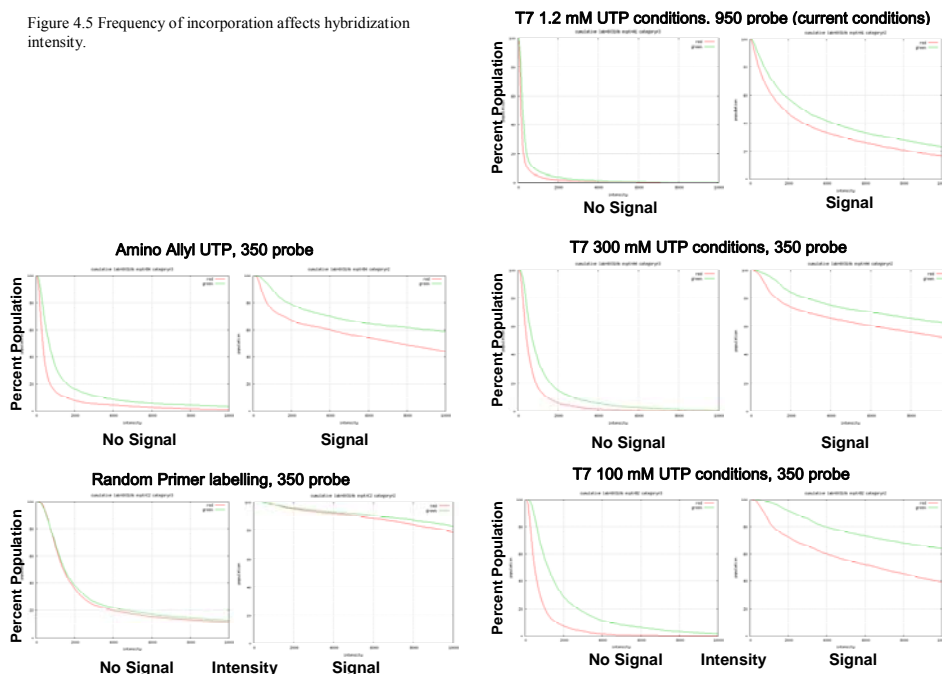
probe hybridized does not significantly affect signal intensity. Changing the probe amount from 1 microgram to 180 nanograms only moderately affected intensity when using T7 in vitro transcribed probes (data not shown). We were interested in using lower amounts of probe because of limited available input genomic DNA.

4.3.3 Frequency of incorporation affects hybridization intensity

(Analysis by Ravi Sachidanadam)

The hybridization intensities from our current protocols for probe labeling methods and sizes are shown in Figure 4.5. Random primer labeling gave the highest signal and best signal to noise ratio compared to all other labeling methods profiled. An analysis by Scott Powers shown in figure 4.6 demonstrates this result. These graphs show the probe intensities of the sequences present (expected) and the biological background of unused spots on the microarray. Note that in this case, “background” refers to spots on the array with no corresponding probe. This analysis also shows the percentage of probes with signals at (n) standard deviations (SD) above the mean background. Of note, more than 78% of the spots from random primed probes had signal at least 1SD above mean background. Our lab-standard method using 950mer probes in a T7 in vitro transcription had only 59% of spot intensities at least 1SD above above the mean background. Intensities of random-primed probes were well distributed across the assay quantitative range; T7 probes generated more spots of lower intensity, although a few features were beyond the quantitative range of the assay.

Figure 4.5 Frequency of incorporation affects hybridization intensity.



4.3.4 Random primer labeling with smaller probes is the best method for doing negative selection screens (Analysis by Scott Powers and Ken Chang)

In addition to studying the intensity of our hybridizations, we were interested in seeing if we could detect actual reductions in barcode abundance on our microarrays, an absolute necessity if we are to do negative selection screens. Ken Chang's analysis of our hybridizations using Spotfire is shown in Figure 4.7. The spots that are expected to hybridize are represented in blue and the spots that are considered background are in red. The c600 set is present at 1/10 the amount of the other hairpins in the Cy5 channel. The hybridizations for 300uM cold UTP and 100uM cold UTP using both direct and aminoallyl T7 *in vitro* transcription had a considerable amount of intensity variability. The random primer labeled 350mer probes had the lowest background signal and showed the depletion of the c600 set of probes clearly. With a reasonable expectation of barcode performance on a microarray, we began to assess our data analysis and hybridization methods.

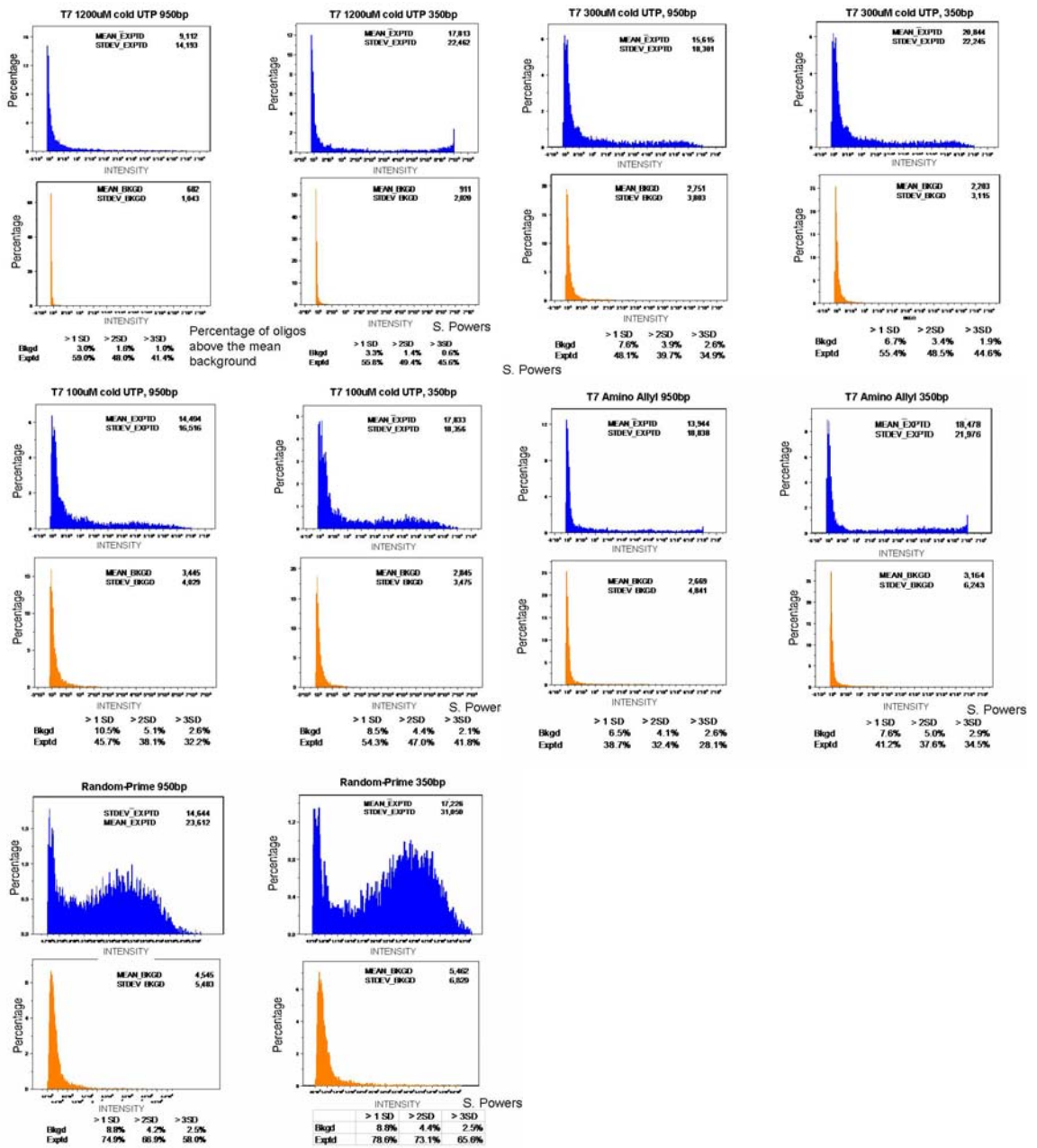
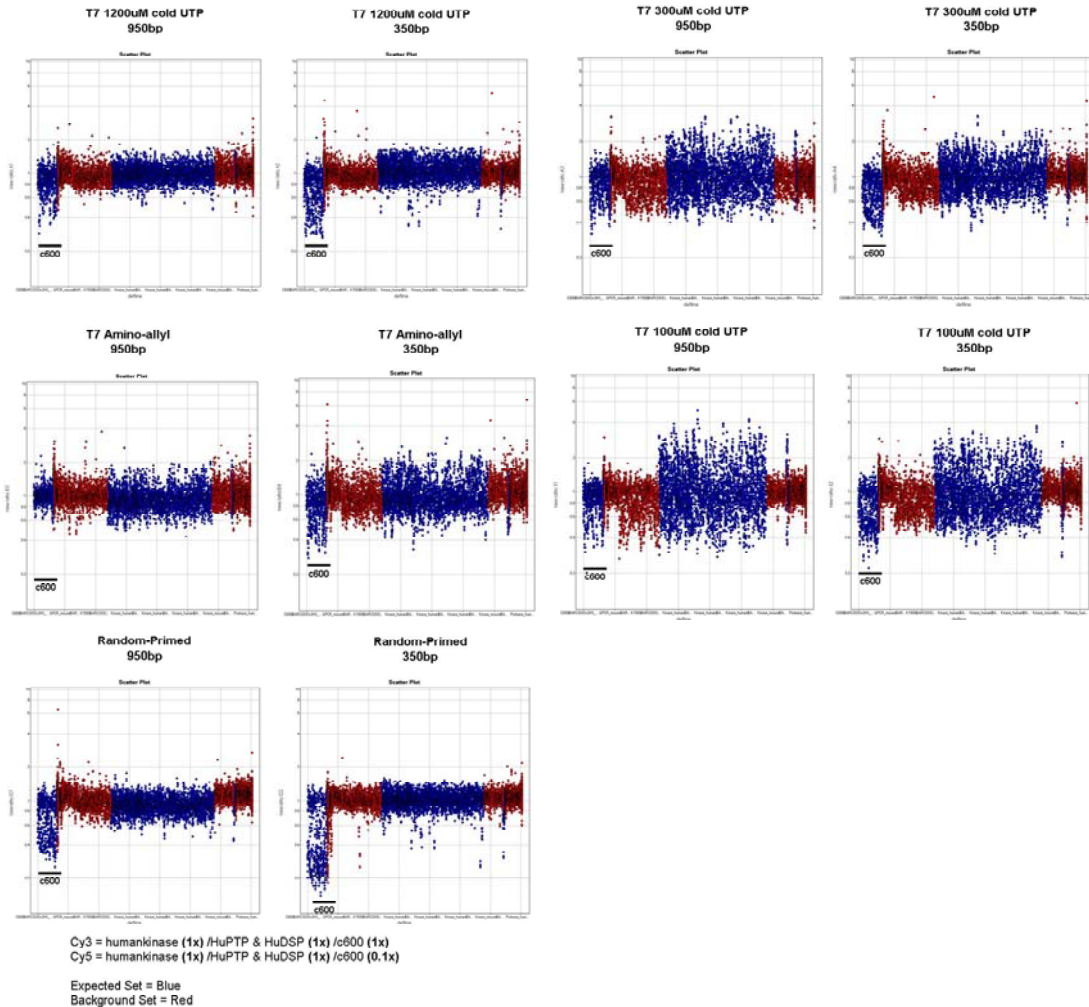


Figure 4.6: Frequency of incorporation affects hybridization intensity. Analysis by Scott Powers shows the random primer labeling method detects the highest amount of probes above background.

Figure 4.7: Random primer labeling with smaller probes is the best method for doing negative selection screens. Analysis by Ken Chang. Probes expected to hybridize are in Blue. Background set of probes represented in red. A group of hairpins called the c600 set were amplified by PCR at 1/10 of the amount of the other probes. This group is shown to be depleted clearly on the microarray using 350mer probes generated by the random-primer labeling.



4.3.5 The biggest source of variability for our hybridizations is the day of hybridization. (Screen by Jose Silva, Analysis by Joel Parker)

Having established labeling and hybridization protocols that allowed us to identify depleted or enriched barcodes from a mixed population, we decided to enlist the help of bioinformaticist Joel Parker to aid our data analysis. Joel analyzed a screen conducted by Jose Silva searching for essential genes in MCF10A cells. These cells were infected with 6000 hairpins and genomic DNA was collected on day 2 or day 4 after

infection and then once each week for 2 weeks.¹⁰⁶ This screen was conducted in biological triplicate. Samples were hybridized to Nimblegen arrays and Joel examined the data by principal component analysis (PCA).¹¹² PCA examines sources of variability in the data by identifying directions called principal components.¹¹³ The first component is the one responsible for greatest variability.¹¹⁴ This method is commonly used in microarray experiments to detect outliers, corroborate expected patterns and identify extraneous factors.¹¹⁵ Principal component analysis (PCA) is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the data set.¹¹⁶ Samples can then be plotted, making it possible to visually assess similarities and differences between samples and determine whether samples can be grouped.¹¹³ PCA can serve as a useful first step for analyzing data because as the principal components are uncorrelated, they may represent different aspects of the samples.¹¹³ As much information will typically be lost in two- or three-dimensional visualizations, it is important to systematically try different combinations of components when visualizing a data set.¹¹³ As shown in Figure 4.8, the day of hybridization is the biggest source of variability for the experiment. 1,157 probes were significantly different at a 5% false discovery rate. False discovery rate is the percentage of genes incorrectly identified as differentially present (depleted or enriched) by chance. Day of hybridization accounts for 57% of the variation among those probes. Using PCA we identified other sources of variability including the time of PCR and genomic DNA preparation (data not shown). This result demonstrated the sensitivity of this procedure to technical variation. Ideally, genomic DNA preparation, PCR amplifications, probe labeling, and hybridizations should all be conducted at the same time for all the samples in a study. Unfortunately,

this may not be practical or feasible when dealing with multiple biological replicates or numerous experimental groups with multiple timepoints. In this case, it would be better to separate the samples by replicate group, not timepoint, so that all samples from each replicate group (A, B, or C) are treated the same. We now frequently use PCA as a method for assessing variability in our screens.

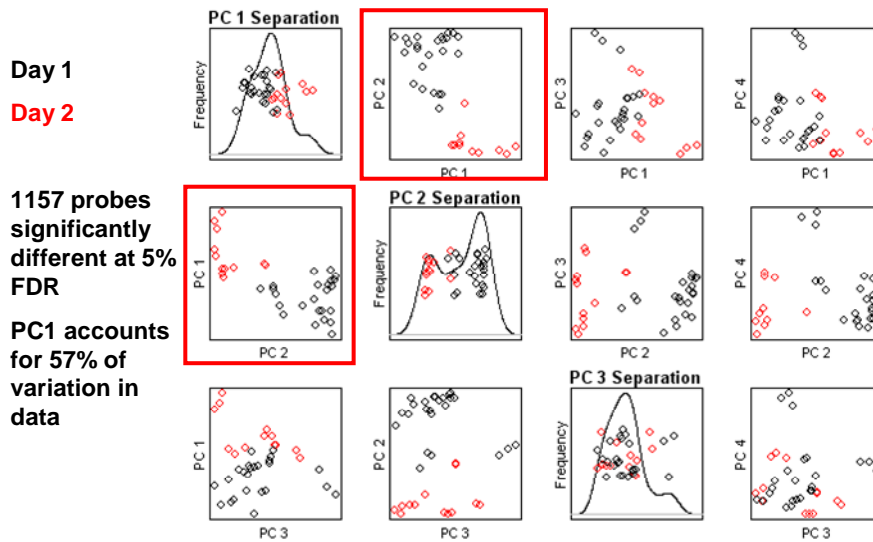


Figure 4.8:
The biggest source of variability for our hybridizations is the day of hybridization. PCA analysis by Joel Parker.

4.3.6 Correlation data can indicate hybridization quality

(Collaboration with Joel Parker)

RNAi genome screens are quite costly and time consuming; it would be beneficial to have quality control metrics to gauge how well the experiment is working. To this end, we can use Pearson correlation coefficients to measure hybridization variability. The correlation coefficient (r) measures the strength and the direction of a linear relationship between two variables.¹¹⁷ Positively correlated data sets will have coefficients approaching 1. Coefficients close to 0 indicate no correlation between sets (as between two random groups). A correlation greater than 0.8 is considered as strong, while correlation less than 0.5 is weak.¹¹⁸

We used Pearson correlation coefficients to examine the reproducibility of our hybridized samples. A small screen of 2300 cancer related hairpins was conducted in search of genes that modify sensitivity to a proteasome inhibitor, Velcade. This experiment was comprised of three biological replicates that were hybridized over two days. Replicates B and C were hybridized on the same day and replicate A was hybridized by itself on a different day. The hybridization had high correlations in reference channel per hybridization for both half hairpin and barcode probes (see figure 4.9). As hypothesized from Joel's PCA analysis, replicates B and C (hybridized on the same day) are more highly correlated than either is to replicate A. The same pattern was noticed when looking at biological replicates in the experimental channel. There were lower correlations than normal between biological replicates for both sense and barcode probes. A reason for this could have been that a dye swap occurred between replicates B, C and A, so that experimental samples for B and C were labeled in Cy3 and biological replicate samples for A were labeled in Cy 5.

These results illustrate the necessity of standardization for our barcode screens. This is especially important for the use of complex assays with multiple experimental groups and time points. In addition, the variability and high background of our hybridizations must be reduced to increase the significance of our results, especially if we want to conduct screens over a panel of different cell lines. To make these screens truly high throughput we must have a standardized method for performing hybridizations with efficient quality control.

Figure 4.9: Correlation coefficients can indicate hybridization quality.

a. Correlation coefficients between biological replicates in the reference channel b. Correlation coefficients between biological replicates in the experimental channel.

a. Correlation coefficients for barcode probes in reference channel				Correlation coefficients for sense probes in reference channel				b. Correlation coefficients between biological replicate barcode probes				Correlation coefficients between biological replicate sense probes							
replicate	B		C		replicate	B		C		replicate	B		C		replicate	B		C	
	T0	T0	T0	T0		T0	T0	T0	T0		T0	T0	T0	T0		T0	T0	T0	T0
C		0.842			C		0.939			C	T0	0.741			C	T0	0.907		
A		0.786		0.790	A		0.819		0.821	A	TO	0.441		0.548	A	TO	0.680		0.675
	T1 DMSO		T1 DMSO			T1 DMSO		T1 DMSO			T1 DMSO		T1 DMSO			T1 DMSO		T1 DMSO	
C	T1 DMSO	0.861			C	T1 DMSO	0.943			C	T1 DMSO	0.817			C	T1 DMSO	0.909		
A	T1 DMSO	0.818		0.855	A	T1 DMSO	0.871		0.880	A	T1 DMSO	0.332		0.476	A	T1 DMSO	0.504		0.544
	T2 DMSO		T2 DMSO			T2 DMSO		T2 DMSO			T2 DMSO		T2 DMSO			T2 DMSO		T2 DMSO	
C	T2 DMSO	0.876			C	T2 DMSO	0.940			C	T2 DMSO	0.543			C	T2 DMSO	0.734		
A	T2 DMSO	0.871		0.908	A	T2 DMSO	0.859		0.865	A	T2 DMSO	0.250		0.281	A	T2 DMSO	0.548		0.543
	T3 DMSO		T3 DMSO			T3 DMSO		T3 DMSO			T3 DMSO		T3 DMSO			T3 DMSO		T3 DMSO	
C	T3 DMSO	0.859			C	T3 DMSO	0.934			C	T3 DMSO	0.583			C	T3 DMSO	0.842		
A	T3 DMSO	0.868		0.895	A	T3 DMSO	0.875		0.848	A	T3 DMSO	0.534		0.632	A	T3 DMSO	0.764		0.848

4.3.7 Specialized data analysis for hairpins with variable knockdown

(Analysis by Joel Parker)

As RNAi screens become more commonplace, it is necessary to standardize how RNAi screening data are reported. This includes the extent to which the primary data are made available and establishing experimental and statistical criteria for defining a “hit”.¹¹⁵ Data analysis is important for establishing candidates. In particular, eliminating false positive hits can minimize the amount of effort needed to validate candidates. Conversely, false negatives reduce the number of relevant hits. Analysis of microarray data is a complex process, and not well standardized; many different methods exist for analyzing microarray data. Significance Analysis of Microarray (SAM) is a freeware software program that assigns scores to each gene on the basis of change in gene expression relative to the standard deviation of repeated measurements.¹¹⁹ Genes with scores greater than a selected threshold are deemed potentially significant. This method measures the strength of the relationship between gene expression and the standard deviation by performing a gene specific modified T-test. SAM determines false discovery

rate (FDR) by analyzing permutations of the measurements. A cutoff is chosen based on false discovery rate, generating a list whose significance depends on user input.

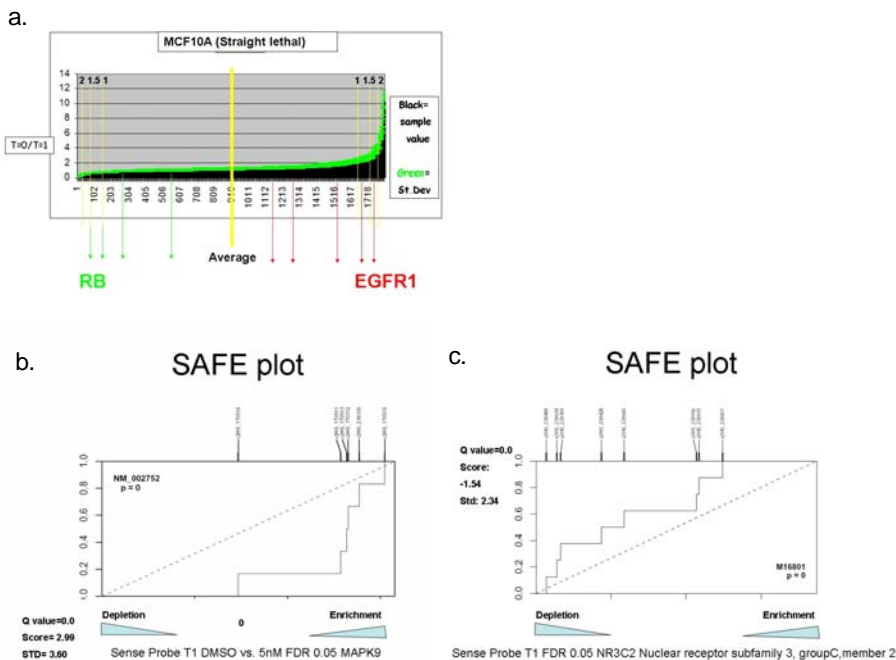
Although SAM is popular among scientists performing microarray experiments, statistical methods are needed for analyzing screening data that explicitly account for differences in RNAi knockdown efficiencies. For example, not all hairpins have the same potency and some may result in stronger phenotypes than others (see figure 4.10a). Hairpins which generate efficient knockdown will be represented by barcodes that are highly enriched or depleted on the array. These probes would be easier to identify than probes from hairpins with moderate knockdown. Such barcodes would show only a modest shift in representation, and thus may not be deemed statistically significant. Because of these differences in silencing efficiency, we have designed multiple hairpins for each gene in our library. This allows us to look at the performance of multiple hairpins for the same gene as an single set of array probes.

To this end, we are developing a procedure for analyzing genome wide RNAi screens with the help of Joel Parker based on the Significance Analysis of Function and Expression (SAFE) method from the Wright lab.¹²⁰ SAFE allows us to increase our ability to identify target genes corresponding to hairpins that have been enriched or depleted. Hairpins that show a modest shift might not be noticeable when examining a large group of probes. A single hairpin may not be statistically significant, unless it is taken together with independent estimates (different hairpins) of the same gene. To account for this, first a local statistic (simple T test) is computed and the p-values are applied to a larger global statistic (SAFE method) with the goal of detecting a shift in the local statistic to more extreme values.¹²⁰ The significance of the global statistic is

assessed by permuting the values and recomputing the statistic. This method results in a unitless “score” which reflects the amount of change that a probe experiences on the array. A high score indicates a greater amount of change. A positive score signifies enrichment and a negative score signifies depletion. A score near zero has no change. Standard deviation is used to measure how similar hairpin probes for the same gene behave on a microarray. Thus, if one hairpin results in a large amount of change on the array, and the another hairpin for the same gene does not produce a large change, this will result in a high standard deviation. A small standard deviation is reflected by multiple hairpins behaving in a consistent manner. See figure 4.10 for a graphical interpretation of the SAFE method called a SAFE plot.

Figure 4.10 Multiple hairpins for the same gene can be analyzed by SAFE

a. Example of how multiple hairpins for a gene can result in different amounts of probe change on an array due to differences in silencing efficiency (data from Jose Silva) b. A SAFE plot. The y axis represents percentage of the population. The x axis represent amount of depletion or enrichment. The diagonal line indicates complete randomness. The majority of hairpins for this genes are enriched and thus their representative line is to the right of the diagonal. c. A SAFE plot where the majority of hairpins are resulting in depletion. Both graphs in b and c have a high standard deviation because their hairpins have a high distribution over the plot.



4.4 Discussion

This chapter outlines the development of methodology for performing microarray hybridizations using barcode probes. This in-depth assessment of our methods was necessary for us to perform complex RNAi screens on a genome-wide scale. It is important to make these procedures standardized and reproducible if we are to expand the scope of our studies and compare screens across multiple cell lines. Ultimately, information from different screens can be used to form a picture of how cancer develops.

Hybridization of shorter labeled molecules containing the barcode sequence resulted in increased signal. This is likely due to a more thermodynamically favorable interaction between probe and target molecule. Furthermore, longer probes have a greater chance of cross hybridization because of extraneous sequence. Random-primed probes generated using Klenow fragment were better able to detect depleted probes than the other methods tested. This is probably in part because random-priming generates small probes regardless of input sequence length. These small probes are able to hybridize better to the arrays. Appreciable signal for all probes is important; in a negative selection screen, one has to detect depletion of probes of interest, a change which is harder to detect on a microarray. Many probes may be depleted enough to fall into a low intensity level that is considered background, and this can skew statistical analysis (See chapter 5).

In addition, we took efforts to determine sources of variability in our screens. This is of paramount importance if we are to conduct complex screens with multiple experimental groups. Variation can be minimized by standardizing the biological and

technical components of the hybridizations. Using the same batch of salmon sperm, primer, enzyme, etc can help reduce variation. Also, it may be interesting if we were to label our probes using sequence-specific primers instead of random nonamers with Klenow. This might help to reduce probe variation. It would help if all the hybridizations were done later in the evening, so that the ozone levels were kept constant and at a minimum. Finally, repeating experiments using dye swaps could also minimize variability. We also recognized the need for a quality control method for our hybridizations. We used correlation coefficients to try to measure reproducibility and efficiency, but other methods of measuring quality of hybridization could be explored, including intra-spot correlation of pixels, use of spike in controls, and inter-spot correlation.

Finally, we recognized that there are many different methods for microarray data analysis already in place. New method tailored for RNAi screens may be necessary, however. This is due to the unique nature of RNAi knockdown, and the fact that we have multiple hairpins which silence with different efficiencies. These differences may be evident on the microarray as modest shifts in fold change. We are modifying the permutation based testing method, SAFE, developed by Barry Wright. While no one in the lab has explored this further, it is important for obtaining the most information possible from RNAi screens. Certainly, our lab has enough data to develop this method further. The rate limiting step would be taking the time to validate the candidates that result from the SAFE method.

A great amount of information can be obtained from performing large scale RNAi screens. The Genetic Cancer Genome Project as envisioned by Steve Elledge and Greg

Hannon in 2005 is a systematic exploration of the genetic alterations that could kill cancer cells through large-scale functional RNAi screens to identify potential anti-cancer drug targets.¹²¹ This project would be come from a synthesis of RNAi technology, ROMA analysis, current cancer therapeutics, cancer mouse models, microarray technology, and bioinformatics. ROMA analysis could be used to detect genes that are amplified or deleted from human cancer samples.¹²² Genes that are deleted may be putative tumor suppressor genes. Hairpins can be created for these genes and used in an RNAi screen to examine their function. These genes are particularly interesting targets as they are obtained from clinical patient samples. They can help us understand why certain tumors respond to certain chemotherapies. Data obtained from cell culture screens could then be explored in *in-vivo* mouse models. The information from this project could be stored in a public database of lethality profiles and genetic interactions.

The technical experiments in this chapter, though tedious, were necessary for the development of RNAi screens. These screens have great potential for examining gene function. Further screens in our lab would not be able to proceed if the experiments in this chapter were not performed, as the intensity levels of our previous microarray experiments were so low, they were considered “noise.” Jose Silva and Ken Chang have gone on to validate a highly scalable approach for screening shRNA libraries by building upon these methods and integrating them into an Agilent platform.¹⁰⁶ Their work illustrates the potential of RNAi libraries to “permit rational searches for lesions that synergize with existing therapeutics to produce a path toward genetically informed combination therapies.”¹⁰⁶

Literature Cited

1. Chang, K., Elledge, S. J. & Hannon, G. J. Lessons from Nature: microRNA-based shRNA libraries. *Nat Methods* 3, 707-14 (2006).
2. Silva, J. M. et al. Profiling essential genes in human mammary cells by multiplex RNAi screening. *Science* 319, 617-20 (2008).
3. Paddison, P. J. et al. A resource for large-scale RNA-interference-based screens in mammals. *Nature* 428, 427-31 (2004).
4. Berns, K. et al. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* 428, 431-7 (2004).
5. Lucito, R. a. W., M. Preparation of Slides and Hybridization. (ed. Sambrook), e. D. B. a. J.) (Cold Spring Harbor Press, Cold Spring Harbor, NY., 2003).
6. website, A.
7. Ramdas, L. et al. Sources of nonlinearity in cDNA microarray expression measurements. *Genome Biol* 2, RESEARCH0047 (2001).
8. GE Healthcare-Amersham. Cy3-dUTP.
http://www1.gelifesciences.com/aptrix/upp00919.nsf/content/71985FA39EE4F72CC1256EB40044AD3A?OpenDocument&Path=Catalog&Hometitle=Catalog&entry=3&newrel&LinkParent=C1256FC4003AED40-2359E91E6518BB7EC1257019004907EE_RelatedLinksNew-C821BEC677D8448BC1256EAE002E3030&newrel&hidesearchbox=yes&moduleid=164367.
9. Ambion, I. in Amino Alkyl MessageAmp Instruction Manual.
10. Wall, M. E., Andreas Rechtsteiner, Rocha, LM. Singular value decomposition and principal component analysis". (Kluwer, Norwell, MA 2003).
11. Alter, O., Brown, P. O. & Botstein, D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A* 97, 10101-6 (2000).
12. Haney, S. A. Increasing the robustness and validity of RNAi screens. *Pharmacogenomics* 8, 1037-49 (2007).
13. Moore, D. The Basic Practice of Statistics (Paper) & Student CD. (W. H. Freeman, San Francisco, 2006).
14. Roberts, F. a. R., D. (2008).

15. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98, 5116-21 (2001).
16. Barry, W. T., Nobel, A. B. & Wright, F. A. Significance analysis of functional categories in gene expression studies: a structured permutation approach. *Bioinformatics* 21, 1943-9 (2005).
17. Elledge, S. J. & Hannon, G. J. An open letter to cancer researchers. *Science* 310, 439-41 (2005).
18. Lucito, R. et al. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res* 13, 2291-305 (2003).

Chapter 5

Synthetic lethality screens for the discovery of cancer
chemotherapeutic targets

Abstract:

I applied our RNAi barcode screening method (developed in chapter 4) in a screen using a complex mixture of 7500 library hairpins in a p53 isogenic HCT 116 colon cancer cells to identify genes that modify sensitivity to a common chemotherapeutic, doxorubicin and paclitaxel. Knockdown of Chk1, a kinase that mediates the G2/M checkpoint, increases sensitivity to doxorubicin in the p53 deficient HCT116 cells but not to the p53 wt HCT116 cell line. In addition, I identified a number of other candidates, including checkpoint related genes, that increase sensitivity to paclitaxel or doxorubicin for p53 null cells. This work illustrates the powerful use of RNAi screens to search for genes that synergize with existing therapeutics, and contributes to genetically informed combination therapies.

5.1 Introduction

The aim of the cancer drug discovery process is the identification of targets that will kill cancer cells but will leave normal cells unharmed. Most chemotherapeutics were discovered based on their ability to kill rapidly dividing cells *in vitro*.¹²³ When these agents are administered to patients, their side effects include also injuring rapidly dividing normal cells such as gastrointestinal mucosal epithelial cells.⁵² These drugs are often toxic to non-cancerous cells that are not dividing. This can result in damage to the heart, lung, etc. New research has focused on finding cancer specific molecular proteins or altered biological processes to help distinguish cancer cells from normal cells.

A way to achieve cancer-cell selectivity is to identify cellular targets that are present in cancer cells but not in normal cells. This has been exploited by the concept of oncogene addiction.¹²⁴ Oncogene addiction occurs when tumors become dependent on

the uninterrupted activity of a single mutated gene for tumor maintenance.¹²⁵ This idea has been exploited in the clinic by the development of Imatinib to treat BCR-ABL derived leukemias.¹²⁶ This concept has also been used as a basis for RNAi screens by Staudt and colleagues to identify genes that are essential for survival in two subtypes of Diffuse large B-cell lymphoma (DLBCL)⁶⁰ and most recently, to identify the dependence of multiple myeloma to the gene IRF4.¹²⁷ The sequencing of the human genome has enabled scientists to find genetic targets that specifically injure cancer cells, but spare normal cells. This has propelled the concept of synthetic lethality in the context of cancer therapy.¹²⁸

Synthetic lethality occurs when simultaneous mutation of two genes results in cell death, where mutation in either gene alone causes the cell or organism to remain viable¹²⁹ (see figure 5.1). A variation of this concept occurs when mutations in two genes result in impaired cellular fitness or “synthetic sick.”⁵² Synthetic lethality has been studied extensively in budding yeast, where approximately 20% of the genome is essential genes, but genetic screens suggest synthetic lethal interactions are common among the remaining 80%.⁵² Lee Hartwell proposed that synthetic lethal interactions could be used to develop anticancer drugs by therapeutically exploiting the contextual differences between cancer cells and normal cells.¹³⁰ The epigenetic and genetic alterations of cancer cells may increase their requirement for a molecular target. Inhibiting gene products that are synthetic lethal to cancer mutations should kill cancer cells and spare normal cells.

The primary goal of my thesis is to use RNAi screening technology to identify genes that would be novel chemotherapeutic targets. The previous chapters outlined the development of all the tools necessary to perform selection screens. We now have a

potent shRNA library and advanced microarray technology that permits negative selection screens. Using mammalian isogenic-paired cell lines, RNAi can be used to identify drug targets that when inhibited will result in the selective death of tumor cells.¹³¹

5.2 Experimental System

I have completed a screen using our shRNA library to identify genes that modify sensitivity to two standard chemotherapeutics in both a p53 wildtype and p53 null genetic background. The p53 protein inhibits tumor cell growth by evoking several cellular responses including cell-cycle arrest, senescence, apoptosis and activation of immune responses.^{132, 133} p53 can also contribute to the repair of genotoxic damage, potentially allowing for the release of damaged cells back into the proliferating pool.¹³⁴ Mutations in the p53 gene occur in approximately half of all human cancers.¹³⁵ In addition, perturbations of the p53 pathway are associated with more aggressive and therapeutically refractory tumors.¹³⁶ By performing this screen using two different p53 genetic backgrounds, we can look for genes that will cause the p53 null cells to die and yet leave the p53 wildtype cells untouched.

RNAi Synthetic-lethal analysis to enhance drug sensitivity

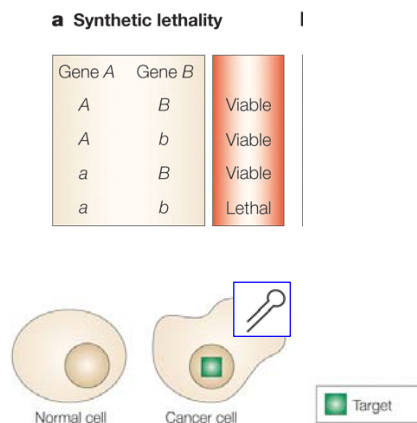


Figure 5.1: Synthetic lethality and Cancer
a. Synthetic lethal interactions from mutations in two genes b. Hairpin and gene target creating synthetic lethal scenario (modified from WG Kaelin 2005)

For my screen, I selected two cells lines that are genotypically identical except for their p53 status. We used HCT116 p53 wild type and HCT116 p53 null colon cancer cell lines.¹³⁷ These well studied cancer cell lines, from the lab of Bert Vogelstein, provide a genotypically and phenotypically well characterized model.¹³⁸ These cell lines were retrovirally infected with our shRNA library focusing on human kinases, phosphatases, cell cycle genes and a set of 1000 genes that been linked to cancer progression, the Harvard Cancer 1000. Kinases and phosphatases are attractive drug targets as their dysfunction can result in a variety of diseases including cancer.⁶ In recent years, several protein kinase inhibitors have performed successfully in clinical studies. They now account for 20–30% of the drug discovery programs of many companies.¹³⁹

Hairpins also contain a puromycin resistance marker allowing us to select cells that have been infected. These hairpin infected cells were separated into three groups. Two groups received specific drug treatment and the third group was treated with DMSO as a control. Each experimental group was treated with one of two different chemotherapeutics, doxorubicin and paclitaxel. Doxorubicin is a DNA damaging agent that works by inhibiting topoisomerase II¹⁴⁰ and is currently in over 900 clinical trials.¹⁴¹ In contrast, paclitaxel causes cell death by hyperstabilizing the cell's microtubules.¹⁴² Both drugs have been approved for use in a variety of cancers. Using drugs with two different mechanisms allows us to uncover broad mechanisms of drug sensitivity. In addition, we selected these two drugs as ways to generate cellular stress that would activate pathways where p53 would be necessary. Hairpin infected cells were treated at an ineffectual dose (IC20) where 80% of the cells survive. This low dose exposes genes that will increase the tumor's sensitivity to the drug by causing an increase in cell death.

Genes that synergize with a suboptimal drug treatment will allow us to make the drug more potent at a lower dose. In addition, since the chemotherapy could potentially be used at lower levels, some of its toxic side effects could be relieved.

Changes in cell growth were monitored through our cutting-edge DNA barcoding technology. Each hairpin is linked to a unique 60 nucleotide sequence, which serves as a barcode, and allows us to virtually count the number of cells that contain a specific hairpin in a cell population. Small changes in barcode copy number can be monitored through the use of microarray technology. The barcode was amplified from the cell's genomic DNA and made into a fluorescent probe that was hybridized to a microarray. In addition, we also used half hairpins as probes. These contain the unique 22mer sequence of the hairpin. Cells that have increased susceptibility to the chemotherapy are reflected as a loss of probe representation on the microarray as compared to untreated control populations. By performing this screen in two cell lines that differ in their p53 status, we can look for genes that will synergize with drug treatment in the p53 null cell line but will leave the p53 wild type cell line relatively unharmed. Depending on the type of analysis performed, this screen can answer different questions (see figure 5.2). If we compare each drug sample to its corresponding DMSO population, we can see what hairpins enhance sensitivity to the drug. Comparisons of DMSO samples between the p53 null and p53 wt cell line will reveal genes that are synthetic lethal with p53. If we compare the drug treated samples between the two cell lines, we can see what hairpins are synthetic lethal with p53 in the presence of a cellular stress represented by the drug.

Different Screen Analyses Answers Different Questions

- What is synthetic lethal with p53?
 - Under cellular stress with **taxol**?
 - Under cellular stress with **doxorubicin**?
- What genes modify sensitivity to chemotherapy?
 - What genes **enhance sensitivity to Taxol**?
 - What genes **enhance sensitivity to Doxorubicin**?

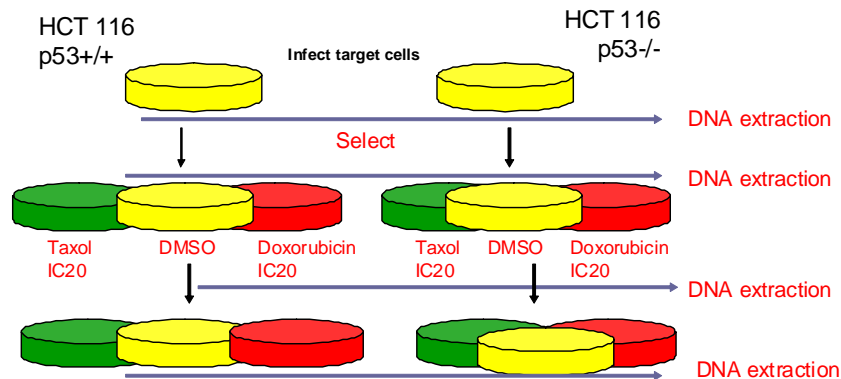


Figure 5.2: Different screen analyses answer different questions. Depending on what samples are analyzed, we can ask different questions.

5.3 Results

5.3.1 Preprocessing of data reveals variability within the screen (with Joel Parker)

The background for our microarrays is determined by calculating the mean intensity for probes that are printed on the chip but are not present in the biological experiment. (see chapter 4). This will help reduce any cross hybridization. The genomic DNA sample taken just after infection (T0) is used as a starting point to analyze what probes are present in the population before puromycin induced hairpin selection. Eighty percent of barcode probes reported above the mean background and 68% of sense probes reported above the mean background (data not shown). Probes less than 1.5x mean of the background probe intensities from each T0 sample are removed. If a probe is not present in two of three replicate T0 samples, then it is also removed. This resulted in 3,495 barcode probes and 3,726 barcode probes in the starting population. We decided to focus on barcode probes for all subsequent analysis, as their intensity levels were higher than half hairpin probes. Cy3 and Cy5 do not emit equal light per molecule at different

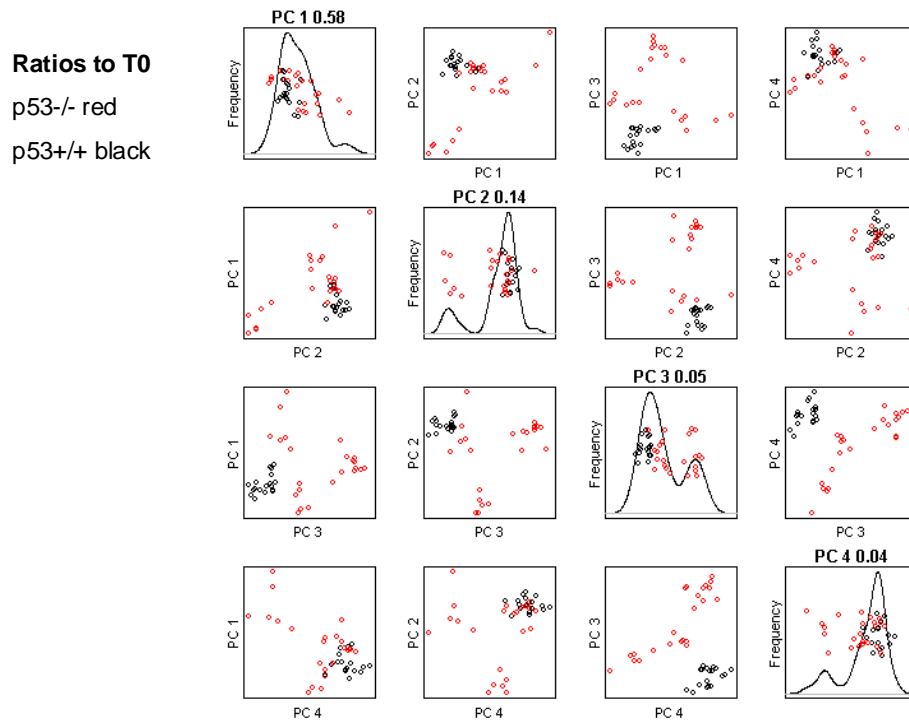
concentrations which can result in a dye bias.¹⁴³ Joel Parker, our collaborator, corrected for any dye intensity dependent effects by performing Loess normalization.¹⁴⁴ Joel also analyzed the Pearson correlation coefficients for the data (see figure 5.3). The correlation coefficients for the reference channel barcode probes are slightly higher than the experimental channel. In addition, while the correlation coefficients for our biological replicates at T0 is good, one confounding variable is our day of hybridization. Two biological replicates (A and B) were hybridized on one day and the third (replicate C) was hybridized on another day. This variable is causing replicate C to have lower correlation coefficients than the other replicates. PCA analysis also showed that p53 status was a component of variability indicating that changes on the microarray were also due to p53 status. This indicates that our screen is working.

Figure 5.3: Correlation coefficients for synthetic lethal screen reveal variabilities. Hybridizations performed on different days have the most variability when compared to the other replicate groups.

Correlation coefficients for Barcode probes in reference channel				Correlation coefficients for Barcode probes in experimental channel				
P53+/+	replicate	B	C	P53+/+	B	C		
replicate		T0	T0	replicate	T0	T0		
	C	T0	0.898	C	T0	0.832		
	A	T0	0.876	0.876	A	T0	0.764	0.775
		<i>P53-/-</i>	B	C		<i>P53-/-</i>	B	C
	C	T0	0.689	C	T0	0.573		
	A	T0	0.926	0.751	A	T0	0.843	0.636

Blue samples different hybridizations/PCR/labelling

Figure 5.4: PCA analysis shows p53 status as a component of variability. Principal component analysis shows that samples can be segregated by p53 status. Analysis by Joel Parker.



5.3.2 Different statistical analyses reveals different candidates.

(Analysis by Joel Parker)

Our collaboration with Joel Parker resulted in exploring different methods for data analysis. Initially, we used the SAM program to generate a candidate list (see Chapter 4). We used the method as described in Tusher et al.¹¹⁹ and used the “two class, unpaired test” with 100 permutations. We used a different statistical method that can calculate trends between multiple conditions over time called Extraction of Differential Gene Expression (EDGE). The Edge analysis first measures if a trend is occurring for each individual cell line under each condition (drug, DMSO, etc) over time. It then calculates if there is a trend between the p53 wt and p53 null cell line. For a detailed description of Edge, please see the following references.¹⁴⁵⁻¹⁴⁷ Joel used the two class test, 100 permutations and default parameters for fitting the cubic spline. Time points

were encoded as 0, 1, 2, 3 (not in actual hours). The Edge test results in a trend ratio, which reflects the amount of change between the two cell lines. A negative number signifies that a hairpin in the p53 null cell line is decreasing at a higher rate than the same hairpin in the p53 wild type cell line. However, scanning through a candidate list by trend ratio can be misleading as a high negative trend ratio can mean the wild type cell line is rapidly increasing while the p53 null cell line is flat and unchanging. A graphical interpretation of the Edge results is quite helpful in identifying candidates (see figure 5.5).

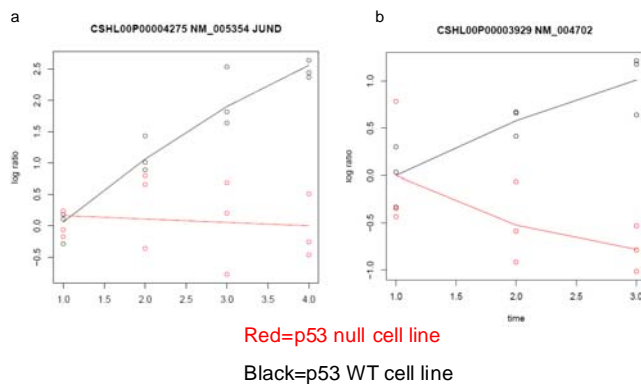


Figure 5.5: Two graphical interpretations of Edge results. The behavior of the probes for p53 WT cell line are in black and the p53 null cell probes are in red. The X axis represents time and the Y axis is the log ratio of probe intensity. a. An example of how a candidate with a high negative trend ratio can have a misleading probe behavior. The WT cell line probe is increasing while the null probe is remaining relatively unchanged. b. An example of the wt probe increasing while the p53 null probe decreases over time.

As probe intensities decrease over time, they can fall into background levels.

When the intensity of a probe is too low, it causes the ratio of (experimental probe/reference probe) to be a large number causing a problem. To correct for this, we added an intensity filter to the Edge results, where the background intensity (unused spots) is used as a guideline. Any probe intensity that is lower than background is set to a specific value. This is done independently for each sub-array as the background noise can be different for each. I used the SAM method and Edge results both with and without an intensity filter to generate candidate lists for DMSO (straight lethal), doxorubicin and taxol treated samples (see Table 5.1, 5.2 and 5.3.) For the SAM analysis, an FDR of 5%

was selected. For the EDGE results, a p or q value of 0.1 was chosen. However, it would be extremely difficult to validate hundreds of genes, so I sought to further narrow down my candidates by focusing on EDGE graphs. I decided to identify candidates by manually searching through EDGE graphs generated both with and without an intensity filter. My selection criteria were a p value of 0.1 and approximately two fold change or greater on the microarray for at least two of three biological replicates. If a candidate demonstrated a convincing change among the biological replicates but had a p value greater than 0.1, this was included on the list. The p value was relaxed to compensate for the variables such as day of hybridization. This resulted in a list of approximately 40 candidates for validation experiments (see Table 5.4).

Table 5.1 Probes depleted in the p53 null cell line but not the p53 wt cell line (DMSO only). SAM tests, EDGE tests with and without intensity filters.

<u>SAM results</u>		DMSO				p53 Null trend vs WT trend	
Description	Hairpin	p53 wt Fold Change	Q value	p53 KO Fold Change	Q-value	Q value	Q value
CDC20	v2HS_112884	0.605	0.002	1.077	0.153		0.001
PTPRR	v2HS_171028	0.513	0.001	1.098	0.073		0.001
JUND	v2HS_37282	0.179	0.001	1.052	0.210		0.001
LTBP2	v2HS_229759	0.549	0.001	1.007	0.093		0.001
DBF4	v2HS_85492	0.550	0.004	1.261	0.058		0.003
CCNH	v2HS_112784	0.503	0.003	1.038	0.098		0.005
STK39	v2HS_71810	0.461	0.001	1.120	0.182		0.006
HP	v2HS_49547	0.418	0.002	1.078	0.057		0.007
MPHOSPH6	v2HS_209736	0.757	0.013	1.273	0.097		0.008
RNF14	v2HS_27621	0.541	0.007	1.156	0.066		0.008
SPP1	v2HS_226915	0.641	0.022	1.121	0.104		0.009
PTPRK	v2HS_171002	0.591	0.002	1.182	0.185		0.009
NRAS	v2HS_152356	0.746	0.022	1.241	0.055		0.011
DDX11	v2HS_24008	0.778	0.023	1.156	0.073		0.016
CDC23	v2HS_41438	0.843	0.029	1.083	0.073		0.016
PRKCA	v2HS_170433	0.622	0.006	1.218	0.089		0.025
XM_293128	v2HS_123952	0.765	0.019	1.169	0.069		0.030
FKBPL	v2HS_209496	0.740	0.074	1.191	0.059		0.047
NRG1	v2HS_84773	0.665	0.044	1.117	0.068		0.049
FGF9	v2HS_113848	0.792	0.047	1.173	0.049		0.053
EPHB3	v2HS_17765	0.749	0.033	1.307	0.084		0.070
PSME3	v2HS_284	0.743	0.033	1.004	0.089		0.077
NFYC	v2HS_196293	0.844	0.079	1.116	0.055		0.078

CCNF	v2HS_150290	0.817	0.099	1.260	0.073	0.096
CHEK1	v2HS_112994	0.875	0.083	1.248	0.065	0.097

EDGE results with Intensity Filter

Description	Accession	Hairpin	Trend ratio	p-value
JUND	NM_005354	v2HS_37282	-2.540	0.000
CCNG2	NM_004354	v2HS_14930	-2.032	0.079
FARSLA	NM_004461	v2HS_17353	-1.479	0.000
AKT1	NM_005163	v2HS_188525	-1.335	0.038
CDC14B	NM_033331	v2HS_203518	-1.311	0.000
STK39	NM_013233	v2HS_71810	-1.297	0.002
ZAK	NM_133646	v2HS_225895	-1.264	0.000
PSMC1	NM_002802	v2HS_5	-1.235	0.048
	AF092810	v2HS_71	-1.216	0.058
FGF11	NM_004112	v2HS_218728	-1.210	0.000
DBF4	NM_006716	v2HS_85492	-1.146	0.000
RNF14	NM_004290	v2HS_27621	-1.095	0.001
SH3BP4	NM_014521	v2HS_257067	-1.080	0.008
PTPRR	NM_002849	v2HS_171028	-1.072	0.000
EPHB3	NM_004443	v2HS_17765	-0.990	0.014
BTC	NM_001729	v2HS_150112	-0.977	0.000
PTPRK	NM_002844	v2HS_171002	-0.935	0.004
CDH1	NM_004360	v2HS_243170	-0.795	0.344
SPP1	NM_000582	v2HS_226915	-0.789	0.005
CDC20	NM_001255	v2HS_112884	-0.766	0.000
RNF14	NM_183400	v2HS_27619	-0.759	0.749
MPHOSPH6	NM_005792	v2HS_209736	-0.743	0.001
Abcb1a	NM_133401	v2HS_50	-0.718	0.590
NRAS	NM_002524	v2HS_152356	-0.694	0.003
CHEK1	NM_001274	v2HS_112994	-0.663	0.005
IGFBP1	NM_000596	v2HS_111616	-0.615	0.343
DDX11	NM_004399	v2HS_24008	-0.598	0.003
RUVBL1	NM_003707	v2HS_14740	-0.587	0.037
CKS1B	NM_001826	v2HS_150607	-0.549	0.393
FGF9	NM_002010	v2HS_113848	-0.504	0.007
CDC23	NM_004661	v2HS_41438	-0.390	0.002
CSNK2A2	NM_001896	v2HS_150997	-0.385	0.620
HRG	NM_000412	v2HS_229439	-0.226	0.074
ORC2L	NM_006190	v2HS_63513	0.815	0.000
PTK2	NM_005607	v2HS_57330	0.820	0.000
ERBB2	NM_004448	v2HS_17671	0.880	0.000
E2F6	NM_001952	v2HS_151297	1.094	0.000
S100B	NM_006272	v2HS_56975	1.149	0.000
MAPK10	NM_138982	v2HS_203028	1.201	0.000
CD34	NM_001773	v2HS_150351	1.262	0.000
ABCG1	NM_004915	v2HS_62693	1.285	0.000
ELF5	NM_001422	v2HS_225958	1.541	0.000

EDGE results

Description	Accession	Hairpin	Trend	DMSO q-value
--------------------	------------------	----------------	--------------	---------------------

			ratio	
JUND	NM_005354	v2HS_37282	-2.554	0.009
F8	NM_000132	v2HS_67564	-1.837	0.026
CCNE2	NM_057735	v2HS_37176	-1.704	0.012
RRM2	NM_001034	v2HS_93887	-1.690	0.018
PRKAA2	NM_006252	v2HS_57674	-1.641	0.010
SPOCK1	NM_004598	v2HS_47114	-1.618	0.037
FARSLA	NM_004461	v2HS_17353	-1.540	0.004
RPS4L	NG_000967	v2HS_21735	-1.527	0.042
HP	NM_005143	v2HS_49547	-1.366	0.010
PSMC1	NM_002802	v2HS_5	-1.335	0.084
ITGA6	NM_000210	v2HS_77132	-1.301	0.010
CDC14B	NM_033331	v2HS_203518	-1.287	0.004
ZAK	NM_133646	v2HS_225895	-1.287	0.005
STK39	NM_013233	v2HS_71810	-1.281	0.050
FGF11	NM_004112	v2HS_218728	-1.225	0.008
DBF4	NM_006716	v2HS_85492	-1.198	0.012
AKT3	NM_181690	v2HS_200305	-1.178	0.033
AKT3	NM_181690	v2HS_200305	-1.121	0.040
PTPPR	NM_002849	v2HS_171028	-1.098	0.004
RNF14	NM_004290	v2HS_27621	-1.095	0.038
SH3BP4	NM_014521	v2HS_257067	-1.075	0.080
PDGFB	NM_002608	v2HS_169797	-1.070	0.096
CDH1	NM_004360	v2HS_243170	-1.016	0.358
PTPRK	NM_002844	v2HS_171002	-0.999	0.073
PRKCA	NM_002737	v2HS_170433	-0.969	0.092
CDC20	NM_001255	v2HS_112884	-0.833	0.010
RNF14	NM_183400	v2HS_27619	-0.822	0.679
SPP1	NM_000582	v2HS_226915	-0.806	0.069
EPHB3	NM_004443	v2HS_17765	-0.804	0.079
MPHOSPH6	NM_005792	v2HS_209736	-0.751	0.059
NRAS	NM_002524	v2HS_152356	-0.735	0.033
CHEK1	NM_001274	v2HS_112994	-0.654	0.062
PRKCB1	NM_002738	v2HS_170441	-0.649	0.848
FGF10	NM_004465	v2HS_17325	-0.599	0.343
DDX11	NM_004399	v2HS_24008	-0.571	0.073
FDF9	NM_002010	v2HS_113848	-0.565	0.050
IGFBP1	NM_000596	v2HS_111616	-0.559	0.623
RUVBL1	NM_003707	v2HS_14740	-0.547	0.274
CKS1B	NM_001826	v2HS_150607	-0.501	0.612
Abcb1a	NM_133401	v2HS_50	-0.494	0.822
CDC23	NM_004661	v2HS_41438	-0.362	0.095

Table 5.2 Probes that were depleted in HCT116 p53 null cells after treatment with Doxorubicin.

SAM Test for Doxorubicin

Description	Accession	Hairpin	p53 WT vs KO fold change	FDR
ERF	NM_006494	v2HS_197559	2.127	0.000
CDKN1C	NM_000076	v2HS_88831	2.123	0.000
TP53BP1	NM_005657	v2HS_56193	2.245	0.000
TNFRSF7	NM_001242	v2HS_112804	2.015	0.000
CTS0	NM_001334	v2HS_225877	2.361	0.000

NEUROD4	NM_021191	v2HS_65748	2.175	0.000
MDM2	NM_002392	v2HS_254556	2.530	0.000
JUND	NM_005354	v2HS_37282	2.214	0.000
F8	NM_000132	v2HS_67564	2.341	0.000
NR0B2	NM_021969	v2HS_72557	2.863	0.000
APP	NM_000484	v2HS_231737	2.280	0.000
JAK1	NM_002227	v2HS_133562	2.138	0.000
BMPR2	NM_001204	v2HS_226281	2.379	0.000
ACVR1	NM_001105	v2HS_94266	2.437	0.000
RYK	NM_002958	v2HS_31984	2.136	0.000
WEE1	NM_003390	v2HS_172009	2.209	0.000
#VALUE!	XM_030845	v2HS_59892	2.160	0.000
GMNN	NM_015895	v2HS_114947	2.247	0.837
FOSL2	NM_005253	v2HS_42865	2.570	0.837
AXL	NM_021913	v2HS_201787	2.107	0.837
STK39	NM_013233	v2HS_71810	2.085	0.837
SNRK	NM_017719	v2HS_173445	2.177	0.837
PRKG1	NM_006258	v2HS_57569	2.013	0.837
PSMD13	NM_002817	v2HS_170854	2.236	1.474
SH3BP4	NM_014521	v2HS_257067	2.065	1.474
MAPRE1	NM_012325	v2HS_50453	2.833	1.474
RAD51	NM_002875	v2HS_171185	2.170	1.474
SYK	NM_003177	v2HS_153705	2.468	1.474
PPAP2B	NM_003713	v2HS_24417	2.752	1.474
GTF2E2	NM_002095	v2HS_114306	2.241	1.474
DLL1	NM_005618	v2HS_56930	2.014	1.474
PSTPIP2	NM_024430	v2HS_136018	2.061	1.474
COL11A1	NM_080630	v2HS_71589	2.258	1.474
ALS2CR2	NM_018571	v2HS_176129	2.326	1.474
MCM5	NM_006739	v2HS_84918	2.558	1.891
GSK3B	NM_002093	v2HS_114293	2.181	1.891
BCAS2	NM_005872	v2HS_250914	2.074	1.891
XPA	NM_000380	v2HS_92696	2.015	1.891
CTSZ	NM_001336	v2HS_226085	2.038	1.891
DCC	NM_005215	v2HS_43698	2.235	1.891
SLC39A6	NM_012319	v2HS_50770	2.700	1.891
FGF7	NM_002009	v2HS_224602	2.039	1.891
DNM1	NM_004408	v2HS_18786	2.292	1.891
ODZ1	NM_014253	v2HS_197083	2.496	1.891
PRKCI	NM_002740	v2HS_170453	2.385	1.891
COL9A3	NM_001853	v2HS_150762	2.156	1.891
JAG1	NM_000214	v2HS_77060	2.340	1.891
FGFR2	NM_000141	v2HS_83540	2.044	1.891
FHIT	NM_002012	v2HS_113859	2.214	1.891
FGF19	NM_005117	v2HS_50187	2.119	1.891
MCM7	NM_005916	v2HS_252457	2.146	1.891
SYK	NM_003177	v2HS_153705	2.493	1.891
DYRK1B	NM_006483	v2HS_194185	2.286	1.891
PRKCI	NM_002740	v2HS_170453	2.146	1.891
EIF2AK4	AB037759	v2HS_202539	2.194	1.891

ADRBK2	NM_005160	v2HS_49036	2.356	1.891
PHKG1	NM_006213	v2HS_58678	2.332	1.891
CHKA	NM_001277	v2HS_113012	2.838	1.891
CDC2L6	NM_015076	v2HS_203055	2.189	1.891
MAPKAPK2	NM_004759	v2HS_36045	2.695	1.891
APOBEC3G	NM_145298	v2HS_18329	2.384	3.339
PRKACG	NM_002732	v2HS_170412	2.163	3.339
FGF18	NM_003862	v2HS_16532	2.464	3.339
CREBBP	NM_004380	v2HS_24251	2.297	3.339
IL20	NM_018724	v2HS_26365	2.055	3.339
RAB35	NM_006861	v2HS_77772	2.064	3.339
MAP4K5	NM_006575	v2HS_2161	2.180	3.339
FLJ25006	NM_144610	v2HS_19098	2.169	3.339
IKBKB	AF080158	v2HS_202180	2.057	3.339
CSNK1A1	NM_001892	v2HS_150971	2.051	3.339
MAP3K13	NM_004721	v2HS_36732	2.146	3.339
KIAA0391	NM_014672	v2HS_79170	2.172	3.464
NR0B1	NM_000475	v2HS_93206	3.227	3.464
SLK	NM_014720	v2HS_73518	2.062	3.464
CCNH	NM_001239	v2HS_112787	2.042	4.802
PRPS1	NM_002764	v2HS_170580	2.216	4.802
ADCK4	BC013114	v2HS_203175	2.058	4.802

EDGE Results

Description	Accession	Hairpin	Trend Ratio	q-value
DGKE	NM_003647	v2HS_27429	-1.320	0.001
SNF1LK2	XM_041314	NM_015191.1	-1.237	0.001
AIF1	NM_001623	v2HS_132494	-1.053	0.002
GPR34	NM_005300	v2HS_38396	-1.037	0.003
FH	NM_000143	v2HS_83494	-1.127	0.006
NOL1	NM_006170	v2HS_248935	-1.204	0.009
RB1	L41870	v2HS_130606	-1.550	0.012
AKT3	NM_181690	v2HS_200305	-1.050	0.016
DGKE	NM_003647	v2HS_27428	-1.123	0.024
NEK3	NM_152720.1	v2HS_83314	-1.543	0.027
AKT3	NM_181690	v2HS_200305	-1.159	0.027
CCNA2	NM_001237	v2HS_112776	-1.224	0.033
NEK3	NM_152720.1	v2HS_83314	-1.392	0.033
TRRAP	NM_003496	v2HS_172540	-0.847	0.034
DNMT1	NM_001379	v2HS_113506	-1.563	0.035
TGFBR3	NM_003243	v2HS_154004	-1.326	0.040
TGFBR3	NM_003243	v2HS_154003	-0.907	0.042
PDGFB	NM_002608	v2HS_169797	-1.351	0.055
STK39	NM_013233	v2HS_71810	-1.177	0.079
EGFR	NM_201283	v2HS_200678	-0.790	0.339
EGFR	NM_201283	v2HS_200678	-0.747	0.395

EDGE results with Intensity Filter

Description	Accession	Hairpin	Trend Ratio	p-value
TRRAP	NM_003496	v2HS_172540	-1.018	0.002

XRCC3	NM_005432	v2HS_69648	-2.025	0.106
MDR1	AF092810	v2HS_71	-1.441	0.006
ACMSD	NM_138326	v2HS_37924	-1.655	0.006
GSK3B	XM_291773	v2HS_104795	-1.927	0.000
CDC14B	NM_033331	v2HS_203518	-1.825	0.001
PDGFB	NM_002608	v2HS_169797	-1.097	0.005
EDD1	NM_015902	v2HS_201754	-1.590	0.010
TBRG4	NM_004749	v2HS_36146	-1.676	0.000
CCNA2	NM_001237	v2HS_112776	-1.692	0.002
SMG1	NM_014006	v2HS_60493	-1.491	0.001
MAP4K3	NM_003618	v2HS_27892	-1.684	0.000
EPHA6	XM_093359	v2HS_88921	-1.620	0.011
DCAMKL3	AB051552	v2HS_201790	-1.366	0.092
DGKE	NM_003647	v2HS_27429	-1.414	0.004
UCK2	NM_012474	v2HS_42846	-1.712	0.010
BCL2	M13995	v2HS_130617	-0.586	0.279
SPIN	NM_006717	v2HS_85454	-1.374	0.000
CHES1	NM_005197	v2HS_176178	-1.479	0.000
CAMKK1	NM_032294	v2HS_138358	-1.361	0.000
C16orf5	NM_013399	v2HS_63848	-1.488	0.001
CKAP4	NM_006825	v2HS_78443	-1.568	0.002
JUND	NM_005354	v2HS_37282	-1.429	0.002
PHKA1	NM_002637	v2HS_169956	-1.347	0.002
CDCA5	NM_080668	v2HS_70776	-1.176	0.003
CCNE2	NM_004702	v2HS_37177	-1.792	0.003
RAD51L1	NM_133509	v2HS_239000	-1.076	0.003
TYRO3	NM_006293	v2HS_56438	-1.845	0.003
FGFR2	NM_000141	v2HS_83539	-0.775	0.003
RAD51L1	NM_133509	v2HS_239000	-1.203	0.003
YWHAE	NM_006761	v2HS_84593	-0.862	0.003
NEUROD2	NM_006160	v2HS_64390	-1.547	0.004
TAZ	NM_000116	v2HS_83935	-0.975	0.004
PCBP4	NM_020418	v2HS_34043	-0.957	0.004
RPS6KC1	NM_012424	v2HS_43908	-1.584	0.005
MAPK10	NM_138982	v2HS_201743	-1.676	0.005
FAM54B	NM_019557	v2HS_52812	-0.723	0.007
ALDH18A1	NM_002860	v2HS_171097	-1.228	0.008
PRKCA	NM_002737	v2HS_170438	-1.809	0.008
PTPRK	NM_002844	v2HS_171007	-1.334	0.008
WDR3	NM_006784	v2HS_84274	-1.380	0.009
JAK2	NM_004972	v2HS_61649	-1.103	0.009
MARK1	NM_018650	v2HS_30754	-1.362	0.010
PRKAA2	NM_006252	v2HS_57674	-1.547	0.010
ALS2CR2	NM_018571	v2HS_176124	-1.169	0.010
ID1	NM_002165	v2HS_235359	-1.694	0.010
NBN	NM_002485	v2HS_152144	-1.948	0.014
TTN	NM_003319	v2HS_171632	-1.441	0.029
PTK9	NM_002822	v2HS_170878	-1.293	0.065
NEK3	Z29067	v2HS_202287	-1.923	0.073
AKT1	NM_005163	v2HS_188525	-1.325	0.081

MYADM	NM_002739	v2HS_170448	-1.810	0.118
KRIT1	NM_004912	v2HS_62788	-1.171	0.131
RAC1	NM_006908	v2HS_232790	-1.935	0.157

Table 5.3 Probes depleted in HCT116 p53 null cells with paclitaxel treatment.

SAM Test for paclitaxel

Description	Accession	Hairpin	p53 WT vs KO fold change	q value
MCM6	NM_005915	v2HS_197429	1.898	0.000
MAPRE1	NM_012325	v2HS_50453	2.788	0.000
FAS	NM_000043	v2HS_89562	2.097	0.000
TNFRSF7	NM_001242	v2HS_112804	1.825	0.000
CTSO	NM_001334	v2HS_225877	1.963	0.000
IBSP	NM_004967	v2HS_61769	1.710	0.000
CRSP3	NM_004830	v2HS_68245	1.839	0.000
MDM2	NM_002392	v2HS_254556	2.191	0.000
NR2F1	NM_005654	v2HS_56174	2.085	0.000
RAD17	NM_133339	v2HS_203352	2.003	0.000
FGF19	NM_005117	v2HS_50187	2.006	0.000
ARHGDI A	NM_004309	v2HS_15549	2.078	0.000
APP	NM_000484	v2HS_231737	1.867	0.000
COL11A1	NM_080630	v2HS_71589	2.663	0.000
MAPK14	NM_001315	v2HS_113216	1.971	0.000
DYRK1B	NM_006483	v2HS_194185	2.241	0.000
STK32A	NM_145001	v2HS_30160	1.935	0.000
PFDN1	NM_002622	v2HS_169872	0.554	2.687
TNFRSF10B	NM_003842	v2HS_16709	0.479	2.687
NOL1	NM_006170	v2HS_248935	1.643	2.943
GJA5	NM_005266	v2HS_42547	1.715	2.943
FGF7	NM_002009	v2HS_224602	2.394	2.943
SYK	NM_003177	v2HS_153705	2.211	2.943
APOD	NM_001647	v2HS_132619	1.777	2.943
PPAP2B	NM_003713	v2HS_24417	2.104	2.943
FOSL2	NM_005253	v2HS_42865	2.127	2.943
JUND	NM_005354	v2HS_37282	2.219	2.943
FHIT	NM_002012	v2HS_225339	1.651	2.943
NR5A2	NM_003822	v2HS_17033	1.835	2.943
BCKDK	NM_005881	v2HS_198098	2.204	2.943
PRKCI	NM_002740	v2HS_170453	2.306	2.943
ALS2CR7	NM_139158	v2HS_64495	1.651	2.943
BMPR2	NM_001204	v2HS_226281	2.012	2.943
MAPKAPK2	NM_032960	v2HS_202459	1.653	2.943
ALPK2	NM_052947	v2HS_223607	1.996	2.943
WEE1	NM_003390	v2HS_172009	1.785	2.943

EDGE Results

Description	Accession	Hairpin	Trend Ratio	q-value
LYN	NM_002350	v2HS_134141	-1.440	0.010
RBBP6	NM_032626	v2HS_159574	-1.124	0.010
SESN2	NM_031459	v2HS_117406	-1.206	0.010

MUC7	NM_006685	v2HS_90306	-1.671	0.028
PSMD5	NM_005047	v2HS_19360	-1.562	0.028
CRKRS	AF227198	v2HS_111182	-1.692	0.033
DGKE	NM_003647	v2HS_27429	-1.971	0.036
TRPM6	NM_017662	v2HS_155076	-1.501	0.038
CDKN1C	NM_000076	v2HS_88831	-1.523	0.042
AKT3	NM_181690	v2HS_200305	-1.287	0.042
PSMD5	NM_005047	v2HS_19361	-1.786	0.042
PSMD3	NM_002809	v2HS_170804	-1.496	0.042
DGKE	NM_003647	v2HS_27429	-1.762	0.042
DNMT1	NM_001379	v2HS_113506	-1.576	0.042
ACVR2A	NM_001616	v2HS_235462	-1.637	0.042
DGKE	NM_003647	v2HS_27428	-1.281	0.042
AKT3	NM_181690	v2HS_200305	-1.431	0.046
LYN	NM_002350	v2HS_134141	-0.944	0.046
CCNA2	NM_001237	v2HS_112776	-1.708	0.048
NEK3	NM_152720.1	v2HS_83314	-1.528	0.048
BRDT	NM_001726	v2HS_150092	-1.624	0.048
ANAPC4	NM_013367	v2HS_254660	-1.007	0.061
SMG1	NM_014006	v2HS_60493	-1.475	0.071
STK39	NM_013233	v2HS_71810	-1.631	0.122
MDR1	AF092810	v2HS_71	-1.688	0.151
NR3C2	NM_000901	v2HS_131200	-1.509	0.193
XRCC3	NM_005432	v2HS_69647	-1.516	0.204
EGFR	NM_201283	v2HS_200678	-1.221	0.220
EGFR	NM_201283	v2HS_200678	-1.220	0.260
XRCC3	NM_005432	v2HS_69648	-2.240	0.288
EPHA6	XM_093359	v2HS_88921	-1.155	0.294
KALRN	NM_007064	v2HS_199006	-0.441	0.369
MAPRE1	NM_012325	v2HS_50453	-1.745	0.443
MARK3	NM_002376	v2HS_151569	-0.833	0.465
LASPL1	NM_006148	v2HS_64684	-1.593	0.490
RNF14	NM_183400	v2HS_27619	-0.760	0.585

EDGE results with Intensity Filter

Description	Accession	Hairpin	Trend Ratio	p-value
LYN	NM_002350	v2HS_134141	-1.402	0.000
MSH2	NM_000251	v2HS_76354	-1.451	0.000
ZW10	NM_004724	v2HS_36687	-1.370	0.000
MUC7	NM_006685	v2HS_90306	-1.530	0.000
DGKE	NM_003647	v2HS_27429	-1.801	0.000
AKT3	NM_181690	v2HS_200305	-1.421	0.001
PSMD5	NM_005047	v2HS_19361	-1.669	0.001
CKAP4	NM_006825	v2HS_78443	-1.656	0.001
CDKN1C	NM_000076	v2HS_88831	-1.399	0.001
PSMD3	NM_002809	v2HS_170804	-1.379	0.001
MARK1	NM_018650	v2HS_30754	-1.365	0.001
DGKE	NM_003647	v2HS_27429	-1.720	0.002
PSMD3	NM_002809	v2HS_170807	-1.856	0.002
CCNE2	NM_004702	v2HS_37177	-1.724	0.002

SHFM1P	NG_001311	v2HS_33357	-1.085	0.002
GSK3A	NM_019884	v2HS_35842	-1.361	0.003
GSK3B	XM_291773	v2HS_104795	-1.859	0.003
PSME1	NM_006263	v2HS_57260	-1.705	0.003
ANAPC4	NM_013367	v2HS_254660	-0.935	0.003
F8	NM_000132	v2HS_67564	-1.653	0.003
CDK8	NM_001260	v2HS_112909	-1.679	0.003
PLCB1	NM_015192	v2HS_86589	-1.418	0.004
MAPK12	NM_002969	v2HS_31747	-1.458	0.004
CCNA2	NM_001237	v2HS_112776	-1.482	0.006
WDR3	NM_006784	v2HS_84274	-1.704	0.006
RB1	L41870	v2HS_130609	-1.375	0.006
AOF2	NM_015013	v2HS_34926	-0.765	0.007
SMG1	NM_014006	v2HS_60493	-1.469	0.008
CHES1	NM_005197	v2HS_176178	-1.514	0.008
TAZ	NM_000116	v2HS_83935	-0.624	0.008
TGFBR1	BC071181	v2HS_55961	-0.879	0.009
CCNA2	NM_001237	v2HS_112776	-1.690	0.009
GRB2	NM_002086	v2HS_114253	-1.325	0.009
PDGFB	NM_002608	v2HS_169797	-1.315	0.009
STK39	NM_013233	v2HS_71810	-1.578	0.010
CDC14B	NM_033331	v2HS_203518	-1.565	0.010
ALS2CR2	NM_018571	v2HS_176124	-1.995	0.016
SMG1	NM_015092	v2HS_201817	-0.867	0.036
PTK9	NM_002822	v2HS_170878	-0.781	0.043
MAPKAPK5	NM_139078	v2HS_201871	-0.970	0.043
MDR1	AF092810	v2HS_71	-1.588	0.052
CCNA2	NM_001237	v2HS_112774	-1.102	0.067
XRCC3	NM_005432	v2HS_69648	-2.236	0.078
MYADM	NM_002739	v2HS_170448	-1.539	0.078
AKT1	NM_005163	v2HS_188525	-1.391	0.098
PCAF	NM_003884	v2HS_53091	-1.408	0.101
RAC1	NM_006908	v2HS_232790	-2.238	0.110
NBN	NM_002485	v2HS_152144	-1.383	0.120
EPHA6	XM_093359	v2HS_88921	-1.271	0.125
DNM1	NM_004408	v2HS_242475	-0.934	0.141
MARK3	NM_002376	v2HS_151569	-0.864	0.162
MAPRE1	NM_012325	v2HS_50453	-1.941	0.175
STAT3	NM_139276	v2HS_203085	-0.967	0.241
NEK3	Z29067	v2HS_202287	-1.594	0.246
EGFR	NM_201283	v2HS_200678	-1.248	0.247
AKT1	NM_005163	v2HS_188525	-0.899	0.255
ALDH18A1	NM_002860	v2HS_171097	-0.957	0.276

5.3.3 Candidates list includes DNA damage response genes, novel genes and checkpoint related genes (Analysis by Joel Parker)

This screen has resulted in a list of candidates that create a unique signature for both drugs. However, a few candidates however overlap for both drugs as well. The

genes in the list include genes we would expect, such as DNA damage response genes (MDR1 and XRCC3) and novel genes of unknown function such as PPP2R3B and MPHOSPH9. MDR1 encodes the P-glycoprotein whose shRNA mediated silencing has been shown to increase sensitivity toward the cytotoxic drugs vincristine, paclitaxel, and doxorubicin in human colon carcinoma HCT-8 and HEK293 cell lines.¹⁴⁸ A few checkpoint related genes such as Nibrin, GSK3 β and DBF4, also scored in the analysis. Genes that scored twice (Nibrin and Chk1) because of multiple hairpins were also given priority. Ingenuity Pathway Analysis is a software program that enables scientists to input a data set of genes or proteins and identify relevant pathways or functions. Unfortunately, when the gene list is put into the Ingenuity Pathway Analysis program, no particular pathway is illuminated. A hairpin that scored in the screen of particular significance is the checkpoint control related gene, Chk1. Chk1 scored using both the SAM and EDGE test for straight lethal. We chose to focus our validation methods on hairpins that created a synthetic lethal situation in the presence of drug for the p53 null cell line. These candidates would be attractive as they would synergize with chemotherapeutics currently being used in the clinic for cells that are defective in p53.

Table 5.4 Candidate gene list resulting from screen for validation experiments. These candidates were selected by examining Edge trend analysis graphs for probes that decreased in the p53 null cell line over time while the p53 wt cell line remained static or increasing. Selection criteria included a p value of approximately 0.1 and a two fold change on the microarray for at least two of three biological replicates. Trend ratios and p values are not listed as these are different for Edge tests depending on the inclusion of an intensity filter. In addition, trend ratios are not reflective of probe behavior on a graph. (see Section 5.3.2).

<u>Paclitaxel Candidates</u>	<u>Hairpin Number</u>
ALDH18A1 ALDH18A1 Aldehyde dehydrogenase 18 family, member A1	v2HS_171097
CCNE2 Cyclin E2	v2HS_37177
DNM1 DNM1 Dynamin 1	v2HS_242475
EGFR EGFR Epidermal growth factor receptor	v2HS_200678
EPHA6 EPH receptor A6	v2HS_88921
KALRN KALRN Kalirin, RhoGEF kinase	v2HS_199006
MAPKAPK5 Mitogen-activated protein kinase-activated protein kinase 5	v2HS_201871

MAPRE1	MAPRE1 Microtubule-associated protein, RP/EB family, member 1	v2HS_50453
MARK3	MARK3 MAP/microtubule affinity-regulating kinase 3	v2HS_151569
MDM4	Mdm4 Transformed mouse 3T3 cell double minute 4	v2HS_539
MYADM	MYADM Myeloid-associated differentiation marker	v2HS_170448
NBN	NBN Nibrin	v2HS_152144
NBN	NBN Nibrin	v2HS_152148
NEK3	NEK3 NIMA (never in mitosis gene a)-related kinase 3	v2HS_202287
PTK9	PTK9 PTK9 protein tyrosine kinase 9	v2HS_170878
PTPRE	PTPRE Protein tyrosine phosphatase, receptor type, E	v2HS_12500
PTPRJ	PTPRJ Protein tyrosine phosphatase, receptor type, J	v2HS_171000
RAC1	Ras-related C3 botulinum toxin substrate 1	v2HS_232790
RNF14	Ring finger protein 14	v2HS_27619
SMG1	PI-3-kinase-related kinase SMG-1	v2HS_201817
SRMS	Src-related kinase lacking C-terminal regulatory tyrosine sites	v2HS_58606
STAT3	Signal transducer and activator of transcription 3	v2HS_203085
TAZ	Tafazzin (cardiomyopathy, dilated 3A (X-linked)	v2HS_83935
TPM2	Tropomyosin 2 (beta)	v2HS_223002
XRCC3	X-ray repair complementing defective repair in Chinese hamster cells 3	v2HS_69648
Chk1	Chk 1	v2HS_112994

Doxorubicin Candidates

AKT1	AKT1 V-akt murine thymoma viral oncogene homolog 1	v2HS_188525
ALDH18A1	Aldehyde dehydrogenase 18 family, member A1	v2HS_171097
CCNE2	Cyclin E2	v2HS_37177
CDCA1	CDCA1 Cell division cycle associated 1	v2HS_68506
DCAMKL3	DCAMKL3 Doublecortin and CaM kinase-like 3	v2HS_201790
EGFR	EGFR Epidermal growth factor receptor	v2HS_200678
EPHA6	EPH receptor A6	v2HS_88921
GRB2	growth factor receptor-bound protein 2	v2HS_137365
GSK3B	Homo sapiens similar to Glycogen synthase kinase-3 beta (GSK-3 beta)	v2HS_104795
KRIT1	KRIT1, ankyrin repeat containing	v2HS_62788
MDR1	multidrug resistance p-glycoprotein	v2HS_71
MPHOSPH9	M-phase phosphoprotein 9	v2HS_98516
MYADM	Myeloid-associated differentiation marker	v2HS_170448
NBN	NBN Nibrin (Nbs1)	v2HS_152144
PDGFB	Platelet-derived growth factor beta polypeptide	v2HS_169797
PPP2R3B	protein phosphatase 2 (formerly A), regulatory subunit B, beta	v2HS_85691
RAC1	Ras-related C3 botulinum toxin substrate 1	v2HS_232790
RNF14	RNF14 Ring finger protein 14	v2HS_27619
SMG1	PI-3-kinase-related kinase SMG-1	v2HS_60493
TTN	TTN Titin	v2HS_171632

XRCC3 X-ray repair complementing defective repair in Chinese hamster cells 3 v2HS_69648

DMSO Candidates (Synthetic Lethal without Drug)

CDH1	CDH1 Cadherin 1, type 1, E-cadherin (epithelial)	v2HS_243170
CTDP1	(carboxy-terminal domain, RNA polymerase II, polypeptide A)	v2HS_36865
DBF4	DBF4 DBF4 homolog (S. cerevisiae)	v2HS_85492
PSMC1	Proteasome (prosome, macropain) 26S subunit, ATPase, 1	v2HS_5
RNF14	Ring finger protein 14	v2HS_27619
RUVBL1	RuvB-like 1 (E. coli)	v2HS_14740

5.3.4 Chk1 knockdown sensitizes p53 null cells to Doxorubicin

Recent experiments by Christian Reinhardt showed that Mapkapk2 (MK2), a component of the p38MapK pathway, increases sensitivity to DNA damaging agents in p53 null MEFs in comparison to p53 wild type MEFs.¹⁴⁹ The p38MAPK/MK2 pathway is a general stress-activated kinase pathway that responds to various cellular stimuli including cytokines, hyperosmolarity and UV radiation.¹⁵⁰ MK2 is a critical downstream effector kinase of p38MAPK and Reinhardt showed MK2 is linked to the G2/M checkpoint.¹⁴⁹ Cells treated with DNA damaging agents result in activation of checkpoint response genes including Chk1.¹⁵¹ Chk1 phosphorylates cdc25B/C, inducing their binding to 14-3-3 proteins and sequestering them from cyclin/Cdk substrates.¹⁵¹ Binding of Cdc25B to 14-3-3 proteins is lost in MK2-depleted cells, although, MK2 depletion does not alter total Chk1 levels or reduce Chk1 activation.¹⁴⁹ It was hypothesized that the combination of MK2 and p53 deletion resulted in a synthetic lethal combination under DNA damage because of their checkpoint control related functions. Upon DNA damage, p53 is activated resulting in G1/S arrest through upregulation of p21.¹⁵² Chk1 is phosphorylated by ATR and inactivates the protein phosphatase, cdc25, preventing the G2/M transition.¹⁵¹ The elimination of these two major checkpoint control proteins in the

presence of a DNA damaging agent can result in mitotic catastrophe. Mitotic catastrophe results from entry of cells into mitosis despite the presence of damaged DNA, resulting in cell death.¹⁵³ A number of authors have shown that Chk1 inhibition shows selectivity between p53 wt and p53 null cells for camptothecin and etoposide.^{154, 155} Thus, we wanted to focus on the Chk1 as a candidate from our screen.

Validating candidate genes is a necessary consequence of RNAi screens to prevent off target effects or artifacts from microarray hybridization. Filtering through candidates on a one by one basis can be a tiring and enduring task. A validation assay should be somewhat reproducible, semi-hi throughput and robust assay that can easily be scored in both the p53 null and p53 wt cell line. We devised a high dose drug/recovery assay, similar to the one conducted in Reinhardt et al.¹⁴⁹ In this assay, cells are treated with a high dose of drug for 4 hours. Cells are then counted using a automated cell counter by Guava Technologies and replated at a low density (2000 cells/10 cm dish) in triplicate. After 10 days, the colonies that form are stained with crystal violet. Images are scanned using a gel camera and colonies are counted using Cell Profiler, a freeware program from MIT.¹⁵⁶ Two Chk1 hairpins that scored in my screen were selected for this assay in addition to a Chk1 hairpin I created using the Novartis biopredsi algorithm design program.¹⁵⁷ Westerns for these hairpins are shown in Figure 5.6. The library Chk1 hairpins silence efficiently while the biopred hairpins do not. All the hairpins were put through this assay at two different drug concentration, 0.15nM doxorubicin and 0.25nM doxorubicin (see figure 5.7). In both cases, Chk1 hairpins with efficient knockdown increase sensitivity of p53 null cells to doxorubicin in comparison to p53 wildtype cells. Studies examining knockdown of Chk1 responses to paclitaxel are pending.

Figure 5.6 Chk1 hairpins knock down at varying efficiency.

Westerns for Chk1 are shown. Exportin 5 is used as a loading control.

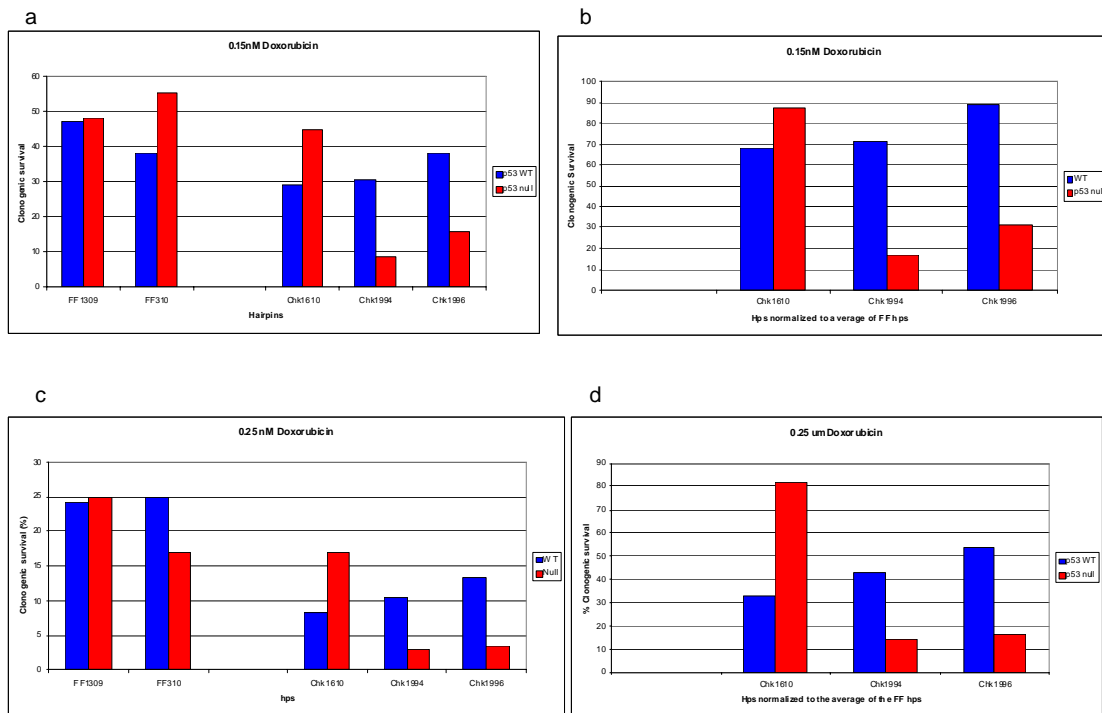
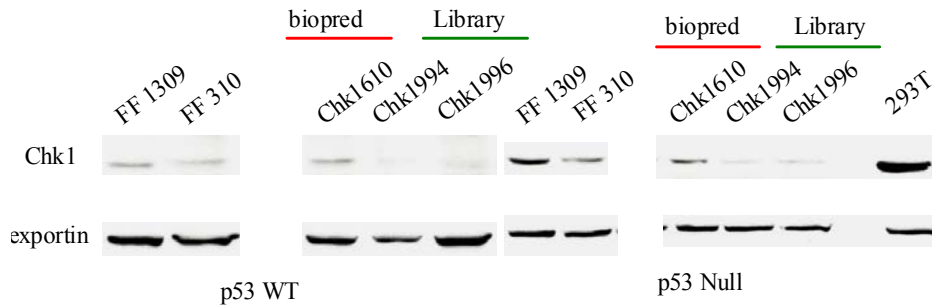


Figure 5.7 Chk1 knockdown increases sensitivity to doxorubicin in p53 null cells. HCT116 p53 WT cell line in blue and HCT116 p53 null cell line in red. a. Clonogenic survival of cell lines with chk1 hairpins or with two different FF control hairpins after treatment with 0.15nM doxorubicin for 4 hours. b. Clonogenic survival calculated from the results presented in a normalized to the average of the FF hairpins. c. Clonogenic survival of cell lines with chk1 hairpins or FF control hairpins after treatment with 0.25nM doxorubicin for 4 hours. D. Clonogenic survival calculated from the results presented in c normalized to the average of the FF hairpins.

5.3.5 Additional validation assays (analysis by Joel Parker)

Numerous different validation methods are being pursued. While screening candidates on one by one basis is the best method, it is also the most time consuming and

least practical method considering the correlation coefficients of the original hybridization. As such, I have performed an additional screen focusing the response of HCT 116 p53 wt and p53 null cells to high and low doses of doxorubicin. The hairpin sets for this screen were paired down to only the Cancer 1000 and cell cycle functional set, as the majority of hits from the original screen came from these cell lines. The number of samples in the re-screen decrease from 48 in the original screen to 24 by focusing on one drug and using only one time point (5 days). This allows all the hybridizations to be performed in one day. Day of hybridization was the primary variable in the original screen. Attempts to label and hybridize all 48 samples at the same time only resulted in poor hybridizations and loss of money. In addition, this focused mini-screen was conducted on an Agilent hybridization platform which is more sophisticated than our previous platform. The barcode probes on the chip were also extended to the full 60mer barcode sequence rather than the previous 50mer sequence. All probes on the chip were printed in both forward and reverse directions. Hairpin probes were also used as not all barcode sequences from the library are available. Joel Parker performed a PCA analysis to identify the causes of variability for this screen. Samples can be analyzed by timepoint, p53 status or biological replicate (see Figure 5.8). In addition, there are strong correlation coefficients between biological replicates (Figure 5.9) in the new screen for both the experimental and reference channel at timepoint 0. These are much higher than the initial screen because we were able to decrease the causes of variability. These coefficients indicate the screen is reproducible.

Joel Parker performed a SAM “two class, unpaired test” with 100 permutations.”¹¹⁹ A candidate list was generated based on a p value of approximately 0.1

for the p53 null cell line. A list of probes that decrease in the doxorubicin samples (normalized to DMSO control) regardless of p53 status are listed in Table 5.5 along with their corresponding p values. All the log ratios are log(DOX/DMSO). A list of probes are depleted in the doxorubicin samples that are p53 null but are not depleted in the samples that are p53 wildtype are shown in Table 5.6. Probes that match genes in the original screen are highlighted in red. Probes appear twice are highlighted in blue.

Table 5.5 Probes depleted in doxorubicin samples regardless of p53 status. These were normalized to their DMSO controls. HR= hairpin reverse sequence BC=Barcode forward sequence BR=barcode reverse. Each column has the log ratio and its corresponding p value. The column on the right looks at DMSO vs Doxorubicin, ignoring the genotype. A negative value indicates the average change across both genotypes is depletion in Doxorubicin relative to DMSO. A positive number signifies an average enrichment in Doxorubicin vs DMSO.

Probe/Hairpin	Symbols	p52 WT		P53 Null		Combined	
		dms0VsDox Log Ratio	dms0VsDox p-value	dms0VsDox Log Ratio	dms0VsDox p-value	dms0VsDox both Log Ratio	dms0VsDox both p-value
HR_v2HS_61652	JAK2	-1.269	0.026	-1.039	0.003	-1.130	0.000
HR_v2HS_6067	eukaryotic translation initiation factor 2-alpha kinase 2	-1.474	0.018	-1.048	0.003	-1.224	0.001
BC_v2HS_202884	cell division cycle 25B	-0.920	0.027	-1.690	0.014	-1.351	0.001
BR_v2HS_68245	CRSP3	-1.540	0.003	-0.635	0.244	-1.071	0.004
BC_v2HS_218486	RAD51	-1.341	0.095	-1.673	0.106	-1.497	0.015
BR_v2HS_218486	RAD51	-1.451	0.069	-1.993	0.137	-1.715	0.018
BR_v2HS_95072	CART1	-1.486	0.051	-0.908	0.219	-1.207	0.019
BR_v2HS_202884	cell division cycle 25B	-0.835	0.021	-1.151	0.120	-1.041	0.023

Table 5.6 Probes depleted in doxorubicin samples for the p53 null cell line with a p value less than 0.1. Joel Parker analyzed data using a two class, unpaired SAM test. Probes that match genes in the original screen are highlighted in red. Probes appear twice are highlighted in blue. The log ratios are normalized to DMSO for each cell line and are shown with their corresponding p values.

Probe/hairpin	Symbols	p53 WT		p53 NULL	
		dms0VsDox Log Ratio	p value	dms0VsDox Log Ratio	p value
BC_v2HS_152436	ODF2	0.636	0.039	-0.962	0.031
BR_v2HS_151981	NUDT1	0.887	0.073	-1.032	0.069
HR_v2HS_5789	APPBP2	-0.533	0.076	-0.988	0.000
BC_v2HS_98516	MPHOSPH9	1.499	0.123	-2.148	0.072
BR_v2HS_203598	NR6A1	1.340	0.134	-1.999	0.037
BR_v2HS_169793	PDGFB	0.565	0.147	-0.890	0.066
BC_v2HS_203598	NR6A1	1.139	0.155	-1.634	0.074
BC_v2HS_131642	RPL30	1.304	0.169	-1.856	0.088
BR_v2HS_131642	RPL30	1.309	0.182	-2.193	0.105
BC_v2HS_151981	NUDT1	0.911	0.225	-0.951	0.089

BC_v2HS_203371	RAD1	0.460	0.348	-1.773	0.018
BR_v2HS_112050	ALDH3A1	0.850	0.400	-1.595	0.016
BR_v2HS_42033	FANCG	-0.437	0.415	-1.040	0.045
HR_v2HS_222459	POLB	-0.166	0.573	-0.900	0.006
BC_v2HS_239527	VDR	-0.451	0.582	-1.398	0.064
BR_v2HS_24953	MAX	0.490	0.642	-1.548	0.090
BC_v2HS_133341	IL13	-0.091	0.649	-1.071	0.091
BC_v2HS_42033	FANCG	-0.215	0.712	-0.976	0.051
BR_v2HS_112920	CDK9	0.261	0.715	-1.116	0.080
BC_v2HS_35223	RBL1	0.426	0.722	-1.066	0.093
BR_v2HS_52026	NANS	-0.487	0.722	-1.826	0.037
BC_v2HS_112050	ALDH3A1	0.325	0.726	-1.382	0.020
BR_v2HS_254646	CUL2	-0.302	0.841	-1.296	0.010
BC_v2HS_15310	RABGAP1	0.296	0.846	-2.502	0.034
BR_v2HS_191980		-0.141	0.858	-1.903	0.043
BR_v2HS_201747	RAD51C	0.230	0.858	-1.135	0.027
BC_v2HS_150245	CAT	0.109	0.882	-1.376	0.068
BC_v2HS_113327	CTSZ	-0.107	0.903	-1.629	0.082
BC_v2HS_188525	AKT1	0.120	0.917	-1.423	0.057

Figure 5.8 PCA analysis of Agilent screen. a. Samples from timepoint 0 are in black and T1 are in red. b. PCA analysis of samples by p53 status. Minus (black) signifies p53 null and plus (red) signifies p53 wildtype. c. Samples analyzed by biological replicate.

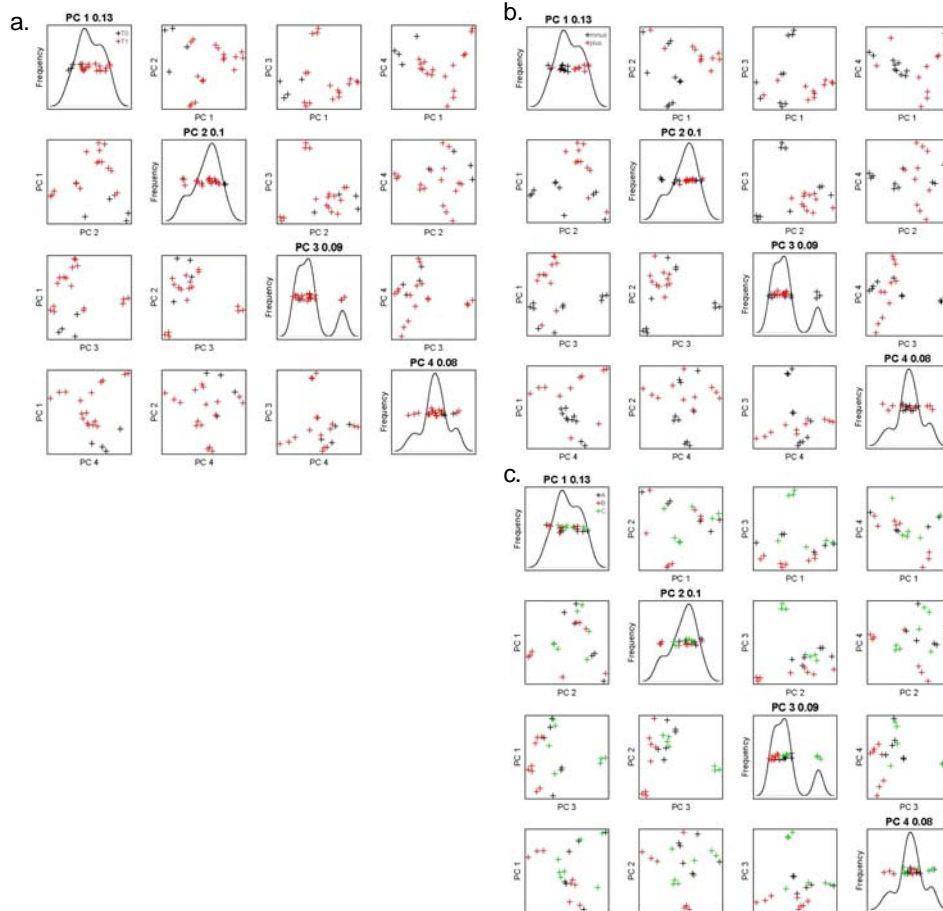


Figure 5.9 Agilent re-screen samples have strong correlation coefficients between biological replicates in both the experimental and reference channel. Correlation coefficients for barcode probes in both the experimental (left) and reference (right) channel at time point 0. There are three biological replicates (A, B and C).

Correlation coefficients for Barcode probes in experimental channel				Correlation coefficients for Barcode probes in reference channel			
	replicate	B	C		replicate	B	C
	<u>P53+/+</u>	<i>T0</i>	<i>T0</i>		<u>P53+/+</u>	<i>T0</i>	<i>T0</i>
A	<u>T0</u>	0.8981	0.9379	A	<u>T0</u>	0.948	0.9142
C	<u>T0</u>	0.9051		C	<u>T0</u>	0.9075	
<hr/>				<hr/>			
	<u>P53-/-</u>	B	C		<u>P53-/-</u>	B	C
A	<u>T0</u>	0.9602	0.9692	A	<u>T0</u>	0.9857	0.9747
C	<u>T0</u>	0.9487		C	<u>T0</u>	0.9781	

In addition, we are taking advantage of new deep sequencing technologies for our validation assay. I created a set of Chk1 related genes in the GIPZ lentiviral vector from our shRNA library. These were put through the high dose/ recovery assay. The remaining colonies were lysed and their genomic DNA was extracted. The half hairpin sequences will be amplified by PCR and sequenced with Solexa technology. Since Solexa provides approximately 2 million reads and our Chk1 related set is miniscule in comparison, we have put a two nucleotide DNA barcode in our primers that will allow us to run multiple samples in the same Solexa flow cell. In addition, I have a pooled approximately 500 hairpins from the first screen that have passed a broad selection criteria. This group will also be put through the high dose/recovery assay and analyzed via Solexa based sequencing. Hairpins candidates that have resulted from these additional assays will be

placed in 96 well plate growth curves and assayed with cell titer glo, a luminescent cell viability assay. Ultimately, it would be interesting to see if any of these candidates can be linked to mitotic catastrophe by measuring the mitotic index through FACS for histone 3 phosphorylation. We can also do FACS analysis for MPM2 proteins which are phosphorylated during mitosis. As the goal of this project is to result in better combination therapy, it would be interesting if we could recapitulate our findings with a small molecule inhibitor for the gene of interest and doxorubicin in a p53 null cancer cells.

5.4 Discussion

5.4.1 Synthetic lethal screen execution, data analysis and validation.

This chapter outlines the execution of a synthetic lethal screen utilizing a library of thousands of hairpins and the challenges associated with both analyzing and validating its data. The true core of any RNAi screen is the experimental system. It is extremely important to use a library that includes genes of known and unknown function with efficient and potent knockdown. Genes of known function can help with the validation procedure as experimental reagents, such as antibodies for western blots, will be readily available. Certainly, a database of hairpins with validated knockdown efficiencies would be helpful in performing screens. It is also crucial to have an assay system that is high throughput and reproducible with known positive and negative controls. These components are key to the production of any viable data.

However, one aspect of RNAi screens that is still highly debated is data analysis. It was estimated that it would take a year for us to develop bioinformatics methods for our RNAi screens (personal communication, Greg Hannon). Through a collaboration with Joel Parker, we explored a number of different analysis methods for these screens in a matter of months. While each method has benefits over the others, a combination of these methods were used to generate a candidate list. It is important to have a candidate list that can be trusted to reduce the number of false positives. False positives can bottleneck the validation process and ultimately result in increased cost of experiments and loss of time. The number of false positives can be decreased by increasing the amount of biological replicates and reducing extraneous variables. Our initial p53 screen had a considerable amount of background because it was hybridized on separate days due to the high volume of samples. It may be better to do less complex screens with more biological replicates. Since this screen was performed, we moved our hybridization platform to Agilent Technologies. This platform is more user friendly and is not as technically overwhelming as the Nimblegen platform. Data from chapter 4 illustrated experiments that attempted to decrease our false negatives by improving our microarray hybridizations. Our data analysis methods from this chapter were attempts to reduce our false positives and focus our energy on significant microarray changes. For this reason, one of my selection criteria was selecting probes that exhibited more than 2 fold changes on a microarray. I wanted to ensure the phenotype I was looking for would be easily scorable in a validation assay. My early validation attempts focused on hairpins that showed moderate change on the microarray but had strong p values. Upon validation, these changes were not readily detectable on the phenotypic level.

A year and a half was spent on generating validating assays for this screen. Part of this could have been expedited if the experimental assay had better positive and negative controls. In addition, all hairpins had to be moved from the library vector PSM2C to MLP, the vector which the screen was performed. Different validation attempts have resulted in certain flexible guidelines such as below.

- Secondary screens: It's often necessary to repeat the screen to narrow down the false positives. This can be done by narrowing down the pool of candidates to a smaller amount and doing a "focused mini-screen." This can help reduce any variables in the first screen focus any secondary assays.
- Secondary assays (one by one approach): No matter how specific the primary screen seems to be, secondary assays are important for identifying the genes that are particularly relevant.¹⁵⁸ First and foremost, it will be best to validate hairpins in the same assay as the screen. After this is accomplished, it is necessary to show the phenotype is not simply an off target effect by generating multiple hairpins for that gene, and testing their knockdown efficiency through western or QPCR. Finally, a rescue experiment or overexpression of the gene of interest can also be valuable in validating its phenotypic effect. For our synthetic lethal screen, we chose to use a high dose/recovery assay, where cells are treated with a high dose of drug and then the amount of surviving colonies are monitored. This system is a robust assay that is different from our initial experiment, but is easily scorable.

After the gene is validated, creative experiments can be designed for studying gene function and interactions with other genes. Computer programs such as Ingenuity Pathway Analysis can help researchers comb current literature to find what pathways the

gene of interest lies in, other proteins or genes it may interact with and the availability of any small molecule inhibitors. While this synthetic lethal screen had many false positives that clouded any pathway analysis, Ingenuity Pathway analysis provided an impressive amount of information about biological molecules and processes.

A complication of performing synthetic lethal screens with an RNAi library, is that full knockdown of the gene of interest is not achieved. RNAi induces a hypomorphic state with a certain basal amount of protein still present. While a decrease in cellular fitness might be evident, total gene deletion may be necessary to see cell death. However, it would be harder to find lethality using stable integrants that need to be selected for. Use of an inducible vector may help finding a synthetic lethal combination that is “truer” as gene expression can be turned off after selection with the addition of doxycycline or restored by removing doxycycline. However, this does not alleviate the problem of working with a hypomorph where cell death may not be readily detectable.

5.4.2 Chk1 inhibitors for combination therapy with doxorubicin in p53 null cells

The identification of Chk1 as a gene that can modify sensitivity for a popular chemotherapeutic in p53 null cells is quite interesting as it solidifies Chk1’s role as a key mediator in the G2/M transition and illustrates its potential as a chemotherapeutic target. There is one commonly used small molecule inhibitor for Chk1 in the lab called UCN01 or 7-hydroxystaurosporine. It is a synthetic derivative of the antibiotic, staurosporine, and its specificity to Chk1 inhibition is questionable. It has also been shown to inhibit AKT, Mapkapk2 (MK2), calcium dependent protein kinase C and cyclin dependent kinases.¹⁴⁹
¹⁵⁹ Studies by Gerardo Ferbeyre showed that ATM, ATR or Chk1 knockdown in p53-defective PC3 prostate cancer cells sensitized them to the killing effects of

doxorubicin.¹⁶⁰ His results also suggest that this is due to an increase in E2F activity and increased expression of E2F target and proliferation marker, PCNA. PCNA expression is associated with high proliferation rates and DNA synthesis. Febreyre hypothesizes PCNA increases drug sensitivity by accelerating DNA replication and inducing p53 independent apoptosis via ING1b, a candidate tumor suppressor.¹⁶⁰ He also suggests E2F may increase drug sensitivity by inducing proapoptotic genes and notes the presence of nucleosomal sized DNA fragments in PC3 cells with ATM knockdown after drug treatment. While this data supports combining DNA damage agents with checkpoint inhibitors to selectively kill p53 deficient tumor cells, no clinical trials exist using doxorubicin and UCN01 as combination therapy. Only 22 clinical trials with UCN01 are listed on clinicaltrials.gov website. Several trials exist between UCN01 and DNA damaging agents, such as cisplatin/carboplatin or Fluorouracil, a pyrimidine analog. Four other chk1 inhibitors are in phase 1 trials.¹⁶¹ Unfortunately, the clinical trials website of the US National Institute of Health lists a total of five trials for all them.¹⁶² Two agents, AZD7762 and XL844 are being evaluated with gemcitabine, a nucleoside analog.

While more validation studies are pending, it will be interesting to see any additional genes that surface from this screen and whether they will be checkpoint related genes. A candidate that is related to Chk1 is Nbs1 (Nibrin). The Nbs1 protein is essential in mammals and functions with in the Mre11/Rad50/Nbs1 (MRN) complex to recruit ATM to the site of double strand DNA breaks.¹⁶³ Nbs1 is mutated in patients with Nijmegen Breakage Syndrome resulting in microcephaly, growth retardation, immunodeficiency, chromosome instability, radiation sensitivity, and a strong predisposition to lymphoid malignancy.¹⁶⁴ The Nbs1 protein plays many roles in the

functions of the MRN complex, some of which are still being elucidated.¹⁶⁵ These include localizing the MRN complex to the nucleus, regulating the MR catalytic functions and activating dimeric ATM.¹⁶⁶ Khanna and colleagues showed that the NBS1 gene product is required for optimal IR-induced phosphorylation of Chk1 on Ser-317 but is dispensable for its phosphorylation in response to Hydroxy urea and UV, suggesting that NBS1 assists ATM in targeting some of its substrates.¹⁶⁶ The association of Nbs1 to Chk1 and ATM make it a likely candidate for my screen. Mutation of NBS1 has been shown to enhance cisplatin-induced DNA damage and cytotoxicity in head and neck cancer.¹⁶⁷ Hopefully, Nbs1 will be a strong mediator of doxorubicin sensitivity and not a false positive. Two hairpins targeting Nbs1 scored in the original screen. Most recently, one of the two hairpins for Nbs1 scored again in the independently re-screen hybridized in the Agilent platform.

RNA interference screens such as the one in this chapter can also help investigate the biology of cancer processes. A hairpin for Claspin scored in my original screen causing an increase in p53 null cells in comparison to p53 wildtype cells in the presence of doxorubicin. Claspin, has been hypothesized to be a scaffolding protein that brings together ATR and Chk1,¹⁶⁸ a sensor for DNA replication and involved in cell proliferation. Most recently Chk1 was found to stabilize Claspin levels in the cell.¹⁶⁹ It was recently discovered that Chk1 regulates the DNA-damaged induced ubiquitination of PCNA, which facilitates continuous replication of damaged DNA.¹⁶⁹ In addition, depletion of Claspin was shown to decrease PCNA ubiquitination after DNA damage.¹⁶⁹ It was proposed that Chk1 may protect replication forks by stabilizing Claspin. While the details of the interactions between these proteins are still being clarified, it would be interesting

to see how Claspin knockdown affects cellular responses to DNA damaging agents. Because Claspin increased the p53 null cells in response to chemotherapy, it may be a marker for cells that are resistant to DNA damaging agents. A small study of 103 primary invasive breast carcinomas demonstrated that Chk1 and Claspin transcript levels were elevated in tumors showing large tumor size (>2 cm), high histologic grade and lack of the estrogen and progesterone receptor expression.¹⁷⁰ While this study did not examine if these patients received chemotherapy or their clinical outcomes, it does reflect how cancer classification and treatment is beginning to be correlated with gene transcript levels. It would be interesting to see if Claspin is a modifier of cancer chemotherapeutics.

A number of genes also scored in the second rescreen focusing on doxorubicin treatment. This is particularly striking because this screen was conducted using a different isolate of HCT116 cells. These genes will be focused on for further validation studies. Also, I examined genes that increased sensitivity to doxorubicin regardless of p53 status. Rad51 and cdc25B both scored using two different probes. These are indicators that the screen is working as rad51 has been shown to increase sensitivity to doxorubicin.^{171, 172} These candidates will be used as controls for further validation experiments.

Literature Cited

1. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* 100, 57-70 (2000).
2. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nat Med* 10, 789-99 (2004).
3. Papac, R. J. Origins of cancer therapy. *Yale J Biol Med* 74, 391-8 (2001).
4. Frei, E., 3rd et al. The effectiveness of combinations of antileukemic agents in inducing and maintaining remission in children with acute leukemia. *Blood* 26, 642-56 (1965).
5. Chabner, B. A. & Roberts, T. G., Jr. Timeline: Chemotherapy and the war on cancer. *Nat Rev Cancer* 5, 65-72 (2005).
6. Sawyers, C. Targeted cancer therapy. *Nature* 432, 294-7 (2004).
7. Roche-Lestienne, C. et al. Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. *Blood* 100, 1014-1018 (2002).
8. Shah, N. P. et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* 2, 117-25 (2002).
9. Sordella, R., Bell, D. W., Haber, D. A. & Settleman, J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* 305, 1163-7 (2004).
10. van 't Veer, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530-6 (2002).
11. Agendia, I.
12. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-11 (1998).
13. Hannon, G. J. RNA interference. *Nature* 418, 244-51 (2002).
14. Birchler, J. A. & Kavi, H. H. Molecular biology. Slicing and dicing for small RNAs. *Science* 320, 1023-4 (2008).
15. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363-6 (2001).

16. Carmell, M. A. & Hannon, G. J. RNase III enzymes and the initiation of gene silencing. *Nat Struct Mol Biol* 11, 214-8 (2004).
17. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293-6 (2000).
18. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843-54 (1993).
19. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858-62 (2001).
20. Lee, Y. et al. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* 23, 4051-60 (2004).
21. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* 431, 343-9 (2004).
22. Han, J. et al. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18, 3016-27 (2004).
23. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415-9 (2003).
24. Song, J. J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* 10, 1026-32 (2003).
25. He, L. & Hannon, G. J. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 5, 522-31 (2004).
26. Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* 7, 719-23 (2005).
27. Liu, J. et al. Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science* 305, 1437-1441 (2004).
28. Tolia, N. H. & Joshua-Tor, L. Slicer and the Argonautes. *Nat Chem Biol* 3, 36-43 (2007).
29. Song, J.-J. & Joshua-Tor, L. Argonaute and RNA -- getting into the groove. *Current Opinion in Structural Biology* 16, 5-11 (2006).

30. Faehnle, C. R. & Joshua-Tor, L. Argonautes confront new small RNAs. *Current Opinion in Chemical Biology* 11, 569-577 (2007).
31. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 9, 22-32 (2008).
32. Parker, J. S. & Barford, D. Argonaute: a scaffold for the function of short regulatory RNAs. *Trends in Biochemical Sciences* 31, 622-630 (2006).
33. Davis, S. & Watson, J. C. In vitro activation of the interferon-induced, double-stranded RNA-dependent protein kinase PKR by RNA from the 3' untranslated regions of human alpha-tropomyosin. *Proc Natl Acad Sci U S A* 93, 508-13 (1996).
34. Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H. & Williams, B. R. Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* 5, 834-9 (2003).
35. Elbashir, S. M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411, 494-8 (2001).
36. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. & Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 16, 948-58 (2002).
37. Sarov, M. & Stewart, A. F. The best control for the specificity of RNAi. *Trends Biotechnol* 23, 446-8 (2005).
38. Paddison, P. J., Caudy, A. A. & Hannon, G. J. Stable suppression of gene expression by RNAi in mammalian cells. *Proc Natl Acad Sci U S A* 99, 1443-8 (2002).
39. Pei, Y. & Tuschl, T. On the art of identifying effective and specific siRNAs. *Nat Meth* 3, 670-676 (2006).
40. Jackson, A. L. et al. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotech* 21, 635-637 (2003).
41. Barik, S. RNAi in moderation. *Nat Biotech* 24, 796-797 (2006).
42. Wolters, N. M. & MacKeigan, J. P. From sequence to function: using RNAi to elucidate mechanisms of human disease. *Cell Death Differ* 15, 809-19 (2008).
43. Bernards, R., Brummelkamp, T. R. & Beijersbergen, R. L. shRNA libraries and their use in cancer genetics. *Nat Methods* 3, 701-6 (2006).
44. Chang, K., Elledge, S. J. & Hannon, G. J. Lessons from Nature: microRNA-based shRNA libraries. *Nat Methods* 3, 707-14 (2006).

45. Root, D. E., Hacohen, N., Hahn, W. C., Lander, E. S. & Sabatini, D. M. Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Meth* 3, 715-719 (2006).
46. Silva J, C. K., Hannon GJ, Rivas FV. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. *Oncogene* 51, 8401-9. (2004).
47. Dickins, R. A. et al. Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nat Genet* 37, 1289-95 (2005).
48. Dickins, R. A. et al. Tissue-specific and reversible RNA interference in transgenic mice. *Nat Genet* 39, 914-21 (2007).
49. Silva, J., Chang, K., Hannon, G. J. & Rivas, F. V. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. *Oncogene* 23, 8401-9 (2004).
50. Burgess, D. J. RNAi Genetic Screens for Chemotherapy Response Modifier Genes. (PhD dissertation, Watson School of Biological Sciences) (2007).
51. Hemann, M. T. et al. An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nat Genet* 33, 396-400 (2003).
52. Kaelin, W. G., Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer* 5, 689-98 (2005).
53. Aza-Blanc, P. et al. Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. *Mol Cell* 12, 627-37 (2003).
54. Westbrook, T. F. et al. A genetic screen for candidate tumor suppressors identifies REST. *Cell* 121, 837-48 (2005).
55. Kolfschoten, I. G. et al. A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. *Cell* 121, 849-58 (2005).
56. MacKeigan, J. P., Murphy, L. O. & Blenis, J. Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. *Nat Cell Biol* 7, 591-600 (2005).
57. Berns, K. et al. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* 428, 431-7 (2004).
58. Paddison, P. J. et al. A resource for large-scale RNA-interference-based screens in mammals. *Nature* 428, 427-31 (2004).

59. Brummelkamp, T. R., Nijman, S. M., Dirac, A. M. & Bernards, R. Loss of the cylindromatosis tumour suppressor inhibits apoptosis by activating NF-kappaB. *Nature* 424, 797-801 (2003).
60. Ngo, V. N. et al. A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* 441, 106-110 (2006).
61. Whitehurst, A. W. et al. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature* 446, 815-9 (2007).
62. Berns, K. et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 12, 395-402 (2007).
63. Soutschek, J. et al. Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* 432, 173-8 (2004).
64. Bitko, V., Musiyenko, A., Shulyayeva, O. & Barik, S. Inhibition of respiratory viruses by nasally administered siRNA. *Nat Med* 11, 50-5 (2005).
65. de Fougerolles, A., Vornlocher, H. P., Maraganore, J. & Lieberman, J. Interfering with disease: a progress report on siRNA-based therapeutics. *Nat Rev Drug Discov* 6, 443-53 (2007).
66. Tara, M. L., Howell, F. M. & Carl, D. N. Not miR-ly small RNAs: Big potential for microRNAs in therapy. *The Journal of allergy and clinical immunology* 121, 309-319 (2008).
67. Krutzfeldt, J. et al. Silencing of microRNAs in vivo with 'antagomirs'. *Nature* 438, 685-689 (2005).
68. He, L. et al. A microRNA component of the p53 tumour suppressor network. *Nature* 447, 1130-1134 (2007).
69. He, L. et al. A microRNA polycistron as a potential human oncogene. *Nature* 435, 828-833 (2005).
70. Bignell, G. R. et al. Identification of the familial cylindromatosis tumour-suppressor gene. *Nat Genet* 25, 160-165 (2000).
71. Lakhani, S. R. Putting the Brakes on Cylindromatosis? *N Engl J Med* 350, 187-188 (2004).
72. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs Exhibit Strand Bias. *Cell* 115, 209-216 (2003).

73. Carmell, M. A., Xuan, Z., Zhang, M. Q. & Hannon, G. J. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* 16, 2733-42 (2002).
74. Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear export of microRNA precursors. *Science* 303, 95-8 (2004).
75. Nakayashiki, H. et al. RNA silencing as a tool for exploring gene function in ascomycete fungi. *Fungal Genet Biol* 42, 275-83 (2005).
76. Tang, J., Erikson, R. L. & Liu, X. Checkpoint kinase 1 (Chk1) is required for mitotic progression through negative regulation of polo-like kinase 1 (Plk1). *Proc Natl Acad Sci U S A* 103, 11964-9 (2006).
77. Dasgupta, R. & Perrimon, N. Using RNAi to catch Drosophila genes in a web of interactions: insights into cancer research. *Oncogene* 23, 8359-65 (2004).
78. Fraser, A. Towards full employment: using RNAi to find roles for the redundant. *Oncogene* 23, 8346-52 (2004).
79. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281-97 (2004).
80. Reinhart, B. J. et al. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901-6. (2000).
81. Ketting, R. F. et al. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* 15, 2654-9. (2001).
82. Grishok, A. et al. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106, 23-34. (2001).
83. Knight, S. W. & Bass, B. L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* 293, 2269-71. (2001).
84. Hutvagner, G. et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834-8. (2001).
85. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231-5 (2004).
86. Landthaler, M., Yalcin, A. & Tuschl, T. The Human DiGeorge Syndrome Critical Region Gene 8 and Its *D. melanogaster* Homolog Are Required for miRNA Biogenesis. *Curr Biol* 14, 2162-7 (2004).

87. Gregory, R. I. et al. The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235-40 (2004).
88. Siolas, D. et al. Synthetic shRNAs as potent RNAi triggers. *Nat Biotechnol* 23, 227-31 (2005).
89. Schwarz, D. S. et al. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199-208 (2003).
90. Zeng, Y., Wagner, E. J. & Cullen, B. R. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* 9, 1327-33. (2002).
91. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. & Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 16, 948-58. (2002).
92. Chen, C. Z., Li, L., Lodish, H. F. & Bartel, D. P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* 303, 83-6 (2004).
93. Zeng, Y. & Cullen, B. R. Sequence requirements for micro RNA processing and function in human cells. *Rna* 9, 112-23 (2003).
94. Kawasaki, H. & Taira, K. Short hairpin type of dsRNAs that are controlled by tRNA(Val) promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells. *Nucleic Acids Res* 31, 700-7. (2003).
95. Brummelkamp, T. R., Bernards, R. & Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296, 550-553 (2002).
96. Zheng, L. et al. An approach to genomewide screens of expressed small interfering RNAs in mammalian cells. *Proc Natl Acad Sci U S A* 101, 135-40 (2004).
97. Stegmeier, F., Hu, G., Rickles, R. J., Hannon, G. J. & Elledge, S. J. A Lentiviral microRNA-based System for Single Copy Pol II Regulated RNAi in mammalian Cells. *Proc Natl Acad Sci U S A*, in press (2005).
98. Li, M. Z. & Elledge, S. J. MAGIC, an in vivo genetic method for the rapid construction of recombinant DNA molecules. *Nat Genet* 37, 311-9 (2005).
99. Li, X. et al. Generation of destabilized green fluorescent protein as a transcription reporter. *J Biol Chem* 273, 34970-5 (1998).
100. Elledge, S. J. & Walker, G. C. Phasmid vectors for identification of genes by complementation of *Escherichia coli* mutants. *J Bacteriol* 162, 777-83 (1985).

101. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97, 6640-5 (2000).
102. Cherepanov, P. P. & Wackernagel, W. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* 158, 9-14 (1995).
103. Chalker, A. F., Leach, D. R. & Lloyd, R. G. *Escherichia coli* sbcC mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* 71, 201-5 (1988).
104. Caudy, A. A., Myers, M., Hannon, G. J. & Hammond, S. M. Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev* 16, 2491-6. (2002).
105. Silva, J. M. et al. Profiling essential genes in human mammary cells by multiplex RNAi screening. *Science* 319, 617-20 (2008).
106. Lucito, R. a. W., M. Preparation of Slides and Hybridization. (ed. Sambrook), e. D. B. a. J.) (Cold Spring Harbor Press, Cold Spring Harbor, NY., 2003).
107. website, A.
108. Ramdas, L. et al. Sources of nonlinearity in cDNA microarray expression measurements. *Genome Biol* 2, RESEARCH0047 (2001).
110. Ambion, I. in *Amino Alkyl MessageAmp Instruction Manual*.
111. Wall, M. E., Andreas Rechtsteiner, Rocha, LM. *Singular value decomposition and principal component analysis*". (Kluwer, Norwell, MA 2003).
112. Alter, O., Brown, P. O. & Botstein, D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A* 97, 10101-6 (2000).
113. Haney, S. A. Increasing the robustness and validity of RNAi screens. *Pharmacogenomics* 8, 1037-49 (2007).
114. Moore, D. *The Basic Practice of Statistics (Paper) & Student CD*. (W. H. Freeman, San Francisco, 2006).
115. Roberts, F. a. R., D. (2008).
116. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98, 5116-21 (2001).

117. Barry, W. T., Nobel, A. B. & Wright, F. A. Significance analysis of functional categories in gene expression studies: a structured permutation approach. *Bioinformatics* 21, 1943-9 (2005).
118. Elledge, S. J. & Hannon, G. J. An open letter to cancer researchers. *Science* 310, 439-41 (2005).
119. Lucito, R. et al. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res* 13, 2291-305 (2003).
120. Chabner, B. A. & Roberts, T. G. CHEMOTHERAPY AND THE WAR ON CANCER. *Nature Reviews Cancer* Nat Rev Cancer 5, 65-72 (2005).
121. Weinstein, I. B. CANCER: Enhanced: Addiction to Oncogenes--the Achilles Heel of Cancer. *Science* 297, 63-64 (2002).
122. Hübner, A., Jaeschke, A. & Davis, R. J. Oncogene Addiction: Role of Signal Attenuation. *Developmental Cell* 11, 752-754 (2006).
123. Sawyers, C. L. Making Progress through Molecular Attacks on Cancer. *Cold Spring Harbor Symposia on Quantitative Biology* 70, 479-482 (2005).
124. Shaffer, A. L. et al. IRF4 addiction in multiple myeloma. *Nature* 454, 226-231 (2008).
125. Willingham, A. T., Deveraux, Q. L., Hampton, G. M. & Aza-Blanc, P. RNAi and HTS: exploring cancer by systematic loss-of-function. *Oncogene* 23, 8392-400 (2004).
126. Lucchesi, J. C. SYNTHETIC LETHALITY AND SEMI-LETHALITY AMONG FUNCTIONALLY RELATED MUTANTS OF DROSOPHILA MELANOGASTER. *Genetics* 59, 37-44 (1968).
127. Hartwell, L. H., Szankasi, P., Roberts, C. J., Murray, A. W. & Friend, S. H. Integrating Genetic Approaches into the Discovery of Anticancer Drugs. *Science* 278, 1064-1068 (1997).
128. Iorns, E., Lord, C. J., Turner, N. & Ashworth, A. Utilizing RNA interference to enhance cancer drug discovery. *Nat Rev Drug Discov* 6, 556-68 (2007).
129. Xue, W. et al. Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature* 445, 656-60 (2007).
130. Riley, T., Sontag, E., Chen, P. & Levine, A. Transcriptional control of human p53-regulated genes. *Nat Rev Mol Cell Biol* 9, 402-12 (2008).

131. Hirao, A. et al. DNA damage-induced activation of p53 by the checkpoint kinase Chk2. *Science* 287, 1824-7 (2000).
132. Gasco, M., Shami, S. & Crook, T. The p53 pathway in breast cancer. *Breast Cancer Res* 4, 70 - 76 (2002).
133. Miller, L. D. et al. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 102, 13550-13555 (2005).
134. Bunz, F. et al. Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science* 282, 1497-501 (1998).
135. Torrance, C. J., Agrawal, V., Vogelstein, B. & Kinzler, K. W. Use of isogenic human cancer cells for high-throughput screening and drug discovery. *Nat Biotechnol* 19, 940-5 (2001).
136. Cohen, P. Protein kinases--the major drug targets of the twenty-first century? *Nat Rev Drug Discov* 1, 309-15 (2002).
137. Froelich-Ammon, S. J. & Osheroff, N. Topoisomerase poisons: harnessing the dark side of enzyme mechanism. *J Biol Chem* 270, 21429-32 (1995).
138. Health, U. N. I. o. in www.clinicaltrials.gov (2008).
139. Kumar, N. Taxol-induced polymerization of purified tubulin. Mechanism of action. *J. Biol. Chem.* 256, 10435-10441 (1981).
140. Smyth, G. K. & Speed, T. Normalization of cDNA microarray data. *Methods* 31, 265-73 (2003).
141. Quackenbush, J. Microarray data normalization and transformation. *Nat Genet* 32 Suppl, 496-501 (2002).
142. Storey, J. D., Xiao, W., Leek, J. T., Tompkins, R. G. & Davis, R. W. Significance analysis of time course microarray experiments. *Proc Natl Acad Sci U S A* 102, 12837-42 (2005).
143. Leek, J. T., Mosen, E., Dabney, A. R. & Storey, J. D. EDGE: extraction and analysis of differential gene expression. *Bioinformatics* 22, 507-8 (2006).
144. Storey, J. D., Dai, J. Y. & Leek, J. T. The optimal discovery procedure for large-scale significance testing, with applications to comparative microarray experiments. *Biostatistics* 8, 414-32 (2007).

145. Pichler, A., Zelcer, N., Prior, J. L., Kuil, A. J. & Piwnica-Worms, D. In vivo RNA interference-mediated ablation of MDR1 P-glycoprotein. *Clin Cancer Res* 11, 4487-94 (2005).
146. Reinhardt, H. C., Aslanian, A. S., Lees, J. A. & Yaffe, M. B. p53-deficient cells rely on ATM- and ATR-mediated checkpoint signaling through the p38MAPK/MK2 pathway for survival after DNA damage. *Cancer Cell* 11, 175-89 (2007).
147. Pimienta, G. & Pascual, J. Canonical and alternative MAPK signaling. *Cell Cycle* 6, 2628-32 (2007).
148. Chen, Y. & Sanchez, Y. Chk1 in the DNA damage response: conserved roles from yeasts to mammals. *DNA Repair* 3, 1025-1032 (2004).
149. Vousden, K. H. & Lu, X. Live or let die: the cell's response to p53. *Nat Rev Cancer* 2, 594-604 (2002).
150. Erenpreisa, J., Kalejs, M. & Cragg, M. S. Mitotic catastrophe and endomitosis in tumour cells: an evolutionary key to a molecular solution. *Cell Biol Int* 29, 1012-8 (2005).
151. Chen, Z. et al. Selective Chk1 inhibitors differentially sensitize p53-deficient cancer cells to cancer therapeutics. *Int J Cancer* 119, 2784-94 (2006).
152. Carrassa, L., Broggin, M., Erba, E. & Damia, G. Chk1, but not Chk2, is involved in the cellular response to DNA damaging agents: differential activity in cells expressing or not p53. *Cell Cycle* 3, 1177-81 (2004).
153. Lamprecht, M. R., Sabatini, D. M. & Carpenter, A. E. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques* 42, 71-5 (2007).
154. Huesken, D. et al. Design of a genome-wide siRNA library using an artificial neural network. *Nat Biotechnol* 23, 995-1001 (2005).
155. Boutros, M. & Ahringer, J. The art and design of genetic screens: RNA interference. *Nat Rev Genet* 9, 554-66 (2008).
156. NCI Drug Dictionary, Definition of 7-hydroxystaurosporine.
<http://www.cancer.gov/Templates/drugdictionary.aspx?CdrID=42305>
157. Mukhopadhyay, U. K., Senderowicz, A. M. & Ferbeyre, G. RNA silencing of checkpoint regulators sensitizes p53-defective prostate cancer cells to chemotherapy while sparing normal cells. *Cancer Res* 65, 2872-81 (2005).
158. O'Connor, M. J., Martin, N. M. & Smith, G. C. Targeted cancer therapies based on the inhibition of DNA strand break repair. *Oncogene* 26, 7816-24 (2007).

159. Health, U. N. I. o.
160. Zhang, Y., Zhou, J. & Lim, C. U. K. The role of NBS1 in DNA double strand break repair, telomere stability, and cell cycle checkpoint control. *Cell Res* 16, 45-54.
161. Kondratenko, I., Paschenko, O., Polyakov, A. & Bologov, A. Nijmegen breakage syndrome. *Adv Exp Med Biol* 601, 61-7 (2007).
162. Lee, J. H. & Paull, T. T. Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. *Oncogene* 26, 7741-7748.
163. Gatei, M. et al. Ataxia-telangiectasia-mutated (ATM) and NBS1-dependent Phosphorylation of Chk1 on Ser-317 in Response to Ionizing Radiation. *J. Biol. Chem.* 278, 14806-14811 (2003).
164. Tran, H. M. et al. Mutant Nbs1 enhances cisplatin-induced DNA damage and cytotoxicity in head and neck cancer. *Otolaryngol Head Neck Surg* 131, 477-84 (2004).
165. Kumagai, A., Kim, S. M. & Dunphy, W. G. Claspin and the activated form of ATR-ATRIP collaborate in the activation of Chk1. *J Biol Chem* 279, 49599-608 (2004).
166. Yang, X. H., Shiotani, B., Classon, M. & Zou, L. Chk1 and Claspin potentiate PCNA ubiquitination. *Genes Dev* 22, 1147-52 (2008).
167. Verlinden, L. et al. The E2F-regulated gene Chk1 is highly expressed in triple-negative estrogen receptor /progesterone receptor /HER-2 breast carcinomas. *Cancer Res* 67, 6574-81 (2007).

Chapter 6

Conclusions and Perspectives

6.1 Experimental Conclusions and their Context in the Field

6.1.1 Synthetic shRNAs are more potent silencing triggers than siRNAs

My initial studies on the biology of Dicer function suggest that shRNAs are more potent silencing triggers than siRNAs because they are processed by Dicer. shRNA enter the RNAi pathway at an earlier step than siRNAs. Previously, it was hypothesized that Dicer may be involved in loading its processed product into the RISC complex for silencing. In 2005, Shiekhattar and colleagues showed that indeed Dicer catalytic activity is coupled with Ago2 cleavage activity. They compared the RISC activity of the Dicer-TRBP-Ago1 complex using increasing concentrations of let-7 miRNA hairpin or a duplex siRNA which can no longer be cleaved by Dicer. The hairpin trigger generated more RISC activity when compared to the siRNA. They estimated programming RISC using a Dicer substrate is approximately 10 fold more efficient than a siRNA.¹⁷³ Further studies showed that Dicer is already associated with Ago2 prior to its encounter with processing substrates by participating in a RISC Loading Complex (RLC).¹⁷⁴ The main components of the RLC are Dicer, TRBP and Argonaute2. It has been reported that *in-vitro* reconstitution of the human RISC-loading complex does not display significant coupling between Dicer activity and active Slicer.¹⁷⁵ This may be due to the presence of other factors or modifications in RLC samples from human cell lines. The experiments above confirm our hypothesis that shRNAs are more potent RNAi triggers than siRNAs because they are loaded into the RISC complex by Dicer.

In addition, my biochemical studies detailed in Chapter 2 indicate that Dicer processes shRNA from the end of a hairpin. This knowledge allows us to determine the

precise small RNA sequence will be loaded into RISC. This is not only important for designing effective synthetic shRNA *in-vitro*, but also for shRNA therapeutic approaches. shRNA can be chemically modified for direct *in vivo* delivery into a mouse or human. These modifications can include changing the oligo backbone, replacement of individual nucleotides with nucleotide analogues, and adding conjugates to the oligo.¹⁷⁶ By chemically modifying shRNAs, they will have higher stability and serum resistance. However, synthetic shRNA still compete *in vivo* with endogenous miRNA pathways for enzymatic machinery. Studies have shown that exogenous RNAi triggers will compete for exportin 5 (for pri-miRNAs) and cytoplasmic Dicer (for pre-miRNAs).¹⁷⁷ This can be alleviated by identifying potent shRNAs that are capable of target down regulation when expressed at low levels.¹⁷⁸ The experiments in Chapter 2 illustrate how more potent triggers can be achieved. By knowing what small RNA sequence will be produced by Dicer, we can apply design algorithms to produce highly potent silencers with fewer off target effects. However, an efficient delivery system is still needed for protecting shRNA oligos from urinary excretion and RNase degradation, transporting them through the physical barriers to the target tissue, and enhancing their cellular uptake.¹⁷⁶

The experiments presented in Chapter 2 illustrate how biological studies of RNA interference are advancing its technological use. Previously, designing effective triggers were entirely empirical, but now mechanistic studies have guided the development of defining criteria by which small RNAs become incorporated into RISC. Our increased knowledge of Dicer processing and RISC activity are enabling us to make better shRNA triggers for studying gene function and therapeutic use.

6.1.2 Establishment of second-generation shRNA libraries covering the mouse and human genomes

The field of functional genomics has been accelerated due to the sequencing of the complete human genome in 2003. The work in Chapter 3 outlines our efforts to create a library of shRNA that mimic natural primary miRNA transcripts. Previously, the ability to screen for complex but specific phenotypes was limited to the realm of model organisms, but can now be realized for mammalian cells due to our construction of genome-wide shRNA libraries.¹¹⁵ The experiments in Chapter 3 demonstrate how the creation of human and mouse libraries allow researchers to silence genes in a sequence specific manner on a one-by-one basis or in a high throughput manner to study gene function. This library is unique in that both biological and technical strategies have been incorporated in its design. Because the processing sites on precursor shRNA-mirs have been mapped, siRNA design algorithms have been incorporated. This will help reduce off-target effects and increase potency. We moved these triggers into a primary miRNA transcript format and established they are more potent than our previous pre-miRNA hairpin library.⁵⁸

The second generation library also can be shuttled into other vectors for the creation of secondary libraries with tissue specific or inducible promoters. Different vector technologies, such as inducible vectors, can be quite useful. While version 1 vectors of the Hannon Elledge library take advantage of RNA polymerase III promoters to drive shRNA expression, work by Ross Dickins has showed that RNA polymerase II promoters are better for primary microRNA transcripts.⁴⁷ Endogenous primary transcripts are normally synthesized by RNA polymerase II.²⁰ Dickins showed that using functional

LTR promoters greatly enhance miR-30-based shRNA expression and target knockdown.

⁴⁷ This is another example of how our biological studies of RNAi have enhanced our technological tools. The MSCV-based vectors with functional LTRs are compatible with our library and our hairpins are routinely subcloned into these vectors. This library has been used by the scientific community to conduct screens in both *in vitro* and *in vivo* models and has also been used to explore individual gene function in mouse models.^{48, 106} As this enormous resource continues to grow, it holds much potential for illuminating genetic pathways and gene functions.

6.1.3 Barcode shRNA screens are tools for discovering gene function.

Chapter 4 details the technical experiments in microarray design and analysis that were necessary for the development of barcode screens. We improved the identification of barcode probes on microarrays and focused on negative selection screens, as these screens are more technically challenging. The use of negative selection screens with pooled hairpins illustrates a genuine advancement of RNAi screening technology as it has been suggested that only gene-by-gene screening approaches could identify such effects.⁴³ In addition, it has been suggested that these negative selection screens could not be done in pools of greater than 500 hairpins, which is a belief that we have since disproven. The methods developed in this Chapter have enabled researchers (Jose Silva and Ken Chang) to screen multiple breast cancer cell lines and identify hairpins that cause lethality.¹⁰⁶ In addition, these methods were used by Darren Burgess to conduct a microarray screen for mediators of doxorubicin resistance and sensitization.¹⁷⁹

We also tried to develop quality control methods that would allow standardization of RNAi screens. This is especially important if we want to increase our screen complexity and also compare results between multiple screens. Future work should control for extraneous variables and thereby standardize probe generation and hybridization methods even further. Adequate internal biological controls must also be used. For example, it may be beneficial to develop a set of hairpins that are known to cause selective advantage/disadvantage over time that could serve as internal controls.

While many reports on RNAi screens have emerged, much experimental variability exists, and few groups publish their methods in detail. Because of this, there is no standardized way to conduct screens and most groups focus on one or two robust hits, thus diluting the deep statistical power of these screens. This problem could be alleviated by creating a public repository for mammalian screens. This repository could hold protocols and raw or normalized data. These databases already exist for *C.elegans* (RNAiDB) and *Drosophila* (*Drosophila* RNAi screening center, Genome RNAi).¹⁵⁸

6.1.4. A synthetic lethal microarray screen identifies Chk1 as a modifier of doxorubicin sensitivity to p53 null cells

The work in Chapter 5 focuses on conducting a synthetic lethal screen and the accompanying data analysis and validation experiments. To avoid the enormous effort associated with validation, we used sophisticated analytical methods to reduce false positives. These methods, such as SAFE and our trend analysis can be improved with more screening data. Chanda and colleagues have devised a statistical analysis method that minimizes the impact of off-target activities.¹⁸⁰ The end product is a score that

models the probability of a gene 'hit' based on the collective activities of multiple siRNAs per gene.¹⁸⁰ This method gives priority to genes that have multiple hairpins that score with moderately strong phenotypes. While this statistical analysis is applied to multi-well screens, no group has published such a method for analysis of pooled barcode screens.

In addition, my screen identified a number of checkpoint related and DNA damage response genes. I reviewed the finding of Chk1 and its clinical implications in Chapter 5. New studies of Chk1 are elucidating its role in the cell cycle.¹⁸¹ Numerous groups are clarifying the roles of p53, MK2 and Chk1 both *in vivo* and *in vitro*.¹⁸²⁻¹⁸⁴ Hopefully, these studies will translate into changes for cancer patient treatment. Chk1 utilizes ATP for its enzymatic functions and is therefore readily accessible for small molecule inhibition at its catalytic site.¹⁶¹ However, there are few specific Chk1 inhibitors in clinical trials and none in combination with doxorubicin. This may be due to pharmacological properties of the chk1 inhibitors and doxorubicin. These problems may be relieved by using a shRNA as a drug (see below).

We are continuing our validation methods using both Solexa deep sequencing technology and microarrays to hone the list of candidates. Candidates are currently being put through growth curves in 96 well plates as a validation assay. It would be interesting if we could recapitulate these results *in vivo* by using the extensive RNAi mouse vectors developed in the Lowe lab.⁴⁸

6.2 Future Directions:

6.2.1 Expansion of the Hannon Elledge Library and our screening capabilities.

Two of the biggest challenges in using RNAi as a tool are off-target effects and variable silencing efficiency. While numerous design algorithms taking into account RNAi biology have attenuated off-target effects, much can be done to ensure we are using efficient hairpins. Experiments to validate efficient library knockdown are currently being conducted. One method is to create a sensor system where GFP is attached to a target sequence and knockdown levels are quantified through FACS analysis. This method can filter underperforming hairpins from the library and transform the library into a potent resource consisting of only high performance hairpins. This project would help decrease any false positives from screens. However, sometimes it is informative to study genes at a hypomorphic level rather than an absolute null. For example, different Trp53 shRNAs produced distinct phenotypes *in vivo*, ranging from benign lymphoid hyperplasias to highly disseminated lymphomas that paralleled Trp53^{-/-} lymphomagenesis in the E(mu)-Myc mouse.⁵¹ It would be interesting to see what hairpins could accelerate tumorigenesis in combination with weak performing p53 hairpins.

Jose Silva and Ken Chang have shown that they can conduct negative selection screens using pools of 20,000 hairpins.¹⁰⁶ Advances in microarray and deep sequencing technologies are allowing researchers to screen even more complex pools of shRNAs. There has been much discussion over when the first mammalian genome-wide screen will occur. The pooled hairpin approach to screening is very adaptable for this and less costly than the gene by gene multi-well approach. The limiting factor to such a screen seems to be tissue culture capabilities and data analysis techniques. While a genome-wide screen is very attractive, it would seem more informative and more practical to take a conservative approach to screening. There is still much to be learned by focusing on

functional sets of hairpins and developing well defined assays that have high clinical and biological impact. Indeed, smaller screens are still illuminating cancer specific genes and functions.¹²⁷ Much of the time devoted to screening occurs after the screen, during data analysis and downstream validation. Screening the entire genome for a simple phenotype will yield fruitful information, but screening 20,000 genes in an assay with high clinical impact will yield effects that may be directly applied to everyday patient treatment.

Data analysis is an area of RNAi screening that needs much improvement. While most researchers validate one or two robust hits from each screen, it would be more informative to see what patterns occur among all the candidates. While numerous databases exist for analyzing gene networks such as the Reactome, and Ingenuity Pathway Analysis, these are highly convoluted and are not specific for the kind of information one obtains from RNAi screens. We need defined statistical analyses that allow our candidates to be ranked not only by their significance and fold change, but also taking into account the number of hairpins per gene that scored in the screen and their knockdown efficiency.

6.2.2 Screens across multiple drugs and genotypes.

RNAi screens pose huge clinical implications for cancer therapy. We will now have the ability to understand what drugs will work well in certain genetic backgrounds. We can use this information to personalize a patient's therapy to the genetic makeup of their tumor. Chemotherapy sensitization/resistance screens and synthetic lethal screens have the ability to transform cancer therapy. Chemotherapy sensitization/resistance screens allow us to understand how our current chemotherapy works and also the genetic

modifiers that can change its efficacy. Synthetic lethal screens have the ability of identifying genetic lesions that can be new chemotherapeutic targets. The creation of a drug from a target is a long, arduous process that has a very low success rate. However, this may be alleviated if we can use small RNAs as drugs. (see below).

In order to harness the power of these screens, they need to be expanded across multiple drugs and cellular genotypes. I discussed briefly the prospect of a “Genetic Cancer Genome Project”¹²¹ in Chapter 4. This massive undertaking has the potential to revolutionize patient treatment. However, a project this size will require enormous financial resources and integration of efforts from numerous labs. A coordinated effort to find genes that modify sensitivity to chemotherapy will be expedited as a shared project between labs rather than sporadic screens by random groups. It will be important to study drugs that have been shown to be clinically effective in specific cancers (ex 5-FU in colon cancer) and also to study drugs that may be effective in a broad range of cancers (proteasome inhibitors, cdk inhibitors, etc).

The synthetic lethal screen outlined in Chapter 5 uses a pair of isogenic cell lines that provide a genotypically and phenotypically well characterized model. These are derived from human colon carcinoma. It may be more desirable to compare cell lines of non-tumor origin or primary cells with cancer cells for synthetic lethal screens. Jose Silva in our lab has done studies searching for genes that are lethal to numerous breast cancer cell lines and a human breast epithelial cell line that retains many characteristics of normal breast epithelium (MCF10A).¹⁰⁶

Another possibility for RNAi screens is to take advantage of rapidly expanding biotechnology. While this thesis has focused on microarrays as a means to assay a

barcode screen, deep sequencing may provide an alternative to hybridization that may expand our detection capabilities and sensitivity. In addition, FACS sorting may be used in a negative selection screen. After chemotherapeutic treatment, dying cells can be sorted for using a caspase-3 fluorescent substrate¹⁸⁵ as an early apoptotic marker and the DNA barcodes can be quantified using Solexa sequencing or microarray hybridization. This is a way of using a positive assay in a negative selection screen.

While this thesis focuses on performing RNAi screens *in vitro*, entire shRNA screens can be performed *in vivo*. Mouse screens have focused on positive selections with care being taken into the number of hairpins used in a pool.⁵⁰ Multiple screens are being pursued in the lab of Scott Lowe. One method involves infecting tumor cells *in vitro*, transplanting them into mice and measuring shRNA changes after chemotherapy treatment.⁵⁰ Alternatively, embryonic stem cells can be infected *in vitro* and allowed to form tumors *in vivo*.¹⁸⁶ Hairpins that increase during tumorigenesis will be present in the resulting tumor. These screens take into account tumor-microenvironment and immune system interactions which are not readily assessable in *in vitro* screens.

6.2.3 RNAi therapeutics

RNAi is an interesting phenomenon because not only can it be used as a tool for research, but it has potential for clinical use as well. It is my hope that one day routine RNAi screens will be used to find patient specific target genes, which subsequently can be silenced with an RNAi drug. Our current drug development pipeline is cumbersome. Each new compound must be tested thoroughly and driven through the drug development process. By contrast, the problems with making small RNAs into drugs are mostly

universal. If we can solve those few problems we will be able to make not just one drug, but many drugs. Our current concerns regarding small RNAs as drugs involves saturation of the cellular RNAi machinery, toxicity phenomena in liver and kidney, and delivery to the target organ/site. In addition, no mammalian RNA dependent RNA polymerase has been found and thus shRNA effects are transient. A study by Zimmerman et al noted that knockdown in non-human primates lasted 11 days.¹⁸⁷ However, this would not be suitable for chronic disease treatment. Besides *in vivo* delivery of shRNA and siRNA, there has been work associated with adenoviral and adeno-associated viral delivery of hairpins.¹⁷⁸ These also have a long way to go before they can be applied as therapy. In addition, there have been studies trying to suppress miRNA function, particularly in the context of cancer, but they are still in their early stages.⁶⁷

Much work is needed to conquer the problems still associated with developing RNAi therapeutics. However, research in both basic and translational RNAi is progressing rapidly and synergistically.⁶⁶ Hopefully, one day RNAi Clinics will emerge and will be based on research products taken directly from the bench.

6.3 Summary

The field of RNAi has advanced considerably in the last five years. Different kinds of small RNAs have been associated with gene regulation. The catalytic machinery of RNAi has been identified and characterized in depth. In addition, large scale RNAi libraries have been created and subsequently both positive and negative selection screens are now almost routine.

The work in this thesis outlines studies of RNAi biology and its technological applications. Chapter 2 describes how the biology of Dicer processing shRNA may be connected to making potent silencing triggers. Because of this work, we can predict what small RNA will enter the RISC complex and subsequently apply silencing algorithms to these hairpins to make effective silencers with minimal off-target effects. We then combined silencing algorithms into a mir30 primary transcript format to make a second generation shRNA library targeting the mouse and human genomes. We then show how this library can be used en masse to conduct negative selection RNAi screens. These screens use microarray technology and data analysis methods described in Chapter 4. Finally, we apply these methods to conduct a synthetic lethal screen for genes that modify sensitivity to doxorubicin in a p53 null background. While this screen requires deeper examination of the identified targets, the foundation has been laid for improving our chemotherapeutic approaches. Hopefully one day, screens such as these will have real and tangible clinical impact and RNAi will not be restricted to merely a laboratory tool and biological phenomenon, but a clinically effective therapeutic tool as well.

Literature Cited

1. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* 100, 57-70 (2000).
2. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nat Med* 10, 789-99 (2004).
3. Papac, R. J. Origins of cancer therapy. *Yale J Biol Med* 74, 391-8 (2001).
4. Frei, E., 3rd et al. The effectiveness of combinations of antileukemic agents in inducing and maintaining remission in children with acute leukemia. *Blood* 26, 642-56 (1965).
5. Chabner, B. A. & Roberts, T. G., Jr. Timeline: Chemotherapy and the war on cancer. *Nat Rev Cancer* 5, 65-72 (2005).
6. Sawyers, C. Targeted cancer therapy. *Nature* 432, 294-7 (2004).
7. Roche-Lestienne, C. et al. Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. *Blood* 100, 1014-1018 (2002).
8. Shah, N. P. et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* 2, 117-25 (2002).
9. Sordella, R., Bell, D. W., Haber, D. A. & Settleman, J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* 305, 1163-7 (2004).
10. van 't Veer, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530-6 (2002).
11. Agendia, I.
12. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-11 (1998).
13. Hannon, G. J. RNA interference. *Nature* 418, 244-51 (2002).
14. Birchler, J. A. & Kavi, H. H. Molecular biology. Slicing and dicing for small RNAs. *Science* 320, 1023-4 (2008).
15. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363-6 (2001).

16. Carmell, M. A. & Hannon, G. J. RNase III enzymes and the initiation of gene silencing. *Nat Struct Mol Biol* 11, 214-8 (2004).
17. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293-6 (2000).
18. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843-54 (1993).
19. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858-62 (2001).
20. Lee, Y. et al. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* 23, 4051-60 (2004).
21. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* 431, 343-9 (2004).
22. Han, J. et al. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18, 3016-27 (2004).
23. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415-9 (2003).
24. Song, J. J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* 10, 1026-32 (2003).
25. He, L. & Hannon, G. J. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 5, 522-31 (2004).
26. Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* 7, 719-23 (2005).
27. Liu, J. et al. Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science* 305, 1437-1441 (2004).
28. Tolia, N. H. & Joshua-Tor, L. Slicer and the Argonautes. *Nat Chem Biol* 3, 36-43 (2007).
29. Song, J.-J. & Joshua-Tor, L. Argonaute and RNA -- getting into the groove. *Current Opinion in Structural Biology* 16, 5-11 (2006).

30. Faehnle, C. R. & Joshua-Tor, L. Argonautes confront new small RNAs. *Current Opinion in Chemical Biology* 11, 569-577 (2007).
31. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 9, 22-32 (2008).
32. Parker, J. S. & Barford, D. Argonaute: a scaffold for the function of short regulatory RNAs. *Trends in Biochemical Sciences* 31, 622-630 (2006).
33. Davis, S. & Watson, J. C. In vitro activation of the interferon-induced, double-stranded RNA-dependent protein kinase PKR by RNA from the 3' untranslated regions of human alpha-tropomyosin. *Proc Natl Acad Sci U S A* 93, 508-13 (1996).
34. Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H. & Williams, B. R. Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* 5, 834-9 (2003).
35. Elbashir, S. M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411, 494-8 (2001).
36. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. & Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 16, 948-58 (2002).
37. Sarov, M. & Stewart, A. F. The best control for the specificity of RNAi. *Trends Biotechnol* 23, 446-8 (2005).
38. Paddison, P. J., Caudy, A. A. & Hannon, G. J. Stable suppression of gene expression by RNAi in mammalian cells. *Proc Natl Acad Sci U S A* 99, 1443-8 (2002).
39. Pei, Y. & Tuschl, T. On the art of identifying effective and specific siRNAs. *Nat Meth* 3, 670-676 (2006).
40. Jackson, A. L. et al. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotech* 21, 635-637 (2003).
41. Barik, S. RNAi in moderation. *Nat Biotech* 24, 796-797 (2006).
42. Wolters, N. M. & MacKeigan, J. P. From sequence to function: using RNAi to elucidate mechanisms of human disease. *Cell Death Differ* 15, 809-19 (2008).
43. Bernards, R., Brummelkamp, T. R. & Beijersbergen, R. L. shRNA libraries and their use in cancer genetics. *Nat Methods* 3, 701-6 (2006).
44. Chang, K., Elledge, S. J. & Hannon, G. J. Lessons from Nature: microRNA-based shRNA libraries. *Nat Methods* 3, 707-14 (2006).

45. Root, D. E., Hacohen, N., Hahn, W. C., Lander, E. S. & Sabatini, D. M. Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Meth* 3, 715-719 (2006).
46. Silva J, C. K., Hannon GJ, Rivas FV. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. *Oncogene* 51, 8401-9. (2004).
47. Dickins, R. A. et al. Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nat Genet* 37, 1289-95 (2005).
48. Dickins, R. A. et al. Tissue-specific and reversible RNA interference in transgenic mice. *Nat Genet* 39, 914-21 (2007).
49. Silva, J., Chang, K., Hannon, G. J. & Rivas, F. V. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. *Oncogene* 23, 8401-9 (2004).
50. Burgess, D. J. RNAi Genetic Screens for Chemotherapy Response Modifier Genes. (PhD dissertation, Watson School of Biological Sciences) (2007).
51. Hemann, M. T. et al. An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nat Genet* 33, 396-400 (2003).
52. Kaelin, W. G., Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer* 5, 689-98 (2005).
53. Aza-Blanc, P. et al. Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. *Mol Cell* 12, 627-37 (2003).
54. Westbrook, T. F. et al. A genetic screen for candidate tumor suppressors identifies REST. *Cell* 121, 837-48 (2005).
55. Kolfschoten, I. G. et al. A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. *Cell* 121, 849-58 (2005).
56. MacKeigan, J. P., Murphy, L. O. & Blenis, J. Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. *Nat Cell Biol* 7, 591-600 (2005).
57. Berns, K. et al. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* 428, 431-7 (2004).
58. Paddison, P. J. et al. A resource for large-scale RNA-interference-based screens in mammals. *Nature* 428, 427-31 (2004).

59. Brummelkamp, T. R., Nijman, S. M., Dirac, A. M. & Bernards, R. Loss of the cylindromatosis tumour suppressor inhibits apoptosis by activating NF-kappaB. *Nature* 424, 797-801 (2003).
60. Ngo, V. N. et al. A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* 441, 106-110 (2006).
61. Whitehurst, A. W. et al. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature* 446, 815-9 (2007).
62. Berns, K. et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 12, 395-402 (2007).
63. Soutschek, J. et al. Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* 432, 173-8 (2004).
64. Bitko, V., Musiyenko, A., Shulyayeva, O. & Barik, S. Inhibition of respiratory viruses by nasally administered siRNA. *Nat Med* 11, 50-5 (2005).
65. de Fougerolles, A., Vornlocher, H. P., Maraganore, J. & Lieberman, J. Interfering with disease: a progress report on siRNA-based therapeutics. *Nat Rev Drug Discov* 6, 443-53 (2007).
66. Grimm, D. & Kay, M. A. Therapeutic application of RNAi: is mRNA targeting finally ready for prime time? *J Clin Invest* 117, 3633-41 (2007).
67. Tara, M. L., Howell, F. M. & Carl, D. N. Not miR-ly small RNAs: Big potential for microRNAs in therapy. *The Journal of allergy and clinical immunology* 121, 309-319 (2008).
68. Krutzfeldt, J. et al. Silencing of microRNAs in vivo with 'antagomirs'. *Nature* 438, 685-689 (2005).
69. He, L. et al. A microRNA component of the p53 tumour suppressor network. *Nature* 447, 1130-1134 (2007).
70. He, L. et al. A microRNA polycistron as a potential human oncogene. *Nature* 435, 828-833 (2005).
71. Bignell, G. R. et al. Identification of the familial cylindromatosis tumour-suppressor gene. *Nat Genet* 25, 160-165 (2000).
72. Lakhani, S. R. Putting the Brakes on Cylindromatosis? *N Engl J Med* 350, 187-188 (2004).
73. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs Exhibit Strand Bias. *Cell* 115, 209-216 (2003).

74. Carmell, M. A., Xuan, Z., Zhang, M. Q. & Hannon, G. J. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* 16, 2733-42 (2002).
75. Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear export of microRNA precursors. *Science* 303, 95-8 (2004).
76. Nakayashiki, H. et al. RNA silencing as a tool for exploring gene function in ascomycete fungi. *Fungal Genet Biol* 42, 275-83 (2005).
77. Tang, J., Erikson, R. L. & Liu, X. Checkpoint kinase 1 (Chk1) is required for mitotic progression through negative regulation of polo-like kinase 1 (Plk1). *Proc Natl Acad Sci U S A* 103, 11964-9 (2006).
78. Dasgupta, R. & Perrimon, N. Using RNAi to catch Drosophila genes in a web of interactions: insights into cancer research. *Oncogene* 23, 8359-65 (2004).
79. Fraser, A. Towards full employment: using RNAi to find roles for the redundant. *Oncogene* 23, 8346-52 (2004).
80. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281-97 (2004).
81. Reinhart, B. J. et al. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901-6. (2000).
82. Ketting, R. F. et al. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* 15, 2654-9. (2001).
83. Grishok, A. et al. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106, 23-34. (2001).
84. Knight, S. W. & Bass, B. L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* 293, 2269-71. (2001).
85. Hutvagner, G. et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834-8. (2001).
86. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231-5 (2004).
87. Landthaler, M., Yalcin, A. & Tuschl, T. The Human DiGeorge Syndrome Critical Region Gene 8 and Its *D. melanogaster* Homolog Are Required for miRNA Biogenesis. *Curr Biol* 14, 2162-7 (2004).

88. Gregory, R. I. et al. The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235-40 (2004).
89. Siolas, D. et al. Synthetic shRNAs as potent RNAi triggers. *Nat Biotechnol* 23, 227-31 (2005).
90. Schwarz, D. S. et al. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199-208 (2003).
91. Zeng, Y., Wagner, E. J. & Cullen, B. R. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* 9, 1327-33. (2002).
92. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. & Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 16, 948-58. (2002).
93. Chen, C. Z., Li, L., Lodish, H. F. & Bartel, D. P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* 303, 83-6 (2004).
94. Zeng, Y. & Cullen, B. R. Sequence requirements for micro RNA processing and function in human cells. *Rna* 9, 112-23 (2003).
95. Kawasaki, H. & Taira, K. Short hairpin type of dsRNAs that are controlled by tRNA(Val) promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells. *Nucleic Acids Res* 31, 700-7. (2003).
96. Brummelkamp, T. R., Bernards, R. & Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296, 550-553 (2002).
97. Zheng, L. et al. An approach to genomewide screens of expressed small interfering RNAs in mammalian cells. *Proc Natl Acad Sci U S A* 101, 135-40 (2004).
98. Stegmeier, F., Hu, G., Rickles, R. J., Hannon, G. J. & Elledge, S. J. A Lentiviral microRNA-based System for Single Copy Pol II Regulated RNAi in mammalian Cells. *Proc Natl Acad Sci U S A*, in press (2005).
99. Li, M. Z. & Elledge, S. J. MAGIC, an in vivo genetic method for the rapid construction of recombinant DNA molecules. *Nat Genet* 37, 311-9 (2005).
100. Li, X. et al. Generation of destabilized green fluorescent protein as a transcription reporter. *J Biol Chem* 273, 34970-5 (1998).
101. Elledge, S. J. & Walker, G. C. Phasmid vectors for identification of genes by complementation of *Escherichia coli* mutants. *J Bacteriol* 162, 777-83 (1985).

102. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97, 6640-5 (2000).
103. Cherepanov, P. P. & Wackernagel, W. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* 158, 9-14 (1995).
104. Chalker, A. F., Leach, D. R. & Lloyd, R. G. *Escherichia coli* sbcC mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* 71, 201-5 (1988).
105. Caudy, A. A., Myers, M., Hannon, G. J. & Hammond, S. M. Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev* 16, 2491-6. (2002).
106. Silva, J. M. et al. Profiling essential genes in human mammary cells by multiplex RNAi screening. *Science* 319, 617-20 (2008).
107. Lucito, R. a. W., M. Preparation of Slides and Hybridization. (ed. Sambrook), e. D. B. a. J.) (Cold Spring Harbor Press, Cold Spring Harbor, NY., 2003).
108. website, A.
110. Ambion, I. in *Amino Alkyl MessageAmp Instruction Manual*.
111. Ramdas, L. et al. Sources of nonlinearity in cDNA microarray expression measurements. *Genome Biol* 2, RESEARCH0047 (2001).
112. Wall, M. E., Andreas Rechtsteiner, Rocha, LM. *Singular value decomposition and principal component analysis*". (Kluwer, Norwell, MA 2003).
113. Ringner, M. What is principal component analysis? *Nat Biotech* 26, 303-304 (2008).
114. Alter, O., Brown, P. O. & Botstein, D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A* 97, 10101-6 (2000).
115. Haney, S. A. Increasing the robustness and validity of RNAi screens. *Pharmacogenomics* 8, 1037-49 (2007).
116. Jolliffe, I. *Principal Component Analysis* (Springer, New York, 2002).
117. Moore, D. *The Basic Practice of Statistics (Paper) & Student CD*. (W. H. Freeman, San Francisco, 2006).
118. Roberts, F. a. R., D. (2008).

119. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98, 5116-21 (2001).
120. Barry, W. T., Nobel, A. B. & Wright, F. A. Significance analysis of functional categories in gene expression studies: a structured permutation approach. *Bioinformatics* 21, 1943-9 (2005).
121. Elledge, S. J. & Hannon, G. J. An open letter to cancer researchers. *Science* 310, 439-41 (2005).
122. Lucito, R. et al. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res* 13, 2291-305 (2003).
123. Chabner, B. A. & Roberts, T. G. CHEMOTHERAPY AND THE WAR ON CANCER. *Nature Reviews Cancer*
Nat Rev Cancer 5, 65-72 (2005).
124. Weinstein, I. B. CANCER: Enhanced: Addiction to Oncogenes--the Achilles Heel of Cancer. *Science* 297, 63-64 (2002).
125. Hübner, A., Jaeschke, A. & Davis, R. J. Oncogene Addiction: Role of Signal Attenuation. *Developmental Cell* 11, 752-754 (2006).
126. Sawyers, C. L. Making Progress through Molecular Attacks on Cancer. *Cold Spring Harbor Symposia on Quantitative Biology* 70, 479-482 (2005).
127. Shaffer, A. L. et al. IRF4 addiction in multiple myeloma. *Nature* 454, 226-231 (2008).
128. Willingham, A. T., Deveraux, Q. L., Hampton, G. M. & Aza-Blanc, P. RNAi and HTS: exploring cancer by systematic loss-of-function. *Oncogene* 23, 8392-400 (2004).
129. Lucchesi, J. C. SYNTHETIC LETHALITY AND SEMI-LETHALITY AMONG FUNCTIONALLY RELATED MUTANTS OF DROSOPHILA MELANOGASTER. *Genetics* 59, 37-44 (1968).
130. Hartwell, L. H., Szankasi, P., Roberts, C. J., Murray, A. W. & Friend, S. H. Integrating Genetic Approaches into the Discovery of Anticancer Drugs. *Science* 278, 1064-1068 (1997).
131. Iorns, E., Lord, C. J., Turner, N. & Ashworth, A. Utilizing RNA interference to enhance cancer drug discovery. *Nat Rev Drug Discov* 6, 556-68 (2007).
132. Xue, W. et al. Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature* 445, 656-60 (2007).

133. Riley, T., Sontag, E., Chen, P. & Levine, A. Transcriptional control of human p53-regulated genes. *Nat Rev Mol Cell Biol* 9, 402-12 (2008).
134. Hirao, A. et al. DNA damage-induced activation of p53 by the checkpoint kinase Chk2. *Science* 287, 1824-7 (2000).
135. Gasco, M., Shami, S. & Crook, T. The p53 pathway in breast cancer. *Breast Cancer Res* 4, 70 - 76 (2002).
136. Miller, L. D. et al. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 102, 13550-13555 (2005).
137. Bunz, F. et al. Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science* 282, 1497-501 (1998).
138. Torrance, C. J., Agrawal, V., Vogelstein, B. & Kinzler, K. W. Use of isogenic human cancer cells for high-throughput screening and drug discovery. *Nat Biotechnol* 19, 940-5 (2001).
139. Cohen, P. Protein kinases--the major drug targets of the twenty-first century? *Nat Rev Drug Discov* 1, 309-15 (2002).
140. Froelich-Ammon, S. J. & Osheroff, N. Topoisomerase poisons: harnessing the dark side of enzyme mechanism. *J Biol Chem* 270, 21429-32 (1995).
141. Health, U. N. I. o. in www.clinicaltrials.gov (2008).
142. Kumar, N. Taxol-induced polymerization of purified tubulin. Mechanism of action. *J. Biol. Chem.* 256, 10435-10441 (1981).
143. Smyth, G. K. & Speed, T. Normalization of cDNA microarray data. *Methods* 31, 265-73 (2003).
144. Quackenbush, J. Microarray data normalization and transformation. *Nat Genet* 32 Suppl, 496-501 (2002).
145. Storey, J. D., Xiao, W., Leek, J. T., Tompkins, R. G. & Davis, R. W. Significance analysis of time course microarray experiments. *Proc Natl Acad Sci U S A* 102, 12837-42 (2005).
146. Leek, J. T., Monsen, E., Dabney, A. R. & Storey, J. D. EDGE: extraction and analysis of differential gene expression. *Bioinformatics* 22, 507-8 (2006).
147. Storey, J. D., Dai, J. Y. & Leek, J. T. The optimal discovery procedure for large-scale significance testing, with applications to comparative microarray experiments. *Biostatistics* 8, 414-32 (2007).

148. Pichler, A., Zelcer, N., Prior, J. L., Kuil, A. J. & Piwnica-Worms, D. In vivo RNA interference-mediated ablation of MDR1 P-glycoprotein. *Clin Cancer Res* 11, 4487-94 (2005).
149. Reinhardt, H. C., Aslanian, A. S., Lees, J. A. & Yaffe, M. B. p53-deficient cells rely on ATM- and ATR-mediated checkpoint signaling through the p38MAPK/MK2 pathway for survival after DNA damage. *Cancer Cell* 11, 175-89 (2007).
150. Pimienta, G. & Pascual, J. Canonical and alternative MAPK signaling. *Cell Cycle* 6, 2628-32 (2007).
151. Chen, Y. & Sanchez, Y. Chk1 in the DNA damage response: conserved roles from yeasts to mammals. *DNA Repair* 3, 1025-1032 (2004).
152. Vousden, K. H. & Lu, X. Live or let die: the cell's response to p53. *Nat Rev Cancer* 2, 594-604 (2002).
153. Erenpreisa, J., Kalejs, M. & Cragg, M. S. Mitotic catastrophe and endomitosis in tumour cells: an evolutionary key to a molecular solution. *Cell Biol Int* 29, 1012-8 (2005).
154. Chen, Z. et al. Selective Chk1 inhibitors differentially sensitize p53-deficient cancer cells to cancer therapeutics. *Int J Cancer* 119, 2784-94 (2006).
155. Carrassa, L., Broggin, M., Erba, E. & Damia, G. Chk1, but not Chk2, is involved in the cellular response to DNA damaging agents: differential activity in cells expressing or not p53. *Cell Cycle* 3, 1177-81 (2004).
156. Lamprecht, M. R., Sabatini, D. M. & Carpenter, A. E. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques* 42, 71-5 (2007).
157. Huesken, D. et al. Design of a genome-wide siRNA library using an artificial neural network. *Nat Biotechnol* 23, 995-1001 (2005).
158. Boutros, M. & Ahringer, J. The art and design of genetic screens: RNA interference. *Nat Rev Genet* 9, 554-66 (2008).
159. NCI Drug Dictionary Definition of 7-hydroxystaurosporine.
<http://www.cancer.gov/Templates/drugdictionary.aspx?CdrID=42305>
160. Mukhopadhyay, U. K., Senderowicz, A. M. & Ferbeyre, G. RNA silencing of checkpoint regulators sensitizes p53-defective prostate cancer cells to chemotherapy while sparing normal cells. *Cancer Res* 65, 2872-81 (2005).
161. O'Connor, M. J., Martin, N. M. & Smith, G. C. Targeted cancer therapies based on the inhibition of DNA strand break repair. *Oncogene* 26, 7816-24 (2007).

162. US National Institute of Health. www.clinicaltrials.gov, search terms: AZD7762, PF477736, XL844, LY2603618 viewed July 28, 2008
163. Zhang, Y., Zhou, J. & Lim, C. U. K. The role of NBS1 in DNA double strand break repair, telomere stability, and cell cycle checkpoint control. *Cell Res* 16, 45-54.
164. Kondratenko, I., Paschenko, O., Polyakov, A. & Bologov, A. Nijmegen breakage syndrome. *Adv Exp Med Biol* 601, 61-7 (2007).
165. Lee, J. H. & Paull, T. T. Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. *Oncogene* 26, 7741-7748.
166. Gatei, M. et al. Ataxia-telangiectasia-mutated (ATM) and NBS1-dependent Phosphorylation of Chk1 on Ser-317 in Response to Ionizing Radiation. *J. Biol. Chem.* 278, 14806-14811 (2003).
167. Tran, H. M. et al. Mutant Nbs1 enhances cisplatin-induced DNA damage and cytotoxicity in head and neck cancer. *Otolaryngol Head Neck Surg* 131, 477-84 (2004).
168. Kumagai, A., Kim, S. M. & Dunphy, W. G. Claspin and the activated form of ATR-ATRIP collaborate in the activation of Chk1. *J Biol Chem* 279, 49599-608 (2004).
169. Yang, X. H., Shiotani, B., Classon, M. & Zou, L. Chk1 and Claspin potentiate PCNA ubiquitination. *Genes Dev* 22, 1147-52 (2008).
170. Verlinden, L. et al. The E2F-regulated gene Chk1 is highly expressed in triple-negative estrogen receptor /progesterone receptor /HER-2 breast carcinomas. *Cancer Res* 67, 6574-81 (2007).
171. Hannay, J. A. F. et al. Rad51 overexpression contributes to chemoresistance in human soft tissue sarcoma cells: a role for p53/activator protein 2 transcriptional regulation. *Mol Cancer Ther* 6, 1650-1660 (2007).
172. Koehn, H., Magan, N., Isaacs, R. J. & Stowell, K. M. Differential regulation of DNA repair protein Rad51 in human tumour cell lines exposed to doxorubicin. *Anticancer Drugs* 18, 419-25 (2007).
173. Gregory, R. I., Chendrimada, T. P., Cooch, N. & Shiekhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* 123, 631-40 (2005).
174. Maniataki, E. & Mourelatos, Z. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev* 19, 2979-90 (2005).

175. MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V. & Doudna, J. A. In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci U S A* 105, 512-7 (2008).
176. Lu, P. Y., Xie, F. & Woodle, M. C. In vivo application of RNA interference: from functional genomics to therapeutics. *Adv Genet* 54, 117-42 (2005).
177. Castanotto, D. et al. Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. *Nucleic Acids Res* 35, 5154-64 (2007).
178. Rossi, J. J. Expression strategies for short hairpin RNA interference triggers. *Hum Gene Ther* 19, 313-7 (2008).
179. Burgess, D. J. et al. Topoisomerase levels determine chemotherapy response in vitro and in vivo. *Proc Natl Acad Sci U S A* 105, 9053-8 (2008).
180. Konig, R. et al. A probability-based approach for the analysis of large-scale RNAi screens. *Nat Methods* 4, 847-9 (2007).
181. Jin, J., Ang, X. L., Ye, X., Livingstone, M. & Harper, J. W. Differential Roles for Checkpoint Kinases in DNA Damage-dependent Degradation of the Cdc25A Protein Phosphatase. *J. Biol. Chem.* 283, 19322-19328 (2008).
182. Levesque, A. A., Fanous, A. A., Poh, A. & Eastman, A. Defective p53 signaling in p53 wild-type tumors attenuates p21/waf1 induction and cyclin B repression rendering them sensitive to Chk1 inhibitors that abrogate DNA damage-induced S and G2 arrest. *Mol Cancer Ther* 7, 252-62 (2008).
183. Vitale, I. et al. Inhibition of Chk1 kills tetraploid tumor cells through a p53-dependent pathway. *PLoS ONE* 2, e1337 (2007).
184. Vitale, I. et al. Chk1 inhibition activates p53 through p38 MAPK in tetraploid cancer cells. *Cell Cycle* 7, 1956-61 (2008).
185. Jixiang, L. et al. Fluorescent molecular probes V: A sensitive caspase-3 substrate for fluorometric assays. *Bioorganic & Medicinal Chemistry Letters* 9, 3231-3236 (1999).
186. Hemann, M. T. et al. Suppression of tumorigenesis by the p53 target PUMA. *Proc Natl Acad Sci U S A* 101, 9333-8 (2004).
187. Zimmermann, T. S. et al. RNAi-mediated gene silencing in non-human primates. *Nature* 441, 111-4 (2006).

Bibliography

1. Aas, T. et al. Specific P53 mutations are associated with de novo resistance to doxorubicin in breast cancer patients. *Nat Med* 2, 811 - 814 (1996).
2. Adams, J. The development of proteasome inhibitors as anticancer drugs. *Cancer Cell* 5, 417-421 (2004).
3. Adams, J. THE PROTEASOME: A SUITABLE ANTINEOPLASTIC TARGET. *Nat Rev Cancer* 4, 349-360 (2004).
4. Adams, J. et al. Proteasome inhibitors: a novel class of potent and effective antitumor agents. *Cancer Res* 59, 2615-22 (1999).
5. Agendia, Inc. Mammaprint Brochure
http://usa.agendia.com/en/for_healthcare_professionals_2.html
6. Akhtar, S. & Benter, I. F. Nonviral delivery of synthetic siRNAs in vivo. *J Clin Invest* 117, 3623-32 (2007).
7. Allinen, M., Huusko, P., Mantyniemi, S., Launonen, V. & Wingvist, R. Mutation analysis of the CHK2 gene in families with hereditary breast cancer. *Br J Cancer* 85, 209 - 212 (2001).
8. Alsner, J., Yilmaz, M., Guldborg, P., Hansen, L. L. & Overgaard, J. Heterogeneity in the clinical phenotype of TP53 mutations in breast cancer. *Clin Cancer Res* 6, 3923 - 3931 (2000).
9. Alter, O., Brown, P. O. & Botstein, D. Singular value decomposition for genome-wide expression data processing and modeling. *Proc Natl Acad Sci U S A* 97, 10101-6 (2000).
10. Ambion website, Amino Alkyl Instruction manual
<http://www.ambion.com/techlib/resources/arrays/labeling.html> viewed July 20, 2008. (2008).
11. Angele, S. et al. Abnormal expression of the ATM and TP53 genes in sporadic breast cancer. *Clin Cancer Res* 6, 3536 - 3544 (2000).
12. Aza-Blanc, P. et al. Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. *Mol Cell* 12, 627-37 (2003).
13. Barbareschi, M. et al. p63, a p53 homologue is a selective marker of myoepithelial cells of the human breast. *Am J Surg Pathol* 25, 1054 - 1060 (2001).
14. Barik, S. RNAi in moderation. *Nat Biotech* 24, 796-797 (2006).

15. Barry, W. T., Nobel, A. B. & Wright, F. A. Significance analysis of functional categories in gene expression studies: a structured permutation approach. *Bioinformatics* 21, 1943-9 (2005).
17. Bartek, J., Iggo, R., Gannon, J. & Lane, D. P. Genetic and immunocytochemical analysis of mutant p53 in human breast cancer. *Oncogene* 5, 893 - 899 (1990).
18. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281-97 (2004).
19. Bell, D. W. et al. Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome. *Science* 286, 2528 - 2531 (1999).
21. Bernards, R., Brummelkamp, T. R. & Beijersbergen, R. L. shRNA libraries and their use in cancer genetics. *Nat Methods* 3, 701-6 (2006).
22. Berns, K. et al. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* 428, 431-7 (2004).
23. Berns, K. et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 12, 395-402 (2007).
24. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363-6 (2001).
25. Bernstein, E., Denli, A. M. & Hannon, G. J. The rest is silence. *Rna* 7, 1509-21 (2001).
26. Bernstein, E. et al. Dicer is essential for mouse development. *Nat Genet* 35, 215-7 (2003).
27. Bignell, G. R. et al. Identification of the familial cylindromatosis tumour-suppressor gene. *Nat Genet* 25, 160-165 (2000).
28. Birchler, J. A. & Kavi, H. H. Molecular biology. Slicing and dicing for small RNAs. *Science* 320, 1023-4 (2008).
29. Bitko, V., Musiyenko, A., Shulyayeva, O. & Barik, S. Inhibition of respiratory viruses by nasally administered siRNA. *Nat Med* 11, 50-5 (2005).
30. Boutros, M. & Ahringer, J. The art and design of genetic screens: RNA interference. *Nat Rev Genet* 9, 554-66 (2008).

31. Brummelkamp, T. R., Bernards, R. & Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* 296, 550-553 (2002).
32. Brummelkamp, T. R., Nijman, S. M., Dirac, A. M. & Bernards, R. Loss of the cylindromatosis tumour suppressor inhibits apoptosis by activating NF-kappaB. *Nature* 424, 797-801 (2003).
33. Bunz, F. et al. Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science* 282, 1497-501 (1998).
34. Burgess, D. J. RNAi Genetic Screens for Chemotherapy Response Modifier Genes. (PhD dissertation, Watson School of Biological Sciences) (2007).
35. Burgess, D. J. et al. Topoisomerase levels determine chemotherapy response in vitro and in vivo. *Proc Natl Acad Sci U S A* 105, 9053-8 (2008).
36. Caplen, N. J., Parrish, S., Imani, F., Fire, A. & Morgan, R. A. Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc. Natl Acad. Sci. USA*. 98, 9742-9747 (2001).
37. Carmell, M. A. & Hannon, G. J. RNase III enzymes and the initiation of gene silencing. *Nat Struct Mol Biol* 11, 214-8 (2004).
38. Carmell, M. A. & Hannon, G. J. RNase III enzymes and the initiation of gene silencing. *Nat Struct Mol Biol* 11, 214-218 (2004).
39. Carmell, M. A., Xuan, Z., Zhang, M. Q. & Hannon, G. J. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* 16, 2733-42 (2002).
40. Carrassa, L., Broggini, M., Erba, E. & Damia, G. Chk1, but not Chk2, is involved in the cellular response to DNA damaging agents: differential activity in cells expressing or not p53. *Cell Cycle* 3, 1177-81 (2004).
41. Castanotto, D. et al. Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. *Nucleic Acids Res* 35, 5154-64 (2007).
42. Catteau, A., Harris, W. H., Xu, C. F. & Solomon, E. Methylation of the BRCA1 promoter region in sporadic breast and ovarian cancer: correlation with disease characteristics. *Oncogene* 18, 1957 - 1965 (1999).
43. Caudy, A. A., Myers, M., Hannon, G. J. & Hammond, S. M. Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev* 16, 2491-6 (2002).

44. Chabner, B. A. & Roberts, T. G., Jr. Timeline: Chemotherapy and the war on cancer. *Nat Rev Cancer* 5, 65-72 (2005).
45. Chabner, B. A. & Roberts, T. G. CHEMOTHERAPY AND THE WAR ON CANCER. *Nat Rev Cancer* 5, 65-72 (2005).
46. Chalker, A. F., Leach, D. R. & Lloyd, R. G. Escherichia coli sbcC mutants permit stable propagation of DNA replicons containing a long palindrome. *Gene* 71, 201-5 (1988).
47. Chang, K., Elledge, S. J. & Hannon, G. J. Lessons from Nature: microRNA-based shRNA libraries. *Nat Methods* 3, 707-14 (2006).
48. Chauhan, D. et al. Blockade of ubiquitin-conjugating enzyme CDC34 enhances anti-myeloma activity of Bortezomib/Proteasome inhibitor PS-341. *Oncogene* 23, 3597-602 (2004).
49. Chauhan, D. et al. Blockade of Hsp27 overcomes Bortezomib/proteasome inhibitor PS-341 resistance in lymphoma cells. *Cancer Res* 63, 6174-7 (2003).
50. Chen, C. Z., Li, L., Lodish, H. F. & Bartel, D. P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* 303, 83-6 (2004).
51. Chen, Y. & Sanchez, Y. Chk1 in the DNA damage response: conserved roles from yeasts to mammals. *DNA Repair* 3, 1025-1032 (2004).
52. Chen, Z. et al. Selective Chk1 inhibitors differentially sensitize p53-deficient cancer cells to cancer therapeutics. *Int J Cancer* 119, 2784-94 (2006).
53. Cherepanov, P. P. & Wackernagel, W. Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* 158, 9-14 (1995).
54. Cleary, M. A. et al. Production of complex nucleic acid libraries using highly parallel in situ oligonucleotide synthesis. *Nat Methods* 1, 241-248 (2004).
55. Cohen, P. Protein kinases--the major drug targets of the twenty-first century? *Nat Rev Drug Discov* 1, 309-15 (2002).
56. Dasgupta, R. & Perrimon, N. Using RNAi to catch Drosophila genes in a web of interactions: insights into cancer research. *Oncogene* 23, 8359-65 (2004).
57. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* 97, 6640-5 (2000).

58. Davidoff, A. M., Humphrey, P. A., Iglehart, J. D. & Marks, J. R. Genetic basis for p53 overexpression in human breast cancer. *Proc Natl Acad Sci USA* 88, 5006 - 5010 (1991).
59. Davis, S. & Watson, J. C. In vitro activation of the interferon-induced, double-stranded RNA-dependent protein kinase PKR by RNA from the 3' untranslated regions of human alpha-tropomyosin. *Proc Natl Acad Sci U S A* 93, 508-13 (1996).
60. De Cremoux, P. et al. p53 mutation as a genetic trait of typical medullary breast carcinoma. *J Natl Cancer Inst* 91, 641 - 643 (1999).
61. de Fougères, A., Vornlocher, H. P., Maraganore, J. & Lieberman, J. Interfering with disease: a progress report on siRNA-based therapeutics. *Nat Rev Drug Discov* 6, 443-53 (2007).
62. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231-5 (2004).
63. Diab, S. G., Yu, Y. Y., Hilsenbeck, S. G., Allred, D. C. & Elledge, R. M. WAF1/CIP1 protein expression in human breast tumours. *Breast Cancer Res Treat* 43, 99 - 103 (1997).
64. Dickins, R. A. et al. Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nat Genet* 37, 1289-95 (2005).
65. Dickins, R. A. et al. Tissue-specific and reversible RNA interference in transgenic mice. *Nat Genet* 39, 914-21 (2007).
66. NCI Drug Dictionary. Definition of 7-hydroxystaurosporine. <http://www.cancer.gov/Templates/drugdictionary.aspx?CdrID=42305>
67. Doi, N. et al. Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. *Curr Biol* 13, 41-6 (2003).
68. Dominguez, G. et al. Wild type p73 over-expression and high-grade malignancy in breast cancer. *Breast Cancer Res Treat* 66, 183 - 190 (2001).
69. Done, S. J., Eskandarian, S., Bull, S., Redston, M. & Andrulis, I. L. p53 mis-sense mutations in microdissected high-grade ductal carcinoma in situ of the breast. *J Natl Cancer Inst* 93, 700 - 704 (2001).
70. Dork, T. et al. Spectrum of ATM gene mutations in a hospital-based series of unselected breast cancer patients. *Cancer Res* 61, 7608 - 7615 (2001).

71. Elbashir, S. M. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411, 494-498 (2001).
72. Elbashir, S. M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 411, 494-8 (2001).
73. Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* 15, 188-200 (2001).
74. Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W. & Tuschl, T. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J.* 20, 6877-6888 (2001).
75. el-Deiry, W. S. Regulation of p53 downstream genes. *Semin Cancer Biol* 8, 345 - 357 (1998).
76. Elledge, S. J. & Hannon, G. J. An open letter to cancer researchers. *Science* 310, 439-41 (2005).
77. Elledge, S. J. & Walker, G. C. Phasmid vectors for identification of genes by complementation of *Escherichia coli* mutants. *J Bacteriol* 162, 777-83 (1985).
78. Erenpreisa, J., Kalejs, M. & Cragg, M. S. Mitotic catastrophe and endomitosis in tumour cells: an evolutionary key to a molecular solution. *Cell Biol Int* 29, 1012-8 (2005).
79. Esteller, M. et al. Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. *J Natl Cancer Inst* 92, 564 - 569 (2001).
80. Faehnle, C. R. & Joshua-Tor, L. Argonautes confront new small RNAs. *Current Opinion in Chemical Biology* 11, 569-577 (2007).
81. Fahy, B. N., Schlieman, M. G., Mortenson, M. M., Virudachalam, S. & Bold, R. J. Targeting BCL-2 overexpression in various human malignancies through NF-kappaB inhibition by the proteasome inhibitor bortezomib. *Cancer Chemother Pharmacol* (2005).
82. Ferbeyre, G. et al. Oncogenic ras and p53 cooperate to induce cellular senescence. *Mol Cell Biol* 22, 3497-508 (2002).
83. Ferguson, A. T. et al. High frequency of hypermethylation at the 14-3-3sigma locus leads to gene silencing in breast cancer. *Proc Natl Acad Sci USA* 97, 6049 - 6054 (2000).
84. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-11 (1998).

85. Fraser, A. Towards full employment: using RNAi to find roles for the redundant. *Oncogene* 23, 8346-52 (2004).
86. Frei, E., 3rd et al. The effectiveness of combinations of antileukemic agents in inducing and maintaining remission in children with acute leukemia. *Blood* 26, 642-56 (1965).
87. Fridman, J. S. et al. Tumor promotion by Mdm2 splice variants unable to bind p53. *Cancer Res* 63, 5703-6 (2003).
88. Froelich-Ammon, S. J. & Osheroff, N. Topoisomerase poisons: harnessing the dark side of enzyme mechanism. *J Biol Chem* 270, 21429-32 (1995).
89. Gasco, M., Shami, S. & Crook, T. The p53 pathway in breast cancer. *Breast Cancer Res* 4, 70 - 76 (2002).
90. Gatei, M. et al. Ataxia-telangiectasia-mutated (ATM) and NBS1-dependent Phosphorylation of Chk1 on Ser-317 in Response to Ionizing Radiation. *J. Biol. Chem.* 278, 14806-14811 (2003).
91. Gayther, S. A. et al. Mutations truncating the EP300 acetylase in human cancers. *Nat Genet* 24, 300 - 303 (2000).
92. Geisler, S. et al. Influence of TP53 gene alterations and c-erbB-2 expression on the response to treatment with doxorubicin in locally advanced breast cancer. *Cancer Res* 61, 2505 - 2512 (2001).
93. Gentile, M., Ahnstrom, M., Schon, F. & Wingren, S. Candidate tumour suppressor genes at 11q23-q24 in breast cancer: evidence of alterations in PIG8, a gene involved in p53-induced apoptosis. *Oncogene* 20, 7753 - 7760 (2001).
94. Greenblatt, M. S., Chappuis, P. O., Bond, J. P., Hamel, N. & Foulkes, W. D. TP53 mutations in breast cancer associated with BRCA1 or BRCA2 germ-line mutations: distinctive spectrum and structural distribution. *Cancer Res* 61, 4092 - 4097 (2001).
95. Gregory, R. I., Chendrimada, T. P., Cooch, N. & Shiekhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* 123, 631-40 (2005).
96. Gregory, R. I. et al. The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235-40 (2004).
97. Grimm, D. & Kay, M. A. Therapeutic application of RNAi: is mRNA targeting finally ready for prime time? *J Clin Invest* 117, 3633-41 (2007).

98. Grimm, D. et al. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* 441, 537-541 (2006).
99. Grishok, A. et al. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106, 23-34. (2001).
100. Grishok, A. e. a. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106, 23-34 (2001).
101. Hamilton, A. J. B., D.C. A species of small antisense RNA in posttranscriptional gene silencing in plants [see comments]. *Science* 286, 950-952 (1999).
102. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G.J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293-296 (2000).
103. Hammond, S. M., Boettcher, S., Caudy, A.A., Kobayashi, R. & Hannon, G.J. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* 293, 1146-1150 (2001).
104. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293-6 (2000).
105. Han, J. et al. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18, 3016-27 (2004).
106. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* 100, 57-70 (2000).
107. Haney, S. A. Increasing the robustness and validity of RNAi screens. *Pharmacogenomics* 8, 1037-49 (2007).
108. Hannay, J. A. F. et al. Rad51 overexpression contributes to chemoresistance in human soft tissue sarcoma cells: a role for p53/activator protein 2 transcriptional regulation. *Mol Cancer Ther* 6, 1650-1660 (2007).
109. Hannon, G. J. RNA interference. *Nature* 418, 244-51 (2002).
110. Hartwell, L. H., Szankasi, P., Roberts, C. J., Murray, A. W. & Friend, S. H. Integrating Genetic Approaches into the Discovery of Anticancer Drugs. *Science* 278, 1064-1068 (1997).

111. He, L. & Hannon, G. J. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 5, 522-31 (2004).
113. He, L. et al. A microRNA component of the p53 tumour suppressor network. *Nature* 447, 1130-1134 (2007).
114. He, L. et al. A microRNA polycistron as a potential human oncogene. *Nature* 435, 828-833 (2005).
116. Hemann, M. T. et al. Evasion of the p53 tumour surveillance network by tumour-derived MYC mutants. *Nature* 436, 807-11 (2005).
117. Hemann, M. T. et al. An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nat Genet* 33, 396-400 (2003).
120. Hemann, M. T. et al. Suppression of tumorigenesis by the p53 target PUMA. *Proc Natl Acad Sci U S A* 101, 9333-8 (2004).
121. Herbst, A. et al. A conserved element in Myc that negatively regulates its proapoptotic activity. *EMBO Rep* 6, 177-83 (2005).
122. Hernando, E. et al. Rb inactivation promotes genomic instability by uncoupling cell cycle progression from mitotic control. *Nature* 430, 797-802 (2004).
123. Hideshima, T. et al. p38 MAPK inhibition enhances PS-341 (bortezomib)-induced cytotoxicity against multiple myeloma cells. *Oncogene* 23, 8766-76 (2004).
124. Hirao, A. et al. DNA damage-induced activation of p53 by the checkpoint kinase Chk2. *Science* 287, 1824-7 (2000).
125. Ho, G. H. et al. Genetic alterations of the p14ARF-hdm2-p53 regulatory pathway in breast carcinoma. *Breast Cancer Res Treat* 65, 225 - 232 (2001).
126. Hübner, A., Jaeschke, A. & Davis, R. J. Oncogene Addiction: Role of Signal Attenuation. *Developmental Cell* 11, 752-754 (2006).
127. Huesken, D. et al. Design of a genome-wide siRNA library using an artificial neural network. *Nat Biotechnol* 23, 995-1001 (2005).
128. Hussain, S. P. & Harris, C. C. Molecular epidemiology and carcinogenesis: endogenous and exogenous carcinogens. *Mutat Res* 462, 311 - 322 (2000).
129. Hutvagner, G. et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834-8. (2001).

130. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 9, 22-32 (2008).
131. Hutvagner, G. e. a. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293 (2001).
132. Iorns, E., Lord, C. J., Turner, N. & Ashworth, A. Utilizing RNA interference to enhance cancer drug discovery. *Nat Rev Drug Discov* 6, 556-68 (2007).
133. Jackson, A. L. et al. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotech* 21, 635-637 (2003).
134. Jin, J., Ang, X. L., Ye, X., Livingstone, M. & Harper, J. W. Differential Roles for Checkpoint Kinases in DNA Damage-dependent Degradation of the Cdc25A Protein Phosphatase. *J. Biol. Chem.* 283, 19322-19328 (2008).
135. Jixiang, L. et al. Fluorescent molecular probes V: A sensitive caspase-3 substrate for fluorometric assays. *Bioorganic & Medicinal Chemistry Letters* 9, 3231-3236 (1999).
136. Johnson, T. R. et al. The proteasome inhibitor PS-341 overcomes TRAIL resistance in Bax and caspase 9-negative or Bcl-xL overexpressing cells. *Oncogene* 22, 4953-63 (2003).
137. Joliffe, I. *Principal Component Analysis* (Springer, New York, 2002).
138. Kaelin, W. G., Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer* 5, 689-98 (2005).
139. Kandioler-Eckersberger, D. et al. TP53 mutation and p53 overexpression for prediction of response to neoadjuvant treatment in breast cancer patients. *Clin Cancer Res* 6, 50 - 56 (2000).
140. Kawasaki, H. & Taira, K. Short hairpin type of dsRNAs that are controlled by tRNA(Val) promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells. *Nucleic Acids Res* 31, 700-7. (2003).
141. Ketting, R. F. et al. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* 15, 2654-9 (2001).
142. Ketting, R. F. e. a. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* 15, 2654-2659 (2001).

143. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs Exhibit Strand Bias. *Cell* 115, 209-216 (2003).
144. Knight, S. W. & Bass, B. L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* 293, 2269-71. (2001).
145. Koehn, H., Magan, N., Isaacs, R. J. & Stowell, K. M. Differential regulation of DNA repair protein Rad51 in human tumour cell lines exposed to doxorubicin. *Anticancer Drugs* 18, 419-25 (2007).
146. Kolfschoten, I. G. et al. A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. *Cell* 121, 849-58 (2005).
147. Kondratenko, I., Paschenko, O., Polyakov, A. & Bologov, A. Nijmegen breakage syndrome. *Adv Exp Med Biol* 601, 61-7 (2007).
148. Konig, R. et al. A probability-based approach for the analysis of large-scale RNAi screens. *Nat Methods* 4, 847-9 (2007).
149. Krutzfeldt, J. et al. Silencing of microRNAs in vivo with 'antagomirs'. *Nature* 438, 685-689 (2005).
150. Kumagai, A., Kim, S. M. & Dunphy, W. G. Claspin and the activated form of ATR-ATRIP collaborate in the activation of Chk1. *J Biol Chem* 279, 49599-608 (2004).
151. Kumar, N. Taxol-induced polymerization of purified tubulin. Mechanism of action. *J. Biol. Chem.* 256, 10435-10441 (1981).
152. Lakhani, S. R. Putting the Brakes on Cylindromatosis? *N Engl J Med* 350, 187-188 (2004).
153. Lamprecht, M. R., Sabatini, D. M. & Carpenter, A. E. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques* 42, 71-5 (2007).
154. Landthaler, M., Yalcin, A. & Tuschl, T. The Human DiGeorge Syndrome Critical Region Gene 8 and Its *D. melanogaster* Homolog Are Required for miRNA Biogenesis. *Curr Biol* 14, 2162-7 (2004).
155. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858-62 (2001).
156. Lee, J. H. & Paull, T. T. Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. *Oncogene* 26, 7741-7748.

157. Lee, N. S. Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nature Biotechnol.* 20, 500-505 (2002).
158. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843-54 (1993).
159. Lee, Y. et al. The nuclear RNase III Droscha initiates microRNA processing. *Nature* 425, 415-9 (2003).
160. Lee, Y. et al. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* 23, 4051-60 (2004).
161. Lee, Y. e. a. The nuclear RNase III Droscha initiates microRNA processing. *Nature* 425, 415-419 (2003).
162. Lee, Y. S. et al. Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* 117, 69-81 (2004).
163. Leek, J. T., Monsen, E., Dabney, A. R. & Storey, J. D. EDGE: extraction and analysis of differential gene expression. *Bioinformatics* 22, 507-8 (2006).
164. Levesque, A. A., Fanous, A. A., Poh, A. & Eastman, A. Defective p53 signaling in p53 wild-type tumors attenuates p21waf1 induction and cyclin B repression rendering them sensitive to Chk1 inhibitors that abrogate DNA damage-induced S and G2 arrest. *Mol Cancer Ther* 7, 252-62 (2008).
165. Li, M. Z. & Elledge, S. J. MAGIC, an in vivo genetic method for the rapid construction of recombinant DNA molecules. *Nat Genet* 37, 311-9 (2005).
166. Li, X. et al. Generation of destabilized green fluorescent protein as a transcription reporter. *J Biol Chem* 273, 34970-5 (1998).
167. Linderholm, B. K. et al. The expression of vascular endothelial growth factor correlates with mutant p53 and poor prognosis in human breast cancer. *Cancer Res* 61, 2256 - 2260 (2001).
168. Ling, Y. H., Liebes, L., Zou, Y. & Perez-Soler, R. Reactive oxygen species generation and mitochondrial dysfunction in the apoptotic response to Bortezomib, a novel proteasome inhibitor, in human H460 non-small cell lung cancer cells. *J Biol Chem* 278, 33714-23 (2003).
169. Lingel, A. & Izaurralde, E. RNAi: Finding the elusive endonuclease. *RNA* 10, 1675-1679 (2004).

170. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* 11, 576-7 (2004).
171. Liu, J. et al. Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science* 305, 1437-1441 (2004).
172. Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* 7, 719-23 (2005).
173. Lu, P. Y., Xie, F. & Woodle, M. C. In vivo application of RNA interference: from functional genomics to therapeutics. *Adv Genet* 54, 117-42 (2005).
174. Lucchesi, J. C. SYNTHETIC LETHALITY AND SEMI-LETHALITY AMONG FUNCTIONALLY RELATED MUTANTS OF DROSOPHILA MELANOGASTER. *Genetics* 59, 37-44 (1968).
175. Lucito, R. et al. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res* 13, 2291-305 (2003).
176. Lucito, R. a. W., M. Preparation of Slides and Hybridization. (ed. Sambrook), e. D. B. a. J.) (Cold Spring Harbor Press, Cold Spring Harbor, NY., 2003).
177. Lukas, J. et al. WAF1/Cip1 gene polymorphism and expression in carcinomas of the breast, ovary and endometrium. *Am J Pathol* 150, 167 - 175 (1997).
178. Luker, G. D., Pica, C. M., Song, J., Luker, K. E. & Piwnicka-Worms, D. Imaging 26S proteasome activity and inhibition in living mice. *Nat Med* 9, 969-73 (2003).
179. Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear export of microRNA precursors. *Science* 303, 95-8 (2004).
180. Ma, J. B., Ye, K. & Patel, D. J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* 429, 318-22 (2004).
181. Ma, M. H. et al. The proteasome inhibitor PS-341 markedly enhances sensitivity of multiple myeloma tumor cells to chemotherapeutic agents. *Clin Cancer Res* 9, 1136-44 (2003).
182. Maass, N. et al. Decline in the expression of the serine proteinase inhibitor maspin is associated with tumour progression in ductal carcinomas of the breast. *J Pathol* 195, 321 - 326 (2001).

183. MacKeigan, J. P., Murphy, L. O. & Blenis, J. Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. *Nat Cell Biol* 7, 591-600 (2005).
184. MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V. & Doudna, J. A. In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci U S A* 105, 512-7 (2008).
185. Malkin, D. et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* 250, 1233 - 1238 (1990).
186. Maniatakis, E. & Mourelatos, Z. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev* 19, 2979-90 (2005).
187. Mazars, R. et al. p53 mutations occur in aggressive breast cancer. *Cancer Res* 52, 3918 - 3923 (1992).
188. McCaffrey, A. P. et al. RNA interference in adult mice. *Nature* 418, 38-9 (2002).
189. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* 431, 343-9 (2004).
190. Melnikova, I. RNA-based therapies. *Nature Reviews Drug Discovery* 6, 863-864 (2007).
191. Metcalfe, S., Wheeler, T. K., Picken, S., Negus, S. & Milner, A. J. p53 autoantibodies in 1006 patients followed up for breast cancer. *Breast Cancer Res* 2, 438 - 443 (2000).
192. Miller, L. D. et al. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 102, 13550-13555 (2005).
193. Mitsiades, N. et al. The proteasome inhibitor PS-341 potentiates sensitivity of multiple myeloma cells to conventional chemotherapeutic agents: therapeutic applications. *Blood* 101, 2377-80 (2003).
194. Mittal, V. IMPROVING THE EFFICIENCY OF RNA INTERFERENCE IN MAMMALS. *Nat Rev Genet* 5, 355-365 (2004).
195. Moll, U. M., Riou, G. & Levine, A. J. Two distinct mechanisms alter p53 in breast cancer: mutation and nuclear exclusion. *Proc Natl Acad Sci USA* 89, 7262 - 7266 (1992).

196. Moore, D. *The Basic Practice of Statistics (Paper) & Student CD.* (W. H. Freeman, San Francisco, 2006).
197. Mukhopadhyay, U. K., Senderowicz, A. M. & Ferbeyre, G. RNA silencing of checkpoint regulators sensitizes p53-defective prostate cancer cells to chemotherapy while sparing normal cells. *Cancer Res* 65, 2872-81 (2005).
198. Nakayashiki, H. et al. RNA silencing as a tool for exploring gene function in ascomycete fungi. *Fungal Genet Biol* 42, 275-83 (2005).
199. Ngo, V. N. et al. A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* 441, 106-110 (2006).
200. Nikrad, M. et al. The proteasome inhibitor bortezomib sensitizes cells to killing by death receptor ligand TRAIL via BH3-only proteins Bik and Bim. *Mol Cancer Ther* 4, 443-9 (2005).
201. Norberg, T. et al. Increased p53 mutation frequency during tumor progression - results from a breast cancer cohort. *Cancer Res* 61, 8317 - 8321 (2001).
202. O'Connor, M. J., Martin, N. M. & Smith, G. C. Targeted cancer therapies based on the inhibition of DNA strand break repair. *Oncogene* 26, 7816-24 (2007).
203. Paddison, P. J. A resource for large-scale RNA-interference-based screens in mammals. *Nature* 428, 427-431 (2004).
204. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J. & Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 16, 948-58 (2002).
205. Paddison, P. J., Caudy, A. A. & Hannon, G. J. Stable suppression of gene expression by RNAi in mammalian cells. *Proc Natl Acad Sci U S A* 99, 1443-8 (2002).
206. Paddison, P. J. & Hannon, G. J. RNA interference: the new somatic cell genetics? *Cancer Cell* 2, 17-23 (2002).
207. Paddison, P. J. et al. A resource for large-scale RNA-interference-based screens in mammals. *Nature* 428, 427-31 (2004).
208. Papac, R. J. Origins of cancer therapy. *Yale J Biol Med* 74, 391-8 (2001).
209. Parker, J. S. & Barford, D. Argonaute: a scaffold for the function of short regulatory RNAs. *Trends in Biochemical Sciences* 31, 622-630 (2006).

210. Paul, C. P., Good, P. D., Winer, I. & Engelke, D. R. Effective expression of small interfering RNA in human cells. *Nature Biotechnol.* 20, 505-508 (2002).
211. Pei, Y. & Tuschl, T. On the art of identifying effective and specific siRNAs. *Nat Meth* 3, 670-676 (2006).
212. Pham, J. W., Pellino, J. L., Lee, Y. S., Carthew, R. W. & Sontheimer, E. J. A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in *Drosophila*. *Cell* 117, 83-94 (2004).
213. Pharaoh, P. D., Day, N. E. & Caldas, C. Somatic mutations in the p53 gene and prognosis in breast cancer: a meta-analysis. *Br J Cancer* 80, 1968 - 1973 (1999).
214. Phillips, K. A. et al. Frequency of p53 mutations in breast carcinomas from Ashkenazi Jewish carriers of BRCA1 mutations. *J Natl Cancer Inst* 91, 469 - 473 (1999).
215. Pichler, A., Zelcer, N., Prior, J. L., Kuil, A. J. & Piwnicka-Worms, D. In vivo RNA interference-mediated ablation of MDR1 P-glycoprotein. *Clin Cancer Res* 11, 4487-94 (2005).
216. Pimienta, G. & Pascual, J. Canonical and alternative MAPK signaling. *Cell Cycle* 6, 2628-32 (2007).
217. Quackenbush, J. Microarray data normalization and transformation. *Nat Genet* 32 Suppl, 496-501 (2002).
218. Quesnel, B., Preudhomme, C., Fournier, J., Fenaux, P. & Peyrat, J. P. MDM2 gene amplification in human breast cancer. *Eur J Cancer* 30A, 982 - 984 (1994).
219. Rabinowicz, P. D. et al. Genes and transposons are differentially methylated in plants, but not in mammals. *Genome Res* 13, 2658-64 (2003).
220. Rajkumar, S. V., Richardson, P. G., Hideshima, T. & Anderson, K. C. Proteasome inhibition as a novel therapeutic target in human cancer. *J Clin Oncol* 23, 630-9 (2005).
221. Raman, V. et al. Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature* 405, 974 - 978 (2000).
222. Ramdas, L. et al. Sources of nonlinearity in cDNA microarray expression measurements. *Genome Biol* 2, RESEARCH0047 (2001).
223. Reinhardt, H. C., Aslanian, A. S., Lees, J. A. & Yaffe, M. B. p53-deficient cells rely on ATM- and ATR-mediated checkpoint signaling through the p38MAPK/MK2 pathway for survival after DNA damage. *Cancer Cell* 11, 175-89 (2007).

224. Reinhart, B. J. et al. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901-6. (2000).
225. Reynolds, A. et al. Rational siRNA design for RNA interference. *Nat Biotech* 22, 326-330 (2004).
226. Richardson, P. G., Mitsiades, C., Hideshima, T. & Anderson, K. C. Proteasome Inhibition in the Treatment of Cancer. *Cell Cycle* 4 (2005).
227. Riley, T., Sontag, E., Chen, P. & Levine, A. Transcriptional control of human p53-regulated genes. *Nat Rev Mol Cell Biol* 9, 402-12 (2008).
228. Ringner, M. What is principal component analysis? *Nat Biotech* 26, 303-304 (2008).
229. Roberts, F. a. R., D. (2008).
230. Roche-Lestienne, C. et al. Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. *Blood* 100, 1014-1018 (2002).
231. Root, D. E., Hacohen, N., Hahn, W. C., Lander, E. S. & Sabatini, D. M. Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Meth* 3, 715-719 (2006).
232. Rossi, J. J. Expression strategies for short hairpin RNA interference triggers. *Hum Gene Ther* 19, 313-7 (2008).
233. Rossi, R., Lidonnici, M. R., Soza, S., Biamonti, G. & Montecucco, A. The dispersal of replication proteins after Etoposide treatment requires the cooperation of Nbs1 with the ataxia telangiectasia Rad3-related/Chk1 pathway. *Cancer Res* 66, 1675-83 (2006).
234. Sachse, C. & Echeverri, C. J. Oncology studies using siRNA libraries: the dawn of RNAi-based genomics. *Oncogene* 23, 8384-91 (2004).
235. Samuels-Lev, Y. et al. ASPP proteins specifically stimulate the apoptotic function of p53. *Mol Cell* 8, 781 - 794 (2001).
236. Sarov, M. & Stewart, A. F. The best control for the specificity of RNAi. *Trends Biotechnol* 23, 446-8 (2005).
237. Sawyers, C. Targeted cancer therapy. *Nature* 432, 294-7 (2004).

238. Sawyers, C. L. Making Progress through Molecular Attacks on Cancer. Cold Spring Harbor Symposia on Quantitative Biology 70, 479-482 (2005).
239. Schwarz, D. S. Asymmetry in the assembly of the RNAi enzyme complex. Cell 115, 199-208 (2003).
240. Schwarz, D. S. et al. Asymmetry in the Assembly of the RNAi Enzyme Complex. Cell 115, 199-208 (2003).
241. Seger, Y. R. et al. Transformation of normal human cells in the absence of telomerase activation. Cancer Cell 2, 401-13 (2002).
242. Sgroi, D. C. et al. In vivo gene expression profile analysis of human breast cancer progression. Cancer Res 59, 5656 - 5661 (1999).
243. Shaffer, A. L. et al. IRF4 addiction in multiple myeloma. Nature 454, 226-231 (2008).
244. Shah, N. P. et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. Cancer Cell 2, 117-25 (2002).
245. Shao, Z. M., Wu, J., Shen, Z. Z. & Nguyen, M. p53 mutation in plasma DNA and its prognostic value in breast cancer patients. Clin Cancer Res 7, 2222 - 2227 (2001).
246. Sigal, A. & Rotter, V. Oncogenic mutations of the p53 tumor suppressor: The demons of the guardian of the genome. Cancer Res 60, 6788 - 6793 (2000).
247. Silva, J., Chang, K., Hannon, G. J. & Rivas, F. V. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. Oncogene 23, 8401-9 (2004).
248. Silva, J. et al. Analysis of genetic and epigenetic processes that influence p14ARF expression in breast cancer. Oncogene 20, 4586 - 4590 (2001).
249. Silva J, C. K., Hannon GJ, Rivas FV. RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. Oncogene 51, 8401-9. (2004).
250. Silva, J. M., Hammond, S. M. & Hannon, G. J. RNA interference: a promising approach to antiviral therapy? Trends Mol Med 8, 505-8 (2002).
251. Silva, J. M. et al. Profiling essential genes in human mammary cells by multiplex RNAi screening. Science 319, 617-20 (2008).

252. Siolas, D. et al. Synthetic shRNAs as potent RNAi triggers. *Nat Biotechnol* 23, 227-31 (2005).
253. Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H. & Williams, B. R. Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* 5, 834-9 (2003).
254. Smith, P. D. et al. Novel p53 mutants selected in BRCA-associated tumours which dissociate transformation suppression from other wild-type p53 functions. *Oncogene* 18, 2451 - 2459 (1999).
255. Smyth, G. K. & Speed, T. Normalization of cDNA microarray data. *Methods* 31, 265-73 (2003).
256. Song, J. J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* 10, 1026-32 (2003).
257. Song, J.-J. & Joshua-Tor, L. Argonaute and RNA -- getting into the groove. *Current Opinion in Structural Biology* 16, 5-11 (2006).
258. Song, J.-J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity. *Science* 305, 1434-1437 (2004).
259. Sordella, R., Bell, D. W., Haber, D. A. & Settleman, J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* 305, 1163-7 (2004).
260. Soutschek, J. et al. Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* 432, 173-8 (2004).
261. Stankovic, T. et al. ATM mutations and phenotypes in ataxia-telangiectasia families in the British Isles: expression of mutant ATM and the risk of leukemia, lymphoma, and breast cancer. *Am J Hum Genet* 62, 334 - 345 (1998).
262. Stegmeier, F., Hu, G., Rickles, R. J., Hannon, G. J. & Elledge, S. J. A Lentiviral microRNA-based System for Single Copy Pol II Regulated RNAi in mammalian Cells. *Proc Natl Acad Sci U S A*, in press (2005).
263. Storey, J. D., Dai, J. Y. & Leek, J. T. The optimal discovery procedure for large-scale significance testing, with applications to comparative microarray experiments. *Biostatistics* 8, 414-32 (2007).
264. Storey, J. D., Xiao, W., Leek, J. T., Tompkins, R. G. & Davis, R. W. Significance analysis of time course microarray experiments. *Proc Natl Acad Sci U S A* 102, 12837-42 (2005).

265. Sui, G. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl Acad. Sci. USA* 99, 5515-5520 (2002).
266. Sullivan, A. et al. Concomitant inactivation of p53 and Chk2 in breast cancer. *Oncogene* (2002).
267. Tan, T. T. et al. Key roles of BIM-driven apoptosis in epithelial tumors and rational chemotherapy. *Cancer Cell* 7, 227-38 (2005).
268. Tang, J., Erikson, R. L. & Liu, X. Checkpoint kinase 1 (Chk1) is required for mitotic progression through negative regulation of polo-like kinase 1 (Plk1). *Proc Natl Acad Sci U S A* 103, 11964-9 (2006).
269. Tara, M. L., Howell, F. M. & Carl, D. N. Not miR-ly small RNAs: Big potential for microRNAs in therapy. *The Journal of allergy and clinical immunology* 121, 309-319 (2008).
270. Tolia, N. H. & Joshua-Tor, L. Slicer and the Argonauts. *Nat Chem Biol* 3, 36-43 (2007).
271. Tomari, Y. et al. RISC assembly defects in the *Drosophila* RNAi mutant armitage. *Cell* 116, 831-41 (2004).
272. Torrance, C. J., Agrawal, V., Vogelstein, B. & Kinzler, K. W. Use of isogenic human cancer cells for high-throughput screening and drug discovery. *Nat Biotechnol* 19, 940-5 (2001).
273. Tran, H. M. et al. Mutant Nbs1 enhances cisplatin-induced DNA damage and cytotoxicity in head and neck cancer. *Otolaryngol Head Neck Surg* 131, 477-84 (2004).
274. Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P. & Sharp, P. A. Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev.* 13, 3191-3197 (1999).
275. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98, 5116-21 (2001).
276. Umbricht, C. B. et al. Hypermethylation of 14-3-3 sigma (stratifin) is an early event in breast cancer. *Oncogene* 20, 3348 - 3353 (2001).
277. Vahteristo, P. et al. p53, CHK2, and CHK1 genes in Finnish families with Li-Fraumeni syndrome: further evidence of CHK2 in inherited cancer predisposition. *Cancer Res* 61, 5718 - 5722 (2001).

278. van 't Veer, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530-6 (2002).
279. Verlinden, L. et al. The E2F-regulated gene Chk1 is highly expressed in triple-negative estrogen receptor /progesterone receptor /HER-2 breast carcinomas. *Cancer Res* 67, 6574-81 (2007).
280. Vitale, I. et al. Inhibition of Chk1 kills tetraploid tumor cells through a p53-dependent pathway. *PLoS ONE* 2, e1337 (2007).
281. Vitale, I. et al. Chk1 inhibition activates p53 through p38 MAPK in tetraploid cancer cells. *Cell Cycle* 7, 1956-61 (2008).
282. Vogelstein, B. & Kinzler, K. W. Achilles' heel of cancer? *Nature* 412, 865 - 866 (2001).
283. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nat Med* 10, 789-99 (2004).
284. Vogelstein, B., Lane, D. & Levine, A. J. Surfing the p53 network. *Nature* 408, 307 - 310 (2001).
285. Vousden, K. H. & Lu, X. Live or let die: the cell's response to p53. *Nat Rev Cancer* 2, 594-604 (2002).
286. Wall, M. E., Andreas Rechtsteiner, Rocha, LM. Singular value decomposition and principal component analysis". (Kluwer, Norwell, MA 2003).
287. Weinstein, I. B. CANCER: Enhanced: Addiction to Oncogenes--the Achilles Heal of Cancer. *Science* 297, 63-64 (2002).
288. Westbrook, T. F. et al. A genetic screen for candidate tumor suppressors identifies REST. *Cell* 121, 837-48 (2005).
289. Whitehurst, A. W. et al. Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature* 446, 815-9 (2007).
290. Willingham, A. T., Deveraux, Q. L., Hampton, G. M. & Aza-Blanc, P. RNAi and HTS: exploring cancer by systematic loss-of-function. *Oncogene* 23, 8392-400 (2004).
291. Wolters, N. M. & MacKeigan, J. P. From sequence to function: using RNAi to elucidate mechanisms of human disease. *Cell Death Differ* 15, 809-19 (2008).
292. Xue, W. et al. Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature* 445, 656-60 (2007).

293. Yang, X. H., Shiotani, B., Classon, M. & Zou, L. Chk1 and Claspin potentiate PCNA ubiquitination. *Genes Dev* 22, 1147-52 (2008).
294. Zamore, P. D., Tuschl, T., Sharp, P.A. & Bartel, D.P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* 101, 25-33 (2000).
295. Zeng, Y., Wagner, E.J. & Cullen, B.R. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* 9, 1327-1333 (2002).
296. Zeng, Y. & Cullen, B. R. Sequence requirements for micro RNA processing and function in human cells. *Rna* 9, 112-23 (2003).
297. Zeng, Y., Wagner, E. J. & Cullen, B. R. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* 9, 1327-33. (2002).
298. Zhang, H. et al. BRCA1 physically associates with p53 and stimulates its transcriptional activity. *Oncogene* 16, 1713 - 1721 (1998).
299. Zhang, Y., Zhou, J. & Lim, C. U. K. The role of NBS1 in DNA double strand break repair, telomere stability, and cell cycle checkpoint control. *Cell Res* 16, 45-54.
300. Zheng, L. et al. An approach to genomewide screens of expressed small interfering RNAs in mammalian cells. *Proc Natl Acad Sci U S A* 101, 135-40 (2004).
301. Zimmermann, T. S. et al. RNAi-mediated gene silencing in non-human primates. *Nature* 441, 111-4 (2006).